DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Reference herein to any social initiative (including but not limited to Diversity, Equity, and Inclusion (DEI); Community Benefits Plans (CBP); Justice 40; etc.) is made by the Author independent of any current requirement by the United States Government and does not constitute or imply endorsement, recommendation, or support by the United States Government or any agency thereof.

INL/RPT-25-87526-Revision-0

Adoption of AI in the Utility T&D Sector: Use Cases, Consequence, Assessment and Benefits

September 2025

Emma Stewart, Patience Christina Yockey, Remy Vanece Stolworthy, Megan Jordan Culler

Changing the World's Energy Future

Idaho National Laboratory

INL is a U.S. Department of Energy National Laboratory operated by Battelle Energy Alliance, LLC

INL/RPT-25-87526-Revision-0

Adoption of AI in the Utility T&D Sector: Use Cases, Consequence, Assessment and Benefits

Emma Stewart, Patience Christina Yockey, Remy Vanece Stolworthy, Megan Jordan Culler

September 2025

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Idaho National Laboratory

Rev 4.
Aug 2025

Adoption of AI in the Utility T&D Sector:

Use Cases, Consequence Assessment, and Benefits

White Paper

Emma M. Stewart, Patience Yockey,
Megan Jordan Culler, Remy Vanece Stolworthy

Battelle Energy Alliance manages INL for the U.S. Department of Energy 's Office of Nuclear Energy

Adoption of AI in the Utility T&D Sector: Use Cases, Consequence Assessment, and Benefits

White Paper

INL/RPT-25-87526

Emma M. Stewart, Patience Yockey, Megan Jordan Culler, Remy Vanece Stolworthy
Idaho National Laboratory

September 2025

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Grid Deployment Office
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

Table of Contents

1. Introduction

Digital transformation and utilization of artificial intelligence (AI) in the electric grid are fundamentally changing the industry's approach to common problems and enabling a broader paradigm shift in grid planning and operations. The change in approach is circularly both enabling and driving modernization, with load growth and reliable management of data center and AI infrastructure shifting away from planning approaches with relatively predictable behaviors and toward a mix of consumer and industrial choices that surpass human cognitive abilities to process. This movement has potential to condition humans to not understand the system on which the AI depends, while requiring it for development of the necessary infrastructure. Approaches which would address most likely grid conditions and events, such as faults, aging of equipment, and weather, now must also account for large loads which shift not based upon weather or time of day, but the computational load. Quantifying computational load is independent of the traditional grid forecasting variables, where a data center's aggregate load is determined by user and AI system behavior and decoupled from normal grid planning and operations. AI is both the cause and solution for these challenges, with new grid planning tools integrating massive amounts of decisions into frameworks.

The race to dominate the AI field between the United States and China has significant positive and negative implications for both technological advancement and national security. Domination of the field is not an isolated function of AI technology; instead, the U.S. must grow its capacity to handle large AI-driven electronic loads, develop new models, and drive adoption. As America invests heavily in AI research and development, the country's goal is to achieve breakthroughs that could redefine industries, economies, and military capabilities. However, the intense competition from adversarial nations also heightens security risks, including the potential for cyber espionage, intellectual property theft, and the misuse of AI technologies from the increased digital attack surface. Ensuring the responsible appropriate use and robust security of AI systems is paramount to preventing malicious exploitation and maintaining global stability. The U.S. must navigate this AI race carefully, balancing innovation with international cooperation to address critical security concerns, specifically when it comes to the integration of AI with critical functions such as the electrical power grid.[1]

The electric grid is evolving into a complex cyber-physical system where AI is increasingly applied to enhance monitoring, control, and optimization. Grid operators are already using AI to monitor transmission lines and isolate faults, and to predict fluctuations in electricity supply and demand.[2] As they look to grow and thrive in this new reality, utilities must also consider their approach to a secure and responsible use of the load on which the AI is driven along with the AI itself. Many AI applications, will be integrated into these tools, along with new requirements to implement cloud computing into operational technology (OT) operations, which also increases utility use of data centers that house cloud computing platforms. Therefore, risk frameworks must also address these interdependent concerns to enable their responsible and secure use as rapidly as AI adoption. Energy vendors and service providers are racing to deploy

[1] Prier, S., Strong, A. M., & Welburn, J. W. (2023). *Interdependence across the national critical functions* (RAND Working Paper WR-A210-1). RAND Corporation.
[2] Frigerio, G. (2024, December 3). *Exploring the role of AI in energy forecasting and grid management*. Plexigrid. https://plexigrid.com/exploring-the-role-of-ai-in-energy-forecasting-and-grid-management/

AI within their tools, products, and services to demonstrate their acceptance and enriched modernization capabilities. Though acting fast (and even failing fast) are qualities needed to help win the AI race, this comes at a cost to early adopters of these technologies, who will bear the burden of poor implementations, evolving models, and nascent application spaces. When these applications are in critical infrastructure sectors like the grid, the costs of these challenges could be significant (e.g., safety or reliability impacts).

This paper focuses on key aspects of AI adoption in the electric energy sector through three modalities: AI model integration inside the organization, purchase of commercial tools with AI integrated, and edge device integration. The authors summarize key risk-informed insights for AI use in electric grid operations and planning, which includes a review of the current use cases for the electric grid and a simple analysis framework to understand AI impact, coupled with application decomposition and consequence-driven analyses of the application's security and impact and in which case security controls can be applied. By doing this style of analysis, risk-informed controls can be placed around technologies, enabling their use and reducing the risk of the highest consequence repercussions. Like cloud data models, AI introduces new shared risk boundaries among utilities, vendors, regulators, AI developers, maintainers and support. Similar to other digital transition challenges in the distributed generation space, ownership transition for companies that don't exist long term is also a challenge. High-consequence actions (e.g., autonomous protection) may require hybrid architectures with fallback and manual conditions, whereas low-consequence actions (e.g., customer energy reports) can tolerate more automation processes. AI must be bound with consequence-driven engineering controls, redundancy, and transparency to prevent systemic failure.

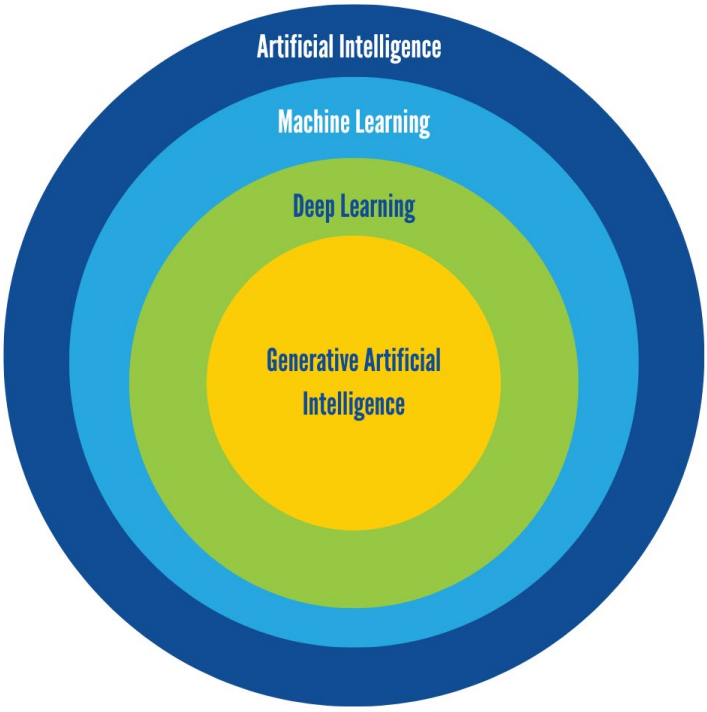

Figure 1: General AI configuration.

1.1 What is AI? A simple primer.

While there is no exact and well-accepted definition of AI, 15 USC 9401(3) defines AI as "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems use machine and human-based inputs to (a) perceive real and virtual environments; (b) abstract, such as perceptions into models through analysis in an automated manner; and (c) use model inference to formulate options for information or action." Using this definition, AI can encompass several

different modeling techniques and can be used in a large variety of contexts. Generally, however, AI is defined to include key disciplines such as machine learning (ML), deep learning (DL), and generative artificial intelligence (GenAI). The generalized set of AI configurations or building blocks is shown in the Figure 1.

The Department of Energy (DOE) defines ML as "the process of using computers to detect patterns in massive datasets and then make predictions based on what the computer learns from those patterns [...] In ML, algorithms serve as rules for how to analyze the data using statistics. ML systems use these rules to identify relationships between data inputs and desired outputs."[3] With respect to grid optimization, ML algorithms have the capability to "analyze vast amounts of data from smart meters, sensors, and other grid components to optimize energy distribution, forecast demand, and detect irregularities that could indicate potential failures."[4] This additional computation power is expected to improve many metrics in time and detection of events. Example ML techniques commonly used in grid applications include support vector machines (SVMs), decision trees (DTs), and random forests (RFs).

DL is a subset of ML in which DL algorithms "can automatically learn representations from data such as images, videos, or text without introducing human domain knowledge."[5] Traditional ML algorithms generally require humans to determine the hierarchy of features within an input dataset and are a more structured approach. DL, on the other hand, automates the feature extraction process and uses many different layers of algorithms known as "hidden layers" to flexibly learn directly from raw data and recognize patterns autonomously, where each layer is composed of "neurons" that perform individual calculations based on inputs and "weights" to calculate a set output. Given the increased flexibility of these architectures, DL generally requires significantly more data than traditional ML and parallelizable processors such as graphics processing units (GPUs) to process a large number of identical neurons and hidden layers. Example DL models used in grid applications include artificial neural networks (ANNs), long short-term memory (LSTM) networks, and convolutional neural networks (CNNs).

As seen in the innermost circle of Figure 2, GenAI generates new statistically probable outputs according to learned patterns from raw data.[6] GenAI requires massive amounts of data, GPUs, and time to produce and train a model. As a workaround for the complexity of producing these models, companies have started producing pre-trained GenAI models where a method called "transfer learning" can be used to reduce the amount of time, data, and computing resources to train a model. Transfer learning is the process of applying knowledge to a DL/GenAI model derived from solving one problem and repurposing the final layers of a model to a related but different problem. This allows for a large pre-trained GenAI model to be used across domains

---

[3] U.S. Department of Energy. (n.d.). *DOE Explains...Machine Learning*. Retrieved from Office of Science: https://www.energy.gov/science/doe-explainsmachine-learning.

[4] Rashid, A., Biswas, P., Al Masum, A., Al Nasim, M., & Gupta, D. K. (2024, October 20). Power Plays: Unleashing Machine Learning Magic in Smart Grids. *4*, p. 16. Retrieved from https://arxiv.org/abs/2410.15423.

[5] Nvidia. (2025). *Deep Learning*. Retrieved from Glossary: https://www.nvidia.com/en-us/glossary/deep-learning/.

[6] Microsoft. (2024, August 29). *Deep learning vs. machine learning in Azure Machine Learning*. Retrieved from Machine Learning: https://learn.microsoft.com/en-us/azure/machine-learning/concept-deep-learning-vs-machine-learning?view=azureml-api-2&viewFallbackFrom=azureml-api-2).

and contexts with reduced resources. GenAI, such as a generative pretrained transformer (GPT) model, has been proposed to improve state estimation in the power grid, where measurements are not available, or allow for flexible, probabilistic forecasting[7] through the use of transfer learning to fit a GPT to grid operations.

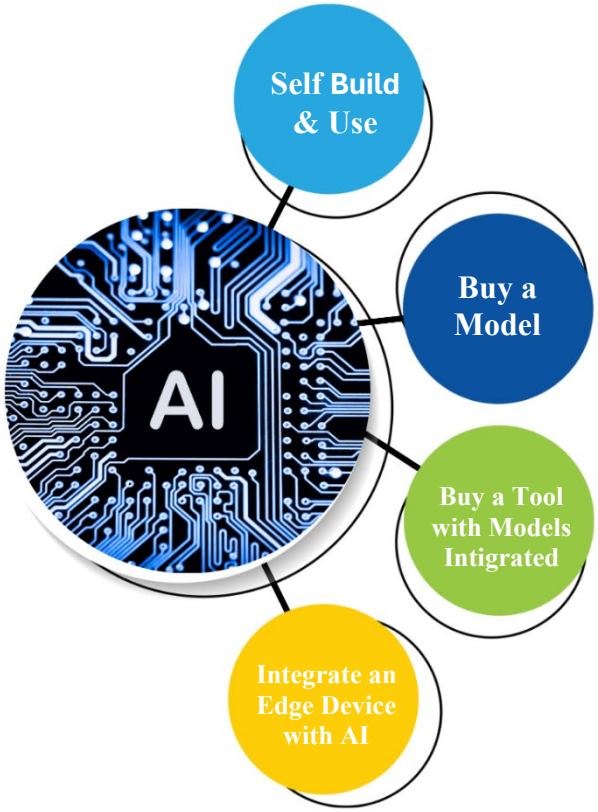

Figure 2: Diagram of adoption pathways for AI in the electric sector.

2. AI Adoption Styles in the Electric Sector

In electric grid infrastructure entities, digital technology is intentionally adopted by (a) building it internally, (b) buying a product with it already integrated, or (c) edge-deployed, such as in smart meters. Utilities with high levels of resources might consider building home-grown products whereas smaller entities will often buy the product and use third-party services. This is evident in cloud transformation where smaller utilities are using third-parties.[8] [9] In many cases, a combination of all three approaches is implemented depending on the use case.

Larger utilities have the ability (or desire, as the ability is established) to test multiple AI tools for multiple applications, both through use of their own testbeds and partnerships with research institutions. They have more structured risk management processes, building in multiple checks for the safe adoption of new technologies. Smaller utilities lack the freedom and flexibility to experiment with different products and see what works best for them. They often do not have the negotiating power to ask providers[10] to adjust their products or services, which would make adoption more palatable in their

---

[7] Lok Choi, S., Jain, R., Feng, C., Emami, P., Zhang, H., Hong, J., . . . Kroposki, B. (2024). *Generative AI for power grid operations*. U.S. Department of Energy. Golden: National Renerable Energy Laboratory. Retrieved from https://www.nrel.gov/docs/fy25osti/91176.pdf.

[8] Stewart, Emma Mary, Morgan, Julia Catherine, Woodruff, Nathan Lee, Combe, Glenn, & Stolworthy, Remy Vanece (2024). Enhancing Cloud Cybersecurity: Prescriptive Controls for Operational Technology. https://doi.org/10.2172/2474851

[9] J. Morgan, E. M. Stewart, R. Stolworthy, N. Woodruff, J. Briones and J. Whitaker, "Consequence Based Framework for Deployment of Cloud Solutions in the Digital Energy Transition," *2025 IEEE PES Grid Edge Technologies Conference & Exposition (Grid Edge)*, San Diego, CA, USA, 2025, pp. 1-5

[10] *Co-ops not ready to use the cloud for electric grid operations.* (n.d.). Cooperative.com. Retrieved August 26, 2025, from https://www.cooperative.com/news/Pages/co-ops-not-ready-to-use-cloud-for-electric-grid-operations.aspx

environments. However, many recognize that they may miss opportunities for efficiency gains, or cost reduction in the grid modernization space, so they will be adopting AI tools despite the perceived and real risks. In contrast, due to the regulatory environment for the electric sector, many smaller utilities have rapidly adopted cloud-based technology (and likely AI), as they are not as closely regulated as the larger critical entities and therefore can make wider digital choices. There is no widespread regulation for use of AI in the electric sector, and the pace of standards and regulation will not hold with the pace of AI adoption. Alternate approaches for its secure and responsible adoption are needed.

Instead, the AI applications that will be deployed in the grid are developed by service and tool providers, some of whom may have experience in the utility sector and are attempting to include AI in their offerings, and some of whom may have AI expertise and see an opportunity in the utility sector. In either case, the dual understanding of how the AI applications work and what they are actually accomplishing in the utility space is needed for success. Models need to be tuned to the environment in which they are deployed to be most impactful. Careful coordination and communication between experts for the individual systems in which technology is deployed and the service providers will enable impactful use of AI. As an example of early adoption, which is both internally developed and offered as something other entities can buy, Dairyland Power Cooperative launched VoltWrite, a private generative AI platform, now offered as a service to other rural electric cooperatives. Introduced by CIO Nate Melby at the NRECA TechAdvantage Conference in March 2025, VoltWrite is represented as a "secure, in-house AI model that avoids the cybersecurity risks of open public models." An example of its use is field inspectors wearing AR glasses linked to VoltWrite and Microsoft Teams for real-time collaboration in remote site inspections.[11] In another example, Itron is integrating NVIDIA's Jetson Orin Nano system-on-module and AI Enterprise software into its Grid Edge Intelligence platform, which supports more than 13 million distributed intelligence endpoints. The addition of GPU-based processing at the meter level enables execution of AI models directly on edge devices, reducing the need for centralized computation. Potential applications include localized load forecasting, voltage optimization, and real-time fault detection using high-resolution consumption and power quality data. This architecture is intended to support faster decision-making in grid operations and improve coordination with distributed energy resources.[12]

Early adopters of AI technologies will bear the costs of the challenges in novel applications. Vendors and service providers are racing to deploy AI in their tools, products, and services to demonstrate their adoption and modernization capabilities. Although acting fast and even failing fast are qualities need to help win the AI race, early adopters experience costs associated with poor implementation, evolving models. When these applications are in critical infrastructure sectors like the grid, the costs of these challenges could be significant (e.g. safety or reliability impacts).

---

[11] Sukow, R. (2025, March 11). *Dairyland Power introduces AI as a service for rural electrics.* Retrieved from NRTC: https://www.nrtc.coop/dairyland-power-introduces-ai-as-a-service-for-rural-electrics/.

[12] Itron Inc. (2025, March 19). Itron to Bring NVIDIA-Powered Artificial Intelligence to the Grid Edge. Retrieved from https://investors.itron.com/news-releases/news-release-details/itron-bring-nvidia-powered-artificial-intelligence-grid-edge.

3. Selection of Models and Considerations for AI in the Electric Grid

There are four primary Operational Technology (OT) domains in which power systems applications for AI reside detection, prediction, control and optimization, and business and customer applications (Figure 4). Some applications fit within multiple domains, and each domain has a different risk profile, which will be discussed in the following sections. The authors address critical considerations for deploying AI in power systems planning, operations, and business contexts. They consider the importance of physical modeling, robust data foundations, and secure automation strategies, in relation to these domains.

3.1 AI Task Domains in Operational Power Systems

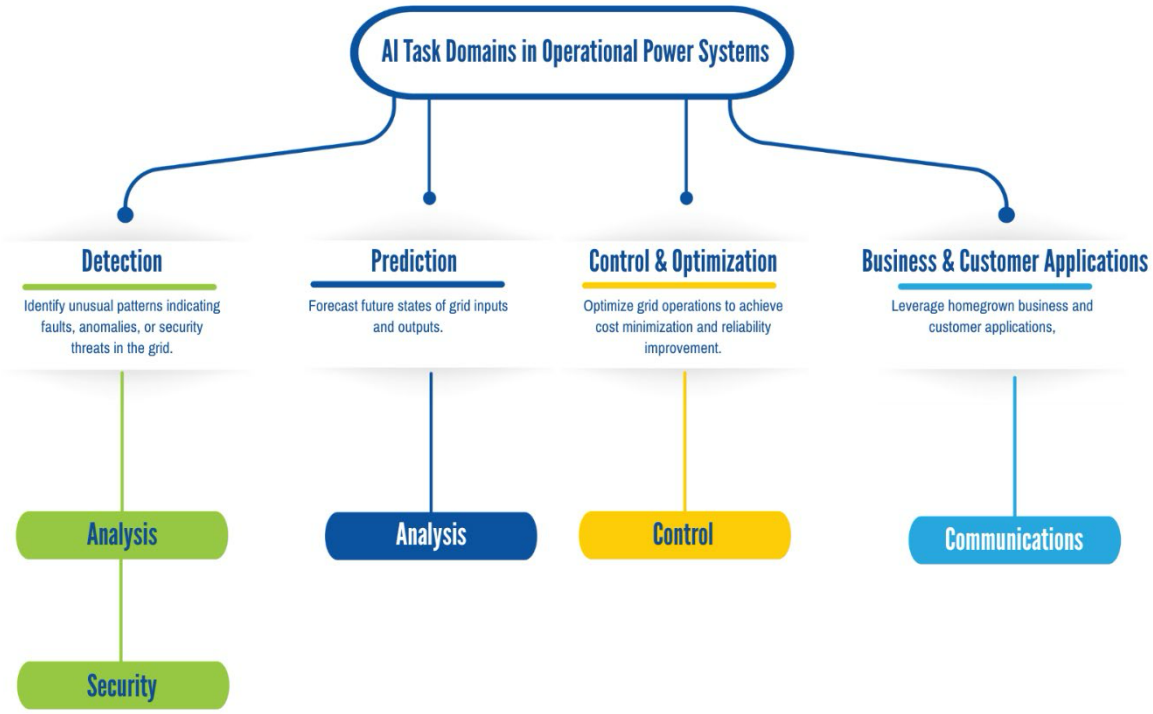

Figure 3: The four AI task domains for power systems are detection, prediction, and control and optimization.

3.1.1 Detection: Anomaly and Fault Detection

In the grid context, detection tasks involve identifying unusual patterns or events that could indicate faults, anomalies, or security threats. Examples include detecting equipment failures, line faults, cyber-intrusions, or abnormal grid oscillations. AI-based detection augments traditional alarms by sifting through high-volume sensor data (e.g. SCADA, PMU, or smart

meter streams) to catch subtle issues that rule-based systems might miss.[13] Many control centers now employ synchrophasor (PMU) analytics for oscillation and event detection, but as data volumes grow, ML techniques are being explored to automatically flag anomalies and provide operator support.[14] [15]

Model Deployment Examples and Case Studies: Unsupervised learning can be used when trying to determine the presence of novel events without explicit labels — i.e., a squirrel or a physical attack causing a fault. Techniques like clustering and outlier detection can learn the "normal" patterns of grid data and highlight outliers. Alternatively, supervised learning is applicable when labeled examples of faults or events are available. Researchers have applied supervised learning for tasks like fault detection and location,[16] asset health monitoring, and security assessment in power grids.[17] For example, image-based deep learning[18] has been used to detect equipment issues from drone imagery, and graph convolutional networks have been tested to detect and isolate faults using the network topology.[19] [20] Physics-informed models for detection are suitable where electrical laws are embedded into an AI model to ensure that an "anomaly" truly violates expected physical behavior.[21] For example, wildfire risk monitoring is an illustration of detection in action utilizing physics based models. In regions like California,

---

[13] Chen, Y., Fan, X., Huang, R., Huang, Q., Li, A., & Guddanti, K. (2024). *Artificial Intelligence/Machine Learning Technology in Power System Applications* (PNNL-35735) (Technical Report). Pacific Northwest National Laboratory. https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-35735.pdf

[14] Stewart, E. M. (2017, September 14). *GMLC 1.4.9 – Integrated Multi Scale Machine Learning*. Retrieved from U.S. Department of Energy: https://gmlc.doe.gov/projects/integrated-multi-scale-data-analytics-and-machine-learning-grid.

[15] Wilson, A., Ekti, A., Follum, J., Biswas, S., Annalicia, C., Joo, J.-Y., . . . Lian, J. (2024). The Grid Event Signature Library: An Open-Access Repository of Power System Measurement Signatures. IEEE Access, 76207-76218. doi:10.1109/ACCESS.2024.3404886.

[16] Chen, Y., Fan, X., Huang, R., Huang, Q., Li, A., & Guddanti, K. (2024). *Artificial intelligence/machine learning technology in power system applications*. U.S. Department of Energy. Oak Ridge: Pacific Northwest National Laboratory. Retrieved from https://www.pnnl.gov/main/publications/external/technical_reports/PNNL-35735.pdf.

[17] Morales-Rodriguez, M., Srinivas, N., Olatt, J., Hahn, G., Wilson, A., Fuhr, P., . . . Chakraborty, I. (2020). Big Data Framework for the Development of a Signature Library for the Power Grid. *IEEE Power and Energy Society General Meeting*. Washington: Lawrence Livermore National Laboratory. Retrieved from https://www.osti.gov/servlets/purl/1798440.

[18] Shashsavari, A., Farajollahi, M., Stewart, E., Cortez, E., & Mohsenian-Rad, H. (2019, February 11). Situational Awareness in Distribution Grid Using Micro-PMU Data: A Machine Learning *Approach*. 10(6). Retrieved from https://www.osti.gov/pages/biblio/1811286.

[19] *Austin Energy announces full deployment of AI-driven Early Wildfire Detection System*. (2024, August 29). Retrieved from Austin Energy: https://austinenergy.com/about/news/news-releases/2024/austin-energy-announces-full-deployment-of-ai-driven-early-wildfire-detection-system2024.

[20] Stewart, E., Chellappan, K., Backhaus, D., Deka, D., Reno, M., Peisert, S., . . . Buckner, M. (2018*). Integrated Multi Scale Data Analytics and Machine Learning for the Grid; Benchmarking Algorithms and Data Quality Analysis*. U.S. Department of Energy, Lawrence Livermore National Laboratory. Grid Modernization Laboratory Consortium. Retrieved from https://www.osti.gov/servlets/purl/1490956.

[21] Boroujeni, S., Razi, A., Khoshdel, S., Afghah, F., Coen, J., O'Neill, L., . . . Vamvoudakis, K. (2024, August). A comprehensive survey of research towards AI-enabled unmanned aerial systems in pre-, active-, and post-wildfire management. *Information Fusion*, 108. Retrieved from https://www.sciencedirect.com/science/article/pii/S1566253524001477.

utilities deploy sensors (for temperature, smoke, etc.)[22] and AI models to detect abnormal conditions near power lines; if an AI model flags a high fire risk (e.g., a line sparking in dry conditions), the system can trigger corrective actions such as de-energizing lines or adjusting protection settings.[23]

AI technologies applied to OT cybersecurity in power grid systems employ advanced techniques, such as deep neural networks (e.g., LSTM, CNN), digital twins, hybrid intrusion detection frameworks, and LLM-enhanced interpretability layers to perform detection for cyber anomalies. These systems process real-time operational data, like phasor measurements and substation communication streams, to detect and distinguish between genuine physical anomalies and cyber-induced disruptions. Innovations like generative AI-based anomaly detection and LLM-supported explanations enhance both adaptability and comprehension, delivering timely, interpretable responses to cyber threats in critical grid infrastructure.[24] Digital-twin approaches create virtual replicas of power systems, enabling machine learning models to learn normal behavior patterns and flag deviations that could detect cyberattacks or system faults, particularly in grids with high penetration of power electronics.[25] Additionally, large language models (LLMs) are being used to bring interpretability into OT cyber detection. In particular, frameworks combining lightweight ML classifiers with LLM-generated explanations are reported to detect attacks on Automatic Generation Control (AGC) systems in near real time and provide human-readable diagnoses of detected cyberattacks.[26]

3.1.2 Prediction: Forecasting and Proactive Predictive Analytics

Prediction tasks involve forecasting future states of inputs and outputs of the grid (in response to inputs of forecast, etc.). Electric utilities have long performed forecasting (e.g., demand or load, renewable generation), and AI can enhance accuracy by learning complex patterns from historical and real-time data, along with enabling multivariate processing. Beyond load prediction, AI-based prediction is used for failure prognostics (predicting equipment failure or maintenance needs), outage forecasting (anticipating where and when outages or stresses will occur), resource forecasting (predicting availability of generation capacity, particularly for wind

---

[22] NASA. (2025). *AI-enabled drone swarms for fire detection, mapping, and modeling*. Retrieved from Earth Science Technology Office: https://esto.nasa.gov/firetech/ai-enabled-drone-swarms-for-fire-detection-mapping-and-modeling/.

[23] Wolfe, S. (2024, August 29*). Another utility adopts AI cameras for wildfire detection*. *This time, it's not as far west*. Retrieved from Factor This Power Engineering: https://www.renewableenergyworld.com/power-grid/outage-management/another-utility-adopts-ai-cameras-for-wildfire-detection-this-time-its-not-as-far-west/.

[24] Qi, R., Rasband, C., Zheng, J., & Longoria, R. (2021). Detecting Cyber Attacks in Smart Grids Using Semi-Supervised Anomaly Detection and Deep Representation Learning. *Information*, *12*(8), 328. Retrieved from https://www.mdpi.com/2078-2489/12/8/328.

[25] Elimam, M., Isbeih, Y. J., Azman, S. K., El Moursi, M. S., & Al Hosani, K. A. (in press, 2022). Deep learning-based PMU cyber security scheme against data manipulation attacks with wide-area damping control (WADC) application**.** *IEEE Transactions on Power Systems*, 1–15. https://doi.org/10.1109/TPWRS.2022.3181353

[26] Coppolino, L., Nardone, R., Petruolo, A., Romano, L., & Souvent, A. (2023). Exploiting Digital Twin technology for Cybersecurity Monitoring in Smart Grids. *Proceedings of the 18th International Conference on Availability, Reliability and Security* (pp. 1-10). ARES'23. Retrieved from https://doi.org/10.1145/3600160.3605043.

and solar resources), and stability or reliability risk forecasting. By analyzing weather data, asset condition data, and usage patterns, AI could potentially help utilities move from reactive operations toward proactive management. For instance, AI-driven models can forecast which transformers are likely to overload or fail during an upcoming heat wave or predict which tree branches will threaten power lines in a storm.[27] This enables pre-failure intervention, dispatching crews or actioning an outage to prevent a fire. Prediction can also be considered using planning tools, such as Google Tapestry or PNNL's ChatGrid.[28] [29] [30]

Model Deployment Examples and Case Studies: Prediction problems are typically approached with supervised learning regression models trained on historical data. Classical statistical models (ARIMA, state-space models) are now often outperformed by ML models, such as gradient-boosted trees, support vector regression, or neural networks, that capture nonlinear relationships. In particular, deep learning models excel at time-series forecasting: recurrent neural networks like LSTMs and GRUs are designed to capture temporal dependencies and have been successfully applied to forecast loads accounting for daily and seasonal patterns.[31] LSTMs, for example, can learn the influence of past demand and weather on future load, improving accuracy for day-ahead and hour-ahead forecasts. These models can incorporate multiple inputs for multivariate forecasting in the grid. Integrating physical knowledge can further improve predictions; for instance, physics-informed neural networks (PINNs) and hybrid models include constraints from power system equations (like power flow or machine dynamics) to ensure that predictions of system states (voltage, frequency, etc.) remain physically plausible. Such models have been proposed for tasks like system frequency prediction and stability assessment, where purely data-driven approaches might violate conservation laws or operational limits.[32] AI-driven forecasting, adequately secured and implemented, may benefit grid operations. One study found that combining machine-learning-based forecasts with real-time sensor data reduced grid imbalance events by up to 30%.[33]

---

[27] *PG&E Applies Enhanced Outage Prediction Models To Ready Crews and Resources Ahead of This Week's Storm*. (2025, March 11). Retrieved from PG&E: https://www.pge.com/en/newsroom/currents/safety/pg-e-applies-enhanced-outage-prediction-models-to-ready-crews-an.html#:~:text=The%20data%20provided%20through%20AI,Storm%20safety%20tips.

[28] Wendel, J. (2024, February 22). *ChatGrid™: A New Generative AI Tool for Power Grid Visualization*. Retrieved from Pacific Northwest National Laboratory: https://www.pnnl.gov/news-media/chatgridtm-new-generative-ai-tool-power-grid-visualization.

[29] *Tapestry*. (2025). Retrieved from Google LLC: https://x.company/projects/tapestry/

[30] von Meier, A., Stewart, E., McEachern, A., Andersen, M., & Mehrmanesh, L. (2917). Precision Micro-Synchrophasors for Distribution Systems: A Summary of Applications, *8*(6). *IEEE*, 2926-2936. doi:10.1109/TSG.2017.2720543.

[31] Kara, E., Roberts, C., Tabone, M., Alvarez, L., Callaway, D., & Stewart, E. (2018, March). Disaggregating solar generation from feeder-level measurements. *Sustainable Energy, Grids and Networks*, 112-121. doi:https://doi.org/10.1016/j.segan.2017.11.001.

[32] U.S. Department of Energy. (2025). *Artificial Intelligence*. Retrieved from Pacific Northwest National Laboratory: https://www.pnnl.gov/artificial-intelligence.

[33] Aslam, S., Aung, P., Rafsanjani, A., & Majeed, A. (2025). Machine learning applications in energy systems: current trends, challenges, and research directions. *Energy Informatics, 8*. Retrieved from https://energyinformatics.springeropen.com/articles/10.1186/s42162-025-00524-6.

3.1.3 Optimization, Control, and Decision-Making

Optimization tasks involve determining the best set of actions or configurations for the grid, often in real time, to meet certain objectives (cost minimization, loss reduction, reliability improvement) under constraints. Traditional grid optimization (like economic dispatch, unit commitment, or network reconfiguration) relies on mathematical programming and expert rules. AI is augmenting these tasks by handling high-dimensional decision spaces and adapting to changing conditions. For example, AI can optimize voltage control on distribution networks by coordinating many voltage regulators and inverters, or optimize power dispatch in microgrids, balancing generation, storage, and loads. During emergencies, optimization algorithms help in outage restoration, finding the fastest way to reroute power or prioritize repairs.[34] AI can also manage distributed energy resources by solving distributed optimization problems, such as deciding how to dispatch hundreds of batteries or loads for peak shaving. In essence, optimization AI seeks the optimal or near-optimal operating point for the grid at any moment, often needing to satisfy physical and reliability constraints.

Model Deployment Examples and Case Studies: Deep reinforcement learning has been studied for grid control tasks such as automatic generation control, battery dispatch, or network switching operations.[35] However, a challenge is ensuring safety of poorly trained AI, adversarial inputs, and bad policies could violate grid limits, and use of these models for real-time operation means the consequences of a bad AI decision are heightened. A potential strategy is to train these models on realistic simulators (digital twins of the grid) and incorporate safety constraints into the learning process, but that process also depends on good grid models, another current challenge.[36] AI is proposed to assist optimization by serving as a fast approximation to traditional power flow solvers. For example, neural network surrogate models have been developed to mimic the results of optimal power flow (OPF) calculations much faster, enabling near real-time dispatch decisions.[37] [38] There are also hybrid approaches, where a neural network proposes an optimal control action and a physics-based check (or simpler optimization) verifies it, leveraging both AI speed and physical law compliance. In real-world deployment, fully autonomous grid optimization by AI is still approached cautiously due to safety, but there are successful examples of closed-loop optimization. One notable case is Entergy Louisiana's deployment of AI in its advanced metering infrastructure (AMI) to predict when distribution

---

[34] Allsup, M. (2025, July 28). *Under the hood of CAISO's AI outage management pilot*. Retrieved from https://www.latitudemedia.com/news/under-the-hood-of-caisos-ai-outage-management-pilot/

[35] Stewart, E., Kiliccote, S., Shand, C., McMorran, A., Arghandeh, R., & von Meier, A. (2014). Addressing the challenges for integrating micro-synchrophasor data with operational system applications. *2014 IEEE PES General Meeting*. National Harbor: IEEE. Retrieved from https://ieeexplore.ieee.org/document/6938994/.

[36] U.S. Department of Energy. (2023). *Inverter-Based Resource Performance Issues Report*. Atlanta: North American Electric Reliability Corporation. Retrieved from https://www.nerc.com/comm/RSTC_Reliability_Guidelines/NERC_Inverter-Based_Resource_Performance_Issues_Public_Report_2023.pdf.

[37] U.S. Department of Energy. (2025). *Artificial Intelligence*. Retrieved from Pacific Northwest National Laboratory: https://www.pnnl.gov/artificial-intelligence.

[38] Soltani, Z., Ma, S., Khorsand, M., & Vittal, V. (2023, May). Simultaneous Robust State Estimation, Topology Error Processing, and Outage Detection for Unbalanced Distribution Systems. *IEEE Transactions on Power Systems, 38*(2), 2018-2034. doi:10.1109/TPWRS.2022.3181118.

transformers are near failure and by proactively replacing them mitigate outages.[39] The AI system processes smart meter data and other inputs to reconfigure the network or dispatch field crews proactively. Since deployment, it has reportedly prevented 536 outages and avoided over 48,000 outage minutes by automatically rerouting power and isolating problems.[40] This illustrates how an AI-driven optimization can directly translate to reliability gains (or losses if availability of the system is a requirement). Generally, AI-driven optimization in the grid is implemented in a hierarchical way: slower, cloud-based AI might optimize day-ahead planning or system configurations, while faster embedded AI at the grid edge (in controllers or devices) handles real-time fine-tuning. As these models are adopted, ensuring they respect operational constraints (voltage, frequency, line limits) is a vital constraint. AI applications should not violate the laws of physics in power systems applications.[41]

California Independent System Operator (CAISO), which oversees the majority of California's power grid and portions of the Western Energy Imbalance Market, is piloting OATI Genie™, an advanced AI assistant designed to streamline and modernize outage management workflows. Traditionally, CAISO operators manually sift through a high volume of structured data and free-form outage reports to assess grid impact, often under time pressure. OATI Genie augments this process by using a hybrid AI architecture that blends generative AI, machine learning, and rule-based logic to extract relevant keywords, correlate outage events with system data (e.g., GIS, SCADA, weather), and deliver real-time insights in natural language. This co-pilot approach is designed not just to automate, but to enhance human situational awareness, reduce manual review time, and support faster, more informed decisions in outage scheduling and grid stability planning.

Southwest Power Pool (SPP) is collaborating with Hitachi and NVIDIA to develop an AI-enabled solution aimed at reducing the time required to complete generator interconnection (GI) impact studies. These studies currently span 18 to 27 months and are critical to evaluating how new generation sources will affect transmission system reliability. The AI platform, based on Hitachi iQ and utilizing NVIDIA-accelerated compute infrastructure, will automate power flow simulations, optimize resource modeling, and support predictive analytics for interconnection scenarios. According to SPP, the AI tool could reduce study durations by up to 80%, and preliminary pilots are expected to begin in winter 2025-2026. The tool is designed to work alongside, rather than replace, existing engineering workflows; applicants will still have access to conventional study tracks as a baseline for comparison.[42]

---

[39] entergy. (2024, July 16). *We're taking steps to advance reliability across our region*. Retrieved from https://www.entergy.com/blog/we-re-taking-steps-advance-reliability-across-our-region.

[40] Clarion Energy. (2025, February 6). *Resilience through intelligence: AI's impact on utility disaster readiness*. Retrieved from Factor This Engineering: https://www.datacenterdynamics.com/en/news/southwest-power-pool-partners-with-hitachi-to-slash-interconnection-study-times-through-ai-tool/

[41] Mahdi, J., Scaglione, A., Roberts, C., Stewart, E., Peisert, S., & McParland, C. (2018, July 4). Anomaly Detection Using Optimally Placed μPMU Sensors in Distribution Grids. *IEEE Transactions on Power Systems*, 33(4), 3611 - 3623. doi:10.1109/TPWRS.2017.2764882.

[42] kidmore, Z. (2025, June 7). *Southwest Power Pool partners with Hitachi to slash interconnection study times through AI tool*. Retrieved from https://www.datacenterdynamics.com/en/news/southwest-power-pool-partners-with-hitachi-to-slash-interconnection-study-times-through-ai-tool/

3.1.4 Business and Customer Applications

GenAI is primarily being used in business applications. Notably, most of these applications are home-grown, rather than commercially purchased as part of another management platform. New applications such as SEW[43] are being sold as a solution for customer AI management.

Model Deployment Examples and Case Studies: Generative AI is being used to assist in drafting social media posts for utilities. While at first glance this might seem a benign application, social media in recent years has become the primary source of communication of outage and customer issues, with rapid communication during events being a gold standard and trained.[44] Modern utility contact center platforms now use AI to proactively notify customers of known outages and estimated restoration times via text or voice, reducing inbound inquiries. AI-driven systems are being linked to outage management to draft messages to customers automatically. Malicious use of social media tools to drive societal panic and challenges in power restoration is also a key feature in large-scale grid exercises such as GridEx[45]; therefore, the automation of these functions through AI should be considered a risk. Data security is less of a concern, but reliability and clarity of AI outputs are critical. Lastly, GenAI is being used to draft repeated processes and manuals. For example, a guide on "How Net Metering Works."[46] While this is a low-risk short-term item, without subject matter expert review, this could be a challenge. GenAI is also being used internally to support utility employees and enhance operational workflows. This ranges from knowledge assistants that help staff quickly find information to AI copilots that assist with coding and analysis.

Many utilities have also implemented AI chatbots to handle routine customer inquiries, utilizing LLMs such as ChatGPT for natural language responses. These systems can be available 24/7 to answer billing questions, account changes, and provide information, and therefore availability. For example, Sulphur Springs Electric Cooperative (AZ) launched a LivePerson AI chat on its website to handle member questions after hours, reducing call center load.[47] UK-based Octopus Energy's GPT chatbot is handling 44% of its customer service inquiries, equivalent to 250 people's work, and has higher satisfaction rates than its human counterparts.[48] In particular, these tools are being used to draft customer service agent email responses. Risks of these types of applications are in data privacy, with customers' information being used in potentially external

---

[43] SEW. (2025). *#1 Industry AI Chatbot WeSmart*. Retrieved from https://www.sew.ai/content/smart-messaging-chatbots

[44] Cohn, L. (2025, August 18). *Emergency communications evolution: How intelligent workflows are reshaping utility crisis response.* American Public Power Association. https://www.publicpower.org/periodical/article/emergency-communications-evolution-how-intelligent-workflows-are-reshaping-utility-crisis-response

[45] North American Electric Reliability Corporation. (2018, March). *Grid Security Exercise GridEx IV: Lessons learned* (TLP: WHITE) (Public Report). NERC. https://www.nerc.com/pa/CI/ESISAC/GridEx/GridEx%20IV%20Public%20Report.pdf

[46] Mullen, M. (2023, October 18). *Here Comes The Meter Man – A Lesson in GenAI Planning.* Deep Analysis. https://www.deep-analysis.net/here-comes-the-meter-man-a-lesson-in-genai-planning/

[47] LivePerson. (2025). AI chatbot agents: Better conversations start with a better AI chatbot. Retrieved from https://www.liveperson.com/products/ai-chatbots/.

[48] Celonis, Inc. (2025). *Connect Octopus Energy and OpenAI (ChatGPT, Whisper, DALL-E) integrations*. Retrieved from Make: http://make.com/en/integrations/octopus-energy/openai-gpt-3.

AI models, and human review is introduced in some cases already due to hallucinatory responses. Chatbots are known to be eager to please and may provide inaccurate results to poorly phrased or intentionally misleading prompts by customers, potentially resulting in billing issues or unintentional panic. Fraud detection may also be a challenge.

Another notable example is Southern California Edison's deployment of a GenAI chatbot with retrieval-augmented generation for its network operations center (Project "Orca"). This internal chatbot lets Network Operations Center (NOC) engineers query network device information, troubleshoot tickets, and get real-time trend analysis by pulling from SCE's telecom manuals and data, improving response times and aiding training of new operators. The tools are also proactively assisting in incident management and data analytics. This use case promotes the use of private network AI with NVIDIA.[49] While these kinds of functions boost workforce efficiency and capture expert knowledge, errors in technical guidance could have serious consequences; strict validation of AI-provided answers is required. Data must stay on secure servers to protect grid information.

3.2 Model Deployment Examples and Case Studies

Table 1 below outlines various AI use cases across different categories, highlighting the functionality and applications they bring to the electric grid. These use cases demonstrate the transformative potential of AI in driving efficiency, reliability, and security in utility operations, paving the way for a more resilient energy future. For instance, in the category of Predictive Maintenance and Asset Health Monitoring, AI analyzes sensor data from transformers and inverters to predict failures, exemplified by tools such as GE Vernova's GridOS[50] and Schneider Electric's EcoStruxure ADMS.[51]

Note: The examples listed are not exhaustive; they represent only those solutions identified in the current context, and many other AI solutions may be available in the market.

Table 1: AI Use Cases in Electric, Categorization and Commercial Use Cases (July 2025).

Use Case	Domain	Description	Example Commercial AI Solutions (July 2025)
DERMS	Prediction and Optimization (Control)	- AI optimizes dispatch of DER assets, balancing grid conditions, market participation, and customer needs.	GE Vernova's GridOS DERMS,[52] Uplight's AutoGrid Flex,[53] and

[49] World Wide Technology. (2025, May 29). *The Orca AI journey with Southern California Edison*. WWT. https://www.wwt.com/blog/the-orca-ai-journey-with-southern-california-edison

[50] GE Vernova. (2025). GridOS® Orchestration Software. Retrieved from https://www.gevernova.com/software/products/gridos.

[51] Schneider Electric. (2025). EcoStruxure™ ADMS. Retrieved from https://www.se.com/us/en/product-range/61751-ecostruxure-adms/#overview.

[52] GE Vernova. (2025). GridOS® DERMS - Unlocking Opportunities for Future DSO Transition. Retrieved from https://www.gevernova.com/software/resources/webinar/gridosr-derms-unlocking-opportunities-future-dso-transition.

[53] Uplight. (2025). AutoGrid Unveils Cutting-Edge EV Grid Services Solution to Help Utilities Spur Adoption and Attain Sustainability Goals. Retrieved from https://uplight.com/press/autogrid-unveils-cutting-edge-ev-grid-services-solution-to-help-utilities-spur-adoption-and-attain-sustainability-goals/.

		- Localized AI processing in edge devices enhances efficiency.	Kraken's AI-powered tools.[54]
Predictive Maintenance and Asset Health Monitoring	Prediction	- AI analyzes sensor data from transformers and inverters to predict failures. - Determines optimal maintenance paths and timelines to minimize downtime.	GE Vernova's GridOS[55] and Schneider Electric's EcoStruxure ADMS[56]
Grid Forecasting (Load, Generation, Weather)	Prediction	- AI predicts future load and renewable generation patterns for improved dispatch planning. - Localized processing in edge devices supports real-time forecasting.	GE Vernova and Uplight's AutoGrid Flex[57]
Model Validation and Data Lake (Central Point of Truth, etc.)/ Digital Twin	Prediction	- AI validates models against a centralized data lake, providing a single source of truth. - Digital twins simulate grid operations for scenario testing and validation.	Opal RT Digital Twin Advanced Simulation[58] and Siemens Electrical Digital Twin.[59]
Power Flow and Interconnection Analyses	Prediction (Analysis)	- AI runs system impact studies, considering diverse energy resource mixes and dynamic model setups. - AI considers larger mixes of resources and more combinatorial situations. - AI sets up the models dynamically.	Google Tapestry,[60] Hitachi iQ with NVIDIA.[61]
Autonomous Grid Control (Real-Time Dynamic Protection and Remediation)	Optimization (Control)	-AI dynamically adjusts grid protection settings and autonomously reacts to frequency, voltage, and topology changes. - AI adjusts protection settings and reacts to grid changes, enhancing stability, and resilience. - Localized AI in edge devices supports real-time adjustments.	Google and Tapestry Grid Aware[62]

[54] Kraken. (2025). Upgrade your utility. Retrieved from https://kraken.tech/.
[55] GE Vernova. (2025). GridOS® Orchestration Software. Retrieved from https://www.gevernova.com/software/products/gridos.
[56] Schneider Electric. (2025). EcoStruxure™ ADMS. Retrieved from https://www.se.com/us/en/product-range/61751-ecostruxure-adms/#overview.
[57] Uplight. (2025). AutoGrid Unveils Cutting-Edge EV Grid Services Solution to Help Utilities Spur Adoption and Attain Sustainability Goals. Retrieved from https://uplight.com/press/autogrid-unveils-cutting-edge-ev-grid-services-solution-to-help-utilities-spur-adoption-and-attain-sustainability-goals/.
[58] OPAL-RT TECHNOLOGIES, Inc. (2025). Digital twins for advanced simulation. Retrieved from https://www.opal-rt.com/industries-and-applications/simulation-and-testing/digital-twins/.
[59] Siemens. (2025). Electrical Digital Twin. Retrieved from https://www.siemens.com/global/en/products/energy/grid-software/planning/electrical-digital-twin.html.
[60] *Tapestry*. (2025). Retrieved from Google LLC: https://x.company/projects/tapestry/.
[61] Southwest Power Pool . (2025). SPP Partners with Hitachi to Develop Advanced AI Solution. Retrieved from https://spp.org/news-list/spp-partners-with-hitachi-to-develop-advanced-ai-solution/.
[62] *Tapestry*. (2025). Retrieved from Google LLC: https://x.company/projects/tapestry/.

Fault Location (precursor to grid control)	Detection (Analysis)	- AI enhances fault location capabilities, aiding in efficient identification and isolation of issues	GE Vernova's GridOS DERMS.[63]
Outage Management System	Detection and Optimization (Control)	-AI OMS systems help utilities detect, locate, and restore power outages. - AI-driven models analyze historical data, weather patterns, maintenance records, and sensor data to predict the likelihood and location of potential outages	OATI Genie[64] and Entergy Louisiana's AI system.[65]
Energy Market Optimization (AI for Trading & Bidding Strategy)	Optimization	- AI optimizes participation in wholesale and retail energy markets through trend and price analysis	Fluence Mosaic[66] and Dexter Energy.[67]
Cybersecurity Ops	Detection (Security)	- AI monitors SCADA/ICS traffic for anomalies to detect cyber threats. - Anomaly detection capabilities in edge devices enhance security	Schneider Electric's EcoStructure ADMS.[68]
Customer Load Management & Flexibility (AI-Powered Demand Response)	Optimization (Demand Response)	- AI learns customer behavior to control smart devices for grid stability during peak demand - Enhance customer interaction and service management, thus contributing to overall customer satisfaction and operational efficiency.	Kraken's AI-powered tools,[69] LivePerson AI Chat[70] and Octopus Energy's GPT Chatbot.[71]

[63] GE Vernova. (2025). GridOS® DERMS - Unlocking Opportunities for Future DSO Transition. Retrieved from https://www.gevernova.com/software/resources/webinar/gridosr-derms-unlocking-opportunities-future-dso-transition.

[64] Open Access Technology International, Inc. (2025). Introducing OATI Genie. Retrieved from https://www.oatiai.com/.

[65] entergy. (2024, July 16). We're taking steps to advance reliability across our region. Retrieved from https://www.entergy.com/blog/we-re-taking-steps-advance-reliability-across-our-region

[66] Fluence. (2025). Intelligent, AI-powered bidding for solar, wind, and energy storage. Retrieved from https://fluenceenergy.com/mosaic-intelligent-bidding-software/?_gl=1*jlivmp*_gcl_au*NzY4MzMyNDUyLjE3NTQ1OTQ0NjA.*_ga*MTM0NDQyODc5OC4xNzU0NTk0NDYw*_ga.

[67] Dexter. (2025). Forecast & Trade Optimization. Retrieved from https://dexterenergy.ai/.

[68] Schneider Electric. (2025). EcoStruxure™ ADMS. Retrieved from https://www.se.com/us/en/product-range/61751-ecostruxure-adms/#overview.

[69] Kraken. (2025). Upgrade your utility. Retrieved from https://kraken.tech/.

[70] LivePerson. (2025). AI chatbot agents: Better conversations start with a better AI chatbot. Retrieved from https://www.liveperson.com/products/ai-chatbots/.

[71] Celonis, Inc. (2025). Connect Octopus Energy and OpenAI (ChatGPT, Whisper, DALL-E) integrations. Retrieved from Make: http://make.com/en/integrations/octopus-energy/openai-gpt-3.

Generation Prediction	Business (Communication)	- GenAI is used for drafting social media posts, managing customer communication, and automating processes. - Internally supports utility employees with knowledge assistants and chatbots. - Requires strict validation of AI outputs to ensure accuracy and data security.	Southern California Edison's Project "Orca"[72]
Wildfire Prediction and Monitoring	Prediction and Detection	- AI monitors real-time hot spots for potential ignition.	Pano Real-time hot spots for potential ignition,[73] LookOut,[74] and WIFIRE.[75]

3.3 Key Constraints for Model Selection and Case Studies in Power Systems AI

AI deployment in power system applications requires careful model selection that balances algorithmic capabilities with the physical realities, operational constraints, and contextual factors of the grid. Unlike purely digital domains, power systems are governed by well-defined physical laws, network topologies, and operational standards that constrain both the data available for analysis and the types of models that can be meaningfully applied. Effective AI design must integrate physics-based understanding, such as power flow equations, stability margins, and protection schemes, with contextual knowledge of regulatory requirements, market structures, and system operator practices.

At the foundation of any AI implementation lies data quality and accessibility. Power system data is often fragmented across supervisory control and data acquisition (SCADA) systems, phasor measurement units (PMUs), advanced metering infrastructure (AMI), and asset management systems, each with different temporal resolutions, sampling rates, and privacy considerations. The choice of model is inherently linked to these data characteristics: the temporal granularity, spatial coverage, and trustworthiness of the available measurements will influence whether statistical learning, deep learning, or hybrid physics-informed approaches are feasible and reliable. This section examines key constraints in selecting AI models for power system applications, emphasizing the interplay between physics and operational context, data foundations, and algorithmic suitability. It also reviews case studies demonstrating how these constraints shape real-world deployments, including examples where domain-aware AI has

[72] World Wide Technology. (2025). The GenAI Chatbot Transforming Network Operations at Southern California Edison. Retrieved from https://www.wwt.com/case-study/genai-chatbot-transforming-network-ops-at-sce.
[73] Pano, Inc. (2025). AI-Powered wildlife detection and situational awareness. Retrieved from https://www.pano.ai/.
[74] Robotics Cats. (2022). AI-powered Wildfire Detection System to save lives and cut loss. Retrieved from https://roboticscats.com/.
[75] University of San Diego. (2025). Workflows integrating collaborative hazard sciences. Retrieved from https://wifire.ucsd.edu/.

enabled successful integration into operational workflows, and cases where overlooking such constraints could lead to limited or unsuccessful outcomes.

3.3.1 Physics and Power Systems Context

Applying AI in the electric grid requires alignment of the model with power system physics and operational realities. The electrical grid's behavior is governed by physical laws (e.g., Kirchhoff's and Ohm's laws) and operational constraints (voltage limits, frequency stability criteria, protection settings). Ensuring the AI model is appropriate for this context is vital for safety and effectiveness. An AI solution must handle domain-specific data like bus voltages, line capacities, frequencies, power flows, and system state variables in a way that reflects physical reality; otherwise, it may produce infeasible or unsafe outputs. One key consideration is choosing or designing models that incorporate power system physics. Generic black-box models might ignore constraints, such as maintaining 60 Hz frequency or keeping voltage within ±5% of nominal, which could lead to recommendations that violate grid limits. The importance of embedding domain physics is underscored by researchers as a way to augment model robustness and reliability, especially in critical functions.

Voltage, Frequency, and State Data: Data representing system state (voltage magnitudes or angles, frequency, currents, etc.) carry physical meaning that AI models should treat appropriately. For instance, frequency deviations in a power system are typically very small (e.g., 59.8–60.2 Hz) but highly significant; an AI model doing frequency prediction or anomaly detection must have the resolution and understanding to detect a 0.2 Hz drop as a major event, not noise. Similarly, bus voltages are coupled by the network; if an AI recommends a control action that would raise voltage in one area, it must account for how it affects neighboring buses and overall system balance. Thus, feature engineering and input selection for grid AI should include physically relevant variables (like including system load, generation, intertie flows, etc.) rather than treating the data as a generic or independent time series. In anomaly detection, distinguishing a true physical anomaly from a sensor error requires understanding physical patterns; for example, a simultaneous dip in voltage and frequency system-wide is likely a real disturbance, whereas a dip in one sensor reading alone might be a bad measurement.

Physics-based relationships can be expressed as constraints on the model but come at a cost of more processing power. For example, checks for physics-based validation can also be checked after the model has come to an output; however, if constraints are not met, the user must to start over instead of making adjustments during the process. Models may "learn" relationships between variables that exist because of the physics-based dependencies, but that does not mean they will always respect that relationship, since there are a variety of conflating objectives for the algorithm. Testing against edge cases is important to verify if a model truly respects the real-world constraints.

Model Selection and Validation: The consideration of physics also influences which AI model types are suitable. For example, a purely data-driven deep learning model might need an enormous amount of data to learn something that a simpler physics-based model already "knows." In cases where data is limited, combining first-principles power system models with AI (for instance, using load flow calculations to generate features, or using residual physics models to guide learning) can improve performance. As noted in a recent DOE-supported

study, integrating domain knowledge into ML frameworks remains a persistent challenge, but it is crucial for trustworthiness.[76] Practitioners should favor model architectures that can easily integrate prior knowledge. For example, using a model that allows adding hard constraints (like linear programming with an ML-based objective) for tasks like dispatch, or using multi-model ensembles that include a physics-based estimator alongside an AI predictor. Finally, any AI meant to act on the grid should be extensively validated against simulations and field data to verify it does not violate physical requirements. Validation guidelines should be developed for detecting issues in drift and other spaces. The use of digital twins (high-fidelity grid simulators mirroring real operations) is emerging as a best practice. This simulation testing, combined with gradual field trials, ensures that the AI's outputs always make physical sense and keep the system within safe operating bounds.[77] [78] [79]

3.3.2 Data Foundation for Utility AI

Data is the backbone of any AI solution, and in the electric grid context, establishing a strong data foundation is a critical consideration. Model performance and reliability are heavily dependent on the data quality, quantity, and accessibility.[80] Key aspects to consider include data volume needs, integration of diverse data sources, creating a "single source of truth," data pre-processing pipelines, latency requirements, and governance or security measures. These issues are not isolated to AI use. Data volume and quality are critical for many new digital grid functions, such as advanced distribution management systems, which require validated and accurate grid models to perform.[81] Some applications require highly accurate data input (in the order of 99.9%) data, while others can approximate. This must be understood along with the volume, quality, and accuracy. Training AI on simulated data must account for simulation inaccuracies and tolerance, which can sum to around 10% margins.[82]

Data Volume and Quality: Modern AI, especially deep learning, is a data and power-intensive application. Utilities generate massive streams of data (SCADA measurements, phasor

---

[76] Chen, Yousu, et al. "Artificial Intelligence/Machine Learning Technology in Power System Applications," Apr. 2024. https://doi.org/10.2172/2340760

[77] Pacific Northwest National Laboratory (PNNL). (2020, November 9). *PNNL Researchers Speed Power Grid Simulations Using AI*. Retrieved from U.S. Department of Energy: https://www.pnnl.gov/news-media/pnnl-researchers-speed-power-grid-simulations-using-ai

[78] OPAL-RT TECHNOLOGIES, Inc. (2025, August 22). *7 real power hardware-in-the-loop examples that sharpen your simulation strategy*. Retrieved from https://www.opal-rt.com/blog/7-real-power-hardware%E2%80%91in%E2%80%91the%E2%80%91loop-examples-that-sharpen-your-simulation-strategy/

[79] McKinsey & Company. (n.d.). *Digital twins and generative AI: A powerful pairing*. Retrieved from https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/digital-twins-and-generative-ai-a-powerful-pairing

[80] U.S. Department of Energy, Grid Modernization Laboratory Consortium. *GMLC Sensing Measurement Strategy Sensor Technology Roadmap Final Report*. April 2019. GMLC, DOE

[81] U.S. Department of Energy. (2015, February). *Voices of experience: Insights into Advanced Distribution Management Systems* (DOE/GO-OT-6A42-63689). National Renewable Energy Laboratory; Energetics Incorporated

[82] E. M. Stewart, S. Kiliccote, D. Arnold, A. von Meier and R. Arghandeh, "Accuracy and validation of measured and modeled data for distributed PV interconnection and control," *2015 IEEE Power & Energy Society General Meeting*, Denver, CO, USA, 2015, pp. 1-5, doi: 10.1109/PESGM.2015.7286484.

measurements, smart meter reads, equipment sensor logs, weather feeds, etc.), but not all of it may be AI-ready. Much is pre-processed before any use in an application. To be used appropriately, data must undergo exhaustive exploratory data analysis and preprocessing to enable machine learning, which could be up to 80% of the model development time.[83] High-quality, labeled datasets might be scarce for certain use cases (e.g., examples of rare faults). Practitioners must assess how much data is needed and whether it is available. For instance, training a transformer-failure prediction model might require many years of historical records across many transformers; if a utility only has a dozen failure examples, a purely data-driven model may not be feasible without augmentation or transfer learning. Data quality is equally important: noisy or erroneous data can mislead AI models. Data cleaning and validation are, therefore, foundational steps. As one technical report notes, the effectiveness of AI models relies on data variety, volume, and recency and obtaining comprehensive, accurate data (potentially from many sources) remains a major hurdle. It is vital to invest in improving data accuracy (calibrating sensors, filtering out corrupt entries) and completeness, along with the implementation of AI in power systems.

Simulated Data for Training: It's not uncommon to see AI models trained on data that comes from simulations.[84] This can create data that is hard to otherwise obtain (e.g., different types of faults at multiple locations within a given system), and it is attractive for researchers because it produces a cleaner source of data (e.g., no gaps in timestamps or erroneous sensor collections). However, few models accurately represent all behaviors of a system (consider the common saying, "no model is right, some are useful"), and this method of data generation often leaves out edge cases which do appear in the real world but which the researcher may not have considered. AI models trained on simulated data may have holes in their understanding which present risks when deployed for operational use.

Data Discovery and Integration (Multimodal Data): Utilities often have "scattered datasets" across different systems. For example, asset health records in one database, outage history in another, weather data in a third, etc. An AI initiative must include data discovery: identifying and bringing together these disparate sources. A multimodal or multivariate data approach can greatly enhance AI capabilities, combining time-series measurements, text reports (e.g. maintenance logs), geospatial data (e.g. GIS of network), images (e.g. line inspection photos), and more to give a holistic view. However, integrating these diverse datasets is challenging. It requires a modern data architecture and governance processes to reconcile different data schemas and ensure everything lines up correctly.[85] Best practice is to establish a single source of truth – e.g., a unified data lake or database where all relevant data is collected, cleaned, and keyed in a

---

[83] Medium. (2025, March 4). *Data Cleaning Why 80 Percent of Data Science Is Spent Fixing Dirty Data*. Retrieved from https://medium.com/@preetikapuria587/data-cleaning-why-80-percent-of-data-science-is-spent-fixing-dirty-data-0d0a214ce5c0

[84] National Renewable Energy Laboratory (NREL). (2025). *Generative Artificial Intelligence for the Power Grid*. Retrieved from U.S. Department of Energy: https://www.nrel.gov/grid/generative-artificial-intelligence-for-the-power-grid

[85] Stewart, Emma, Chellappan, Karthik, Backhaus, Scott, Deka, Deepjyoti, Reno, Matthew, Peisert, Sean, Arnold, Dan, Chen, Chen, Florita, Anthony, & Buckner, Mark (2018). Integrated Multi Scale Data Analytics and Machine Learning for the Grid; Benchmarking Algorithms and Data Quality Analysis. https://doi.org/10.2172/1490956

consistent way. This prevents the situation of different teams using slightly different data (or definitions) that lead to inconsistent AI outputs. AI for outage prevention[86] [87] emphasizes that data from asset health, vegetation management, fault history, and weather must be integrated, cleansed, and made accessible for AI to distinguish between these types of events, and this demands clarity on who owns each data source, how it is validated, and how often it is updated. In practice, many utilities start with a data audit and build pipelines to stream and merge data. For example, linking smart meter outage flags with feeder SCADA telemetry and weather stations for a storm prediction model.

Pre-Processing Pipelines: Once data sources are identified, robust pre-processing is needed before feeding data to AI models. This includes steps such as filtering out bad or irrelevant data (e.g., communication glitches, out-of-range values), handling missing data (through interpolation or using domain-specific proxies), normalizing or scaling features (especially important for mixing variables like power, voltage, temperature with very different units and ranges), and aligning data of different time granularities. Multivariate time synchronization is often needed – e.g., aligning PMU data at 30 samples per second with SCADA data at 1 sample every 4 seconds and weather data every 5 minutes.[88] [89] Feature engineering is another part of pre-processing; creating composite features that capture domain knowledge, such as load to capacity ratio of a transformer or rate-of-change of frequency can significantly improve model performance. Additionally, if a project involves supervised learning, a pipeline to label the data is needed. For example, automatically labeling historical events as "fault" or "no fault" based on relay logs when training a fault detector. Automating these pipelines ensures consistency and allows data to flow into AI models continuously. It is also critical to incorporate feedback loops in the pipeline. For instance, flagging when incoming data falls outside the range seen in training to warn the model might be extrapolating. All these steps add latency, so they must be designed to meet the operational timeline.

Latency and Real-Time Requirements: It is not just training data that needs to be considered. Various grid applications have different latency needs, which affects the quality and quantity of data needed for these applications to run in operation. Some analyses can be done offline (e.g., a long-term asset risk model might be run monthly on batch data), but many operational applications require real-time or near-real-time data streaming and processing, such as fault applications.[90] For example, an AI-based anomaly detector in the control room might need to ingest PMU data with only a few seconds of delay to be useful for grid operators. This necessitates a data pipeline that can handle streaming data and perform inference quickly. It

---

[86] M. Jamei *et al.*, "Anomaly Detection Using Optimally Placed μ PMU Sensors in Distribution Grids," in *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3611-3623, July 2018, doi: 10.1109/TPWRS.2017.2764882

[87] Q. Huang, R. Huang, W. Hao, J. Tan, R. Fan and Z. Huang, "Adaptive Power System Emergency Control Using Deep Reinforcement Learning," in *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1171-1182, March 2020

[88] Pattanaik, V., Malika, B. K., Panda, S., Rout, P. K., Sahu, B. K., Samanta, I. S., Bajaj, M., Blazek, V., & Prokop, L. (2024). *A critical review on phasor measurement units installation planning and application in smart grid environment*. *Results in Engineering, 24*, Article 103559. https://doi.org/10.1016/j.rineng.2024.103559

[89] M. Madadi *et al.*, "Distributed Hierarchical Sensing and Analytics for Grid-Edge Behind the Meter Visibility and Grid Resilience," *2023 IEEE Energy Conversion Congress and Exposition (ECCE)*, Nashville, TN, USA, 2023, pp. 6580-6582, doi: 10.1109/ECCE53617.2023.10362941.

[90] J. Morgan and E. Stewart, "Evaluation and Use Cases in Cloud Implementation for Future Electric Grid Technologies," *2024 IEEE Power & Energy Society General Meeting (PESGM)*, Seattle, WA, USA, 2024, pp. 1-7.

might mean placing certain computations at the network edge, like in substation or field devices, to avoid round-trip delays to a central cloud. Indeed, edge computing and distributed AI are emerging so that critical decisions, such as isolating a fault or regulating voltage, can be made in sub-seconds and at the grid edge.[91]

When designing AI systems, engineers must account for the end-to-end latency: from data generation, through communication networks and any data bus or broker (e.g., an enterprise SCADA historian or message queue), into the model, and back out to a control command. If the total latency is beyond the requirement (say, an AI that adjusts a battery should respond within one second, but the data path takes five seconds), then architectural changes are needed. Sometimes this involves model simplification or optimization for speed, using a smaller neural network or compiling models to faster runtimes. It may also involve prioritizing which data is truly needed in real-time (sending high-frequency essential signals and leaving less critical data for batch updates). In summary, aligning the AI's data pipeline with operational timing is crucial for success.

Data Governance and Security: Given the sensitive and critical nature of grid operations data, governance and security cannot be overlooked. Data governance includes defining clear ownership of data streams, setting up data quality audits, and managing access controls so that only authorized systems and personnel can use the data. Utilities often deal with customer data (from smart meters), which has privacy implications, as well as critical infrastructure data (substation, generation status), which has national security implications. A single source of truth helps enforce governance rules uniformly. Additionally, metadata (data about the data) should be maintained; knowing the source, timestamp, units, etc., for each data point to avoid misinterpretation by AI models.

On security, the introduction of AI and networking in grid operations expands the attack surface. AI models themselves can be targets of adversarial attacks or data poisoning. An attacker who feeds manipulated data into an AI-based control system could cause it to misoperate the grid in a worst-case scenario or create long-term small errors contributing to financial issues. Therefore, utilities must secure data pipelines, such as through encryption in transit and at rest and authentication of sensors to ensure data integrity, and possibly monitor AI outputs for signs of tampering. Recent assessments highlight risks like false data injection, where maliciously altered sensor data causes AI models to make incorrect decisions.[92] Strong security measures, following standards such as NERC CIP for cyber protection of critical systems, are required to mitigate these risks. In practice, this means including the IT or security teams in AI project planning and using secure gateways and implementing fail-safes or fallbacks if data is suspicious. Data resilience is also part of security, ensuring that if one data source fails (say, a PMU is offline),

---

[91] Clarion Energy. (2025, February 6). *Resilience through intelligence: AI's impact on utility disaster readiness*. Retrieved from Factor This Engineering: https://www.renewableenergyworld.com/power-grid/smart-grids/resilience-through-intelligence-ais-impact-on-utility-disaster-readiness/#:~:text=Automated%20Actions%3A%C2%A0%20AI%20empowers%20utilities,AMI%20in%20enhancing%20grid%20reliability.

[92] U.S. Department of Homeland Security. (2023, December 22). *Risks and mitigation strategies*. Science and Technology Directorate. https://www.dhs.gov/sites/default/files/2023-12/23_1222_st_risks_mitigation_strategies.pdf

the AI system can gracefully degrade or switch to backup data. Lastly, compliance with privacy regulations is key. For example, if using smart meter data for an AI model, it's important to aggregate or anonymize as needed to protect customer identities. Establishing proper governance, privacy handling, and security from the start not only protects the utility and customers but also builds confidence that the AI's insights can be trusted and are legally compliant.

4. Balancing Risk and Benefit of Adoption of AI in Electric Utilities

There are several key risks associated with AI adoption in grid planning and operations. These challenges range from technical considerations of the models and data themselves to the ways in which using AI will change human behaviors and reactions, to the potential for these tools to be misused or adversarial manipulated. This section will identify several key risks categorized as cybersecurity, operator and planning, regulatory, supply chain, privacy, and human-related risks. AI adoption and, therefore, risk can be grouped into functions similar to how cybersecurity is approached in these spaces. In the IT spaces, primary functions revolve around human processes, such as data and billing management, and customer engagement. These kinds of applications in GenAI are considered simple models that have low operational risk but potentially higher customer data security risk.

In OT spaces, functional and mission assurance is key, in particular, safety, reliability, and performance. Accuracy is paramount in life-saving applications, and this is why so many safety-critical functions are manually verified, such as lockout or tagout. In planning applications, AI may be trusted to make decisions that have longer-term implications for the reliability, performance, and economics of the power grid. In operations applications, timing and performance, as expected, are key. Whether humans are part of the decision loop or actions are fully autonomous, trusting the AI to make decisions or recommendations that are most appropriate in the moment can make or break the usability of the model. Utilities in California and elsewhere have faced real-world scenarios where cutting power, despite the significant inconvenience and economic disruption, was deemed the lesser risk compared to potential loss of life or catastrophic infrastructure damage. Public Safety Power Shutoff (PSPS) events implemented by Pacific Gas and Electric (PG&E) and other utilities exemplify this trade-off. In high wind and drought conditions, proactively de-energizing lines can prevent equipment faults from igniting wildfires, a danger that has historically caused billions in damage and claimed many lives. The same principle applies in unusual operational incidents, such as the case of a woman who climbed onto an energized transformer, where operators had to quickly disable equipment to prevent injury, despite the cost of localized outages.[93]

[93] Jennings, A. (2024, November 13). *More than 800 Utah homes lose power after woman climbs transformer*. Retrieved from https://www.abc4.com/news/digital-exclusives/800-homes-lose-power-trespass/

This trade-off reflects an "AI trolley problem"[94] dynamic when decision support systems or automated controls are integrated into grid operations. The system may be required to choose between two unfavorable outcomes: continuing service under unsafe conditions versus interrupting service to avert a potentially more severe event. In human decision-making, these judgments draw on experience, training, and safety regulations; in AI-assisted environments, the same logic must be encoded into model objectives, risk thresholds, and control algorithms. In both human and machine contexts, the guiding principle is that temporary loss of power is preferable to the irreversible consequences of wildfire ignition, severe injury, or death. Embedding such prioritization into AI systems ensures that safety remains the dominant factor, even when automated decision-making is involved in real-time grid control. Key risk themes include cybersecurity, operational resilience, physical security, regulatory, supply chain, and privacy (Table 2). An additional risk to consider is human talent and knowledge drain in power operations. As AI automation is used to replace human tasks, humans may lose the skills, expertise, and contextual awareness required to execute these tasks, creating risk if AI is unavailable or makes decisions that humans cannot identify as risky behaviors.

Utilities must take a secure and responsible approach when integrating AI and cloud technologies. Generative AI can greatly improve data management and customer engagement, but it also introduces serious challenges to data security. On the operational side, AI plays a pivotal role in supporting decisions that affect grid reliability and economic performance. However, the adoption of AI in grid planning and operations brings distinct risks, ranging from technical vulnerabilities and shifts in human behavior to opportunities for misuse. These risks can be grouped into the categories identified in Table 2 below.

Table 2: Summary of key risk themes.

Category	Example Risk
Cybersecurity	Adversarial attacks on AI models, causing grid disruption
Operational	Loss of human oversight due to black-box AI decisions, Synchronized responses destabilizing frequency or voltage
Regulatory	Lack of standards for AI validation in critical infrastructure
Supply Chain	Dependency on foreign AI models, chips, and firmware
Privacy	Customer, PII, and CEII data

4.1 Cybersecurity Risks

Several potential vulnerability classes of AI and ML are being researched. To determine how these may result in consequences to power grid operations and planning, the authors utilize the OWASP Top 10.[95] Table 3 outlines the critical cybersecurity threats associated with the adoption of AI, highlighting potential vulnerabilities and risks that utilities must address to safeguard their operations.

[94] The Alan Turning Institute. (2025). *AI's "Trolley Problem" Problem*. Retrieved from https://www.turing.ac.uk/blog/ais-trolley-problem-problem

[95] OWASP. (n.d.). *ML01:2023 Input Manipulation Attack*. Retrieved from https://owasp.org/www-project-machine-learning-security-top-10/docs/ML01_2023-Input_Manipulation_Attack

Table 3: OWASP's Top 2023 LLM vulnerability list and descriptions.

OWASP vulnerability Name	Description of OWASP Vulnerability
ML01:2023 Input Manipulation Attack	A type of attack in which an attacker deliberately alters input data to mislead the model.
ML02:2023 Data Poisoning Attack	An attacker manipulates the training data to cause the model to behave in an undesirable way.
ML03:2023 Model Inversion Attack	Model inversion attacks occur when an attacker reverse-engineers the model to extract information from it.
ML04:2023 Membership Inference Attack	Membership inference attacks occur when an attacker manipulates the model's training data in order to cause it to behave in a way that exposes sensitive information.
ML05:2023 Model Theft	Model theft attacks occur when an attacker gains access to the model's parameters.
ML06:2023 AI Supply Chain Attacks	In ML Supply Chain Attacks, threat actors target the supply chain of ML models. This category is broad and important, as software supply chain in ML includes even more elements than in the case of classic software.
ML07:2023 Transfer Learning Attack	Transfer learning attacks occur when an attacker trains a model on one task and then fine-tunes it on another task to cause it to behave in an undesirable way.
ML08:2023 Model Skewing	Model skewing attacks occur when an attacker manipulates the distribution of the training data to cause the model to behave in an undesirable way.
ML09:2023 Output Integrity Attack	In an Output Integrity Attack scenario, an attacker aims to modify or manipulate the output of a machine learning model in order to change its behavior or cause harm to the system it is used in.
ML10:2023 Model Poisoning	Model poisoning attacks occur when an attacker manipulates the model's parameters to cause it to behave in an undesirable way.

4.2 Operational and Planning Risks

These risks can be systematically grouped into two broad categories, operational and planning, providing a structured framework for assessing the potential impacts of AI integration on grid reliability (Table 4Table 4).

Table 4: Safety risk and description.

Risk	Description
Loss of Operator Understanding	AI models may make decisions that operators cannot interpret, reducing situational awareness and impairing crisis response.
Operations Model Drift	Models trained on historical data may fail when grid conditions change.

Automation Dependency	Over-reliance on AI may reduce manual system checks, leading to latent failures that go unnoticed until critical.
Inflexible Failover Paths	AI-based control systems might lack robust manual override or fail-safe options in case of malfunction.
Model Validation Challenges	There is no standardized way to test AI for safe grid deployment under all scenarios, making certification difficult.
Load and Generation Forecast Errors	AI systems may mispredict load or renewable generation, leading to imbalances and reserve margin failures.
Safety	If AI is used in protective relaying or fault isolation, incorrect decisions could lead to cascading outages.
Power Outage	Failure in safety, or other automation or operations results in a power outage.

4.3 Regulatory and Policy Risk

These challenges can also be systematically grouped into categories that highlight their implications for standards, jurisdiction, and accountability, providing a structured framework for evaluating regulatory and policy risks (Table 5).

Table 5: Regulatory risk and description.

Risk	Description
Standards Lag	Existing grid reliability standards (NERC CIP, IEEE, etc.) are not fully equipped to address AI-specific risks, leaving gaps in regulation.
Policy	Policies at federal, regional, and state levels may restrict or change the way that AI can be used in different applications and may not be informed by the state-of-the-art technical best practices. Additionally, policies may vary across jurisdictional territories, making it difficult to standardize the adoption of best practices.
Liability, Responsibility, and Accountability	It may be unclear who is liable if an AI-enabled device causes a grid event—manufacturer, AI vendor, utility, or operator

4.4 Supply Chain

These vulnerabilities can be systematically grouped into categories that reflect risks in sourcing, hardware dependencies, and development pipelines, providing a structured framework for evaluating supply chain security (Table 6).

Table 6: Supply chain risk and description.

Risk	Description
Foreign Adversary Control Points	AI software or hardware may be sourced from or maintained by foreign entities, raising concerns about embedded malicious code or operational dependencies.
Hardware-Software Coupling Risks	AI functions may require specialized chips (e.g., GPUs, NPUs) that have their own security and supply chain issues, especially if sourced from non-U.S. manufacturers.

Software Development Pipeline	AI applications that are used in the grid may involve multiple stakeholders, including not only the primary tool provider, but also software integrators, developers of models on top of which utility applications are built, and more. If not all of these actors understand the utility's primary objectives, the deployed model may not meet all of the safety and reliability needs.

4.5 Privacy and Security

These concerns can be grouped into categories spanning customer privacy, financial integrity, and data security, providing a structured framework for assessing the vulnerabilities AI introduces to sensitive information (Table 7).

Table 7: Security risk and description.

Risk	Description
Customer Data Breach/Manipulation/Loss	Customer data, including personally identifiable information (PII) or power usage-data (AMI data breach [96]) that reveals personal habits, may be accidentally revealed by AI if proper protections are not in place. Manipulation of such data could lead to billing errors.
Planning Data Breach/Loss/Manipulation	Loss of Critical Data, Manipulation of Data used nefariously

5. Consequence Analysis and Attribute Decomposition of Cyber Physical Risk of the Use of AI

In AI-enabled power system applications, consequence analysis must evaluate not only the projected physical and operational impacts of an event on the grid, but also the performance constraints of the AI system itself. This additional layer is essential because the reliability of AI-derived outputs directly influences the quality and timeliness of operator actions. The decomposition into availability, accuracy, and speed or delay provides a structured method for characterizing the AI system's operational profile (Figure 4). From the system perspective, availability refers to the proportion of operational time the AI application remains functional and able to process inputs, accounting for hardware uptime, data pipeline continuity, and model service stability, and the application availability itself, for example if a utility were to lose access to its fault location tooling, they likely can manually process, but if the utility lost access to its OMS, there may be more significant events such as longer outage processing times and therefore longer times with lights out for customers. Accuracy measures the system's ability to generate correct classifications, forecasts, or anomaly detections under diverse operating conditions and data quality states, including during contingencies or degraded measurements. An error in fault location may result in a significant cost in digging the wrong space for access to the power line. Speed/delay encompasses end-to-end latency from data acquisition through AI inference to

[96] Kovacs, E. (2025, July 9). *Canadian Electric Utility Says Power Meters Disrupted by Cyberattack*. Retrieved from SecurityWeek: https://www.securityweek.com/canadian-electric-utility-says-power-meters-disrupted-by-cyberattack/

delivery of actionable output, with consideration of worst-case execution times under high computational load. Restoration and fast reaction tooling such as line drop detection, to prevent a wildfire, cannot tolerate latency, but power systems planning can.

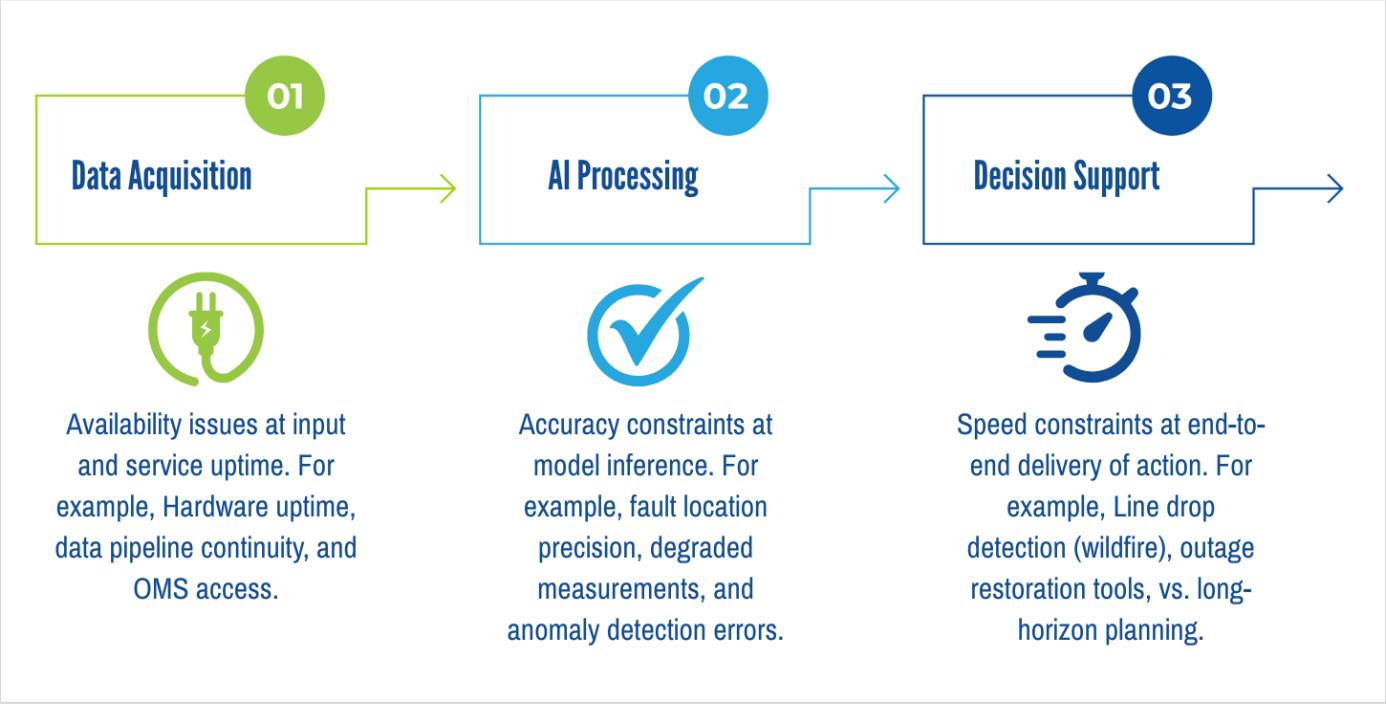

Figure 4: Performance Constraints of AI for Grid.

From the grid impact perspective, availability failures in the AI system may translate into delayed fault isolation, missed protection coordination windows, or reduced situational awareness. Accuracy deficiencies may result in incorrect dispatch signals, misidentification of incipient instability, or failure to detect cyber-physical intrusions, each carrying distinct safety, reliability, and resilience consequences. Excessive speed or delay may render AI outputs non-actionable in time-critical operations such as remedial action schemes, fast frequency response, or wildfire-related PSPS decision-making. By mapping AI system performance attributes to specific grid consequence categories, including safety, reliability, resilience, reputation, and economic cost, operators can quantify the dependency of operational risk on AI system behavior. This approach ensures that consequence analysis remains physically grounded in grid performance metrics while explicitly incorporating the technical constraints and capabilities of the AI system delivering the analysis

For each use case or AI application domain, it's possible to evaluate key operational constraints such as latency, availability, and accuracy of the solution and evaluate these in a matrix. This evaluation and remediation or protection approach is evaluated against the top 10 OWASP styles of AI attack or failure (see Table 3).

The authors selected several AI-enabled applications for this analysis and categorized them using a subset of OWASP's Top Risks and Vulnerabilities for LLM applications. A heatmap was then created to show which applications pose the highest consequence if the identified OWASP vulnerabilities and/or risks are exploited.

5.1 Consequence Evaluation

The consequences of each AI event style, on the use case are evaluated utilizing the following criteria, noting that this is not exhaustive, but intended to determine a prioritized range of consequences for ranking and mitigating risk. As a framework, individual entities may consider alternative features, but as a starting point these are the primary concerns of many utilities.

Note: The applications were rated on a scale from green to red, with green indicating the lowest consequence or risk and red indicating the highest consequence or risk.

Table 8: Criteria and Mapped Consequence of AI Event Styles.

Criteria	Consequence of AI Event Style
Ability to Serve Load	- Size of Outage: Ranging from single home to whole region or territory depending on entity assessing. - Duration of Outage: How long the power would likely be impacted (momentary to days).
Safety	- Potential for a safety event, from low to loss of life.
OT Asset Integrity	- Can be reinstated quickly, or needs fully replaced and/or no longer trusted.
Privacy	- Data remains private, to full loss of data.
Critical Data Loss	- None to all.
Cost of Event	- Low loss event to entity bankruptcy.
Reliability of Grid	- Impact to metrics, such as CAIDI and SAIDI.
Reputation or Customer Satisfaction	- Many utilities are ranked on customer satisfaction and reputational damage can significantly impact their business operations.
Ability to Reason about System	- Potential complexity and timescale impact to reason about that system.

Putting together the evaluation, the style of use case and the OWASP Top 10, the authors created an extensive list of potential consequences of the style of AI issue. An example is provided in Table 9, with a complete table available in Appendix A: Consequence Table Heatmap, along with additional descriptions for that table in Appendix B: Consequence Table Descriptions. The heatmap in Appendix A categorizes AI vulnerabilities across use cases by severity, using a color-coded scale from green (lowest consequence) to red (highest consequence). This evaluation is informed by the 2023 OWASP Top 10 for LLMs, which highlights key risks such as prompt injection, model poisoning, and insecure feedback loops.

Disclaimer: While this consequence ranking provides a structured view of potential impacts, the actual severity of AI vulnerabilities depends heavily on the specific model architecture, deployment environment, operational dependencies, and organizational context. A use case ranked as "moderate" in one system may be "high" in another, depending on how the AI is integrated and relied.

Table 9: Example of consequence ranking with varying severities.

Use Case	ML1: Input Manipulation	ML8: Model Skewing	ML10: Model Poisoning	Total Consequence Ranking
Predictive Maintenance and Asset Health Monitoring	Injected false sensor data (temperature, vibration) causes premature or missed maintenance.	Maintenance team interventions bias model to ignore outliers, degrading predictive performance.	Weights are tampered to reduce model sensitivity to critical failure indicators like thermal spikes.	Green = Lowest Consequence
Model Validation and Data Lake	Erroneous or mislabeled data leads to incorrect model acceptance or rejection.	Spurious correlations or feedback loops reinforce historical biases or flawed assumptions.	Poisoned or inconsistent data schemas degrade model traceability and long-term performance.	Yellow = Moderate Consequence
Fault Location (Precursor to Grid Control)	Injected harmonics or noise spoof true location, delaying restoration.	Field technician error feedback retrains model(s) to ignore certain anomalies.	Internal model logic gradually diverges, misplacing faults systematically over time.	Orange = Moderate - High Consequence
Autonomous Grid Control (Real-Time Dynamic Protection and Remediation)	Compromised frequency or phase angle data triggers false grid separation or reclosure.	Operator feedback to dampen transient events shifts control away from robust configurations.	Small, persistent changes in training degrade grid stabilization ability over time.	Red = Highest Consequence

Ultimately, the evaluation underscores the importance of applying the greatest rigor to AI applications that fall within the "red" or high-consequence category. Rigor, in this context, refers to the disciplined application of robust safeguards: comprehensive testing and validation, continuous monitoring for drift or manipulation, alignment with evolving standards, and structured contingency planning. Utilities should ask not only whether AI is needed for a given use case, but also whether the benefits of adoption meaningfully outweigh the potential risks. Choosing not to adopt carries its own risks, such as foregone efficiencies, slower response times or missed opportunities for resilience. However, adoption without sufficient mitigations could expose the grid to unacceptable vulnerabilities. Safe deployment therefore requires striking a balance: prioritizing high-value use cases, investing in layered defenses, and implementing governance mechanisms that ensure AI augments rather than undermines grid reliability.

6. Repercussions of Non-AI Deployment or Limited Deployment

While much attention is given to the risks induced by AI in power systems, the risks of not deploying AI, or deploying it in a limited, fragemented, or delayed manner, must also be evaluated. Some business drivers for reducing cost in utilities, and therefore not increasing rates,

include increasing reliability and increasing customer satisfaction will outweigh the risks posed in utilizing the modernized functionality. This has been one of the reasons cloud adoption in smaller utilities is outpacing large IOUs, and there is potential for the same to occur in AI adoption.

Using the availability, accuracy, and speed or delay decomposition provides a consistent framework to assess operational risk in both scenarios. From the AI system performance perspective, availability refers to the operational uptime of the AI service; accuracy is the rate of correct outputs under varying data and grid conditions; and speed or delay measures inference latency relative to the time-criticality of the function. These characteristics map directly to grid repercussions (Table 10).

Table 10: AI Characteristics Mapped to Grid Repercussions.

Characteristics	Grid Repercussions
Availability	- Failures can delay protective actions
Accuracy	- can lead to incorrect operational decisions
Speed or Delay	- can make outputs non-actionable in time-sensitive scenarios

From the non-adoption perspective, absence of AI can introduce comparable risk in a future modernized grid. In terms of availability, reliance on purely manual processes can bottleneck operations, especially under high workload or during emergencies. In terms of accuracy, human operators, typically managing a strained workload, are more prone to misinterpret complex system states, increasing error rates. In terms of speed or delay, manual analysis can be significantly slower, leading to delayed interconnections, slower contingency response, or missed real-time mitigation opportunities. Non-adoption can also limit the organization's ability to leverage advanced grid technologies such as adaptive protection schemes, predictive maintenance, or DER orchestration, constraining system flexibility, resilience, and integration capacity. With grid deployments and operational environments becoming more complex, often surpassing the threshold of human cognitive processing capacity, AI becomes increasingly critical for managing the scale, speed, and multidimensional trade-offs inherent in modern power system decision-making.

6.1 Consequence Analysis

By scoring both AI-enabled and non-AI scenarios qualitatively—classifying the consequence in each attribute as low, medium, or high—operators can compare not only the risks of using AI but also the operational and strategic costs of forgoing it. This approach ensures decision-making accounts for the full spectrum of impacts, enabling a more complete viability assessment when selecting solutions.

In the example of consequence ratings below for select non-adoption of AI application for power systems, the same scale of evaluation is used as for evaluating the risks of AI adoption, namely the same criteria in Table 8.

Table 11: Example use cases of consequence of non-AI adoption.

Use Case	Consequence of Non-AI Deployment	Ranking
Fault Detection	Without AI-driven fault location or predictive analytics, outage sizes may increase and restoration times may lengthen. Manual processes may not scale during widespread events. Delayed detection of hazardous conditions (e.g., line drops, thermal overloads) increases the risk of safety incidents, especially in wildfire-prone or high-voltage environments.	**High**
Predictive asset maintenance	Lack of AI-driven predictive maintenance may lead to situations of undetected degradation, resulting in catastrophic failures and costly replacements.	**Low**
Outage Management	Operational inefficiencies, prolonged outages, and reactive maintenance can escalate costs significantly, especially during high-impact events. Longer outages and slower service restoration can erode public trust, especially in regions where utilities are benchmarked on customer satisfaction. missed optimization opportunities.	**Moderate**
Autonomous Grid Control	As grid complexity increases, human-only reasoning becomes insufficient to manage real-time trade-offs, leading to suboptimal or delayed decisions.	**Moderate**

The evaluation of non-AI deployment consequences reveals that inaction is not a neutral or risk-free decision—it introduces its own spectrum of operational, safety, and strategic vulnerabilities. As power systems grow in complexity and interdependence, the limitations of automated or fixed processes become more pronounced, particularly in areas requiring speed, precision, and adaptability.

7. Engineering Controls and Approach to Develop Solutions

Using this approach, an AI user in a utility setting can evaluate their risk tolerance and develop appropriate controls. By applying this approach, utilities can align AI deployment strategies with their defined risk tolerance and operational priorities. This alignment enables selection of appropriate deployment architectures—for example, on-premises, hybrid, or edge-based AI systems—based on the criticality of the function and associated risk profile. Risk-informed solution deployment can be further reinforced through structured methodologies such as Cyber Informed Engineering (CIE), [97] which integrates cybersecurity considerations into system design

[97] Office of Cybersecurity, Energy Security, and Emergency Response (CESER). (n.d.). *Cyber-Informed Engineering*. Retrieved from U.S. Department of Energy: https://www.energy.gov/ceser/cyber-informed-engineering

from inception, and applicable NIST frameworks—for example, the *NIST Cybersecurity Framework (CSF)* for risk management or *NIST AI Risk Management Framework (AI RMF)* for trustworthy AI practices. These frameworks help ensure that AI systems are not only technically robust, but also secure, resilient, and compliant with industry standards, while providing clear checkpoints for governance, validation, and continuous monitoring.

To mitigate systemic failures, utilities should implement consequence-driven engineering controls (ECs) that bound AI within safe operating parameters. Redundancy, transparency, and fail-safe mechanisms are essential to ensure resilience when AI systems malfunction or produce incorrect recommendations. At a minimum, every system should include a manual override or "off-switch" that enables operators to assume control seamlessly. If an AI application must be taken offline—whether for maintenance or in response to abnormal behavior—grid operations should transition smoothly to manual or traditional controls. Likewise, when outputs deviate from expected bounds or when the AI model signals low confidence, the system should default to a safe mode, reverting control to human operators or conservative, rule-based backups. This mirrors aviation autopilot systems, which hand control back to pilots during complex or uncertain conditions to preserve safety.

Equally important is designing redundancy into AI systems to avoid single points of failure. A lone AI agent should never hold sole responsibility for critical functions. Using multiple models with diverse architectures, or maintaining traditional control systems as backups, significantly enhances reliability. For instance, some utilities deploy "watchdog" systems that monitor AI health by tracking factors like input data integrity and processing time. These watchdogs issue alerts when anomalies occur, enabling proactive responses before failures escalate. By embedding these engineering safeguards into AI deployments, utilities can ensure that grid operations remain resilient and customer safety is preserved—even in the face of AI-related disruptions.

7.1 Deployment Taxonomy

As electric utilities begin integrating artificial intelligence (AI) capabilities, including large language models (LLMs) and agentic AI systems, into operational workflows, it is necessary to categorize deployment environments based on the location of compute and data, the required level of control, and the associated operational and regulatory risk. The deployment taxonomy presented herein establishes a structured framework to align AI system implementation with established consequence thresholds in electric utility operations. Each deployment category (e.g., public cloud, hybrid, on-premises, edge, federated, and private cloud) presents a distinct profile in terms of latency, data governance, scalability, and infrastructure control. These factors must be evaluated in the context of system criticality, regulatory compliance requirements (e.g., NERC CIP), and the potential operational impact of AI system performance, error, or degradation.

7.1.1 Taxonomy Options for Deployment

The taxonomy options described briefly below are summarized in more detail in Table 12.

Full Public Cloud

- Model Location: Cloud (e.g., OpenAI, Azure OpenAI, AWS Bedrock)

- Data Location: Cloud-native (utility-hosted on cloud or SaaS environments)
- Advantages: Scalable, fastest innovation cycles, access to cutting-edge APIs
- Challenges: NERC-CIP data restrictions, sovereignty risks, network latency, compliance limitations

Hybrid Cloud (Cloud Model + On-Prem/Private Data)

- Model Location: Cloud-hosted model (e.g., OpenAI API)
- Data Location: Utility-owned infrastructure on-prem
- Advantages: Secure data control + advanced models
- Challenges: Complex integration, data tunneling, latency, increased need for cyber governance

Fully On-Prem

- Model Location: On-site (using open-source LLMs or fine-tuned local models)
- Data Location: On-prem data centers or secure utility networks
- Advantages: Total control, compliance-aligned (NERC CIP, IT/OT segmentation)
- Challenges: Infrastructure and maintenance burden, slower to scale, high CapEx

Edge-Based AI (at Substations, DERs, Field Devices)

- Model Location: Embedded in devices (e.g., ARM chips, rugged edge servers)
- Data Location: Localized (on-device, at edge nodes)
- Advantages: Low latency, operates during comms loss, fast decisions at the edge
- Challenges: Limited compute/storage, requires distilled models

Federated / Distributed AI

- Model Location: Multi-location (edge + regional + cloud inference nodes)
- Data Location: Local to each node, coordinated for training or inference
- Advantages: Privacy-preserving model sharing, no centralized data exposure
- Challenges: Federated orchestration complexity, high coordination effort, expanded attack surface and ownership models

Private Cloud / Co-Located Data Center (e.g., Vendor or Utility-Owned)

- Model Location: Hosted in utility-controlled or vendor-managed secure cloud
- Data Location: Private cloud storage or connected utility systems
- Advantages: Can align with NERC/FERC compliance, scalable, cloud-like control
- Challenges: Slower AI toolchain updates, requires close vendor partnerships

Table 12: Deployment taxonomy and data residence example.

Deployment Model	Model Resides	Data Resides	Model Resides	Data Resides	Consequence Alignment
Public Cloud	Cloud	Cloud	Cloud	Cloud	Low consequence
Hybrid Cloud	Cloud	On-Prem	Cloud	On-Prem	Moderate consequence

On-Premises	On-Prem	On-Prem	On-Prem	On-Prem	High-moderate consequence
Edge AI	Field Devices	Local	Field Devices	Local	High consequence
Federated	Mixed	Distributed	Mixed	Distributed	High-real-time
Private Cloud	Private Utility DC	Private Cloud	Private Utility DC	Private Cloud	Medium-to-high

8. Responsible Practice Recommendations

Best-practice recommendations for selecting and deploying AI models in the electric grid context. These practices aim to maximize the value of AI while safeguarding reliability and safety. The items below are the authors' Top 10 recommendations for responsible deployment and should be considered for all applications. Future work includes a practice guide for CIE-informed AI deployments for the grid.[98]

- **Align AI Solutions with Use-Case Criticality:** Evaluate the consequences of AI decisions in each application and choose the level of human oversight accordingly. Not all grid AI applications carry equal risk. For example, an AI that writes customer outage texts can be fully automated, whereas an AI that controls a transmission breaker should be under human supervision. A consequence-driven approach is advised: for high-impact, safety-critical actions, keep humans in the loop or in approval roles; for lower-risk optimizations, more autonomy can be allowed. In practice, categorize potential AI use cases (forecasting, outage restoration, market bidding, etc.) by operational risk and adopt stricter controls on those with high impact. The new "technology" is moving and integrators should make efforts to harmonize with the existing before replacing to bring the human's along.
- **Choose the Right Model Type for the Task and Function:** Base model selection on the nature of the problem, available data, and need for interpretability. For instance, use supervised learning models for well-defined prediction tasks with historical labels (e.g. load forecasting), unsupervised models for anomaly detection and pattern discovery when you lack labels (equipment anomaly detection), and reinforcement learning for sequential decision or control problems (voltage control, dynamic energy dispatch). Within these, consider advanced architectures (like LSTMs for time-series, CNNs for image-based asset inspections, GNNs for networked data) as needed. However, balance complexity with transparency – if a simpler model (like a decision tree or linear model) achieves the goal, it may be preferable for ease of explanation to engineers and regulators. Physics-informed models or hybrid approaches should be favored in domains where physical consistency is paramount (state estimation, stability analysis). Evaluate model performance not just on accuracy but also on whether outputs make sense physically (no negative loads, no frequency beyond limits, etc.).
- **Data Readiness First:** Successful AI deployment begins with robust data infrastructure. Utilities should ensure they have a *single, reliable source of truth* for operational data –

[98] Idaho National Laboratory (INL). (n.d.). *Managing Cyber Risk from Concept to Operation* . Retrieved from Cyber-Informed Engineering: https://inl.gov/national-security/cie/

typically by establishing data lakes or federated data platforms that integrate SCADA, AMI, asset management, outage management, weather and other datasets. Implement strong data governance: assign data stewards, document data lineage, and enforce data quality checks regularly. Before model development, spend ample time on data cleaning and feature engineering, as these often yield bigger improvements than fancy algorithms. It's also recommended to create a cross-functional data team that includes domain experts (power engineers) working with data scientists, so that data is interpreted in context (e.g. knowing that a "0" in a SCADA field might mean sensor offline, not true zero). In terms of volume, leverage historical archives and consider augmenting with simulated data if real data is insufficient for rare events. Additionally, ensure data update pipelines are in place – models might need periodic retraining as new data comes in (to avoid model drift). Having automated pipelines for data ingestion and model retraining (with proper validation gates) will keep the AI solution accurate over time. Diverse data sources can unlock powerful insights but requires tackling compatibility and sharing challenges upfront.

- **Embed Security and Privacy Measures/Policy:** Treat AI models as part of the critical infrastructure – secure them and their data. This includes protecting training data and real-time input data from manipulation (through encryption, authentication, intrusion detection) because compromised data can lead to dangerous outputs. Establish processes to detect anomalous model outputs that might indicate either model issues or malicious interference (for example, a sudden recommendation to shed a large load could be flagged for review). Incorporate adversarial testing during development – expose the model to slightly corrupted or adversarial inputs to see if it behaves robustly. On the privacy front, comply with all regulations: use anonymization or aggregation for customer data in AI models, and restrict access to sensitive data. Data governance policies should clearly define what data can be used for AI and ensure regulatory reporting (some regulators require disclosure of how customer data is used or if an AI makes certain decisions). By proactively addressing cyber and privacy aspects, utilities can avoid setbacks later and build trust with stakeholders that the AI is safe to use.
- **Ensure Explainability and Transparency:** One challenge with AI, especially complex ML models, is their "black box" nature. In a critical domain like the grid, operators and engineers need to understand (at least at a high level) why an AI model is recommending a certain action. So, incorporate explainability tools and requirements in model selection. This could mean using inherently interpretable models where possible or adding explanation layers for deep models. For grid operations, even a simple textual explanation can help, such as "Voltage regulator tap changed by AI because voltage at node X exceeded threshold due to high load." Providing these insights helps humans validate and trust the AI's reasoning. As noted in a NERC report, humans often assume AI is reasoning like they do, which is not true, and this mismatch can be problematic if not addressed. Thus, design the AI system's UI to present relevant context: confidence levels, key factors influencing the decision, and any model uncertainty. Transparency also extends to documenting the model's validation: keep a record of how the model was tested (scenarios, performance metrics) and any limitations discovered. This documentation not only aids internal understanding but also is crucial if regulators or other external parties inquire about the AI's use. In summary**, no black boxes in the**

control room – operators should feel the AI is a tool they comprehend and control, not a mysterious oracle.

- **Use a Phased and Pilot-Based Deployment Strategy:** When introducing AI into operations, start small and simple, then progressively expand. Begin with pilot projects in a controlled environment – for example, run the AI decision support in parallel with human decision-making for some months ("shadow mode") to evaluate its suggestions without risk. Use these pilots to gather performance data, user feedback, and to calibrate the model. Phased deployment gradually increase the AI's autonomy and the scope of its actions as confidence grows. A best practice is to start with human-in-the-loop: let the AI recommend actions and have humans approve them. This phase allows operators to get comfortable and also serves as on-the-job training. Over time – and with clear success criteria met (e.g. the AI's recommendations proved correct 99% of the time in pilot) – then closed-loop automation for specific functions. Maintain fallback: the AI should have an easy manual override or an off-switch in case it behaves unexpectedly. Also plan for contingency operations – if the AI system has to be taken offline (for maintenance or due to an issue), ensure the grid operations can smoothly revert to manual or traditional control.
- **Train and Upskill Personnel (Human-Centric Deployment):** The success of AI in the grid is not just a technical matter, but also a human one. Invest in *operational training programs* so that engineers and operators learn how to use the new AI tools effectively, along with continuing to learn core power system operations, one should not replace the other. Establish human-machine teaming framework: redefine roles so that staff know what the AI will handle and what their own responsibilities are.
- **Implement Continuous Monitoring and Improvement:** Deployment is not the end – monitor the AI system's performance continuously in production. Set up metrics and KPIs to track its impact (e.g. forecasting error reduction, fewer outages, cost savings, etc.). Also monitor for errors or drift: for example, if a load forecasting AI starts to show bias as new electrification trends emerge, trigger a retraining or recalibration. Establish a process for periodic review of the AI's decisions by experts, at least in the early stages, to catch any odd behavior. This can be formalized as an *AI governance committee* within the utility that reviews AI outcomes and any incidents. Integrate real-time validation where feasible – one recommendation is to use parallel models or rules as a check on the AI's recommendations. For instance, if the AI suggests a power flow that violates a constraint, a simple rule-based checker could catch it before execution. Over time, these validations can be relaxed as trust increases, but it's good to have them as a safety net. Keep models up to date with the latest data and system configuration (for example, if a new solar farm is added, ensure the forecasting model knows about it). Moreover, plan for model maintenance: machine learning models can degrade over time (data drift, changes in usage patterns), so schedule periodic retraining or re-validation (perhaps every year or after major system changes). Document any model updates and ensure operators are informed of changes in behavior or new features. In essence, treat the AI system as a living part of the grid ecosystem that requires ongoing care, not a one-off install.
- **Leverage Industry Lessons and Collaborate:** The electric utility community is actively sharing lessons learned from AI deployments. Engaging with industry groups (e.g., EPRI, IEEE smart grid groups, NERC working groups) and learning from case studies can greatly benefit a utility's AI journey. Many recent whitepapers and reports (like NERC's

2024 AI in operations white paper, DOE's AI for grid research, etc.) highlight common pitfalls and successful strategies. For example, one RAND study pointed out that introducing AI widely in grid operations can have unexpected systemic effects if not coordinated (they found that if some operators have AI assistance and others do not, it can lead to misaligned decisions). Such insights suggest that industry-wide standards and best practices may be needed for smooth integration. Utilities should advocate for and adopt emerging standards for AI in critical infrastructure – whether it be in model validation, data exchange formats, or reliability testing protocols. Additionally, sharing non-sensitive data and results with research partners or consortia can accelerate innovation (for instance, participating in open datasets or challenges for grid AI). Many utilities have found value in partnering with universities or national labs to develop AI models tailored to their needs, benefiting from cutting-edge research. By staying connected to the broader community, utility leaders can ensure they are following the most up-to-date and proven practices, rather than reinventing the wheel.

- **Focus on Resilience and Fail-safes:** Finally, users should prepare for the scenario where AI might fail or yield an incorrect recommendation. Build resilience into the AI deployment. This means having fail-safe mechanisms: if the AI output is outside normal bounds or the system detects the AI is not confident, it should default to a safe mode (e.g., revert control to human or to a conservative rule-based backup). It's analogous to autopilot systems handing control back to pilots in complex situations – the transition must be smooth. Conduct "what-if" analyses for worst-case AI failures (like what if the AI mis-classifies a major fault as normal? what if it fails to forecast a peak load?) and have mitigation plans (alarms, redundant checks, etc.) for those. Additionally, consider the redundancy of AI systems: a single AI agent should not be a single point of failure for critical control. Using multiple models (with diversity in design) or keeping the traditional control as a backup initially can improve reliability. Some utilities implement a "watchdog" that monitors the AI's health (looking at factors like input data health, processing time, etc.) and raises an alert if something seems off. By engineering these safety features, the grid can tolerate an AI issue without impacting customers or safety – which is absolutely essential given the high stakes.

9. Conclusion

In conclusion, the selection and deployment of AI models in the electric grid should be approached with a blend of technical rigor and operational pragmatism. By focusing on models well-suited to the task domains (detection, prediction, optimization), grounding them in power system physics, ensuring robust data foundations, and carefully managing the human/automation interface, utilities can harness AI's capabilities to create a smarter, more resilient grid. Artificial intelligence could in the next 5 to 10 years become a core component of electric grid business planning and operations. This paper outlined how AI is being proposed or piloted to support functions across detection, prediction, optimization, and customer engagement, while also identifying the risks associated with its adoption. These risks include cybersecurity vulnerabilities, data quality and governance challenges, model drift, regulatory gaps, and potential reductions in operator expertise.

A consequence-based framework provides a systematic method to evaluate AI applications. By assessing adoption against operational criticality, performance attributes, and the consequences of non-adoption, utilities can align AI deployment with their risk tolerance. Effective integration requires models that reflect power system physics, reliable and well-governed data, and clear processes for monitoring and validation. Engineering controls such as redundancy, fail-safes, and manual overrides remain necessary to ensure continuity of safe operation.

AI deployment in the electric sector should be treated as an incremental, managed process with responsible and appropriate adoption. Phased implementation, workforce training, and engagement with regulatory and industry standards will be important for achieving reliable outcomes. By applying consequence-based assessment and cyber-informed engineering practices, utilities can incorporate AI in ways that improve operational effectiveness while maintaining safety, security, and compliance.

Appendix A: Consequence Table Heatmap

Table 13: Heatmap example of Top 10 OWASP AI risks and vulnerabilities by grid application/ use case.

Note: While this consequence ranking offers a structured lens for evaluating AI vulnerabilities, the actual impact of a given issue can vary significantly. Factors such as model architecture, deployment environment, operational dependencies, and how the AI is integrated into decision-making all influence the real-world severity. A use case rated as "moderate" in one context may be "high" in another.

		OWASP Top 10 LLM/AI Vulnerabilities (2023)										
Grid Application / Use Case	**Possible Risk(s)**	ML1: **Input Manipulation**	ML2: **Data Poisoning**	ML3: **Model Inversion**	ML4: **Membership Inference**	ML05: **Model Theft**	ML06: **AI Supply Chain Attack**	ML07: **Transfer Learning Attack**	ML8: **Model Skewing**	ML09: **Output Integrity**	ML10: **Model Poisoning**	**Overall Ranking of Consequence from AI Issues**
DERMS	• Incorrect AI predictions causing over/under-dispatch • Adversarial inputs (e.g., spoofed data) leading to cascading failures • Loss of operator trust due to black-box decision-making											
Predictive Maintenance& Asset Health Monitoring	• False positives causing unnecessary maintenance • Missed failures leading to unexpected outages • Data privacy concerns from vendor-managed AI • Long-term incorrect expenditures											
Grid Forecasting (Load, Generation, Weather)	• Model drift causing forecasts to degrade over time											

Model Validation & Data Lake	• Erroneous model acceptance or rejection due to mislabeled or poor-quality training data • Spurious correlations or reinforcement of historical biases • Data governance issues (e.g., inconsistent schema, time alignment, lack of metadata) • Degraded model performance and traceability											
Power Flow & Interconnection Analyses	• Misrepresentation of grid physics under atypical or stressed conditions • Oversimplification of complex network interactions • Regulatory non-compliance due to lack of traceability • Over-reliance on AI outputs without rigorous validation											
Autonomous Grid Control (Real-Time Dynamic Protection and Remediation)	• Latency issues during rapid grid events • Difficulty in post-event analysis due to black-box models • Synchronization risks between AI controllers											
Fault Location	• Inaccurate predictions due to poor data quality • Failure to generalize to novel fault conditions • Operator over-reliance on AI outputs • Cybersecurity vulnerabilities • Lack of explainability for regulatory compliance											

Outage Management System (OMS)	• Inaccurate outage detection or misclassification of events • Delayed restoration due to incorrect prioritization by AI • Over-reliance on AI recommendations during emergencies • Data integration issues across legacy and modern systems • Lack of explainability in outage cause analysis • Vulnerability to cyberattacks targeting outage response systems											
Energy Market Optimization (AI for Trading & Bidding Strategy)	• Synchronization risks between AI-driven controllers											
Cybersecurity Operations	• Market manipulation or regulatory compliance violations • Misalignment between AI bidding strategies and grid reliability needs • Overconcentration of AI-driven strategies											
Customer Load Management & Flexibility	• High false positive rate leading to alert fatigue • AI missing novel attack patterns • Vulnerability to adversarial evasion											

Appendix B: Consequence Table Descriptions

Table 14: Consequence descriptions per use case.

Use Case	ML1 Input Manipulation	ML2 Data Poisoning	ML3 Model Inversion	ML4 Membership Inference	ML05 Model Theft	ML06 AI Supply Chain Attack	ML07 Transfer Learning Attack	ML08 Model Skewing	ML09 Output Integrity	ML10 Model Poisoning
DERMS	Spoofed DER telemetry (e.g., fake voltage/frequency data) causes AI to miscalculate dispatch, leading to overloading or underutilization of assets.	Malicious training data skews DERMS optimization, prioritizing unreliable or adversary-controlled DERs.	Attackers extract DER operational patterns, revealing sensitive data like load curves or customer behavior.	Inference that specific customer or DER behavior was used during training, exposing private load profiles.	An adversary replicates the DERMS logic to predict utility dispatch in real-time, enabling manipulation or gaming.	Compromised DERMS model or library introduces logic that preferentially dispatches unstable or adversary nodes.	Imported DER optimization model includes a hidden rule that under-dispatches critical feeders under stress.	Operator overrides or feedback loops skew training, reinforcing suboptimal DERMS decisions over time.	Dispatch signals are modified in transit, e.g., switching DERs off when grid needs support.	Gradual corruption of DERMS internal weights leads to misprioritizing dispatch during high-demand periods..
Predictive Maintenance & Asset Health Monitoring	Injected false sensor data (temperature, vibration) causes premature or missed maintenance.	Training data modified to reflect misleading asset failure patterns, skewing prioritization.	Reconstruction of asset operational profiles could expose proprietary equipment stress data.	Disclosure of whether a specific asset or substation failure history was part of training.	Attackers clone predictive logic, using it to guess upcoming outages or equipment replacements.	Third-party maintenance model is embedded with code to suppress alerts on select failures.	Pretrained health model contains logic to ignore transformer degradation beyond a set threshold.	Maintenance team interventions bias model to ignore outliers, degrading predictive performance.	Maintenance recommendations are altered post-model to defer necessary repairs.	Weights are tampered to reduce model sensitivity to critical failure indicators like thermal spikes.
Grid Forecasting (Load, Generation, Weather)	Spoofed weather or generation data skews forecasts, leading to reserve shortfalls.	Training sets altered with fake load/generation patterns degrade future predictions.	Reconstructs load or generation patterns of sensitive sites, risking profiling or targeting.	Reveals which weather or demand events (e.g., heatwaves) were used to train forecasting models.	Forecasting models are stolen and used to game market positions against the utility.	Corrupted forecasting module subtly shifts predictions to destabilize dispatch planning.	Poisoned pretrained forecast model suppresses peak event predictions when reused.	Human edits to forecasts fed back into training bias future predictions under similar conditions.	Forecast outputs are modified before use, misleading operators and reserve planners.	Weights are corrupted to flatten high-risk/variance predictions, masking extremes like cold snaps.
Model Validation and Data Lake (central point of truth etc)	Fake data enters validation set, producing incorrect baseline performance results.	Intentional corruption of training data within the lake leads to systemic model flaws.	Data lake models reveal sensitive system-wide telemetry like topology or	Attackers deduce whether specific fault events or datasets were	Validation engines or centralized models are copied, providing insights into	Compromised data lake pipeline ingests poisoned models or spreads	Pretrained validation models from vendors include logic to overlook performance	Human corrections to model outputs fed back into the lake bias all future	Validation metrics are altered in dashboards to conceal model underperformance.	Training metadata is subtly manipulated to consistently approve

			outage response.	stored in central lake.	system thresholds.	scripts across departments.	dips/set up wrong upgrades.	training datasets.		underperforming models.
Power Flow & Interconnection Analyses	Manipulated network topology inputs mislead power flow analysis, leading to unsafe conditions.	Training with fake interconnection scenarios results in flawed contingency analysis.	Extracts sensitive grid topology or interconnection queue data from the model's responses..	Discovers if specific project interconnection studies were used to train ML models.	Reconstructed power flow models can be used to simulate and evaluate attacker playbooks against the grid	network analysis tools misroute or suppress warnings for risky interconnects	Pretrained models repurposed for interconnection ignore voltage or thermal constraints	Mislabeling of fault or overload cases during training reduces future detection accuracy.	Model-calculated limits or flows are altered in route to engineers, causing unsafe approvals.	Contaminated model weights underrepresent edge-case violations, risking unstable interconnections.
Autonomous Grid Control (Real-Time Dynamic Protection and Remediation)	Compromised frequency or phase angle data triggers false grid separation or reclosure.	Distorted fault event data leads to incorrect automated grid recovery logic.	Attacker reconstructs dynamic control logic for critical substations and breakers.	Infers use of specific transient or voltage collapse events in training, exposing grid weak points.	Stolen real-time control model allows adversary simulation of grid behavior under disturbance.	Backdoored controller logic disables key protection during defined conditions (e.g., high DER penetration).	Embedded logic in inherited model suppresses protective action	Operator feedback to dampen transient events shifts control away from robust configurations.	Control setpoints modified midstream, sending unstable signals to voltage regulators or relays.	Small, persistent changes in training degrade grid stabilization ability over time.
Fault Location (precursor to grid control)	Injected harmonics or noise spoof true location, delaying restoration.	Labeled historical fault data is corrupted to favor incorrect fault paths.	Attacker deduces network impedance and feeder topology from model outputs.	Confirms which line segments had faults in training, exposing high-risk infrastructure.	Reconstructed model can be used to test stealthy grid disruptions offline.	Model routes fault detection away from true location under overload.	Residual logic in reused models ignores faults behind specific device types.	Field technician error feedback retrains model to ignore certain anomalies.	Location results are altered en route to outage management, causing misallocated crews.	Internal model logic gradually diverges, misplacing faults systematically over time.
OMS	False outage calls or sensor inputs trick OMS into dispatching to wrong areas.	Training with fake or improperly labeled outage records degrades event triage.	Exposes historical restoration decisions and prioritization logic for critical loads.	Confirms which customer complaints or outages were included in OMS training.	Cloned OMS model can be used to mimic utility response for malicious load shifting.	Backdoored OMS module removes outage visibility under defined loading.	Legacy logic disables alerts if restoration exceeds X hours under certain temperatures.	Repeated manual dispatcher changes bias AI to deprioritize rural regions.	Crew dispatch orders are tampered before field receipt, delaying recovery.	Continuous retraining erodes fast restoration recommendations during peak outages.
Energy Market Optimization (AI for Trading & Bidding Strategy)	Fake prices or resource forecasts distort real-time bidding outputs.	Falsified past market actions lead to risky bidding strategies.	Competitors deduce pricing strategy and supply curve from query responses.	Confirms participation of specific generators or bids in training history.	Stolen AI model used to undercut utility bids or cause financial imbalance.	Corrupted optimizer skews bids toward underpriced or infeasible blocks.	Leftover bidding logic disables high price responses during scarcity.	Manual overrides to AI bids retrain system to avoid profitable dispatch.	Final bid stack is altered during transmission, creating exposure to market penalties.	Model begins self-sabotaging high-return bids after long horizon exposure.

Cybersecurity Ops	Spoofed occupancy or thermostat data leads to disjointed demand response events.	Training with falsified comfort or usage data weakens DR effectiveness.	Infers appliance schedules or home occupancy patterns via smart control models.	Reveals which homes or behaviors trained the DR logic risking privacy.	Stolen load model lets adversaries simulate control sequences for sabotage or exploitation.	Compromised model favors third-party devices or delays DR for select users.	Legacy HVAC logic suppresses override when load shed is required.	Biased homeowner feedback retrains model away from effective load control.	Thermostat control signals tampered, reducing event alignment with grid needs.	DR response curves decay over time due to corrupted retraining episodes.
Customer Load Management & Flexibility (AI-Powered Demand Response)	Adversarial inputs (e.g., spoofed voltage/frequency signals) could mislead AI systems, causing false dispatches or incorrect protection settings.	Compromised training data from sensors or historical logs leads to poor asset health predictions, skewed forecasting, or biased market strategies.	Attackers may extract sensitive operational data (e.g., customer loads or SCADA patterns), risking privacy and system insights.	An attacker could infer which feeders or substations were involved in past grid events, leaking grid behavior or incident response data.	Stolen models (e.g., DERMS optimization logic) could be replicated or exploited by competitors or adversaries to predict system behavior.	Pretrained ML libraries or packages embedded with malicious payloads could corrupt dispatch or monitoring systems at scale.	Poisoned pre-trained models reused in utilities could embed hidden logic, only triggering erroneous actions under certain conditions.	MLOps feedback loops from operator override or manipulated event logs may result in performance drift and decreased grid reliability.	If AI outputs (e.g., dispatch orders or alerts) are modified in transit, operators may act on false recommendations, risking safety.	Deliberate parameter corruption at training or deployment alters behavior over time—e.g., predicting normal operations during faults.