EPRI | ELECTRIC POWER RESEARCH INSTITUTE

DIGITAL TWINS – A PROMISING TECHNOLOGY FOR DATA-CENTRIC CYBER SECURITY

January 2021

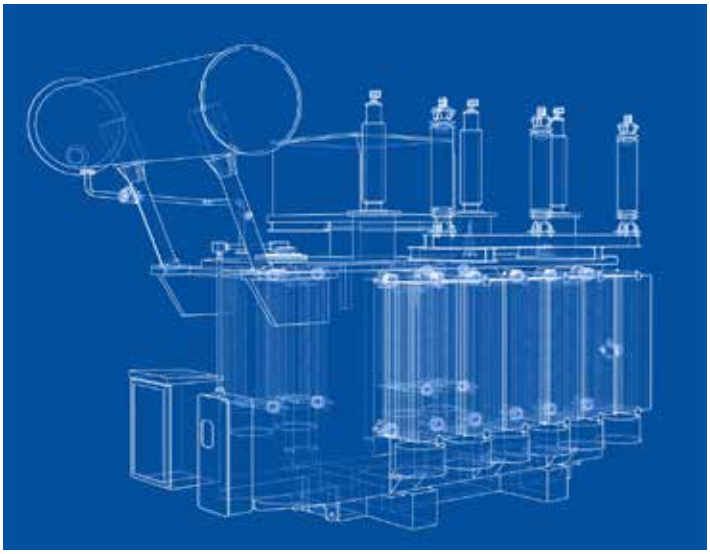
Digital utility initiatives can transform production environments through the deployment of digital technologies and data-centric operations. These transformations significantly impact operations technology (OT) cyber security policies, practices, and technologies that are tasked with protecting mission-critical grid operations. For OT cyber security, data-centricity is a dual-edged sword. It increases the number and type of vulnerabilities that OT cyber security resources must address. However, data-centric operations also offer new opportunities for utilities to improve their OT cyber security capabilities. One of those data-centric opportunities is the use of digital twins.

This white paper explores intriguing questions for electric utility executives and operations technology (OT) cyber security professionals planning digital transformation initiatives. Could digital twin technologies, with proven value in applications as varied as predictive maintenance and process design, help utilities manage cyber security risks in OT environments? What ramifications does a data-centric approach to OT cyber security engender for the development and deployment of digital twins? This white paper discusses important industry drivers and cyber security trends and explains our conclusions and research recommendations.

The research required to answer these questions included a review of industry literature, podcasts and webcasts; interviews with digital twin technology vendors; and exploration of existing digital twins that have parallels to cyber security applications. EPRI's conclusion is that digital twins have probable value to electric utilities in OT cyber security management. Applied demonstrations and additional research are needed to validate this conclusion, investigate use cases, and measure the full benefits utilities could expect from OT cyber security digital twins.

Digital Twins – Demonstrating Value in Industry Sectors

What is a digital twin? Definitions vary based on which company or individual is providing the answer and are often contextual. The Digital Twin Consortium offers this definition: *A digital twin is an abstraction of something in the real world. It may be physical (a device, product, system or other asset) or conceptual (a service, process or notion). A digital twin captures the behavior and attributes of its physical sibling with data and life cycle state changes potentially moving in either, or both, directions. A digital twin may be used for simulation, as a kind of prototype to understand expected behavior, existing before*



there is a physical twin. It can also capture real-world behavior so that, for example, analytics and learning can be performed. Digital twins can also be used in virtual reality (VR) and augmented reality (AR).

Bentley Systems, a vendor with experience in the electricity subsector, defined digital twin more specifically for electric systems: *An electrical digital twin provides a digital representation of the grid and the major assets in it. The twin facilitates the simulation of all technical and economical aspects relevant for reliable, efficient and secure electrical system planning, operation and maintenance.*

EPRI researchers have proposed a couple of definitions too. This version is from a recent whitepaper that focused on digital workers: *A digital twin is a machine-readable, 3D, interactive, geospatially accurate model of an asset or assembly of assets, continuously updated from datasets representing past, present, and future configuration and operating parameters.*

A digital twin definition contextualized for cyber security will be offered in this paper to help spur thinking about its potential in this specific functional application. Here are three facts about digital twins.

- Products, processes, and systems can be twinned.
- Twins can support a range of functions from design to operations of product, processes, and systems.
- They have proven value in industrial applications for complex and/or expensive assets and processes.

A digital twin models complex equipment like turbines. Siemens and General Electric are among the companies that utilize digital twins to improve management of electric generation equipment. The digital twins that are deployed within utilities are typically used for predictive maintenance of generation components. Digital twins are also used for complex processes found in industrial chemical manufacturing.

IDC market research[1] made predictions in 2019 about utility technologies that included one focused on digital twins. By 2023 50% of utilities will use digital twins for transmission and distribution asset management. The anticipated benefit was a reduction in simulation model development time with a concomitant reduction in engineering costs. In somewhat of a contrasting view, a recent digital twin webinar asked utility attendees this question: Have you taken any steps to adopt the digital twin technology?
The answers were:

> Digital twins are slowly gaining ground in electric utilities. Can they play a part to manage OT cyber security vulnerabilities and mitigate risks?

- 0% - Yes, have a digital twin strategy and plan.
- 25% - Yes, on a small scale and in some areas of the business.
- 50% - No, but it is being evaluated.
- 25% - No, we are not sure how and where to start.

Utilities are slowly climbing the digital twin learning curve, but the adoption rate may lag from the most optimistic industry predictions. Gartner[2] conducted a survey in 2019 about digital twins among a range of companies deploying Internet of Things (IOT) projects. While only 13% already had a digital twin in production, 62% planned to or were establishing a digital twin.

Cyber security receives little discussion in existing literature as an area that may benefit from digital twin applications. But the increasing product, process and system complexity of OT cyber security will require innovative solutions, and digital twins deserve a serious look as data-centric risk management tools.

[1] https://www.idc.com/research/viewtoc.jsp?containerId=US44327819
[2] https://www.gartner.com/en/newsroom/press-releases/2019-02-20-gartner-survey-reveals-digital-twins-are-entering-mai

Drivers for Digital Twin Adoption

Digital transformations are occurring in utilities today and impacting OT cyber security technologies, policies, and practices. Some of the transformations are deploying the very same technologies and capabilities that are responsible for the growth of digital twin technologies. These include sensors used for IoT and IIoT projects, artificial intelligence (AI) and machine learning (ML), augmented, virtual, and mixed reality, cloud computing, and the telecom networks that support realtime data traffic demands and other high speed/high volume data requirements.

These are the most significant technology trends influencing the growing deployment of digital twins across industries.

- IoT deployments are a common factor for the adoption of digital twin technology. Digital twins can help manage existing networks of IoT components and model changes prior to rolling out in production environments.
- Sensors are increasing in functionality as their costs decline which fuels digital twin deployments as well as IoT projects. Sensors can gather realtime data that is consumed by digital twins for a variety of use cases.
- Increased capacity and reduced costs of wireless network solutions, which transport sensor data to IoT applications and digital twins.
- Cloud solutions, specifically cloud-based data management and storage migrations are accelerating deployment of digital twins in industry segments such as manufacturing and healthcare. Cloud providers like Amazon Web Services and Microsoft Azure have IOT and digital twin platforms along with partner ecosystems to offer a wide range of as-a-service solutions. Some of these as-a-service solutions are tailored to specific industry verticals.
- Industry 4.0 is another term for digitalization and is focused on sensor proliferation, network capacities, and cloud migrations. This trend is creating digital twin opportunities in a wide range of industrial applications.

Simulations, Emulations, and Digital Twins: Similar But Different

How is a digital twin different than a simulation or an emulation? A simulation offers a model that generally resembles but does not replicate a product or process. An emulation offers a realistic replica of a product or process. A digital twin combines an emulation with

artificial intelligence, specifically machine learning, to exactly model performance and responses to different data inputs. For example, Google Maps is a simulation of a journey. It is a static representation of data inputs for the start and end points of a trip. A simulation is not a model of the real world. A navigation app like Waze, which combines map data with real-time traffic conditions and driving speed, emulates the real world. An unexpected detour or delay produces recalculations of directions and arrival times. An app like Waze is a basic digital twin of your trip.

Digital Twin Industry Overview

This section briefly describes three vendors with interesting solutions or approaches that could have potential for electric utility OT cyber security applications. It also lists other digital twin solution providers that may have a presence in electric utilities.

General Electric

General Electric (GE) has a conceptual technology called the "digital ghost" for cyber security. GE has over one million digital twins deployed today to emulate products, processes, and systems. The digital ghost uses a product twin's data to teach the digital ghost what is normal versus abnormal activity or performance. Digital ghost is not a digital twin. Instead, a product digital twin trains the digital ghost algorithm. The algorithms perform real-time intrusion detection, localization, and neutralization. It's a biomimicry[3] concept of how the human immune system attacks a virus. Upon detection of a virus, the immune system attempts to isolate the offending organism and then neutralize it. GE describes the digital ghost as an additional layer of defense and offense to existing IT and OT cyber security technologies and practices. While GE has discussed this technology with utilities, it has not been deployed with any as of the publication of this white paper. GE has 5 grants from the Department of Energy (DOE) focused on digital twins for cyber security. DOE is particularly interested in real-time neutralization to help build grid resiliency.

GE defines a digital twin as a model plus machine learning. Modeling is done on a per asset basis. In 2018 GE launched Digital Twin for Power Transformers that would enable asset performance management for these products. The configuration settings for the transformers would be part of the model.

[3] Biomimicry definition – the emulation of models, systems, and elements of nature for the purpose of solving complex human problems. The practice of biomimicry recognizes that evolution has performed innumerable iterations on organisms and may have achieved the optimized design or process that can be applied to other unrelated challenges.

Microsoft Azure

Microsoft (MS) has a platform-as-a-service for developers of digital twins[4]. MS IOT Hub is a prerequisite to build MS Digital Twins, so there is a built-in assumption that an IOT project is driving the digital twin development. The platform includes an open modeling language and an environment that takes data inputs from IOT devices and systems and outputs reports and analytics. The platform can also store data. Models are built on the Digital Twins Definition Language (DTDL), which is in turn based on JSON LD (JavaScript Object Notation for Linked Data). A model describes a digital twin in terms of telemetry, properties, commands, relationships, and components. A model defines semantic relationships to map a digital twin into a knowledge graph. Relationships could be "node A sends data to node B and node C, but only node C can send data back to node A". Azure Digital Twin allows queries on property values, relationships, and relationship properties. Could that mean easy location of every sensor that sends data to a specific PLC or other convenient relationship mappings to aid cyber security? That question needs to be answered. Microsoft's documentation also indicates that external data sources can be connected using REST APIs or a Logic Apps connector. That connectivity is useful to their partner ecosystem.

There's one other observation about Microsoft Azure. Microsoft acquired CyberX, an OT cyber security company in June 2020 to integrate the IOT and IIOT cyber security capabilities into the Azure platform. That will increase their familiarity with ICS environments and the special requirements of OT cyber security.

Bentley Systems

Bentley is well-known in the utility sector and has experience with digital twin technology. This company created an open source tool, model, and hub called iTwin.js[5]. Their approach lets their digital twin customers compare actual states of an asset with the virtual twin, and then identify problems, track progress and perform what-if analysis on proposed changes. Bentley has connectors to traditional data sources like GIS and asset maintenance data and uses a rules-based methodology to ensure data accuracy. They have familiarity with realtime data like SCADA. Bentley partners include Microsoft Azure and Siemens, both with digital twin expertise and/or significant presence in electric utilities. The iTwin solution leverages Microsoft Azure's digital twin. Their work with Siemens

[4] https://www.youtube.com/watch?v=PM10Q9HDnTo&app=desktop
[5] https://www.itwinjs.org/

includes building asset digital twins and electrical model digital twins for utilities.

> Data preparation is critical to the success of any digital twin.

Bentley offers pragmatic information based on their experiences about the groundwork required to design, deploy, and maintain digital twins. Data preparation is critical to success. Utilities are well known for their functional silos, and that includes the data supporting those functions. Utilities also suffer from bad data quality, which compounds the data availability issue. The first tasks in any utility digitalization initiative should assess and address data quality and make data painlessly available. These tasks benefit all utility analytics projects, because all rely on the same basic data foundation. It is also very important to understand digital workflows to accurately describe data sources and relationships to create digital twins that remain true to their real products, processes, and systems. EPRI published a white paper[6] about data foundations for Automated Threat Mitigation (ATM) that reviewed data issues and mitigations in support of data-intensive OT cyber security initiatives. Bentley participates in an EPRI research activity focused on grid model data management, so their attention to data preparation is consistent across different digitalization activities. Their summary advice for digital twin projects is very applicable to any initiative involving emerging technologies: Think big, start small, and scale fast.

ANSYS - uses MS Azure Digital Twin to create digital twins for predictive maintenance of physical assets and equipment.

Hitachi ABB Power Grids – has some digital twin-like capabilities in the ABB Lifecycle Simulator for Process, Safety & Power Automation or ABB Process Power Simulator.

IBM Corporation – Digital Twin Exchange[7] lists the products from their partner ecosystem that focus on digital twins for specific industry segments, including electric utilities.

Oracle Corporation – IOT, predictive asset maintenance, digital worker, and facilities management are some of the applications described for digital twins.

SAS – there's an understandable emphasis on the importance of data analytics in digital twins in the SAS documentation on digital twins.

Siemens AG – has a digital twin approach and focus in their Industrial group. Their partnership with Bentley Systems targets asset digital twins and electrical model digital twins.

Digital Twin Research in Electric Utility Scenarios

EPRI is already researching applications[8] for digital twins in nuclear applications and spatial computing applications for utilities[9] in addition to this investigation focused on OT cyber security applications. Oak Ridge National Laboratory conducted one exploratory study of digital twin technology applied to utility cybersecurity.[10] Their research included building a digital twin for a lab scale model of a low voltage grid. The digital twin received real-time sensor data and ran simultaneously with the grid model. When the researchers created a man-in-the-middle cyberattack on the grid model, the digital twin detected the attack and recommended mitigation steps. One key learning was that a digital twin buildout takes time and expertise to develop, and experience building simulations was helpful. The research demonstrated that a digital twin can identify attacks. The digital twin continued normal operations in contrast to the hacked system, and proved it can deliver automated detection of anomalous data activity. That conceptual validation of an OT cyber security use case is important because the growing volumes and velocities of cyber security data in increasingly complex OT systems will challenge existing practices and technologies for fast detection and recovery.

On the academic front, researchers at the University of Illinois – Urbana/Champaign experimented with digital twin technology in the industrial control systems (ICS) environment. They used a digital twin to validate a command before executing it in ICS production. One key learning reported from this research is that every parameter modeled in a digital twin must be maintained, and thus it is extremely important to select the right parameters to develop the digital twin. Another important point concerned the

[6] Data Foundations for Operations Technology Cyber Security Analytics, AI and other Data-Intensive Applications (product ID 3002017455)

[7] https://shop.exchange.se.com/en-US/apps/60331/ibm-digital-twin-exchange

[8] Quick Insight Brief: Digital Twin Activities at EPRI (product ID: 3002020014)

[9] Standing Up The Real Digital Twin (product ID: 3002018703)

[10] https://www.ornl.gov/blog/seeing-double-digital-twin-secure-resilient-grid

use of existing product simulations as sources for data to create a digital twin. Product simulations are made by their vendors and are closed, and thus cannot readily serve as data sources to help accelerate the buildouts of digital twins.

Digital Twins for OT Cyber Security Use

Digital twin definitions are contextual. Here is EPRI's digital twin definition customized for OT cyber security. A cyber security digital twin is a data description of physical and logical assets, processes, and systems to enable realtime emulations of cyber security performance for autonomous and human-managed decisions.

OT cyber security is focused on protecting the availability and integrity of data for grid operations. Therefore, the data that describes an OT cyber security digital twin should include information about the hardware, software, and firmware in solutions deployed to protect OT assets and the availability and integrity of OT data. An OT cyber security digital twin needs data about the networks that connect OT assets and cyber security solutions and performance parameters. That data should include information about network addresses for all mission-critical gear that has a wired or wireless connection, including communications protocols. A normal operating parameter could include the typical data flows for a connection point. Access control might rely on a list of resources with authorization to read or edit configuration files. GIS data, a common data source for digital twins, could supply location information for physical assets. Data from asset management systems such as product serial number, hardware, software, and firmware versions, and normal operating parameters may be useful for an OT cyber security digital twin.

Potential Use Cases for OT Cyber Security Digital Twins

Digital twins are used to compare real-time performance against expected performance in asset and process applications in other business sectors. For instance, ABB described a digital twin that enabled a "virtual factory acceptance test" and significantly reduced costs and time incurred in a typical acceptance test. What are potential use cases for OT cyber security digital twins? Here are some problems that digital twins could address.

- Zero-day exploit detection. A digital twin offers a comparison to the real product, process, or system that is modeled and may accelerate detection of new attacks.
- Attack resiliency. A digital twin may provide the last good configuration of an asset that is compromised by malware, particularly if that malware erased logs. The digital twin could accelerate recovery to normal operations.
- Attack scenario exercises. A digital twin may be an improved tool to run "what if" scenarios of different types of attacks and train machine learning systems for autonomous responses such as the capabilities envisioned for ATM.
- Patch management. A digital twin may help utilities plan and deploy software updates for OT environments in a timely basis at reduced cost in contrast to existing methods.
- Zero trust authorization. A digital twin may reduce data latency and network capacity issues that would otherwise be introduced in production environments using zero trust security architectures for access and authorization.
- Data flow and network capacity planning. A digital twin of OT networks may proactively identify traffic bottlenecks and latency issues before new systems or applications are deployed.
- Deception traps. A digital twin may create more powerful deception technology solutions to protect mission-critical environments.

The initial OT cyber security use cases for digital twins will be applied to solve problems that cannot be more cost-effectively resolved through other solutions. Digital transformation initiatives will continue to increase the volumes, velocities, and varieties of data, along with attack surfaces, and digital twins may serve to help humans manage realtime cyber security operations. Digital twins may prove to be an important technology that supports or enables automated threat mitigation solutions. Research focused on proof

of concepts can help determine the use cases that derive the most benefits from digital twins.

Digital Twins Are Data-Intensive Projects

Regardless of the industry or the type of digital twin, the consensus across vendors[11] is that building and maintaining a digital twin is a data-intensive project, and the data challenges are similar in any digital transformation or projects that include advanced analytics or AI. The overarching success factor for any type of digital twin is data preparation. Utilities, like many business sectors, operate today with multiple application-oriented siloes. Bad or poorly organized data incurs costs to organizations now. One vendor estimates that 20-30% of today's operating expenses are related to repairing data quality and managing data changes. Another estimate is that 40% of engineering time is spent in data manipulation. Data preparation ensures the quality of data or data veracity. It also ensures that appropriate data governance and management principles and practices are defined and followed to establish real data-centric operations. There are some mitigations that utilities may consider that pave the way for digital twin deployments, and these would also aid in many digital transformations in utility production environments.

> Successful digitalization initiatives like a digital twin require utility commitment to data-centric operations.

There is no open platform of federated data that can integrate all required data sources and manage data quality and serve as the single source of truth for OT cyber security digital twins. However, this is not a challenge unique to design of digital twins for OT cyber security or its data. This lack of federated data is a challenge to many digital initiatives deploying AI or advanced predictive analytics. Resolving this issue would help mitigate project risks in digital initiatives.

Templates that digitally define assets like intrusion detection and intrusion prevention systems, Security Information and Event Management systems (SIEMs) or other OT cyber security solutions can accelerate adoption of digital twins by reducing deployment costs.

[11] https://event.webcasts.com/viewer/event.jsp?ei=1294623&tp_key=39e1e1d39b

The workforce skill requirements for digital twin buildouts will require a blend of skills that include simulation modeling and practical OT cyber security knowledge. This requirement is common to many digital transformations. Leveraging EPRI's collaboratively-developed knowledge in emerging technologies can help address this ongoing issue for utilities.

EPRI Research Question

The potential use cases and challenges for OT cyber security digital twins present interesting investigation opportunities for EPRI research. This list offers a good starting point for utilities interested in collaborating with EPRI in proofs of concept with this emerging technology.

1. What are the highest value use cases for digital twins in utility OT cyber security?
2. Can a digital twin support automated incident response?
3. What data must be collected to describe a cyber security digital twin?
4. What data must be collected to support machine learning in an OT cyber security use case?
5. What is the work effort involved to create a digital twin?
6. Can EPRI create a common vocabulary to describe cyber security assets and processes for digital twins?
7. What are the best practices for digital twin data management and machine learning for OT cyber security?
8. How often must data be updated to ensure that the digital twin accurately represents reality?
9. What are the network requirements to support a digital twin's consumption of realtime data?
10. What skill sets are required to support cyber security digital twins?

Conclusions

Lord Kelvin said "If you can't measure it, you cannot improve it" in the nineteenth century. In the 21st century, he might say, "If you can't emulate it, you can't proactively manage it." Digital twins, a combination of emulation and machine learning, may be an important technology to realize the Automated Threat Mitigation (ATM)

vision for OT cyber security. The Department of Energy's work with General Electric on threat neutralization suggests some alignment to the ATM vision. OT cyber security use cases may benefit from this emerging technology but there's an absence of hands-on knowledge about the real, measurable benefits of digital twins in cyber risk mitigation. A reasonable research starting point is to explore the work effort required to prototype a digital twin in EPRI's cyber security lab and demonstrate its capabilities. A subsequent research activity could then knowledgeably prioritize the highest value use cases for OT cyber security functions in collaboration with utilities. There are optimistic market forecasts about the future of digital twins for a variety of industries and applications. The greatest challenge for digital twins can be summed up in one sentence: *Successful digital transformations require utility commitment to data-centric operations.* That includes data preparation, data management, and data governance policies that enable friction-less access to data. This data-centric approach is also required for successful advanced analytics projects. Therefore, digital transformations offer an interesting entry point for inclusion of digital twin use cases. OT cyber security presents opportunities to determine the value of digital twins in risk mitigation, threat detection, and resiliency.

EPRI RESOURCES

Christine Hertzog, *Principal Technical Leader*
650.314.8111, chertzog@epri.com

Information, Communication and Cyber Security (ICCS)

The Electric Power Research Institute, Inc. (EPRI, www.epri.com) conducts research and development relating to the generation, delivery and use of electricity for the benefit of the public. An independent, nonprofit organization, EPRI brings together its scientists and engineers as well as experts from academia and industry to help address challenges in electricity, including reliability, efficiency, affordability, health, safety and the environment. EPRI also provides technology, policy and economic analyses to drive long-range research and development planning, and supports research in emerging technologies. EPRI members represent 90% of the electricity generated and delivered in the United States with international participation extending to nearly 40 countries. EPRI's principal offices and laboratories are located in Palo Alto, Calif.; Charlotte, N.C.; Knoxville, Tenn.; Dallas, Texas; Lenox, Mass.; and Washington, D.C.

Together…Shaping the Future of Electricity

Product ID 3002020292 January 2021

Electric Power Research Institute
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com