A Review and Framework for Modeling Complex Engineered System Development Processes

John Meluso, Jesse Austin-Breneman, James P. Bagrow, and Laurent Hébert-Dufresne

***Abstract*—Developing complex engineered systems (CES) poses significant challenges for engineers, managers, designers, and businesspeople alike due to the inherent complexity of the systems and contexts involved. Furthermore, experts have expressed great interest in building a foundation of "theories" that describe the patterns underlying how development process qualities shape system outcomes. This article contributes to that foundation in two ways. First, it identifies the core elements of CES development processes (CESDPs) through a literature review. Then, it proposes the ComplEX System Integrated Utilities Model (CESIUM), a novel framework for exploring how numerous system and development process characteristics may affect the performance of CES. CESIUM creates abstract representations of a system architecture, the corresponding engineering organization, and the new product development process through which the organization designs the system. It does so by representing the system as a network of interdependent artifacts designed by agents. Simulated agents iteratively design their artifacts through optimization and share information with other agents, thereby advancing the CES toward a solution. Hence, this article makes it possible for researchers to compare how development process characteristics shape system outcomes.**

***Index Terms*—Complexity theory, design methodology, graph theory, large-scale systems, networks, product development, simulation, systems engineering.**

I. Introduction

Society increasingly relies on healthcare systems, stock markets, automotive systems, national defense programs, and countless other designed systems; however, these systems continually grow more difficult to develop and manage due to the complexity within and around them. These complex engineered systems (CES) are large sets of highly interacting engineered artifacts with a defined purpose.[1] Often "difficult to describe, understand, predict, manage, design, or change" [1] they are composed of many heterogeneous elements, characterized by nonlinear interactions at multiple levels of organization and abstraction, and exhibit complex behaviors which emerge from those interactions [2].

To date, government and industry organizations alike manage system development via systems engineering techniques [3], Deming's quality control methods [4], and six sigma principles [5]. Researchers construct workflow models through Petri nets [6], [7] and integrated design-and-management approaches [8] to help practitioners incorporate uncertainty and feedback. However, traditional methods have not kept pace with the varying scales and increasing interactions between elements of systems. The cost and schedules of engineering projects have grown exponentially [9], [10], leading to frequent overruns [11], [12], billions of dollars in losses, and even in lives [13]. Such pervasive failures mean that rather than "exceptional," failures have instead become "normal" or expected [14]. Made more challenging, organizational factors often create the adverse conditions in which society feels the effects of technical and "human" errors [15].

While traditional systems engineering methods still hold value, leading government, industry, and academic voices have expressed a "dire need" [16] for deeper *theoretical* understandings of CES, CES development processes (CESDPs), and the organizations that create them [12], [16]–[18].[2] Researchers have begun to build theory for traditional systems engineering processes [19] and frameworks for particular CES contexts, such as autonomous vehicles [20] or Internet of Things

Manuscript received 4 November 2021; accepted 17 March 2022. Date of publication 11 April 2022; date of current version 18 November 2022. The work of John Meluso, James P. Bagrow, and Laurent Hébert-Dufresne was supported in part by Google Open Source under the Open-Source Complex Ecosystems And Networks (OCEAN) Project. This article was recommended by Associate Editor C. Zhang. *(Corresponding author: John Meluso.)*

John Meluso is with the Vermont Complex Systems Center, University of Vermont, Burlington, VT 05405 USA (e-mail: john.meluso@uvm.edu).

Jesse Austin-Breneman is with the Department of Mechanical Engineering, University of Michigan at Ann Arbor, Ann Arbor, MI 48109 USA.

James P. Bagrow is with the Department of Mathematics and Statistics, University of Vermont, Burlington, VT 05405 USA.

Laurent Hébert-Dufresne is with the Department of Computer Science, University of Vermont, Burlington, VT 05405 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2022.3163019.

Digital Object Identifier 10.1109/TSMC.2022.3163019

[1]An *artifact* is any piece of technology designed to serve a specific purpose, often used as "an umbrella term for any technical product of human minds including physical parts, software, processes, information, etc." [1], [33], up to and including the designed roles of people in those systems [89].

[2]Throughout his work on behalf of the United States National Science Foundation (NSF), the United States Department of Defense (DoD), and the International Council on Systems Engineering (INCOSE), Collopy [23] and Topcu *et al.* [90] identified several areas "ripe for exploration" to address these needs, including abstraction and model-based system design. They suggest that "theory could make a real difference" toward understanding CESDPs [16]. Importantly, what "theory" means goes beyond a need for better tools for understanding how systems behave [90] or better processes for developing systems [91]. The greater need is for a foundation of the abstract, cross-cutting patterns (i.e., "theories") that describe how methods of guiding, facilitating, and coordinating teams shape system outcome [16], [17], [90].

devices [21], [22].[3] Nonetheless, few efforts have suggested abstract frameworks through which to build theoretical understandings of the complexity involved in CES more broadly and their requisite development processes [23].[4]

Fortunately, recent advances make such analysis possible.[5] Systems engineering research has explored alternative system formulations, such as value-driven design [24] while engineering design research has formulated Multidisciplinary Design Optimization [25], each of which lends itself to abstract representations of CES development.[6] Likewise, organizational research is beginning to explore computational methods of representing organizational processes [26]–[29] as engineering follows suit [30]–[32].

Given these advances, this article builds on the work of Meluso *et al.* [33] to propose a modeling framework for analyzing variations in how organizations develop CES. The framework combines techniques from systems engineering, complex systems, engineering design, and organizational theory to form the complex system integrated utilities model (CESIUM). The remainder of this article is organized as follows. First, it summarizes the literature defining CES, the constituent elements of CESDPs, and the methods for scientifically studying CESDPs (Section II). It then describes the model and simulation methods of CESIUM as an abstract, but grounded, representation of a CESDP (Section III), before characterizing the model through parameter sweeps, discussing validation, and exemplifying its potential for theory formation (Section IV). This article closes by generalizing the framework to facilitate the development of a more robust theory of CES development.

II. Background

Several disciplines touch, and uniquely define, the complex systems constructed by people. As such, this section starts by delineating the definitions of CES (Section II-A). From this common understanding, it describes the constituent elements of CESDPs (Section II-B). Then, it details the literature specific to the methods of CESIUM (Section II-C).

A. Definitions of Complex Engineered Systems

Scholars continue to debate what to call systems of interacting artifacts that people design, and with good reason. Different definitions emphasize different system characteristics, the relationships between the artifacts, and the relationship between the system and the outside world.[7] A few particular terms are worth highlighting, though, as they reveal the core elements necessary to describe CESDPs.

1) Complex Systems: Conceptually, complex systems appear in disciplines from biology to sociology [35]. The interdisciplinary field of complex systems grew out of these and several other disciplines in the latter half of the 20th century with a few core methods for understanding adaptation, information, and collective behaviors [36]. Newman [35] notes that while there is no precise definition, most researchers define complex systems as sets of "many interacting parts, such that the collective behavior of those parts together is more than the sum of their individual behaviors." An important characteristic of complex systems is what are called *emergent behaviors*, the unexpected macroscopic behaviors that often result from simple microscopic interactions between constituent elements [35], [36]. These characteristics—significant interactions and emergent behaviors—collectively suffice to begin defining engineered versions of such systems.[8]

2) Complex Engineered Systems: As with complex systems, consensus does not exist about the definition of a CES [37]. The simplest definition is "networks of interconnected components" [38], though the focus therein limits itself to networks of physical artifacts.[9] Put another way, they are complex systems composed of designed artifacts with interactivity and potential for emergence, thereby making CES a subset of complex systems.

3) Engineering Systems: So far, the definitions of CES do not consider the significant role that people play in the creation, operation, and evolution of these systems. The term *engineering system* takes a more social perspective of engineered systems, defined as "a class of systems characterized by a high degree of technical complexity, social intricacy, and elaborate processes, aimed at fulfilling important functions in society" [1], underscoring the social purposes while acknowledging the technical, procedural, and social interaction.[10]

[3]This work draws from the definition of Fortino *et al.* [22] for our definition of a *framework*, which they define as, "A general guideline, to be specialized or extended, outlining the steps that should be followed in analyzing, implementing or maintaining a solution to a number of similar problems." Frameworks often solve a specific class of problem (e.g., Liu *et al.* [92] framework created a process for safely managing constrained, networked resources). In this work, the framework creates a process for identifying relationships between CESDP characteristics and CES outcomes across a large number of systems.

[4]Increasingly, scholars recommend utilizing *systems theory* to assess such problems, a set of concepts rooted in complex systems and management scholarship [93], [94], acknowledging the need for agent-based methods to explore relative independence, discrete events, and interactive dynamics [95].

[5]Instead of undertaking the extremely difficult task of collecting data on a statistically significant number of complex systems, on both their design processes and outcomes, an alternative approach is to "generat[e] large numbers of unique complex system models and simulat[e] design processes in those complex systems, all grounded in evidence from empirically verified phenomena" [33]. This article takes this generative approach. Here, *simulations* are computational representations of real-world systems, processes, and events [96]. *Representations* are individual or group-level knowledge structures that provide simplified mental maps of concepts [97]. Network theory facilitates simulated representations of CES architectures [38], [60], and agent-based modeling grants access to macro-level outcomes of micro-level decisions and interactions [98].

[6]*Abstraction* refers to the process of describing a concept through the identification of a common set of features [90]. Therefore, *abstract representations* are simplified knowledge structures that describe common features of a concept, in this case CES and CESDPs.

[7]At the least, it creates something of a "branding problem" when university departments adopt more than 20 unique names for CES [1].

[8]The term complexity also appears throughout the complex systems literature, along with the field of complexity science. For the purposes of this article, *complexity* refers to the characteristics of a complex system and are thus roughly synonymous with complex systems, though entire texts exist to define the term and field more completely [36].

[9]Braha *et al.* [2] defined CES as "engineered systems [that] are composed of many heterogeneous subsystems and are characterized by observable complex behaviors that emerge as a result of nonlinear spatio-temporal interactions among the subsystems at several levels of organization and abstraction."

[10]Much like how information scholars embrace social subsystems of systems via the term socio-technical system [99], the term engineering system ensures that human interaction with technology and other people remains integral, and consequently unique among the definitions.

4) Systems of Systems: Moving closer to practitioners, a *system of systems* is "a set or arrangement of systems that results when independent useful systems are integrated into a larger system that delivers unique capabilities" [39]. International standards bodies provide similar definitions [40]. Papers on systems of systems often refer to such standards [41] reflecting the industry orientation of the term and practice-based attempts to address complexity [42], though academics utilize similar definitions [43].[11]

5) Large-Scale Complex Engineered Systems: Bridging theory and practice, large-scale CESs (LSCES or LaCES) are "engineering projects with significant cost and risk, extensive design cycles, protracted operational timelines, a significant degree of complexity, and dispersed supporting organizations" [32]. The term grew out of a confluence of experts at the NSF, the National Aeronautics and Space Administration (NASA), academia, and industry to address the challenges described in Section I [9], [12], [44], [45].[12]

6) Working Definition of Complex Engineered Systems: Given these definitions and the various concepts they encompass, this article uses the term *complex engineered system* (still abbreviated CES) defined as large sets of highly interacting engineered artifacts with a defined purpose. This definition includes the interactivity and potential for emergence of complex systems. While the definition does not explicate human influence, it presumes that people design, operate, and evolve artifacts. Similarly, while environments, contexts, and non-designed objects also interact with the system, they remain peripheral until integrated or modified by people, at which point they too become artifacts. Through this definition, it becomes possible to advance theory of how CES develop.

B. Constituents of CES Development Processes

Rather than describing widely used development processes,[13] this section draws on Section II-A to identify the abstracted elements that compose CESDPs regardless of process generality or specificity. These elements include:

1) the *organization*[14] developing the CES;
2) the *context* in which the organization develops the CES;
3) the *process* an organization uses to develop the CES;
4) the *CES* itself.

To fully understand how CES develops, researchers, and practitioners must understand all four of these elements. The following sections address these elements in turn.

1) Organizational Theory: Management decisions affect organizational performance, and consequently, system performance. This holds true with subsystem complexity [46], the balance between project and functional management in a matrix organization [47], employee incentives [48], and even employee perceptions of procedural justice in top management decisions [49]. The complexity of these systems necessitates organizational involvement, so leading scholars argue organizational theory is necessary for creating systems engineering theory as well [17] because engineering is a fundamentally social activity [9].[15]

2) Context: Defining a system's boundaries is one of the most important tasks in system definition, and no less important are understanding inputs, outputs, and external forces or "externalities" [1]. The breadth or narrowness of a project's scope can affect performance outcomes [50]. Just as the context of the system affects its performance, so too does the context of the organization, such as through information technology selection [51], the influence of marketing on project decisions [52], whether the companies involved in development are vertically or horizontally integrated [53], and organizational culture [54].[16]

3) Development Processes: Practitioners have long suggested that an organization's choice of development process matters. Development processes are sets of interacting tasks that collectively accomplish a specified objective [55]. Unsurprisingly, system performance depends on the chosen development process [50]. Currently, insufficient theory exists to answer why processes perform differently [16]. Some knowledge exists in the management sciences about "new product development" [56] and engineering design researchers are conducting experiments on micro-level processes that produce better designs [45], [57]. Still, much remains unknown.

4) Theory on Complex Engineered Systems: The majority of existing theory on CES grounds itself in network theory. Network scholars observed that the Internet, the world wide Web, power grids, and other CES could be represented as networks [58]–[60]. Engineering scholars have characterized the networks of several CES as well, including spacecraft, vehicles, and software. Theories have begun to form around structure [38], [61], modularity [62], quality [38], [46], and the limitations of systems engineering for CES [63]. A small contingent of work also studies how uncertainty affects system performance via game theory [32], [64].

[11] Valid critiques exist on both sides: Bar-Yam [63] noted that traditional systems engineering practices struggle to manage the many interactions and interdependencies between systems, though Alderson and Doyle [100] similarly critique complex systems for not addressing issues of practice.

[12] Proponents advocated for the development of systems engineering theory [9], [44], of engineering design [45], optimization techniques [9], [11], [101], and socio-technical understandings of systems [12], [101]. Despite efforts to incorporate the various interests of the academy and industry, the technical and the social, the term remains limited to the mechanical design community.

[13] For example, the Systems V [43], the Toyota Production System [102], or processes specific to individual organizations [103].

[14] *Organizations* are collectives of actors who pursue both shared and disparate interests, but recognize the value of perpetuating the collective as a shared resource [104].

[15] One of the core organizational theory texts is Scott and Davis' [104] *Organizations and Organizing*. In organizational theory contexts, "organizations are *systems* of elements, each of which affects and is affected by the others" [104] much as in technical systems, though here the term includes people, information, and artifacts. Historically, three views of organizations emerged: 1) as *rational systems* that presume formal relational structures between people who share a common goal; 2) as *natural systems* wherein people have multiple (shared and individual) goals and the relational structure serves as a resource; and 3) and as *open systems* through which resources pass and interact with shifting coalitions of people. Each theoretical perspective bears similarity with CES concepts and may serve as foundations upon which to build theory.

[16] Theoretical work in this space is primarily limited to the organization and management sciences to date, though it certainly merits further inquiry given its importance.

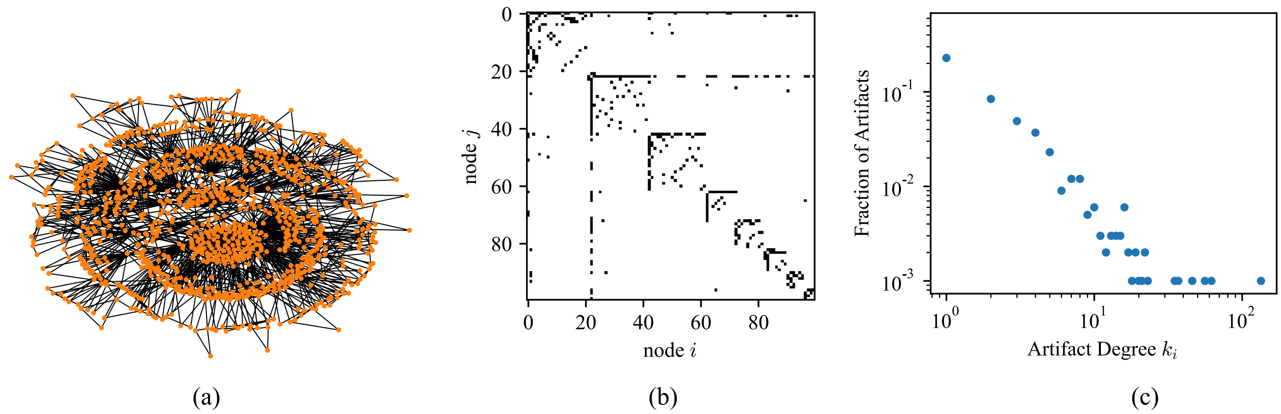


Fig. 1. Visual representations of undirected networks generated with a Holme–Kim preferential attachment algorithm and probability that new edges will form a triangle $p_t = 0.9$. (a) Graph of a network with $n = 1000$ artifacts. The orange dots represent nodes or artifacts, and the black lines represent edges or interfaces. (b) Adjacency matrix A_{ij} for a network with $n = 100$ artifacts where $A_{ij} = 1$ (black) if an edge exists between nodes i and j and 0 (white) if not. Also, known as a design structure matrix. (c) Scale-free degree distribution of a network with $n = 1000$ artifacts. Note the approximately linear, negatively sloping form of the distribution on a log-log scale, characteristic of a scale-free degree distribution [34].

C. Methods Specific to CESIUM

Increasingly, diverse methods exist for studying CESDPs (see Section I in the supplementary material for a brief review) thereby creating opportunities to combine methods into representations of CESDPs, as with CESIUM. The following sections describe the contributions of network theory, agent-based modeling, and design optimization to CESIUM.[17]

1) Network Models: Network theory represents systems of people or artifacts as nodes and edges, as in Fig. 1.[18] Artifacts in many (though not all) complex systems follow a *scale-free* degree distribution, also called power-law or inverse exponential distributions [38], [60], [61], [65]. A more-recent subject in network theory with significant potential is that of generative network models. These models algorithmically construct networks out of basic rules [34] and so facilitate simulations of various dynamics (e.g., agent synchronization [66]). One of the most common network generation algorithms is called preferential attachment which builds a network by connecting new nodes to existing nodes with an attachment probability proportional to the degree of the existing node [34]. Several such algorithms exist, including those of Barabási and Albert [65], Holme and Kim [67], and Carlson and Doyle [59], all of which generate networks with scale-free degree distributions [34], [67]. The Holme–Kim preferential attachment algorithm includes a parameter for tuning node clustering [67], making it useful for simulating the meso-scale practice of subsystem formation. On the other hand, Carlson and Doyle's highly optimized tolerance (HOT) generates bow-tie structures, yielding macro-level realism. An approach combining the Holme–Kim and HOT approaches would likely provide the most realistic CES structures, albeit similarly presumptive of solution structures.

2) Agent-Based Modeling: Agent-based modeling is a widely used, effective, and tested method for simulating CES [68], [69]. An ABM creates a system of autonomous decision-making entities called agents which individually assess their situations and make decisions based on a set of rules [68]. Agents affect their surroundings through their actions and in doing so, self-organization, patterns, structures, and behaviors "emerge" from the "ground up" that were not explicitly programmed into the models but nevertheless arise through agent-interactions [69].[19] The ability of ABMs to explore how outcomes emerge from system or development process properties makes it ideal for understanding highly interacting systems without impacting the performance of real systems.

Recent applications of ABMs include system design [32] and organization studies [70]. Because CES is often composed of many smaller engineered systems that are designed, developed, and operated by organizations of dispersed, loosely connected people [11], ABMs facilitate simulation of aggregated artifact development in ways that top-down models cannot [71], making them one of the primary methods through which "to inform tradeoff decisions" regarding "complexity in system design and development" [18].

3) Design Optimization: Engineers in various disciplines use *design optimization* to maximize the performance of a system, a process of selecting the relative "best" alternative from among a set of possible designs called the *design space* [72]. They do this through *objective functions* (or *utility functions* in agentic contexts), sets of evaluation criteria typically constructed as functions describing the relationships between independent or *decision variables* [72]. Optimization algorithms then explore the design space to find a global or

[17] Section II-C was partially adapted from [33, Sec. 2.2].

[18] For a brief introduction to networks, see in the supplementary material Section II.

[19] Macal and North [105] defined "multiagent simulation [as] the simulation of an agent-based model. [Then,] multiagent simulations take place in a simulated world, whereas multiagent systems are usually considered to have interactions with the real world." While multiagent systems are valuable for optimization, this work deals with agent-based models and simulations for their descriptive and exploratory benefits within such simulated worlds. This "ground up," agent-centered approach differentiates ABMs from other system modeling methods, such as system dynamic models, process-centered approaches like discrete event simulation or Petri nets [6], and from optimization-oriented approaches like multiagent systems [105].

local minimum (or maximum depending on problem construction) as efficiently as possible to identify a solution [25].

While the methods of constructing system objectives are beyond the scope of this article, one method for searching design spaces remains relevant. Recent studies [31] have validated hypotheses that engineers [73] and organizations [74] sample their design spaces comparably to *simulated annealing* which can therefore serve as an abstract modeling representation of human decision making.

Given this background, the next section combines these concepts to form CESIUM through the framework established in Section II-B.

III. Methodology

The CESIUM framework simulates a theoretical CESDP by modeling its constituent elements (Section II-B). This section describes how the *base instance* of CESIUM represents each element: the system architecture and system boundary (Section III-A), the constituent artifacts (Section III-B), the organization of interacting agents (Section III-C), agents' design processes (Section III-D), and the system development process (Section III-E). Thereafter, an example execution of the base instance follows for a vizualizable system (Section III-F) along with a brief introduction to the framework's flexibility in simulating variations on the elements of CESDPs (Section III-G).[20,21]

A. System Architecture

First, assume that a CES is composed of n interacting artifacts where each artifact $i \in \{1, \dots, n\}$. The n artifacts interact with one another in a technical network generated from a scale-free degree distribution via a Holme–Kim preferential attachment algorithm by adding h edges to each new node i.[22] The result is a network of n interacting artifacts described by adjacency matrix A_{ij}, in this case with no formal hierarchy and clustering specified by p_t.[23]

B. Artifact Construction

In a real-world setting, the design of each artifact i in the system would depend on numerous contextual and specific factors, say $\mathbf{v}_i = [v_{i1}, v_{i2}, \dots]$. Because these factors cannot be known *a priori* for innumerable real systems, the model representatively parameterizes these variables such that the design of each artifact is defined by a single decision variable $x_i(\mathbf{v}_i)$. Therefore, each x_i parameterizes a complex set of inputs,

[20]The implementation of CESIUM described herein was developed using Python 3 with the NumPy 1.18.1, SciPy 1.4.1, and NetworkX 2.4 packages. The full code for this article is available at https://github.com/meluso/cesium-framework, and a clean version of code for the base instance of CESIUM is available at https://github.com/meluso/cesium-base.

[21]Sections III-A–III-E were partially adapted from [33, Sec. 3].

[22]The probability that the first edge connecting node i to the network will attach to a specific node j is proportional to that node's degree k_j; the probability that subsequently added edges will be placed to form a triangle by connecting i to a node l that is already connected to j is p_t [67].

[23]Generally, artifacts could also interact with nodes outside the system boundary. The base instance assumes that the system boundary contains all factors with sufficient impact on the CES leaving no exogenous variables, and hence no interaction between the CES and its context.

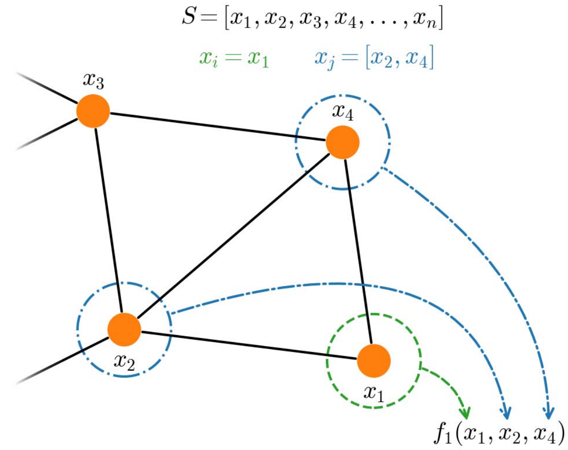


Fig. 2. Example of artifact interaction. In this case, the ith artifact is artifact 1 with variable $x_i = x_1$. Artifact 1 interacts with $j \in \{2, 4\}$ and so $x_j = [x_2, x_4]$. Therefore, artifact 1's objective function is $f_1(x_1, x_2, x_4)$.

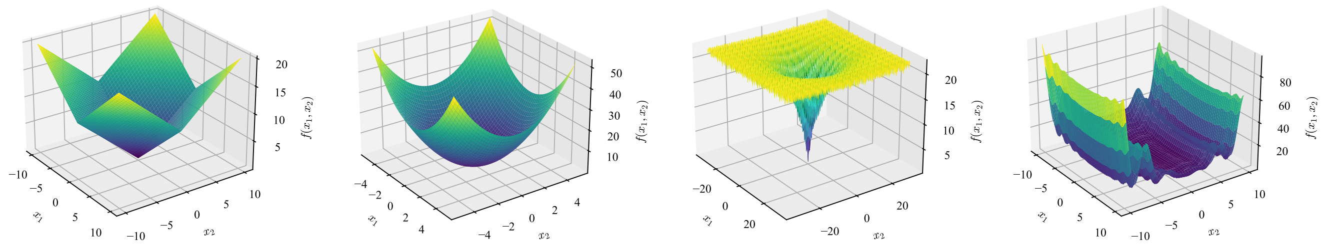
allowing the performance of each artifact to be represented as an objective function $y_i = f_i(x_i, \mathbf{x}_j)$, where $j \in \{1, \dots, k_i\}$ represents the set of artifacts interfacing with artifact i, and $\mathbf{x}_j$ is a vector of the parameterized decision variables of the k_i artifacts as exemplified in Fig. 2. These variables can be expressed using a combined notation $\mathbf{x}_i = [x_i, \mathbf{x}_j]$.

Objective functions can take countless forms in the framework, but building up from the simplest relationships between artifacts provides a foundation for understanding more complex relationships. Because $\mathbf{x}_i$ can be parameterized to any mapping, the simplest objective topologies can be identified through a Taylor expansion of $f_i(\mathbf{x}_i)$. At some point $\mathbf{x}_{i0}$, with $\partial\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}_{i0}$ and the Hermitian matrix $H_i(\mathbf{x}_i)$, the Taylor expansion of $f_i(\mathbf{x}_i)$ is

$$f_i(\mathbf{x}_i) \approx f_i(\mathbf{x}_{i0}) + \nabla f_i(\mathbf{x}_{i0})\partial\mathbf{x}_i + \frac{1}{2}\partial\mathbf{x}_i^T H_i(\mathbf{x}_{i0})\partial\mathbf{x}_i + \cdots \quad (1)$$

If $\mathbf{x}_i$ is parameterized such that the optima all occur at $\mathbf{x}_i^* = \mathbf{0}$, then $f_i(\mathbf{x}_i \neq \mathbf{0}) > f_i(\mathbf{x}_i^* = \mathbf{0})$ for every i. The simplest such relationship occurs under the condition of linearity such that $|\nabla f_i(\mathbf{x}_{i0})\partial\mathbf{x}_i| = \sum_{m=1}^{k_i+1} |x_m|$ with all higher order terms equal to 0. The next simplest then is a quadratic formed by taking only the second term of the Taylor expansion such that $|(1/2)\partial\mathbf{x}_i^T H_i(\mathbf{x}_{i0})\partial\mathbf{x}_i| = \sum_{m=1}^{k_i+1} x_m^2$. More complex relationships, with multiple minima and asymmetrical objective functions, will of course prove more realistic though may prove more challenging to formulate and analyze. However, those complexities can similarly be overcome through the initial parameterizations of each x_i.

To examine both simple and complex relationships, the base instance of CESIUM considers four objective functions (see Fig. 3). The first two represent simple relationships between artifacts. They correspond to the first-order Taylor expansion term (the Absolute Sum function) and the second-order Taylor expansion term (the Sphere function). The first complex function is a symmetrical multiple-minima function (the Ackley function) which represents problems with multiple solutions of varying quality. The final function is an asymmetrical multiple-minima function (the Levy function) which



(a) (b) (c) (d)

Fig. 3. Graphical representations of the selected objective functions with two decision variables. (a) Absolute Sum function. (b) Sphere function. (c) Ackley function. (d) Levy function.

represents instances where agents have a greater influence on each others' artifacts than on their own artifact. Each function is formulated in Section III of the supplementary material.[24]

C. Organization

CESIUM assumes that agents in an agent-based model represent members of an organization (engineers or otherwise). In the base instance, one agent represents one engineer. While, the pairings of agents to artifacts take many forms in real life, an organization's structure approximately reflects the structure of the technical artifacts that those organizations create. This phenomenon, called the mirroring hypothesis or Conway's Law [75], means the simplest mapping is one in which the network of agents in the organization and the network of artifacts in the CES are synonymous. Each agent is then responsible for one artifact in the system. Hence, an organization exists wherein engineers pass information via the technical network.[25]

D. Agent Design Process

Next, the model incorporates a design process for the artifacts. Given the mirroring hypothesis, each agent uses the technical objective function of its artifact as its utility function, so the objectives will be spoken of as belonging to the agents. Each agent seeks to optimize (minimize) its objective function over a number of *turns* to reach the best performance.[26]

During each turn of the model, each agent engineer receives a set of constant input values $\mathbf{x}_j$ from its k_i interacting agents to utilize when adjusting x_i to optimize their objective functions. As real engineers do [31], agents explore the design space using a simulated annealing algorithm in search of a local optimum $y_i^* = f_i(x_i^*, \mathbf{x}_j)$ with a random initial position in the domain of x_i, ω iterations per optimization, initial temperature of τ, and cooling rate of ρ.[27] No cognitive factors affect agent decision making.

E. System Development Process

With a technique established through which members of an organization develop each artifact, it becomes feasible to simulate the development of the complete system. At the beginning of the CESDP, the ABM first initializes a new system following the method outlined in Section III-A. Then, the model steps through a series of turns where one turn in the ABM represents one design cycle in a CESDP. Each turn, agents exchange information. To facilitate the exchange, CESIUM stores the latest reported designs of all agents in a system vector $\mathbf{S}$ as a central repository. At the beginning of each design cycle, each agent receives $\mathbf{S}$ as a constant input before proceeding to optimize their variable x_i using only the values from their networked neighbors $\mathbf{x}_j$. Then, each agent passes their updated value of x_i back to the system vector for storage in $\mathbf{S}$ and a new design cycle begins with the updated values as constants.

The model performs these design cycles, iterating through all of the agents in each cycle, until either the system design converges or the model performs d design cycles. System convergence is calculated from a metric for system performance F. While many formulations of performance are possible, the base instance defines F as a sum of the reported objective evaluations of all of the agents during the current design cycle

$$F(t, \mathbf{f}) = \sum_{i=1}^{n} f_i(t, \mathbf{x}_i(t)). \tag{2}$$

Although the n objective functions f_i have different magnitudes depending on the degree k_i of each artifact i, the system performance is assumed to have a greater dependence on components which are more highly connected so simply adding their contributions exemplifies this behavior. Then, the base instance of CESIUM assumes that system convergence occurs when the system performance consistently changes less than a specified convergence threshold ε over each of three turns,

[24] These objective are n-dimensional, meaning they scale to incorporate the k_i decision variables for each neighbor j of i. This leaves n coupled objective functions $\{f_1, \ldots, f_n\}$ that comprise the CES being designed.

[25] Section V-A discusses more complex mappings, including variation in the numbers of agents artifacts, as part of the generalization of CESIUM.

[26] Stepping forward in time through the use of turns is a common practice in ABMs. Turns may represent a common development process technique known as the Shewhart and Deming Cycle [106] wherein members of an organization iteratively improve and share the design for their subset of a system; though, turns also serve as a discretized representation of design refinement more generally.

[27] While other design space search algorithms remain possible, the base instance of CESIUM uses simulated annealing for agent optimization due to its established validation.

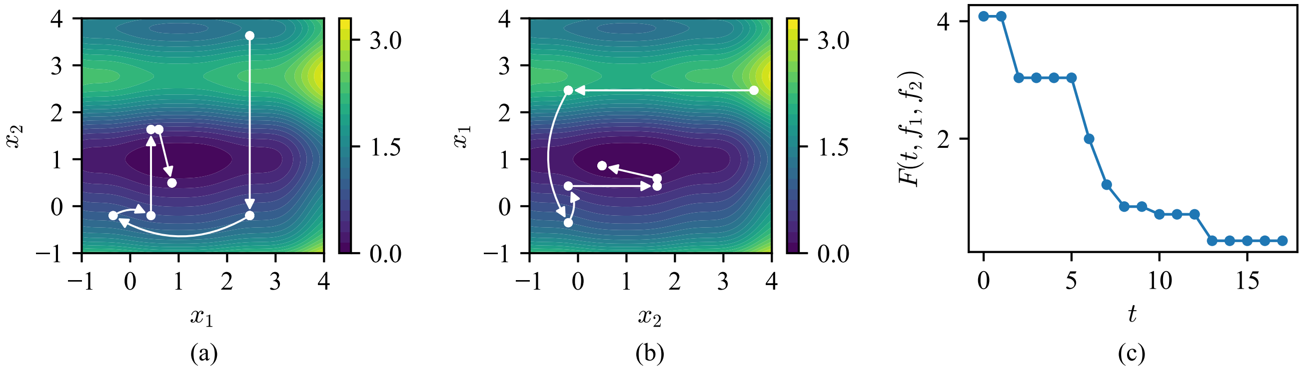


Fig. 4. Example execution of CESIUM for a system with two agents using the Levy function. White dots and arrows illustrate the progress of $\mathbf{x}$ over time, thereby affecting the objective evaluations. (a) f_1 of agent 1, (b) f_2 of agent 2, and (c) F the overall system performance. (a) $f_1(x_1, x_2)$. (b) $f_2(x_2, x_1)$. (c) $F(t, f_1, f_2)$.

meeting the following condition:

$$\frac{1}{3}\sum_{t'=1}^{3}\left|F(t,\mathbf{f}) - F\left(t-t',\mathbf{f}\right)\right| < \varepsilon. \tag{3}$$

Every execution of the model completes a minimum of three design cycles. Upon convergence, the number of design cycles N measures how long it took the organization to converge to a design solution, and therefore to complete the CESDP. This completes the base instance of CESIUM.

F. Example Execution

To illustrate how CESIUM simulates a complete development process, this section provides an example execution. Individuals in an organization explore their design spaces thereby affecting system outcomes. Large numbers of agents may be difficult to visualize, so consider a system with only two agents (Fig. 4). Each of the $i \in \{1, 2\}$ agents controls one decision variable, x_1 and x_2, respectively. Assume their artifacts are neighbors in the system network and the agents' objective functions $f_1(x_1, x_2)$ and $f_2(x_2, x_1)$ are Levy functions, as in Fig. 3(d), with system performance $F(t, f_1, f_2)$.

Unlike the other objectives presented, the Levy function is not symmetric meaning that some $f_i(x_i, x_j)$ is not necessarily equal to $f_i(x_j, x_i)$. Given the variable ordering for each agent's objective, Agent 1 has limited control over its outcomes because changes in x_1 tend to yield smaller changes in f_1 than changes in x_2 due to the shape of the Levy function, and vice versa for Agent 2. This configuration represents a system in which the performance f_i of each artifact i is disproportionately influenced by neighboring artifacts making convergence to the global optimum more challenging.

Fig. 4(a) and (b) shows the progress of each agent toward its individual objective as a function of the two decision variables. Fig. 4(c) shows the system convergence and progress toward the global optimum. Observe how changes in x_2 tended to improve f_1 even as they produced limited improvement in f_2; however, those changes in x_2 eventually led x_1 to identify improved values of f_1 and consequently x_2 to do the same for f_2. The iterative process of simulated annealing—a combination of necessarily accepting values of x_i that improve f_i and probabilistically accepting values of x_i that do not improve f_i—gradually improved system performance as in Fig. 4(c). Hence, this example execution demonstrates how CESIUM simulates both individual and organization exploration of a design space toward collective system outcomes.

G. Framework Flexibility

Throughout Sections III-A–III-E, the base instance makes a number of simplifying assumptions. However, the CESIUM framework *requires* very few of these. Only the elements described in Section II-B and forthcoming Section V-A need accounting for, thereby creating flexibility for researchers to alter and explore other characteristics, from system architecture topologies to representations of established development processes.

For example, Meluso *et al.* [33] recently utilized the framework to simulate miscommunication in a CESDP. Building on prior findings that practitioners define the "estimates" they communicate with one another either as representations of their *current* design status or *future* design outcome [76], the authors introduced a variable p_e specifying the probability that an agent would share future estimates (or conversely 1-p_e that they would share current estimates) finding that miscommunication can affect system performance. Such investigations follow from Section V, which generalizes the modeling framework and recommends further avenues of inquiry.

IV. Model Analysis

Model design decisions affect outcomes with any model. But as this section will demonstrate, the strength of CESIUM as a modeling framework lies in its ability to expose patterns that emerge across variation in design decisions. Abstracting away from particular CESDPs to their constituent elements (Section II-B) creates the opportunity to explore the fundamental characteristics of those elements, such as the number of artifacts in a system, the resources provided to engineers, or the qualities of the phases of a development process.

To that end, this section fills three purposes. First, Section IV-A characterizes the model through sensitivity analysis. In doing so, the analysis reveals a tremendous range of possible outcomes contingent on variables representative of practical system and design process qualities. Section IV-B

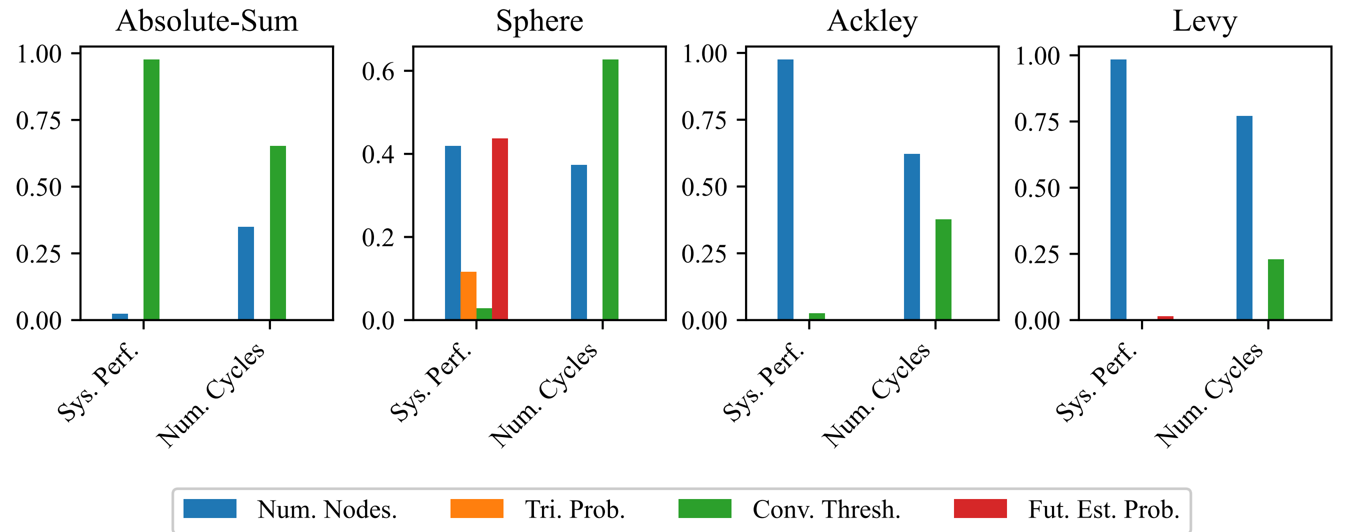


Fig. 5. Random forest feature importances for the objective function subsets of the executions. Higher bars indicate a more important feature (greater reduction in error when predicting the outcome). The Number of Nodes and the Convergence Threshold tended to have the greatest importance on System Performance and the Number of Design Cycles for most objective functions, the notable exception being the Sphere function, the outcomes of which were significantly affected by the Future Estimate Probability as well.

validates the framework. Section IV-C closes by demonstrating the framework's power to identify testable hypotheses and propose theories describing how real phenomena in practice affect system outcomes.

A. Sensitivity Analysis

Broeke *et al.* [77] characterized the performance of CESIUM via sensitivity analysis. The analysis swept independent CES and CESDP variables, examined their main and interaction effects via regression (Section IV-A in the supplementary material), and determined overall feature importances with a random forest. Table I in the supplementary material contains the independent variables, constants, and dependent variables; their definitions; and the values varied as part of a fractional factorial experiment. The subset of independent variables were selected because 1) they easily map to real-world concepts (e.g., system size and development time) and 2) they were assessed most likely to produce variation in the dependent variables. The Future Estimate Probability p_e (Section III-G)—the probability that agents will exchange estimates of their future design projections instead of their current design values [33], [76]—was also included to demonstrate how deviating from assumptions can reveal more and less optimal strategies for CESDPs contingent upon other variables (see Section IV-C). Combined, the independent variable sets consisted of 13 552 unique combinations, each of which ran 100 times for a total of 1 355 200 executions of CESIUM.

According to the regression analysis, nearly all of the system outcomes depend strongly on the Number of Nodes, the convergence threshold, or both, though rarely through interaction effects. The number of nodes tended to increase the number of design cycles, while the convergence threshold consistently revealed an inverse relationship between system performance and the number of design cycles, regardless of the objective function selected. In contrast, the main and interaction effects of the triangle and future estimate probabilities proved scattered with only select cases generating significant effects. For example, the Absolute Sum function (Model II.3), Sphere function (Model II.4), and Ackley function (Model II.5) regressions identified statistically significant effects between the probability of triangle formation in the system network and the probability of sharing future estimates (Tri. Prob. × Fut. Est. Prob.) with magnitudes comparable to their respective constants, an effect that was limited in the Ackley and absent from the Levy function.

To further support the regressions and capture any nonlinear or higher-order effects of each variable, a feature importance analysis using a random forest regressor was applied, shown in Fig. 5. Random forests fit an ensemble of decision trees to predict the outcome variable, and feature importances describe which variables best reduce errors in outcome predictions when trees split the data on that variable [78]. Random forests were fit using the default parameters of sci-kit learn v0.24.1 with a maximum depth of four. The results proved largely consistent with the regressions with a few noteworthy exceptions. Again, the number of nodes and convergence thresholds proved most important. However, the triangle and future estimate probabilities only substantially affected the Sphere function performance, suggesting their effects may be limited to particular topologies. This somewhat contradicts the regression results which identified an interaction effect for the Absolute Sum and Ackley functions, suggesting that the regressions for those functions may have missed nonlinearities, higher-order effects, or excluded uncaptured variation in either the network structure or simulated annealing searches. Still, the Triangle and Future Estimate Probability contributions remained significant for the Sphere function.

B. Validation

Because CESIUM models organizational processes, this work draws from organization science for its definitions of micro- and macrovalidation. Microvalidation ensures that programmed behaviors and mechanisms sufficiently represent, or are "grounded" in, real-world analogs. Macrovalidation is divided into face validation—whether results are superficially distinguishable from real-world conceptual phenomena—and empirical validation—whether results are statistically calibrated to match real-world data [79], [80].

To satisfy the conditions of microvalidation, Section III grounds each construction decision of CESIUM in either established research or simplified abstractions of real-world practices. To satisfy face and empirical macrovalidation, simulation results would need to correspond to empirical outcomes. While face validation occasionally proves possible, insufficient data exists about CESDPs to perform empirical validation (discussed in the supplementary material Section I). Indeed, CESIUM is designed to address the absence of sufficient data by reconstructing mechanisms underlying development processes and generating data instead.

Consequently, it is beneficial to view modeling frameworks like CESIUM from a different perspective. Carley refers to computational models as "hypothesis generating machine[s]" [79], in CESIUM's case generating data with which to form hypotheses about causal variables in CESDPs. With sufficient evidence from multiple sources of knowledge, researchers can validate those hypotheses establish systems engineering theories. Thereafter, systems engineers can devise methods for improving the outcomes resulting from those relationships.

Contributing to that process, the following section demonstrates both the creation of hypotheses and formation of theory by validating the sensitivity analysis results.

C. Theory Building Demonstration

Consider the results of the sensitivity analysis for each independent variable:

Result 1: The number of design cycles tended to increase proportional to the *Number of Nodes*, while System Performance varied with the objective function.

Result 2: Both system outcomes varied depending on the *Objective Functions*, via main and interaction effects.

Result 3: Triangle Probability had small or statistically insignificant effects on system outcomes for all but the Sphere function, where system performance varied as a function of the triangle probability, future estimate probability, and interaction between the two.

Result 4: Increasing the size of *Convergence Threshold* tended to decrease the number of design cycles while degrading the system performance.

Result 5: Increasing the *Future Estimate Probability* had small or no effect on system outcomes for all but the Sphere function, where system performance varied through interaction with the triangle probability.

Result 1 suggests that even if system performance varies differently for different objective functions, (larger) systems with more artifacts will tend to take longer to converge on system designs. This observation can be posed as a hypothesis for further examination.

Hypothesis 1: The time it takes an organization to converge on a system design increases as a function of system size.

Scholars can examine this hypothesis with existing and new studies to determine whether it merits conversion into theory.

Next, note that Results 1–3 and 5 involve variation in system outcomes depending on the objective functions of the agents. This, too, could form a hypothesis. However, the finding is consistent with extensive research from game theory, decision science, and forms the underlying premise of value-driven design [24], thereby providing face validation for the hypothesis. With abundant evidence from multiple disciplines [31], [32], [38], [48], [81]–[85], the claim is more accurately stated as an established piece of knowledge. Recall, CESIUM assumes that the utilities of agents designing system artifacts are synonymous with the objective functions of the artifacts. This yields the following.

Proposition 1: The outcomes of a CES development process depend on the objective functions of the agents who develop system artifacts.

Variation in the other regression coefficients coincident with the objective functions leaves a question as to the relationship between even the most significant of the remaining variables and system outcomes. Nevertheless, CES outcomes certainly depended on the selected function giving this first proposition in combination with existing evidence.

The results lend themselves to two further hypotheses meriting examination. Result 4 suggests the following.

Hypothesis 2: Given a system objective function in negative null form, system performance varies inversely proportionately to system convergence time as a function of maximum uncertainty required in system performance.

Finally, recall that: 1) the triangle probability is a proxy for modularity or clustering in the system architecture's technical network and 2) the future estimate probability examines the likelihood that an agent will communicate information about either current design statuses or projected design outcomes. Combining these abstractions with Results 3 and 5 give the following.

Hypothesis 3: System performance varies as a function of interactions between modularity in the system architecture and the temporal references of information exchanged within the organization.

V. Future Directions

As a framework, CESIUM's conceptualization of CESDPs presents numerous opportunities for further exploration, including system architectures, system development processes, artifact design, patterns and contents of interaction, and objective functions. To facilitate exploration, this section begins by generalizing the framework (Section V-A). Then, it outlines recommended directions for future inquiry based on the framework to facilitate the development of theory (Section V-B).

A. Framework Generalization

The base instance of CESIUM makes simplifying assumptions about CES and CESDPs, including a scale-free degree distribution of artifacts, no exogenous variables, one agent per artifact via the mirroring hypothesis, simulated annealing artifact design, a design cycle development process, and a simple sum of artifact performances as representative of CES performance. The following paragraphs relax those assumptions and, in doing so, open opportunities for exploration.

Assume that a CES is composed of n_a artifacts interacting in a technical network and system architecture described by adjacency matrix A_{ij}. The n_b members of an organization (engineers and otherwise) can be represented by agents in an agent-based model which exchanges information with one another in a communication network described by adjacency matrix B_{qr}. Each of the n_b agents acts upon or "works on" some set of the n_a artifacts according to matrix C_{iq}. Summing over q then yields the number of agents that contribute to the design of each artifact $\lambda_i = \sum_q C_{iq}$.

Agents apply design decisions to each artifact they work on, influenced by specific and contextual factors. Describe the decisions made by agent $w \in \{1, \ldots, \lambda_i\}$ about the ith artifact with a decision variable vector $\mathbf{v}_{iw} = [v_{iw1}, v_{iw2}, \ldots]$ of contributions made by each agent working on the artifact. The contributions of any variables exogenous to the system can be described by $\mathbf{u}_i = [u_{i1}, u_{i2}, \ldots]$. As before, these variables can be representatively parameterized so that each artifact is modeled by a single decision variable $x_i(\mathbf{v}_{i1}, \ldots, \mathbf{v}_{i\lambda_i}, \mathbf{u}_i)$ (but this time of vectors $\mathbf{v}_{iw}$ and $\mathbf{u}_i$), thereby mapping a set of complex factors to a more manageable form.[28]

This parameterized specification for the design process of each artifact makes it possible to describe the performance for each as y_i which is a function of the artifact's own design x_i, the vector of designs $\mathbf{x}_j$ corresponding to the k_i interacting artifacts, and a vector of any exogenous variables $\mathbf{z}_i$ that influence performance. The combined notation $\mathbf{x}_i = [x_i, \mathbf{x}_j, \mathbf{z}_i]$ then allows the artifact's performance to be described in terms of one or more objective functions $\mathbf{y}_i = \mathbf{f}_i(\mathbf{x}_i)$. Over time, agents can decrease uncertainties in their designs, improve their own objective evaluations, improve CES objective evaluations, etc. Hence, given functions for each artifact and another vector of exogenous contributions to the system ξ, one or more objective or multiobjective functions can be formed quantifying overall outcomes for the CESDP, including system performance at different points in time and for context-specific measurement conditions. This gives us $\mathbf{F}(t, f_1(\mathbf{x}_1), \ldots, f_i(\mathbf{x}_i), \ldots, f_n(\mathbf{x}_n), \xi)$.

While abstract in this form, these generalizations enable researchers to vary assumptions within the CESIUM framework, from which it becomes possible to explore variable dependencies, pose hypotheses, and form theories about CESDPs. The following section enumerates some such options.

B. Recommended Investigations

Identifying the existence of trends is an important first step toward establishing a body of theories. From hypotheses and theories, other hypotheses and theories may arise which explain mechanisms through which these trends occur in practice. So the trends posed in Section IV-C should serve as starting points for investigating relationships between development process variables, through CESIUM and other approaches, both qualitative and quantitative.

Many other possibilities also remain. Real technical network distributions vary, so exploring bow-tie structures generated by HOT (Section II-C1) and complete graphs (in which every node is connected to every other node) could yield novel insights about how artifact interaction networks shape system outcomes. Again, the clustering parameter of the Holme–Kim algorithm leaves it well-suited to study subsystem formation and interaction, closely connecting it to current practices. Of course, artifact networks are rarely static in practice, so incorporating network evolution into CESIUM could shed light on how temporal variation affects both agent-level and system-level outcomes. And because CES can vary dramatically in size, exploring different scales may reveal how outcomes manifest as a result of hierarchical decomposition of systems and system of systems interactions.

Development processes themselves also merit exploration, including different permutations of the relationships between artifacts, variable composition, agent contribution, design space exploration, and development process stopping conditions. Artifacts undergo different development processes depending on the industry, design process, designer identity, and designer context [86], each of which could be studied through CESIUM. Likewise, many CESDPs employ processes that are more involved than the turn-based Shewhart and Deming Cycle (Section III-E), including Systems V, agile development, scrum, etc. The ability to explore such varied development processes may prove one of the greatest opportunities afforded by CESIUM.

The framework also opens possibilities for studying the contents of interactions between contributors. While researchers have begun to investigate communication frequency in other contexts, the contents of communication and interpretation have received limited treatment to date, thereby lending themselves to questions of feedback, synchronization [66], coordination between designers [87], and different mediums of information exchange. Direct communication between agents may also merit the introduction of communicative imprecision (as statistical uncertainty, ambiguity, missing information, etc.) [88] which prior works have shown affect outcomes [33], [76].

Different objective constructions could prove extremely informative at the agent, subsystem, system, and system of systems levels. The functions selected for this work began to explore concepts, such as linear relationships between variable contributions (Absolute Sum), parabolic relationships (Sphere), local optima (Ackley), and disproportionate influence by collaborators (Levy), though innumerable topologies remain possible at each level. System performance certainly takes forms other than sums of its parts, creating opportunities to study questions of collective innovation, high-reliability organizations, and evolving landscapes, from completely abstract to precisely representative of real-world systems. Hopefully these suggestions will contribute to the beginning of theory building for CES and CESDPs.

[28]Using a vector of decision variables instead of a parameterization is functionally equivalent and may prove more suitable for certain applications.

Finally, these investigations leave the important task of connecting theory to practice. While Section IV-B demonstrates that CESIUM is grounded in established knowledge, significant effort will be required to tune the model for predicting the performance of actual systems. This would necessitate identifying network structures, objective functions of real-world artifacts, system objectives, network evolution, exogenous variables, etc. While challenging, data science analyses can probably identify these qualities from existing systems to predict how the characteristics of systems will change. Coupling that knowledge with CESIUM could then facilitate more accurate and reliable predictions about CES and development process outcomes.

VI. Conclusion

CESs play essential roles in society. However, decades of increasing system complexity and development process inadequacies have garnered substantial concern about the capabilities of established systems engineering processes. Leading experts from government, academic, and industry organizations now call for the development of theories to guide solutions to these growing needs. But to date, the notorious difficulty of gathering data from real-world CESDPs has stymied efforts to form and validate generalizable hypotheses about how systems develop.

This work responds to that challenge by uniting recent advances from systems engineering, complex systems, engineering design, and organization science to form the CESIUM. Beginning from a review of terminology, the piece identifies the abstracted elements of development processes and integrates techniques for simulating each aspect into a modeling framework. It then demonstrates the ability of CESIUM to generate sizeable datasets representing the outcomes of system development processes from which researchers can form theory. The underlying flexibility of the framework leaves numerous opportunities for researchers to explore abstracted phenomena and real-world examples alike toward the construction of systems engineering theory.

Acknowledgment

Thanks to their colleagues who reviewed this work and supported them, including Abigail Jacobs, Vanessa Rojano, Mojtaba Arezoomand, Lynette Shaw, and Maciej Kurant.

References

[1] O. L. de Weck, D. Roos, and C. L. Magee, *Engineering Systems: Meeting Human Needs in a Complex Technological World*. Cambridge, MA, USA: MIT Press, 2011.

[2] D. Braha, A. Minai, and Y. Bar-Yam, *Complex Engineered Systems: Science Meets Technology*. Heidelberg, Germany: Springer, 2006.

[3] S. J. Kapurch (Nat. Aeronaut. Space Admin., Washington, DC, USA). *NASA Systems Engineering Handbook*, (2007). [Online]. Available: http://adsabs.harvard.edu/full/1995NASSP6105.....S

[4] W. E. Deming, *Out of the Crisis: Quality, Productivity and Competitive Position*. Cambridge, MA, USA: Massachusetts Inst. Technol. Center Adv. Eng. Study, 1986.

[5] R. G. Schroeder, K. Linderman, C. Liedtke, and A. S. Choo, "Six sigma: Definition and underlying theory," *J. Oper. Manag.*, vol. 26, no. 4, pp. 536–554, 2008.

[6] W. M. P. Van Der Aalst, "The application of petri nets to workflow management," *J. Circuits Syst. Comput.*, vol. 8, no. 1, pp. 21–66, Feb. 1998.

[7] S. Modi, M. K. Tiwari, Y. Lin, and W. J. Zhang, "On the architecture of a human-centered CAD agent system," *Comput.-Aided Des.*, vol. 43, no. 2, pp. 170–179, Feb. 2011.

[8] W. J. Zhang, J. W. Wang, and Y. Lin, "Integrated design and operation management for enterprise systems," *Enterp. Inf. Syst.*, vol. 13, no. 4, pp. 424–429, Apr. 2019.

[9] P. D. Collopy, "A research agenda for the coming renaissance," in *Proc. 50th AIAA Aerosp. Sci. Meeting Including New Horizons Forum Aerosp. Exposit.*, Jan. 2012, pp. 1–10.

[10] I. Maddox, P. Collopy, and P. A. Farrington, "Value-based assessment of DoD acquisitions programs," *Procedia Comput. Sci.*, vol. 16, pp. 1161–1169, Jan. 2013.

[11] C. L. Bloebaum, P. D. Collopy, and G. A. Hazelrigg, "NSF/NASA workshop on the design of large-scale complex engineered systems—From research to product realization," in *Proc. 12th AIAA Aciation Technol. Integr. Oper. (ATIO) Conf.*, Sep. 2012, pp. 1–20.

[12] C. L. Bloebaum and A.-M. R. McGowan, "The design of large-scale complex engineered systems: Present challenges and future promise," in *Proc. AIAA Aviation Technol. Integr. Oper. (ATIO) Conf. 14th AIAA/ISSM*, Sep. 2012, pp. 1–19.

[13] Y. Bar-Yam, "When systems engineering fails-toward complex systems engineering," in *Proc. Conf. IEEE Int. Conf. Syst. Man Cybern. Conf. Theme Syst. Security Assur. (SMC)*, vol. 2, 2003, pp. 2021–2028.

[14] C. Perrow, *Normal Accidents: Living With High Risk Technologies-Updated Edition*. Princeton, NJ, USA: Princeton Univ. Press, 2011.

[15] A. M. Madni and S. Jackson, "Towards a conceptual framework for resilience engineering," *IEEE Syst. J.*, vol. 3, no. 2, pp. 181–191, Jun. 2009.

[16] P. D. Collopy, "Systems engineering theory: What needs to be done," in *Proc. Annu. IEEE Syst. Conf. (SysCon)*, 2015, pp. 536–541.

[17] K. P. Triantis and P. D. Collopy, "A comprehensive basis for systems engineering theory," in *Proc. IEEE Int. Syst. Conf.*, 2014, pp. 97–102.

[18] D. Salvucci-Favier, *A Complexity Primer for Systems Engineers*, INCOSE Found., San Diego, CA, USA, Jul. 2016, pp. 1–5.

[19] S. Safarkhani, I. Bilionis, and J. H. Panchal, "Toward a theory of systems engineering processes: A Principal–Agent model of a one-shot, shallow process," *IEEE Syst. J.*, vol. 14, no. 3, pp. 3277–3288, Sep. 2020.

[20] S. H. Young, T. A. Mazzuchi, and S. Sarkani, "A framework for predicting future system performance in autonomous unmanned ground vehicles," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 7, pp. 1192–1206, Jul. 2017.

[21] J. Cecil, S. Albuhamood, A. Cecil-Xavier, and P. Ramanathan, "An advanced cyber physical framework for micro devices assembly," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 1, pp. 92–106, Jan. 2019.

[22] G. Fortino, C. Savaglio, G. Spezzano, and M. Zhou, "Internet of Things as system of systems: A review of methodologies, frameworks, platforms, and tools," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 1, pp. 223–236, Jan. 2021.

[23] P. D. Collopy, "Report on the science of systems engineering workshop," in *Proc. 53rd AIAA Aerosp. Sci. Meeting*, 2015, p. 1865.

[24] P. D. Collopy and P. M. Hollingsworth, "Value-driven design," *J. Aircraft*, vol. 48, no. 3, pp. 749–759, May 2011.

[25] J. R. R. A. Martins and A. B. Lambe, "Multidisciplinary design optimization: A survey of architectures," *AIAA J.*, vol. 51, no. 9, pp. 2049–2075, 2013.

[26] K. M. Carley, "Computational organization science: A new frontier," *Proc. Nat. Acad. Sci.*, vol. 99, no. s3, pp. 7257–7262, May 2002.

[27] J. R. Harrison, Z. Lin, G. R. Carroll, and K. M. Carley, "Simulation modeling in organizational and management research," *Acad. Manag. Rev.*, vol. 32, no. 4, pp. 1229–1245, Oct. 2007, doi: 10.5465/amr.2007.26586485.

[28] G. Fioretti, "Agent-based simulation models in organization science," *Organ. Res. Methods*, vol. 16, no. 2, pp. 227–242, Dec. 2012, doi: 10.1177/1094428112470006.

[29] G. P. Morgan and K. M. Carley, "Computational organizational theory," in *International Encyclopedia of the Social & Behavioral Sciences*, 2nd ed. Amsterdam, The Netherlands: Elsevier, 2015, pp. 485–492.

[30] R. M. Crowder, M. A. Robinson, H. P. Hughes, and Y.-W. Sim, "The development of an agent-based modeling framework for simulating engineering team work," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 42, no. 6, pp. 1425–1439, Nov. 2012.

[31] C. McComb, J. Cagan, and K. Kotovsky, "Optimizing design teams based on problem properties: Computational team simulations and an applied empirical test," *J. Mech. Des.*, vol. 139, no. 4, 2017, Art. no. 41101.

[32] J. Meluso and J. Austin-Breneman, "Gaming the system: An agent-based model of estimation strategies and their effects on system performance," *J. Mech. Des.*, vol. 140, no. 12, pp. 121101–121109, Sep. 2018, doi: 10.1115/1.4039494.

[33] J. Meluso, J. Austin-Breneman, and L. Shaw, "An agent-based model of miscommunication in complex system engineering organizations," *IEEE Syst. J.*, vol. 14, no. 3, pp. 3463–3474, Sep. 2020.

[34] M. Newman, *Networks*. Oxford, U.K.: Oxford Univ. Press, 2018.

[35] M. E. J. Newman, "Resource letter CS1: Complex systems," *Amer. J. Phys.*, vol. 79, no. 8, pp. 800–810, Jul. 2011.

[36] M. Mitchell, *Complexity: A Guided Tour*. Oxford, U.K.: Oxford Univ. Press, 2009.

[37] T. W. Simpson and J. R. R. A. Martins, "Multidisciplinary design optimization for complex engineered systems: Report from a national science foundation workshop," *J. Mech. Des.*, vol. 133, no. 10, 2011, Art. no. 101002.

[38] M. E. Sosa, J. Mihm, and T. Browning, "Degree distribution and quality in complex engineered systems," *J. Mech. Des.*, vol. 133, no. 10, 2011, Art. no. 101008.

[39] Office of the Deputy Under Secretary of Defense for Acquisition and Technology Systems and Software Engineering, "Systems engineering guide for systems of systems, version 1.0," U.S. Dept. Defense, Washington, DC, USA, Rep., 2008. [Online]. Available: https://www.sebokwiki.org/wiki/Systems_Engineering_Guide_for_Systems_of_Systems

[40] *System Life Cycle Processes*, ISO/IEC/IEEE International Standard 15288, May 2015.

[41] A. Gorod, B. Sauser, and J. Boardman, "System-of-systems engineering management: A review of modern history and a path forward," *IEEE Syst. J.*, vol. 2, no. 4, pp. 484–499, Dec. 2008.

[42] J. Dahmann and G. Roedler, "Systems of systems engineering standards," *Insight*, vol. 19, no. 3, pp. 23–26, Oct. 2016, doi: 10.1002/inst.12102.

[43] C. Keating *et al.*, "System of systems engineering," *Eng. Manag. J.*, vol. 15, no. 3, pp. 36–45, Sep. 2003, doi: 10.1080/10429247.2003.11415214.

[44] A. Deshmukh and P. Collopy, "Fundamental research into the design of large-scale complex systems," in *Proc. 13th AIAA/ISSMO Multidiscipl. Anal. Optim. Conf.*, Sep. 2010, pp. 1–20, doi: 10.2514/6.2010-9320.

[45] K. E. Lewis and P. D. Collopy, "The role of engineering design in large-scale complex systems," in *Proc. 12th AIAA Aviation Technol. Integr. Oper. (ATIO) Conf. 14th AIAA/ISSM*, 2012, pp. 1–9.

[46] B. Gokpinar, W. J. Hopp, and S. M. R. Iravani, "The impact of misalignment of organizational structure and product architecture on quality in complex product development," *Manag. Sci.*, vol. 56, no. 3, pp. 468–484, 2010.

[47] R. Katz and T. J. Allen, "Project performance and the locus of influence in the R&D matrix," *Acad. Manag. J.*, vol. 28, no. 1, pp. 67–87, 1985.

[48] M. Natter, A. Mild, M. Feurstein, G. Dorffner, and A. Taudes, "The effect of incentive schemes and organizational arrangements on the new product development process," *Manag. Sci.*, vol. 47, no. 8, pp. 1029–1045, 2001.

[49] H. Li, J. B. Bingham, and E. E. Umphress, "Fairness from the top: Perceived procedural justice and collaborative problem solving in new product development," *Org. Sci.*, vol. 18, no. 2, pp. 200–216, 2007.

[50] D. Gerwin and N. J. Barrowman, "An evaluation of research on integrated product development," *Manag. Sci.*, vol. 48, no. 7, pp. 938–953, 2002.

[51] N. Melville, K. Kraemer, and V. Gurbaxani, "Information technology and organizational performance: An integrative model of IT business value," *Manag. Inf. Syst. Quart.*, vol. 28, no. 2, pp. 283–322, 2004.

[52] K. Atuahene-Gima and F. Evangelista, "Cross-functional influence in new product development: An exploratory study of marketing and R&D perspectives," *Manag. Sci.*, vol. 46, no. 10, pp. 1269–1284, 2000.

[53] R. Kapoor and R. Adner, "What firms make vs. what they know: How firms' production and knowledge boundaries affect competitive advantage in the face of technological change," *Org. Sci.*, vol. 23, no. 5, pp. 1227–1248, Aug. 2011, doi: 10.1287/orsc.1110.0686.

[54] G. S. Saffold, "Culture traits, strength, and organizational performance: Moving beyond 'strong' culture," *Acad. Manag. Rev.*, vol. 13, no. 4, pp. 546–558, Oct. 1988, doi: 10.5465/amr.1988.4307418.

[55] S. T. Collins, A. A. Yassine, and S. P. Borgatti, "Evaluating product development systems using network analysis," *Syst. Eng.*, vol. 12, no. 1, pp. 55–68, Mar. 2009, doi: 10.1002/sys.20108.

[56] B. J. Zirger and M. A. Maidique, "A model of new product development: An empirical test," *Manag. Sci.*, vol. 36, no. 7, pp. 867–883, 1990.

[57] J. Austin-Breneman, "Observations of designer behaviors in complex system design," Ph.D. dissertation, Dept. Mech. Eng., Massachusetts Inst. Technol., Cambridge, MA, USA, 2011.

[58] A.-L. Barabási, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Phys. Stat. Mech. Appl.*, vol. 272, no. 1, pp. 173–187, 1999.

[59] J. M. Carlson and J. Doyle, "Highly optimized tolerance: A mechanism for power laws in designed systems," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 60, no. 2, pp. 1412–1427, Aug. 1999.

[60] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *SIAM Rev.*, vol. 51, no. 4, pp. 661–703, Nov. 2009, doi: 10.1137/070710111.

[61] D. Braha and Y. Bar-Yam, "The statistical mechanics of complex product development: Empirical and analytical results," *Manag. Sci.*, vol. 53, no. 7, pp. 1127–1145, 2007.

[62] M. E. Sosa, S. D. Eppinger, and C. M. Rowles, "A network approach to define modularity of components in complex products," *J. Mech. Des.*, vol. 129, no. 11, pp. 1118–1129, 2007.

[63] Y. Bar-Yam, "Engineering complex systems: Multiscale analysis and evolutionary engineering," in *Complex Engineered Systems*. Heidelberg, Germany: Springer, 2006, pp. 22–39.

[64] J. Austin-Breneman, B. Y. Yu, and M. C. Yang, "Biased information passing between subsystems over time in complex system design," *J. Mech. Des.*, vol. 138, no. 1, 2015, Art. no. 11101.

[65] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[66] L. M. Pecora and T. L. Carroll, "Master stability functions for synchronized coupled systems," *Phys. Rev. Lett.*, vol. 80, no. 10, pp. 2109–2112, Mar. 1998.

[67] P. Holme and B. J. Kim, "Growing scale-free networks with tunable clustering," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 2, 2002, Art. no. 26107.

[68] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proc. Nat. Acad. Sci. U. S. A.*, vol. 99, no. S3, pp. 7280–7287, May 2002.

[69] C. Macal and M. North, "Tutorial on agent-based modelling and simulation," *J. Simul.*, vol. 4, no. 3, pp. 151–162, 2010.

[70] B. Jamshidnezhad and K. M. Carley, "Agent-based modelling of quality management effects on organizational productivity," *J. Simul.*, vol. 9, no. 1, pp. 73–82, Feb. 2015, doi: 10.1057/jos.2014.26.

[71] C. M. Macal, "Everything you need to know about agent-based modelling and simulation," *J. Simul.*, vol. 10, no. 2, pp. 144–156, May 2016, doi: 10.1057/jos.2016.7.

[72] P. Y. Papalambros and D. J. Wilde, *Principles of Optimal Design: Modeling and Computation*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[73] J. Cagan and K. Kotovsky, "Simulated annealing and the generation of the objective function: A model of learning during problem solving," *Comput. Intell.*, vol. 13, no. 4, pp. 534–581, 1997.

[74] K. M. Carley and D. M. Svoboda, "Modeling organizational adaptation as a simulated annealing process," *Sociol. Methods Res.*, vol. 25, no. 1, pp. 138–168, Aug. 1996, doi: 10.1177/0049124196025001005.

[75] L. J. Colfer and C. Y. Baldwin, "The mirroring hypothesis: Theory, evidence, and exceptions," *Ind. Corp. Change*, vol. 25, no. 5, pp. 709–738, Oct. 2016, doi: 10.1093/icc/dtw027.

[76] J. Meluso, J. Austin-Breneman, and J. Uribe, "Estimate uncertainty: Miscommunication about definitions of engineering terminology," *J. Mech. Des.*, vol. 142, no. 7, p. 9, Feb. 2020.

[77] G. T. Broeke, G. van Voorn, and A. Ligtenberg, "Which sensitivity analysis method should I use for my agent-based model?" *J. Artif. Soc. Soc. Simul.*, vol. 19, no. 1, p. 5, 2016.

[78] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[79] K. M. Carley, "Validating computational models," Dept. School Comput. Sci., Inst. Softw. Res., Carnegie Mellon Univ., Pittsburgh, PA, USA, Rep. CMU-ISR-17-105, 2017. [Online]. Available: http://www.casos.cs.cmu.edu/publications/papers/CMU-ISR-17-105.pdf

[80] U. Wilensky and W. Rand, *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with Netlogo*. Cambridge, MA, USA: MIT Press, 2015.
[81] J. C. Hershey, H. C. Kunreuther, and P. J. H. Schoemaker, "Sources of bias in assessment procedures for utility functions," *Manag. Sci.*, vol. 28, no. 8, pp. 936–954, Aug. 1982.
[82] K.-J. Kim and D. K. J. Lin, "Simultaneous optimization of mechanical properties of steel by maximizing exponential desirability functions," *J. Roy. Stat. Soc. C, Appl. Stat.*, vol. 49, no. 3, pp. 311–325, 2000.
[83] Y. Kainuma and N. Tawara, "A multiple attribute utility theory approach to lean and green supply chain management," *Int. J. Prod. Econ.*, vol. 101, no. 1, pp. 99–108, May 2006.
[84] Z. Laslo and A. I. Goldberg, "Resource allocation under uncertainty in a multi-project matrix environment: Is organizational conflict inevitable?" *Int. J. Project Manag.*, vol. 26, no. 8, pp. 773–788, Nov. 2008.
[85] A. E. Abbas, J. E. Matheson, and R. F. Bordley, "Effective utility functions induced by organizational target-based incentives," *Manag. Decis. Econ.*, vol. 30, no. 4, pp. 235–251, 2009.
[86] S. Chou, "Exploring 'designer context' in engineering design: The relationship between self, environment, and design methods," Ph.D. dissertation, Dept. Eng. Des., Univ. Michigan, Ann Arbor, MI, USA, 2020. [Online]. Available: http://hdl.handle.net/2027.42/163100
[87] G. Wen, X. Yu, W. Yu, and J. Lü, "Coordination and control of complex network systems with switching topologies: A survey," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 10, pp. 6342–6357, Oct. 2021.
[88] M. Cai, Y. Lin, B. Han, C. Liu, and W. Zhang, "On a simple and efficient approach to probability distribution function aggregation," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 9, pp. 2444–2453, Sep. 2017.
[89] J. Meluso, S. Johnson, and J. Bagrow, "Flexible environments for hybrid collaboration: Redesigning virtual working through the four orders of design," *Des. Issues*, vol. 38, no. 1, pp. 55–69, 2022.
[90] T. G. Topcu, K. Triantis, R. Malak, and P. Collopy, "An interdisciplinary strategy to advance systems engineering theory: The case of abstraction and elaboration," *Syst. Eng.*, vol. 23, no. 6, pp. 673–683, 2020.
[91] W. N. Felder and P. Collopy, "The elephant in the mist: What we don't know about the design, development, test and management of complex systems," *J. Aerosp. Oper.*, vol. 1, no. 4, pp. 317–327, Jan. 2012.
[92] X. Liu, W. Li, Y. L. Tu, and W. J. Zhang, "An expert system for an emergency response management in networked safe service systems," *Expert Syst. Appl.*, vol. 38, no. 9, pp. 11928–11938, Sep. 2011.
[93] K. M. Adams, P. T. Hester, J. M. Bradley, T. J. Meyers, and C. B. Keating, "Systems theory as the foundation for understanding systems," *Syst. Eng.*, vol. 17, no. 1, pp. 112–123, 2014.
[94] C. B. Keating, P. F. Katina, and J. M. Bradley, "Complex system governance: Concept, challenges, and emerging research," *Int. J. Syst. Syst. Eng.*, vol. 5, no. 3, pp. 263–288, Jan. 2014.
[95] P. F. Katina, A. Tolk, C. B. Keating, and K. F. Joiner, "Modelling and simulation in complex system governance," *Int. J. Syst. Syst. Eng.*, vol. 10, no. 3, pp. 262–292, Jan. 2020.
[96] J. P. Davis, K. M. Eisenhardt, and C. B. Bingham, "Developing theory through simulation methods," *Acad. Manag. Rev.*, vol. 32, no. 2, pp. 480–499, Apr. 2007.
[97] T. W. Zuzul, "'Matter battles': Cognitive representations, boundary objects, and the failure of collaboration in two smart cities," *Acad. Manag. J.*, vol. 62, no. 3, pp. 739–764, Aug. 2018, doi: 10.5465/amj.2016.0625.
[98] E. B. Smith and W. Rand, "Simulating macro-level effects from micro-level observations," *Manag. Sci.*, vol. 64, no. 11, pp. 5405–5421, 2017.
[99] P. M. Leonardi, "Materiality, sociomateriality, and socio-technical systems: What do these terms mean? How are they different? Do we need them," in *Materiality and Organizing: Social Interaction in a Technological World*, vol. 25. Oxford, U.K.: Oxford Univ. Press, 2012.
[100] D. L. Alderson and J. C. Doyle, "Contrasting views of complexity and their implications for network-centric infrastructures," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 40, no. 4, pp. 839–852, Jul. 2010.
[101] A.-M. R. McGowan, S. Daly, W. Baker, P. Papalambros, and C. Seifert, "A socio-technical perspective on interdisciplinary interactions during the development of complex engineered systems," *Procedia Comput. Sci.*, vol. 16, pp. 1142–1151, Jan. 2013.
[102] P. S. Adler, B. Goldoftas, and D. I. Levine, "Flexibility versus efficiency? A case study of model changeovers in the Toyota production system," *Org. Sci.*, vol. 10, no. 1, pp. 43–68, 1999. [Online]. Available: http://www.jstor.org/stable/2640387
[103] W. Zhang and J. Wang, "Design theory and methodology for enterprise systems," *Enterp. Inf. Syst.*, vol. 10, no. 3, pp. 245–248, Mar. 2016.
[104] W. R. Scott and G. F. Davis, *Organizations and Organizing: Rational, Natural, and Open System Perspectives*, 1st ed., W. R. Scott, Ed. Upper Saddle River, NJ, USA: Pearson, 2007. [Online]. Available: http://www.worldcat.org/oclc/70839752
[105] C. Macal and M. North, "Agent-based modeling and simulation," in *Proc. Winter Simul. Conf. (WSC)*, 2009, pp. 86–98.
[106] J. C. Anderson and M. Rungtusanatham, "A theory of quality management underlying the deming management method," *Acad. Manag. Rev.*, vol. 19, no. 3, pp. 472–509, Jul. 1994.

John Meluso received the B.S. degree in engineering physics from Cornell University, Ithaca, NY, USA, in 2014, and the M.S. and Ph.D. degrees in design science from the University of Michigan at Ann Arbor, Ann Arbor, MI, USA, in 2017 and 2020, respectively.

He is currently the OCEAN Postdoctoral Fellow with the Vermont Complex Systems Center, University of Vermont, Burlington, VT, USA. His research interests include systems engineering, complex systems, collective problem solving, organization science, communication, and diversity, equity, and inclusion.

Jesse Austin-Breneman received the B.S. degree in ocean engineering and the M.S. and Ph.D. degrees in mechanical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2005, 2011, and 2014, respectively.

He is an Assistant Professor with the Mechanical Engineering Department, University of Michigan at Ann Arbor, Ann Arbor, MI, USA. His research interests include complex system design, multidisciplinary design optimization, modeling of team-based design decision making, design for emerging markets, socially engaged design, and design for entrepreneurship.

James P. Bagrow received the A.S. degree in liberal arts and sciences from SUNY Cobleskill, Cobleskill, NY, USA, in 2001, and the B.S., M.S., and Ph.D. degrees in physics from Clarkson University, Potsdam, NY, USA, in 2004, 2006, and 2008, respectively.

He is currently an Associate Professor with the Mathematics and Statistics Department and the Vermont Complex Systems Center, University of Vermont, Burlington, VT, USA. He uses mathematical models and large-scale data analysis to study the underlying rules and organizing principles of complex physical and social systems. Other interests include dynamical systems, data science, and machine learning.

Laurent Hébert-Dufresne received the B.Sc., M.Sc., and Ph.D. degrees in physics from Université Laval, Quebec City, QC, Canada in 2009, 2011, and 2014, respectively.

He is currently an Associate Professor with the Department of Computer Science and the Vermont Complex Systems Center, University of Vermont, Burlington, VT, USA, and a Professeur Associé with the Départment de Physique, Université Laval. His research interests include complexity, network theory, and nonlinear dynamics in epidemiology, biology, sociology, and ecology.