

18th Annual Conference on Systems Engineering Research

October 08 - 10, 2020

Investigating Model Credibility within a Model Curation Context

Donna H. Rhodes

Massachusetts Institute of Technology

rhodes@mit.edu



Model Curation Context

Early Research

Last Decade - Research

Heuristics

Enablers & Precursors

Looking Forward

Model Curation

Curation practices promote formalism and provide for governance of models (and digital artifacts)

Motivations for curated model collections

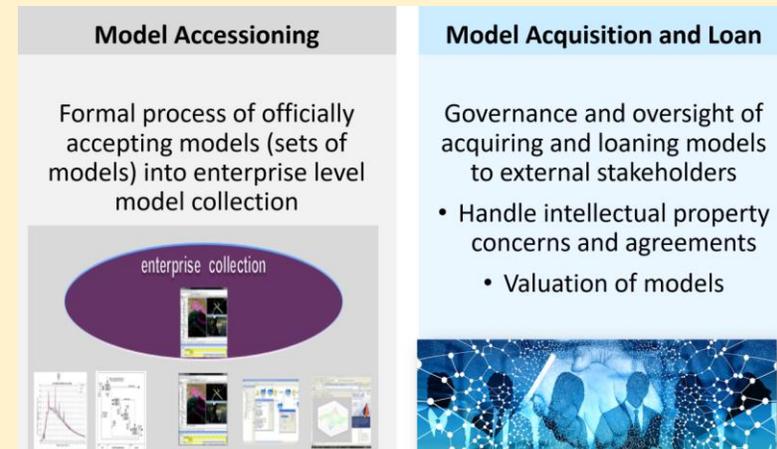
Models are *increasingly valuable*, some are used throughout system lifespan

Significant *IP challenges* but model sharing is imperative

Models seldom *managed as enterprise collection*

As model collections grow, *model discovery becomes difficult*

....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

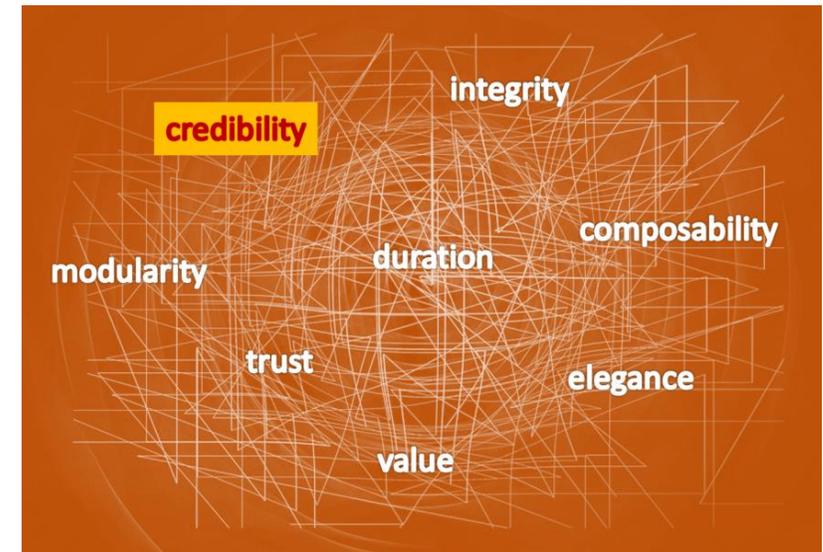


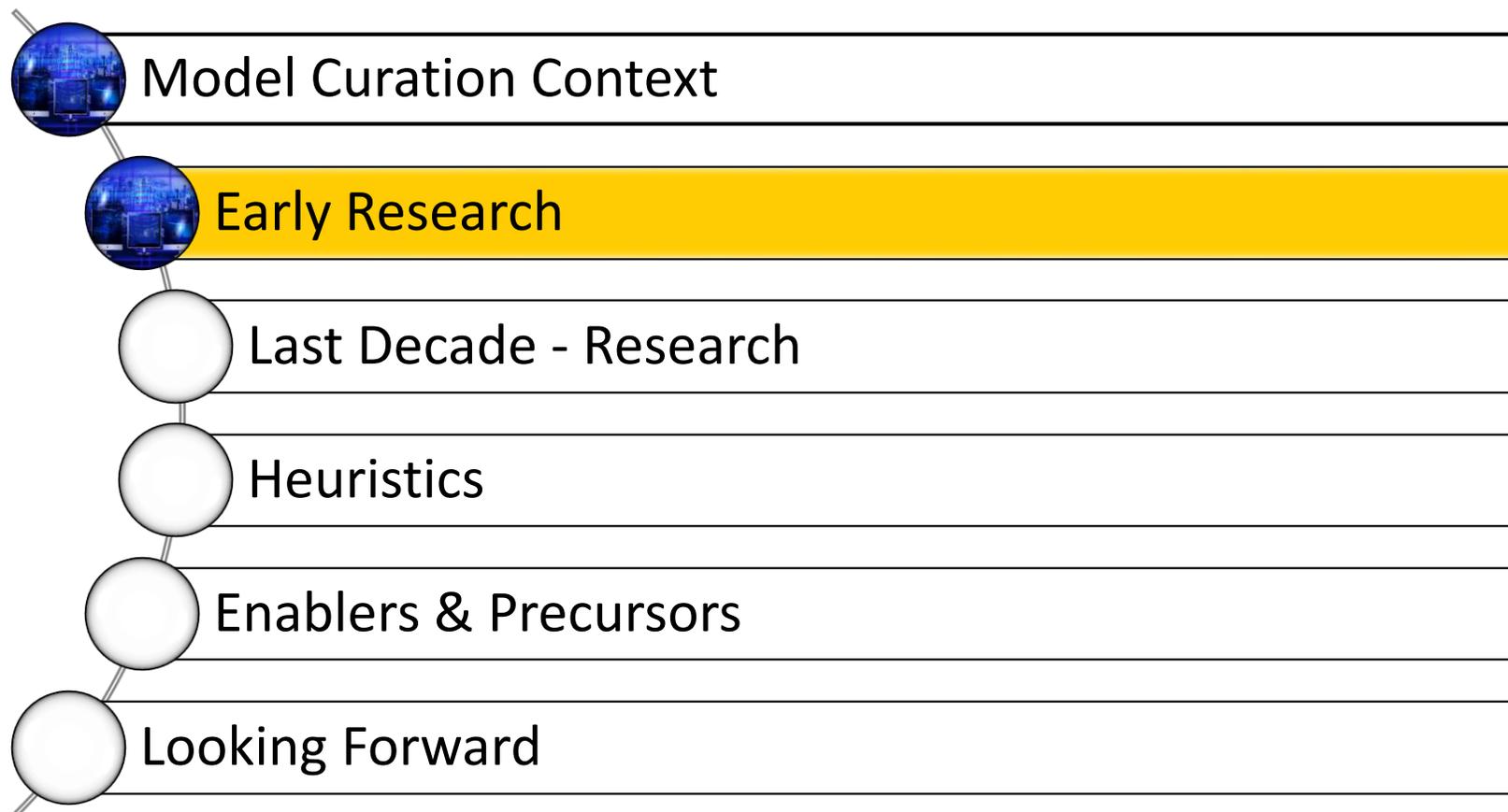
Sharing results of research investigation...

... motivated by the belief that *foundational ideas and approaches for model credibility are worth re-examining* respective to a model curation context

Model credibility is a construct that underpins facets of model curation:

- curation practice
- requisite leadership
- supporting infrastructure
- curating for model consumer needs
- innovations to enhance model discovery





Kahne (1976) proposed a new approach for examining model credibility for large-scale systems

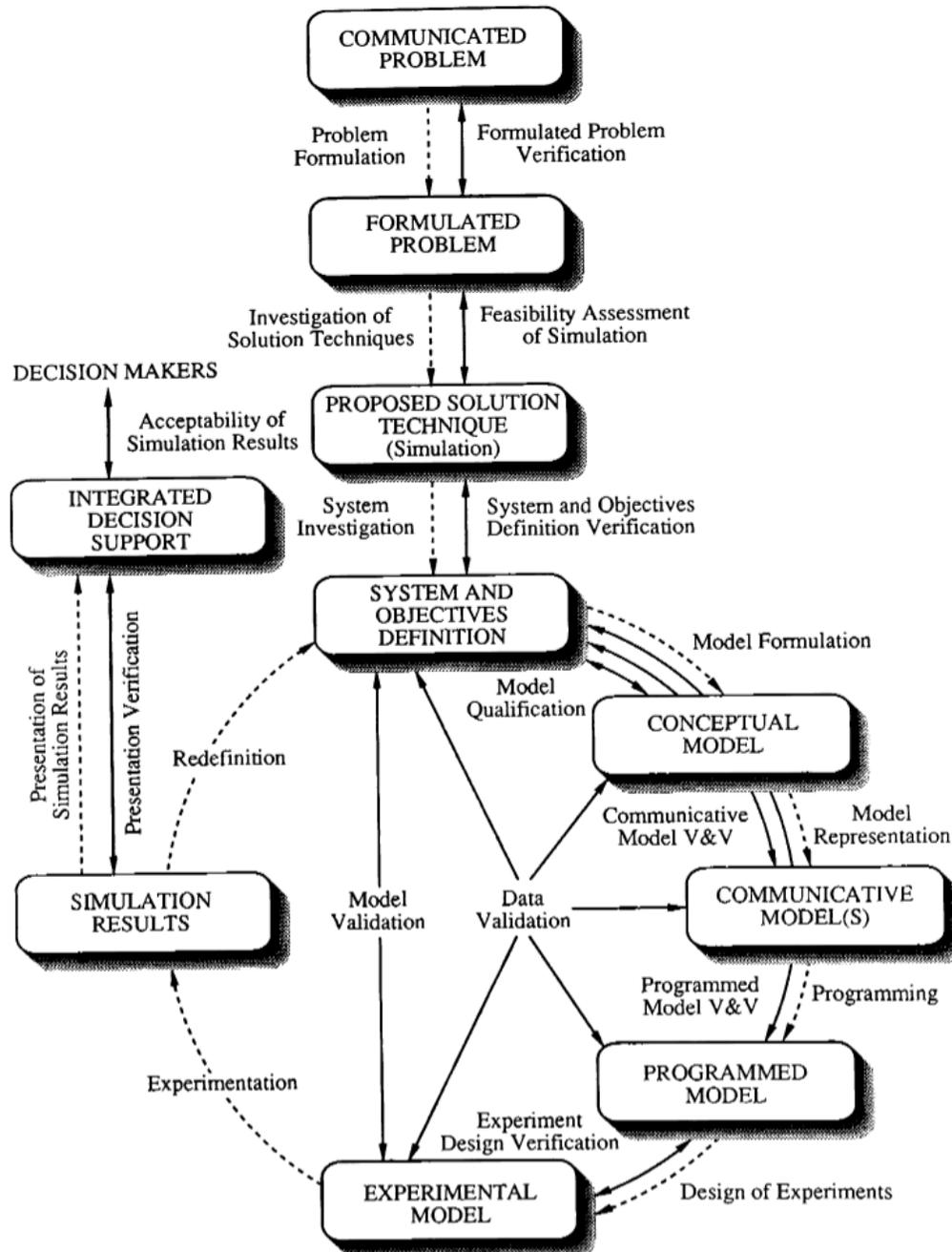
- Asserts model credibility is **separated into two distinct**, if not independent, issues: **validity and value** (verification is assumed)
 - *...to be a bit more specific, we are contrasting verisimilitude (having the appearance of truth) with worth (value)*
- Credibility of a model will depend, among other things, upon the **quality of the match** between model and model user
 - **model reflects the biases and outlook of the modeler**

A novelty of his approach is to take viewpoint of **buyer/seller**, with a subjective approach to credibility-type questions (credibility is defined as *capable of being believed*)

Gass & Joel (1981)

- Model confidence **not an attribute of a model**, but of the model user
- **Seven confidence criteria** proposed are
 1. model definition
 2. model structure
 3. model data
 4. computer model verification
 5. model validation
 6. model usability
 7. **model demographics** (later called model pedigree)

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”



- Balci (1986) proposed comprehensive guidelines for assessing credibility of simulation results
 - Characterizes a life cycle of simulation study as 10 phases, 10 processes and **13 credibility assessment stages**
 - Balci's important work demonstrates that **credibility assessment is complex and involves staged assessment** through the lifespan
 - Demonstrates **acceptance of the model is a result of the model consumer's cumulative perception of validation efforts**
- Suggests **importance of giving a model consumer transparency into the series of validation activities that went into the original development, not only the end result**
- Original paper: Osman Balci. 1986. Credibility assessment of simulation results. In *Proceedings of the 18th conference on Winter simulation (WSC '86)*. Association for Computing Machinery, New York, NY, USA, 38–44.
- **Modified paper:** Balci, O. (1990). *Guidelines for successful simulation studies*. Institute of Electrical and Electronics Engineers (IEEE).

Fogg et al. (2001)

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**
- Several findings insightful for model curation:
 - Web credibility found to **increase when users perceive a real world organization and real people behind a website**
 - **Small errors had a large negative impact** on credibility
 - Users view websites as **less credible if they experience technical problems** (e.g. delays in information download)

“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 “

An early study (Fogg, et al., 2000) aimed to assess a broad range of elements that impact varying aesthetic, context, and technical factors on credibility of websites

Websites at the time of the studies were dense and content-rich static information (contrasted with today’s highly visual, more interactive experience)

As such, this is an informative body of work for digital engineering in that the websites in the early website credibility studies are more akin to digital artifacts

NASA

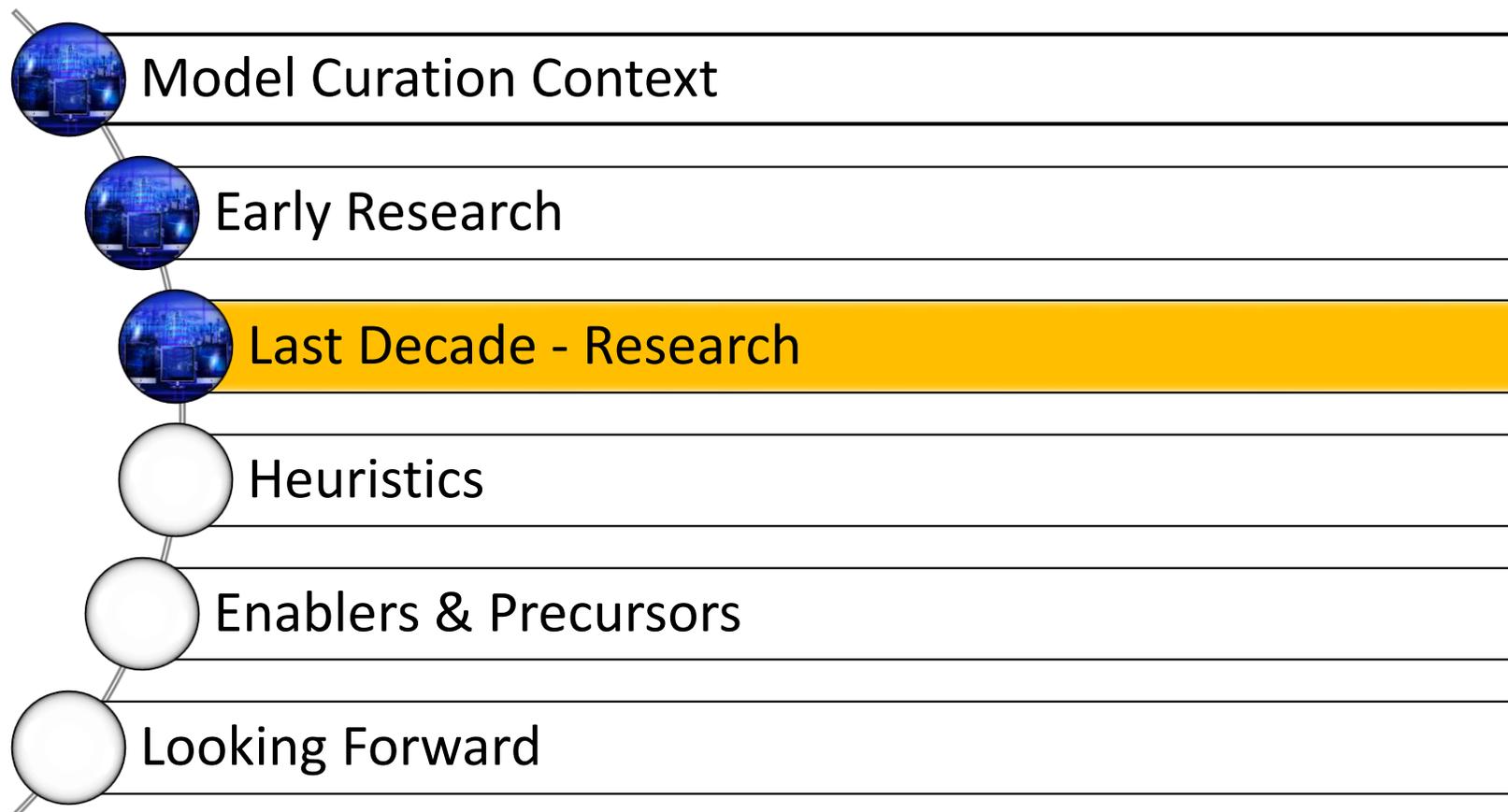
defines credibility as the quality to elicit belief or trust in M&S results

Steele (2008) reveals the insights and thinking behind the NASA's standard for models and simulations (M&S)....

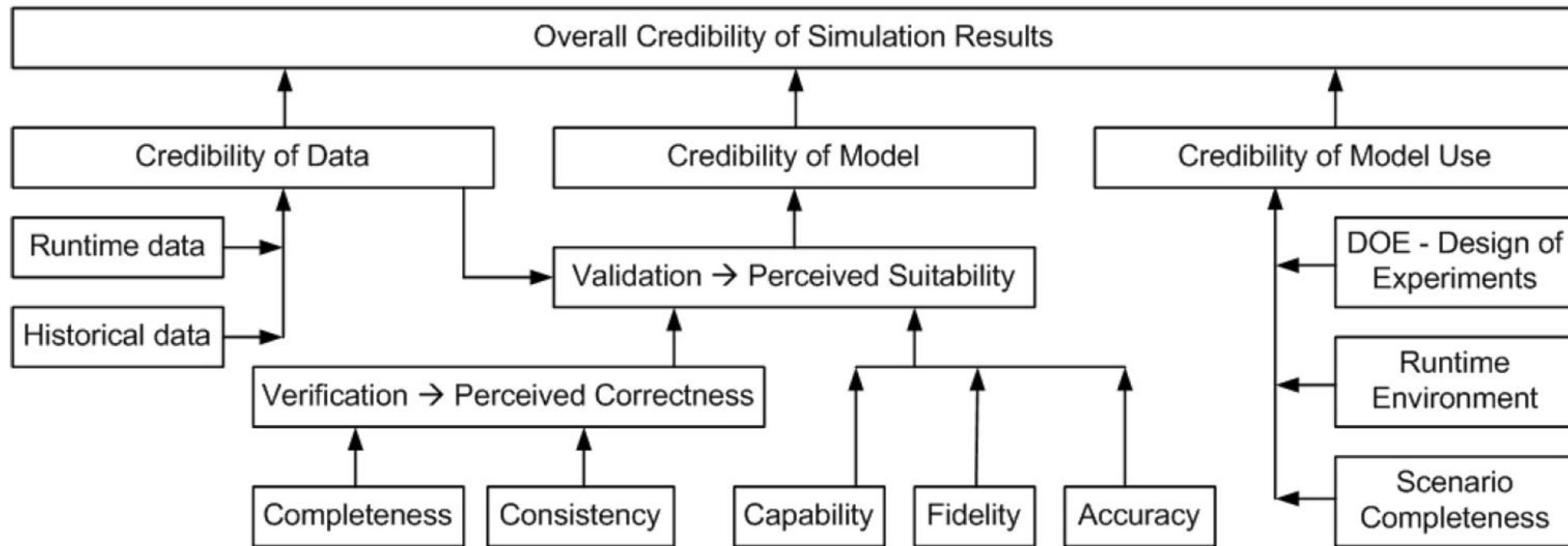
- **Eight relevant factors** of credibility were identified during the development of the standard
- Evolution of the NASA standard surfaced various dimensions of credibility, and more recently an assessment approach

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

A method for M&S credibility assessment is described in Appendix E of the 2016 update of ***NASA Standard for Models and Simulations (2016)***



De Vin's Schematic of Factors Influencing Overall Credibility of Simulation Results



Recent work by De Vin (2015) provides a significant discussion on credibility of simulation, stating it "...is thus influenced by three factors: Credibility of the model, credibility of the data, and credibility of the model use." He notes that without credible data (also called Data Pedigree as discussed in, it will not be possible to carry out 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"**.

Flanagan (2012)

Objective: understand factors that cause perception of model quality to differ from actual quality

- Uses case studies and a web-based experiment to investigate key challenges to model-based design: **distinguishing model confidence from model validation**
 - She proposes **eight factors as the key variables** to misaligned model confidence, and tests hypotheses for six of these to illustrate effect on perception of model credibility.
- Flanagan: 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”**
- One of the hypotheses validated in the experiment, **a more trustworthy model author and transparent governing equations will improve model perception**
 - ***Her finding: for cases where the source was important to the decision, there was a significant difference in decision outcome where untrustworthy sources caused reduced confidence.***

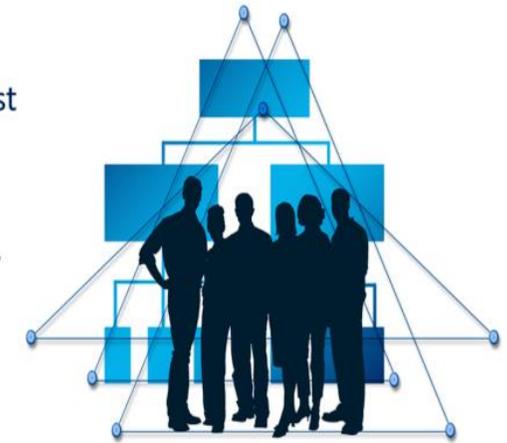
German & Rhodes (2017)

Interview-based study on model-centric decision making and trust

Voice of systems engineering experts:

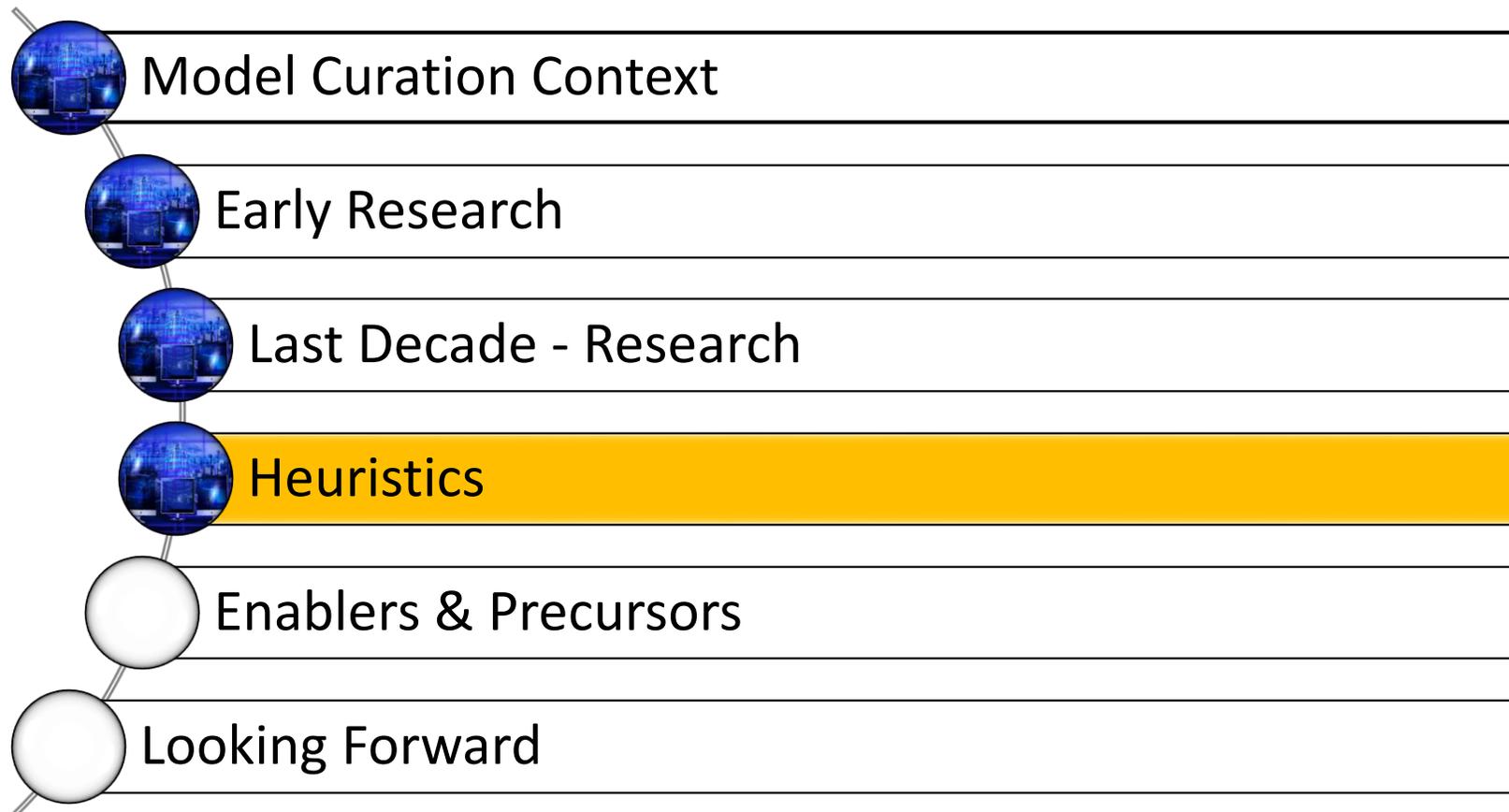
- Model credibility was confirmed to be a perceived quality
- Positively impacted by tailorable transparency and “model pedigree”
- Model consumers must have available model transparency when determining if a model should be trusted in making a decision
- Consistent with Flanagan’s findings

Three actor decision flow
 Importance of intercommunication
 Understanding of assumptions and uncertainty
 Technological and social factors influencing trust
 Importance of model-related documentation
 Need for model pedigree
 Using models as primary versus supplementary
 Non-advocate role in reviews
 Transparency and trust
 Model investment bias and confirmation bias
 Factors limiting model-centric decisions
 Real-time interaction with models
 Viewing humans as endogenous



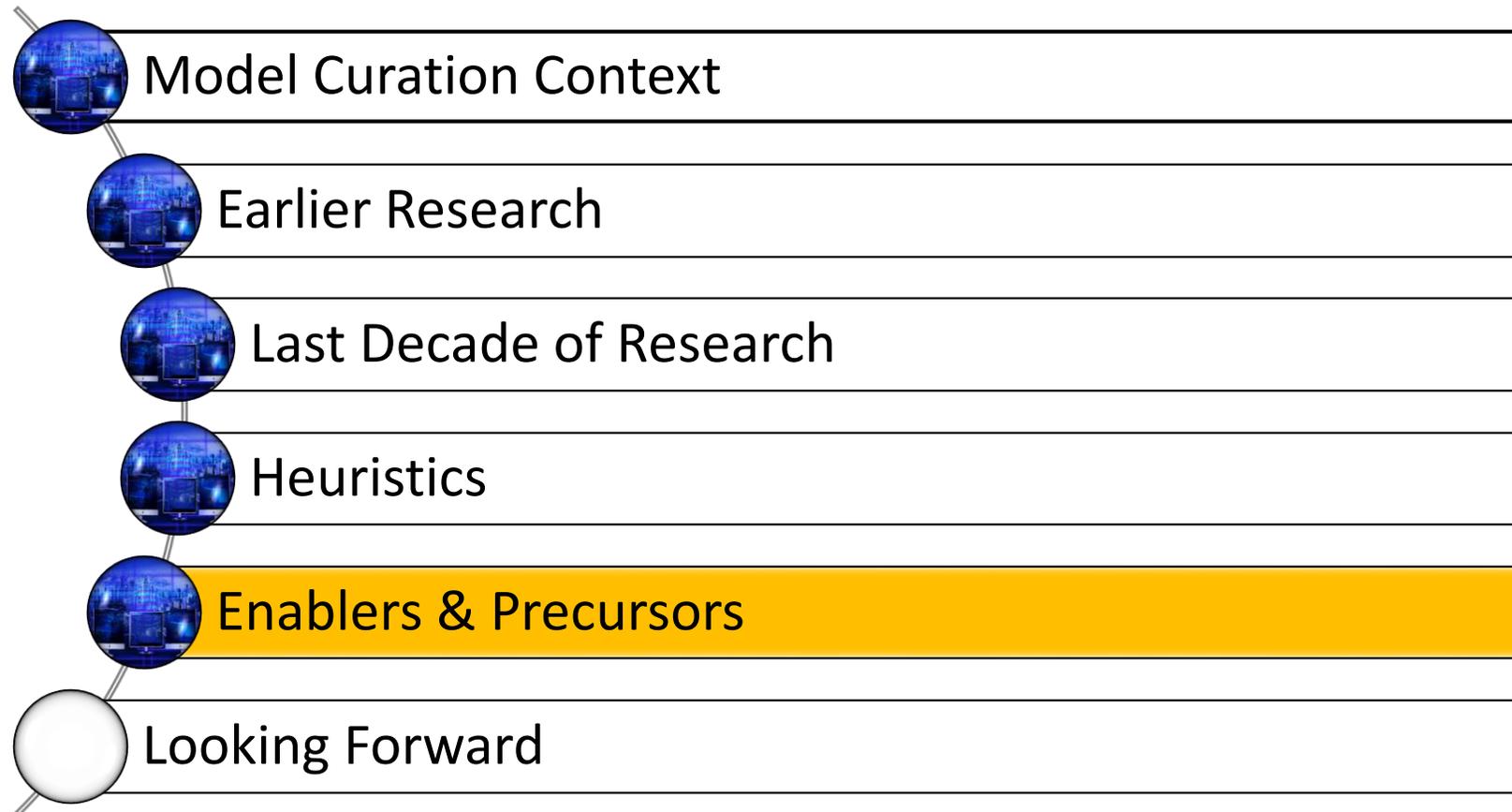
30 recognized experts

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



Heuristics from literature and studies

- Model credibility is an **attribute of the model consumer**, not the model.
- Model credibility is **positively influenced by communication between modeler and model consumer**, both active and passive.
- Credibility of digital artifacts is **influenced by both trustworthiness and expertise of a model consumer**.
- **Credibility of models** in a collection influences a model consumer's trust in the enabling infrastructure.
- A model consumer's **experience in discovering and retrieving models from a repository influences perceived credibility** of the model.
- 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**.
- Model credibility is influenced by a **model consumer's capacity for transparency into the validation activities throughout its development and enhancement**.
- Credibility of the model collection is **influenced by a model consumer's perception of expertise of the governance authority that accepted the model into the collection**.



Model Pedigree as Enabler for Trust

A **model pedigree** contains all of the information about a model, its origins and use over time

PURPOSE

...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)



Pedigree has information not always included in engineering model documentation and metadata: model origin, originators and developers, development process used, assumptions and limitations, related expert knowledge, investment in model, etc.

Precursors to Authoritative Source of Truth (proposed)

SERC WRT-1009 Model Curation Innovation and Implementation (Rhodes et al., 2020)

1. Model credibility is an attribute of the model consumer, not the model itself.

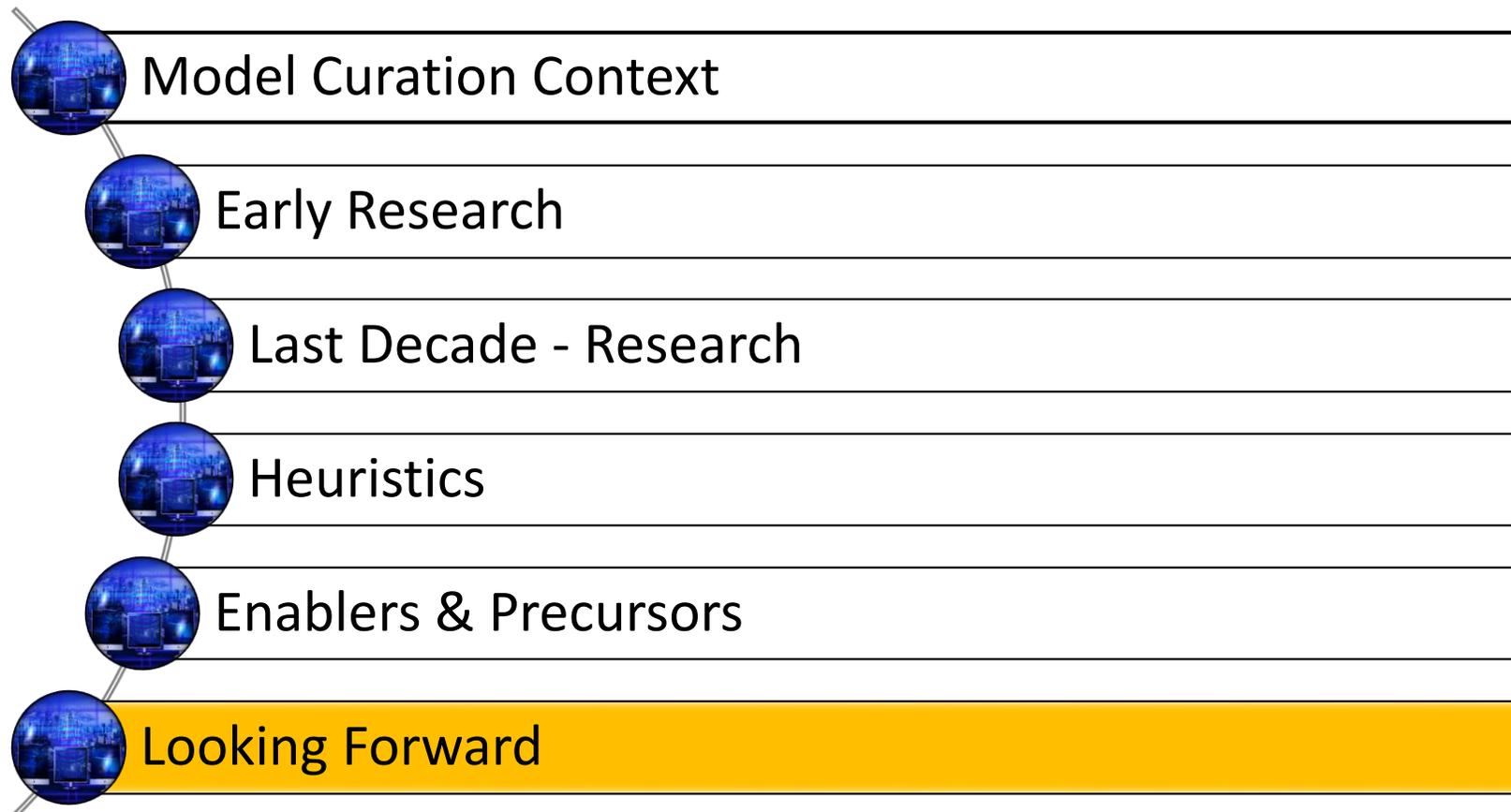
- a. Perceived credibility of a model is influenced by the expertise of the model consumer.
- b. Model credibility is influenced by a model consumer's trust in the expertise of the model originator.
- c. Perceived credibility of model is influenced by the model consumer's individual propensity for model 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.



Ongoing Research in Systems Community

Model Verification, Validation, and Uncertainty Quantification (VVUQ) Pattern (Schindel)

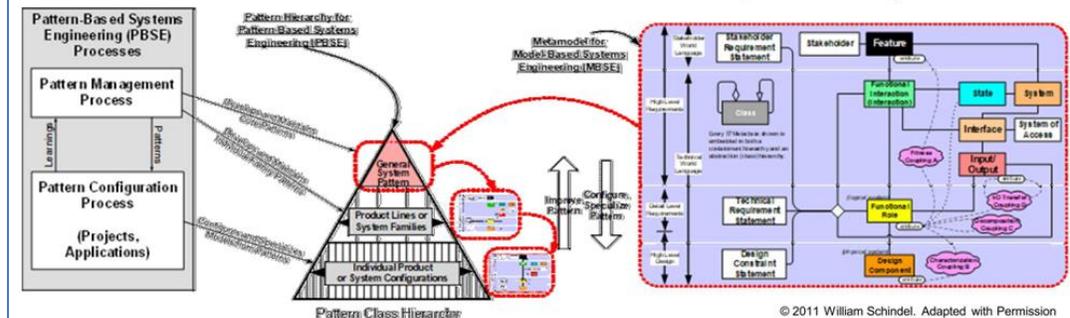
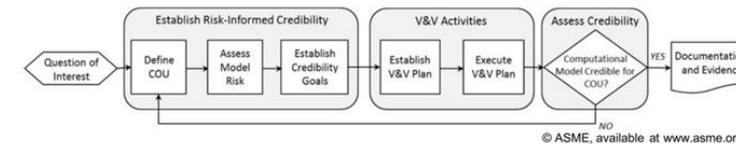
Model Credibility															
					Model Envelope			Validated Conceptual Model Credibility			Verified Executable Model Credibility				
					MODEL APPLICATION ENVELOPE			Quantitative Accuracy Reference			Quantitative Accuracy Reference				
								Function Structure Accuracy Reference			Function Structure Accuracy Reference				
								Uncertainty Quantification (UQ) Reference			Uncertainty Quantification (UQ) Reference				
								Model Validation Reference			Model Validation Reference				
								Speed			Speed				
								Quantization			Quantization				
								Stability			Stability				
								Model Validation Reference			Model Validation Reference				
Feature Group	Feature Name	Feature Definition	Feature Attribute	Attribute Definition	Model User	Model Developer	Model Maintainer	Model Deployer	Model User Supporter	Regulatory Authority	Model Investor	Physics Based	Data Driven	Model Type	
Model Credibility	Verified Executable Model Credibility	The verified capability of the executable portion of the model to represent the System of Interest with acceptable Credibility.	Quantitative Accuracy Reference	The specification reference describing the quantitative accuracy of the executable model to the conceptual model.	X	X			X	X		X	X		
			Structural Accuracy Reference	The specification reference describing the structural (presence or absence of elements) accuracy of the executable model to the conceptual model.	X	X			X	X		X	X		
			Uncertainty Quantification (UQ) Reference	The specification reference describing the degree of uncertainty of the executable model to the conceptual model.	X	X			X				X	X	
			Speed	The specification reference describing the execution run time (speed) for the executable model.	X	X			X	X		X	X		
			Quantization	The specification reference describing the quantization error of the executable model.	X	X			X	X		X	X		
			Stability	The specification reference describing the level of stability of the accuracy and uncertainty of the executable model error characteristics.	X	X			X	X		X	X		
			Model Validation Reference	The reference documenting the verification of the executable model's Credibility to the conceptual model.	X	X			X	X		X	X		

Schindel W.D. (2020) System Patterns in Engineering and Science. In: Metcalf G., Kijima K., Deguchi H. (eds) Handbook of Systems Sciences. Springer, Singapore. http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-981-13-0370-8_19-1

A Research Proposal: Assessing Credible Systems Engineering Models Through Rigorous Augmented Intelligence Analysis (Carroll)

How do we trust a digital system model?

What if we automated the validation (VVUQ) of new models to known patterns using augmented intelligence rules?



Bill Schindel has proposed the creation of patterns formulated in a Pattern-Based-Systems-Engineering S*Model – similar to how a modular architecture library of design models is implemented. A pattern catalog and processes for pattern capture are developed to manage the verification, validation, and uncertainty quantification (VVUQ-regression test and analyses), reuse, and integration of model patterns based on a shared ontology. – Analyzing for completeness, traceability, compliance, recognition, reusability, and more.

Carroll, E. R. (2019). A Research Proposal: Assessing Credible Systems Engineering Models Through Rigorous Augmented Intelligence Analysis (No. SAND2019-3729C). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

Closing Thoughts

- Model credibility and its associated constructs (model confidence, model trust, model validation, model value, etc.) have been investigated and discussed in the literature for **more than four decades**
- Findings suggest that **prior work on credibility has significant relevance** for model-based systems engineering and model curation
- Formulation of **heuristics** that draw from prior work and other fields can be **inform the strategies and enablers** for model-based systems engineering and model curation
- Many areas for future inquiry: e.g., credibility in regard to **data curation, design of repositories and interaction, and augmented intelligence**

