Unlocking AI's Potential for the Energy Transition Through Open Source

January 2025

LF ENERGY

Table of Contents

Executive summary

As the global energy sector undergoes a critical transformation driven by decarbonization, digitalization, and decentralization, artificial intelligence (AI) has emerged as a key enabler for optimizing energy systems. This whitepaper demonstrates how open source innovation is essential to accelerate AI adoption in the energy industry and deliver on its promises.

Open source is essential for delivering AI solutions in power systems

AI holds great promise and potential to help transform power systems and optimize their operation at a time of unprecedented change and challenges. However, to deliver on these promises, a different collaboration model is necessary.

This is evident from the experience and trends in other sectors, where open source dominates AI, and from a closer examination of the challenges in applying AI to power systems:

- **Speed and scale:** Vendors will need to work with utilities and system operators as well as research laboratories and academia to develop the AI solutions their customers need. They need access to data and operational power system expertise from utilities and system operators, which will have to take a leadership role in driving AI innovation. Compared to the traditional approach of consortia, non-disclosure agreements, and closed solutions, open source is far superior for effectively structuring such digital collaboration and scaling it up while still preserving a sustainable business model and revenue.
- **Trust and compliance:** Critical applications of AI in power systems will require transparent, secure, and auditable models, which utilities and operators will most likely run and control themselves. Most of the data will remain behind firewalls to comply with privacy and critical infrastructure protection (CIP) laws and regulations. Open source fosters transparency, security, and trust, which will streamline meeting these requirements and those of AI regulations (for example, "high-risk systems" in the E.U. AI Act).
- **Standardization:** Open source tools and community specifications can drive forward standardization, interoperability and data exchange at a much faster pace, which will be instrumental in enabling and deploying AI.

If the energy industry does not apply the best practices from other sectors and leverage open source strategically, it will not deliver on the promises and potential of AI with speed and scale. Open source is the key practice that turns hype into reality.

AI readiness framework

Energy stakeholders must adopt a strategic approach to AI readiness by:

- Establishing robust data governance to enable AI innovation and deployment
- Investing in digital twins and data platforms based on open source shared components to support AI initiatives
- Supporting open, realistic datasets to fuel AI model development with third-party innovators and researchers
- Promoting AI literacy in the organization and workforce to help power systems make the most of AI tools and technologies and support AI experts in navigating and understanding energy-specific knowledge and challenges

LF Energy is building open source solutions for key AI use cases

LF Energy has launched or is setting up projects for six priority AI application areas in power systems:

- **Forecasting:** Enhancing short-term load forecasting
- **Accelerated Simulation & Optimization:** Speeding up operational decision making by leveraging AI in simulation and optimization tools
- **Asset Management:** Optimizing maintenance and reducing failures by supplementing predictive analytics with AI models
- **Long-Term Planning:** Developing optimal decision-making tools for long-term infrastructure investment while tackling uncertainty and the changing climate
- **AI Assistants:** Supporting AI-powered decision making in operations
- **Grid-Interactive Smart Communities:** Enabling flexible, decentralized energy management at the grid edge

Other projects focus on horizontal needs such as foundational models, digital twins, and synthetic smart meter data.

LF Energy is the leading collaborative framework for open source AI

The Linux Foundation is the world's leading platform for large-scale, industry-grade open source projects in AI. LF Energy is bringing this experience and knowledge to support open source AI initiatives in the energy sector. We're eager and ready to partner with utilities, system operators, vendors, academia, research labs, and other collaborative organizations to scale up their open source efforts.

Introduction

This white paper explores the intersection of three major trends shaping the future of energy and technology:

- The worldwide efforts to decarbonize energy systems
- The recent spread of AI technologies
- The adoption of open source practices in all technology sectors

While synergies between each pair of these trends are clear, a comprehensive discussion that encompasses all three has been missing. This white paper seeks to address that gap.

It serves as a guide for energy stakeholders navigating these transformative shifts, helping organizations become AI-ready. These organizations face intricate challenges, such as sharing and managing sensitive energy data to drive AI innovation, fostering collaboration between AI and power system experts, and ensuring trust, security, and explainability when deploying AI in critical processes. We will explore how embracing open source and open innovation serves as a key enabler for overcoming these hurdles.

Successfully developing and deploying AI for energy systems will require large-scale, collaborative efforts rooted in open source projects. This whitepaper identifies priority areas for such projects and proposes specific actions — either by leveraging existing initiatives or establishing new ones where necessary.

What is LF Energy?

LF Energy is a community for collaboration on open industrial-grade technology platforms and agile specifications to accelerate the digital energy transition. It is a member-funded non-profit that is part of the larger Linux Foundation, the largest open source steward in the world. LF Energy has an ecosystem of nearly 35 open source projects, with the support of over 30 members. By providing a neutral home for utilities, vendors, and others to collaborate on foundational digital technology and specifications, the energy industry can confidently work together on production-ready software and trust its sustainability for decades to come.

Open source is critical to AI adoption in the energy industry

AI is transforming industries by enabling machines to perform complex tasks such as understanding natural language and generating creative content. Open source AI has emerged as a powerful force, driving innovation and democratizing access to cutting-edge technology.

For example, LLaMA, developed by Meta, is a family of AI models advancing natural language processing. Meta's decision to release LLaMA models for research has spurred global AI development. Hugging Face serves as a central hub for open source AI, hosting models, datasets, and tools, fostering collaboration and innovation.

The Model Openness Framework[1] defines what makes an AI model genuinely open by assessing its completeness and openness. Completeness ensures that essential components such as training data, model parameters, and evaluation methods are available. Openness requires releasing these components under licenses that allow sharing and modification.

Open source AI fosters collaboration and accelerates innovations across all sectors and industries. The open source model dominates[2] AI and provides exclusive benefits to AI and data, such as fairness (methods to detect and mitigate bias in datasets and models), robustness (methods to detect alterations and tempering with datasets and models), explainability (methods to enhance persona's or role's ability to understand and interpret AI model outcomes, decisions, and recommendations), and lineage (methods to ensure the provenance of datasets and AI models). Open source data-specific licenses also make data freely accessible for use without mechanisms of control, and open source helps with data governability (a governance structure and tools to clean, sort, tag, and govern data and datasets). The characteristics of the open source model make it ideal for cooperating on enabling technologies regardless of domain or industry.

This key role of open source in AI extends to the energy industry. Yet, as the sector faces an unprecedented era of transformation, energy stakeholders tend to look at their own specific challenges. Therefore, it is useful to focus on how open source and open innovation are poised to accelerate the deployment and use of AI, specifically in the energy sector.

Open source collaboration and innovation enable speed and scale

The current transformation in energy systems is systemic and far reaching — while its pace varies between countries or systems, it is everywhere and touches almost every component, process, or regulation at once. Whereas power used to flow from large-scale dispatchable generation units to passive end users through centrally managed transmission and distribution networks, modern systems are based on non-dispatchable carbon-free distributed generation and consumer-centric flexible technologies on the demand side, with bidirectional flows and increased electrification as means to decarbonize overall energy use. These distributed systems leverage advanced sensors and controls, digital edge and IoT solutions, and now, AI to handle the associated complexity and variability, maintain high reliability even under extreme conditions, and optimize overall system costs to ensure affordability and equitable access to electricity.

To tackle these issues, innovation and collaboration are essential, and they are key to speed and scale in an economical and equitable way. As everyone faces similar challenges and issues, developing solutions together is a must. The energy industry has always relied on shared innovation and collaboration to achieve success. Our electric grids, the "most complex machine ever built," are the result of the collaborative efforts of many. As we invent and deploy AI as part of this massive transformation, this is more important than ever.

Open source is the modern approach to collaboration[3] in digital and AI:

- While traditional collaboration projects relied on consortia, open source is a frictionless and scalable way to set up a collaborative effort with generic, well-tested intellectual property and governance frameworks.
- Open source allows stakeholders and innovators to work directly together on software and AI model implementations without having to align first on specifications or standards before writing code. In addition to speeding up the process, this also reduces the inherent ambiguity in implementing specifications and standards since implementation is built-in along the way, making it accessible for everyone to use and modify consistently.
- Interoperable systems are created through well-documented data standards and interfaces, benchmarks and test cases, conformance programs, and cross-pollination between projects.
- Sustainable, future-proof, and universal solutions are built as any stakeholder can propose a modification to software or AI models, and there is a strong incentive to integrate such proposed modifications as part of a consistent, high-quality solution to maintain the largest user and contributor base possible.
- Vendors and innovators collaborate on shared foundational solutions, upon which they can build an ecosystem of commercial products and services.
- Competition regulation is respected, which is essential for utilities as many are public companies or subject to public procurement laws and

regulations. Through open source, utilities can collaborate with vendors without the risk of barring them from future public tender offers, whereas bilateral collaboration may appear to give them an unfair advantage.

Stakeholders can leverage such benefits whether they develop software or AI models in house or procure them through vendors. In the first case, stakeholders can collaborate and share development and maintenance costs while benefiting from innovation and use cases from others and keeping the solution targeted to their needs. In the second case, stakeholders can specify solutions based on open source projects, in which case they will reap most of the benefits of the open source ecosystem without investing directly in the required expertise and infrastructure to deploy and maintain their specific implementation. In both cases, stakeholders can influence the course and priorities of the projects in accordance with their business strategy and needs.

For regulators and policymakers, open source helps with the adoption and mutualization of efforts in a fragmented energy industry. Many energy stakeholders will not have the resources to invest in AI innovation, and even if they do, there will be massive duplication of effort, a waste of rate and taxpayer money at a time when affordable energy is a key priority for communities and businesses. Centralized R&D initiatives, whether driven by the government, large utilities, or vendors, may not immediately address all the diverse needs and use cases of end users and will give rise to competing products that require more explicit standards and regulations to resolve interoperability issues. In addition, in the case of large or foundation models, no single entity can build them by themselves, as a large quantity of diverse data, power systems expertise, AI expertise, and compute resources is needed. Open source and open innovation offer an attractive alternative in which all stakeholders can make their voices and needs heard, large entities can invest and recover their investments through innovative products and solutions, and an ecosystem of third-party innovators and startups can flourish. When developing digital solutions based on AI, this is the most effective path from research and innovation to market and implementation.

Open source fosters transparency and trust in AI applications

Electric systems are critical to our modern way of life, from day-to-day critical applications to supporting our modern global economy. Like any other innovative solution that has transformed energy systems, deploying AI must happen in a way that keeps the lights on.[4] This hinges on trustworthiness, which "in the context of AI adoption by electrical power utilities refers to the system's ability to ensure reliability and performance, to inspire confidence and transparency among stakeholders in a context of energy transition with more volatility and an evolving system complexity."

Open source and innovation foster transparency, which is the pillar of both trust and security. Both stakeholders and independent organizations can audit open solutions, exposing and quantifying risks, biases, and vulnerabilities and contributing to minimizing them through shared innovation. This is relevant for real-time AI applications ranging from protection to control systems and tools used in operations but also extends to other non-real-time applications such as asset management and long-term planning studies.

Protection systems must operate reliably and predictably, as they underpin the safe and secure operation of the system at timescales that are far too rapid (typically less than one second) for operators to intervene. Compared to proprietary black boxes, open source AI solutions are easier to certify across any range of parameters and operational conditions, and their behavior is simpler to model as part of more complex studies.

Control systems and automatons can optimize the response of various devices and protection systems based on local measurements and parameters. They typically operate more slowly than those systems but still on timescales between one second and one minute, which are usually too fast for operators to process inputs and react in a relevant way. Such automatic systems have been used for frequency or voltage control and will become more sophisticated through the application of AI. In a way, where human operators used to operate the grid directly, the operators of tomorrow will control those automatic systems that will, in turn, operate the grid, much like a pilot controls the plane through an autopilot system rather than directly. This will also be true as organizations deploy intelligent devices in home and building environments.

Transparency is essential in this context to understand how these systems will respond, affect the behavior of underlying protection systems and devices, and interact with each other under a wide variety of conditions and parameters. Without transparency, there is a huge risk of adverse behavior and cascading effects that operator action cannot remedy at a much larger scale compared to established N-1 frameworks. As with protection systems, accurately representing the behavior of these systems in larger system studies is essential, with an adequate compromise between fidelity and computational burden— this is already a challenge for experts when they have transparent access to the relevant information, and thus almost inconceivable when dealing with a black box. Organizations cannot properly assess or address vulnerabilities and exposure to cyberattacks without a detailed understanding of the system's behavior. Without open source solutions, this will constitute a barrier against AI adoption, as stakeholders will not be willing to take such risks or severely constrain the specifications of such systems to limit them. As AI applications in this layer are an essential component of grid modernization efforts, open source is a critical enabler in this context.

For systems operating in timescales slower than a minute, interaction with operators and human decision processes becomes relevant. Explainability is a key element here for the adoption of AI, as is the case in other industry sectors involving critical infrastructure or decision making, such as transportation or telecommunication. This is very much an area for research and innovation in AI, involving academic and industry experts, social and cognitive science, and specialized knowledge of the underlying systems. Cross-industry open source collaboration will be essential in this context to speed up adoption while addressing specific industry needs. The energy sector is already relying heavily on optimization and simulation tools for its operational processes, and the adoption of new tools by operators has always required careful design and change management. Open source projects will allow vendors and operators to collaborate, leveraging generic solutions and experts from academia and research organizations. This will also require data sharing and proper modeling across organizational silos, which has always been difficult — open source collaboration and the use of open data licenses can help circumvent these barriers as well.

This extends to software used outside of the control room for asset management, system planning, and investment decisions. Here, more time is available to challenge the outputs of AI models and their design and training. However, the software helps make or inform decisions involving enormous financial stakes, liability, or exposure to risk. Ensuring that the organization understands the AI models it uses in this context, their training data, hyperparameters, and biases will therefore be critical for decision makers. Compared to in-house or proprietary black box solutions, open source is the best compromise — it allows using well-established and adopted "state of the art" tools and performing a specific strategy-driven implementation while having transparent and auditable access to its source code, training data, and parameters.

A similar argument applies to AI-powered tools meant to inform long-term energy policy and public debate. Regulators and policymakers can leverage open source to make expert studies and recommendations accessible to all stakeholders, with the possibility of reproducing, auditing, and challenging the results, methodologies, and assumptions.

Open source strengthens privacy, cybersecurity, and ethical safeguards

Openness and transparency also allow stakeholders to assess AI models in terms of privacy, cybersecurity, and ethical concerns. Machine learning (ML) relies on data for training and inference, and for energy applications, such concerns are common and critical for industry adoption.

Many granular datasets or data sources in energy involve some degree of privacy or confidentiality, from meter data potentially revealing individual behavior and habits or strategic business information to detailed models of components and systems with underlying intellectual property and trade secrets. Open source allows code to be fully transparent, meaning that the public, security experts, and stakeholders can review and audit any algorithms handling sensitive energy-related data (such as customer usage patterns or infrastructure information). This level of scrutiny can ensure that the AI respects data privacy requirements and does not unintentionally or intentionally expose sensitive information.

Open source projects benefit from a wide array of contributors who can quickly identify and patch security vulnerabilities. This helps protect AI systems used in energy applications from evolving cyber threats such as data breaches, ransomware attacks, or system manipulation. Open source AI models and tools can be subject to rigorous peer review, where cybersecurity experts around the world audit the code for weaknesses. This community-based scrutiny reduces the chances of malicious code or backdoors being introduced into AI systems, which are particularly critical for protecting energy infrastructure. Collaborative open source platforms provide real-time information on emerging cyber threats and responses, creating a collective defense mechanism. Thus, energy companies can stay ahead of potential cybersecurity challenges.

Reviewing open source AI models for bias or unethical behavior ensures accountability in energy decision-making processes. With full access to model architectures and training data, it is easier for independent researchers or watchdogs to identify bias or discrimination in AI applications, such as those used for grid management or energy distribution. Open source projects often involve diverse contributors from different sectors (academia, industry, civil society), which can lead to the development of ethical guidelines or best practices that prioritize fairness, sustainability, and human rights in AI. This is particularly crucial for energy systems, which have broad societal impacts.

Open frameworks help comply with AI regulations (e.g., the E.U. AI Act)

The E.U. AI Act, enacted in mid-2024, is one of the first comprehensive frameworks for AI regulation, and it is expected to have global influence beyond the E.U. Much like the General Data Protection Regulation set a precedent for privacy laws worldwide, the AI Act is poised to shape AI regulatory frameworks, especially for industries such as energy that operate across borders. Open source AI models can serve as a baseline for companies that must comply with different regulations across various jurisdictions. By leveraging open source, energy companies can more easily adapt to future regulatory requirements that are likely to emerge outside the E.U.

The AI Act uses a risk-based approach, classifying AI applications into prohibited, high-risk, and minimal-risk categories. Many energy-related AI models fall under the high-risk category because they relate to critical infrastructure, such as managing power grids or optimizing energy distribution. Open source models offer flexibility, allowing developers to tailor systems for each risk level, incorporating specific compliance layers that align with the classification.

For high-risk systems, the AI Act mandates strict requirements in areas such as risk management, data governance, documentation, record keeping, and transparency. Open source AI tools can help energy companies meet these requirements by integrating automated processes for managing risk, securing data, and ensuring full traceability. For example, open source projects can develop risk assessment dashboards that monitor AI systems continuously, flagging potential compliance issues in real time.

When it comes to documentation, record keeping, and transparency, open source AI models provide an inherent advantage. Because the code is open for review, these systems naturally offer transparency, allowing regulators and stakeholders to audit the algorithms and understand how decisions are made. Automated documentation tools can also integrate with open source frameworks, helping energy companies track system decisions and ensure that they comply with the Act's requirements for traceability and transparency.

Open source projects can collaborate on shared compliance frameworks, where developers, industry stakeholders, and regulators work together to create standardized tools for managing cybersecurity, data governance, and model accuracy. These frameworks are adaptable to specific use cases in energy, such as grid management or predictive maintenance, ensuring that high-risk systems are not only compliant but also secure and robust.

By emphasizing the open and collaborative development of compliance tools, open source platforms enable energy companies to build AI systems aligned with the AI Act's requirements from the ground up. This approach not only facilitates compliance but also promotes continuous improvement, as the community can contribute updates, patches, and new features that respond to emerging regulatory changes or security threats.

Open source helps attract, develop, and focus talent

AI experts are in high demand, and for many energy companies, finding specialists in AI and ML can be a significant challenge. The steep learning curve involved in applying AI to complex energy systems further complicates this issue. However, by using AI solutions developed across industries or by opening energy-specific challenges to the broader AI/ML community, energy companies can tap into the surge of innovation in AI and ML.

An open source approach helps build a global ecosystem of AI and energy experts where independent researchers, small organizations, and large utilities can collaborate. This collaborative environment fosters innovation and accelerates problem solving for the most critical issues in the energy sector, such as optimizing grid performance, integrating renewable energy, and enhancing energy efficiency. For smaller organizations, the opportunity to work alongside larger vendors and utilities is invaluable, allowing them to contribute cutting-edge solutions while benefiting from the shared expertise of a diverse community.

By becoming part of this ecosystem, energy companies can leverage the contributions of others, reducing the time and resources needed to develop proprietary AI solutions in house. This access to a broader talent pool adds strategic value, enabling organizations to stay competitive in the rapidly evolving energy landscape. Open source AI projects also offer a collaborative platform where companies can share challenges, co-develop solutions, and ensure that the AI systems they use are not only effective but also compliant with emerging regulations and standards.

AI readiness for energy stakeholders

After demonstrating that open source is a critical enabler for AI adoption in energy, we now turn to practical guidance for energy stakeholders. How can organizations — more specifically, you as a key decision maker and AI champion within your organization — prepare to make the most of AI?

From improving data access and governance to leveraging digital twins, readiness in AI involves several strategic steps that energy stakeholders — you — must consider.

From data access and governance to digital twins

Put your "data" house in order: In the age of AI, data quality, governance, and digital twin investments are more critical than ever.

At the core of AI's potential lies data — particularly, the ability to collect, manage, and make use of high-quality data. The current boom in AI, especially in ML, hinges on large volumes of relevant data and on being able to cross-analyze different data sources to derive new insights or improve decision making. For AI readiness, this means addressing long-standing issues surrounding data access, governance, and quality.

AI applications demand not only data quantity but also data quality. Energy stakeholders must prioritize establishing robust data governance frameworks that address the fragmented and siloed nature of energy data. This includes investing in secure, scalable platforms that allow for real-time access to operational data while maintaining compliance with privacy and security regulations. Efficient data governance ensures data flows freely across your organization, eliminating bottlenecks and enhancing decision making. Investing in a solid data governance program and data platforms will be instrumental in making your organization AI-ready.

There is a strong synergy between AI and the push to create digital twins in the energy sector. A digital twin — a virtual model of a physical energy system — can offer rich, real-time data streams that fuel AI models. The integration of AI with digital twins enables predictive and condition-based maintenance, real-time optimization, and scenario planning. Stakeholders should champion the development of digital twins as part of their AI strategy, as these models offer a dynamic testing ground for AI algorithms and help bridge the gap between operational data and actionable insights.

AI innovation in energy requires open, realistic datasets

Access to data fuels AI innovation — break the data barriers with realistic open benchmarks.

Energy data is siloed between actors or even within organizations not only due to the cost of establishing efficient and reliable data platforms but also because of the confidential and sensitive nature of the underlying data. The latter requirements derive from privacy and CIP laws and regulations, which explicitly prevent or set limits to sharing data with third parties. Many utilities will also limit the sharing of sensitive data internally to limit the risk of leaks or unauthorized access.

This creates a situation where many innovators and researchers cannot easily gain access to real data sources, and even if they manage to do so, it is under non-disclosure agreements or other conditions that prevent them from publishing their work for peer review and open collaborative research. This severely hinders innovation and limits the scalability of collaboration effects. Due to such data gaps, researchers often have to rely on smaller, simplified, or outdated data inputs, which limits the applicability of their work to real use cases. It is also difficult to compare approaches because of the lack of established benchmarks representing such use cases with proper characteristics.

To circumvent this limitation, we advocate for realistic open datasets based on a mix of real, simulated, and synthetic data, the latter being derived from real data. These datasets should embed the features and characteristics of the real system so that any technique or AI model developed based on them will be likely to perform well when applied to real data. At the same time, they utilize statistical techniques, modeling tools, algorithms, and AI models to limit the risk of exposing confidential or sensitive information.

For energy stakeholders, investing in building open, realistic datasets is a strategic move to attract and nurture a strong AI innovation ecosystem around the real priority topics and challenges the organization faces. This approach will establish your company as a leader in AI innovation while also maximizing the speed of transition from innovation to actual implementation and value creation from AI.

Open commons and inner source

Don’t reinvent the wheel alone—build your AI stack using open source as a foundation.

Much of the AI innovation in energy already leverages open source repositories and tools such as PyTorch, Jupyter, or TensorFlow as a foundation. As a utility or system operator, you may see that your internal data scientists and AI specialists commonly use these state-of-the-art libraries as part of their work, but this also applies to your vendors and innovation partners when procuring software. By engaging in open source, you can see under the hood in both cases and help secure this supply chain of critical libraries and tools according to your strategic priorities, leveraging the resources of many actors. Extending this to models developed specifically for energy applications makes sense and will also provide strong benefits. In this context, setting up open commons projects is a strategic investment.

For energy applications, while AI models need to be specialized, there are a lot of common elements, parameters, and efforts that are shareable. This not only avoids duplication of effort but also allows training models on a broader set of parameters and contexts across regions, events, or system and market designs.

Within organizations, open sharing of information on AI models, their training and validation data, hyperparameters, and other documentation can help mutualize development efforts and create consistency in processes and tools. This approach is called "inner source" and has proven extremely successful, just as open source is the best approach to collaboration in AI.

Strategic investment in open commons, inner source, and collaboration will be the key to adopting and deploying AI in a rapid, secure, sustainable, and cost-effective manner.

AI literacy and governance

The new tool in the toolbox — build comfort with AI through openness and transparency.

While collaboration and technology are essential for AI readiness, energy stakeholders must also focus on developing the right internal culture and governance to support these innovations.

To capitalize on AI, energy organizations must foster AI literacy within their workforce. This means upskilling teams to understand how AI tools work, what data is needed, and how to interpret AI-driven insights. Stakeholders should support comprehensive training programs to ensure staff at all levels — from data scientists to operational managers — can interact effectively with AI systems. Cross-training between energy domain experts and AI specialists is also crucial to fostering collaboration and maximizing the value of AI systems — many AI experts do not understand power systems, and most power system experts are not AI specialists. While both may think the other field is easy to master, the most effective path to AI innovation and deployment in power systems is through close communication and dialogue between them. Open collaboration and innovation can bridge this gap, as well as help educational and academic institutions best prepare the workforce of tomorrow to understand both facets.

A data culture comes with AI literacy. In the age of AI, it is crucial to recognize data as useful and valuable even if the “what” and “how” are not obvious at first. AI-driven organizations invest in promoting a culture that focuses on bringing its data in order and in an accessible location. This is a requirement to foster AI innovation and rapid deployment, with a strong return on investment, and is very much part of AI literacy all the way to the strategic and executive level of the organization.

As AI systems integrate into energy operations, solid governance frameworks are necessary. AI deployment must be ethical, with clear guidelines for accountability, transparency, and bias mitigation. Stakeholders should prioritize establishing governance structures that ensure the responsible use of AI systems in compliance with regulations such as the E.U. AI Act and with careful attention to security and privacy concerns. Open frameworks can be a great resource to set up such governance structures.

Energy stakeholders that successfully integrate AI into their operations will be those that combine technological innovation with robust governance and a culture of continuous learning. By focusing on these areas, organizations can position themselves to lead the next phase of the energy transition driven by AI.

Priority use cases for open source AI projects in energy

There are many potential AI/ML applications in the energy sector, and recently, as in many other industries, R&D and innovation in this area have been booming. As an illustration, Entezari et al.[5] looked at the number of scientific papers related to AI and ML in the energy field (1970–2020).

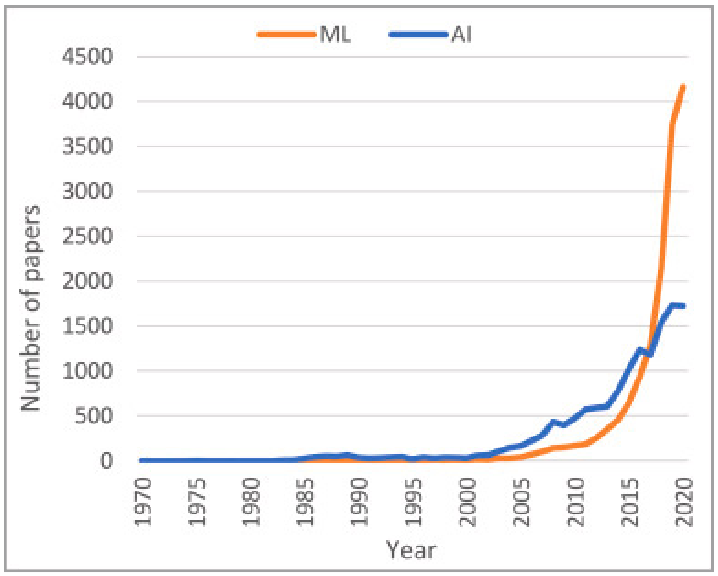
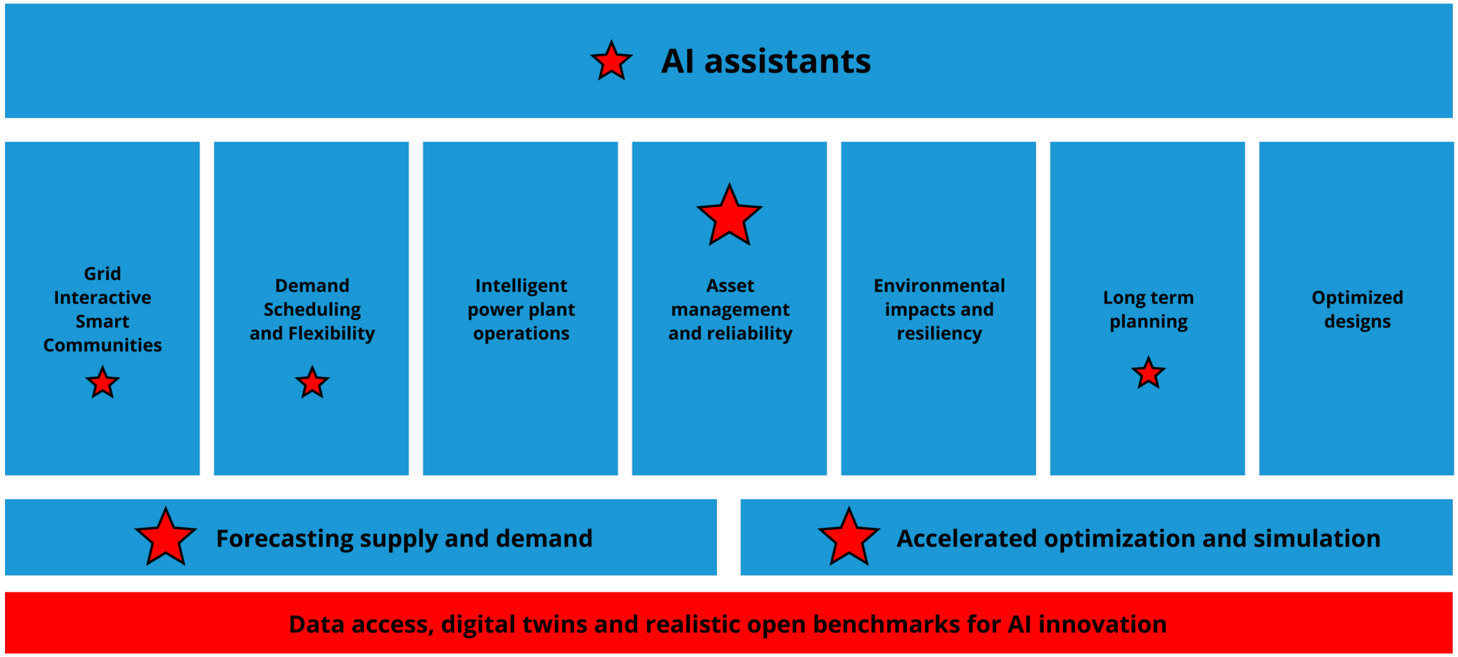


Image copyright 2023; "*Artificial intelligence and machine learning in energy systems: A bibliographic perspective*"; Ashkan Entezari, Alireza Aslani, Rahim Zahedi, Younes Noorollahi; Energy Strategy Reviews (CC BY 4.0)

While the field is evolving rapidly, some good review articles have been recently published. For example, Rolnick et al.[6] reviewed ML applications for climate change and accounted for about 900 publications, many of them in the last 10 years. Other papers focus on specific use cases, such as Antonopoulos et al.,[7] which reviewed AI and ML approaches to energy demand-side response.

Despite all these research efforts, actual industry adoption has been relatively slow, largely given the concerns described above, lack of AI readiness, and difficulty in putting together the necessary collaborations between industry and academia. In AI and ML, the transition from low technology readiness level[8] (TRL) innovation to high TRL and actual deployment requires open collaboration and open source projects, supported by best practices and an industry-grade focus from the outset. LF Energy is the premier organization that can enable this.

Through LF Energy, energy stakeholders can join forces to develop industry-grade open source AI solutions with a rich and sustainable ecosystem to support and commercialize them. With our members, we have prioritized some use cases as the most promising areas for such collaborations in the energy sectors, using the rich academic literature as input.



Forecasting

Forecasting has long been a cornerstone of effective energy grid management. In the traditional energy landscape, forecasting was primarily concerned with predicting electricity demand. Factors such as weather conditions, time of day and year, special events, and other variables served to predict the expected electricity consumption. This demand forecast was the basis for planning and scheduling large-scale power plants. The goal was to meet the projected demand optimally and reliably, taking into account the capabilities and constraints of these plants and of the transmission and distribution systems.

With the integration of variable and non-controllable distributed renewable sources, the emergence of demand-side flexibility and distributed storage, and the advent of new energy usage, such as electric vehicles, this forecasting process becomes even more complex and critical. Weather impacts demand and supply, along with other old and new factors, and more local effects need accounting for. For example, heat island phenomena in cities influence electric load and require careful evaluation. Predicting the resulting “net” demand curves (variable demand minus variable supply) with enough accuracy and confidence is essential for planning and scheduling remaining controllable power plants, storage cycles, and load flexibility and predicting and handling distribution and transmission grid limits and congestions. Beyond supply and demand, some applications seek to forecast other parameters, such as prices and CO_2 emissions, either as a byproduct of supply and demand predictions or through direct methods.

It is important to realize that a forecasting application is not an end but a means to inform a decision problem with its associated assumptions, objectives, and risk parameters. Therefore, as a general use, it may be beneficial to start from the actual decision problem one seeks to inform, like reducing cost or emissions, anticipating constraints, and evaluating the risk from extreme events rather than focusing on forecast accuracy. Developers may need to tailor underlying weather models for specific energy applications, depending on the most impactful phenomena, their spatial and temporal features, and the influence of uncertainties on the final decisions being informed.

Academic literature documents AI methods for forecasting demand, supply, or other parameters, and industries have already implemented them, often achieving performance similar to or better than traditional methods. Forecasting is a recurrent exercise, and input data is generally available, either from historical sources or outputs from weather models. Other sources, such as imagery or geographic databases, can provide additional information when relevant data, such as the distribution of solar panel installation, is not directly available.

The OpenSTEF[9] project offers a solid foundation for such use cases as it offers automated, open source ML pipelines for accurate short-term load forecasting. Developed by energy experts and constantly refined as part of an open industry development process, the OpenSTEF forecasting pipelines combine inputs of measured historical grid data with relevant external predictors such as weather and market pricing. Users can forecast the future load on any grid with the ability to look at energy consumption, (renewable) generation, or a combination of both.

OpenSTEF outputs forecasts via either an API or an (expert) graphical user interface. The entire technology stack leverages open source technology and open standards, and its microservice architecture optimizes performance for cloud native deployment. Alliander uses it for congestion management and to prohibit the exceedance of grid limitations. However, OpenSTEF also has other uses, such as solar farms using it to forecast peak shaving. Other use cases could be an aggregator steering its customers to adhere to a contract, forecasting as a service, or an energy trader that implements it to maximize returns on a specific market.

Many open source forecasting libraries exist in the research community, and utilities have been experimenting with AI/ML techniques and benchmarking them against their legacy solutions.

Building upon OpenSTEF, stakeholders can collaborate to advance the use of AI/ML techniques and improve forecasting in power systems, and LF Energy is the best environment to do this. As highlighted above, adapting generic tools as needed for specific use cases is necessary, taking into account the final decisions one seeks to inform and the risks one seeks to address. Open source enables this effectively, allowing each stakeholder to contribute to and modify the frameworks according to their needs while leveraging collaboration and community efforts.

Accelerated simulation and optimization

Optimization and simulation of energy systems have been key activities since the dawn of computing itself. The optimal dispatch and power flow problem is at the core of operations and markets, and it is a very complex problem with many parameters, both continuous and integer (mix integer programming), which creates difficult non-convexities. The integration of variable renewables sources and storage makes market clearing more demanding, with more frequent runs closer and closer to real time (and therefore less available time to run), larger regions to study with more granularity, etc. Simulation of power systems is also more demanding, mainly due to new components and controls based on power electronics and digital control algorithms. These require accounting both for large-scale, long-duration phenomena with step-by-step-based controls and rapidly responding devices.

AI has proven very effective here by speeding up traditional optimization and simulation tools through approximate solutions or parameter estimations. AI can train offline on a wide range of expected conditions and then draw from this training online very fast, therefore providing valuable input.

AI can also serve as proxies to approximate the behavior of subsystems that do not need precise simulations, retaining the most important features without the need for computationally intensive simulations. This can be a great tradeoff between very simple models, which may fail to capture relevant behavior, and more involved approaches that couple and co-simulate detailed models or tools. An example may be simulating electricity and gas system interactions or transmission and distribution grids—if the focus is on detailed simulations of one system, i.e., electricity or transmission, then a proxy of the gas or distribution system may be a very efficient approach.

There is a lot of research going on both in academia and national labs to improve optimization and simulation with AI, and much of this work is published and made available through open repositories. However, as stated in an article by H. Khaloie et al.[10] , "The shift from a beneficial machine-learning model to a secure and widely adopted platform is a path yet to be traversed." An LF Energy project could very well provide this path by creating the right collaborative environment involving industry stakeholders. The availability of standardized datasets and testbeds is a specific challenge to address. "The reviewed literature indicates that a vast majority of prior research predominantly depends on synthetic datasets generated within simulation environments. This reliance on synthetic data is largely attributed to the scarcity of real-world datasets". This is also an area where an LF Energy project would bring a lot of benefits.

Asset management

Reliability in energy systems refers to the consistent availability of power to meet the demands of consumers. A reliable energy system ensures a consistent supply of electricity or other forms of energy without significant interruptions or deviations of key parameters such as voltage and frequency. As decarbonization drives electrification, one can expect that the desired reliability will increase as electricity becomes even more central in our lives and economies. This creates additional challenges for asset management and operations.

Energy system assets have long life cycles and can be subject to harsh environmental conditions. In Western countries especially, most of the current infrastructure has been in place for a long time, and its continued performance is key even as they age and new or upgraded infrastructure takes time to build. Due to the variability of renewables and the need to mutualize supply over large areas, it's necessary to push assets closer to their limits. Ambient conditions need accounting for in real time to avoid unnecessary curtailment and extract the most from clean sources. The minimization of unplanned outages and optimization of maintenance and replacement are challenging in this dynamic landscape of aging infrastructure, variable loading, and limits. Supply chain management is also becoming increasingly critical as the system undergoes a profound transformation, and resources for manufacturing, extracting raw materials, transporting, and assembling equipment can experience stress, leading to shortages and long wait times.

AI helps leverage and analyze historical failure data, information from sensors, and measurements of ambient conditions. It allows us to improve the information on the present systems and make predictions about equipment performance, anticipate catastrophic component failure, and optimize maintenance while giving close to real-time information about acceptable stress levels, margins, and risks, and hence, determine the best maintenance and operational policies and trade-offs. It can inform on supply chain bottlenecks and best strategies to maintain and improve reliability.

Open source projects in this area will allow utilities, operators, vendors, and innovators to collaborate to build efficient solutions.

Long-term planning

Determining optimal long-term planning strategies has always been an important process in the energy sector. Assets are costly and generally have a long life — good planning decisions are crucial across the whole sector in ensuring reliable energy supply at affordable costs. Planning has proven more complex, mainly due to decentralized variable energy sources and longer permitting processes, which makes future supply more uncertain, while transmission and distribution infrastructure takes longer to build.

AI can help deal with uncertainties, generate scenarios, and propose solutions. There is a large amount of data that AI can leverage, but the difficulty lies in the overall methodology, with many non-technical elements factoring into the decision process. As an illustration, synthetic data generation has been used[11] in distribution planning by generating conditioned future profiles with ML models and later simulating via the power flow problem and identifying when investments may be necessary.

AI assistants

As grids evolve with renewables and prosumers, operators face growing challenges in real-time decision making under high uncertainty. Traditional multi-screen monitoring systems are reaching their limits, prompting the industry[12] to explore AI-driven hypervision and human-in-the-loop interfaces that unify data and support real-time decisions by operators. By combining human expertise with AI, operators can retain control over critical choices while offloading routine, data-heavy tasks to the AI, enhancing situational awareness and decision accuracy.

Open source frameworks are ideal for building such adaptable, collaborative AI tools. They enable grid operators to tailor algorithms and interfaces to specific operational needs, with the added benefit of continuous improvements from a community of researchers and engineers. Open source also promotes transparency, allowing operators to understand and trust the AI's recommendations. By openly developing and testing models, the community can address potential issues, such as over-reliance on AI, and foster adaptive systems that support human autonomy and accountability in grid management.

Grid interactive smart communities and demand side flexibility

At the grid edge, AI-powered devices, appliances, and control systems can help optimize energy use, lower costs for consumers, and promote flexibility to help operate the system in stressed conditions.

Interoperability and system integration will be critical for such applications. The components of grid-interactive smart communities will have to rely on local information, connection to signals from the distribution grid and system operator, and address cybersecurity and privacy imperatives.

While not directly focused on AI, open source projects such as LF Energy Grid Edge Interoperability & Security Alliance (GEISA)[13] can help develop interoperable AI/ML solutions that will meet these requirements.

Initiatives to accelerate AI adoption in the energy sector

In addition to the use case-specific projects above, LF Energy is also exploring more horizontal initiatives destined to enable AI adoption across a wider range of applications.

Digital twin platforms

Digital twins are transforming the energy sector by enabling AI to enhance grid reliability, optimize asset management, and integrate renewable energy sources seamlessly. These sophisticated digital models replicate real-world physical systems, providing a dynamic platform for AI-driven insights that improve both operational efficiency and strategic planning.

Digital twins serve as robust platforms for AI research and deployment by aggregating diverse data streams. This capability is crucial for developing AI models that require comprehensive datasets to analyze and predict complex grid behaviors accurately. By simulating real-world conditions and allowing for scenario testing, digital twins provide a unique environment where we can train, test, and refine AI algorithms under controlled yet realistic conditions. The TwinEU project,[14] supported by the European Commission, exemplifies this by aiming to create a digital twin of the European electricity infrastructure, illustrating the potential and applicability of open source methodologies in large-scale, complex systems.[15]

A recent IEEE PES Technical Report[16] provides a framework for understanding and deploying digital twins within the utility sector. It outlines the core functions and considerations for digital twins, emphasizing model maintenance, simulation engines, interactive visualization, and model validation—key areas where AI can play a transformative role. This report helps frame how to structure digital twins to support AI integration effectively, ensuring that AI applications developed using digital twins are both practical and scalable across different utility operations.

Adopting an open source approach to developing digital twins can significantly accelerate AI innovation across the energy sector. Open collaboration allows utilities, system operators, technology vendors, and innovators to share developments, enhance methodologies, and deploy AI technologies more rapidly. This collective approach not only speeds up technological advancements but also ensures that solutions are robust, scalable, and adaptable to various operational contexts. Open source also enables the reuse of software components from operational applications in the digital twin, saving development effort and ensuring greater consistency between the real world and its twin.

The TwinEU project serves as an excellent illustration of how to implement open source collaboration in the development of digital twins. By assembling a large consortium of partners from industry and academia, TwinEU is setting the foundations for a European power system digital twin ecosystem. All components developed within this project are open source, promoting transparency, interoperability, and shared innovation. For instance, TwinEU reuses industrial-grade software from LF Energy's PowSyBl[17] and Dynaωo[18] projects, which saves resources by avoiding reinventing the wheel and ensures greater accuracy and confidence in the digital twin by using software components already validated on operational use cases. Such collaboration is critical for tackling the complexities of modern energy systems and demonstrates the practical benefits of open source development in accelerating technological advancements.

LF Energy, with its commitment to fostering open innovation in energy systems, provides an ideal platform for pioneering the development of digital twins. By facilitating global collaboration, LF Energy can help standardize data models and interfaces, ensuring that digital twins developed under its umbrella are interoperable and meet the industry's complex needs. Stakeholders across the energy landscape should engage with this effort, contributing to the collective advancement of technology that will shape the future of energy systems worldwide.

Open datasets, synthetic data, and open benchmarks

As stressed above, access to high-quality, use-case-driven, and well-documented energy datasets will be a key enabler for AI applications in the energy industry. Open datasets and benchmarks will be critical to accelerating and scaling up AI innovation while respecting privacy and commercial confidentiality.

In particular researchers, modelers and policymakers need to understand how energy demand profiles are changing, in a system that requires greater real time optimization of demand and supply on the grid. Current global energy modeling and policymaking is still largely based on static and highly aggregated data from the past – when energy flowed in one direction, consumer profiles were relatively predictable, and power generation was highly controllable. Better data exists through Advanced Metering Infrastructure (AMI), or smart meters. However, access to energy consumption data is limited, in large part due to highly restrictive access to demand data as a result of privacy protections.

The OpenSynth[19] project illustrates this realistic open benchmark approach. It is a new, global, open community designed to democratize synthetic data and accelerate the decarbonization of global energy systems.

OpenSynth leverages synthetic data generation as the fastest way to achieve widespread, global access to smart meter datasets. OpenSynth is building a community for holders of raw smart meter (i.e. demand) data to generate and share synthetic data and models, as well as a community for quality assurance, so that researchers, industry innovators and policy-makers can use the data with confidence. OpenSynth partners Centre for Net Zero have undertaken research to define what constitutes 'good' synthetic smart meter data. Developed in partnership with academics from MIT, the University of Oxford and Georgia Tech, 'Defining 'Good':

Evaluation Framework for Synthetic Smart Meter Data' investigates how common frameworks used by other industries leveraging synthetic data (including financial services and healthcare), can be applied to synthetic smart meter data, such as fidelity, utility and privacy. The emerging nature of synthetic data means this kind of research is critical to its widespread adoption.[20]

The smart meter data shared through OpenSynth is not limited to synthetic approaches, and will contain important metadata such as property type, energy performance certificates, and low carbon technology ownership, including heat pumps, electric vehicles, and batteries. This will enable a better understanding of behind-the-meter changes and inform the ongoing development of future demand profiles for different demographics.

This approach also applies to other types of sensitive data, especially characteristics of electrical networks or operational information, as they often involve critical infrastructure protection aspects. LF Energy is exploring potential initiatives on realistic open datasets for research and innovation on optimal power flow, topology optimization, or asset management.

Foundational models for power systems

Foundational models represent[21] a new frontier in AI, leveraging self-supervision to learn from vast amounts of data and adapt to a variety of tasks. These models can autonomously extract structural information from complex systems, making them particularly suitable for dynamic and intricate systems such as electrical grids.

The electric power grid, being one of the most complex machines ever constructed by humanity, stands to benefit significantly from foundational models due to their ability to manage and simplify this complexity.

LF Energy has recently formed the GridFM[22] project to harness the potential of foundational models for the electric power grid, addressing the increasing uncertainties and complexities brought about by the energy transition and climate change. The project focuses on developing a robust foundational model specifically tailored for the electric power grid. These models leverage large datasets covering diverse grid scenarios for pre-training and undergo fine-tuning to meet specific operational needs, enhancing computational efficiency and facilitating the integration of renewable energy sources. The project team will build it on a modular architecture to support various phases of a model's lifecycle—from training on synthetic and real grid data (using federated learning to address privacy concerns) to deploying models for operational use. GridFM could speed up computations by several orders of magnitude, facilitating faster and more efficient grid management, planning, and operation.

The GridFM initiative at LF Energy represents a strategic opportunity for stakeholders across the energy sector to leverage a global collaborative environment. This project unites utilities, technology vendors, researchers, and other industry participants in co-developing foundational models that enhance grid management and accelerate the energy transition.

GridFM provides a platform where stakeholders can enhance their capabilities, drive innovation, and expand their knowledge and influence. Utilities have the opportunity to securely integrate their data into foundational models, gaining insights that are not possible in isolation. Technology vendors can develop and refine commercial solutions based on robust, open source foundational models, aligning with market needs and regulatory standards. Academics and researchers contribute to cutting-edge research that shapes the future of energy systems, influencing both policy and practice.

Engaging with the GridFM project allows participants to access a world-class research and development ecosystem, drawing insights from leading experts and advanced technologies in AI and grid management. Participation ensures that organizational needs and perspectives help shape the project's direction and outcomes. Moreover, stakeholders can realize economic and operational efficiencies by contributing to and utilizing GridFM, optimizing operations, reducing costs, and improving service delivery in the energy sector.

Conclusion

Embracing open source and open collaboration is not merely a strategic advantage but a necessary evolution in the rapidly transforming energy sector. LF Energy stands at the forefront of this transformation, driving initiatives that leverage AI to accelerate the decarbonization of global energy systems and enhance grid management.

To advance these efforts, LF Energy has formed a special interest group (SIG) dedicated to exploring and expanding the role of AI in energy. This SIG focuses on identifying opportunities for open source projects and initiatives in AI and energy, as well as sharing best practices on AI readiness. This whitepaper provides a summary of the group's activities so far and its roadmap ahead, along with specific projects and actions that complement or derive from them.

Stakeholders can tap into a rich ecosystem of knowledge, tools, and collaborative opportunities. This engagement allows for direct influence on project directions and priorities, addressing your unique challenges and needs. The benefits of participating in this collective effort include accelerated innovation cycles, shared development costs, and access to cutting-edge AI applications tailored to the energy sector.

We invite all energy stakeholders to join us in this endeavor. Whether you are a utility, a technology provider, a policymaker, or an academic entity, your contribution is valuable. Together, we can not only meet the challenges of today's energy demands but also shape the future of energy to be more sustainable, efficient, and resilient.

Join LF Energy and its members in pioneering the integration of open source AI into the energy sector, reaping the benefits highlighted throughout this whitepaper, and contributing to a sustainable energy future.

Acknowledgments

This whitepaper is the product of the work and discussion of the AI SIG at LF Energy, coordinated by Alexandre Parisot, LF Energy's director of ecosystem, AI, and energy systems.

We would like to especially thank the following members of the group for their input, comments, and insights: Gus Chadney, Sheng Chai, Boris Dolley, Virginie Dordonnat, Abder Elandaloussi, Mital Kanabar, David Lamers, Vincent Lefieux, Francois Miralles, Camille Pache and Pedro Vergara Barrios.

Special thanks also go to Lucian Balea and Alex Thornton for their help and support. The section on trustworthiness originated largely from expert discussions with Patrick Panciatici and Pascal Van Hentenryck.

End Notes

1 *https://arxiv.org/abs/2403.13784*

2 *Artificial Intelligence and Data in Open Source, I. Habbad, LF Research in partnership with LF AI & Data, March 2022*

3 *Alex Thornton, "Accelerating grid modernization with open technology and standards" https://www.utilitydive.com/news/accelerating-grid-modernization-with-open-technology-and-standards/735363/*

4 *D. Slate et al., Adoption of Artificial Intelligence by Electric Utilities, Energy Law Journal, Vol 45:1, 2024*

5 *Ashkan Entezari et al., Artificial intelligence and machine learning in energy systems: A bibliographic perspective, Energy Strategy Reviews, Volume 45, 2023, 101017, https://www.sciencedirect.com/science/article/pii/S2211467X22002115*

6 *David Rolnick et al., Tackling Climate Change with Machine Learning. ACM Comput. Surv. 55, 2, Article 42 (February 2023) https://dl.acm.org/doi/full/10.1145/3485128*

7 *Antonopoulos, Ioannis et al. "Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review." Renewable and Sustainable Energy Reviews 130 (September 2020): https://dspace.mit.edu/handle/1721.1/130282*

8 *Lavin, A., Gilligan-Lee, C.M., Visnjic, A. et al. Technology readiness levels for machine learning systems. Nat Commun 13, 6039 (2022). https://doi.org/10.1038/s41467-022-33128-9*

9 *https://lfenergy.org/projects/openstef/*

10 *Khaloie, H. et al. Review of Machine Learning Techniques for Optimal Power Flow (May 31, 2024). Available at SSRN: https://ssrn.com/abstract=4681955 or http://dx.doi.org/10.2139/ssrn.4681955*

11 *E. M. S. Duque et al. "Risk-Aware Operating Regions for PV-Rich Distribution Networks Considering Irradiance Variability," IEEE Transactions on Sustainable Energy, vol. 14, no. 4, pp. 2092-2108, Oct. 2023, 10.1109/TSTE.2023.3281890.*

12 *Antoine Marot at al. Towards an AI assistant for power grid operators, arXiv:2012.02026*

13 *https://lfenergy.org/projects/geisa/*

14 *https://twineu.net/*

15 *A. Monti et al. "Building a Digital Twin for European Energy Infrastructure: The role of open source", IEEE Electrification Magazine, vol. 12, no. 3, pp. 78-84, Sept. 2024*

16 *Digital Twins for Electric Utilities: Definition, Considerations, and Applications, IEEE PES Technical report PES-TR 122, April 2024*

17 *https://www.powsybl.org/*

18 *https://dynawo.github.io/*

19 *https://lfenergy.org/projects/opensynth/*

20 *https://arxiv.org/pdf/2407.11785*

21 *H. Hamann et al. A Perspective on Foundation Models for the Electric Power Grid, Joule, 2024,ISSN 2542-4351, https://doi.org/10.1016/j.joule.2024.11.002*

22 *https://lfenergy.org/projects/gridfm/*

LF ENERGY

www.lfenergy.org