Task Order: WRT-1009, Model Curation Innovation & Implementation

Technical Report SERC-2020-TR-003

June 04, 2020

Principal Investigator: Dr. Donna H. Rhodes, Massachusetts Institute of Technology

Research Team:

Massachusetts Institute of Technology: Dr. Eric Rebentisch, Dr. Adam M. Ross

Sponsor: DASD (SE)

The Systems Engineering Research Center (SERC) is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Office of the Assistant Secretary of Defense for Research and Engineering (ASD(R&E)) under Contract [HQ0034-19-D-0003, TO#0242].

TABLE OF CONTENTS

Table of Contents iii

List of Figures v

List of (Tables, Sequences) v

Executive Summary 6

PERSONNEL 7
- Research Team: 7
- Government: 7

Background and Research Objectives 8
- Research Needed 9

Model Curation Implementation Practices 10
- Curation in the Model Life Cycle 10
  - Early Efforts on Curation of Engineering Models 11
- Types of Models in an Enterprise Model Collection 11
  - Exploring Suitability of Analysis Models for Enterprise-Level Curation 12
- Criteria for Placing Models under Curation 16
- Enterprise Model Collection Object 18
- Implementation & Governance Guiding Principles 20

Model Curation and Authoritative Source of Truth (AST) 22
- Sources of Authority 23
- Examining Accreditation in a Model Curation Context 24
  - General Requirements Associated With Accreditation 26
- Model Credibility 30
  - Credibility of models and simulations 30
  - Assessment of credibility of models and simulations 32
  - Website Credibility 33
  - Recent research on model confidence and trust 33
  - Toward design guidelines for model curation 34
- Precursors to Authoritative Source of Truth 35

Innovation 36
- Curation of BioModels 36
- Display of Model Information 37
- Augmented Model Discovery 38

Conclusion and Future Research 39
- Model Curation Interview-based Empirical Study 39
- Curating for Model Consumers 39
- Relationship of Model Credibility and Data Credibility 40
- Valuation of Enterprise-level Models/Model Collections 40

Accreditation and Enterprise Model Collections 40
Digital Demonstrators 41
Project Timeline & Transition Plan 42
Appendix A: List of Publications Resulted 44
Appendix B: Cited and Related References 45
Appendix C. Model Curation Lexicon 49

LIST OF FIGURES

Figure 1 M&S Life Cycle from NASA-STD-7009A W/CHANGE 1 [NASA 2016, p. 70] 10
Figure 2 Tradeoff analysis elements (left) with associated models (right) 13
Figure 3 Enterprise Model Collection Elements (notional) 18
Figure 4 M&S Options and VV &A Steps (source: Air Force Instruction 16-001, 29 April 2020).. 26
Figure 5 SCS Framework to Review Credibility of a Simulation [SCS, 1979] 31
Figure 6. Balci's Life Cycle of a Simulation Study 32
Figure 7 DeVin's schematic of factors influencing overall credibility of simulation results 34
Figure 8 Two Branches of the Model Production Pipeline of BioModels (source, Juty et al., 2016) 37
Figure 9 DARPA Data-Driven Discovery of Models (D3M) 38
Figure 10 SERC Workshop on Model Curation (postponed to Nov 2020).................................. 42

LIST OF (TABLES, SEQUENCES)

Table 1 Description of Model Types for Tradeoff Analyses... 14
Table 2. Criteria for Placing Models Under Curation (adapted from: DCC) 17
Table 3 Transition Actions and SERC Principles Implemented 43

EXECUTIVE SUMMARY

SERC research project WRT-1009 is a continued investigation of model curation as a topic that spans implementation of model curation practice, new roles and responsibilities of individuals and organizations, and approaches to curate models for intended purpose. Knowledge gathering and continued interaction with research stakeholders has furthered the understanding of elements of implementation practice. Criteria for placing models under enterprise-level curation were adapted based on curation in other fields. Preliminary implementation and governance guiding principles were derived from research findings. An exploration of enablers, barriers and precursors for model curation and authoritative source of truth motivated a deeper exploration of model credibility and accreditation. As a result, precursors to credibility were formulated. Several areas of innovation were identified as candidates for use of newer technologies to advance model curation. Six recommendations for future research are identified.

Model Curation can be defined as the *lifecycle management, control, preservation and active enhancement of models and associated information to ensure value for current and future use, as well as repurposing beyond initial purpose and context*. Curation practices promote formalism and provide for the strategic management and control of models and associated digital artifacts, particularly when managed as a collection at the enterprise level. Curation activities include model governance, accession, acquisition, valuation, preservation, active enhancement, model discovery, deaccessioning, and archiving. Not all models are suitable for enterprise-level curation. Curation applies to longer duration models, rather than those developed for a quick study or to simply work out a problem. A first category of models suitable for curation includes models that will be used throughout the lifespan of a major program, for example models comprising a digital twin. A second category includes models designed (or enhanced) to be intentionally reused for a new purpose and/or within a new context. Examples are reference architectures and models, and "platform" models that enable the enterprise to effectively re-purpose and reuse models. Model curation requires supporting infrastructure to enable an enterprise to establish and actively enhance a collection of models of value to the larger enterprise. As evidenced by curation practice in institutional collections (e.g., museums, historical society, libraries), dedicated leadership, governance and support functions are essential. Prior phase investigation of model curation, especially on organizational aspects, was accomplished in SERC RT-199 (2019).

PERSONNEL

Research Team:

Name	Organization	Labor Category	Contact
Donna H. Rhodes	MIT	Principal Investigator	rhodes@mit.edu
Eric Rebentisch	MIT	Research Scientist	erebenti@mit.edu
Adam M. Ross	MIT	Research Scientist	adamross@mit.edu

Government:

Name	Organization	Role	Contact
Philomena Zimmerman	OUSD(R&E)	Deputy Director	Philomena.m.zimmerman.civ@mail.mil
Scott Lucero	OUSD(R&E)	SERC Program Manager	don.s.lucero.civ@mail.mil
Rhyan Johnson	OUSD(R&E)	Support to SERC PM	rhyan.d.johnson.ctr@mail.mil

Background and Research Objectives

As engineering practice becomes increasingly model-centric, models become increasingly valuable assets for designing and evolving systems. Their continuing existence throughout the program lifespan makes the management and control of models and digital artifacts imperative. There is a growing recognition of the need for curation, and the DoD Digital Engineering Strategy (2018) explicitly states one focus area is to" formally develop, integrate and curate models."

Model Curation can be defined as "the lifecycle management, control, preservation and active enhancement of models and associated information to ensure value for current and future use, as well as repurposing beyond initial purpose and context". The SERC's initial investigation of model curation and the curation role has indicated that the systems engineering community will benefit from defined curation practices. At present, a significant challenge exists; model curation is a desired activity but there are no standards or exemplar practices available to the systems community. While useful practices from other fields can be found, there is a need for research to determine which of these existing practices are effective and appropriable for the engineering context, and what additional practices are required for digital engineering. Lacking research on model curation implementation, it is likely that curation practices will emerge in isolated pockets through ad-hoc approaches. As a result, standardization and maturation of curation practices would be prolonged, and programs having multiple constituents will spend resources on developing and aligning curation practices rather than tailoring standard implementation practices. Pragmatic approaches to curation are needed in the near term, and early adopters are challenged with finding information and knowledge to inform their efforts. Further, there has been minimal exploration of how new technologies may offer capabilities for performing model curation in the future. As in the case of digital curation, active management and enhancement is an essential aspect of model curation, distinguishing it from curation practices that focus solely on collecting and storing data and information. This is because models in digital engineering enterprises have potential to be reused and repurposed beyond their original context and use. The lack of access to models, mistrust of models, and perception of legitimacy of models are all barriers in model reuse and longevity, potentially mitigated by model curation.

Model curation implementation practices will provide practical formalism to ensure both technical and non-technical data and information concerning a model is created and maintained, including model metadata and model pedigree. With digital engineering transformation, model curation can be expected to involve unique practices and responsibilities at the enterprise level. A key benefit of model curation is that it increases the potential for effective reuse of a model on a future program, as the model and associated information (metadata, technical data, and pedigree) will be available to assess its fitness for purpose and context. Effective curation of models will have positive impact on model trust, credibility and integrity. In addition to defined model curation implementation practices, several enablers for implementation offer potential to further enhance the practice. Innovative approaches and technologies may be beneficial in curation as digital engineering practices and infrastructure mature.

The WRT-1009 research project builds on SERC research findings from prior SERC research (Rhodes & Ross 2015; Rhodes, 2018b; Rhodes, 2019b). Five points on curation emerging from prior research that inform this investigation are:

1. SERC's initial investigation of model curation has indicated the systems community will benefit from formal curation practices.
2. Lack of access to models, mistrust of models, and perception of legitimacy of models are barriers to reuse and longevity that are potentially mitigated through model curation.
3. There is potential to adapt practices from other fields once model curation-specific needs in the digital engineering context are clearly understood.
4. Enablers for implementation have potential to further enhance curation practice, including lexicon, pedigree standards, curation criteria, and many others.
5. Model curation requires both implementation and governance practices.

Research Needed

Three areas were identified for investigation in WRT-1009.

1. Investigate and Adapt/Develop Model Curation Implementation Practices. Research is needed to identify and adapt/develop practices through an iterative approach. Building on initial prior research, this activity examines successful practices from other fields and adapts these as appropriate for digital engineering. The investigation of precursor practices will inform curation practices and identify gaps between present practice and curation objectives.

2. Investigate Model Curation Precursors, Enablers and Barriers. There is a need to understand what curation-like practices presently exist across various types of engineering organizations and how these are presently being implemented. There is a need to investigate existing and non-existing enablers for curation implementation (e.g., standards, model-related data capture templates, etc.). The relationship of model curation and authoritative source of truth needs to be explored, including model validity and trust.

3. Investigate Innovative Technologies. Research is needed to investigate opportunities for future application of newer technology (e.g., machine learning, augmented intelligence, data science techniques, etc.) in model curation. This includes investigating technology innovation in other curation fields and emerging approaches that may enhance curation practice (e.g., ontology, natural language processing, automated feature engineering, visual analytics, etc.). A desired outcome is to formulate recommendations for future investigation.

MODEL CURATION IMPLEMENTATION PRACTICES

Model curation implementation practices are expected to evolve over time. While the specific practices will be unique to the organization, there is a need for fundamentals that can inform the overall practice. Accordingly, this research project considers four areas related to commonality in practice, and has performed preliminary investigation. It is expected that continued efforts of the systems community will result in convergence on these topics:

- A conceptual view of where curation fits into the model lifecycle.
- Types of models that would be appropriate for an enterprise model collection.
- Criteria for the decision to place models under enterprise-level curation.
- Elements that would comprise an *enterprise model collection object*.

Curation in the Model Life Cycle

The latest version of NASA Standard for Models and Simulations (NASA, 2016) illustrates the life cycle of M&S as shown in Figure 1. This is a useful depiction to consider where model curation would overlay existing life cycle models. An enterprise model collection would come into play at the start the life cycle, where reuse/repurposing of an existing model would be considered as an option to replaces the development of a new model (blue text in the figure). As completed models are selected for curation, model curation can be thought of as overlapping the latter two phases, model use/operations and model archiving (green text in the figure). In the final phase, the decision would be made as whether to archive the model as an enterprise-level asset.

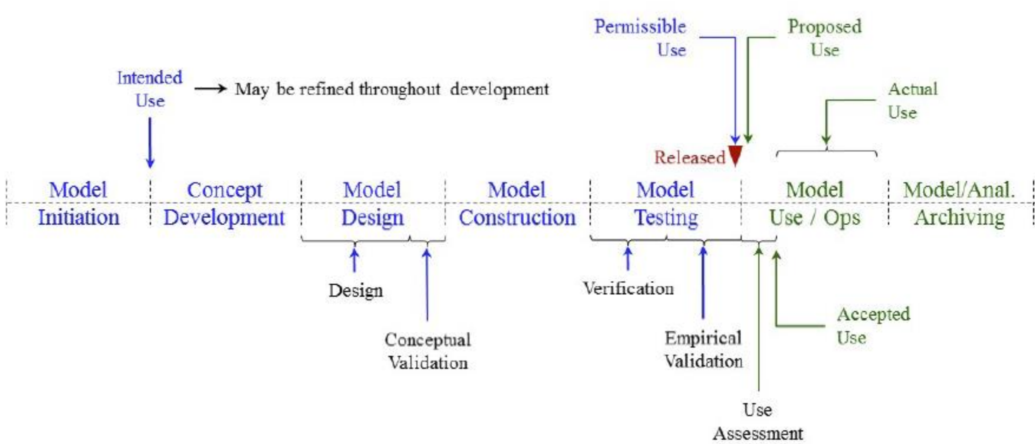

Figure 1 M&S Life Cycle from NASA-STD-7009A W/CHANGE 1 [NASA 2016, p. 70]

Model curation applies to the numerous types of models used in digital engineering practice. The terms *model* and *simulation* are used in this report; useful definitions from NASA (2016) are:

> ***Model****: A description or representation of a system, entity, phenomenon, or process (adapted from Banks (1998).* ***Note:*** *A model may be constructed from multiple sub-models; the sub-models and the integrated sub-models are all considered models. Likewise, any data that goes into a model are considered part of the model.*
>
> ***Simulation:*** *The imitation of the behavioral characteristics of a system, entity, phenomenon, or process.*

Early Efforts on Curation of Engineering Models

The earliest evidence of curation applied to engineering models was performed by Patel et al. (2009) for CAD models. The authors noted that the information to be dealt with in curation of CAD engineering models is diverse and particularly complex. They observe challenges of communicating with a wide range of different stakeholders, each having unique information needs and access rights. Their work primarily focuses on technical strategies for curation; but Patel et al. also state “there is a need for best practice guidelines and cost-benefit models to aid in choosing appropriate curation strategies since the business of deciding a suitable path is non-trivial and contingent on many factors.”

Types of Models in an Enterprise Model Collection

Curation applies to longer duration models, rather than those developed for a quick study or to simply work out a problem. A first category of models suitable for curation includes models that will be used throughout the lifespan of a major program, for example models comprising a digital twin. A second category includes models designed (or enhanced) to be intentionally reused for a new purpose and/or within a new context. Examples are reference architectures and models, and “platform” models that enable the enterprise to effectively re-purpose and reuse models.

There are many types of models involved in engineering that are typically owned and controlled by individual programs and/or units within a program area. Some of these models are developed and used/reused for analyzing performance and making interim decisions. The value of collecting and maintaining some of these models at the enterprise-level is likely not worth the return on investment. The decision to elevate a model to an enterprise collection must be purposeful and well-informed. *Accession* is the curation activity that accepts a model into the enterprise collection, using appraisal of value to the enterprise and specific acceptance criteria.

There would appear to be clear-cut instances of models that belong in an enterprise collection, for instance, a systems model used on a major program. The return on investment for the enterprise to maintain a model in an enterprise repository is an open area of investigation. This necessitates having a process for placing a monetary value on the model and being able to assess the cost of maintaining it in a model repository as either an active or archived asset. Initially, it would be expected that enterprises will determine value using some proxy measure or ordinal scale, rather than a monetary value.

Empirical research would be useful to gather expert opinion on what models are best suited for enterprise-level curation. The various types of models need to be analyzed to inform this. An analysis model is an example of a type of model where it is more difficult to discern if it belongs in an enterprise model repository. For purposes of this report, we use the term "model package" to refer to the complete set of model, data and information that would comprise a model collection object. In the following subsection some considerations are explored for analysis models, including what a "model package" might include for such a model.

EXPLORING SUITABILITY OF ANALYSIS MODELS FOR ENTERPRISE-LEVEL CURATION

An important use case for models is to support the conduct of tradeoff analyses. In such analyses, users seek to generate data-driven evidence that describe coupling between decisions and outcomes. For example, how design decisions impact performance and cost of potential alternative systems. For cases where more than one outcome measure is considered, it is likely the case that a single "best" answer does not exist (i.e. the "best" answer for a design choice). Usually increases in one measure also come at the expense of decreases in another measure (e.g. the proverbial cost-benefit tradeoff where higher cost usually corresponds to higher benefit, and lower cost to lower benefit, with no clear "best" answer). When tradeoffs exist, there must be a decision logic for determining how to go about discovering the tradeoff relationships (i.e. how do outcomes X and Y relate to one another), as well as how to go about making a selection decision (i.e. how to aggregate the outcomes into a single dimension metric, such as ranking). In fact, the structure of tradeoff analyses can be roughly characterized into four elements: decisions, outcomes, evaluations, and other factors (Figure 2, left).

Decisions encompass the factors under control of the analyst or engineer, while outcomes are consequences of the decisions. It is the relationships between the decisions and outcomes that are typically investigated in tradeoff analyses. For example, investigating the impact of wheel base, engine, and chassis design on vehicle weight, fuel efficiency, and safety. In order to generate evidence (e.g. modeled-derived data) on the relationships, we need to be able to evaluate the decisions in terms of the outcomes, but this often requires taking account of other factors, such as assumptions on operations and environment. Each of these four factors are in fact models themselves (Figure 2, right), with the evaluations (i.e. evaluative models) most closely resembling I/O models, such as physics-based models or parametric models, such as cost-estimating relationships. The other three factors in fact are also model-based, although not through traditional I/O models, but rather through representation-type models where the analyst or engineer is making conceptualization decisions in how to numerically represent the decisions (input models), outcomes (output models), and other factors (other models).

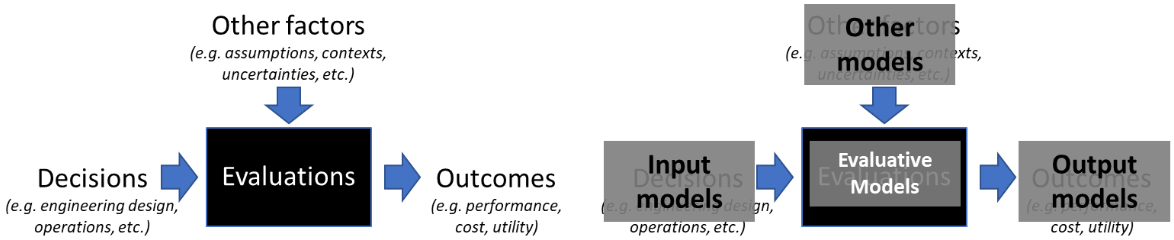

Figure 2 Tradeoff analysis elements (left) with associated models (right)

The input models include not only the meta-data associated with a tradeoff analysis (such as goals, timeframe, and context for the study), but also the representational model for how the decisions will be represented as input variables. For example, parameterization of a satellite design in terms of its physical subsystem choices, orbit elements, and launch date. Additionally, decisions around what space of alternative inputs will be considered in the tradeoff analysis is typically derived using a structured approach or algorithm, such as Design of Experiment (DOE) approaches for covering a design space. Reproduction of the study would necessitate some capture of these algorithms and sampling decisions, not only in terms of outcome goals, but also in terms of predicted resource usage or constraints for running the evaluation of the input space (a very common constraint on the scale of a tradeoff analysis is the time/resource cost of evaluating the number of alternatives considered).

The output models describe how one or more outcome measures will be interpreted in order to draw conclusions, or support decisions around the nature of tradeoffs that exist among outputs or between inputs and outputs. Sometimes this may involve use of one or more multi-objective algorithms (including optimization techniques) or other means for aggregating outcome measures into one or more decision criteria. Such models could also include value models, which transform the raw outputs into "goodness" scores according to stakeholder or market priorities (e.g. utility models, or economic models).

The last category, other models, includes capturing of not only data needed for the execution of the other models, but also representational models of the external factors that may change and alter the outputs of the evaluative models. For example, models describing uncertainty in environment or technology performance that are needed by the evaluative model. Different means for representing that uncertainty could be undertaken, each with associated costs and benefits for collection and maintenance. The explicit consideration of these four types of models in a tradeoff analysis not only help to ensure that the analysis is grounded in data, but also increases the likelihood of potential reuse and appropriate interpretation of results.

Table 1 summarizes these four model types.

Table 1 Description of Model Types for Tradeoff Analyses

Model Type	Description	Examples
Inputs	The approach by which the study inputs are defined, enumerated and sampled	Data describing existing systems, design parameterization, sampling algorithms such as DOE and associated data and assumptions
Other	The supporting data and assumptions needed for execution of evaluative models	Data about the context for the tradeoff analyses, any assumptions or fixed data inputs for the evaluative models
Evaluative	The models used to predict outputs based on inputs and other needed data	Performance and cost models
Outputs	The approach by which the study utilizes or transforms the output data in order to support tradeoff decisions	Value models or other multi-objective algorithms for aggregating multiple output metrics into decision criteria, along with associated decision logic for making decisions or drawing tradeoff conclusions

Given these four types of models of relevance for a tradeoff analysis study, under what conditions does it make sense to manage such "packages" at the enterprise level rather than within a program? The key motivations for curating these at the enterprise level is reuse across programs, and alignment with organizational context and priorities.

From a reuse perspective, clearly there may be value in being able to reuse evaluative models where such models may be relevant across programs. For example, in the DoD context, there are many instances where AoAs are performed on similar classes of systems, such as ground or air platforms. To the extent evaluative models, such as cost-estimating relationships, are relevant beyond a given program, there is opportunity for substantial time and cost savings if reuse is supported with proper documentation and computational environments. This type of reuse is likely already considered within organizations. Reuse of the input models, output models, and other models may be less common. In particular, reuse of input sampling algorithms, as well as output value models and decision algorithms are likely to be relevant across programs (as evidenced by existing reusable libraries of optimization and sampling algorithms in open source[1],[2]). Furthermore, the other models type encompasses a broad set of both data and context representation that could be relevant across an enterprise, not only for other programs of similar nature, but also for organizational insights into assumptions about factors outside of a given program. This latter case provides an opportunity for the enterprise to inform alignment of these models that may not be possible from within a particular program.

---

[1] Dakota, Sandia National Laboratories, "Dakota toolkit provides a flexible, extensible interface between simulation codes and iterative systems analysis methods… The Dakota project delivers both state-of-the-art research and robust, usable software for optimization and UQ." https://dakota.sandia.gov/, [last accessed June 2020]

[2] OpenMDAO, OpenMDAO is "an open-source framework for efficient multidisciplinary optimization… The OpenMDAO project is focused on supporting gradient-based optimization with analytic derivatives. This allows you to explore design spaces with hundreds or thousands of design variables very fast." https://openmdao.org/, [last accessed June 2020]

Organizational alignment in this sense relates to assumptions or even strategic considerations that may be implied in how a tradeoff analysis is conducted. For example, the scope of system concepts considered may be biased by organizational strategic interest, such as competitive advantages or workforce latent experiences. Such may be the case for a program oriented around platform development where the concepts considered (e.g. satellite) may be grounded or limited by organizational decisions (e.g. focus on geosynchronous systems due to past history or experience in developing such systems). The enterprise could impose or pass down these constraints to a program, or the enterprise could inherit this information across programs to gain situational awareness of perceptions at the program level, informing strategic planning activities. Additionally, commonality in representational models for inputs, outputs, and other models could help with economies of scope and scale for model and data reuse across programs. For example, multiple different Army programs could draw from a common representational model for how to parameterize ground vehicles for tradeoff studies, as well as common contextual uncertainties that were shown to be relevant in prior studies. Enterprise-level value models could inform strategic tradeoff decisions at the program level, where the "best" solutions are less likely to be locally optimized for the program and take into account cross-program considerations.

Another potential need for an enterprise-level repository of tradeoff analysis model packages is for maintaining historical records of past decision-making to support retrospective inquiries, such as accident investigations or defense of IP-related claims. Such a use case can not only provide ex post facto defense of actions, but also ab initio protection from repeating "past mistakes."

Formal prescription around what should be contained in tradeoff analysis model packages for an enterprise is still an open area of research. A number of factors should be considered in shaping the form of such prescription for a particular enterprise. These could include questions such as:

- What types of decisions were considered appropriate at the time of the study? (e.g. design decisions, operations decisions)
- What representational model was used in formulating the study and how reusable? (e.g. parameterization of an engineering design)
- What were the priorities and mechanisms for collapsing the output tradeoffs into decision advice? (e.g. decision models such as multi-disciplinary optimization or multi-attribute utility functions)
- What were the consequential assumptions/sensitivities in previous studies? (e.g. key contextual constraints such as budgets, available technologies and infrastructure)
- How to account for organizational competitive advantages and strategies? (e.g. we make "geo-satellites" or we build "wheeled ground vehicles")

Looking toward the future, curating tradeoff analysis model packages at the enterprise level likely would provide currently unrealized value because it could be leveraged to augment tradeoff analyses in future programs. The specific content of such packages and the criteria for determining whether such packages should be moved up to the enterprise level warrants further investigation and may be a valuable early example use case for the curation of model packages.

Looking toward the future, curating tradeoff analysis model packages at the enterprise level likely would provide currently unrealized value because it could be leveraged to augment tradeoff analyses in future programs. The specific content of such packages and the criteria for determining whether such packages should be moved up to the enterprise level warrants further investigation and may be a valuable early example use case for the curation of model packages.

Criteria for Placing Models under Curation

Digital engineering involves many types of models, such as descriptive system models, simulations, and tradeoff models. It remains an open question as to which of these would be retained in an enterprise-level model repository, as well as how these might be bundled or interrelated within a model repository. An enterprise model collection could include models for programs under development, models used by active programs in operations phase, models archived for historical or objective evidence purposes, reference models, surrogate models, demonstration models, and others.

The decision to place a model under enterprise-level curation will be driven by myriad factors. Some of the decision drivers include:

- Level of importance of the model to a major program
- Level of importance to the strategic business strategy
- Control and protection of enterprise IP
- Retention of model as objective evidence
- Controlled multi-program access and/or multi-enterprise access
- Enabling knowledge sharing and training

Once a decision to place a model under curation is made, specific criteria will be used to determine its readiness and acceptability. As experience is gained with enterprise model curation, the systems community can evolve a standard set of criteria.

This research project has defined a preliminary set of criteria for the purpose of trial use by early adopters. A model selected for enterprise model curation would be expected to have documented responses and evidence that satisfy the selection criteria, which would then be associated with the model in the enterprise model collection repository. At present, there are six criteria categories, each having associated sub-questions to be answered.

Initial criteria were adapted from criteria used by the digital engineering community (DCC, n.d.). These were then refined through testing on a surrogate program system model in collaboration with another SERC research project (WRT-1008). SERC developed a surrogate pilot for NAVAIR in support of systems engineering transformation (SET), with a focus on characterizing, assessing, and refining its SET framework, using an experimental system called Skyzer. (Blackburn et al., 2019 May). The research in phase I involving the UAS Skyzer system was furthered in Phase II adding a ship-based launch and recovery capability and a landing gear deep dive (Blackburn, M. et al., 2020 Feb). The criteria are applied to the Skyzer system model, and results can be found in (Blackburn, M., 2020 June). The process of developing responses for each of the criteria tested the clarity of the questions, as well as prompted additional ones.

The proposed preliminary criteria are shown in Table 2.

Table 2. Criteria for Placing Models Under Curation (adapted from: DCC)

Criteria for Placing a Model under Curation (preliminary)	
Relevance to the Enterprise and/or Program Mission	• *Is the model relevant to the overall enterprise mission?* • *Is the model relevant to specific current or future program mission?* • *Does model (including metadata, data, model representation, documentation) fall within the model collection/repository's scope?* • *Are there legal requirements or guidelines that require placing the model under curation?* • *Is there authoritative evidence of current value to engineering field?* • *Is there future value in having evidence of the model's use/reuse?*
Economic Business Case	• *Does benefit of placing model under curation exceed required cost?* • *Has the total cost of retaining the model package over active lifespan been considered?* • *Has the funding source for model retention and performance of curation activities been determined and agreed upon?* • *Have security and safety been considered in the economic case?* • *Has cost of archiving model after deaccession been considered?*
Completeness of Metadata, Data, Pedigree, and Documentation	• *Does model documentation span the lifecycle phases during which the model was conceived, generated and used?* • *Is the model metadata and pedigree information complete?* • *Is there sufficient documentation to support sharing, access and re-use of the model?* • *Is there sufficient information to judge the integrity and credibility of the model package?* • *Is there a sufficient set of data associated with the model to enable understanding and replication of model results?*
Potential for Redistribution, Reuse, and/or Repurposing	• *Are there any IP issues, data rights issues, human subject issues or restrictions that are not addressable?* • *Is there evidence of model reliability and usability?* • *Does the model have evidence of verification and validation?* • *Is the model package complete (model, data, metadata, documentation, digital artifacts, etc.)?* • *Has the data been stored in a way that ensures its integrity has not been compromised? Does the model meet standards and other technical criteria that allow its easy redistribution?*
Uniqueness of Model/ Non-Replicability	• *Is the model the only sole source of its content?* • *Can the model be easily replicated, recreated or re-measured?* • *Is the cost of replicating the model financially viable?* • *Is there historic value and/or education value for future workforce?*

It is expected that the criteria for placing a model under curation will be specific to the individual enterprise, as consistent with governance structure and practices. It is possible in the future that there could be community-level model repositories. As discussed in the Innovation section of this report, there is evidence of the value of such repositories in the systems biology community. Accordingly, converging on standard criteria for acceptance is important.

Enterprise Model Collection Object

Based on research to date, it is envisioned that there will be a set of elements that are associated with a model collection object. For the purposes of the report, the model with model-related data, digital artifacts and metadata is referred to as a "model package" and when accepted for the enterprise model collection it is termed a "curated model package." When the model package is accepted for the enterprise model collection, an accession record is created. Depending on implementation, a pedigree record may be independently linked to the curated model package. In the future, there may be other enterprise collection elements, such as "model consumer viewpoints." It is expected that the contents of the collection elements and terminology will be revised following additional research within the larger systems community. Figure 3 shows these notional model collection elements.

Accession Record	**Curated Model Package**
• Informational record of what model and when it was accepted into the enterprise collection (accession) and when it is removed from the collection (deaccession) • Retained as a permanent record of the enterprise regardless of status of the model • Specifies the curated model package and repository where currently available or archived • Linked to curated model package(s) and pedigree	• Bundled model set (model, metadata. associated data, digital artifacts) stored in a repository • Designated as active (in use) or archived • Version and baseline information • Linked to accession record and pedigree record • Retained as a permanent enterprise collection object • Has associated measures (maturity, value, etc.)
Pedigree Record	**Model Consumer Viewpoints (Future)**
• Informational record associated with accessioned model that includes originator, origination date, context of use, reuse/repurpose information • Retained as a permanent record of the enterprise regardless of status of model • Updated throughout the accession lifecycle through deaccession • Linked to (or part of) accession record • Associated with curated model package(s)	• Specifies reference viewpoints of model consumers based on defined personas and context of use • Used to dynamically generate displayed information for a model consumer viewpoint for models in the collection • Linked to curated model package

Figure 3 Enterprise Model Collection Elements (notional)

The curated model package has two linked records. The primary record is the "accession record" that documents the acceptance into the collection and links to the curated model package in the repository. The "pedigree record" is an informational record about the model's originators and other "non-technical" aspects. It may be implemented as part of the curated model package, but this depends on how collection objects are defined and managed by the enterprise. Model pedigree was first described by Gass & Joel (1980) as "model demographics" and the term pedigree was subsequently used by Gass. A pedigree contains the information about a model, its origins and use over time. Gass & Joel state the purpose is to "enable the decision maker to determine the model's status with respect to past achievements, theoretical and methodological state of the art, and the expert advice that went into its development" (Gass & Joel, 1980). While model documentation is typically developed, the pedigree may contain information not always included in model documentation. Model pedigree provides non-technical information concerning the model origins, extensions and applications. Given that SERC research on model-centric decision making has shown that trust is a key determinant in use of models (Rhodes, 2019a), a pedigree provides information that engenders trust.

A model collection is envisioned as having both active and archived objects. A curated model package accepted into a repository might be superseded by a new version, in which case the original version is archived. In such a case, the enterprise approach could be to retain pedigree information independently, but traceable to the original curated model package and the superseding active curated model package. This would avoid duplication of pedigree information if there is a superseding model. Research is needed to explore such practices to a greater extent, and it is possible that the approach to this would be enterprise-specific based on preference.

Beyond the scope of this project, additional investigation is needed to determine how data pedigree is handled in the case of an enterprise model repository. NASA has defined and uses data pedigree and input pedigree in Modeling & Simulations (M&S) efforts. NASA Standard 7009A cites data pedigree and input pedigree as two of eight credibility factors for Modeling & Simulation. Data pedigree is defined in the standard as "A record of traceability from the data's source through all aspects of its transmission, storage, and processing to its final form used in the development of an M&S" (NASA, 2016). NASA associates data pedigree with M&S Development, and input pedigree with M&S Operations. Input pedigree is defined as "A record of traceability from the input data's source through all aspects of its transmission, storage, and processing to its final form when using an M&S." Further, the NASA Standard notes that "changes from real world source data may be of significance to its pedigree" (NASA, 2016). It specifies four levels for data/input pedigrees in terms of credibility assessment levels. A model pedigree standard may benefit from using a similar scale. While data pedigree and input pedigree share some common content with model pedigree, there are unique aspects that must be addressed.

Also beyond the scope of this research project, there is a need to investigate the requirements for enterprise model repositories from a technology perspective, as well as from the model consumer's perspective. There are many questions to be explored, such as: Is an enterprise repository a scaled-up version of a program repository? Are there various levels of security and

IP protection required? What technology exists to implement the repository to enable ease of use and visualization of model information?

Looking to the future of curating model information, planned research investigates use cases for model consumers, used to define reference "model consumer viewpoints." These could be associated with the curated model package, and used to "curate" displayed information for a model consumer.

Implementation & Governance Guiding Principles

Investigation of foundational literature and prior research studies across multiple fields provided knowledge used to formulate guiding principles for model curation implementation and governance. These principles aim to inform development of organization-specific model curation policies, practices and operating procedures. The proposed principles should be considered to be preliminary and incomplete. It is anticipated that a series of workshops and follow-on activities will review and extend these over time as experience with model curation grows.

1. **The enterprise must establish and enforce policies for model accreditation, model accession, model valuation, model deaccession, and model archiving.**
    a. Models must undergo period enterprise-level valuation assessment to justify the cost of maintaining it in the active model collection.
    b. Model packages selected for deaccession from the active model collection must be archived and moved to a designated long-term retention storage option.
2. **The enterprise must appoint the enterprise authority who will be responsible for model curation leadership and governance.**
    a. Enterprise policy must specify the enterprise authority for model accession and deaccession.
    b. Enterprise policy must specify the enterprise authority for model accreditation and model valuation.
    c. Enterprise policy must specify the enterprise authority for acquiring models and loaning models from/to other enterprises.
    d. Enterprise policy must specify the enterprise authority for model curation infrastructure (repository, software, interfaces, etc.), including responsibility for the security, protection and maintenance.
3. **Enterprises must establish standard criteria and measures used to conduct, monitor and control model curation practice.**
    a. Enterprise policy must establish standard criteria and measures for model accession and deaccession.

   b. Enterprise policy must specify standard criteria and measures for model accreditation and model valuation.

   c. Enterprises must establish standard criteria and measures for monitoring and assessing model access and use.

4. **Model consumers need access to sufficient information to make decisions on making use of existing models in the collection.**

   a. Model consumers need access to sufficient information to judge suitability for use for a new purpose.

   b. Model consumers need access to sufficient information to judge suitability of using the model under a new operational context.

5. **Models must belong to one or more *curated model packages* (as defined by the enterprise), including model, associated data, metadata, pedigree, and accession record.**

   a. Models must be fully traceable to the data sets used, as well as history of use of the data sets.

   b. Models must be fully traceable to other models of relevance (archived baseline model, parent model, etc.)

6. **Model consumers must have access to information that allows credibility to be judged.**

   a. Model consumers must be able to know when a model originated and the subsequent time periods of model use, model modification, and model information updates

   b. Models must exhibit authoritative source credentials to model consumers.

MODEL CURATION AND AUTHORITATIVE SOURCE OF TRUTH (AST)

The Department of Defense (DoD) *Digital Engineering Strategy* outlines five strategic goals for digital engineering. DoD defines digital engineering as "an integrated digital approach that uses authoritative sources of system data and models as a continuum across disciplines to support lifecycle activities from concept through disposal" (DoD, 2018).

Alexander & Coleman (2018) define Authoritative Source of Truth (AST) as "… an entity such as a person, governing body, or system that applies expert judgement and rules to proclaim a digital artifact is valid and originates from a legitimate source." They state:

> *The authoritative source of truth for a digital artifact serves as the primary means of ensuring the credibility and coherence of the digital artifact that its creators share with a variety of stakeholders. It gives stakeholders from diverse organizations and distributed locations the authorization to access, analyze, and use valid digital artifacts from an authoritative source. The owners of digital environments or the community for digital engineering ecosystems provides stakeholders with an authoritative source of truth that assures confidence in the quality of the digital artifact across disciplines, domains, and life cycle phases.*

McDermott et al. (2018) provide a rich narrative discussion of AST using a Systemigram, envisioning AST as residing in a central data storage repository, providing "data and models that are instantly available to program personnel in the cloud server." The DoD Digital Engineering Strategy (2018) views governance of AST as involving responsibilities at various levels. Understandably, the focus of AST has largely been on program-level, which remains a significant challenge given program complexity, relative immaturity of digital engineering transformation, and lack of sufficient infrastructure. While model curation practices are implementable at the program-level, the greatest value is envisioned when implemented at the enterprise level.

Accordingly, this research places significant focus on the future enterprise, envisioning program-level models elevated to an enterprise collection of models with enterprise-level governance. In the prior phase of research (Rhodes, 2019b), the future role for enterprise-level model curation leadership was explored. The envisioned role and responsibilities are discussed in detail in Rhodes (2019a), along with seven proposed forms for enterprise level leadership. The particular organizational form for implementing an enterprise-level role may vary based on situational factors and the current state of transformation. Knowledge gathering and semi-structured interviews with executives informed the seven proposed alternative forms for an enterprise to execute the role and responsibilities of an enterprise model curator function. Future studies are needed to understand how effective these forms are for various enterprises, and under which conditions one might select the form (Rhodes, 2019b).

Validity and truthfulness of the digital artifacts in an enterprise collection depend upon effective governance, processes and controls as well as supporting infrastructure and environments.

Additionally, it is worthwhile to examine fundamentals of trust and credibility of models in the model curation context.

The return-on-investment business case for developing and maintaining enterprise model collections is not presently well-defined. What is clear is that a curated model collection is only worthwhile if there is evidence of model integrity and truthfulness, and sufficient information for model consumers to judge that a model is credible. In prior SERC research, an empirical study explored model trust. In this phase of the research, credibility was re-examined in context of model curation (Rhodes, 2020). Accepting curated model packages into a repository as authoritative source of truth is complex, and while the enterprise cannot fully control this it can provide enablers. One important aspect is the model consumer must have evidence that appropriate authorities have judged a model as acceptable for use for a specific purpose. Another is to use precursors and approaches to enhance perception of credibility.

Sources of Authority

For a source of truth to be authoritative, legitimacy must be conferred upon it from those who will acknowledge it as authoritative. This authority may be derived from an existing governance structure, or simply because those who will consent to the authority find it in their self-interest to do so. Governing authority may be assumed by diktat or granted by consent of those who will operate under that authority. In the case of models and an authoritative source of truth in an enterprise modeling environment, a pivotal question is by what means is the legitimacy obtained to declare models authoritative, and by what means is it obtained? Is a declaration of authority alone enough for it to be compelling to all participants? This will drive decisions in the means used to define authoritative sources so that the greatest proportion of enterprise participants accept and defer to that authority.

Authority in an enterprise can be subdivided into formal authority and informal authority. The formal authority is typically imposed top-down through an accepted hierarchy of roles and responsibilities. Someone is ultimately in charge, and delegates rights and decision authorities in an orderly fashion through the hierarchical structure of the enterprise. Those who participate in that enterprise give their consent to be governed by the defined formal structures and rules of that enterprise. If they are unwilling to recognize that authority structure, they may choose to exit unless they are constrained to continue to participate in the enterprise against their preferences (e.g., authoritarian enterprises control the behavior of the governed and make exit difficult or impossible).

Informal authority, on the other hand, is the legitimacy that enterprise participants confer to governance structures and authoritative sources at their discretion. This reflects the semi-autonomous nature of participants in many enterprises. They may choose to comply with the authority structure completely, somewhat, or not at all, depending on the degree to which they are compelled to conform (or the inverse, their level of agency) to the formal enterprise authority structure. Those who choose not to participate will likely exit the enterprise. Those who are compelled to remain must participate, albeit with potentially questionable true willingness to defer to the authority. Those who have discretion may or may not defer to some authorities, or may choose to defer to the authority at one time and not another. It may not be readily apparent who will or won't participate when the actors have agency, as they will likely affirm compliance when prompted by representatives from the authority structure.

The extent to which an authoritative source of truth is considered authoritative may then depend on the degree of agency and the alternatives available to some of the participants in the enterprise. Absent a clear and compelling mandate to defer to the authority (e.g., a military command structure, a contractual obligation, desire for continuing employment), actors with agency rely more on a variety of motivators to decide to accept the authoritative source of truth. Additional means will need to be employed to appeal to their interests to defer to and embrace the authority. These means include social norms imposed by fellow actors, economic rewards and/or sanctions to incentivize compliance, the offer of non-economic value like an improved process that produces less frustration in completing a task, selection of participants with personal intrinsic motivators aligned with the authority, and so forth. Other incentives to accept authority include norms around membership and standing in a community, for instance identity as an engineering professional or reputational status within an interest community.

Beneficial collaboration is another reason to embrace authority. Use of standards, codes, or rules facilitates work and has economic benefits. By adopting generally-accepted standards, producers can leverage the knowledge of the collective by externalizing the cost of maintaining best practices, reduce transaction costs with suppliers or customers, and network externalities or the benefit from being compatible with a larger ecosystem of products and services. The IEEE 802.11 standard, known more commonly as wi-fi, illustrates how embracing a common standard enables lower costs, easier development, and wider access to market opportunities for producers of electronic products than if a series of proprietary interface protocols were maintained by competing producers.

Examining Accreditation in a Model Curation Context

Models in an enterprise model collection will need to demonstrate clear evidence that they are deemed fit for purpose. A model consumer looking to use or reuse a model would be expected to look for evidence that an authority has judged this to be so. Depending on the type of model, the authority that authorizes model as fit for purpose could be a formal accreditation authority[3] or a less formal authority (e.g., an expert panel).

Accreditation falls into the category of beneficial collaboration. Accreditation is defined as the process in which certification of competency, authority, or credibility is presented. In this sense, accreditation represents an authoritative recognition of a capability. Accreditation is employed

[3] For example, Air Force Instruction 16-001 (2020, April 29) identifies the roles and responsibilities for accreditation of Air Force owned or managed models and simulations.

in a wide range of sectors and activities, including education, financial investment, healthcare, professional certifications, systems engineering, and electronic communications (e.g., email). In the DoD modeling and simulation (M&S) context, the accreditation of models and simulations is "the official certification that a model or simulation and its associated data are acceptable for use for a specific purpose" [DoDI 5000.61][4].

In this investigation of accreditation, it is assumed that the objective of the enterprise that might embrace M&S accreditation is not because it has intrinsic value, but rather because M&S accreditation is a way to achieve a higher organizational objective.

Accreditation is considered an optional gate at the end of the M&S verification and validation process. Some, but not all models and simulations are accredited, since not all models have an accreditation requirement (Cook & Skinner, 2005). It is common to refer to three processes together as model verification, validation, and accreditation (VV&A). The accreditation of a model is the responsibility of the sponsor of the modeling project, but is generally performed by an independent party to ensure an unbiased investigation. The authors state possible outcomes of the accreditation process are:

1. The model or simulation will be used as described.
2. The model or simulation will be used as described with limitations.
3. The model or simulation will be used as described with modifications.
4. The model or simulation requires additional V&V to be considered suitable for accreditation.
5. The model or simulation will not be used for this application. (Cook & Skinner, 2005).

The most common M&S accreditation recommendation is that the model will be used as described with limitations (Cook & Skinner, 2005, p.23). The limitations referred to in the statement reflect the limitations intrinsic to all models as approximations of reality. The resulting accredited model would then typically be stored in a repository of accredited models (Balci, 2012). The terms accreditation and certification relating to model VV&A are used interchangeably by some. Both imply that the certifying authority has authorized the model to be used for the purpose for which it was designed. Published sources consistently agree that VV&A should be performed through the M&S lifecycle (with accreditation being done when the model development effort is completed). That way, the documentation of the sources and assumptions used in the model can be captured and available for the accreditation process.

Air Force Instruction 16-001 (2020, April 29) is an example of the most recent guidance on M&S Verification, Validation & Accreditation (VV&A). It applies to all Air Force (owned or managed) models and simulations that qualify as federation elements, common-use, general-use, or joint modeling & simulation. It provides a VV&A framework, as shown in Figure 4, with associated VV&A steps to meet an M&S requirement based upon the existence and type of VV&A already accomplished for the application of that model or data for M&S.

[4] DoD Instruction 5000.61, "DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A)," USD(R&E), 02/09/2009.

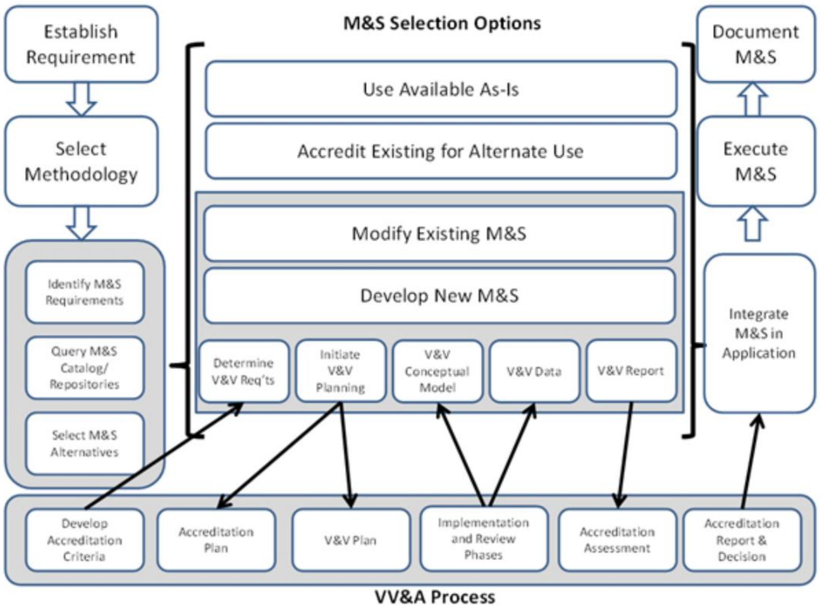

Figure 4 M&S Options and VV &A Steps (source: Air Force Instruction 16-001, 29 April 2020)

Formal accreditation practices in DoD have continued to mature and be enhanced for the future of digital engineering enterprises. Given enterprise model collections may include models that do not go through a formal process for accreditation, it is worthwhile to consider general requirements associated with accreditation across different fields.

General Requirements Associated With Accreditation

Accreditation is an important legitimacy-building mechanism employed in a number of different settings, with the terms accreditation and certification often being used interchangeably. For instance, in the finance sector, individuals are certified as financial planners or accountants, and audits are performed on public companies to ensure that their record keeping is in compliance with accepted accounting practices. Academic institutions are periodically accredited to ensure that their educational programs meet consistent standards and that an attainment at one institution is equivalent in scope to that of another. In the health care sector, hospitals and care facilities are accredited to demonstrate that they meet standards of care for their patients. Manufacturing facilities are certified to show that they comply with standards and procedures for quality (e.g., ISO 9000). The potential list of accreditation examples is long, but the foundation of the practice is to assess whether an entity or product is compliant with requirements or accepted standards of capability, attributes, or performance that have been established and assessed by independent third parties. Others rely on that accreditation declaration to make decisions about purchasing a product or service when there is incomplete or inconsistent information available about potential providers. Accreditation provides legitimacy or credibility; all accredited providers of a product or service are expected to offer roughly equivalent scope and quality consistent with the best practices in their domain, while those that are not accredited are unknown with possibly higher levels of variation in scope and quality.

Because accreditation is so important in a wide range of economic activity, general standards for the accreditation process have been created. Two important international standards for accreditation are:

- ISO/IEC 17065:2012 - *Conformity assessment — Requirements for bodies certifying products, processes and services*
- ISO/IEC 17067:2013 - *Conformity assessment — Fundamentals of product certification and guidelines for product certification schemes*

ISO/IEC 17067 addresses the process to be used in certifying a product (or process or service). The standard is informative rather that prescriptive, and provides guidelines for entities involved in product certification schemes. Certification schemes are the protocol or the requirements, specific rules, and procedures that are applied consistently to groups of people, products or services. The use of consistent schemes by impartial third party evaluators is designed to inspire confidence to consumers, regulators, industry and other interested parties that products conform to specified requirements.

A key part of accreditation is that independent third parties are performing the assessment and granting the certification of compliance. Where do these important third parties come from? ISO/IEC 17065 establishes the requirements for entities that provide certification. The requirements in the standard are considered as general criteria for certification bodies operating product, process or service certification schemes and are intended to ensure that certification bodies operate certification schemes in a competent, consistent and impartial manner. Among many requirements listed in the standard, these are relevant to the topic of M&S accreditation:

- Management of impartiality, including periodic assessments of risks to impartiality and associated mitigations
- Non-discriminatory conditions, including equal access to certification services
- Publicly available information, providing transparency into process and methods used
- Structural requirements, including organizational structure and management, and mechanisms for safeguarding impartiality
- Resource requirements, including having competent personnel and the financial resources needed to conduct the certification evaluation
- Process requirements, to include a certification scheme that covers activities from application for certification to evaluation to issuing certification products and surveillance

Another important element of the accreditation process outlined in the ISO/EIC standards is the body of requirements against which to evaluation compliance. Publications on accreditation of M&S mention the requirements by which models are evaluated only in passing, but the accreditation process is one of assessing compliance with predefined requirements. That pushes the accreditation requirements to a central role in the process, and raises a question about the nature of these requirements, how they are generated, and whether they are themselves part of a reuse strategy. One can find lengthy lists of V&V practices in the publications (see Balci, 1997; Cook & Skinner, 2005), which are no doubt helpful in the V&V process, but are not requirements. DoDI 5000.61, MIL-STD-3022, and NASA-STD-7009A reference the need for accreditation

requirements to be defined and to varying degrees provide templates for those documents, but defer to the system sponsor for the content of the requirements themselves.
In conventional system development, V&V requirements are derived from the system requirements. The system solution elements are verified to have been implemented in accordance with their design, and aggregated solution elements are validated to demonstrate that they deliver the intended results at their respective levels in the system.

For M&S, the models and simulations should accurately predict system behavior. At first order, V&V requirements are most likely to be system–centric. There may also be non-functional requirements for many systems, reflecting higher-order objectives for the sponsoring organization regarding system *ilities*. These may include reusability, extensibility, adaptability, etc. Are these additional requirements currently reflected in M&S accreditation requirements, either for the system they represent, or for the model(s) themselves? That is to say, should models that are accredited meet requirements other than just accurately predicting the behavior of the system? If there are additional requirements of M&S, for instance, stemming from enterprise-level guidelines regarding development of models and simulations, do these also appear in accreditation requirements? Do such requirements currently exist in an organized structure? The list of requirements to meet for a complex model or simulation could become quite long and involve more than just accurately predicting behavior of the immediate system.

The M&S project sponsor is designated as responsible for accreditation of the resulting models and/or simulations. Does that sponsor also have obligations to the larger M&S community or just to his or her immediate system? The currently available evidence suggest that it is the latter – the accreditation is focused on the immediate M&S project solution, and there is no or limited requirements input that reflect enterprise-level concerns or objectives. In addition to accreditation requirements potentially representing enterprise-level interests at the policy or strategy level, do they also accurately reflect the current state of technical reality and knowledge of the experts? An ontology-driven requirements process (OntoRem) has been developed in recent years and has shown potential for significant improvements in both quality and cycle time to manage system requirements (Kossmann & Odeh, 2010). The process is knowledge-driven and solicits domain knowledge from experts through an interactive and repeating process to construct ontological databases. These databases are then used to construct requirements. This approach has several potential benefits for the development accreditation requirements for M&S, based on the findings from the studies to date. First, since it relies on domain experts (who are likely to be M&S experts), it reflects a current awareness of what is possible, realistic, and doable. Second, since the experts themselves contributed to creating the ontological database, it may have enhanced credibility since it reflects their own knowledge, and may be easier to encourage them to use it and its products. Finally, since the construction of the ontological database is collaborative, it may foster relationships across discipline boundaries that may be beneficial to more effective M&S, including encouraging domain experts to embrace and enterprise-level perspective on M&S priorities.

Accreditation is a fairly straightforward extension of V&V that adds an evaluation and decision, assuming the evaluation requirements are clearly-defined. It does constitute an additional formal activity that would be included in a standard development lifecycle. The literature does not provide a clear description of how accreditation requirements are defined, and the extent to which they also include requirements that reflect the overarching objectives of the enterprise. The quantitative discussion of the benefits of accreditation are curiously not found in sources, so it is difficult to establish the ROI of adopting accreditation, although anecdotal claims are positive that the accreditation process yields benefits overall.

The picture of the mature operating model that emerges from the standards and descriptions describes the need to have accrediting agents that are impartial and competent, among other attributes. Autonomy, impartiality, transparency, competency, and non-discrimination are valued attributes. The typical model from other sectors is a fee for service model, usually as a stand-alone organization that provides accreditation services. An illustration would be a consultancy or extension service that can aggregate experience opportunities and that is able to keep a competent professional staff busy and productive. The accreditation process becomes somewhat more challenging as M&S begin to cross organizational boundaries in the extended enterprise. Within a unitary command structure, authority can be defined (as in an authoritative source of truth database, or a structure that defines compliance requirements for M&S projects) with the expectation that it will be honored. Outside of direct command structure (e.g., in the larger enterprise), people may have non-aligned motivations and will likely require additional incentives to recognize the stated authority structure.

Assuming the enterprise objective is model reuse, accreditation is potentially an enabler that alone is not sufficient except as part of a larger M&S transformation effort. Such an effort would include M&S artifact requirements such as reusability, adaptability, and so forth, that are independent of the functional requirements of the M&S project. Accreditation would be used as a means to assess and encourage the M&S behaviors that are necessary to enable the enterprise to achieve its overarching product development performance objectives. The solution will include an integrated process and product strategy with models that interoperate across the product portfolio based on a shared infrastructure. In a sense, the accreditation process serves as a final check of the model meeting the additional requirements imposed by that overarching strategy, and does not by virtue of being an accreditation process enable that. The strategy and imposed requirements must exist as well, but would be defined apart from the accreditation process. Most likely, the enterprise would need to employ a collective-oriented structure with the resources to develop and implement an accreditation agenda. The accreditation body is positioned to see across the collective of M&S project artifacts that must interoperate. The strategy, requirements, and necessary infrastructure would exist at the level above individual programs in order to coordinate. Design of governance and incentives are an important consideration since the enterprise may include participants from multiple organizations. IP policies will have to be established to prevent IP challenges becoming a barrier to achieving the overall objective and therefore negating the potential beneficial impact of accreditation.

A future area of research is to more deeply explore accreditation in context of enterprise model collections and model curation practices.

Model Credibility

Digital engineering is elevating the importance of models and their value to the enterprise, but only if they are trusted as a basis for engineering decisions. As enterprises begin to develop large model repositories, model credibility becomes a central concern and has major impact on program success. Recent studies show the decision to use, reuse and repurpose models is contingent on the model consumer's perception of validity and trustworthiness of the model.

During this research period, selected foundational works on credibility of models, simulations and websites were investigated to inform model curation by leveraging findings and strategies from prior work, and identifying useful heuristics (Rhodes, 2020). Model credibility and its associated constructs (e.g., model confidence, model trust, model validation, model value, etc.) have been investigated and discussed in the literature for more than four decades. The earliest works come from the operations research and simulation communities. Prior investigations of model and simulation credibility suggest there are actions that can be taken in model curation practice and enabling infrastructure that have potential to increase likelihood that a model consumer will perceive a model as trustworthy and valid. Early studies on website credibility were also investigated, given the websites at that time were dense and content-rich static information akin to digital artifacts, as contrasted with today's websites that ae highly visual and more interactive.

CREDIBILITY OF MODELS AND SIMULATIONS

Kahne (1976) proposed a new approach for examining model credibility for large-scale systems, asserting "model credibility is separated into two distinct, if not independent, issues: validity and value" (he assumes verification). He asserts credibility of a model will depend, among other things, upon "the quality of the match between the model and the model user", as a model reflects biases and outlook of the modeler. A novel viewpoint Kahne takes is "buyer/seller", with a subjective approach to credibility-type questions where credibility is defined as *capable of being believed* (Kahne, 1976).

The SCS Technical Committee issued a 1979 report on *terminology for model credibility,* motivated by the desire to develop a standard set of terminology to facilitate effective communication between the builder of a simulation model and its potential users, believed to be the "cornerstone for establishing the credibility of a computer simulation" (SCS, 1979). This SCS committee provided a framework to review credibility of a simulation (Figure 5), dividing the simulation environment into three basic elements: *reality, conceptual model*, and *computerized model*. Inner arrows describe processes (*analysis, programming, computer simulation*) that relate the elements to one another, while outer arrows refer to three procedures (*model qualification, model verification, model validity*) supporting credibility of the processes.

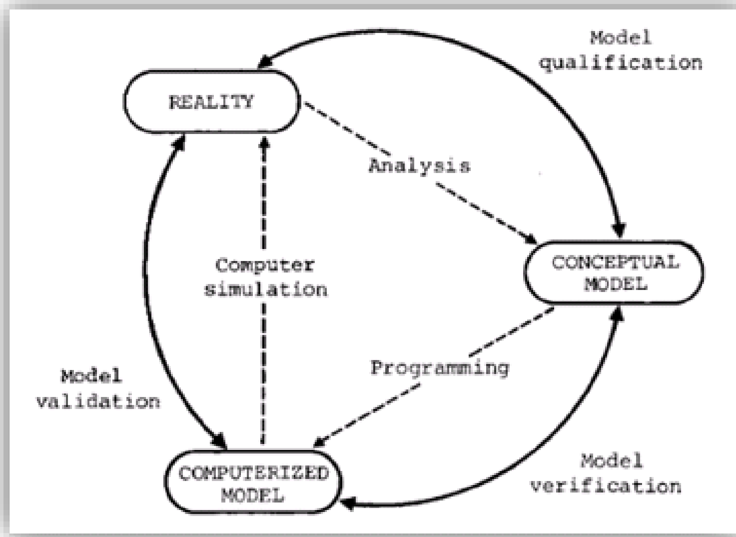

Figure 5 SCS Framework to Review Credibility of a Simulation [SCS, 1979]

Gass & Joel (1981) investigated the concepts of *model confidence*, showing model confidence to be not an attribute of a model, but of the model user. They propose seven confidence criteria: *model definition, model structure, model data, computer model verification, model validation, model usability,* and *model pedigree*. Model pedigree (originally called *model demographics*) is especially pertinent to perception of credibility. They state pedigree "should enable the decision maker to determine the model's status with respect to past achievements, theoretical and methodological state-of-the-art, and the expert advice that went into its development} (Gass & Joel, 1981). Gass (1993) states that critical to use of a model is "the credibility or confidence that the decision maker has in the model and its ability to produce information that would be of value to the decision makers" (Gass, 1993)

Balci (1986) proposed comprehensive guidelines for assessing credibility of simulation results, characterizing a life cycle of simulation study as richly characterized with 10 phases, 10 processes and 13 credibility assessment stages (Figure 6). He demonstrates that credibility assessment is complex and involves staged assessment through the lifespan of a model or simulation (Balci, 1986). Balci's work shows that during development, the acceptance of the model is a result of the model consumer's cumulative perception of validation efforts. Of importance to digital engineering, his work demonstrates the importance of giving a model consumer transparency into the series of validation activities that went into the original development, not only in the end result. Accordingly, it appears important to include such information in the model pedigree of a curated model.

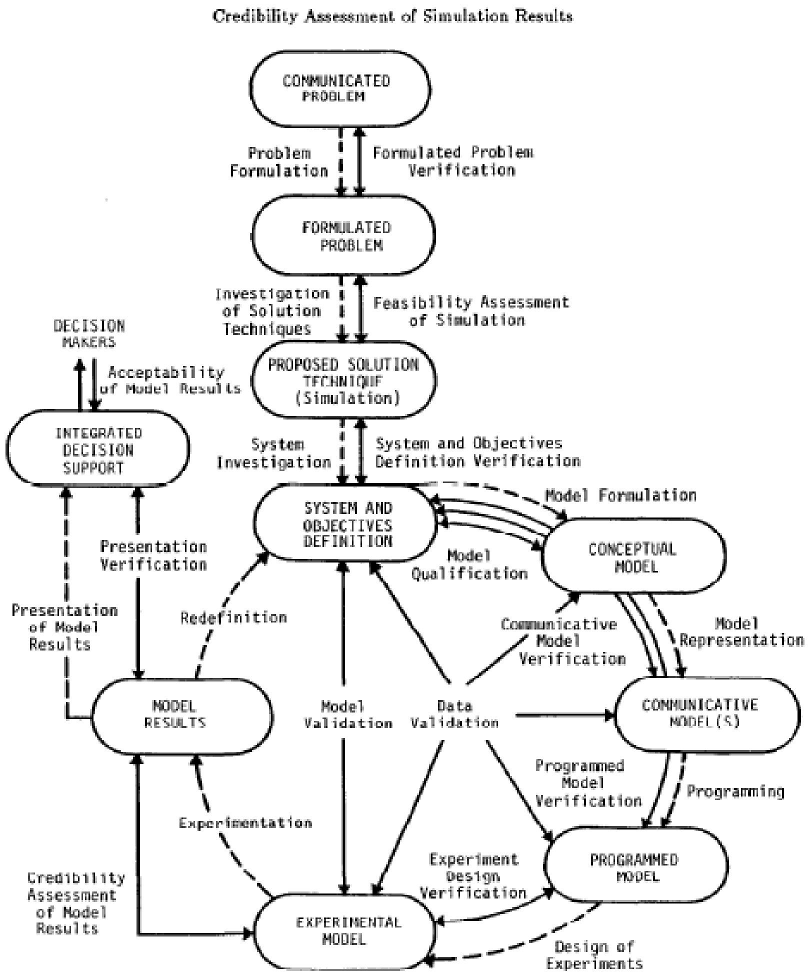

Figure 6. Balci's Life Cycle of a Simulation Study (Balci 1986)

ASSESSMENT OF CREDIBILITY OF MODELS AND SIMULATIONS

State of the practice on model credibility assessment in the systems field has emerged as part of the NASA efforts over more than a decade. Steele (2008) discusses insights and thinking behind NASA's standard for models and simulations (M&S). Eight relevant factors of credibility were identified during the development of this standard, which defines credibility as *the quality to elicit belief or trust in M&S results* (Steele, 2008). Evolution of the NASA standard raised various dimensions of credibility, and led to an assessment approach. A method for M&S credibility assessment is described in Appendix E of the 2016 update of *NASA Standard for Models and Simulations* (2016). [NASA, 2016, pp. 55- 72]. Ahn et al. propose a formal procedure based on the NASA standard to assess the credibility of an M&S in an objective way using the opinions of an expert group for credibility assessment and a Delphi approach (Ahn, et al., 2014), initially piloted on an M&S platform called SpaceNet by Ahn & de Weck (2007). Quantitative measurement of credibility continues to be an important topic of research (Olsen & Raunak, 2019).

Website Credibility

Early studies on website credibility (Fogg, et al., 2000) aimed to assess a broad range of elements that impact varying aesthetic, context, and technical factors on credibility of websites. Fogg et al. (2001) state “simply put, credibility can be defined as believability” and is a perception based on two factors: trustworthiness + expertise. They view credibility is a perceived quality; it does not reside in an object, person or piece of information, therefore when discussing credibility of a computer product, one is always discussing a perception of credibility (Fogg, et al., 2001).

Several findings of their work are insightful for model curation. First, web credibility was found to increase when users perceive a real world organization and real people behind a website. Second, small errors had a large negative impact on credibility of a website. Third, the users view websites as less credible if they experience technical problems (e.g. delays in download of information). Fogg states, “if users think a site lacks credibility – that the information and services cannot be trusted – they will abandon the site and seek to fill their needs in other ways” (Fogg, et al., 2001). This appears to suggest that a poorly designed model repository itself would be a barrier to successful model curation in an enterprise. Findings on website design guidelines appear to be extensible to design of model curation infrastructure and enablers.

Recent research on model confidence and trust

Research performed by Flanagan (2012) uses case studies and a web-based experiment to investigate key challenges to model-based design: distinguishing model confidence from model validation. The objective of her research is to understand factors that cause perception of model quality to differ from actual quality. She proposes eight factors as the key variables to misaligned model confidence, and tests hypotheses for six of these in the experiment to illustrate the effect of the factors on perception of model credibility.

According to Flanagan, these factors can potentially help explain behavior of decision makers, especially in the situation where “the model would be a good tool to help solve a problem; however, the decision-maker does not agree and continues without input from the model, effectively dismissing its predictions” (Flanagan, 2012). One hypothesis that was validated in the experiment, and is most relevant to model curation, concerns source and transparency of the model. This hypothesis is: *A more trustworthy model author and transparent governing equations will improve model perception*. Her finding was that for cases where the source was important to the decision, there was a significant difference in the decision outcome where untrustworthy sources caused reduced confidence.

Flanagan’s research, while preliminary, demonstrates the value of further research of this nature. These findings are consistent with findings of a more recent SERC sponsored empirical study by on model-centric decision making and trust German and Rhodes (2017). Model credibility was found to be a perceived quality, positively impacted by tailorable transparency and available model pedigree (i.e., detailing who originated the model, who subsequently enhanced the model, assumptions made, expertise of modelers, etc.). They also found that while not always needed, model consumers must have available model transparency when determining if a model should be trusted in making a specific decision.

De Vin (2015) provides a significant discussion on credibility of simulation, stating it is "influenced by three factors: credibility of the model, credibility of the data, and credibility of the model use." [De Vin, 2015, p. 152]. He notes that without credible data (also called Data Pedigree as discussed in the NASA Standard (2016), it will not be possible to perform trustworthy validation of the model. De Vin's paper "uses the NASA CAS model for credibility assessment of simulations to arrive at a schematic representation of how overall credibility as composed of aspect related to the model, the data, and the model's use" (De Vin, 2015). His schematic of factors is shown in Figure 7.

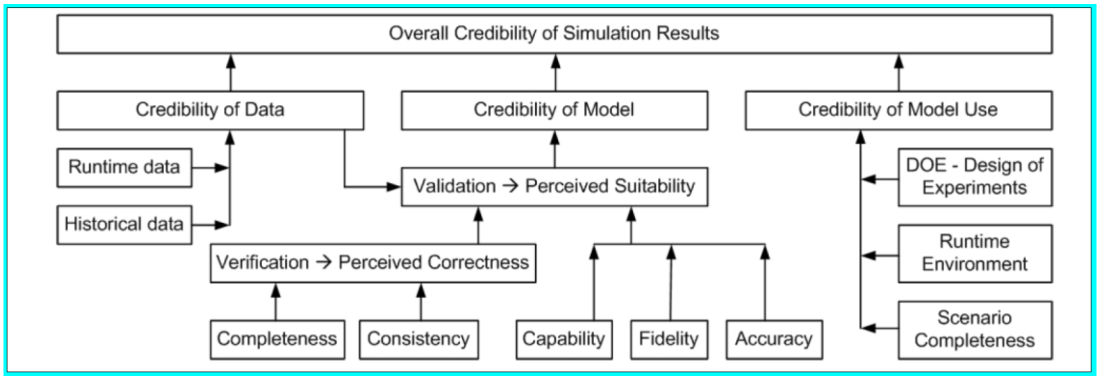

Figure 7 DeVin's schematic of factors influencing overall credibility of simulation results [De Vin, 2015 p. 154]

TOWARD DESIGN GUIDELINES FOR MODEL CURATION

Model curation requires both initial and ongoing investment, which must be outweighed by the payoffs (e.g., lifecycle model usability, re-purposing models for new contexts, re-use of models at enterprise level, etc.). This investment will fail to pay-off if model credibility is neglected in regard to the various aspects of model curation (e.g., practice, repositories, model discovery, etc.). Kahne's buyer/seller approach suggests the approach of thinking about model curation from a model acquirer (consumer)/model curator approach (Kahne, 1976). And, the model acquirer's perception of the expertise and authority of the model curator will have influence on perception of credibility. Model curation practices and infrastructure needs to be designed to support individual model consumer quality of experience, as this impacts perceived trustworthiness of the model. Model curation infrastructure (e.g., model repositories, interfaces for repository access, etc.) need to be designed for cost effectiveness and security, as well as for the quality of the experience of human interaction. As can be inferred by Fogg et al. (2000, 2003), a poorly designed model repository would negatively impact perceived credibility, as would a poorly designed user experience with model access and interaction with a repository. Design implications resulting from the website credibility studies of Fogg et al. may inform practical guidance for model curation. Specific implementation enablers could include markers of expertise and markers of trustworthiness (for example, as model pedigree information).

Precursors to Authoritative Source of Truth

Investigation of foundational literature and prior research studies across multiple fields provided knowledge used to formulate several precursors for authoritative source of truth, particularly as related to credibility of models. The proposed precursors should be considered to be preliminary and incomplete. It is anticipated that further investigation, empirical research and designed experiments will extend this set.

1. Model credibility is an attribute of the model consumer, not the model itself.
    a. Perceived credibility of a model is influenced by the expertise of the model consumer.
    b. Model credibility is influenced by a model consumer's trust in the expertise of the model originator.
    c. Perceived credibility of model is influenced by the model consumer's individual propensity for model trust.
    d. Acceptance of a model for (re)use is influenced by the complex interrelationship of credibility of model, credibility of data, and credibility of model use.
2. Perception of credibility can change over time through the model consumer's interactions.
    a. Model credibility is positively influenced by effective communication between modeler and model consumer, both active and passive.
    b. Model credibility is influenced by a model consumer's trust in the expertise of the model originator, as well as modelers who subsequently enhance and maintain the model over time.
    c. Credibility of a model is influenced by transparency into the cumulative assessment of interim activities during its development lifecycle.
    d. Model credibility is influenced by a model consumer's capacity for transparency into the validation activities throughout its development and enhancement.
    e. A model consumer's experience with ease of discovering and retrieving models from a repository influences perceived credibility of the model.
3. Perceived model credibility shifts with the model consumer's experienced context.
    a. Credibility of models in a collection is influenced by a model consumer's trust in the enabling infrastructure used for that collection.
    b. Credibility of the model is influenced by a model consumer's perception of expertise of the governance authority that accepted the model into the collection.
    c. Acceptance of a model for (re)use is influenced by a model consumer's belief that the model has the ability to produce information of current value to them.

Innovation

The DoD Digital Engineering Strategy (2018) articulates a goal of incorporating technological innovation to improve the engineering practice. Respective to this goal, this research has taken an initial look at what innovations are emerging from other fields that may benefit model curation. There are many areas that could be explored, including the application of data science, visual analytics, machine learning, natural language processing, UX design, augmented intelligence, and lightweight formal methods. Many potential innovations from fields such as data science, visual analytics, and software engineering could be applicable in curation of models.

Ongoing work in other fields such as BioModel curation and digital curation suggest interesting pathways to explore. Machine learning and feature engineering may provide means to automate tasks that curation would involve, which are presently done by humans. For example, Lee et al. (2018) investigate a means for a machine-learning assisted triage method to replace the manual curation of biomedical knowledge, presently done by humans querying and reading articles.

Curation of BioModels

Curation of models in the systems biology community has been ongoing for many years, in part driven by the highly collaborative nature of the field. There are multiple repositories that exist. The largest of these is BioModels, established in 2005, for computational models of biological and biomedical processes. BioModels was established to offer a platform to exchange published, peer-reviewed models between researchers across the globe.

According to Juty et al. (2015), “BioModels is a reference repository hosting mathematical models that describe the dynamic interactions of biological components at various scales … provides access to over 1,200 models described in literature and over 140,000 models automatically generated from pathway resources.” The authors noted that most model components are cross-linked with external resources to facilitate interoperability. A large proportion of models are manually curated to ensure reproducibility of simulation results. According to Malik-Sheriff et al., (2020), “In 2018, every month on average over 23,000 unique hosts accessed BioModels approximately 816,000 times, downloading 232 GB data from BioModels.”

Figure 8 shows two types of models: those that have been described in scientific literature, and those that are generated through automated processing of ‘pathway resources’ (a biological pathway is a series of interactions among molecules in a cell that leads to a certain product or a change in a cell). The models can vary widely across fields, which has prompted “efforts to standardize model encoding, interoperability, distribution, and reuse” (Juty et al., 2019).

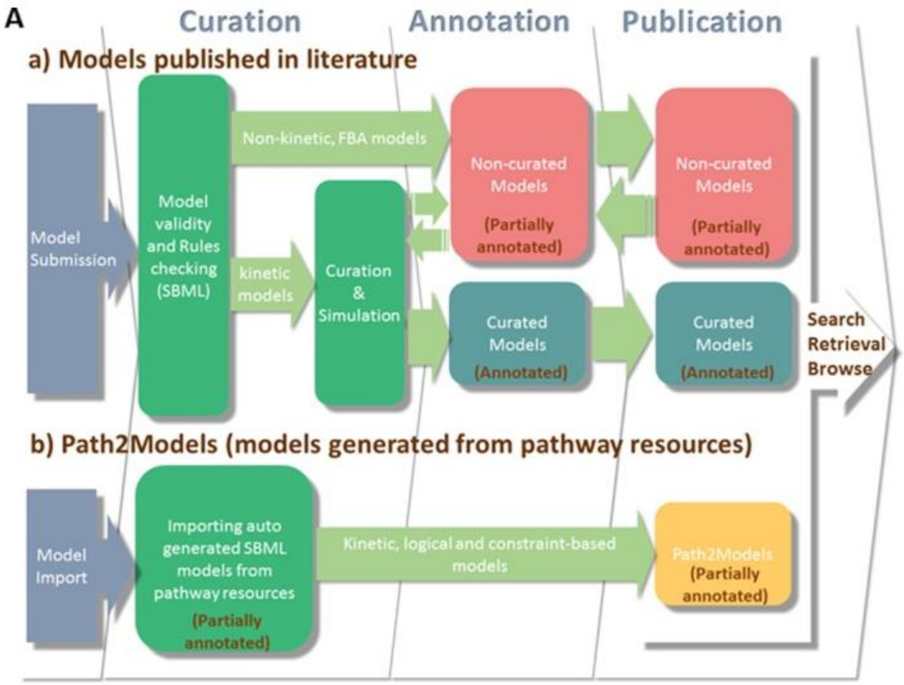

Figure 8 Two Branches of the Model Production Pipeline of BioModels (source, Juty et al., 2016)

While these biomodels are very different from engineering models, the experience in this field appears to suggest that the engineering community would be well-served by developing community conventions and standards that would support the goals for model collections of the future. This will be especially important if the future may involve open model repositories that are used across enterprises. Some of the approaches for automating curation and practices for model sharing in the biological field may be worth examining for general lessons and approaches to inform model repositories for digital engineering.

Display of Model Information

New technologies enable better display of information and more interactive visualizations. Xu et al. (2014) investigated the use of interactive visualization to provide curators with a means to analyze large-scale digital collections, resulting in the development of an interactive visual analytics application. In their work, the authors describe putting methods in place to summarize large and diverse information about the collection and to present it as integrated views. In the area of library science, Wissel and DeLuca (2018) investigate visualizations created from data, used as an aid to tell the story of a library collection. Research in the computer science community opens new possibilities for automatic extraction and methods to structure and analyze comprehensible text information (e.g., Natural Language Processing (NLP) for extraction; Lightweight Formal Methods (LFM) for structuring and analyzing).

Augmented Model Discovery

Large collections of models will require more than manual search. Model consumers will be interested in searching repositories to discover models of interest. Without some augmented capability this would not be possible in very large collections of models. Many opportunities will become available as approaches and tools for intelligent augmented search mature.

As model collections grow, there will be increasing need for searching model repositories, with use of technology to augment the process of discovering existing models. The DARPA Data-Driven Discovery of Models (D3M) program (Figure 9) has pioneered work in developing automated model discovery, and funded research projects to explore this topic.

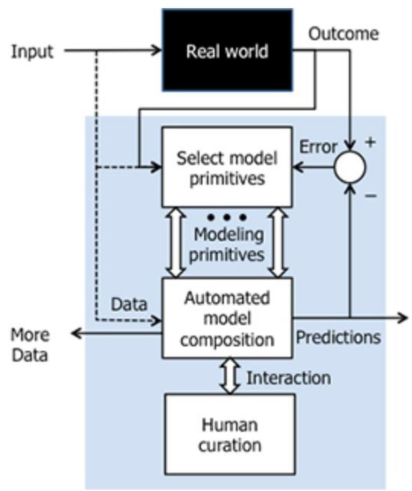

Figure 9 DARPA Data-Driven Discovery of Models (D3M) https://www.darpa.mil/news-events/2016-06-17

Opportunities to compose new models from existing models will become possible as data science advances automated model discovery. Digital engineering can benefit from the new technologies of the future. There is a need to examine composability methods in support of digital engineering, including investigating how models can be better designed to be discoverable by humans and augmentation technologies.

Reuse of models in model collection repositories is a desirable goal for model curation, with the vision that enterprises will in the future have large collections of models. Model composability research is essential to the future of model curation. As Petty & Weisel (2019) state, “Model composition and reuse are closely related. Composability is an important enabler for reuse, and reusing a software implementation of a model may require that it be composable”.

CONCLUSION AND FUTURE RESEARCH

Model curation is a broad topic that spans implementation of model curation practice; the roles and responsibilities of involved individuals and organizations; approaches to curate models for intended purpose and model consumer preference; and options for new technologies that enable curation and curating. Building on prior efforts, this research has further explored implementation practice and the relationship of model curation to authoritative source of truth. A preliminary set of guiding principles for governance and implementation practices were derived from research findings, along with precursors to credibility in the model curation context. Several areas of innovation are summarized as a look ahead to possible uses of newer technologies to advance curation beyond a largely manual endeavor.

Six areas of desired future research are highlighted below, including:

- Model curation interview-based empirical study
- Curating for model consumers
- Relationship of model credibility and data credibility
- Valuation of enterprise-level models/model collections
- Accreditation and enterprise model collections
- Digital demonstrators

Model Curation Interview-based Empirical Study

Over the past several years, the needs and requirements for model curation have been gathered based on literature and knowledge gathering in technical exchanges. As model curation moves forward, it will be important to have a set of needs and requirements that are gathered through a designed study to ensure consistency in questions and topics, diversity of stakeholder types and organizations represented, and use of qualitative research methods in analysis of gathered information. Results of a well-designed and executed research study could inform the formulation of implementation policies and practices, reveal barriers and enablers for curation, and provide better understanding to inform curation-related innovation. The recent SERC WRT-1001 on Digital Engineering Metrics (McDermott et al., 2020) has many useful insights to inform the design of a study.

Curating for Model Consumers

It is a widely acknowledged that the complexity and size of digital engineering models (as a function of system complexity/size) makes it impossible to "see" the entire model. Further, digital models are increasingly used by a more diverse set of stakeholders who aim to communicate through models. Research is needed to understand how model information can be accessed and displayed to a model consumer in such a way that it is useful for the intended purpose, as well as suited to the model consumer preferences, cognitive style, etc. Foundational and recent research could inform the approaches and initiatives for curating for the model consumer. This includes including generating model consumer personas for use in extraction and display of model information. Empirical investigation will be needed to elicit the information needed to develop use cases for model consumers, as well as construct personas. Newer technologies, such as machine learning, provide opportunities for implementation.

Relationship of Model Credibility and Data Credibility

Findings on model credibility in context of model curation indicate useful strategies and practices from other fields can be adapted for digital engineering, so continued knowledge gathering of legacy work may be beneficial. Data credibility has not yet been explored in context of model curation and is an area for future research. This should include the linkage between model credibility and data credibility, which could possibly be investigated through an experiment-based approach. Additionally, other aspects that can be investigated include credibility influences regarding information retrieval process, data curation, human-model interaction behavior, and augmented intelligence. Continued research on credibility is important given myriad negative impacts of untrusted models and data in an enterprise collection.

Valuation of Enterprise-level Models/Model Collections

As models become more valuable to enterprises, the question arises as to how to assess the value of a model. Potential approaches from other fields could provide consideration and approaches for valuation of digital engineering models and model collections. Significant work in the value of big data provides a foundation for work in the area, in addition to legacy approaches in valuation of institutional collections. In the future there is likely to be increasing acquisition and loan of models developed by one enterprise for use by another, raising the question of how compensation for this would be determined. Further, as enterprises begin to build up substantial collections of models, the overall value of the collection will need to be determined. This would be important, for instance, in mergers and acquisitions.

Accreditation and Enterprise Model Collections

The research suggests that one factor in whether model consumers will use/reuse a model is whether they see evidence of a trusted authority having deemed a model suitable for specific purpose. Examining accreditation processes and authorities specific to an enterprise model collection is an area of future research. Open questions remain that concern the role of accreditation in conferring credibility and legitimacy to enterprise model collection objects. Open questions include: What are the accreditation practices used for various types of models across different industry sectors? Can the value of accreditation be quantified? How is the decision to reuse or repurpose a model influenced by accreditation? What is the best approach to maximize acceptance and use of accredited models across the enterprise? From an enterprise transformation perspective, are there case studies that illustrate a successful implementation of accreditation and how did it evolve over time? In a rapidly changing environment, what is the most useful level at which to structure models and interfaces to ensure their reusability, and how does this impact the accreditation process?

Digital Demonstrators

A digital demonstrator is envisioned as generated using a technology-augmented approach that enables model discovery and model composability to produce a constructed digital engineering model that could be used to demonstrate novel system capability to model consumers. Such a demonstrator could be very valuable in research on new systems and technologies. It could enable better assessment of options in a bid and proposal activity. And, digital demonstrators could be useful in a process of analyzing whether existing models could be used to deliver a desired capability in rapid acquisitions. Subtopics of interest include model composability, model discovery, augmented consumer experience, application of technologies such as machine learning, use of simulations, and many others.

PROJECT TIMELINE & TRANSITION PLAN

The long-term transition goal for model curation research is to accelerate the implementation of effective model curation practice in the government. This will be accomplished through continued active technical exchanges with research stakeholders; workshops; development of reference practices and useful guidance material; and generating knowledge on enabling technologies for the government, as well as the broader DoD systems community. The WRT-1009 research project has addressed foundational knowledge and findings from initial interaction with research stakeholders. Continued research, empirical investigation, and observational studies are needed as parallels to early adopter implementation.

SERC planned and developed a workshop to engage stakeholders from across government and industry. This was scheduled for 30 April 2020 (see Figure 10 for flyer); due to the COVID-19 situation this was postponed and tentatively rescheduled for 19 November 2020.

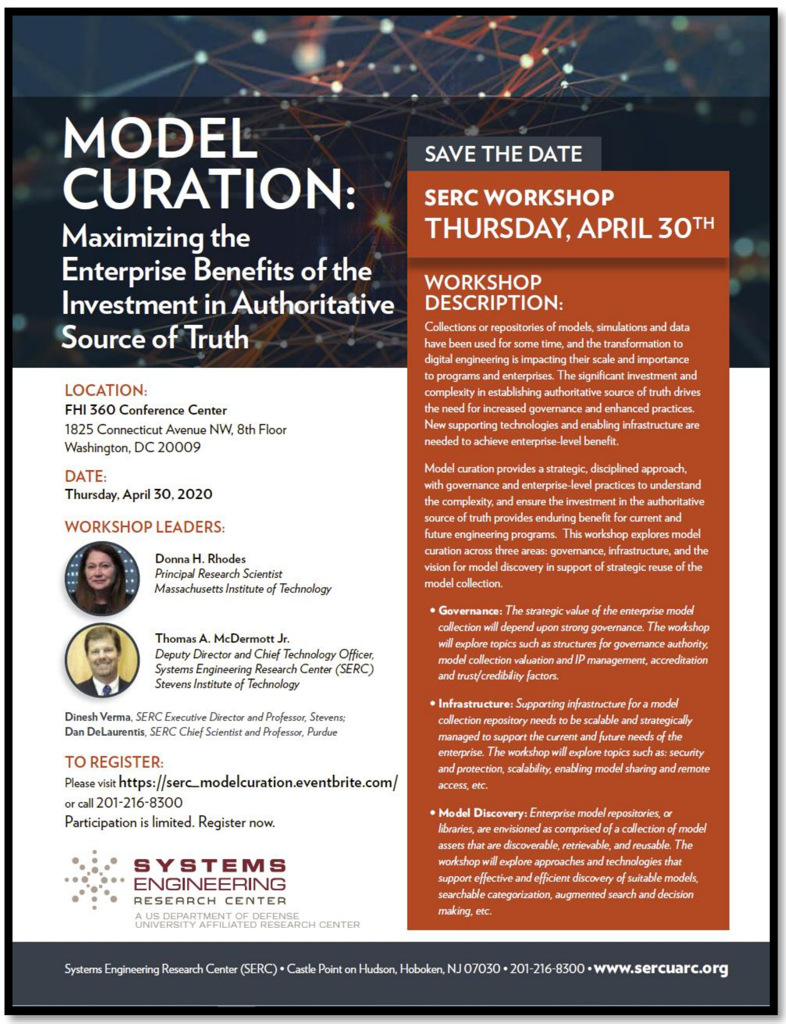

Figure 10 SERC Workshop on Model Curation (postponed to Nov 2020)

The SERC WRT-1009 project has included technical exchanges and interactions with government agencies that have been exploring and piloting model curation-related approaches and practices. It is expected that early adopters of the research in government will be DoD agencies with active digital engineering transformation initiatives. Cross-project collaboration between WRT-1009 and WRT-1008 has provided the opportunity for illustrative application of curation criteria, with potential for future use in education and training. Several publications and webinars have resulted from this research that have raised awareness in this topic and have resulted in useful feedback to further research goals.

The team has interacted with many stakeholders in industry and government during this research. Specific interaction in support of transformation relevant to model curation occurred through technical exchange meetings, workshops/forums, and working group support of several organizations, especially Army CCDC-Armaments Center, NAVAIR, SPAWAR Atlantic, Sandia National Laboratories, and SAF CMSO.

Table 3 summarizes the transition actions accomplished in this research, and identifies the respective SERC principles for successful transition.

Table 3 Transition Actions and SERC Principles Implemented

	Transition Action	Principles Implemented
1	Actively engaged with government stakeholders on relevant initiatives to establish model curation role/organization and practices.	• Plan Early • Pilot Continuously
2	Participation in government and industry working groups and forums to transition knowledge and findings.	• Engage Community • Pilot Continuously
3	Transitioned findings and principles through publications, forums and webinars, and obtained stakeholder feedback.	• Engage Community • Pilot Continuously
4	Planned and developed a SERC workshop to bring together stakeholders from government, industry and academia to share early initiatives and conduct breakout discussions to gather needs for both short and long term research.	• Engage Community • Balance Long and Short Term • Support Centrally

APPENDIX A: LIST OF PUBLICATIONS RESULTED

The following publication was completed during this research:

1. Rhodes, D.H. (2020). *Investigating Model Credibility within a Model Curation Context*. 18th Conference on Systems Engineering Research.

The following publications were completed in prior SERC research related to model curation:

1. Rhodes, D.H. (2019) *Model Curation: Requisite Leadership and Practice in Digital Engineering Enterprises.* 17th Conference on Systems Engineering Research.
2. Rhodes, D.H. (2018). *Using Human-Model Interaction Heuristics to Enable Model-Centric Enterprise Transformation*. IEEE SysCon.
3. German, E.S., Rhodes, D.H. (2017). Model-Centric Decision-Making: Exploring Decision-Maker Trust and Perception of Models. 15th Conf. on Systems Engineering Research. Los Angeles, CA.
4. Reymondet, L., Rhodes, D.H., Ross, A.M. (2016). *Considerations for Model Curation in Model-Centric Systems Engineering*, IEEE SysCon.

Appendix B: Cited and Related References

Ahn, J., de Weck, O, Steele, M. (2014). Credibility Assessment of Models and Simulations Based on NASA's Models and Simulation Standard Using the Delphi Method. *Systems Engineering*, 17 (2), pp. 237 – 248.

Ahn, J., de Weck, O.L. (2007). Pilot study: Credibility assessment of SpaceNet 1.3 with NASA-STD-(I)-7009, Tech Rep. MIT, Cambridge, MA.

Alexander, T. and Coleman, J. (2018, Dec). Authoritative Source of Truth. https://www.omgwiki.org/MBSE/doku.php?id=mbse:authoritative_source_of_truth

Balci, O. (1986). Credibility Assessment of Simulation Results. Proceedings of the 1986 Winter Simulation Conference.

Balci, O. (1997). Verification validation and accreditation of simulation models. In Proceedings of the 29th conference on Winter simulation (pp. 135-141).

Blackburn, M.R., Bone, M. Dzielski, J. Kruse, B., Peak, R., Cimtalay, S. Baker, A., Ballard, M., Carnevale, A., Stock, W., Ramaswamy, A., Austin, M., Coelho, M/, Rhodes, D., Rouse, W. (2020, June). Transforming Systems Engineering through Model-Centric Engineering. Final Technical Report WRT-1008 (NAVAIR).

Blackburn, M., R., Bone, M. Dzielski, J. Kruse, B., Peak, R., Cimtalay, S. Baker, A., Ballard, M., Carnevale, A., Stock, W., Ramaswamy, A., Austin, M., Coelho, M/, Rhodes, D., Rouse, W. (2020, Feb 1). Transforming Systems Engineering through Model-Centric Engineering. Interim Technical Report WRT-1008 (NAVAIR).

Blackburn, M.R., Bone, M., Dzielski, J., Kruse, B., Peak, R., Edwards, S. Baker,A., Ballard, M., Austin, M., Coelho, M., Rhodes, D., Smith, B. (2019, May 28). Transforming Systems Engineering through Model-Centric Engineering. Final Technical Report SERC-2019-TR-103, RT-195 (NAVAIR).

Cook, D.A., Skinner, J.M. (2005). How to perform credible verification, validation, and accreditation for modeling and simulation. The Journal of Defense Software Engineering, 18(5), 20-24.

DCC, Digital Curation Centre. website, http://www.dcc.au.uk.

De Vin, L. (2015). Simulation, Models, and Results: Reflections on their Nature and Credibility. Proceedings of FAIM 2015. (pp. 148-155).

Department of Defense. (2018, June). Digital Engineering Strategy. https://www.acq.osd.mil/se/docs/2018-DES.pdf.

DARPA. DARPA Goes “Meta” with Machine Learning for Machine Learning. https://www.darpa.mil/news-events/2016-06-17 (accessed Dec 2019).

Flanagan, G. (2012). Key Challenges to Model-Based Design: Distinguishing Model Confidence from Model Validation. Master’s Thesis. Massachusetts Institute of Technology.

Fogg, B.J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P. and Treinen, M., (2001). What makes Web sites credible?: a report on a large quantitative study. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 61-68). ACM.

Fogg, B.J., Marshall, J., Osipovich, A., Varma, C., Laraki, O., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P. and Treinen, M., (2000). Elements that affect web credibility: early results from a self-report study. In CHI'00 extended abstracts on human factors in computing systems (pp. 287-288). ACM.

Gass, S. (1993). Model accreditation: A rationale and process for determining a numerical rating. European Journal of Operational Research, 66(2), 250-258.

Gass, S.I. and Joel, L. (1981). Concepts of Model Confidence. Computers and Operations Research. 8(4), 341-346.

German, E.S., Rhodes, D.H. (2017). Model-Centric Decision-Making: Exploring Decision-Maker Trust and Perception of Models. 15th Conf. on Systems Engineering Research. Los Angeles, CA.

Kahne, S. (1976). Model Credibility for Large-Scale Systems. IEEE Transactions on Systems, Man and Cybernetics. Aug. 587- 590.

Kossmann, M., Odeh, M. (2010, July). 7.4. 3 Ontology-driven Requirements Engineering—A case study of OntoREM in the aerospace context. In INCOSE International Symposium (Vol. 20, No. 1, pp. 1000-1012).

Juty, N., Ali, R., Glont, M., Keating, S., Rodriguez, N., Swat, M. J., Wimalaratne, S. M., Hermjakob, H., Le Novère, N., Laibe, C., & Chelliah, V. (2015). BioModels: Content, Features, Functionality, and Use. CPT: pharmacometrics & systems pharmacology, 4(2), e3. https://doi.org/10.1002/psp4.3

Lailli, Y., Zhang, L., Yang, G. (2019). A Comprehensive Method for Model Credibility Measurement. In Model Engineering for Simulation (pp. 189-207). Academic Press.

Lee K., Famiglietti, M.L., McMahon A., Wei C-H., MacArthur JAL., Poux S., et al. (2018) Scaling up data curation using deep learning: An application to literature triage in genomic variation resources. PLoS Comput Biol 14(8).

Malik-Sheriff,R. et al. (2020, January 08). BioModels—15 years of sharing computational models in life science, Nucleic Acids Research, Volume 48, Issue D1, Pages D407–D415

McDermott, T., Hutchison, N., Clifford, M., Van Aken, E., Salado, A., Henderson, K. (2020). Digital Engineering Metrics, WRT-1001. Systems Engineering Technical Report, SERC-2020-SR-001. Systems Engineering Research Center, Hoboken NJ.

McDermott, T., Collopy, P., Paredis, C., Nadolski, M. (2018). Enterprise System-of-Systems Model for Digital-Thread Enabled Acquisition. Technical Report SERC-2018-TR-109.

NASA. (2016). NASA-STD-7009A w/CHANGE 1, Standard for Models and Simulations https://standards.nasa.gov/standard/nasa/nasa-std-7009..

Olsen, M., Raunak, M., (2019). Quantitative Measurements of Model Credibility. In Model Engineering for Simulation (pp. 163-187). Academic Press.

OMG. MBSE Wiki, Digital Engineering Information Exchange Working Group (DEIX WG). http://www.omgwiki.org/MBSE/doku.php?id=mbse:deix.

Patel, M. Ball, A., Ding, L. (2009). Strategies for the Curation of CAD Engineering Models. The International Journal of Digital Curation. Issue 1. Vol 4.

Petty, M. D., Weisel, E. W. (2019). Model Composition and Reuse. In Model Engineering for Simulation (pp. 57-85). Academic Press.

Reymondet, L., Rhodes, D.H., Ross, A.M. (2016, April). Considerations for Model Curation in Model-Centric Systems Engineering. 10th Annual IEEE Systems Conference. Orlando, FL.

Rhodes, D.H. & Ross, A.M. (2015). Interactive Model-Centric Systems Engineering Pathfinder Workshop Report. MIT, Cambridge, MA.

Rhodes, D.H. (2018a). Using Human-Model Interaction Heuristics to Enable Model-Centric Enterprise Transformation, IEEE Systems Conference. Vancouver, CA.

Rhodes D.H. (2018b). Interactive Model-Centric Systems Engineering Technical Report, Phase 5. (No. SERC-2018-TR 104). Systems Engineering Research Center Hoboken NJ.

Rhodes, D.H. (2019a). Model Curation: Requisite Leadership and Practice in Digital Engineering Enterprises. 17th Conference on Systems Engineering Research, Washington, DC.

Rhodes D.H. (2019b). Interactive Model-Centric Systems Engineering Technical Report, Phase 6. (No. SERC-2019-TR 003). Systems Engineering Research Center Hoboken NJ.

Rhodes, D.H. (2020). Investigating Model Credibility within a Model Curation Context. 18th Conference on Systems Engineering Research.

SCS (1979). Terminology for Model Credibility. Reports of the SCS Technical Committees.
Steele, M. J. (2008, June). Dimensions of credibility in models and simulations. Proceedings of the 2008 Summer Computer Simulation Conference (57). Society for Modeling & Simulation International.

Wissel, K. DeLuca, L. (2018) Telling the Story of a Collection with Visualizations: A Case Study, Collection Management, Taylor & Francis Online https://www.tandfonline.com/doi/full/10.1080/01462679.2018.1524319 accessed Jan 5, 2018.

Xu, W., Esteva, M., Jain, S. D., Jain, V. (2014). Interactive visualization for curatorial analysis of large digital collection. Information Visualization, 13(2), 159–183

APPENDIX C. MODEL CURATION LEXICON

The table below shows a model curation lexicon, which continues to evolve through research investigation and discussions with research stakeholders. Terms are defined as used in context of model curation.

Additional work is needed to complete and validate the usefulness of the terms.

Table A.1 Model Curation Lexicon (preliminary)

Term	Definition (in development)
Digital Demonstrator	A composed set of models with interactive interfaces for the purpose of demonstrating context-specific capability. Demonstrators enable a modeled system to be experienced by a stakeholder (model consumer) through conveying cogent information.
Model Accessioning	The formal process of accepting and recording a model as a collection object in the enterprise level model portfolio. Accessioning addresses, the legal, IP, data rights, and ethical issues.
Model Acquisition	The act of acquiring a model through an arrangement with the model owner (e.g., through purchase, trade, or other business transaction).
Model Cataloging	The formal process of making a model available for use through recording it in a catalog or directory, and tracking it throughout the model lifecycle.
Model Collection	The collection of model-based assets that is possessed by an enterprise, including those developed by the enterprise, acquired by the enterprise, and temporarily resident in the collection (e.g., leased, on loan).
Model Collection Object	A model or model-related object that is a unique asset in the enterprise's collection. An object is assigned a unique identifier.
Model Composability	The characteristic of an interrelated set of models that enables them to be combined in accordance with given modeling formalisms.
Model Composition	The process of composing models and model-related information that provides collective value beyond the individual models.
Model Consumer	An individual or group who uses a model for a specific purpose related to their need for information
Model Curator	A designated professional role entrusted with the ownership, tracking and use of model collection objects, and possessing designated authorities for managing and controlling models.
Model De-accessioning	The formal process of removing a model as a collection object from the enterprise level model portfolio.

Model Loan	The act of temporarily loaning a model through an agreement whereby the model owner agrees to share the model with the model acquirer for a specified time and specified terms (e.g., terms of use, remuneration, etc.).
Model Metadata	Descriptive metadata is contextual data about the model object(s). Metadata documents characteristics and used for indexing, discovering, identification. Provides discovery, access to, and management of an object.
Model Pedigree	Model-associated information that describes model origin, development process, originators and developers, assumptions, expert knowledge, model enhancements, investment costs, versions, change history, etc.
Model Repository	A central location where models are stored and managed.