Alliance for AI, IoT and Edge Continuum Innovation

AI in Energy

Release 1.0

AIOTI WG Energy

21 November 2024

Executive Summary

The rapid integration of Artificial Intelligence (AI), 5G, and upcoming 6G technologies into the energy sector signifies a transformative and long-term transitional shift towards a more sustainable and efficient future enabled by myriads of digital technologies.

This paper focuses on the potential and challenges of leveraging AI across various segments of the energy landscape and stakeholders, including generation, distribution, and consumption within smart grids. Through comprehensive and dynamically evolving digital infrastructures the advanced data analytics, real-time monitoring, and predictive modelling, AI enhances the flexibility, reliability, and efficiency of energy systems. The convergence of AI and 6G technologies is particularly critical in optimising the performance of renewable energy sources, rolling out smart grid solutions to ensure grid stability, and fostering a resilient energy ecosystem.

Key insights from this paper highlight the rapidly evolving role of AI in driving energy innovations, addressing challenges and policy and standardisation aspects, the ethical and cybersecurity considerations that accompany the deployment of AI solutions and technologies. The system and digital technologies stakeholders are embracing the full potential of AI to create a cleaner, smarter, and more resilient energy.

Table of Content

Table of Figures

List of Tables

Introduction

The rapid integration of Artificial Intelligence (AI), 5G, and upcoming 6G technologies into the energy sector signifies a transformative and long-term transitional shift towards a more sustainable and efficient future enabled by myriads of digital technologies.

This paper focuses on the potential and challenges of leveraging AI across various segments of the energy landscape and stakeholders, including generation, distribution, and consumption within smart grids. Through comprehensive and dynamically evolving digital infrastructures the advanced data analytics, real-time monitoring, and predictive modelling, AI enhances the flexibility, reliability, and efficiency of energy systems. The convergence of AI and 6G technologies is particularly critical in optimising the performance of renewable energy sources, rolling out smart grid solutions to ensure grid stability, and fostering a resilient energy ecosystem.

Key insights from this paper highlight the rapidly evolving role of AI in driving energy innovations, addressing challenges and policy and standardisation aspects, the ethical and cybersecurity considerations that accompany the deployment of AI solutions and technologies. The system and digital technologies stakeholders are embracing the full potential of AI to create a cleaner, smarter, and more resilient energy future.

The following key insights and innovations are analysed:

AI-driven energy innovations:

AI enables advanced data analytics and real-time monitoring, which are essential for improving the efficiency and reliability of energy systems. Predictive modelling facilitated by AI helps in anticipating and mitigating potential disruptions in energy supply and demand.

Optimisation of renewable energy sources:

The integration of AI with 6G technologies is particularly effective in optimizing the performance of renewable energy sources. Enhanced connectivity and data processing capabilities of 6G support real-time adjustments and efficient energy distribution.

Grid stability and resilience:

AI contributes to maintaining grid stability by predicting and responding to fluctuations in energy supply and demand. The deployment of AI-driven solutions ensures a more resilient energy infrastructure capable of withstanding various challenges.

Policy and standardisation:

Addressing policy and standardisation issues is crucial for the seamless integration of AI technologies in the energy sector. Developing comprehensive standards ensures interoperability and enhances the overall efficiency of energy systems.

Ethical and cybersecurity considerations:

The deployment of AI technologies raises important ethical and cybersecurity concerns that must be addressed to ensure safe and responsible use. Implementing robust cybersecurity measures is essential to protect energy systems from potential threats and vulnerabilities.

Further addressing these critical topics dynamically leads towards adopting and advancing potential of AI to create a cleaner, smarter, and more resilient energy future in Europe in line with Fit for 55 and Green Industrial Deal. The convergence of AI with 5G and 6G technologies not only enhances the efficiency and reliability of energy systems but also plays a pivotal role in driving innovations and ensuring sustainable development in the energy sector.

3C's: collaborative connectivity and computing as mentioned in the **European Commission White Paper How to master Europe's digital infrastructure needs?** discusses the advanced digital network infrastructures and services that will become a key enabler for transformative digital technologies and services such as AI, Virtual Worlds and the Web 4.0, and for addressing societal challenges such as those in the fields of energy, transport or healthcare and for supporting innovation in creative industries.

This represents an opportunity for new pilots closing demand and supply side, i.e. strengthen industrial demand side and deliver an enablement for the path towards sustainability of key sectors. It can enable new innovations and services for European digital supply industry while strengthening the industrial ecosystems in areas like mobility, energy, health, smart communities, logistics and manufacturing, drawing on our experiences from the IoT Pilot projects funded by the EU.

In addition, synergies with Industry 5.0 aim to bring Industry 4.0 towards a more responsible manufacturing, driven by human-centric, resilient and sustainable production. I5.0 is not in reality in opposition with I4.0 and inherits from it the basic concepts of Cyber Physical Systems and the way to implement it through digital technologies, what is now called IoT and edge computing.

Energy management is a key aspect of the sustainability pillar in Industry5.0 and regards both the product lifecycle (sustainability by design till circularity in end-of-life), the production sites (measuring and controlling the energy consumption by manufacturing systems and human operators) and the whole value chain, including considerations of resilience to cyber attacks, to financial disruptions but also to sudden changes in the demand and offer which could be detected and addressed by IoT and edge computing technologies

Managing and governing the Industry5.0 transition, in particular the aspects regarding energy management, requires new roles, characterised by new skills and competencies for managers, engineers and workers which could be assessed and improved in companies by proper methods and tools. While large companies do have the necessary human and financial resources to develop such new competencies and skills and to remain competitive on the market, SMEs need partnerships and intermediaries which could provide such skills "as a Service". This is one of the four fundamental pillars of the (E)DIH services, beyond "support to find investments", "ecosystem building" and "test before invest".

1. Continuity from AIOTI Edge driven Digital Twins in distributed energy systems paper

As the industry embrace cutting-edge technologies such as digital twins, the incorporation of artificial intelligence (AI) becomes indispensable. Digital twins, which are virtual replicas of physical systems, depend on robust data infrastructure to ensure accuracy. AI plays a crucial role in maximizing the potential of digital twins by enabling sophisticated data analysis, predictive capabilities, and real-time decision-making.

The integration of AI significantly enhances digital twins through efficient data collection, secure storage, and advanced management systems. This infrastructure guarantees that digital twins mirror the physical world with exceptional fidelity. AI algorithms process extensive amounts of data, continuously updating the digital twin with current and pertinent information.

In the realm of energy optimization, AI-driven digital twins offer detailed models of energy systems. AI algorithms simulate and analyse consumption patterns, leading to improved energy efficiency and cost reduction. AI also forecasts demand and supply fluctuations, allowing for proactive adjustments that optimize energy usage.

Predictive maintenance reaps significant benefits from AI integration. AI-powered digital twins accurately forecast equipment failures by analysing historical and real-time data, identifying patterns, and providing early warnings. This proactive approach prevents downtime, optimizes maintenance schedules, and extends equipment lifespan.

Another area where AI excels is process optimization. AI-powered digital twins model production processes, identifying inefficiencies and suggesting improvements. Continuous learning from operational data enables ongoing enhancements in productivity and efficiency.

AI ensures that digital twins are dynamic systems that provide actionable insights, enabling informed decision-making and operational optimization. Investing in AI and data infrastructure is strategic and impacts revenue and stability.

AI-driven digital twins also play a pivotal role in improving environmental sustainability by optimizing resource use and managing renewable energy sources efficiently, thereby reducing reliance on fossil fuels and lowering carbon footprints.

AI integration transforms digital twins, enhancing accuracy, efficiency, and predictive capabilities. This strategic investment enables unprecedented levels of efficiency, sustainability, and reliability, positioning organizations at the forefront of technological innovation.

2. References to AIOTI SRIA

AIOTI has published in January 2023 **Strategic Research and Innovation Agenda** (SRIA).

In AIOTI SRIA there are the following references to AI, digital twins and Energy:

2.1 AI:

Federated Learning, Artificial Intelligence technologies and learning for edge IoT Systems

To perform according to the devised expectations[1], the new distributed IoT architectures for computing optimisation across the edge continuum need to improve responsiveness by reducing decision-making latency, to increase data security and privacy, to decrease power consumption, using less network bandwidth, thus maximizing efficiencies, reliability, and autonomy.

The IoT edge contains computing capabilities scaled across the micro-, deep- and meta-edge to process workloads, including the latest technology like AI model training and ML inference and signal processing, using signal conditioning[2] followed by neural networks[3] computing. The neural network computing and memory requirements are significantly reduced by using signal conditioning on the raw data.

In this context, federated learning s a distributed ML technique, which creates a global model by learning from multiple decentralised edge clients, is a significant technological development that can be implemented in distributed IoT architectures across the edge continuum.

Federated learning uses complex methods for handling distributed data training by enabling the cooperative training of common AI models, by combining and averaging locally calculated updates submitted by edge IoT devices. Federated learning permits training new models on multiple edge IoT devices simultaneously without the need to have data stored in a central cloud.

Technological developments

Integrating AI-based techniques across the edge continuum requires a new layer of edge processing infrastructure and scalable, energy-efficient modules for AI-based processing[4].

Federated learning methods offer several advantages, including scalability and data privacy, with ML and DL algorithms that can be executed on edge IoT devices, delivering faster real-time insights for increased IoT application efficiency. Bringing AI to the edge increase the efficiency of processing the data locally and reduces latency and the cost of connectivity for many IoT applications.

[1] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang and W. Zhao, "A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications," in IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1125-1142, Oct. 2017, doi: 10.1109/JIOT.2017.2683200.

[2] R. Tirupathi and S. K. Kar, "Design and analysis of signal conditioning circuit for capacitive sensor interfacing," 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), Chennai, India, 2017, pp. 1717-1721, doi: 10.1109/ICPCSI.2017.8392007.

[3] Z. Li, F. Liu, W. Yang, S. Peng and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 12, pp. 6999-7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.

[4] Vermesan, O. and Nava, M.D. (Eds.). Intelligent Edge-Embedded Technologies for Digitising Industry. River Publishers Series in Communications, June 2022. ISBN: 9788770226110, e-ISBN: 9788770226103. Online at: https://www.riverpublishers.com/pdf/ebook/RP_E9788770226103.pdf

In the distributed data exchange environment, federated learning/training combined with the IoT heterogeneous compute based on various underlying processing architectures (CPUs, GPUs, NPUs, neuromorphic, etc.) can provide the solution for future IoT edge intelligent heterogeneous systems.

The federated learning/training must consider the underlying connectivity systems, including Bluetooth, Wi-Fi, LPWAN, mesh, 5G and beyond applied to the IoT application topology, range, and power requirements. The convergence of connectivity, AI/ML, and processing can ease the implementation of different IoT federated learning/training solutions for industrial and consumer applications.

The concept of IoT federated learning/training is combined with the IoT edge mesh, which allocates the decision-making tasks among edge devices within the network. The computation and processing tasks and data are shared using an IoT networks of edge devices, routers, PLCs, and micro-servers.

The IoT edge mesh combined with federated learning/training provide distributed processing, low latency, fault tolerance, scalability, security, and privacy, as required by IoT applications, which demand higher reliability, real-time processing, mobility support, and context awareness.

Cooperative IoT computing based on federated learning/training provides better usage of resources, reduced latency, due to easy access to local resources, better services, as IoT devices can cooperate to get better information, reduced communication with centralised entities, and improved security and privacy as data remain most of the time within a local network.

Enabling distributed intelligence in IoT/edge computing or swarm computing applications is difficult due to a set of known problems, e.g., synchronisation, consensus, cooperation etc. In addition, due to scalability and complexity issues of IoT systems, it is challenging to determine how to generate, coordinate and federate the intelligence, which edge IoT device provides intelligent functionality and how different edge IoT devices cooperate, transfer, and acquire intelligence.

Main Trends, Issues and Challenges

The challenge for many IoT applications is that federated learning/training uses multiple entities collaborating to solve ML problems under the coordination of a central server. This approach for edge networks creates many issues concerning security and privacy and the data itself.

Federated learning raises several risks and weaknesses in terms of computational complexity in the case of heterogeneous edge IoT devices that may have limited computing resources, inadequate wireless connectivity quality, or may use different OSs.

Another edge IoT federated learning challenge relates to communication delays expressed as the latency between edge IoT devices and the ML system. Decreasing latency is critical for AI-based edge IoT devices operating in real-time applications such as industrial equipment.

Replacing the client-server process of the federated learning/training model with fully decentralised learning replaces communication with the server by peer-to-peer communication between individual clients.

Optimisation algorithms are necessary to implement federated learning/training at the edge, considering edge IoT devices' constraints and resource limitations as part of the edge continuum.

The IoT devices limited bandwidth can restrain scalability, but this is solved in IoT edge mesh architectures, as data is sent to multiple edge IoT devices that share data with other devices. The communication bottleneck issue is resolved due to the distributed nature of the system.

At the same time, the computation tasks are offloaded to different edge IoT devices, operation which speeds up the processing time and increases the efficiency of federated learning/training, leading to better response time, reduced make span, and higher throughput. The distribution of loads drives the edge IoT systems to be more flexible and robust, as, in the case of a device failure, other devices can share the load of the failed IoT device.

IoT systems are dynamic, as devices can be mobile, added, removed, or changed in configuration; all of the above require new context-aware solutions for distributed security and privacy algorithms.

The heterogeneity of computing, communication and AI technologies requires AI-based algorithms that are portable across different IoT edge environments. Communication technology is intrinsically heterogeneous for what concerns data rate, transmission range, and bandwidth. The IoT SW solutions depend on the HW, and programming models are needed to execute workloads simultaneously at multiple HW levels.

Research on standard protocols and interfaces should address the integration of AI-based algorithms and lightweight protocols for communicating with different devices in a heterogeneous environment.

More research is needed to develop distributed learning/training algorithms and to maximise the average time between errors and optimise the availability by minimising the failure probability and average recovery time in the OTA learning/training process.

The IoT federated learning/training must consider the hybrid computing method that combines HW/SW and AI techniques across the edge continuum. Hybrid computing implies the integration of specialised advanced AI processors at different computing levels for both high-level and low-level operations.

Further research is needed to develop optimised algorithms for federated learning/training to complement and effectively leverage the computing capabilities with the AI-based processors and other types of processor architectures.

Research Priorities Timeline

Table 1 - Federated learning and AI for edge IoT systems research priorities

Topic	Short Term 2023-2024	Medium Term 2025-2027	Long term 2028-2030
Federate learning approaches	Techniques and methods to integrate federated learning into IoT/edge computing systems. Management of edge IoT systems, by addressing mesh network security and management by leveraging ML. Research on central training data sets and edge IoT local data sets.	Edge IoT intelligence architectures and AI frameworks for federated learning. Development of tools and tool chains for dedicated edge IoT federated learning. Methods for providing reference training datasets for performing standard federated learning application tuning.	Benchmarking techniques and methods for edge IoT federated learning. Scalability and portability of AI-based models for federated learning across the edge granularity.
Federated learning architecture and frameworks	Advanced architectural approaches for the federated learning server integrated into mesh networking environments.	Extend the capabilities of open-source federated learning frameworks. Communication and computation efficiency of the federated learning architectures, synchronisation optimisation among edge IoT devices.	Federated learning architectures addressing task scheduling, dynamic resource allocation to achieve low-latency services.
Hardware platforms for federated learning	HW requirements for implementing federated learning in edge IoT computing environments.	HW heterogenous solutions to minimise memory transfer, increase energy efficient and improve computational speed. Computation offloading and content caching using dynamic cache allocation techniques, context-aware offloading algorithms adapted to resource-constrained (e.g., limited storage and capacity) edge IoT devices.	HW/SW/AI algorithms heterogeneity management. Understanding of the effect of system heterogeneity on the AI model aggregation efficiency, accuracy and the divergence or convergence of optimisation processes.

2.2. Digital Twins

IoT Digital Twins, Modelling and Simulation Environments

An IoT DT is a virtual representation of an IoT device that models the device's characteristics, properties, environmental conditions, behaviours, and functions over the operational lifetime, based on real-time data and information synchronised automatically and bi-directionally at a specified frequency and accuracy. An IoT DT uses simulation, ML, and reasoning to simulate various scenarios in different IoT applications and help optimise and improve the overall IoT system functionalities and services.

The real-time feature represents a vital characteristic to define IoT DTs, considering that the real-time instances vary according to IoT applications. In many IoT applications, time values are not defined identically, and such issue should be carefully considered when designing IoT DT instances. The synchronisation between the physical IoT device and its virtual representation in the simulation environment and the synchronisation of the events and scenarios in the simulation platform is critical for the performance of the whole IoT system.

Technological developments

IoT DTs support optimal decision-making and effective and efficient actions of the IoT devices in IoT applications, using real-time and historical data to represent the past and present and simulate predicted activities tailored to use cases based on domain knowledge, and implemented in Information Technology / Operation Technology (IT/OT) systems. The IoT DTs can expose a set of services to execute certain operations and produce data describing the physical activity of the IoT devices that they virtualise.

How to choose and optimize characteristics, properties, environmental conditions, behaviours, and functions over the operational lifetime of an IoT device that are mapped to the IoT DT is a matter that needs further research; moreover, further analysis is needed also to properly develop IoT DT user interfaces that can connect the IoT DTs with other devices, with humans, and with other DTs coming from different domains (e.g., telco DTs or connectivity DTs).

The lifecycle evolution of the IoT DTs must take into consideration the updatability and upgradability of the IoT devices, including new features addressing the dependability characteristics.

Operational intelligence is used to build IoT DTs as it supports digitising the IoT and edge computing infrastructure, monitoring operations in real-time, predicting events, taking actions based on intelligence and engaging with different stakeholders.

The virtual representation of an IoT DT reflects all the relevant dynamics, characteristics, critical components, and essential properties of the IoT physical device throughout its life cycle. The creation and update of IoT DTs rely on timely and reliable multi-sense wireless sensing, while the cyber-physical interaction relies on timely and dependable wireless control over many interaction points, where wireless interfaces of the IoT device are embedded.

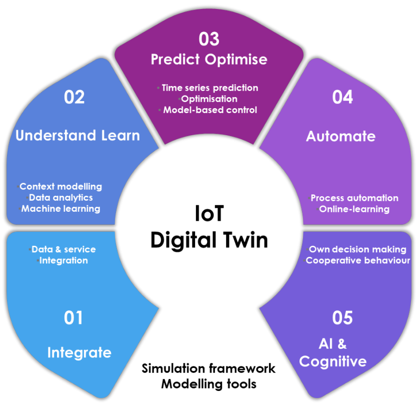
In future networks, IoT DTs will be a valuable tool to create novel and disruptive solutions, especially for vertical industries, that are enabled by a large scale of real-time, robust, and seamless interactions among, for example, machines, humans, and environments.

IoT DTs need to be scaled up in IoT applications, thus enabling for the broadest possible population a sustainable living with systematic climate mitigation measures, improving society's resilience in crises by actively monitoring and simulating a huge number of future scenarios, and potentially helping transform the whole societal structure, so to make it more robust and capable of addressing the environmental (but not only) challenges of the future.

IoT DTs must possess a minimum set of attributes to be integrated into IoT applications and platforms to optimise the functions and services of these applications. These IoT DT attributes are summarised below:

- Abstractness – free from details which are specific to implementations.
- Correctness – give a correct replication of the IoT ecosystem and its devices.
- Completeness – updated vis a vis the functionality in the real-world system.
- Expandability – adapt easily to emerging technologies and applications.
- Parameterised – accessible for analysis, design, and implementation.
- Reproducible – be able to replicate the same result for the same input as the real system.
- Scalability – must be able to operate at any scale.
- Soundness – exhibit only the functionality available in the real-world system.
- IoT DTs will continue to grow in industrial and production environments, leading to the new designing approach called massive twinning. It will enable to go beyond the current levels of agility of production, thus allowing more efficient interaction of production means to encompass a more significant extent of the respective processes, and achieve the transfer of massive volumes of data, as well as, often, reach unprecedented performance and reliability levels. The evolution of edge IoT digital twin technology is illustrated in **Figure 1.**

Intelligent IoT twins advance process automation by providing decision-making based on actual and simulated scenarios by implementing cooperative behaviour based on the information exchanged in real-time with the edge IoT physical systems.



IoT digital twins implementation collects real-world data about edge IoT devices or systems as inputs. It produces outputs simulations or predictions of how that edge IoT physical devices or systems are affected by those inputs

Figure 1 Edge IoT digital twin technology evolution

IoT DTs, as part of IoT technologies and applications, are being expanded to support more applications, use cases and vertical industries, as well as combined with more technologies, such as speech, augmented reality for an immersive experience and AI capabilities, enabling to look inside the IoT DT by eliminating the need to go and check the physical IoT device.

Main Trends, Issues and Challenges

The research today focuses on developing the virtual model representation of an IoT device, the evolving data sets relating to the IoT domain, the mechanisms to dynamically synchronise and adjust the virtual representation following the changes into the physical IoT device, and the simulation environment in which the IoT DTs will operate.

The IoT DT mapping of the physical environments into the digital world is facilitated by IoT simulation platforms and SW leveraged to create a digital model and a virtual representation of the physical IoT device. IoT digital virtual representation can be used to manipulate and control the real-world IoT device through a teleoperation DT modelling solution.

The research areas for IoT DTs must consider the scope and augmentations of the IoT DT and the operational environment combined with the functions needed to realise the IoT DTs' communication capabilities, and the update frequency required for providing the optimal precision of the IoT DT, based on data measured and acquired during the operation and use of physical IoT devices.

Further research needs to investigate the different levels of IoT and edge computing intelligence through cognitive functions, implemented in the physical/digital and virtual devices.

Understanding, defining, and designing the simulation capabilities of the future IoT platforms, which will be able to provide different fidelity levels of simulation tuned by input parameters, time dependency, behaviour, and prediction aspects, intelligence, and IoT device complexity, are very challenging tasks, which require further investigation.

Research Priorities Timeline

Table 2 - IoT Digital Twins, Modelling and Simulation Environments research priorities

Topic	Short Term	Medium Term	Long term
	2023-2024	2025-2027	2028-2030
IoT DT Models	Aggregation of heterogeneous IoT DT models.	Energy-efficient models. E2E features and optimisation.	Horizontal and vertical integration of IoT DT models. IoT DT that is capable of modelling and simulating the future state and behaviour of the IoT device.
IoT DT Modelling and Simulation Platforms	IoT DT platforms at the edge. Virtual sensing and actuation functions and simulations.	Predictive modelling platforms. Modelling and simulation of energy efficiency.	Integrated IoT platforms with virtual simulation environments including XR.
IoT DT Security	IoT DT security features integrated.	IoT DTs counterfeiting identification and mitigation.	Automatic recognition of fake DTs and their isolation or elimination.
IoT DT Connectivity	Simulation and modelling of the communication channels.	Define influence of the environments on the communication parameters of the IoT DTs.	Virtual platforms for the connectivity of IoT devices.

2.3 Energy

Energy-Efficient Intelligent IoT and Edge Computing Systems

The number of IoT applications is constantly increasing, due to the fact that more and more IoT devices are being deployed in different industrial sectors.

Such IoT applications are consuming increasing amounts of energy, and new technologies and methods need to be developed to increase the energy-efficiency of the IoT devices, AI algorithms, architectures and IoT systems to reduce the overall power consumption.

Next-generation IoT edge applications and networks need greater flexibility to implement edge utilisation mechanisms to maximise energy-efficiency, latency, processing, data transfer, and dependability.

From the IoT system architecture perspective, the technological trend is to move the data processing and analysis from cloud to edge. This shift requires that the edge IoT devices in the edge micro-, deep- and meta-edge domains (e.g., sensors/actuators, microcontrollers, end-devices, gateways, edge servers) become more energy-efficient and support AI techniques at low power consumption to ensure high autonomy/longer battery life, system availability and reliability.

To make such shift happen, a new generation of more performant processing units and new architectures (e.g., neuromorphic and hybrid) are needed, which can guarantee the best trade-off between communication power and ultra-low power consumption and increased intelligent processing needs.

Technological developments

The move of the data processing to the edge and the implementation of distributed IoT computing architectures require optimising the location of processing and the transfer of intelligence where the application needs it (e.g., micro-, deep, meta-edge).

Energy-efficient and green IoT requires a holistic E2E strategy through the IoT architectural layers across the information value chain to address the entire edge IoT systems energy-efficiency continuum and energy management. This energy-efficient design optimisation is required for green IoT components and algorithms at each IoT architectural layer level.

Combining energy-efficient AI and IoT technologies at the edge can maximise the IoT capabilities across the architectural layers and optimise the whole IoT system, including the application domain.

Edge IoT and AI green designs (e.g. advanced and adequate semiconductor technologies, efficient design, energy-efficient SW/HW platforms) are needed for providing environmentally reliable components at all IoT architectural layers and functions, energy-efficient and low CO_2 footprint at IoT infrastructure and technical solutions (edge, hybrid edge-cloud, AI-based learning/training, etc.), and finally also allowing the deployment of green manufacturing (e.g. manufacture IoT electronic components, HW/SW platforms, and IoT systems with minimal or no impact on the environment).

The implementation of green IoT and AI energy-efficient techniques and methods (optimisation, trade-off analyses among crosscutting functions/system properties vs. energy, green IoT/AI metrics, performance, measurement, testbeds, energy harvesting, wireless power transfer, etc.) depends on the functions performed by different HW/SW/algorithms components integrated into the IoT architectural layers.

These techniques include energy management, wake-up scheduling mechanism and selective sensing, HW/SW partitioning, energy-efficient methods/algorithms, communication techniques and distribution of task, efficient IoT nodes and resources on multi-core, minimisation of data path length, data buffer delivery, wireless communication, processing of trade-off communication.

Main Trends, Issues and Challenges

More edge IoT devices and applications are deployed together with intelligent edge IoT platform solutions used to collect, process, analyse the IoT device-collected data while making decisions and taking actions. AI is applied to most of these edge IoT devices to implement the future Internet of Intelligent Things.

Wireless and cellular communication technologies enable for collection of even more significant amounts of data from intelligent edge IoT devices.

The increase in the connected edge IoT devices, combined with the distributed computing model introduced by edge computing, reduce latency and the amount of data that needs to be sent between the central cloud and the edge IoT devices, saving bandwidth costs.

The IoT edge model extends security benefits, and the localised edge processing allows autonomous control of devices when the communication networks are jammed, or the connection is lost.

The capabilities, performance, responsiveness, and energy-efficiency of the IoT edge processing models are increased due to reduced data transmission and distributed processing.

The next-generation IoT and AI edge solutions should focus on novel energy management techniques to select energy sources, energy harvesting techniques, HW/SW/algorithms optimisation for data sensing, monitoring, filtering, prediction, and compression.

Processing combined with sleep/wake-up techniques, energy-efficient task scheduling algorithms, selection of Quality of Information (QoI), allocation of workload distribution at the edge are used together with wireless communication optimisation (send/receive), power down mechanisms to improve the energy-efficiency of IoT systems.

The dynamic wireless network behaviour (e.g., IoT devices-move-in and IoT devices-move-out) is monitored, and the cooperation/information exchange between the edge IoT devices (with optimised green and energy profiles) is optimised to increase the overall IoT system energy-efficiency.

The integration of ML and AI methods support the optimisation of IoT/edge computing-based green and energy-efficient functions providing solutions for moving the processing optimally from cloud to the edge and decarbonising the whole value chain of IoT information.

The optimisation for energy-efficiency and green IoT requires the use of federation and orchestrations techniques that create dynamic and distributed energy control frameworks for edge IoT applications.

The implementation of energy efficient IoT intelligent search engines, cooling systems, and energy harvesting techniques and renewables must be considered when the HW/SW/algorithms components of the IoT application layer are evaluated.

New IoT applications, including AR/VR, DTs, virtual simulations, real-time searching engines and discovery services, bring new challenges to optimising energy-efficiency as the virtual simulation and the AI, learning/training algorithms, are increasing the energy consumption of the whole IoT system that will have two components one physical and one digital/virtual.

The complexity of intelligent IoT applications at the edge requires designing, analysing and optimising the energy-efficiency at the IoT system level by considering the aggregation, over the technology stack, of the functions required to fulfil a given IoT task. This includes estimating the energy used for learning/training of different algorithms implemented in various IoT architectural layers, both during inference and learning by employing real data sets from different databases, the energy consumption of the edge IoT devices (micro-, deep-, meta-edge), the energy consumption of the communication networks and the other processing and storage units by the IoT application, for performing different tasks and services.

Currently, most of the embedded IoT devices and as well low-power IoT sensors are powered by batteries which need to be replaced every few years due to their limited lifespans. Usually, the replacement of these batteries can be costly and therefore solutions on enabling energy efficiency for communicating IoT devices and embedded systems can be very beneficial for the energy footprint of future IoT systems.

One of the most promising approaches to remove the dependency of batteries is the harvesting of energy from naturally or artificially available environmental resources.

Another approach is to increase the energy efficiency by decreasing the battery power needed by IoT devices and embedded systems.

Further needed research activities include improved energy management approached to reduce the energy footprint of IoT devices and embedded IoT devices by enabling the use of both mentioned approaches: energy harvesting and increasing energy efficiency.

Research Priorities Timeline

Table 3 - *Energy-efficient intelligent IoT and edge computing systems research priorities*

Topic	Short Term	Medium Term	Long term
	2023-2024	2025-2027	2028-2030
Energy harvesting	Research on hybrid solutions combining ultra-low power connectivity with energy harvested from ambient radio frequencies (RF), thermal, kinetic, and photovoltaic (e.g., solar, and indoor/outdoor lighting) energy sources.	Multi energy harvesting, wireless power for edge IoT devices. Energy harvesting solutions at mesh network edge IoT devices.	Energy harvesting for edge IoT devices integrating positioning and sensing. Cognitive energy management orchestration in edge IoT systems for data processing energy optimisation.
Energy-efficient hardware	Research on the next-generation of energy harvesting ultra-low-power devices with on-demand wake-up feature integrated into edge IoT applications	Edge IoT devices base on printed electronics (e.g., conductive inks, metal etching, laser-direct structuring (LDS) for printable circuits and batteries) to be embedded in objects and products. Energy harvesting for edge IoT devices integrating machine-vision camera systems using AI and ML.	Research on energy-harvesting interfaces for kinetic energy harvesting from heterogenous generators (piezoelectric, triboelectric etc.).
Energy-efficient data processing	System-level optimisation techniques combining lower power consumption and energy harvesting technologies. E2E energy methods and models for data compression and exchange in edge-cloud IoT platforms.	Benchmarking methods for energy-efficient and low CO_2 footprint of edge IoT infrastructure and technical solutions.	Energy-efficient data aggregation mechanisms in intelligent edge IoT systems considering the associated processing capabilities across the computing continuum.

3. EU and national funded projects

OMEGA-X

The OMEGA-X project focuses on establishing an energy data space that ensures data sharing among various stakeholders in the energy sector. This initiative aims to enhance the availability and use of data for innovation, supporting the development of AI models and new business models, particularly for SMEs and startups. It emphasizes privacy, security, and data sovereignty, aiming to increase energy autonomy and reduce CO2 emissions.

PEDVOLUTION

The PEDvolution project focuses on the development and integration of Positive Energy Districts (PEDs) across Europe to enhance energy efficiency and sustainability. It involves analyzing, co-developing, and demonstrating interoperable solutions for the evolution of PEDs, including design tools, readiness assessments, and business models. The project aims for cross-sectoral integration, addressing the dynamic nature of urban energy transitions influenced by various factors like social context and legislation.

ODEON

ODEON introduces a sound, reliable, scalable and openly accessible federated technological framework (i.e. ODEON Cloud-Edge Data and Intelligence Service Platform and corresponding Federated Energy Data Spaces. AI Containers, Smart Data/AIOps orchestrators) for the delivery of a wealth of services addressing the complete life-cycle of Data/AIOps and their smart spawn in federated environments and infrastructures across the continuum. It will integrate highly reliable and secure federated data management, processing, sharing and intelligence services, enabling the energy value chain actors and 3rd parties to engage in data/intelligence sharing, towards the delivery of innovative data-driven and intelligence-powered energy services in accordance to the objectives set by the DoEAP. ODEON results will be extensively validated in 5 large-scale demonstration sites in Greece, Spain, France, Denmark and Ireland involving all required value chain actors, diverse assets, heterogeneous grid and market contexts, and multi-variate climatic and socio-economic characteristics to support its successful replication and market uptake.

GECKO

The GECKO project aims to support regulatory bodies in developing frameworks that accommodate emerging transportation technologies and business models. It focuses on facilitating adaptable and forward-looking regulatory approaches that contribute to sustainable mobility. By providing guidance, recommendations, and case studies, GECKO helps in the transition towards cooperative, inclusive, competitive, sustainable, and interconnected mobility across various modes. The project's emphasis on evidence-based research suggests a reliance on data analytics, where AI could play a crucial role in analyzing trends, predicting outcomes, and informing policy decisions.

iFLEX

The iFLEX project aims to empower energy consumers by simplifying their participation in demand response programs, adjusting energy consumption in response to various signals. It develops an intelligent personal assistant, the iFLEX Assistant, to manage system interactions, optimize consumer benefits like comfort, cost, and environmental impact, and automate daily operations. This project focuses on a user-centred approach, engaging consumers in design and considering motivational drivers. AI is central, developing based on AI technologies to manage and automate energy flexibility efficiently.

Onenet

The OneNet project aims to create a cohesive, integrated electricity network across Europe, optimizing the energy system and fostering an open, fair market. It emphasizes collaboration among electricity network actors to improve system efficiency and market structure. While the project's description doesn't explicitly mention AI, the integration and optimization goals suggest the potential use of AI technologies for data analysis, system management, and enhancing grid operations, aligning with the broader objective of creating a sustainable, efficient energy market across Europe.

EDDIE

The EDDIE project aims to foster the development of new data-driven services within and beyond the energy sector by establishing a decentralized, open-source data space. This initiative addresses the challenges posed by the lack of standardized procedures across the EU, which hinders interoperability and limits growth opportunities. By significantly reducing data integration costs, EDDIE enables energy service companies to operate seamlessly across a unified European market. AI can be leveraged within EDDIE to analyse energy data, optimize grid operations, and enhance decision-making processes, ensuring secure and reliable access to data based on customer consent.

Data Cellar

The DATA CELLAR project aims to create a federated energy dataspace to support the development and management of local energy communities (LECs) across the EU. By introducing an innovative rewarded private metering approach, it focuses on simplifying onboarding and interaction while ensuring smooth integration with other EU energy data spaces. AI plays a critical role in this project by providing access to AI models and energy services, facilitating advanced data analytics, and evaluating new business models within the energy sector.

Synergies

The SYNERGIES project aims to transform the energy sector by creating a data-driven intelligence ecosystem. It focuses on leveraging data for innovation, improving operational efficiency, and engaging prosumers in the market. The project addresses challenges in transitioning to a renewable-based and decentralized system by integrating diverse data sources and technologies. SYNERGIES' solution includes promoting knowledge sharing, integrating data intelligence among all energy actors, and implementing digital twins for network operators. It represents a collaborative effort with 23 European partners to decarbonize energy systems and empower consumers.

Enershare

The Enershare project focuses on uniting the energy and data value chains to facilitate the energy transition in Europe. It aims to implement and validate its framework through 7 real-life pilots and 11 use cases across 7 countries, promoting EU-wide replicability of energy data sharing technologies and data-driven services. This approach ensures adaptability to various socio-economic contexts. AI could be instrumental in analyzing energy data across these pilots, optimizing energy distribution, and enhancing the overall efficiency of the energy sector.

BD4NRG

BD4NRG artefacts combines DLTs/blockchain technologies with edge processing, Federated Machine Learning and Artificial Intelligence, to operate the data-driven energy ecosystem. Also, the project will make extensive adoption of open sources technology components and tools and Open APIs.

Smart Grid 2.0

The Smart Grid 2.0 project is a Finnish national project aims at investigating new approaches to sustain the reliability of the electricity grid in light of energy decarbonization shift which is changing energy generation into renewables, which have less inertia and thus affect the capability of the transmission power network more vulnerable to power imbalances. The project investigates how emerging AI approaches can be used to produce more accurate model of energy assets, both at the generation and consumption side, predict and forecast the state of power balances and flexibility.

INDECS

INDECS project is a Finnish national project aiming to investigate and develop solutions to reduce the fuel consumption and the greenhouse gas emissions on cruise ships. This is done by design improvements on the HVAC (Heating, Ventilation and Air Conditioning) system, in combination with optimal operation of all hotel functionalities, propulsion and the engines. Significant improvements can be done by considering each part separately, and further improvements are possible by taking a holistic view on the problem, as the different parts are not independent. INDECS will also provide a simulation-based test bench for engines of different sizes and types, which can speed up the introduction of new engine types with very low GHG emissions to cruise ships.

NATWORK

NATWORK is a EU HE SNS 2024 funded project under call B-01-04, investigating energy sustainability of 6G cybersecurity frameworks enabled by AI. The project focuses on working towards realising net-zero targets in AI-enabled 6G cybersecurity. It will develop energy-aware/-efficient AI cybersecurity solutions for providing security-by-design cybersecurity solutions in 6G. The project will analyse, benchmark and optimise the application of AI in 6G cybersecurity with respect to the energy consumption of such AI services. Various AI architectures: central, distributed, federated will be explored and compared in various use-cases. Moreover, the project will investigate the vulnerability of AI components to attacks that might impact its energy profile (consumption, CO2 footprint due to operation on green/brown energy, and others). The project has a set of use-cases that will demonstrate the performance of AI and its energy profile. In particular, "*use-case1: Sustainability and reliability of 6G Slices and services*" will demonstrate the project innovation in defending against Economic Denial of Sustainability of 6G services, with particular focus on AI-based 6G services. Because, of their higher vulnerability to energy-oriented sustainability attacks, being resource-intensive services.

Irish Energy Poverty Observatory (IREPO)

The IREPO project, funded by the Sustainable Energy Authority of Ireland (SEAI), aims to go beyond the conventional factors that define energy poverty by developing a federated database that consolidates all relevant factors impacting energy poverty to overcome data limitations by integrating various data sources in near-real time, including surveys, interviews, historical databases, census data, and more. Using Machine Learning (ML) and Artificial Intelligence (AI) algorithms, IREPO will identify patterns within these datasets, allowing for data-driven decision-making and policy support. By combining diverse data sources such as energy commodity fluctuations, renewable energy data, and economic and social indicators, IREPO will provide a comprehensive understanding of the risk factors and impacts of energy poverty. The IREPO platform's functionality will improve over time as the ML and AI benefit from the accumulation of historical data. Continuous data integration will ensure timely and standardized data in near-real time, supporting powerful decision-making tools and simulations. IREPO aims to become a reference tool for decision-making, data-driven policy support, and simulations in the field of energy poverty. By fostering collaboration among academia, industry, policymakers, and other stakeholders, IREPO seeks to have a lasting impact in addressing energy poverty and its social implications.

BD4OPEM

Technology offers a huge range of opportunities for the energy market. By combining the old metering, operation and control devices with smart systems, there's a huge amount of data available that is unused or underused. This data offers a range of possibilities to improve existing energy services and create new ones. The EU-funded "Big Data for Open innovation Energy Marketplace" (BD4OPEM) project develops an analytic toolbox based on Big Data techniques, providing tools for enabling efficient business processes in the energy sector. By extracting more value from available data, a range of innovative services will be created in the fields of grid monitoring, operation and maintenance, network planning, fraud detection, smart energy management for houses/buildings/industries, blockchain transactions and flexibility aggregation for demand-response in smart grids.

E-LAND

The continued decarbonization of the energy sector through the use of renewable energy sources provides both interesting opportunities for local energy systems and challenges for existing electricity networks. Mainland regions such as isolated villages, small cities, urban districts or rural areas oftentimes have issues with weak or non-existing grid connections. These areas are known as energy islands. The goal of the European-funded H2020 project "Integrated multi-vector management system for Energy islands" (E-LAND) is to provide a synergistic solution between technological, societal and business challenges that the energy sector faces. The main concept is the E-LAND toolbox – a modular set of methodologies and ICT tools to optimize and control multi energy islands and isolated communities. The modular toolbox can be customized to meet local requirements and expandable to incorporate new tools as new challenges arise.

HEDGE-IoT

HEDGE-IoT aims to empower the digitalization of the Energy Ecosystem and to facilitate increased integration of renewable energy sources through an holistic approach. The project will implement a novel Digital Framework which will use different IoT assets (from behind-the-meter, up to the TSO level), to add intelligence to the edge and cloud layers through advanced AI/ML tools and to bridge the cloud/edge continuum.

The project will exploit the computational sharing by offloading applications on the grid edge, towards providing a set of AI/ML federated learning and swarm computing applications. The two other very important aspects of the project are the interoperability pillar through data space, and the standardization pillar for data exchange through semantic ontology and widely used standards. AI systems are central in this project and are key components in most of the use cases. Project results will be implemented and validated among 6 pilots in Finland, Greece, Italy, Netherlands, Portugal, and Slovenia.

A practice on AI trustworthiness is under development to support pilots in their development and ensure that their AI systems are adequately developed and managed throughout the project and throughout the AI lifecycle. This practice takes into account regulation and standardization. At this stage it covers the following high level aspects: AI system information, AI impact assessment, risk management, data/algorithm/model information and quality, regulatory compliance and standards alignment, AI trustworthiness, and continuous improvement. This practice will be included in a deliverable scheduled in March 2025.

SUNRISE-6G

6G represents the future of technology, promising exceptional performance through innovative access technologies. It is poised to prompt a complete re-evaluation of network architecture design, with new stakeholders entering the value chain of future networks. The EU-funded SUNRISE-6G project aims to develop a massively scalable internet-like architecture for all public and private infrastructures. Additionally, it seeks to establish a pan-European facility to support converged workflows and tools, offering a unified catalogue of 6G enablers and facilitating cross-domain vertical application onboarding through a Tenant Web Portal. The project will focus on four pillars: implementing new 6G enablers, a compliant Federation solution, a Federated AI plane, and a commonly adopted Experimentation Plane.

INTENSE

The 6G Smart Networks of the future will provide the high-performance and energy-efficient infrastructure on which next generation internet and other services can be developed and deployed. 6G will foster an Industry revolution and digital transformation and will accelerate the building of smart societies leading to quality-of-life improvements, facilitating autonomous systems, haptic communication and smart healthcare. To achieve the aforementioned objectives in a sustainable way, it is well understood that new approaches are needed in the way the telco infrastructures are architected, federated and orchestrated. These new approaches call for multi-stakeholder ecosystems that promote synergies among MNOs and owners of all kinds or computational and networking resources that will share the extraordinary costs of yet another generation upgrade from 5G to 6G, while facilitating new business models. It is clear, that the new architecture paradigms bring unprecedented complexity due to the sheer scale and heterogeneity of the orchestration domains involved, that should be matched by equally capable automation capabilities, thus 6G is aiming for the "holy grail" of pervasive AI-driven intelligence, termed as Native AI.

However, the multi-stakeholder infrastructures foreseen in 6G as per the "network of networks" concept, will add a level of unprecedented management complexity due to the sheer scale and heterogeneity of the orchestration domains involved. 6G-INTENSE aims to abstract and federate all kinds of computational and communication resources under an internet-scale framework, that is governed by an intelligent orchestration paradigm, termed as DIMO.

6G-BRICKS

6G-BRICKS is a Research and Innovation project in the 6G Smart Networks and Services Joint Undertaking of Horizon Europe. Its aims to deliver a new 6G facility, which builds on the baseline of mature "Smart Connectivity Beyond 5G" projects such as to bring to life breakthrough technologies, i.e., cell-free and RIS technologies, that will play a key role in beyond 5G networks. Moreover, novel unified control paradigms based on Explainable AI and Machine Reasoning will be explored. All enablers will be delivered in the form of reusable components with open APIs, termed "bricks". Finally, initial integrations with O-RAN will be performed, aiming for the future-proofing and interoperability of 6G-BRICKS outcomes. In 6G-BRICKs Intracom Telecom will lead the implementation of a resource management and abstraction framework to offer intelligent management of edge cloud resources and services, extended seamlessly to Far Edge and IoT devices. In addition, Intracom Telecom will contribute to the demonstrator with a disaggregated X-Haul solution based on its UltraLink™ XR, which, leveraging its hardware capabilities, will serve as an additional MEC hosting platform for the optimized deployment of 6G-BRICKS cloud native applications.

ARIADNE

ARIADNE, a "5G PPP Phase 3, Part 4: 5G Long Term Evolution" project, plans to bring together a novel high frequency advanced radio architecture and an Artificial Intelligence (AI) network processing and management approach into a new type of intelligent communications system Beyond 5G. The new intelligent system approach is necessary because the scale and complexity of the new radio attributes in the new frequency ranges cannot be optimally operated using traditional network management approaches. The ARIADNE project will enable efficient high-bandwidth wireless communications by developing three complementary but critical new technologies for Beyond 5G networks in an integrated and innovative way: ARIADNE will develop new radio technologies for communications using the above 100GHz D-Band frequency ranges, (Pillar 1), will exploit the opportunities emerging for advanced connectivity based on meta-surfaces where objects in the environment can become tuneable reflectors for shaping the propagation environment in D-band (Pillar 2) and will employ Machine Learning and Artificial Intelligence techniques to management necessary for the high frequency communications and dynamic assignment and reconfiguration of the meta-surfaces to provide continuous reliable high bandwidth connections in the Beyond 5G scenario (Pillar 3).

5G-VICTORI

5G-VICTORI is addressing the well-defined European objective of providing 5G solutions for verticals. In order to give verticals, the opportunity to verify their use cases in large scale deployments, 5G-VICTORI will conduct large scale trials for advanced vertical use case verification, making use of infrastructures made available by 5G PPP Phase 3 Infrastructure projects (5G-VINNI, 5GENESIS and 5G-EVE) as well as the advanced national UK test-bed 5GUK, focusing on Transportation, Energy, Media, Factories of the Future and cross-vertical use cases. It will employ 5G network technologies developed in 5G-PPP phase 1 and 2 projects 5G-XHaul and 5GPICTURE and will exploit extensively existing facilities interconnecting main sites of all involved infrastructures, enhancing them towards integration of a large variety of vertical and cross-vertical use cases.

5G-VICTORI's platform aims to transform current closed, purposely developed and dedicated infrastructures into open environments where resources and functions are exposed to ICT and vertical industries through common vertical and non-vertical specific repositories. These functions can be accessed on demand and deployed to compose very diverse sets of services.

VirtuWind

SDN and NFV have been originally proposed as architectures for the efficient management of clouds and have recently found their way into the networks of ISPs as well. It is not clear however, how SDN and NFV affect the QoS of networks in general and of industrial networks (manufacturing, oil & gas, building & construction, etc.) in particular. VirtuWind aims at introducing SDN and NFV into such networks so that deterministic (or almost-deterministic) performance is achieved for monitoring and control applications that span multiple network domains. A prototype system will be implemented and will be tested in a real wind park, as a representative example of industrial networks. Intracom Telecom leads the design and development of the inter-domain system.

5GZORRO

5GZORRO envisions the evolution of 5G to achieve truly production-level support of diverse Vertical applications, which coexist on a highly pervasive shared network infrastructure. 5GZORRO uses distributed Artificial Intelligence (AI) for Zero-Touch Automation in end-to-end network slicing, across multiple operators and infrastructure/resource providers. Distributed Ledger Technologies (DLT) are implemented for distributed security and trust across the various parties in a 5G end-to-end service chain. An evolved 5G Service Layer for Smart Contracts allows SLA monitoring, spectrum sharing, intelligent and automated data-driven resource discovery and management in multi-tenant and multi-stakeholder environments. 5GZORRO will be validated on use cases of Smart Contracts, Dynamic Spectrum Allocation and Pervasive vCDN in 5GBarcelona and 5TONIC/Madrid test facilities. The consortium consists of 13 top 5G players from 7 different EU countries.

FEVER

FEVER implements and demonstrates solutions and services that leverage flexibility towards offering electricity grid services that address problems of the distribution grid, thus enabling it to function in a secure and resilient manner. The project encompasses technologies and techniques for extraction of energy flexibility from energy storage assets and implements a comprehensive flexibility aggregation, management and trading solution. In addition, a DLT-based flexibility trading toolbox will be implemented enabling autonomous peer-to-peer trading. FEVER also implements goal-oriented applications and tools that empower DSOs with optimal grid observability and controllability. FEVER will carry out extensive demonstration and testing activities in multiple settings. Intracom Telecom is the Project Coordinator of FEVER. From a technical viewpoint, it provides the platform that integrates novel (i.e. implemented in the project) and legacy systems of the Distribution System Operator, as well as a permissioned blockchain network on top of which p2p flexibility trading services will be running.

OMEGA-X

Large amounts of valuable data are available in energy systems but are often underused, hindered by the lack of interoperability and of proper mechanisms and policies that ensure secure, sovereign and fair data sharing. Relying on European common frameworks, such as those of GAIA-X and IDSA, the EU-funded OMEGA-X project aims to implement an energy data space for "unlocking" data sharing between different stakeholders of the energy domain and demonstrating its value.

The project will develop the federated infrastructure as well as data and service marketplace, enabling four different use cases families: Renewables, Local Energy Communities, Electromobility and Flexibility. In this project Intracom Telecom leads the architectural design of the Data Space and development of the marketplace.

BD4OPEM

Technology offers a huge range of opportunities for the energy market. By combining the old metering, operation and control devices with smart systems, there's a huge amount of data available that is unused or underused. This data offers a range of possibilities to improve existing energy services and create new ones. The EU-funded BD4OPEM project develops an analytic toolbox based on Big Data techniques, providing tools for enabling efficient business processes in the energy sector. By extracting more value from available data, a range of innovative services will be created in the fields of grid monitoring, operation and maintenance, network planning, fraud detection, smart energy management for houses/buildings/industries, blockchain transactions and flexibility aggregation for demand-response in smart grids. In this project Intracom Telecom is responsible for data integration and management as well as for implementation of integrated energy data and analytics services' marketplace platform.

E-LAND

The continued decarbonization of the energy sector through the use of renewable energy sources provides both interesting opportunities for local energy systems and challenges for existing electricity networks. Mainland regions such as isolated villages, small cities, urban districts or rural areas oftentimes have issues with weak or non-existing grid connections. These areas are known as energy islands. The goal of the European-funded H2020 project E-LAND is to provide a synergistic solution between technological, societal and business challenges that the energy sector faces. The main concept is the E-LAND toolbox – a modular set of methodologies and ICT tools to optimize and control multi energy islands and isolated communities. The modular toolbox can be customized to meet local requirements and expandable to incorporate new tools as new challenges arise. In E-LAND Intracom Telecom is responsible for implementing an integration middleware to facilitate the interplay among the various tools of the E-LAND toolbox. Also, it implements a web-based energy planning application serving as front-end of a multi-vector energy simulator.

DSO Toolbox

Modern DSOs are in need of novel tools and methodologies that will enable them: to perform cost / benefit analysis regarding the incorporation of different technologies and the adoption of various emerging added-value services into the Distribution Grid modus operandi; to decide on the technically credible and economically viable operation of the Distribution Grid, especially in light of the presence of a large number of intermittent production units and while operating in island mode; to decide on the composition and functional characteristics of the various types of production units, as well as to the planning, scheduling, operation, supervision and control of the various assets of the Distribution Grid; and, to dynamically shape the load-curve in order to respond to the production conditions and dynamic conditions of the Distribution Grid. In the light of the above, the aim of the project is to develop a toolbox that will facilitate DSOs in the fulfilment of their tasks, both in the planning and at the operational phase. Intracom Telecom is the Coordinator of DSO Toolbox. From a technical perspective it participates in the implementation of a tool for techno-economic analysis of the impact of incorporating different technologies and added-value energy services into the Distribution System modus operandi; a tool for the operation of the Distribution Network at emergency conditions; and a tool for power quality monitoring in a Distribution Network.

Finally, Intracom Telecom implements an Enterprise Service Bus for information? and services? integration and interoperation within the DSO?s IT ecosystem.

iFLEX

Consumer response is a crucial factor for the economy. It allows consumers to influence prices through demand and thus lower or increase them, and is based on the needs, wants and economic capabilities of the public. It is also especially important for the energy sector. Unfortunately, consumer response and demand response cannot easily influence demand for the energy sector. The EU-funded iFLEX project aims to change this, not only by making it easier for consumers to participate in demand response but also by increasing its reach and effect. It plans to achieve this by developing an innovative software agent that allows for better management of energy and demand response by acting between energy systems and various stakeholders. In this project Intracom Telecom is responsible for implementing a flexibility aggregation and management solution incorporating customizable incentive and aggregation mechanisms and data-efficient deep learning methods for modelling and optimization of consumer flexibility management. It also focuses on mechanisms for user engagement with natural user experience and incentives.

GIFT

GIFT is an innovative project that aims to decarbonise the energy mix of European islands. Through the development of multiple innovative solutions, such as a virtual power system, energy management systems for harbours, factories, homes, better prediction of supply and demand and visualisation of those through a GIS platform, GIFT will increase the penetration rate of renewable energy sources into the islands' grid, reducing their needs for diesel generation and thus decreasing the greenhouse gases emissions directly related to it. Intracom Telecom is responsible for implementing an integration middleware to facilitate the interplay among the various different applications and services built in the project.

RESOLVD

RESOLVD is a 3-years collaborative project, co-funded by European Commission, under Grant Agreement 773715. The project aims to improve the efficiency and the RES hosting capacity of the electricity distribution networks, in a context of highly distributed renewable generation and storage, by proposing innovative Distributed Storage Systems that enhance flexibility and control in the low voltage grid, as well as hardware and software solutions that improve its observability with wide area monitoring capabilities, demand/production forecasting and automatic fault detection and isolation. Intracom Telecom's role is to provide a platform that integrates legacy systems of the Distribution System Operator (e.g. SCADA, AMI) and newly developed hardware and software applications, and handles data management and analytics processes.

SmarterEMC2

The project SmarterEMC2 implements ICT tools that support Customer Side Participation and RES integration, and facilitate open access in the electricity market. These tools take into account the SGAM as well the future structure of the Distribution Network as described by the relevant EU bodies and organizations. The project supports standardization activity by proposing adaptation to data models of market-oriented standards (e.g. IEC 62325-351) and field level standards (e.g. IEC 61850). Intracom Telecom is the coordinator of this project, and is responsible for the implementation of the ICT infrastructure related to a Demand Response service.

4. AI in various energy sectors

Energy systems consist of processes and artifacts for the conversion, transmission and utilization of energy, combined to meet a specific need. The energy sector itself is diverse and includes a wide range of other systems covering a wide range of industries and activities related to the production, distribution and consumption of energy. Due to continuous technological advancements and an increasing emphasis on sustainability, the energy sector landscape continues to evolve rapidly. The benefits of AI in the energy sector are many.

One of the most promising applications of AI is the deployment of AI algorithms that can optimize the integration of intermittent renewable energy sources such as solar and wind into the grid by forecasting generation, predicting output fluctuations and managing energy storage systems. Compared to smart grids, traditional grids have limited possibilities to manage, integrate and optimize energy production and energy consumption. Energy storage is crucial to enable a better utilization of variable energy resources through, among other things, energy storage optimization. In this case, AI can play a pivotal role in integrating intermittent renewable energy sources into the traditional grid by improving forecasting, optimizing operations, improving demand response, managing energy storage, and increasing the overall resilience and reliability of the grid.

Additionally, AI is used to optimize the charging and discharging of energy storage systems such as batteries based on factors such as energy prices, demand patterns and weather forecasts, thereby maximizing their efficiency and lifespan. In the context of smart grids, AI algorithms are important in optimizing energy distribution and managing demand more efficiently, leading to improved reliability and efficiency. This is done by analysing data from various sources, including sensors, weather forecasts and energy consumption patterns.

This is important for companies to balance supply and demand more effectively, reduce losses and increase integration of renewable energy sources. In terms of security, AI systems monitor the grid for abnormalities and potential cyber threats, detect and respond to security breaches in real time, resulting in improved grid security and resilience. In the planning phase, the use of AI algorithms makes it possible to identify the optimal location for renewable energy projects related to variable energy sources such as solar and wind by analysing geographic and environmental data.

On the demand and supply side, demand and supply forecasts are very important by enabling energy supplier and grid operators to better plan for variations and optimize resource allocation. Here, AI techniques can be used to forecast demand and supply of energy and thus facilitate improved and optimized resource allocation and improved plan for fluctuations. AI technologies can also be used in the field of demand response management by controlling, monitoring and adjusting consumers' energy consumption in response to supply conditions, thereby contributing to reduced peak loads, improved grid reliability, optimized energy consumption and reduced costs for both suppliers and consumers. and avoid costly infrastructure investments.

In the field of predictive maintenance which is highly important for continuous operation of generation, transmission and distribution components, AI is used to predict when equipment that is part of the energy system will fail, enabling preventive maintenance. In this case AI algorithms can analyse the data from sensors and historical maintenance records to predict when equipment is likely to fail.

This enables preventive maintenance, which reduces downtime, maintenance costs and increases efficiency. In daily operation, it is important to have fault detection system. AI algorithms detect faults and inefficiencies in energy systems, contributing to improved overall performance, reliability and reduced maintenance costs.

In terms of energy trading, which is an important part of cost-effective energy systems, AI algorithms are used in energy trading platforms to analyse market data connected to supply and demand dynamics, energy prices, policy instruments and regulatory constraints, to make trading decisions in the energy markets. This is beneficial for utilities and traders optimize profits and manage risk.

4.1 AI for Energy-Efficient Production Equipment

The manufacturing sector, as one of the world's largest energy consumers, faces an urgent need to address energy waste amid rising industrial energy demands and global economic growth. Energy efficiency has thus become essential—not only as an environmental and regulatory focus but also as a critical factor in cost reduction and competitive advantage. Achieving high efficiency in manufacturing requires optimizing production equipment to reduce non-value-adding tasks, cut excess supply, lower base-load energy demands, and fine-tune process parameters, all while maintaining or improving productivity [Zhou et al. 2016, Sihag et al. (2020)].

Energy management plays a central role in the sustainability vision of Industry 5.0, which addresses the entire product lifecycle, from sustainable design to end-of-life circularity. This comprehensive approach emphasizes not only minimizing the environmental impact of products but also enhancing the operational efficiency of production sites. Effective energy management in manufacturing facilities, involving the precise measurement and control of energy usage across systems and human operations, is critical for reducing costs, improving processes, and lessening environmental impact, ultimately contributing to a sustainable and resilient industrial ecosystem.

This chapter focuses on how AI has been recently exploited to support a more efficient management of production systems by providing practical examples of applications addressing some of the key challenges and strategies for energy efficiency in manufacturing. The key lever for AI is to provide advanced models supporting Smart Energy Management Systems to ensure that machinery and processes are respecting energy efficiency criteria and KPIs.

4.1.1 Key focus areas and levers for Energy Efficiency in Manufacturing

Improving energy efficiency in manufacturing involves several strategic approaches that focus on optimizing both the design and operation of machinery and processes, e.g. Renna and Materi (2021). In more details:

Machine and process design. The design of machines and processes plays a critical role, as machinery built with energy efficiency in mind reduces weights and friction, applies more efficient energy transformation and minimizes waste; thus reducing the overall energy input while providing the same output. Additionally, energy and heat recovery systems are deployed to capture and reuse waste energy generated during production, reducing overall consumption and lowering costs. Similarly, process re-design aims at developing new alternative, and more energy efficient, processes or at finding the optimal combination of process parameters to minimize energy consumption.

Green scheduling and planning. Production planning also contributes significantly, with green scheduling techniques that minimize energy use by aligning production schedules with lower-cost, off-peak energy periods, e.g. Gahm et al. (2016). These efforts might be coordinated with multiple energy sources and variable energy prices adapting the production toward increased energy flexibility.

Energy Efficient Control. Similarly to green scheduling and planning, low-power modes of manufacturing equipment are used in real-time, adapting to the current state of the system. Off/on policies, sleeping modes, energy opportunity windows and resource hibernation are examples of such approaches, e.g. Brundage et al (2014), Frigerio and Matta (2015), Jia et al. (2016), Frigerio et al. (2023).

To enhance efficiency further, manufacturers employ various levers within their operations. Accurate metering, monitoring, and modelling of energy consumption provide the real-time data necessary for informed decision-making. By continuously tracking energy use and identifying deviations, manufacturers gain insight into the fluctuating and stochastic nature of energy needs across production cycles. Furthermore, learning and production control systems leverage data-driven modelling and optimisation to predict energy demands, tailoring production rates and resource allocation under specific constraints. This approach enables manufacturers to make multi-level, multi-objective decisions that improve efficiency, reduce waste, and support sustainable production goals.

The integration of Artificial Intelligence (AI) has improved these efforts, enabling more sophisticated approaches to both equipment design and operational management. A critical area where AI contributes is in the monitoring, and modelling of power signature and energy consumption essential for predicting energy requirements. AI-powered learning and production control systems further enhance energy management. The major challenges for energy management for manufacturing include:

Reducing energy costs while maintaining productivity.

Improving energy efficiency without compromising system performance.

Increasing user awareness about the environmental impacts of manufacturing.

Overcoming barriers to integrating energy management with existing production planning and control strategies.

4.2 The evolution of Energy-Efficient Control (EEC)

Energy-Efficient Control (EEC) refers to a set of strategies aimed at minimizing energy consumption during non-productive machine states. These strategies focus on optimizing machine operation by dynamically adjusting power modes—such as switching machines from operational to standby or off when they are idle. This approach ensures that energy is not wasted when machines are not directly involved in production. Several strategies have been proposed and implemented to improve energy efficiency in manufacturing systems, including:

Buffer-Based Control: This strategy relies on monitoring buffer occupancy to optimize machine states, e.g. Frigerio et al. (2023).

Time-Based Control: These approaches switch machines on or off based on pre-defined idle times, ensuring that machines only operate when necessary, e.g. Frigerio and Matta, 2015.

Each of these strategies offers a different approach to energy efficiency, with varying levels of complexity and applicability depending on the manufacturing setup.

An example of application can be found in Loffredo et al. (2023), where extended EEC strategies are applied to industrial parts washers, a process that is particularly energy-intensive in the automotive sector. Using discrete event simulations, an automotive production line is modelled under different policies. The results show significant energy savings by implementing standby modes during idle periods, without negatively affecting production times.

Unlike traditional offline policies, this AI-driven approach allows the system to learn and adapt as new data is acquired. The control strategy proposed in [Frigerio et al., 2021] is based on online learning from real-time shop floor data; thus the algorithm continuously monitors part arrivals and machine states, dynamically adjusting control parameters to optimize energy efficiency. The AI model predicts part arrival times and uses this information to delay the switch-off decision if the next part is expected soon. This ensures that the machine does not incur unnecessary energy costs from frequent transitions between states. The AI-based control agent continuously updates the control parameters by interacting with the manufacturing environment, learning from real-time data, and balancing energy savings with production throughput and, to deal with non-stationary production environments, achieving a balance between minimizing energy consumption and maintaining throughput. In simulation experiments, the proposed AI-driven policy achieved energy savings up to 25% compared to the Always-On policy.

4.3 Reinforcement Learning to Optimize Energy Efficient Control

Machine learning, and in particular reinforcement learning (RL), offers powerful tools for optimizing energy efficiency in manufacturing systems. RL involves training an agent that interacts with the environment (in this case, a manufacturing system) and learns an optimal control policy by receiving rewards based on its actions (in this case, an EEC policy) [Loffredo et al. 2023b].

As breakthrough example, Loffredo et al. (2023b) used RL to optimize energy efficiency in parallel machine production lines, where identical machines perform the same task in parallel. In such systems, the RL agent learns to manage the number of machines that should be active at any given time, depending on the workload and production demands. This ensures that only the necessary machines are consuming energy, while others remain in standby mode. In a real-world case study in the automotive sector, RL was used to manage the operation of identical parallel machines involved in the production of powertrains.

In the real case example presented in Loffredo et al. (2023b), the EEC policy was applied to a manufacturing workstation in the automotive sector, specifically involving the production of cylinder heads. The implementation of the RL agent resulted in an energy saving of 7.72% compared to the AS-IS scenario while the throughput loss was limited to 1.20%, indicating that the RL-based policy maintained near-optimal production rates while achieving energy savings. The RL agent demonstrated fast convergence to an optimal policy, requiring less than 2 hours of simulated production time to learn the optimal control actions. The results highlight the effectiveness of RL-based control.

4.4 Data-driven for Energy Monitoring and Modeling

The key problem addressed in this section focuses on the accurate modelling and predicting the energy consumption of manufacturing machines, particularly CNC machining centres, in a way that is flexible, data-driven, and does not require extensive prior knowledge or costly, time-consuming experiments.

On one side, we have the challenge of classifying machine energy states during operation to develop energy models that are adaptable to varying machine conditions. The traditional methods require extensive pre-configured data and experiments. On the other side, the estimation of energy consumption using minimal prior information about the machines, parts, or processes is also challenging. Traditional methods for energy modelling are resource-intensive, requiring detailed machine-specific data and testing. In both cases, the problem is centred on the need for a more automated, flexible, and generalisable method for energy modelling in manufacturing, leveraging AI and machine learning to overcome the limitations of traditional empirical and analytical approaches.

Recent developments in machine learning (ML), particularly in classification algorithms, have enabled more precise identification of machine energy states. Frigerio et al. (2023b) demonstrated the application of MLA to classify the operational states of machines based on power signal data may help manufacturers in better understand when machines are consuming excessive energy, allowing them to implement appropriate control actions.

A more advanced data-driven approach was developed by Lestingi et al. (online 2024), who introduced a method for modelling energy consumption in CNC machining centres through automata learning. This method provides manufacturers with the ability to estimate energy consumption for specific tasks or part-programs, offering greater flexibility and accuracy in controlling machine power states.

4.5 Conclusions and Beyond

The shift toward energy-efficient manufacturing presents both challenges and opportunities for industries globally. As Industry 4.0 technologies continue to advance, IoT (Internet of Things) and digital twins will play a pivotal role in energy optimisation. By integrating real-time sensor data with digital models, manufacturers can create accurate simulations to predict energy consumption and fine-tune control systems in real time.

The application of AI in industrial energy management systems promises to reduce energy costs, improve efficiency, and provide real-time performance insights. With AI-driven models and control systems, manufacturers can make informed decisions that promote sustainable and cost-effective production processes. For example:

Real-time data continuously updates digital twins, keeping system models accurate and aligned with physical processes, improving monitoring, predictions, and optimisation.

Machine learning enables precise real-time recognition of machine states, enhancing decision-making for actions like reducing energy use during idle times and optimising production.

AI-driven optimisation algorithms, like reinforcement learning, balance energy efficiency and productivity by learning optimal strategies in dynamic, multi-goal environments.

References:

Brundage, M., Q. Chang, Y. Li, G. Xiao, and J. Arinez. 2014. “Energy Efficiency Management of An Integrated Serial Production Line and Hvac System.” IEEE Transactions on Automation Science and Engineering 11: 789–797. doi:10.1109/TASE.2013.2284915.

Frigerio, N., and A. Matta. 2015. “Energy-efficient Control Strategies for Machine Tools with Stochastic Arrivals.” IEEE Transactions on Automation Science and Engineering 12:50–61. doi:10.1109/TASE.2014.2344507.

Frigerio, N., C. F. Cornaggia, and A. Matta (2021). An adaptive policy for on-line energy efficient control of machine tools under throughput constraint. Journal of Cleaner Production 287, 125367.

Nicla Frigerio, Barış Tan & Andrea Matta (2023): Simultaneous control of multiple machines for energy efficiency: a simulation-based approach, International Journal of Production Research, DOI: 10.1080/00207543.2023.2175175

Frigerio, N., Albertelli, P., Matta, A. Data-Driven State Classification for Energy Modeling of Machine Tools Using Power Signals and Part-Program Information (2023b). In: Kayakutlu, G., Kayalica, M. ¨O.(eds) Decision Making Using AI in Energy and Sustainability. Applied Innovation and Technology Management. Springer, Cham. DOI: 10.1007/978-3-031-38387-8 10

Gahm, C., Denz, F., Dirr, M. and Tuma, A. (2016) Energy-efficient scheduling in manufacturing companies: A review and research framework. European Journal of Operational Research, 248(3), 744–757.

Jia, Z., L. Zhang, J. Arinez, and G. Xiao. 2016. “Performance Analysis for Serial Production Lines with Bernoulli Machines and Real-time Wip-based Machine Switch-on/off Control.” International Journal of Production Research 54:6285–6301. doi:10.1080/00207543.2016.1197438.

L. Lestingi, N. Frigerio, M. M. Bersani, A. Matta and M. Rossi, (2024) "Data-Driven Energy Modeling of Machining Centers Through Automata Learning," in IEEE Transactions on Automation Science and Engineering, doi: 10.1109/TASE.2024.3430394.

Alberto Loffredo, Nicla Frigerio, Andrea Matta, Energy efficiency improvement of industrial parts washers using state control, Manufacturing Letters, Volume 35, Supplement, 2023, Pages 1303-1311, doi:10.1016/j.mfglet.2023.08.022.

Alberto Loffredo, Marvin Carl May, Louis Schäfer, Andrea Matta, Gisela Lanza, Reinforcement learning for energy-efficient control of parallel and identical machines, CIRP Journal of Manufacturing Science and Technology, Volume 44, 2023(b), Pages 91-103. doi:10.1016/j.cirpj.2023.05.007.

Renna, P., and S. Materi. 2021. “A Literature Review of Energy Efficiency and Sustainability in Manufacturing Systems.” Applied Sciences 11: 7366. doi:10.3390/app11167366.

Sihag, N., and K. S. Sangwan. 2020. “A Systematic Literature Review on Machine Tool Energy Consumption.” Journal of Cleaner Production 275: Article ID 123125. doi:10.1016/j.jclepro.2020.123125.

Zhou, L., J. Li, F. Li, Q. Meng, J. Li, and X. Xu. 2016. “Energy Consumption Model and Energy Efficiency of Machine Tools: a Comprehensive Literature Review.” Journal of Cleaner Production 112: 3721–3734. doi:10.1016/j.jclepro.2015.05.093.

5. AI for Marketplaces

5.1 AI agents

An AI agent can be described as a digital entity that continually evaluates its surroundings, learns from interactions, and makes decisions to accomplish specific goals. This encapsulates the essence of an AI agent, functioning akin to a digital assistant that carries out tasks, comprehends context, adapts strategies, and even devises new approaches to achieve its objectives. The spectrum of AI agents varies from simple programs designed for singular tasks to intricate systems overseeing complex processes.

In contrast to traditional automation systems, which operate based on pre-set scripts and have limited scopes, AI agents demonstrate a higher level of intelligence and adaptability. They excel in unpredictable settings by leveraging their flexibility and capacity for learning. This dynamic nature empowers them to navigate the web, interact with applications, process vast volumes of data, and even partake in transactions, all the while evolving their approaches based on feedback and outcomes.

The emergence of AI agents represents a significant leap towards achieving Artificial General Intelligence (AGI) – the stage at which machines can undertake any intellectual task that a human can. Even though AGI is still a futuristic goal, today's AI agents are already making a substantial impact by providing increasingly dynamic and intelligent solutions. They have become essential partners in businesses' pursuit of innovation, efficiency, and improved customer experience.

The essential traits of AI agents encompass autonomy, adaptability, learning, and decision-making. These agents function independently, minimizing the need for continual human intervention. They dynamically adjust their actions based on real-time data and feedback. Through the utilization of machine learning methodologies, AI agents iteratively enhance their performance. They possess the capability to make well-informed decisions to optimize processes and attain predetermined objectives.

The incorporation of AI agents into the energy sector signifies a revolutionary opportunity. AI agents excel in analysing extensive datasets to optimize energy generation and distribution. By accurately forecasting demand and promptly adjusting output, they contribute to maintaining a harmonious balance between energy supply and demand, thus reducing wastage and enhancing overall efficiency. Furthermore, AI agents play a crucial role in monitoring the condition of machinery and infrastructure, effectively predicting potential failures. This proactive approach in maintenance leads to reduced downtime and prolonged equipment lifespan, ultimately resulting in substantial cost savings.

Within the dynamic landscape of energy trading markets, AI are able to analyse market trends, forecast prices, and execute trades independently. Their capacity to process vast volumes of data provides them with a notable competitive advantage. AI agents able to play a pivotal role in mitigating the variability of renewable energy sources, such as wind and solar. By predicting fluctuations in energy production, they can dynamically adjust the grid to ensure a consistent and reliable energy supply.

AI agents can benefit in monitoring emissions and environmental metrics in real-time to ensure energy production activities comply with regulations and minimize environmental impact. These agents also contribute to improved customer service by delivering timely information, addressing inquiries, and offering energy-saving suggestions based on consumption patterns.

The integration of diverse energy sources and management of demand fluctuations present complex challenges in maintaining grid stability. AI agents provide real-time monitoring and control of the grid, ensuring stability and reliability. They can also autonomously manage power rerouting, load shedding, and supply balancing from different sources. In case of disruptions, AI agents swiftly identify and resolve issues to prevent widespread outages.

AI agents represent a remarkable advancement for the energy sector, offering solutions to several pressing challenges. Leveraging their capabilities in data analysis, real-time monitoring, predictive maintenance, and autonomous decision-making significantly enhances efficiency, reliability, and sustainability in energy production and distribution.

5.2 Shared Algorithms

The concept of shared AI algorithms signifies a groundbreaking vision in the realm of artificial intelligence. This vision heralds a future where numerous individual AI units have the capacity to independently learn over their lifetimes and share their accumulated knowledge with each other. The fusion of lifelong learning and knowledge sharing holds the promise of establishing a community of AI systems, wherein each unit contributes to and gains from the collective intelligence.

Shared AI algorithms empower units to gradually acquire multiple skills throughout their lifetimes. This continual learning process enables AI systems to adeptly adjust to new tasks and environments without necessitating complete retraining. AI units interact and exchange knowledge through a universal language, facilitating the sharing of insights and learnings. This communication elevates the overall intelligence of the network by pooling the experiences of individual units. By leveraging both local data and communication networks, AI units can efficaciously learn and adapt. Local data furnishes immediate context and relevance, while communication with other units enriches the learning process with diverse perspectives and information. Edge devices host the essential decentralized computation and data storage, empowering AI units to operate efficiently and autonomously. This decentralization fortifies the resilience and expandability of the AI network.

The aggregated knowledge of multiple AI units expedites the learning process. Each unit gains from the experiences and insights of others, resulting in rapid adaptation to new tasks and environments. A network of AI units collectively possesses more knowledge than any single unit could attain. This diversity fosters innovative solutions and approaches to problem-solving, as units can draw from a broader range of experiences and data. Decentralized AI systems demonstrate inherent resilience to failures and disruptions. The distributed nature of knowledge and computation ensures that the network can continue to function effectively even if individual units encounter issues. Shared AI algorithms promote sustainability by optimizing resource use and reducing redundancy. Knowledge sharing mitigates the need for repeated training on similar tasks, conserving computational resources and energy.

The utilization of shared AI algorithms in the energy sector offers significant advantages, particularly in terms of accelerating learning and adaptation. Unlike traditional setups where AI systems function independently, each learning from its own experiences and data, shared algorithms allow for the collective accumulation of a vast repository of information. This shared learning reduces the time needed for individual units to adapt to new tasks or changing environments. For example, an AI unit managing energy production in one region can share insights on demand patterns or efficiency strategies with units in other regions, leading to a more synchronized and optimized energy network.

Furthermore, shared AI algorithms contribute to a more diverse knowledge base within the energy sector. As each AI unit interacts with unique data sets and operational contexts, it leads to diverse experiences and solutions.

When these varied insights are pooled together, the resulting collective intelligence is far richer and more nuanced than the knowledge of any single unit. This diversity enables the development of more innovative and effective strategies for energy production, distribution, and management. For instance, insights from renewable energy management in one geographic area can inform similar efforts in another, leading to more effective integration of renewable sources across different regions.

Moreover, shared AI algorithms enhance the resilience and scalability of the energy sector. Decentralized AI systems, supported by shared algorithms, are inherently more robust against failures and disruptions. If one unit encounters an issue, others can quickly compensate by sharing their knowledge and resources. This decentralized approach ensures that the energy grid remains stable and functional even in the face of localized problems. Additionally, as the number of AI units increases, the system can scale seamlessly, with each new unit contributing to and benefiting from the collective knowledge base.

In the energy sector, sustainability is a critical concern, and shared AI algorithms have significant benefits in this area. These algorithms optimize resource use, reduce redundancy, and contribute to more sustainable operations. Knowledge sharing minimizes the need for repetitive training and computational processes, leading to substantial energy savings. Additionally, shared AI algorithms facilitate the integration of renewable energy sources by sharing strategies for managing variability and intermittency, resulting in a more reliable and sustainable energy supply while reducing reliance on fossil fuels and lowering the overall carbon footprint of energy production.

The predictive capabilities of shared AI algorithms are also crucial in enhancing operational efficiency and reliability. By analysing data from various sources and sharing predictive models, AI units can anticipate equipment failures, demand surges, and other critical events. This predictive maintenance approach reduces downtime and maintenance costs and ensures a more reliable energy supply. For example, an AI unit in one power plant can share predictive insights about turbine performance with other plants, allowing for predictive maintenance actions that prevent costly breakdowns.

Despite the numerous advantages, the integration of shared AI algorithms in the energy sector poses certain challenges that must be tackled. Establishing effective communication protocols for knowledge exchange, ensuring data security and privacy, and standardizing communication languages are crucial steps in fully realizing the potential of shared AI systems.

Additionally, further research is necessary to ascertain the optimal strategies for knowledge sharing, striking a balance between the immediacy of local data and the broader insights of collective intelligence.

The adoption of shared AI algorithms in the energy sector holds the promise of fundamentally transforming how we generate, distribute, and oversee energy. By nurturing a collaborative and perpetually learning network of AI units, the energy sector stands to achieve unprecedented levels of efficiency, sustainability, and reliability. As technological and scientific progress continues to enhance these systems, shared AI algorithms are poised to become a cornerstone of the energy industry's future, propelling innovation and excellence in an increasingly intricate and demanding global landscape.

5.3 BD4NRG Toolbox

BD4NRG Open modular Energy Analytics Toolbox combines and makes available to end-users the advanced user-friendly graphical working environment capabilities to support custom selection of local and/or third-party assets.

It includes:

- discovery of heterogeneous data sources formats
- a huge variety of analytics techniques via an AI-based machine learning algorithms/models
- a large variety of presentation modalities
- edge-level available storage and computing resources

BD4NRG Toolbox enables the development of proactive analytics centred applications, seamless integrated with near real time operational and/or with grid planning systems. The Toolbox supports analytics-driven improved decision making and other energy stakeholders business frameworks, hence contributing to the improvement of reliability of the grid network and to the provisioning of innovative energy services.

6. Distributed data management for AI

Several European countries have established national data hubs or platforms to facilitate data exchange and communication within the energy sector. These platforms could play a crucial role in enabling the integration of renewable energy sources, managing grid operations, and supporting the transition to smart grids:

- ENTSO-E Transparency Platform (European Union): The European Network of Transmission System Operators for Electricity (ENTSO-E) operates a transparency platform that provides access to electricity market data, including generation, consumption, and grid information across the EU.
- Energinet (Denmark): Energinet operates as Denmark's transmission system operator (TSO) and facilitates data exchange and communication in the Danish energy sector, including electricity and gas grid information.
- Energie Data Services Nederland (Netherlands): EDSN is responsible for data communication in the Dutch energy sector, managing data related to electricity and gas, supporting market processes, and ensuring data quality and security.
- Elexon (United Kingdom): Elexon operates as the Balancing and Settlement Code (BSC) Administrator for the electricity market in the United Kingdom. It manages data related to electricity generation, consumption, and settlement processes.
- RTE ECO2mix (France): RTE ECO2mix is a platform operated by Réseau de Transport d'Electricité (RTE) in France, providing real-time data on electricity production, consumption, and grid status.
- Terna (Italy): Terna operates an Energy Data Hub that collects and manages data related to Italy's electricity grid, including information on generation, consumption, and grid stability.
- E-Control (Austria): E-Control, the Austrian Regulatory Authority for electricity and gas markets, operates a data hub that provides access to energy market information.
- Statnett (Norway): Statnett operates a data platform that offers real-time data on electricity generation, consumption, and grid conditions in Norway.
- Terna (Spain): Terna, the Spanish Transmission System Operator, manages a platform that provides data on electricity generation, consumption, and grid operation in Spain.
- Datahub (Finland): Fingrid, the Finnish Transmission System Operator, administers a centralised data exchange system for the electricity retail market.

The entities responsible for national platform operations vary, ranging from national authorities to Transmission System Operators (TSOs) in other countries. Moreover, the collected data lacks standardisation and predominantly comprises historical data. However, many flexibility solutions require real-time or near real-time data.

The non-profit organisation, Center Denmark, which now is a European non-profit organisation is currently experimenting with real time data for smart grid solutions. For example, using a real-time price via a flexibility function (MIM) that is designed in a way that it solves the problem that the TSO or the distributor system operators (DSO) or Balanced Responsible Party have.

On the other hand, it is not always necessary to send data or unprocessed data to a cloud. Additionally, it is expensive and, in many cases, not possible due to lacking network capacity. The amount of sensor data from DSOs, for example, is in many cases very big and security sensitive. It is more efficient to process the data on the edge or near the edge and then communicate only the fraction of processed data to the (next) higher level. This concept can be called edge management and orchestration platform, or edge-to-cloud, as opposed to cloud-centric data processing and analytics.

7. Edge AI

Edge AI, which refers to the engineering of AI algorithms towards edge devices, brings immediate advantages when compared with Cloud-based AI solutions. By reducing the volume of data required to be transmitted to the Cloud, it intuitively brings advantages in terms of overall energy consumption and data security/sovereignty [Himeur, Yassine, Aya Sayed, Abdullah Alsalemi, Faycal Bensaali, and Abbes Amira2023].

As the processing occurs closer to data sources, it also brings a partial solution to storage of sensitive data and private data.

Processing data at the edge may provide advantages such as, for instance in the case of Industrial IoT (IIoT), potential reduction in energy consumption, costs, and enhanced data security.

Edge AI is also crucial to assist in decentralized energy systems, as it lowers latency (to transmit data), and enables dynamic resource allocation. As processing occurs closer to data sources, data transmissions are not so energy intensive.

Hence, for Cloud-Edge-IoT environments and providers (Edge, hybrid Cloud), the availability of open functional AI models is paramount. These models can be trained by considering decentralized approaches, thus reducing the required communication between Edge devices [Himeur et al 2024] and preventing vendor lock-in while enabling continuous improvement of the AI foundation model. Moreover, interoperable models are optimal for seamless integration.

While TinyML approaches have been growing and are the de facto approach when considering Edge AI, the application of foundational models (large AI models) is **growing**, despite the fact that data remains an issue. Specifically, models such as Llama 2 (Meta), Mistral 7B (Mistral AI) or Phi (Microsoft) are open-source but fine-tuned for specific applications. The use of large models towards Edge AI application in the Energy sector is therefore growing.

A further example of using generative AI at the Edge is demonstrated by Pratexo, a technology platform provider that specializes in supporting IoT and AI initiatives, particularly those requiring computational power at the edge. Collaborating with ABB, Pratexo developed a Generative AI-Based Root Cause Analysis Expert System by utilizing an open-source large language model.

To illustrate, consider a scenario where a machine expert responsible for power transformer stations utilizes this system. By furnishing the AI with pertinent context regarding the machinery or equipment they can create a fault cause graph tailored to a specific machine type, predict its performance, and automatically generate Python-based prediction algorithms using generative AI. These algorithms can identify issues like overheating and oil leaks using diverse sensors, subsequently triggering suitable actions, such as sending messages via MQTT streams or SMS, managed by an Edge Management and Orchestration platform. Deploying such AI-enabled systems to the power grid edge facilitates swift responses to issues and bolsters grid resilience.

References:

[Himeur, Yassine, Aya Sayed, Abdullah Alsalemi, Faycal Bensaali, and Abbes Amira]: "Edge AI for Internet of Energy: Challenges and perspectives." Internet of Things (2023): 101035.

8. Sustainable AI

AI services are not only CPU intensive but more so data Input/Output (I/O) intensive, and reliant on consuming large data assets (datasets). This implies novel dependencies between functional and data assets, which impacts the sustainability of running AI in terms of: energy consumption, green ratio of consumed energy a.o. CO2 footprint, economic cost, performance, privacy and cybersecurity.

From an infrastructure perspective, the energy profile of an AI workflow can be broken down to the combination of sub-profiles of:

- functional components of a workflow (e.g. data I/O, data preprocessing and model training);
- storage components of datasets feeding to a workflow; and,
- data-function mapping in a workflow, which implies data transmission that can be instigated over the network.

The above translates to: CPU energy (functional), Memory energy (storage) and Transmission energy (network). Existing work so far has proven that minimising all three – for maximum environmental and economic sustainability – is not feasible. Instead, trade-offs need to be considered to achieve efficient operation that supports AI sustainability.

Moving towards cloudified AI services over an edge-to-cloud continuum, renders these trade-offs non trivial. Because the energy cost, supply and CO2 footprint of edge-cloud Points of Presence (PoPs) differ and so is the different edge/core networks that connect them. The edge is known to have higher unitary cost of energy (e.g. per CPU rack, per sqm) [Bertoldi2017, Shehenaz2022]. The geographic density and distribution of PoPs varies too, particularly in the European landscape of edge-cloud and network infrastructure [Kamiya2024]. On the other hand, the emerging trends of localised production of green energy can play a major role in driving down energy costs of the edge.

Moving from the edge towards larger clouds, energy unitary costs (e.g. per rack, per sqm, per application) typically decreases. But that comes with the trade-off of longer distance between end-users/-devices (i.e. data generation points) and AI workflow components. This automatically introduce network artefacts, observed through data transmission from generation points to storage/processing counterparts. Consequently, increases the role of the network energy in the trade-off. Although, the opposite does not automatically hold. Meaning, bringing AI workflow to the edge does not automatically decrease exposure to the network. Because CPU and memory storage of edge PoPs are constrained. This increases the likelihood of having to distribute workflow components across multiple PoPs, introducing inter-component data transmission within a workflow. Our work in [Ejaz2024] provides an elementary benchmark of resource utilisation, when distributing cloud workflows. It showed that distributing CPU-light microservices is likely to degrade the efficiency of a workflow, as the added CPU usage for routing across components outweighs the CPU usage of the components themselves. Therefore, from the perspective of AI workload management, placement/distribution policies need to systematically factor these trade-offs to support AI sustainability. This need, along with potential clustering of AI workflow components in larger data centres, is recognised by the World Economic Forum as strategic techniques to drive energy demand of AI lower [WEF2024].

Form a ML perspective, different data ingestion and learning architectures results in different energy profile of the respective AI service. From an architectural perspective, federated and distributed learning can yield significant advantages in distributing and balancing the training load across multiple agents. The breakdown of such a large task into smaller sub-tasks allows for higher flexibility and larger opportunities for optimising the three energy profiles of computation, storage and networking. However, the flexibility is not without a cost that arise in the form of: fractional knowledge loss due to limited training per agent and to model aggregation, higher vulnerability to cyber threats/disruptions and higher management complexity due to the larger number of opened interfaces [Galaz2021]. The trade-off between load balancing, optimisation flexibility and risk exposure are often influenced/constrained by the landscape of infrastructure. In Europe, the higher dispersion of resources and larger number of autonomies fosters larger adoption of federated and distributed ML architectures; compared to international competitors elsewhere.

At the granularity of a single training process, data patterns directly correlate with the number of training cycles, required to balance the 'over-under fitting' trade-off and achieve a certain quality threshold. Higher entropy within datasets may induce a larger number of training cycles, to identify a specific pattern, as shown by [ALNaday2024] for anomaly detection use-cases in cybersecurity. Added to that, hyper parameters used in training are next in influencing the computational demand, as the speed with which a model can learn.

Emerging Large Language Models (LLMs), such as OpenAI and ChatGPT, introduce new variables into their sustainability. Most prominent of which, are: the number of parameters they train upon, the modality of data and the inference decisions they make [Wu2022, Luccioni2024]. These works show the correlation between the number of parameters (i.e. LLM size) and the CO2 emissions. The work of [Luccioni2024] further shows a breakdown of CO2 budget per LLM operation. Inference is found to have a significant contribution into CO2 emissions, in some cases outweighs that of model training and/or fine-tuning. Moreover, image training and inference have much higher CO2 footprint, compared to other data models. These efforts illustrate the trade-off between model size, data modality and computation demand, which present an orthogonal optimisation direction to that of workload management and even ML architecture to enable holistic optimisation of trade-offs that collectively support AI sustainability goals.

References

[Bertoldi2017] Bertoldi, P., Avgerinou, M. and Castellazzi, L., Trends in data centre energy consumption under the European Code of Conduct for Data Centre Energy Efficiency, EUR 28874 EN, Publications Office of the European Union, Luxembourg, 2017, ISBN 978-92-79-76445-5, doi:10.2760/358256, JRC108354.

[Shehenaz2022] Shehenaz Shaik, Sanjeev Baskiyar, "Distributed service placement in hierarchical fog environments," Sustainable Computing: Informatics and Systems, Volume 34, 2022, 100744, ISSN 2210-5379, https://doi.org/10.1016/j.suscom.2022.100744.

[Kamiya2024] Kamiya, G. and Bertoldi, P., Energy Consumption in Data Centres and Broadband Communication Networks in the EU, Publications Office of the European Union, Luxembourg, 2024, doi:10.2760/706491, JRC135926.

[Ejaz2024] S. Ejaz and M. Al-Naday, "FORK: A Kubernetes-Compatible Federated Orchestrator of Fog-Native Applications Over Multi-Domain Edge-to-Cloud Ecosystems," 2024 27th Conference on Innovation in Clouds, Internet and Networks (ICIN), Paris, France, 2024, pp. 57-64, doi: 10.1109/ICIN60470.2024.10494435.

[WEF2024] https://www.weforum.org/agenda/2024/04/how-to-manage-ais-energy-demand-today-tomorrow-and-in-the-future/

[ALNaday2024] AL-Naday, M., Dobre, V., Reed, M. et al. Federated deep Q-learning networks for service-based anomaly detection and classification in edge-to-cloud ecosystems. Ann. Telecommun. 79, 165–178 (2024). https://doi.org/10.1007/s12243-023-00977-4

[Galaz2021] Victor Galaz, Miguel A. Centeno, Peter W. Callahan, Amar Causevic, Thayer Patterson, Irina Brass, Seth Baum, Darryl Farber, Joern Fischer, David Garcia, Timon McPhearson, Daniel Jimenez, Brian King, Paul Larcey, Karen Levy, "Artificial intelligence, systemic risks, and sustainability," Technology in Society, Volume 67, 2021, 101741, ISSN 0160-791X, https://doi.org/10.1016/j.techsoc.2021.101741.

[Luccioni2024] Luccioni, Sasha, Yacine Jernite and Emma Strubell. "Power Hungry Processing: Watts Driving the Cost of AI Deployment?" Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (2023)

[Wu2022] Wu, Carole-Jean, Ramya Raghavendra, Udit Gupta, Bilge Acun, Newsha Ardalani, Kiwan Maeng, Gloria Chang, Fiona Aga Behram, James Huang, Charles Bai, Michael K. Gschwind, Anurag Gupta, Myle Ott, Anastasia Melnikov, Salvatore Candido, David Brooks, Geeta Chauhan, Benjamin Lee, Hsien-Hsin S. Lee, Bugra Akyildiz, Maximilian Balandat, Joe Spisak, Ravi Kumar Jain, Michael G. Rabbat and Kim M. Hazelwood. "Sustainable AI: Environmental Implications, Challenges and Opportunities." Proceedings of Machine Learning and Systems 4 (MLSys 2022)

9. Evolving related eco systems and initiatives

The adoption if industry 5.0 and still remain competitive in terms revenue and profit in the market landscape is a challenging task. The reason being these aspects will eventually create traction on top tier technological innovation based investment as well as quick qins, faster GTM with new product and services. This means it's strategic and tactical balance to grow and become thought leader. Some of the fundamental pre-requisites for success are-

Governance – Not only corporate governance but overarching architecture governance in harmony ensuring end to end solution value without architectural debt and sunken cost is very important. This is because in industry 5.0 adoption journey there will be several players in ecosystem, e.g. Network operators, Equipment Manufacturer, Cloud-Infra provider, Service and integrators, OT providers/machine builders, facility constructures, sensor and device providers, Consultation providers etc. There are many more working along with these major players. The end to end value need to be generated by all and all of them has their won responsibility hence overall governance is key to success for the organization without which investment will soon be seen as a waste. The roles needed for this are primarily Strategist or Senior Enterprise architects working in co-operation with CIO and CTO offices.

Investment Rationale – In early phases in the journey there is a always the interest in technical innovation and that needs investment. That investment if not rationalized by use cases, business cases, value stream and value realization strategy in a longer perspective, will provide serious negative consequence. An example is of investment of AI centric technology in cloud or On-prem or in Hybrid. Thoughtful thinking around this will help decide OpEx and possible non-usage GPU expenses.

Composability – Composability is the key thrive and sustain in cases of disruption. This needs to be the DNA of the Organizational Operating Model. Composable thinking, Composable business model, composable architecture and composable design will enable lowest level solution stack to t be interoperable to any combination of solution, process, data and system which will ensure adaptable business model. This is an approach similar to LEGO building blocks. Meaning the business model outcome can be adapted and reconfigured and improved with any existing smaller solution building blocks by interchanging them and reconstruction instead of creating ne solutions.

Technical & Architectural Skill sets – Advanced skills like data scientist, prompt engineers, data engineers, devSecOp engineers will play key role in advancement. Senior Architects and Enterprise Architects will play the key role in in bridging all the gaps and making sure solution providers and vendors are providing the expected outcome.

Business Office/Innovation Hub – A parallel Cloud Infra business office and/or Innovation Hub is essential to apply small scale innovation to keep the technological competitive edge. This will be cross cutting organization providing new use case realization at lowest cost and most innovative way which are pilot scenarios.

6P – Product. Process, Platform; People, Partnership, Performance are essential which will provided the merged single view of IT and OT. This is key to bridge the gap between data, process and platform and connected people which will create clear boundary of data usage, consumptions and governance.

Critical role will be fulfilled by HPC Infrastructures.[i]

Several European networks are playing a critical role in advancing AI in energy. **AI Coalition NL**, **BRIDGE initiative** is setting up a Taskforce aimed at DIHs (**SCoDIHNet**), new infrastructure **Digital Spine**.

10. Virtual Worlds, DT, AI in Energy

AIOTI has published in April 2024 the paper on **Edge IoT Industrial Immersive Technologies and Spatial Computing Continuum**.

The paper aims to provide the vision of the convergence of edge IoT, artificial intelligence (AI), digital twins (DT), immersive triplets (IMT), intelligent mesh connectivity, IoT of senses (IoTS), software-defined automation (SDA) and spatial computing technologies to create an industrial real-digital-virtual continuum. Such continuum is made of immersive environments, which are computer-generated virtual worlds where users can sense as if they were physically embodied in that generated perception context.

The convergence of these technologies into industrial immersive solutions advances the integration and application of edge intelligent immersive technologies combining augmented reality (AR), virtual reality (VR), mixed reality (MR), and extended reality (XR) with concepts like metaverses, omniverse, multiverses, next generation spatial web, Web 4.0 as part of future virtual worlds.

Such convergence of industrial immersive technologies at the edge can improve efficiency, reduce downtime, enhance safety, and better decision-making in industrial settings. However, to be effectively deployed, it both requires a strong interdisciplinary collaboration and presents challenges like robust hardware (HW) design and cost-effective availability, data security and privacy preserving methods, and effective industrial workflow integration. As technology advances, the adoption of such convergence in industry is expected to grow, offering transformative benefits across various sectors and vertical markets, including industrial manufacturing, product operations, design and maintenance, training and collaboration, data visualisation, mobility and logistic, energy, automotive, aerospace, and healthcare.

Edge Computing

Edge computing moves service provisioning closer to producers and users of such services. It can provide reduced latency, mobility support, and facilitate data analytics to be done close to the data source and creates the possibility for reduced energy consumption.

The convergence of the IoT and edge computing with immersive technologies, such as VR, AR, and MR, is creating a transformative synergy that significantly advances the capabilities and applications of each technology. This integration brings a new era of immersive experiences more interactive, responsive, and integrated with the real world.

The synergies created by IoT, edge computing, AI, and immersive technologies enhanced real-time interactivity and provided low latency by processing data closer to the source. This allows users to experience real-time interaction with virtual environments with minimal delay, making experiences like VR and AR-guided applications more seamless and effective.

With IoT devices generating vast amounts of data, edge computing enables real-time data analytics. This lets immersive applications dynamically adjust content based on immediate user interactions and environmental conditions, enhancing personalisation and responsiveness.

IoT devices and edge computing platforms can provide real-time information about the physical environment, which can be integrated into AR and VR applications to create more contextually aware immersive experiences and improve contextual and spatial awareness.

Spatial computing, IoT, and immersive technologies combine to enable advanced spatial computing capabilities, where the physical and digital spaces are more tightly integrated. Users can interact with digital objects that are aware of and responsive to their environment's physical layout and objects, enabling more natural and intuitive interactions.

IoT devices and edge computing can gather data on user preferences, health metrics, and environmental conditions. This allows immersive applications to tailor experiences in real-time to the individual's needs and context, improving accessibility, personalisation, and user satisfaction.

Wearable IoT devices, edge computing processing, and AI can enhance immersive experiences through biometric data, enabling applications to adjust based on the user's physical responses. This can lead to more engaging and personalised content.

Edge processing reduces the need for constant high-bandwidth connectivity to the cloud, making it more feasible to deploy immersive technologies in bandwidth-constrained environments that increase the scalability and efficiency of the technology. Edge computing can reduce the energy consumption of data processing for IoT devices, extending the battery life of wearable and portable immersive technology devices.

AI can boost IoT and edge computing by offering more intelligent data analysis tools and facilitating real-time autonomous decision-making.

It also foresees maintenance requirements in industrial contexts or tailors' user interactions in virtual settings according to behaviour and preferences. AI's role in comprehending and handling the complexity of interactions within multiverse environments, where several virtual worlds coexist, is essential.

The synergies between IoT, edge computing, and immersive technologies drive significant advances across various sectors by enhancing interactivity, contextual awareness, user experience, and scalability. This integration is paving the way for innovative applications that were previously challenging or impossible to achieve, predicting a new era of digital interaction that muddies the lines between the physical and virtual worlds.

Edge Artificial Intelligence

AI creates inclusive interfaces that will make the users' journeys convenient for everyone, including people with disabilities (e.g., improve implementation of universal design). Thus, AI makes the immersive applications user-friendly and easy-to-use platforms. Technologies such as NLP, speech recognition, computer vision, translation, and AR enable users to interact with the metaverse in their native language through images and videos and enhance user-verse interactions.

Integrating edge AI with immersive technologies like VR, AR, MR, and verses (e.g., metaverse, omniverse, multiverse) fosters a new wave of innovation and enhanced capabilities. This synergy improves immersive applications' performance and user experience and opens new possibilities for use in various fields.

Real-time and pseudo-real-time processing, decision-making and responsiveness combined with analytics performed by edge AI models and algorithms means user actions can be processed and reflected in the virtual environment almost instantaneously. This is crucial for maintaining immersion and preventing motion sickness in VR applications. It allows for immediate recognition and overlay of digital information on real-world objects, enhancing user interaction with their environment.

Edge AI and generative AI can improve dynamic content adjustment by leveraging AI models that run on the edge; immersive applications can dynamically adjust content based on the user's behaviour, preferences, and immediate environment, and an AR-based learning application can modify its teaching methods in real-time based on the learner's engagement level and understanding.

Edge AI can process data from various sensors embedded in the environment or the device to better understand the context, the environment and the spatial location and adapt the immersive experience accordingly.

This is particularly useful in AR applications for indoor and outdoor navigation and interactive autonomous systems, robotics, and human interaction, where the digital content and the immersive triplets need to be aware of the physical space.

AI, ML, and computer vision algorithms integrated at the edge enhance immersive technologies' capabilities. Gesture recognition and object detection can be performed locally on AR glasses, enabling more natural interactions with digital content without needing external HW.

In this context, edge AI plays a crucial role in spatial computing, where the technology understands and interacts with the physical space around it. This is fundamental for creating coherent and interactive AR and MR experiences seamlessly blending digital content with the real world. In addition, edge AI can be used for proper and adequate identification of objects function, position, dynamic state and thus allows for enhanced interaction possibilities with real world through digitalisation.

Edge AI can directly implement real-time security measures, such as anomaly detection and immediate response to potential threats, on the device. This adds an extra layer of security for immersive applications that may be processing sensitive information.

While integrating edge AI with immersive technologies offers numerous benefits, it also presents challenges, such as the need for advanced hardware that supports AI computations, the complexity of developing and deploying AI models on edge and ensuring the interoperability of devices and platforms. Addressing these challenges requires ongoing technological advancements, standardisation efforts, and a focus on developing lightweight AI models that can run efficiently on edge devices.

Next to the development of lightweight, classical AI models, neuromorphic and event-based AI models can also address the challenges faced by edge AI implementations. These approaches are biologically inspired and research novel neuron models. By exploiting sparsity in the sensor data and the computations of the resulting AI model, these AI systems can run with less energy and memory requirements but face new training challenges.

The synergy between edge AI and immersive technologies significantly enhances the capabilities, performance, and applications of VR, AR, MR, and immersive verses (e.g., metaverse, omniverse, multiverse). By enabling real-time processing, personalised experiences, and advanced interactive features, this integration is paving the way for the next generation of edge IoT immersive applications across various sectors.

Spatial Generative Edge Artificial Intelligence Technologies

As technology advances, the AI-generated content is making its way into edge immersive technologies through AI-powered algorithms, and the ability to understand and respond to context combined with the capabilities to generate coherent, relevant, and appropriate scenarios and novel situations.

Generative AI technologies can generate text and audio descriptions of the virtual spatial environments, which can help make the experience more immersive and engaging for human and machine users combined with generated interactive dialogue for virtual characters that users might encounter on the virtual world.

Heuristics, as a methodology, facilitate immersive systems' abilities to solve problems using practical and intuitive methods. Generative heuristics can employ rules of thumb, educated guesses, and simplified strategies to approximate solutions, particularly in complex and non-deterministic situations.

Heuristics steer the search for solutions in areas where exhaustive exploration is impractical or impossible, thus aiding in the discovery of acceptable solutions within a reasonable timeframe.

One major challenge is ensuring that the AI-generated content in immersive applications is accurate and relevant as generative AI models are trained on a large dataset of text, images, videos. This may create problems to generate content that is entirely accurate or relevant to a specific edge location, context or scenario and may not always be able to create appropriate content for all users, which can lead to issues with accessibility.

Immersive technologies require the use of multimodal generative AI technologies trained on different types of media, like text, images, voices, audio, and videos and capture multi-lingual elements.

Combining AI-generated content in immersive environments raises certain ethical questions. As generative AI technologies are based on an autonomous system, and it can generate text, images, videos, voices based on what it has learned from the data it was trained on, which could include biases and stereotypes. This requires continuous monitor and review of the generated content to avoid any unwanted biases or stereotypes.

Spatial generative edge AI technologies represent a cutting-edge fusion of spatial computing, generative AI models, IoT and edge computing, explicitly tailored for immersive technologies like VR, AR, MR, and verses (e.g., metaverse, omniverse, multiverse). This integration aims to enhance the creation, interaction, and personalisation of digital content within physical spaces in real-time, leveraging the power of AI directly on edge devices.

Understanding spatial generative edge AI involves using spatial computing to provide edge IoT devices with the ability to understand and interact with their physical environment in three dimensions.

In immersive technologies, spatial computing enables devices to map environments, recognise objects and spaces, and place digital content in the physical world so that it appears to coexist with real-world objects.

Generative AI models can create new data that resembles the training data and create new scenarios, sceneries, and simulation environments. Applied to immersive technologies, these models can generate realistic textures, objects, or entire environments on the fly, enhancing the richness and diversity of immersive experiences. By processing data on local devices or nearby computing resources, edge computing reduces latency, conserves bandwidth, and improves privacy. This is crucial for immersive technologies that require real-time processing to maintain immersion and user engagement.

Spatial generative edge AI can dynamically generate content based on the user's environment and interactions. The environment could adapt in real-time to the physical space around the user, generating obstacles or items based on the space's layout.

For immersive AR environments, spatial generative AI can be used to seamlessly fuse the real surroundings with generated virtual objects and entities, leading to hyper realistic fusion of virtual and real. With generative AI, photorealistic rendering of extended reality is possible, lowering the acceptance barrier of augmented and mixed reality for the user.

The technology can generate realistic scenarios and environments for training and simulation applications that adapt to the user's actions and decisions, providing a highly personalised and practical learning experience.

Spatial generative edge AI can personalise immersive experiences in real-time by understanding the user's preferences and environment context. This could range from adjusting the difficulty level of a simulation to changing the ambience of a virtual meeting space based on the participants' preferences.

In the same manner, interactive immersive triplets and avatars can be created in different contexts, spaces, and time dimensions. Users can interact with these interactive immersive triplets and avatars through VR, AR, MR, and verses (e.g., metaverse, omniverse, multiverse), enabling them to visualise data, simulate changes, and understand complex systems intuitively.

Running sophisticated spatial generative edge AI models requires energy-efficient and performant HW that can handle intensive computations while maintaining low energy consumption, which can be challenging for portable and wearable devices. To address this challenge, edge AI models, especially generative ones, must be highly optimised for edge deployment to ensure they operate efficiently in real-time without compromising the quality of the generated content.

Ensuring seamless interaction between different edge immersive IoT devices, platforms, and data formats is crucial for the widespread adoption of these technologies, requiring ongoing standardisation and development efforts to provide interoperability and scalability.

Spatial generative edge AI technologies are poised to revolutionise immersive experiences. By leveraging the synergy between spatial computing, generative AI, and edge computing, they will offer unprecedented levels of personalisation, realism, and interactivity. As these technologies evolve, they will unlock new possibilities for creating immersive, intelligent, and trustworthy environments that seamlessly integrate with the physical world.

Integration of Sensing and Communications

Integrating advanced IoT sensory technologies and intelligent mesh connectivity is essential for converging edge immersive technologies. Edge AI and immersive intelligent mesh connectivity allow for the classification, detection, localisation, and estimation of an object's attributes and other functions. The IoT is expected to combine sensing and communication technologies into a fully integrated system. Systems incorporating joint communication and sensing (JCAS) functionalities represent a significant innovation for 6G that will facilitate numerous emerging immersive technologies and applications and enable the concept of a perceptive network.

The communication component of an JCAS system would mean connecting the digital, physical, and human worlds in real-time with all extreme requirements coming from the immersive technology value chain (**Figure 2**). The sensing component, on the other hand, will provide us with the capability to sense the world and to provide context information to create the digital map of the environment, which is a prerequisite for creating the DTs and seamless integration of the physical, digital, virtual, and cyber worlds. With JCAS in 6G and beyond, we will have the capability to not only localise objects that are part of the network but also to sense and integrate objects that are not connected to the network.

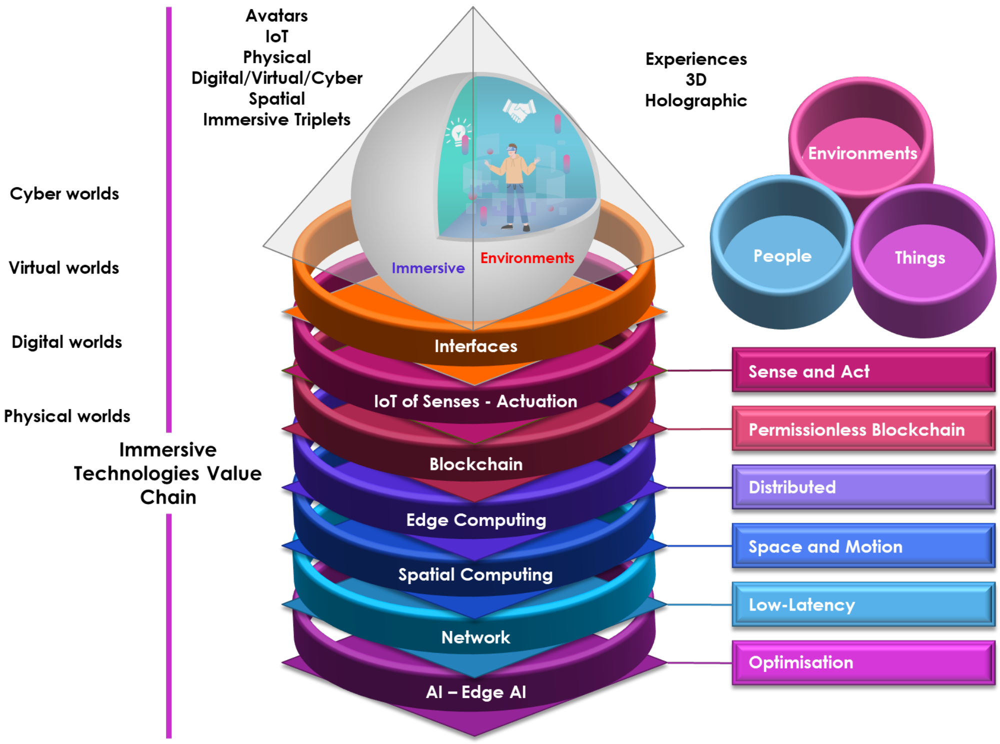


Figure 2 Edge immersive technologies convergence across the value chain

For real-time immersive applications, integrating a human model with a DT representation of the physical world entails exchanging multimodal data, such as haptics, position, velocity, and interactions, as well as the senses in the IoT context, including human gestures, head movements and posture, eye contact, facial expressions, emotions, etc. [One6G Association]

The JCAS integration is envisioned in different ways, from loosely coupled to fully integrated, shared spectrum, shared HW, to shared signal processing module and network protocol stacks, and even using the same waveform for both communications and sensing. Sensing functionality can be introduced as a service with a low incremental cost as it leverages equipment and spectrum deployed for communication purposes. JCAS can extend the capability of cellular networks by adding see-and-feel functionalities. The edge devices can sense their surroundings and exchange their sensing results through communication links. From a network perspective, the widely deployed base stations for legacy cellular service could be re-used for wide-area seamless RF sensing. Incorporating AI, communications, positioning, and sensing capabilities, the cellular network could intelligently fuse the physical world with the digital world and provide various new services for consumers and industry customers [5G Americas Report].

Systems with JCAS functionalities are still in their early stages, and during the next few years, a substantial amount of additional research will be needed. They will also share resources in the time, frequency, and space domains and essential components, including HW, waveforms, and signal processing. This implies that the signals of the cellular system will be used to significantly increase the sensing capabilities in addition to the communication network transporting sensor data.

Integrating communication and sensing can add ambient IoT to budget-friendly and power-efficient edge devices that facilitate the connection of numerous objects and items to networks, allowing for many applications. The driving forces behind ambient IoT are energy harvesting, storage, and backscattering techniques. Energy harvesting and storage can include diverse energy sources, including RF signals, light, vibrations, and thermal energy. Different energy sources have varying levels of availability and energy density.

The backscattering technique can involve a few other factors. Backscattering, combined with active signal generation (e.g., power amplifiers), minimises power consumption in Ambient IoT devices. In backscattering, the transmitted signal reflects continuous waves from a reader, modulated with information for communication. Ambient IoT's reliance on energy harvesting and backscattering holds the potential to advance low-cost, maintenance-free, and environmentally sustainable IoT solutions.

The combined functionality of communication and sensing also facilitates emerging secure proximity services. Such technology ensures the secure and private exchange of data between devices in close physical proximity, enabling applications like mobile payment, intelligent access control, communication to everything and IIoT. The sensing feature acts as the authentication mechanisms to safeguard the communication and prevent unauthorized access or data breaches. JCAS also allows secure communication links by localising potential eavesdroppers, preventing the decoding of malicious data, and activating focussed counter-attack measures based on encryption and authentication protocols in combination with dynamic beamforming or special filtering.

The ambient IoT edge devices can be used as part of the immersive spaces to provide real-time information about the environmental conditions in the physical world.

Future Technology Trends

Sustainable system: The integration of immersive technology, IoT and Edge computing, correspondent to the technological growth of devices and communication system enables the decentralisation and remote operations in industry, with significant impact on emissions of greenhouse gasses. Immersive technologies can facilitate the shift towards using fewer environmentally damaging and non-recyclable materials by minimising physical waste in industrial processes. Developments in IoT, edge computing, AI, and industrial immersive technologies promote a significant transition to business models that substitute production and lifecycle management of physical items with digital services and assets. These technologies play a vital role in decreasing the demand for energy uses that heavily emit greenhouse gases, such as transportation, buildings, and heating and cooling, by enhancing mobility organisation, improving building energy efficiency, and optimising industrial processes. Sustainable immersive technologies require standards encompassing various sustainability facets, including energy efficiency, resource optimisation, circularity, and social responsibility. These measures extend across multiple disciplines and technologies to facilitate sustainability across different industries.

Conclusion:

The convergence of IoT, edge computing, AI, and industrial immersive technologies holds great promise for transforming industries and enhancing human capabilities. However, realising this potential requires overcoming significant challenges, including interoperability, privacy, energy consumption, security, legal and ethical considerations.

Success in this endeavour depends on collaborative efforts among research communities, technologists, industry leaders, standardisation bodies, policymakers, and the broader society to develop innovative solutions and frameworks that can effectively address these challenges while maximising the benefits of these converging immersive technologies.

Concentrated and aligned efforts in technological development, orchestration, standardisation, interoperability, and research at the European level are needed to overcome the challenges and use the opportunities that arise from the convergence of IoT, edge computing, AI, and industrial immersive technologies.

The convergence and fusion of technologies such as IoT, AI, intelligent connectivity with DTs, immersive triplets, and cloud data exchange facilitate the development and deployment of VR, AR, MR, XR, and various verses (metaverse, omniverse, multiverse), inherently require a highly interdisciplinary approach. The multidisciplinary nature of these technologies involves multiple fields, such as computer science, electrical engineering, data science, user experience design, cognitive science, and more.

The systems-of-systems integration of these technologies signifies a technological shift, as well as a cultural and economic one, potentially altering how we perceive and interact with the digital and physical worlds. This convergence promises a more integrated, immersive, and interactive future.

European Industrial Leadership: European leadership in sectors such as automotive, aerospace, manufacturing, energy, and health provides a solid foundation for the adoption and integration of these converging technologies. This existing industrial base can pioneer the development of systems-of-systems integration crucial for realizing industrial verses.

References:

[One6G association], "6G Technology Overview - One6G White Paper". Zenodo, Jun. 10, 2022. doi: https://doi.org/10.5281/zenodo.6630706

[5G Americas] 5G Americas Report, "3GPP Technology Trends". https://www.5gamericas.org/wp-content/uploads/2024/01/3GPP-Technology-Trends-WP.pdf

11. Testing facilities, computing capacity and infrastructures

Experimental Pilot approaches are evolving to address need for computing capacity and testing facilities. Cloud edge infrastructures (CEI) cross-sector large-scale experimental environment that allows for CEI cross-sector experimentation could address needs of different stakeholders and integrate selected AIOTI testbeds out of the 40 existing testbeds[1]. Such facilitation can assist cross-sector interaction and data interoperability. The AIOTI testbed catalogue integrates a total of 40 collected testbeds, where 4 are in a use-case stage (TRL2); 8 are in a proof-of-concept stage (TRL3-4); 28 correspond to deployed testbeds (TRL4-6), focused on specific domains. As shown, most testbeds cover use-cases in Smart Cities (37.5%) and Manufacturing domain (35%). The IoT-X-Edge P5 shall provide support to use-cases with different levels of testbed maturity[4]; This approach can also provide articulation towards various domains TEF initiatives: TEF energy; TEF agriculture (e.g., agrifood TEF); AI-Matters (Manufacturing); Mobility/Smart Cities (e.g., Citcom.AI). One of the initiatives is HEDGE-IoT that was mentioned in the Section 3.

[1] https://aioti.eu/about-us/our-groups/testbeds/
[2] https://aioti.eu/wp-content/uploads/2022/06/AIoTI-Testbed-Methodology-2021-Published.pdf
[3] https://www.sciencedirect.com/science/article/abs/pii/S1474034623000046#b0135
[4] https://www.sciencedirect.com/science/article/abs/pii/S0278612522001108

12. AI and Energy Poverty

Energy poverty context

In the past, there was a distinction between Energy Poverty (EP) and Fuel Poverty (FP). EP was considered the difficulty or lack of access to energy in developing countries, and FP was considered the issue of affordability of such services in developed countries. However, in recent years, the terms EP and FP have been used interchangeably, and today, the EP issue concerns the situation where households are not capable of meeting their respective energy needs in terms of limited supply, affordability, quality, quantity, reliability, or a combination of these factors[5].

Although there is still no standardised way to define EP, several authors offer different approaches to EP classification. For example, the framework[6] that categorises into six groups of potential drivers of EP as follows: access (or the lack of it), affordability issues, energy flexibility, energy efficiency, specific needs, and practices. Another example[1] categorises the EP issue metrics into primary (considering consumer-level info) or secondary (weighted scoring and info from utilities), and relative (comparing multiple households) or absolute (via strict thresholds).

Maybe the two most popular metrics are primary-relative and subjective metric (originated from household surveys to assess their ability to meet energy needs) and objective and primary-absolute metric (how much of the household income, in percentual values, should not be exceeded in energy expenses). The latter metric[7] is popularised as the "10% rule", where no more than 10% of the household income should be used for energy services.

AI approach

AI's ability to provide accurate predictions by handling non-obvious relations between variables and identifying complex relations in large and complex datasets can be a powerful tool to help the development of data-based decision-making supporting tools to address the EP issue. Despite all these advantages, AI can present one limitation that jeopardises its adoption: its potential operation as "black boxes" or, in other words, how the decision rules within algorithms are often obscured and not easily understandable.

This issue, also known as the eXplicability of Artificial Intelligence (XAI)[8], has become an important aspect of AI implementation in a broad aspect of applicability and, in special, for EP implications. With proper XAI, developers would be able to implement improvements and asses the AI algorithm accuracy, GDPR demands[9] (e.g., if segmentation and scores are free from hidden biases) can be fulfilled, and a better understanding of the reasoning behind the algorithms' decisions may help broader adoption of AI algorithms. Understanding how complex results are achieved grants XAI (and AI) transparency and data traceability.

[5] S. Cong, D. Nock, Y. L. Qiu, and B. Xing, 'Unveiling hidden energy poverty using the energy equity gap', Nat Commun, vol. 13, no. 1, p. 2456, May 2022, doi: 10.1038/s41467-022-30146-5.

[6] S. Bouzarovski and S. Petrova, 'A global perspective on domestic energy deprivation: Overcoming the energy poverty–fuel poverty binary', Energy Research & Social Science, vol. 10, pp. 31–40, Nov. 2015, doi: 10.1016/j.erss.2015.06.007.

[7] E. Dogan, M. Madaleno, R. Inglesi-Lotz, and D. Taskin, 'Race and energy poverty: Evidence from African-American households', Energy Economics, vol. 108, p. 105908, Apr. 2022, doi: 10.1016/j.eneco.2022.105908.

[8] Melanie, 'XAI or eXplainable Artificial Intelligence: What is it about?', Data Science Courses | DataScientest. Accessed: Mar. 03, 2024. [Online]. Available: https://datascientest.com/en/xai-or-explainable-artificial-intelligence-what-is-it-about

[9] gdpr.eu, 'GDPR.eu', General Data Protection Regulation (GDPR). Accessed: May 14, 2022. [Online]. Available: https://gdpr.eu/tag/gdpr/

AI adoption, including Machine Learning (ML), to address the EP issue is still in its initial stages[10]. However, the increasing availability of AI techniques and growing relevance of EP points in the direction of a broader acceptance with a special focus on identifying the most significant predictors of the EP issue.

IREPO use case

IREPO (Irish Energy Poverty Observatory) aims to establish a replicable platform to facilitate Research, Development, and Deployment (RD&D) at national and international levels. It focuses on creating an advanced, data-driven toolset to support decision-making and policymaking regarding energy poverty. This toolset integrates diverse data from reliable sources to address challenges faced by energy-insecure households. IREPO's collaborative effort ensures a dynamic tool continuously updated with multi-dimensional data, including social, economic, gender unbiased insights, and market trends, to guide entities like public bodies and policymakers. Leveraging machine learning and AI algorithms, IREPO offers simulation environments for analysing data, validating policies, and enhancing renewable energy adoption.

The IREPO project aims to surpass common factors contributing to energy poverty[11],[12] by establishing a federated database that integrates diverse data sources beyond traditional factors like dwelling characteristics and energy prices. This database will provide a comprehensive and up-to-date resource to support data-driven decision-making and policy formulation. By incorporating various data formats and sources such as surveys, interviews, open-source data APIs, and eventual IoT data, IREPO ensures a continuous and dynamic integration process for timely and precise insights into energy poverty-related aspects. This innovative approach addresses data limitations[13] and enhances the effectiveness of strategies to combat energy poverty.

IREPO utilises machine learning and AI algorithms to uncover patterns within datasets through automated processes. By training machine learning models on representative datasets, IREPO can analyse relationships between features to identify correlations, trends, and complex structures within the data. This approach enables the system to learn patterns and make predictions without explicit programming, offering valuable insights into multidimensional factor correlations[14]. By integrating diverse data sources such as energy market fluctuations, renewable energy history, census data, and socio-economic information, IREPO enhances understanding of the risk factors and impacts of energy poverty, providing a comprehensive view of this multifaceted issue[15] [16].

The IREPO platform's machine learning algorithms will become more accurate over time as historical data grows, despite variations in precision and timeliness across different data sources. The continuous data integration process maps, translates, and transforms data from various sources, as shown in Figure 2, into a standardised federated database, ensuring timely data availability for powerful decision-making, simulations, and policy support tools.

---

10 A. López-Vargas, A. Ledezma-Espino, and A. Sanchis-de-Miguel, 'Methods, data sources and applications of the Artificial Intelligence in the Energy Poverty context: A review', Energy and Buildings, vol. 268, p. 112233, Aug. 2022, doi: 10.1016/j.enbuild.2022.112233.

11 C. Spandagos, M. T. Reaños, and M. Á. Lynch, 'Energy poverty prediction and effective targeting for just transitions with machine learning', ESRI, Working Paper 762, Sep. 2023.: https://esri.ie/publications/energy-poverty-prediction-and-effective-targeting-for-just-transitions-with-machine

12 D. Lawlor and A. Visser, 'Energy Poverty in Ireland', Library & Research Service, Mar. 2022. Accessed: Apr. 26, 2023. [Online]. Available: https://bit.ly/40EUuoM

13 C. Lowans, D. Furszyfer Del Rio, B. K. Sovacool, D. Rooney, and A. M. Foley, 'What is the state of the art in energy and transport poverty metrics? A critical and comprehensive review', Energy Economics, vol. 101, p. 105360, Sep. 2021, doi: 10.1016/j.eneco.2021.105360.

14 C. Lowans et al., 'What causes energy and transport poverty in Ireland? Analysing demographic, economic, and social dynamics, and policy implications', Energy Policy, vol. 172, p. 113313, Jan. 2023, doi: 10.1016/j.enpol.2022.113313.

15 J. P. Gouveia, P. Palma, and S. G. Simoes, 'Energy poverty vulnerability index: A multidimensional tool to identify hotspots for local action', Energy Reports, vol. 5, pp. 187–201, Nov. 2019, doi: 10.1016/j.egyr.2018.12.004.

16 A. Arsenopoulos, V. Marinakis, K. Koasidis, A. Stavrakaki, and J. Psarras, 'Assessing Resilience to Energy Poverty in Europe through a Multi-Criteria Analysis Framework', Sustainability, vol. 12, no. 12, p. 4899, Jun. 2020, doi: 10.3390/su12124899.

Based on a set of Key Performance Indicators (KPIs), this standardised approach to metrics and procedures will guarantee IREPO's reproducibility and long-term adoption by policymakers and stakeholders to support and expand current initiatives[17] [18] in energy poverty alleviation.

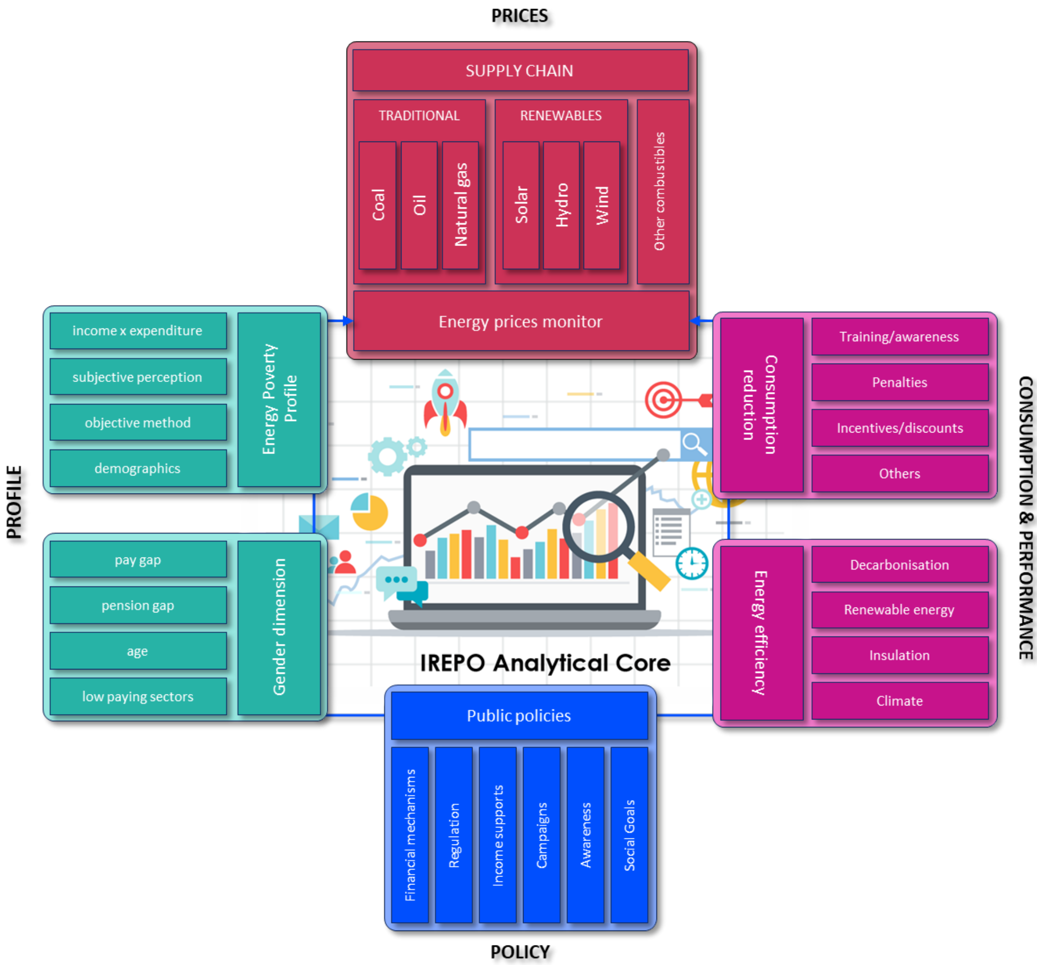


Figure 3 IREPO platform: heigh-level architecture

IREPO's combination of continuous data integration, comprehensive monitoring of energy poverty-related aspects, advanced analytical core using machine learning and statistics, robust expert methodology, and reproducibility, along with a user-friendly web interface, positions it as a reference tool for decision-making, data-driven policy support, and simulations. This innovative approach will establish IREPO as a leader in its field, enabling it to extend its functionalities and services beyond the project's end by fostering long-lasting collaborations across academia, industry, and policymaking. Furthermore, IREPO will enhance the Sustainable Energy Authority of Ireland's (SEAI) prominence and Ireland's scientific capacity and relevance through intense collaboration on the critical issue of energy poverty and its social impact.

17 DCENR, 'A Strategy to Combat Energy Poverty'. Department of Communications, Energy & Natural Resources (DCENR), 2016.: https://bit.ly/3LaOyy1

18 GoI, 'Energy Poverty Action Plan', Government of Ireland (GoI), Ireland, Action Plan 2022, Dec. 2022.: https://bit.ly/423Tk7z

13. Policy and standardisation aspects

In the domain of the energy sector, policy frameworks intersect with artificial intelligence (AI) in unique ways to optimize resource utilization, enhance operational efficiency, and accelerate the transition towards sustainable energy solutions. The integration of AI technologies within energy systems necessitates tailored policy interventions that address sector-specific challenges while harnessing the full potential of advanced data analytics and automation.

The AI Act, tailored for the energy sector, seeks to streamline regulatory procedures and facilitate the adoption of AI-driven solutions for optimizing energy generation, distribution, and consumption. By promoting interoperability and standardization of AI applications, this legislation aims to foster innovation while ensuring compliance with safety, reliability, and environmental standards.

Similarly, the Data Act assumes critical significance in the energy domain by delineating guidelines for the collection, sharing, and utilization of vast volumes of data generated by smart grids, renewable energy sources, and IoT-enabled devices. By promoting data sovereignty and facilitating secure data exchanges, this regulatory framework enables stakeholders to leverage AI algorithms for predictive maintenance, demand forecasting, and grid optimization, thereby enhancing overall system resilience and sustainability.

Privacy considerations in the energy sector revolve around safeguarding sensitive consumer data collected through smart meters, energy management systems, and IoT devices. Policies aimed at protecting consumer privacy rights while facilitating data-driven innovation are essential for fostering trust and promoting widespread adoption of AI-enabled energy solutions.

13.1 Ethics

The ethical and moral implications of AI are a pressing concern demanding immediate attention from diverse stakeholders. While AI ethics is still in its early stages, it extends beyond simplistic notions of right and wrong or good and bad. Addressing these complex issues demands a collaborative effort that transcends the confines of a single group or discipline. Therefore, the need to prioritise the ethics and morality of AI agents in the ongoing development of intelligent systems is essential. To achieve it, the formulation of ethical frameworks to guide the creation of ethical AI is also urgent in the current rapidly evolving AI landscape. As such, it is imperative that we engage in critical discussions, establish robust ethical guidelines, and work collectively to ensure that AI systems are designed and deployed in a manner that aligns with our moral values and principles. To achieve this, some actions are suggested:

International or European Framework[19]: Emphasizing the need for developing AI ethics within an international and/or European framework to ensure consistent ethical standards across borders and industries. Also, the Cybersecurity requirements for high-risk AI systems outlined in the European Commission's proposed AI Act[20] offers a high-level analysis and guiding principles for compliance with the Act. In general, the AI Act applies to AI systems as a whole, not just AI models and, as such, its compliance requires a security risk assessment considering system design to identify and mitigate risks. Even though limitations may exist in securing AI models, compliance can be achieved by mitigating risks through other measures (e.g., an integrated, continuous approach using cybersecurity practices and AI-specific controls).

[19] K. Siau and W. Wang, 'Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI', Journal of Database Management, vol. 31, pp. 74–87, Mar. 2020, doi: 10.4018/JDM.2020040105.

[20] European Commission. Joint Research Centre., Cybersecurity of artificial intelligence in the AI Act: guiding principles to address the cybersecurity requirement for high risk AI systems. LU: Publications Office, 2023.: https://data.europa.eu/doi/10.2760/271009

Government Initiatives[21]: Mentioning the efforts of the German Federal Government's Data Ethics Commission, Digital Council, and Bundestag's Artificial Intelligence Study Commission in working on the ethical positioning of AI, showcasing the importance of governmental involvement in shaping AI ethics such as privacy and informed consent (going beyond GPDR as AI can also reveal personal patterns and habits), energy security (cyber-attacks to the energy grid, and to eventually to the infrastructure), energy equity and affordability (biased AI in smart energy grids), and sustainability (AI models demanding huge processing power, energy consumption, and data storage).

European Commission Guidelines[22]: Referring to the European Commission's presentation of ethical guidelines for trustworthy AI in April 2019, underscoring the significance of regulatory bodies in setting ethical standards for AI systems. More recently, the EU Artificial Intelligence Act, commonly known as the AI Act[23], is a groundbreaking legal framework that aims to regulate the use of AI systems in the European Union. This act is designed to address the risks associated with AI while positioning Europe as a global leader in the development and deployment of AI technologies. The key objectives of the AI Act include providing clear requirements and obligations for AI developers and deployers, reducing administrative and financial burdens for businesses, especially small and medium-sized enterprises, and promoting the development of trustworthy AI that respects fundamental rights, safety, and ethical principles. This comprehensive legislation is part of a broader set of policy measures, including the AI Innovation Package[24] and the Coordinated Plan on AI[25], all geared towards ensuring the safety, rights, and innovation in AI across the EU. The AI Act sets out to create a regulatory environment that fosters responsible AI practices and addresses the potential risks posed by powerful and impactful AI models, ultimately aiming to establish a foundation for trustworthy AI both in Europe and globally.

Discussion and Implementation[26]: Highlighting the necessity to discuss, refine, and translate ethical guidelines into concrete requirements for AI systems, emphasizing the practical enforcement of established AI ethics in real-world applications. Particularly in the context of the EU AI Act, emphasis is on the importance of responsible development and deployment of AI systems. The EU AI Act aims to regulate AI[27,28] systems based on their level of risk, with a focus on ensuring safety, transparency, and accountability. Some ethical key aspects can be highlighted:

- *Transparency and Explainability*: users interacting with AI systems must be informed that they are engaging with automated systems, not humans. High-risk AI systems are required to provide detailed information about their capabilities and limitations to enable informed decision-making by users.

- *Data Quality and Bias Mitigation*: importance of data quality and the mitigation of bias in AI systems by making data be transparent, traceable, and of high quality to avoid biased outcomes and discrimination.

[21] i-nergy, 'I-NERGY AI Ethics Considerations | I-NERGY'. Accessed: May 13, 2024. [Online]. Available: https://i-nergy.eu/aritifial-intelligence-ethics-consideration
[22] H. Bleher and M. Braun, 'Reflections on Putting AI Ethics into Practice: How Three AI Ethics Approaches Conceptualize Theory and Practice', Sci Eng Ethics, vol. 29, no. 3, p. 21, 2023, doi: 10.1007/s11948-023-00443-3
[23] European Commission. (2024, April 23). AI Act | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai
[24] European Commission. (2024, January 24). Commission launches AI innovation package. European Commission - European Commission. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_383
[25] European Commission. (2024, April 30). Coordinated Plan on Artificial Intelligence—Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/policies/plan-ai
[26] UNESCO, 'Ethics of Artificial Intelligence'. Accessed: May 13, 2024.: https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
[27] European Commission, 'Proposal for a Regulation laying down harmonised rules on artificial intelligence | Shaping Europe's digital future'. Accessed: Feb. 07, 2024. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence
[28] European Parliament, 'EU AI Act: first regulation on artificial intelligence'. Accessed: Feb. 07, 2024. [Online]. Available: https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

- *Human Oversight*: high-risk AI systems are mandated to have adequate human intervention and control mechanisms. This ensures that critical decisions are not solely reliant on AI algorithms, maintaining human control over outcomes.

- *Accountability and Compliance*: Developers and providers of AI systems are held accountable for their products under the AI Act. Non-compliance can lead to significant fines, encouraging ethical considerations and responsibility for the societal impact of AI technologies.

- *Innovation and Trustworthiness*: The EU AI Act aims to foster innovation while protecting end-users, promoting the development of trustworthy AI that respects fundamental rights, safety, and ethical principles.

Through explainability, acceptance is created. When decision space is transparent, it is possible to evaluate the process – and also allow for creating or adjusting guardrails, as deemed necessary. An example in the energy sector would be how data sanitation could be applied to sort out or filter missing values from smart meters due to noise or errors in sensor data. Addressing compliance and quality of training data for machine learning models ensure situations like data cleaning through standardisation and imputation are managed and monitored adhering to the proper regulative. Acceptance is also established through informed consent. Governance providing fairness can be demonstrated through how distribution networks may prioritise underserved areas using more accurate predictions.

Another aspect of ethical management is based on responsibility. Not only the moral responsibility of managing energy and assigning processes, but also the corporate responsibility of which practices may cause both societal and environmental harm if not managed properly. Basically, unethical decisions could lead to biased energy access and exploitation of vulnerable communities. This brings up the need for supporting ethical AI development within the energy sector – not only within the design and training process, but within the entire development and maintenance cycle. Thus, inclusivity and empathy through human-centred design should address the needs of the energy grid and stakeholders alike.

Research Projects[29]: Suggesting that research projects should focus on how to enforce established AI ethics in practice, indicating the importance of ongoing research and development in ensuring ethical AI implementation. Examples of initiatives that can accomplish this include good practices and adherence to specific EU Standardisation[30] for AI which aims to create a unified framework, preventing fragmentation and promoting harmonization of AI regulations across sectors. It is crucial to address sector-specific risks during the standardization process to support the upcoming AI Act effectively. Also, alternatives for potential certifications, such as IEEE CertifAIEd[31] - which evaluates the ethics of Autonomous Intelligent Systems (AIS) to enhance product adoption – can be an option to obtain guidance, assessment, and verification, improving AIS quality, building trust with stakeholders, and reaping associated benefits.

These aspects underscore the importance of establishing clear ethical guidelines and frameworks for AI development and deployment, emphasizing the need for collaboration between governmental bodies, regulatory agencies, and research initiatives to ensure ethical AI practices are upheld.

[29] N. K. Corrêa et al., 'Worldwide AI ethics: A review of 200 guidelines and recommendations for AI governance', Patterns (N Y), vol. 4, no. 10, p. 100857, Oct. 2023, doi: 10.1016/j.patter.2023.100857

[30] European Commission, 'AI Standards - European Commission'. Accessed: Feb. 07, 2024. [Online]. Available: https://ai-watch.ec.europa.eu/topics/ai-standards_en

[31] IEEE SA, 'IEEE CertifAIEd', IEEE Standards Association. Accessed: Feb. 07, 2024. [Online]. Available: https://engagestandards.ieee.org/ieeecertifaied.html

13.2 Compatibility with upcoming regulatory framework

The use of general-purpose AI models is becoming more and more common. These models can adapt to perform numerous tasks, but they also pose challenges to oversee all their capabilities. The AI Act[23] introduces transparency obligations for all general-purpose AI models to improve understanding and additional risk management obligations for impactful models. These additional obligations include self-assessment and mitigation of systemic risks, incident reporting, conducting test and model evaluations, and cybersecurity requirements.

For a fast-evolving technology such as AI, the AI Act proposes sort of a "*future-proof*" approach that allows rules to adapt to the evolution of the technology, as long as the applications remain trustworthy even after hitting the market - which will demand constant quality and risk management by solution providers.

The AI Act, which must have entered into force after its publication in March 2024 and voting in May 2024 in the Official Journal[32], should be fully applicable 2 years later. Prohibitions take effect after six months, the governance rules and the obligations for general-purpose AI models become applicable after 12 months, and the rules for AI systems embedded into regulated products become applicable after 36 months.

The European AI Office[33], established in February 2024, supervises the AI Act's enforcement and implementation with the member states, aiming to create an environment where AI technologies respect human rights and trust. It also aims to promote collaboration and research in AI involving various stakeholders to encourage international and potential global engagement towards AI governance. Also, to support future implementation, it invites AI developers from Europe and beyond (a voluntary initiative) to follow the key obligations of the AI Act in advance and promote the new regulatory framework. Similar initiatives are also under development in other jurisdictions, such as the UN's advisory body on AI governance[34], the White House Executive Order and the US AI Safety Institute[35], and the UK's Guidelines[36] for secure AI system development.

13.3 Cybersecurity considerations

It's widely recognised that the development, adoption, and accessibility of AI technology will significantly impact cybersecurity[37]. Threat actors will exploit these technologies to enhance their cyberattack capabilities, improving existing tactics, techniques, and procedures[38] while lowering the technical barriers for cybercriminals, making it easier to launch attacks with less expertise.

This trend also applies to social engineering, which will be bolstered by new large language models (LLMs). These advancements will enable threat actors to create increasingly sophisticated spear-phishing campaigns.

[32] European Commission, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS. 2021. Accessed: Jun. 15, 2024. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206
[33] European Commission, 'European AI Office - Shaping Europe's digital future'. Accessed: Jun. 15, 2024. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/ai-office
[34] United Nations, 'High-Level Advisory Body on Artificial Intelligence | Office of the Secretary-General's Envoy on Technology'. Accessed: Jun. 15, 2024. [Online]. Available: https://www.un.org/techenvoy/ai-advisory-body
[35] NIST, 'U.S. Artificial Intelligence Safety Institute', National Institute of Standards and Technology: https://www.nist.gov/aisi
[36] NCSC, 'Guidelines for secure AI system development', National Cyber Security Centre: https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development
[37] WEF, 'AI and cybersecurity: Navigating the risks and opportunities', World Economic Forum: https://www.weforum.org/agenda/2024/02/ai-cybersecurity-how-to-navigate-the-risks-and-opportunities/
[38] 'The near-term impact of AI on the cyber threat': https://www.ncsc.gov.uk/report/impact-of-ai-on-cyber-threat

Additionally, as AI technology advances, distinguishing between synthetic media and human-generated content will become increasingly difficult. For example, in an assessment developed by the National Cyber Security Centre (NCSC)[39], as per Table 4, the results demonstrate how AI can alter the effectiveness of cyber operations and also how it will impact the cyber threat on over the next two years.

This assessment does not (and cannot) consider significant breakthrough in transformative AI in the referred timeframe. However, such a position must be kept under constant revision as those breakthroughs could bring meaningful impacts for malware and zero-day exploit development. Conversely, the use of AI (or any other AI breakthrough in that aspect) can also impact cyber threats by enhancing cybersecurity resilience with improved detection and overall security design.

While AI will assist in malware development, vulnerability research, and lateral movement, these areas will still heavily rely on human expertise in the short term. AI could potentially create malware that evades detection, and the rapid identification and exploitation of vulnerabilities will strain network managers' ability to patch systems in time.

As AI technology adoption increases, even less-skilled cyber actors will gain significant capabilities, especially through commoditised cyber-crime tools available for purchase.

Overall, the increasing use of AI in cyber operations will heighten the complexity and impact of cyber threats, intensifying challenges for UK cyber resilience in the near term.

Table 4 - Extent of capability uplift caused by AI over next two years[38].

	Highly capable state threat actors	Capable state actors, commercial companies selling to states, organised cyber-crime groups	Less-skilled hackers-for-hire, opportunistic cyber criminals, hacktivists
Intent	High	High	High
Capability	Highly skilled in AI and cyber, well resourced	Skilled in cyber, some resource constraints	Novice cyber skills, limited resource
Reconnaissance	Moderate uplift	Moderate uplift	Uplift
Social Engineering, phishing, passwords	Uplift	Uplift	Significant uplift (from low base)
Tools (malware, exploits)	Realistic possibility of uplift	Minimal uplift	Moderate uplift (from low base)
Lateral Movement	Minimal uplift	Minimal uplift	No uplift
Exfiltration	Uplift	Uplift	Uplift
Implications	Best placed to harness AI's potential in advanced cyber operations against networks, e.g., in advanced malware generation.	Most capability uplift in reconnaissance, social engineering and exfiltration. Will proliferate AI-enabled tools to novice cyber actors.	Lower barrier to entry to effective and scalable access operations - increasing volume of successful compromise of devices and accounts.

[39] NCSC, 'National Cyber Security Centre - NCSC.GOV.UK': https://www.ncsc.gov.uk/

14. Unleashing AI and 6G for Next-Generation Energy Innovations

The energy landscape is undergoing a significant transformation with Renewable Energy Sources (RES) widespread penetration. Driven by the imperative to tackle environmental concerns and reduce greenhouse gas emissions, renewable technologies like solar panels, wind turbines, and hydro generators are stepping into the spotlight as primary energy sources, displacing traditional fossil fuel-based power plants which use coal, oil and gas which are not only environmental unfriendly, but also due to their limited sources can cause energy crisis in the future.

This transition towards sustainable energy solutions is propelled by the convergence of connectivity and technology, facilitating seamless data exchange and management. Arrival of the 6th generation cellular network (6G), poised to revolutionize connectivity with its promise of ultra-fast communication and advanced functionalities beyond communication services.

In the realm of renewable energy, 6G serves as a catalyst for innovation and optimization. Augmented Reality (AR) and Virtual Reality (VR) technologies are transforming maintenance and troubleshooting procedures, empowering technicians to address issues remotely with unprecedented precision. Peer-to-Peer (P2P) energy trading platforms harness 6G's robust connectivity to facilitate seamless transactions between renewable energy producers and consumers, fostering a decentralized energy ecosystem.

The integration of connected Internet of Things (IoT) devices, energy monitoring systems, and Energy Management Systems (EMS) into renewable energy infrastructure is enabled by 6G's unmatched bandwidth and reliability. Automation streamlines operations, optimizing energy production and consumption patterns for maximum efficiency. Cloud computing resources, coupled with real-time weather forecasting capabilities, empower predictive maintenance and dynamic energy management strategies, ensuring resilience in the face of environmental fluctuations.

The anticipated data rates of 6G, are expected to be in the range of terabits per second (Tbps) with ultra-low latency below 1 millisecond (ms), position it in line for supporting United Nations Sustainable Development Goals (UN SDGs). The projected replacement of 5G by 6G around 2030, contingent upon fewer technical constraints, underscores its transformative potential.

The integration of Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) within the 6G network facilitates tailored optimizations and automation of RES power plants. Ultra-fast connectivity underpins communication among RES components, enhancing energy usage and efficiency through techniques like transfer learning (TL) and compressed sensing. AI-driven algorithms, including deep reinforcement learning (DRL), enhance 6G capabilities for diverse applications like vehicular networks, industrial IoT, and ultra-reliable low-latency communications (URLLCs).

DL further amplifies 6G's potential, addressing challenges like visible light communication (VLC) deployment and user mobility limitations. By facilitating intelligent resource provisioning and integrated control, DL enables the seamless integration of IoT applications, driving a data-driven AI system through 6G. These advancements signify a significant leap towards a sustainable, technologically adept future, where AI and 6G synergize to reshape the energy landscape for generations to come.

As we delve deeper into the potential of AI and 6G, novel avenues for innovation emerge. The convergence of these technologies not only revolutionizes energy production and distribution but also opens doors to transformative applications across sectors. In the realm of RESs, AI algorithms are poised to optimize energy production schedules, predict maintenance needs, and enhance grid stability.

Meanwhile, 6G's lightning-fast speeds and ultra-low latency enable real-time monitoring and control of energy assets, ensuring efficient utilization and seamless integration with the broader energy network. AI-driven optimization algorithms possess the agility to dynamically adjust energy production and distribution parameters in response to live data streams, maximizing the utilization of renewable energy sources while concurrently minimizing costs and environmental impact. For instance, ML algorithms scrutinize historical energy usage patterns, weather forecasts, and market dynamics to finely calibrate the scheduling of renewable energy generation and storage resources, ensuring a dependable and economically viable energy supply.

Moreover, AI-fuelled predictive maintenance systems exhibit the foresight to anticipate equipment failures in advance, thus curtailing downtime and optimizing asset management practices. Through the analysis of sensor data gleaned from energy infrastructure components like wind turbines or solar panels, AI algorithms detect anomalies and discern patterns indicative of performance deterioration, enabling pre-emptive maintenance interventions. This proactive approach not only averts costly breakdowns but also ensures optimal system functionality.

Beyond optimization and maintenance, AI and ML applications facilitated by 6G bolster grid resilience and stability by enabling real-time monitoring and regulation of energy flows. Through sophisticated data analytics and control algorithms, energy grids can dynamically adapt to shifts in demand patterns, variations in renewable energy generation levels, and unexpected grid perturbations. This ensures a reliable and steadfast energy supply, even amidst unforeseen contingencies or emergencies.

Moreover, the security and privacy enhancements brought about by 6G are pivotal in fostering trust and confidence in AI-driven systems. As AI becomes increasingly pervasive across critical sectors, robust security measures provided by 6G are essential for safeguarding sensitive data and mitigating cybersecurity threats. By ensuring the integrity and confidentiality of data transmissions, 6G lays a secure foundation for the widespread adoption of AI-powered solutions, driving innovation while maintaining privacy and trust.

Flexibility in energy systems is crucial for accommodating the intermittency of RESs and the variability of energy demand. AI and 6G technologies play a pivotal role in enhancing flexibility on both the demand and generation sides. On the demand side, AI-driven demand response systems empower consumers to adjust their energy consumption patterns in response to supply fluctuations and price signals. Through real-time data analysis and predictive modelling, these systems enable users to optimize their energy usage, reduce peak demand, and participate in demand-side management programs, thereby contributing to grid stability and reliability.

Similarly, on the generation side, AI and 6G technologies enable renewable energy producers to enhance the flexibility and responsiveness of their energy generation assets. AI-based forecasting algorithms leverage weather data, historical patterns, and real-time observations to predict renewable energy generation output with greater accuracy. This allows producers to anticipate fluctuations in energy supply and optimize the operation of their generation assets accordingly. Additionally, 6G connectivity facilitates seamless communication and coordination between distributed energy resources, enabling them to operate in concert to meet fluctuating energy demand and grid requirements.

By empowering users on both the demand and generation sides, AI and 6G technologies foster a more flexible, resilient, and adaptive energy ecosystem. This flexibility not only improves the integration of renewable energy sources but also enhances the overall efficiency and reliability of the grid. As we continue to harness the potential of AI and 6G in energy systems, we pave the way towards a more sustainable and dynamic energy future, where users play an active role in shaping and optimizing the energy landscape.

In essence, the synergy between AI and 6G represents a paradigm shift in how we harness technology to address complex challenges and shape a more sustainable and prosperous future. It drives efficiency, reliability, flexibility and sustainability to unprecedented heights. Further By using the prowess of these cutting-edge technologies, we propel the transition towards a cleaner, smarter, and more resilient energy future. This not only safeguards the environment for future generations but also unveils boundless prospects for technological advancement in the energy domain. By unlocking the full potential of these transformative technologies, we stand poised to usher in an era of unprecedented innovation, where AI and 6G converge to redefine possibilities and propel humanity towards a brighter tomorrow.

Figure 4 ***6G and AI in Energy systems Vertical***

15. Conclusions and Recommendations

Based on the EU and national funded projects analysed in the paper, main conclusions are listed below:

Enhanced data sharing and privacy

Projects like OMEGA-X and Data Cellar focus on creating federated energy data spaces that facilitate data sharing among stakeholders while ensuring privacy, security, and data sovereignty. This approach aims to boost innovation, particularly for SMEs and startups, by providing accessible and secure data environments that can facilitate AI driven services enablement in line with EU AI Act.

Integration of AI in energy management

Several projects, such as ODEON and iFLEX, BD4NRG highlight the critical role of AI in managing energy systems based on large set of use cases. These projects aim to develop AI-driven tools for data processing, demand response management, and the optimization of energy consumption, enhancing overall energy efficiency and consumer engagement.

Support for sustainable energy solutions

Initiatives like PEDVOLUTION and Smart Grid 2.0 are dedicated to developing sustainable energy solutions, such as Positive Energy Districts and resilient electricity grids. These projects focus on enhancing energy efficiency, integrating renewable energy sources, and ensuring grid stability amidst the transition to decarbonized energy systems that can be enhanced with data continuum.

Regulatory frameworks and policy support

The GECKO project addresses the need for adaptable regulatory frameworks to accommodate emerging technologies in transportation and energy sectors, including AI. Project provides guidance and recommendations to ensure the sustainable and competitive development of new business models and technologies. AI driven business models need to evolve.

Combating energy poverty

The Irish Energy Poverty Observatory (IREPO) project aims to tackle energy poverty by developing a comprehensive database that integrates diverse data sources. This initiative uses AI and machine learning to identify patterns and provide data-driven policy support, aiming to alleviate energy poverty through informed decision-making. This is a practical and policy support tool based on AI technologies.

Innovation in energy services

Projects like BD4OPEM, BD4NRG and SYNERGIES focus on leveraging big data and AI to create innovative energy services. These projects aim to enhance grid monitoring, operation, and maintenance, as well as develop new business models and services that drive the energy market towards efficiency and sustainability addressing many use cases.

AI in cybersecurity for energy systems

The NATWORK project explores the intersection of AI, 6G technology, and cybersecurity, aiming to develop energy-efficient AI cybersecurity solutions. This initiative seeks to protect AI-enabled 6G services from energy-oriented sustainability attacks, ensuring the resilience and security of future energy systems.

Collaboration and market integration

Projects such as OneNet and Enershare are enabling extended ecosystems of stakeholders, use cases and digital solutions enabled governance tools where the importance of collaboration among energy actors to create an integrated and efficient electricity network across Europe is strategic core. These projects are already fostering an open and fair energy market approach, interoperable services, and forming "digital spine" to enable AI services, while supporting the broader goal of a sustainable and interconnected energy system. Interconnectors, digital governance mechanisms, energy data spaces focusing on main use cases are at the core.

Role of IIoT and edge technologies

Advancing the EU energy sector through AI services and solutions, IIoT and edge technologies can go hand in hand with focus on sustainability, efficiency, and the integration of new business models and advancements of regulatory frameworks. Dynamic policy feedback is essential for this rapidly evolving field of innovation.

The discussion covers several key impacts of AI on edge and IoT systems, focusing on efficiency, data management, security, and energy consumption.

Efficiency and scalability

IoT devices often have limited bandwidth which can constrain scalability. However, IoT edge mesh architectures resolve this by distributing data across multiple edge devices, which can share and offload computational tasks. this results in faster processing times, better response times, reduced make span, and higher throughput. additionally, if a device fails, other devices can take over its load, enhancing system robustness.

AI/ML techniques tailored to decentralisation, such as hierarchical federated learning, swarm learning, or MARL, play a crucial role in the context of providing energy-efficiency, and improving scalabilities.

Data management and security, Edge AI

Impact of processing data at the edge rather than sending it to the cloud is more efficient and secure, especially for large and sensitive datasets. Edge AI driven approach, known as edge management and orchestration, reduces the volume of data that needs to be transmitted, which in turn lowers energy consumption and enhances data sovereignty and security.

Energy efficiency

the integration of ai methods with IoT/edge computing supports the development of sustainable, optimised and energy efficient systems, helping to optimise energy consumption across the entire value chain of stakeholders through grid operations. As more and more transactions will be taking place at the grid edge, the related solutions and large data management will need to be enhanced and optimised.

Technological integration

The heterogeneity of IIoT energy networks systems, which include various devices and technologies, already requires AI based algorithms that are portable across different environments as demonstrated in a number of projects. This includes the need for lightweight communication protocols and standard interfaces for effective integration and interoperability.

Research is needed to further develop distributed learning algorithms, the enabling platforms and necessary infrastructures that maximise system reliability and availability.

AIOTI has **capacity and ecosystem** for large scale pilots and linking digital infrastructures, computing facilities and DIHs for testing and experimentation.

In addition, Digital Twins and AI can enable optimizing energy management applications. It can be achieved through use of **Federated Digital Twins** to simulate complex interactions in the deregulated digitized energy value chain. To achieve this federation energy companies and solution integrators can leverage the emerging energy data spaces[40] for trusted and interoperable sharing of data between diverse actors in different administrative domains.

Experimentation LLMs and RAG

In an era where **LLMs** are evolving in a rapid pace, further experimentation can be done with using foundational AI models, such as GPT4o, Llama 3.1, Mistral, and RAG (Retrieval Augmented Generation) for synthetic data generation. In addition, forecasting of energy parameters like demand and supply could be greatly facilitated on a large geographical area and in coordination with many actors. It is likely that the results could be surprising when compared to conventional deep learning techniques that are trained and developed on limited domain specific datasets only.

Industrial and consumer metaverse

Further recommendation is for exploring the intersection of **industrial and consumer metaverse** in the energy sector to simulate business models involving prosumers. Such simulations are expected and projected to capture quite well the human acceptance and interaction aspect of the prosumer while blending it with complex industrial processes.

AI sustainability

AI is a resource intensive service and the proliferating adoption of AI in critical infrastructure rapidly increases the need to address the sustainability of AI services and the infrastructure hosting them. That is edge-cloud infrastructure, particularly with the cloudification of AI. The sustainability of AI itself requires addressing challenges in establishing reliable data supply chain with definable trustworthiness levels and preserving of business confidentiality and end-user privacy. Complementary, there is need for developing scalable machine learning architectures, building on existing efforts in federated learning and edge learning. The sustainability of edge-cloud infrastructure training ML pipelines requires consideration of 'where to train model' as the cost, resources and energy types vary. Unnecessary running on the edge risk straining its sparse resources as well as inflate the cost per training task. On the other hand, constraints on data sensitivity and latency for tactile applications may leverage their need to run on the edge. Ultimately, training requirements coupled with user constraints and infrastructure need to be jointly considered in deciding on the deployment of AI workloads in the edge-cloud.

Energy Poverty

In 2022, energy poverty impacted around 40 million Europeans, or approximately 9.3% of the European Union's total population, who struggled to keep their homes adequately warm. This percentage more than doubles among lower-income groups. Currently, over 250 organizations, as well as local, national, and international projects and policies, are working to address energy poverty across Europe.

---

[40] https://op.europa.eu/en/publication-detail/-/publication/21b0260e-a2d5-11ee-b164-01aa75ed71a1/language-en

Despite these efforts, most data and expertise are utilized at a "local" level, with real integration and collaboration remaining weak. AI has the potential to revolutionize the development of more effective energy poverty policies. By integrating the aforementioned data and expertise and leveraging AI algorithms, further research in this area is crucial.

Future research directions

Further research is necessary to enhance federated learning and ai for edge IoT systems. this includes developing architectures and frameworks that support these technologies, improving hardware platforms, and ensuring scalability and efficiency while focusing on security by design and advanced connectivity.

In summary, AI significantly enhances edge and IoT driven systems across energy sector and network topologies by improving efficiency, scalability, and energy consumption while ensuring better data management and security.

The integration of AI at the edge enables faster and more secure data processing, which is essential for the dynamic and heterogeneous nature of IIoT environments.

Contributors

Editor:

Natalie Samovich, AIOTI WG Energy Chair, Enercoutim

Rute Sofia, AIOTI WG Energy co-chair, Fortiss

Reviewer:

Damir Filipovic, AIOTI Secretary General

Contributors:

Aleksandr Zavodovski, University of Oulu

Alemayehu Gebremehdin, NTNU

Andrea Matta, Polimi

Antonio Kung, Trialog

Artur Krukowski, Intracom Telecom

Asbjorn Hovsto, Hafenstrom

Avishek Mukherjee, Hewlett Packard Enterprise

Carolin Zachäus VDI/VDE-IT

Damir Filipovic, AIOTI Secretary General

Eva Pongracz, University of Oulu

Fabio Silva, Tyndall National Institute

Farid Hamzeh Aghdam, University of Oulu

Flemming Svenn, Hafenstrom

Ghazal Etminan, Austrian Institute of Technology GmbH(AIT)

Hannah Funk, VDI/VDE-IT

Ioannis Soldatos, NetCompany Intrasoft

Léo Cornec, Trialog

Mays AL-Naday, University of Essex

Mehdi Rasti, University of Oulu

Natalie Samovich, AIOTI WG Energy Chair, Enercoutim

Nicla Frigerio, POLIMI

Ondrej Cerny, E.DSO

Ovidiu Vermesan, SINTEF

Sergio Gusmeroli, POLIMI

Udayanto Dwi Atmojo, Aalto University

Valerio Frascolla, Intel

Vasileios Karagiannis, Austrian Institute of Technology

Acknowledgements

All rights reserved, Alliance for AI, IoT and Edge Continuum Innovation (AIOTI). The content of this document is provided 'as-is' and for general information purposes only; it does not constitute strategic or any other professional advice. The content or parts thereof may not be complete, accurate or up to date. Notwithstanding anything contained in this document, AIOTI disclaims responsibility (including where AIOTI or any of its officers, members or contractors have been negligent) for any direct or indirect loss, damage, claim, or liability any person, company, organisation or other entity or body may incur as a result, this to the maximum extent permitted by law.

About AIOTI

AIOTI is the multi-stakeholder platform for stimulating AI, IoT and Edge Continuum Innovation in Europe, bringing together small and large companies, academia, researchers, policy makers, end-users and representatives of society in an end-to-end approach. We strive to leverage, share and promote best practices in the AI, IoT and Edge Continuum ecosystems, be a one-stop point of information to our members while proactively addressing key issues and roadblocks for economic growth, acceptance and adoption of the AI, IoT and Edge Continuum Innovation in society. AIOTI contributions go beyond technology and address horizontal elements across application domains, such as matchmaking and stimulating cooperation by creating joint research roadmaps, defining policies and driving the convergence of standards and interoperability.