CCC's Response to the [Networking and Information Technology Research and Development Request for Information on Digital Twins Research and Development]

This response is prepared by the Computing Research Association (CRA)'s Computing Community Consortium (CCC) by inviting CCC Council members with interest and knowledge of the use of digital twins to a roundtable discussion. The participants discussed the RFI and contributed to this written response document. CRA is an association of over 270 North American computing research organizations, both academic and industrial, and partners from six professional computing societies.

The mission of the CCC, a subcommittee of CRA, is to enable the pursuit of innovative, high-impact computing research that aligns with pressing national and global challenges. Please note any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the authors' affiliations, or of the National Science Foundation, which funds the CCC.

This document is approved for public dissemination. The document contains no business-proprietary or confidential information. Document contents may be reused by the government in the National Digital Twins R&D Strategic Plan and associated documents without attribution.

July 28th, 2024

Written by: *David Danks (University of California, San Diego), Catherine Gill (Computing Community Consortium), Chandra Krintz (University of California, Santa Barbara), Brian LaMacchia (Farcaster Consulting Group), Daniel Lopresti (Lehigh University), Mary Lou Maher (Computing Community Consortium), and Pamela Wisniewski (Vanderbilt University).*

Digital twin models have the potential to revolutionize the ways we conduct R&D and to evaluate and measure the performance of systems and physical entities across research sectors. Digital twins can significantly reduce the costs and risks associated with conducting testing, speeding the rate of development and improving overall product design and safety considerations. Digital twins can also be programmed to offer personalized solutions which can be tailored to specific projects or individuals, and the application areas are nearly limitless. However, these models can also cause harm when they are poorly designed and/or implemented, or when we assume the results from these simulated environments seamlessly translate to real world outcomes. Below, we enumerate several critical considerations when developing and implementing digital twins.

To begin with, the term “digital twin” can be misleading for a number of reasons. “Digital” implies that these models live online, only in the digital sphere. However, this implication is incorrect, as seen by the definition used by this Request for Information. Digital twins utilize sensors in the real world to more accurately mimic their physical counterparts. In addition, the term “twin” is misleading as well, because though incredible advances have allowed digital models to progress, we should not be under the assumption that these models are perfect mimics. Digital twins may overvalue data received by certain sensors and conversely may not have enough sensors to perfectly reflect their physical counterparts. Furthermore, there are many varieties of digital twins, such as digital human twins versus twins of environments (e.g., virtual reality). A well-defined taxonomy of the different forms of digital twins should be established to create a shared language between different communities.

Digital twins could potentially be created for almost any physical entity, from modeled organs to digital twins of entire cities. Specific constraints of these systems, however, should be taken into account. Data limitations especially are something to consider. Digital twins cannot take into account every single piece of data in some instances, as this would overload the model. A digital twin meant to measure erosion, for example, could not take into account the minute movements of every individual grain of sand on a beach - this volume of data would overload the system. Instead, a model meant to measure erosion of a certain beach would have to approximate or group together thousands of grains of sand, meaning a certain amount of precision will be lost by the model. The same goes for climate models and models of ocean currents. Learned history by a digital twin can practically only be stored for a certain amount of time or at a certain granularity, leading to a degree of precision loss over time. Digital twins are closed systems, which are constrained to underlying assumptions (i.e. assuming a certain climate model or scenarios under which the model will operate), so we cannot hold all values at a constant and assume the model is robust and will always give

accurate results when physical models exist in complex systems that may affect actual results.

Alternatively, expecting too many specific requirements and capabilities of a single digital twin may overload the model or lead to inaccurate results. For example, a single digital twin meant to mimic the world's climate may perform very well, but if it is asked also to begin accurately modeling the weather patterns of specific locales, this may confuse the model, leading to inaccurate results across all of its approximations. This is also a concern with digital twins for cities, which attempt to model a variety of diverse processes at once (i.e., traffic flow, energy consumption, weather patterns, impacts of recent policy measures, etc.).

We also advise caution on using digital twins which rely heavily on surveillance data, such as CCTV footage. Video footage of individuals, while important to the efficacy of many digital twin models, especially digital twins of cities, can capture and store sensitive personal information without individuals' consent.

We believe the **Trustworthy** topic area identified in the RFI will be a critical component of any Digital Twins R&D Strategic Plan, and we recommend expanding and renaming this topic area to "**Security, Integrity, and Trust"** of digital twins. As the NASEM report points out on p. 36, the tight integration between a physical system and its digital twin creates a new attack surface for the physical system. Digital twins, especially those that store sensitive data or are coupled to physical critical infrastructure, will be prime targets for adversarial attacks[1]. In instances such as these, digital twins can act as massive security vulnerabilities, allowing attackers unfettered access to both the digital twin and physical entity. To prevent this kind of unauthorized access, developers must secure all endpoints, which can be tricky, especially when digital twins may be receiving data from outside systems operating on outdated legacy code. Digital twins must also be developed with security considerations in mind, not as an afterthought, and every aspect of the system, from where the data is coming from to the actual digital twin model itself must be secured.

Similarly, the coupling between a digital twin and its physical system counterpart also creates a new attack surface for the digital twin, because it may be possible to maliciously manipulate or corrupt the digital twin via the physical system. In the digital domain, cryptographic algorithms and protocols may be used to ensure the integrity, authenticity, and confidentiality of all the components of a digital twin; similar security guarantees will have to exist for coupled physical systems.

[1] For example, a digital twin of a power plant that has access to the control mechanisms of the physical power plant.

We also note that there will need to be close coordination between the **Security, Integrity, and Trust** and the **Standards** portions of the Strategic Plan. Creating secure and interoperable digital twins will require agreement on standards for data encryption, digital signatures, authentication protocols, authorization models, and policy language. New standards and language for constrained delegation models (e.g., the ability for the owner of a digital twin to delegate a portion of their access to other entities for legitimate purposes) will need to be developed, coded, and standardized. Standardizing security considerations alone will be very complex given a lack of interoperability standards.

Additionally, digital twins cannot be developed in siloed environments. Though only a small team may be necessary to write the code for a digital model, experts across disciplines need to be consulted to make sure the models are implemented accurately. Engineers can inform developers of optimal sensor placements and can verify that the digital representation accurately matches the physical one. Security experts and data privacy officers can help prevent adversarial attacks and data leaks. Depending on the application area of a digital twin (i.e., agriculture, economics, social sciences, etc.) experts from those domains must also be consulted. The NASEM report on digital twins enumerates several recommendations for federal agencies to improve cross-agency and cross-community collaborations.

Co-design of these systems is also incredibly important. When we use the term “co-design”, we are referring to the conceptualization of the key features of a digital twin as well as the end user or primary stakeholders who will be using the model. Co-design cannot focus solely on design of the model without considering who will use the model and how it will be deployed. Users must be informed of the system constraints, possible edge cases where the model may not deliver reliable results, and unacceptable use cases that the system was not designed for. Developing key features while keeping end users in mind can ensure the models are accurate, comprehensible to those outside of the development team, and most importantly, useful to the organization.

While digital twin models offer transformative potential across numerous research sectors by reducing costs, minimizing risks, and enhancing the precision of product designs, their implementation demands careful consideration. The CCC strongly advises that digital twins be viewed as what they are: tools for prediction and approximation, not prophetic devices that should replace decision makers. These tools should also not be viewed as being so important that they are deserving of funding without robust methodologies or evaluation plans. A nuanced understanding of these models' constraints, such as data overload and the necessity of approximations, is critical to their effective deployment. Furthermore, the security and privacy concerns

associated with digital twins necessitate robust safeguards to protect sensitive information and prevent adversarial attacks. Again, digital twins are not novel, they are an existing technology that has recently been bolstered through recent innovations, including the development of real time sensors, AI, and virtual reality. For digital twins to be truly valuable resources, we need to assess their value and what affordances are needed before implementing them widely across sectors.