

## 1 Description of the use case

### 1.1 Name of the use case

ID	Application Domain(s)	Name of Use Case
UC1.Power Grid	Power grid	AI assistant supporting human operators' decision-making in managing power grid congestion

### 1.2 Version management

Version Management			
Version No.	Date	Name of Author(s)	Changes
0.1	29.01.2024	Bruno Lemetayer (RTE)	Initial document (copy from last version of short template document)
0.2	01.03.2024	Bruno Lemetayer (RTE)	Process of all workshop's feedback
0.3	05.04.2024	Bruno Lemetayer (RTE)	Preparation of final version
0.4	11.04.2024	Bruno Lemetayer (RTE)	Finalization of the document
0.5	20.04.2024	Ricardo Bessa (INESC TEC)	Non-functional requirements from ALTAI
0.6	24.04.2024	Cyrill Ziegler (FHNW)	Insertion of Human Factors KPI's
1.0	06.07.2024	Ricardo Bessa (INESC TEC)	Final version

### 1.3 Scope and objectives of use case

Scope and Objectives of Use Case	
<b>Scope</b>	Power grid real-time operation and operational planning (hours-ahead)
<b>Objective(s)</b>	<p>The goal of a Transmission System Operator (TSO), and thus human operators in the control room, is to control electricity transmission on the electrical infrastructure (transmission grid) while pursuing multiple objectives, firstly to keep the system state within acceptable limits and:</p> <ul style="list-style-type: none"> <li>• keeping people and grid components safe,</li> <li>• meeting the production/consumption balance and avoid blackouts,</li> <li>• minimizing operational costs (control actions, energy losses, etc.),</li> <li>• facilitate energy transition (e.g., integration of renewables) by coping with greater uncertainty in forecasts and greater complexity of events and context.</li> </ul> <p>In this context, this use case describes an AI assistant that provides a human operator with recommendations for actions and/or strategies, considering the following objectives:</p> <p><u>Functional aspects</u></p> <ol style="list-style-type: none"> <li>1. Aimed at safely managing overloads on the electrical lines and, more specifically, remedial action recommendations</li> <li>2. Making the most of the renewable energies installed by limiting the emergency redispatching call to thermal power plants emitting greenhouse gases</li> </ol> <p><u>Behavioral and social aspects</u></p> <ol style="list-style-type: none"> <li>3. Easing the workload of the human operator needed to fulfill his/her missions,</li> <li>4. Integrate explainability, transparency, and trust considerations for the human operator.</li> </ol>

	The AI assistant shall also act in a “bidirectional” manner, i.e. capitalize on the actions and the feedback from the operator with an “online” learning process running continuously.
<b>Deployment model</b>	Cloud services, on-premises systems.

#### 1.4 Narrative of use case

<b>Narrative of Use Case</b>	
<b>Short description</b>	
<p>The AI assistant oversees the transmission grid, using SCADA data and available EMS tools to identify issues and categorize them for human intervention. It monitors power flow, voltage, and balance, adhering to defined operational conditions. Anticipating problems, it sends binary alerts to the operator with confidence levels, avoiding excessive alerts to maintain operator focus (i.e., controls attention budget). Action recommendations include topological changes, storage adjustments, redispatching, and renewable energy curtailment. The human operator selects an action or seeks more information, exploring alternatives. After the operator's decision, the AI-assistant provides feedback through load flow calculations, logging decisions for continuous learning and interaction improvement.</p> <p><b>This use case only addresses congestion issues, even if other types of issues can arise on the Transmission Grid and are handled by the operators (e.g., voltage).</b></p> <p><i>Note: Different modes of interaction are possible between AI assistant and human operator, ranging from “full human control” to “full AI control”. The selected mode depends on the industry domain and context. In this use case, an ex-ante choice is made to apply a hybrid interaction where the human operator gets the final word on AI assistant recommendations.</i></p>	
<b>Complete description</b>	
<ol style="list-style-type: none"> <li> <p>The AI assistant monitors the situation of the transmission grid by using the available data from SCADA (Supervisory Control And Data Acquisition) and Energy Management System (EMS) tools and categorizes issues by distinguishing the ones needing intervention by the human operator.</p> <p>The situation of the transmission grid is monitored at the appropriate horizon (e.g., a few hours ahead to 30 minutes ahead) by using relevant forecasts (generation, consumption). Issues correspond to deviations from acceptable operation conditions of the electric system, mainly defined by:</p> <ul style="list-style-type: none"> <li>• Power flow on electric lines not exceeding thermal limits (considering, for instance, a tolerance for temporary overload).</li> <li>• Voltage maintained within a defined range.</li> <li>• Generation and load are always balanced (frequency is maintained around 50 Hz).</li> </ul> <p>The AI assistant monitors these operating conditions and considers a predefined list of contingencies according to the operational policies of the TSO, which include:</p> <ul style="list-style-type: none"> <li>• The nominal grid, i.e., the “N” situation (in which all grid elements are available).</li> <li>• Cases in N situations where overload duration exceeds allowed thresholds: depending on TSO’s operational policies, it can be indeed allowed to let transit flows exceed a temporary threshold on a given line (e.g., flows can be higher than <math>x \cdot A</math> for 20 minutes, after which line will automatically trip). <i>Note: such equipment is used on all lines of RTE’s grid</i></li> <li>• A list of possible “N-1” (electric system’s state after the loss of one grid element and possibly several grid elements depending on the TSO’s policy).</li> </ul> </li> <li> <p>When anticipating issues requiring intervention, the AI assistant raises alerts for decisions at the appropriate horizon (e.g., a few hours ahead down to 30 minutes ahead) to the human operator in time to carry out corresponding actions. These alerts are “binary” in the sense that either the AI assistant sends a persistent alert or not, and they are associated with a level of confidence, i.e., the level of certainty of the AI assistant that the electric system won’t remain within acceptable operation conditions if no action is performed. The level of confidence is based on the uncertainty in the forecasts. The AI assistant should not send too many alerts to keep the human operator concentrated on his or her tasks and thus ease his or her workload.</p> </li> <li> <p>For a given alert, the human operator receives action recommendations from the AI assistant, with information on the predicted effect and reasons for the decision. Possible actions are:</p> </li> </ol>	

- Topological action: topology can be changed by switching power lines on and off or reconfiguring the busbar connection within substations.
  - Redispatching action: change the flexibility's (generator, load, battery, etc.) active setpoint value. Redispatching actions include therefore storage actions (e.g., define the setpoint for charging and discharging storage units such as batteries)
  - Renewable energy curtailment: limits the power output of a given generation unit to a threshold, defined, for example, as the ratio of maximal production Pmax (a value of 0.5 limits the production of this generator to 50% of its Pmax).
4. The human operator chooses a proposed recommendation or requests new information or explanations, or looks for a different action guided by an exploration agent or via manual simulation using other specific tools (that aren't part of the AI assistant).
  5. The human operator performs needed actions according to his/her decision. The AI assistant provides feedback to the human operators on the corresponding effects: this is performed afterward (1 hour or more after the facts) by running a load flow calculation.

The decisions made are logged with their corresponding context to continuously learn from realized actions and improve the interactions between the human operator and the AI assistant (e.g., relevance of proposed recommendations for actions).

#### **Stakeholders**

**TSO:** The transmission system operator is in charge of maintaining and operating the electricity transmission grid, which is monitored by the human operator and the AI assistant.

*Note: This stakeholder includes all the people working for it. For example, the human operator in charge of the operation liaises with other colleagues working, e.g., in maintenance teams on the field.*

**Other TSOs:** Neighboring TSOs are connected to the TSO via its transmission grid.

**Regional Control centers:** Control centers in charge of European operational services and TSO coordination for grid security analysis processes (e.g., TSCnet, Coreso).

**Human operator:** A member of TSO's team who monitors the grid and takes action.

**Transmission grid users:** Any party connected to the transmission grid in a contractual relationship with the TSO. This also includes Distribution System Operators (DSOs) and other critical infrastructures like railways, airports, and water treatment and distribution.

**Market participants:** Any party involved in a market whose physical underlying is electricity delivered to or from the electricity transmission grid, such as (but not limited to) wholesale markets and balancing markets.

#### **Stakeholders' assets, values**

##### **TSO, Other TSOs, Regional Control Centers**

- Legal and regulatory framework of action (e.g. Energy law defining role and missions of the TSO, European network codes).
- The AI system must enhance rather than hinder the TSO's operational competence. Risks involve misinterpretation of data, leading to incorrect decisions that impact the overall efficiency and reliability of the power transmission.
- Use of an AI Assistant by human operators must not lead to a progressive deskilling of human operators, who could lose (or won't acquire in the case of junior operators) the knowledge needed to handle more complex situations where the AI assistant can't provide any recommendation (i.e. ability to provide feedback to the AI)
- Stakeholders (in particular grid users) must trust the AI system's capabilities. Any malfunction or lack of transparency in the AI decision-making process (e.g., excessive curtailment of a renewable energy producer) can erode trust in the TSO and its ability to manage the transmission grid effectively.  
It is, therefore, important to have a recurrent ex-post analysis process within TSOs to analyze the outputs of an AI system to improve confidence and also detect any bias or malfunctions.
- If the AI system's deployment is not communicated effectively or if there are public concerns regarding its use, the TSO's reputation may suffer, potentially affecting public and Energy Regulator support.  
The AI system should contribute to operational efficiency and cost-effectiveness. Moreover, the AI system's recommendations should align with sustainable energy goals.

##### **Human operator**

- Procedures and operation policies that define:
  - Critical boundaries, i.e., events that must be avoided (blackout or electrocution).

- Conditions to be met by the actions (or applicable constraints/limitations), e.g., a given time must be respected between actions on a given line and changes in a generation are limited by ramp-up/down constraints.
- The human operator's decision-making authority is a significant asset. The AI system should complement human expertise.
- The integration of AI may require additional training for human operators.
- The AI system should aim to alleviate the human operator's workload rather than exacerbate it.
- The integration of AI can present opportunities for professional growth.

#### **Transmission Grid users**

- Depend on a reliable power supply, and the AI system must contribute to maintaining grid reliability.
- Sensitive to energy costs, and the AI system's impact on grid operations should aim to optimize efficiency and minimize operational costs.
- Expect transparency in grid operations.

#### **Market Participants**

The AI system's decisions should not favor specific producers unfairly, ensuring a level playing field in the energy market and promoting fair competition.

#### ***System's threats and vulnerabilities***

**Planned and unexpected outage events:** The planned maintenance of the power grid implies that some lines are switched off for some (fixed) duration to allow their maintenance in safe conditions. Even if these events are planned and thus known in advance, they a) degrade the transmission grid's security state and b) increase the probability of damage to the grid device (e.g., the circuit breaker used to switch back on the line). Planned events can also include regular maneuvers on grid devices to check their operating status. Grid operation can be affected by events related to equipment failures on the network (e.g., unplanned line tripping) due to aging or extreme weather events or by cyber-attacks that can disconnect the grid's equipment. Both events are external to the AI system and can increase the complexity of the solutions to solve the technical problems. The AI system will be more "exposed" to operating conditions, and the human operator will demand faster and more accurate recommendations.

#### **Dependency on external systems**

1) *Forecasting system:* The uncertainty of forecasts over a look-ahead horizon is intrinsically part of the base decision-making problem (or "MDP" for Markov Decision Process, which defines the environments with states and states transitions) and, therefore, part of this use case. There are several sources of uncertainty, such as weather forecast errors, interpolation errors for higher temporal resolution, or elasticity of demand to market prices. Thus, the AI-assist will make decisions under forecast uncertainty (i.e., forecast errors), which can impact its performance (e.g., generate false alerts) and require expensive corrective actions with forecast updates.

2) *SCADA measurements:* Reliance on SCADA data quality and availability in terms of nodal injections and current grid topology, which introduces vulnerabilities if those sources are compromised or unavailable.

**Adversarial data attacks:** Malicious actors might attempt to manipulate the AI system by introducing misleading data or injecting false information into the recommendation process, e.g., feeding deceptive information about the state of a particular grid node, causing it to recommend inefficient solutions or worsening congestion; or, injection of a sequence of false information to flood the human with requests during peak grid operation times.

**Trust from human operators:** The operational performance of the AI assistant will not be close to 100% of problems solved, which may hinder the confidence and trust of the human operator in the AI recommendations. This will introduce a negative cognitive bias in humans.

**Progressive deviation of environment behavior:** Not only can the system conditions evolve (production type, consumption pattern, etc.), but also the operational rules, the human operators' behavior, or other applicable regulations. This can progressively alter the efficiency of the AI assistant if it is not regularly "updated". The issue can be exacerbated by the fact that such changes happen very incrementally in time.

**A mismatch between AI training and deployment:** Related to UC2. Power Grid "*Sim2Real, transfer from simulation to real-world*", where significant differences exist between the digital environment used to train the AI model and the real operating conditions. This could lead to low robustness and poor performance during execution, e.g., recommendations based on inaccurate assumptions about grid observability and controllable resources.

## 1.5 Key performance indicators (KPI)

Note: the table below is intended to give an exhaustive list of possible KPIs. This list will be narrowed down during the course of the project, and especially during WP4 for evaluation works.

Name	Description	Reference to the mentioned use case objectives
Total operational cost	<p>It is based on the cost of operations of a power grid that includes the cost of a blackout<sup>1</sup>, the cost of energy losses on the grid<sup>2</sup>, and the cost of remedial actions<sup>3</sup>.</p> <p>In order to simplify the computation and without hindering future improvements, it is proposed to define it as a vector whose dimensions represent different units, at least:</p> <ul style="list-style-type: none"> <li>• Number of real-time topological actions (switching actions, etc.) Only unitary actions at each timestep are considered, which means that a tuple action would be counted as two separate actions</li> <li>• Number of redispatching actions (including but not limited to storage)</li> <li>• Sum of redispatched energy volumes</li> <li>• Number curtailment action</li> <li>• Sum of curtailed energy volumes</li> <li>• Immediate Financial costs</li> <li>• Long-term financial costs (e.g., indirect costs due to lifetime decay of circuit breakers)</li> </ul> <p>Further details about cost calculation might be given during the course of the project (e.g., in WP4). <i>Note: The cost of AI system execution is not evaluated here. See requirement E-2.</i></p>	Objectives: 1
Network utilization	<p>It is based on the relative line loads of the network, indicating to what extent the network and its components are utilized.</p> <p>This can be quantified by:</p> <ul style="list-style-type: none"> <li>• For each timestamp, the highest encountered N-1 line's load N line's load</li> <li>• The average of the maximum N-1 line's load and N line's load</li> <li>• For each timestamp, the number of lines where the N-1 line's load is greater than a given threshold (e.g., 1.0)</li> <li>• For each timestamp, the number of lines where the N line's load is greater than a given threshold (e.g., 0.9)</li> <li>• For all timestamps, the energy of overloads, calculated as the power exceeding the line capacity, integrated over the concerned timestamps (in N and N-1 state)</li> </ul>	Objectives: 1
Topological action complexity	<p>It is used to give insights into how many topological actions are utilized: performing too complex or too many topology actions can indeed navigate the grid into topologies that are either unknown or hard to recover from for operators.</p> <p>Metrics for quantifying the topological utilization of the grid:</p> <ul style="list-style-type: none"> <li>• The average number of split substations (gives an indication of the distance to the reference topology)</li> <li>• The average number of substations modified in one timestamp (gives an indication of the complexity of the topological actions)</li> <li>• Number of unique split substations</li> </ul>	Objective: 1

<sup>1</sup> calculated by multiplying the remaining electricity to be supplied by the market price of electricity.

<sup>2</sup> determined by multiplying the energy volume lost due to the Joule effect by the market price of electricity.

<sup>3</sup> the sum of expenses incurred by the actions using flexibilities (e.g. balancing products, curtailment or redispatching), based on the energy volume and underlying flexibility cost.

<i>Name</i>	<i>Description</i>	<i>Reference to the mentioned use case objectives</i>
Assistant alert accuracy	<p>It is based on the number of times the AI assistant agent is right about forecasted issues (e.g., overloads) ahead of time. Moreover, a confusion matrix can be calculated to show:</p> <ul style="list-style-type: none"> <li>• True positive cases: forecast alerts were raised by the AI assistant, and the problem did occur on the transmission grid,</li> <li>• False positive cases: forecast alerts were raised by the AI assistant, but no problem occurred on the transmission grid,</li> <li>• False negative cases: no forecast alert was raised by the AI assistant, but problems occurred on the transmission grid.</li> </ul>	Objectives: 3, 4
Assistant relevance	<p>It is based on an evaluation by the human operator of the relevance of action recommendations provided by the AI assistant and measured by the number of recommendations from the AI assistant effectively used by the human operator. It ranges in [0, 100] with:</p> <ul style="list-style-type: none"> <li>• 0 meaning that no action recommendation from the AI assistant was considered useful by the human operator,</li> <li>• 100 that all action recommendations from the AI assistant were considered useful by the human operator.</li> </ul> <p>The KPI can have values different from 0 and 100 if only a part of the action recommendations from the AI assistant were used by the human operator.</p> <p>The KPI shall distinguish between the “best decision given the information available at the time” and the “best decision in hindsight.” The evaluation shall focus on the first case, i.e., it shall not be done after the facts with full knowledge of the human operator, which was not available at the time.</p>	Objectives: 4
Action recommendation selectivity	<p>This KPI measures how recommended actions from AI assistants contrast among KPIs used for human decisions: this allows us to put recommended actions in perspective with trade-offs used in human decisions.</p> <p>For each recommended action from the AI assistant, this KPI consists of calculating the increase of each of the following KPIs (see above) due to action implementation:</p> <ul style="list-style-type: none"> <li>• Network utilization</li> <li>• Topological action complexity</li> <li>• Total operational costs</li> </ul>	Objectives: 3, 4
Assistant disturbance	<p>It aims to measure if the notifications raised by the AI assistant are disturbing the activity of the human operator. For each notification, the score ranges in [0, 5] with:</p> <ul style="list-style-type: none"> <li>• 0 meaning that the notification was not considered disturbing at all by the human operator,</li> <li>• 5 meaning that the notification was considered as fully disturbing by the human operator.</li> </ul>	Objectives: 3
Workload	<p>It is based on a workload assessment of the AI assistant by the human operators. It shall be determined according to the NASA-TLX<sup>4</sup> methodology or similar<sup>5</sup>.</p>	Objectives: 3

<sup>4</sup> <https://humansystems.arc.nasa.gov/groups/tlx/index.php>

<sup>5</sup> See more recent works about design recommendations to create algorithms with a positive human-agent interaction and foster a pleasant user-experience: <http://hdl.handle.net/1853/61232>

Name	Description	Reference to the mentioned use case objectives
Total decision time	It is based on the time needed to decide overall, thus including the respective time taken by the AI assistant and human operator. This KPI can be detailed in a way that allows distinguishing specifically the time needed by the AI assistant to provide a recommendation.	Objectives: 3, 4
Carbon intensity	<p>It is based on the overall carbon intensity of the action recommendation, calculated as follows:</p> <ul style="list-style-type: none"> <li>• The amount of energy curtailed (or decreased following redispatching action) is split according to generation type with a negative sign</li> <li>• The amount of additional energy yielded by redispatching action is split according to generation type with a positive sign</li> <li>• The netted amount of energy <math>E_i</math> (MWh) is calculated per generation type <math>i</math></li> <li>• Each amount <math>E_i</math> is multiplied by the corresponding emission factor (kgCO<sub>2</sub>/MWh) <math>F_i</math></li> <li>• The score is then calculated as:  <math display="block">\frac{\sum_i E_i \times F_i}{\sum_i E_i}</math> </li> </ul>	Objectives: 2
Trust towards the AI Tool	<p>“(Dis)trust is defined here as a sentiment resulting from knowledge, beliefs, emotions, and other elements derived from lived or transmitted experience, which generates positive or negative expectations concerning the reactions of a system and the interaction with it (whether it is a question of another human being, an organization or a technology)” (Cahour &amp; Forzy, 2009, p. 1261).  The human operators' trust towards the AI tool can be measured using the Scale for XAI (Hoffman et al., 2018) or similar.</p>	Objectives: 3, 4
Human motivation	<p>“Intrinsic motivation is defined as the doing of an activity for its inherent satisfaction rather than for some separable consequence. When intrinsically motivated, a person is moved to act for the fun or challenge entailed rather than because of external products, pressures, or rewards” (Ryan &amp; Deci, 2000, p. 54).  The human operators perceived internal work motivation can be measured by using the Job Diagnostic Survey (Hackman &amp; Oldham, 1974) or similar. The questionnaire needs to be adapted to the AI context (e.g., problem detection with AI assistant).</p>	Objectives: 3, 4
Human control / autonomy over the process	<p>“Autonomy is the degree to which the job provides substantial freedom, independence, and discretion to the employee in scheduling the work and in determining the procedures to be used in carrying it out” (Hackman &amp; Oldham, 1975, p. 162). It consists of three interrelated aspects centered on freedom in decision-making, work methods and work scheduling (Morgeson &amp; Humphrey, 2006). Parker and Grote (2022) view job autonomy interchangeably with job control.  The human operator's perceived autonomy over the process can be measured by using the Work Design Questionnaire (Morgeson &amp; Humphrey, 2006) or similar. The questionnaire needs to be adapted to the AI context (e.g. problem detection with AI assistance).</p>	Objectives: 3, 4

<b>Name</b>	<b>Description</b>	<b>Reference to the mentioned use case objectives</b>
Human learning	Human learning is a complex process that leads to lasting changes in humans, influencing their perceptions of the world and their interactions with it across physical, psychological, and social dimensions. It is fundamentally shaped by the ongoing, interactive relationship between the learner's characteristics and the learning content, all situated within the specific environmental context of time and place, as well as the continuity over time (Alexander et al., 2009). The human operators perceived learning opportunities working with the AI-based system can be measured by using the task based workplace learning scale (Nikolova et al., 2014) or similar. The questionnaire needs to be adapted to the AI context.	Objectives: 3, 4
Decision support for the human operator	Decision support tools should be aligned with the cognitive the decision-making process that people use when making judgments and decisions in the real world and ensure that the human operator retains agency (Miller, 2023). AI decision support tools should, therefore, help people to remain actively involved in the decision-making process (e.g., by helping them critique their own ideas) (Miller, 2023). The decision support for the human operator can be measured based on the criteria for good decision support (Miller, 2023) or similar. The instrument needs to be further developed.	Objectives: 3, 4
Ability to anticipate	“The ability to anticipate. Knowing what to expect, or being able to anticipate developments further into the future, such as potential disruptions, novel demands or constraints, new opportunities, or changing operating conditions” (Hollnagel, 2015, p. 4). The human operator’s ability to anticipate further into the future can be measured by calculating the ratio of (proactively) prevented deviations to actual deviations. In addition, the extent to which the anticipatory sensemaking process of the human operator is supported by AI-based assistant can be measured by using the Rigor-Metric for Sensemaking (Zelik et al., 2010) or similar. The instrument needs to be further developed and adapted to the AI context.	Objectives: 3, 4
Situation awareness	“Situation Awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future” (Endsley, 1988, p. 12). The human operator’s situation awareness can be measured by using the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988) or similar.	Objectives: 3, 4

### 1.6 Features of use case

<b>Task(s)</b>	Planning, prediction, interactivity, and recommendation.
<b>Method(s)</b>	Reinforcement learning has been applied to this use case, but other AI approaches are possible.
<b>Platform</b>	Grid2Op digital environment, completed by an interactive tool allowing human operators to interact with the environment and the AI assistant



**1.7 Standardization opportunities and requirements**

<i>Classification Information</i>
<p><b>Relation to existing standards</b></p> <p><i>ISO/IEC 23894:2023, Information technology — Artificial intelligence — Guidance on risk management.</i> Operating the power grid is a high-stakes task, and therefore, risk management specifically related to AI is fundamental. This standard describes the principles applied to AI, risk management framework, and processes. It is intended to be used in connection (i.e., provides additional guidance for AI) with <i>ISO 31000:2018, Risk management – Guidelines.</i></p> <p><i>ISO/IEC 38507:2022, Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations.</i> This use case aims to augment the human operator (not only skills and knowledge but also its role), not replace him, by recognizing the complementary differences between humans and AI and leveraging them for humans. This will require an analysis of governance implications on the use of AI, namely data-driven problem-solving and adaptive AI systems (i.e., retraining during the operational phase) to new operating conditions and/or human feedback, culture, and values with respect to stakeholders, markets, and regulation.</p> <p><i>ISO/IEC 42001:2023, Information technology – Artificial intelligence – Management system.</i> This standard is the world’s first AI management system standard, providing valuable guidance for this rapidly changing field of technology. It addresses the unique challenges AI poses, such as ethical considerations, transparency, and continuous learning. For organizations, it sets out a structured way to manage risks and opportunities associated with AI, balancing innovation with governance.</p> <p><i>IEEE 7000-2021, IEEE Standard Model Process for Addressing Ethical Concerns during System Design.</i> This standard defines a framework for organizations to embed ethical considerations in concept exploration and development. It promotes collaboration between key stakeholders and ensures ethical values are traceable throughout the design process, impacting the operational concept, value propositions, and risk management. It is applicable to all organizations, regardless of size or life cycle model.</p>
<p><b>Standardization requirements</b></p> <p>Application ontology that leverages agent-oriented AI recommendations to aid power grid operators in solving future problems based on past observations stored in a knowledge database. The first work in this direction was initiated in the French project CAB (Cockpit and Bidirectional Assistant), reference: Amdouni, E., Khouadjia, M., Meddeb, M., Marot, A., Crochepierre, L., Achour, W. (2023, April). Grid2Onto: An application ontology for knowledge capitalization to assist power grid operators. In International Conference On Formal Ontology in Information Systems- Ontology showcases and Demos.</p> <p>In other domains of the energy sector, a good example of the use of ontologies is the Smart Applications REference (SAREF) ontology, a family of standards that enables interoperability between solutions from different providers and among various activity sectors on the Internet of Things and therefore contributes to the development of the global digital market.</p>

**1.9 Societal concerns**

<i>Societal concerns</i>
<p><b>Description</b></p> <p><b>Integration of renewable energy sources (RES):</b> Enable higher integration levels of RES and decarbonization of the economy while maintaining (or improving) the reliability and resilience of the electric power system.</p> <p><b>Resilience to extreme (natural or man-made) events:</b> Climate change is increasing the fragility of the power grid, as well as impacting the power produced by RES. Also, the digitalization of energy systems brings additional cybersecurity concerns to TSOs. These extreme events and cyber threats have not traditionally been considered in reliability standards, which typically consider reasonably probable events and neglect very improbable situations. Presently, power systems might not be sufficiently resilient to high-impact-low-probability events, which are becoming more probable.</p> <p><b>Degree of system autonomy:</b> The power grid is a critical infrastructure impacting the economy, the safety of other infrastructures, and the comfort of humans. Therefore, the type of action space is relevant, particularly if AI is providing recommendations or direct action in the environment. Furthermore, the human operator’s sole ability to operate the grid and associated knowledge shall not be hampered by the AI assistant and should, on the contrary, improve thanks to interaction with the AI assistant: deskilling must be avoided.</p> <p><b>Supervision:</b> External supervision and regulator conformity assessment are present.</p> <p><b>Explainability and transparency:</b> the human operator shall be able to understand the ground basis of action recommendations provided by the AI assistant.</p>
<p><b>Sustainable Development Goals (SGD) to be achieved</b></p> <p>SGD7. Affordable and clean energy / SGD13. Climate action</p>

## 2 Environment characteristics

<i>Characteristics</i>	
<i>Observation space</i>	<p>Partially observable.</p> <p>Mixed: discrete (e.g., for switching device states) and continuous (e.g., for transit flows)</p> <p>Data update rate: real-time (modeled with a 5 min resolution in Grid2Op digital environment)</p> <p>Size: very large (a network with around 100 nodes has more than 4,000 dimensions. For instance, RTE's grid is composed of more than 25,000 nodes and 10,000 lines.)</p>
<i>Action space</i>	<p>Mixed actions (discrete and continuous).</p> <p>Size: large (for a network with around 100 nodes, it has &gt; 65,000 different discrete actions &amp; &gt; 200 continuous actions. For instance, RTE's grid is composed of more than 25,000 nodes and 10,000 lines.)</p> <p>All scenarios happen in an intraday time horizon, meaning not more than a 24-hour forecast period.</p>
<i>Type of task</i>	<p>Human operators and AI assistants act in a sequential environment: the previous decisions can affect all future decisions. The next action of these agents depends on what action they have taken previously and what action they are supposed to take in the future. For example, a choice of short-term remedial action can make a planned future action unavailable.</p>
<i>Sources of uncertainty</i>	<p>Stochastic (load and renewable energy forecasts, unplanned outages).</p>
<i>Environment model availability</i>	<p>Yes (physical laws of electricity).</p>
<i>Human-AI interaction</i>	<p>Full-human control (AI-assisted) for all scenarios. Co-learning (between humans and AI) is specific to scenario 3.</p>

### 3 Technical details

#### 3.1 Actors

<i>Actor Name</i>	<i>Actor Description</i>
AI assistant	AI agents provide assistance to human operators. It takes information from the environment to search for recommendations and aid the human operator. In the training phase, it can act on the environment to evaluate its recommendations. In the evaluation/testing phase, the actions on the environment should be performed by the human operator only.
Human operator	A member of TSO's team is in charge of monitoring the grid and taking action on the environment (see "stakeholders" paragraph).
Environment	The human operator will interact with the Digital Environment (illustrated in the Figure below) and the AI assistant through an interface. It can be a digital environment, which is a digital model of the transmission grid, which includes unplanned events that are modeled as events appearing in predefined moments (defined directly in time series). In a real-world implementation, it is the physical environment.

### 3.2 References of use case

<i>References</i>						
<i>No.</i>	<i>Type</i>	<i>Reference</i>	<i>Status</i>	<i>Impact on use case</i>	<i>Originator / organisation</i>	<i>Link</i>
1	Research paper	“Towards an AI Assistant for Power Grid Operators” DOI: 10.3233/FAIA220191	Public	Framework and principles for designing an AI assistant with bidirectional interactions for control room operators	Antoine Marot, Alexandre Rozier, Matthieu Dussartre, Laure Crochepierre, Benjamin Donnot	In book: HHAI2022: Augmenting Human Intellect <sup>6</sup>
2	AI competition	Paris Region AI Challenge for Energy Transition, Low-carbon Grid Operations, April 2023	Public	The track “Assistant” has inspired the use case	Paris Region, RTE	Paris Region <sup>7</sup>

<sup>6</sup> [https://www.researchgate.net/publication/363763107\\_Towards\\_an\\_AI\\_Assistant\\_for\\_Power\\_Grid\\_Operators](https://www.researchgate.net/publication/363763107_Towards_an_AI_Assistant_for_Power_Grid_Operators)

<sup>7</sup> <https://www.iledefrance.fr/toutes-les-actualites/entreprises-et-chercheurs-participez-au-challenge-ia-pour-la-transition-energetique>

## 4 Step-by-step analysis of use case

### 4.1 Overview of scenarios

Notes regarding scenario and environment data:

- It is specific to scenario #1 and scenario #2.
- Scenario #3 uses scenario 1 data.

Note regarding requirements: The column “requirement” for the scenarios’ steps has been left empty for the moment. That column will get more relevant in later stages of implementation/integration when moving for a field demonstration or to demonstrate a technology with higher maturity.

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
1	Preventive action to grant N or N-1 situation security in case of unplanned outage	<p>The AI assistant raises warnings in anticipation of the human operator and provides associated action recommendations.</p> <p>The AI assistant considers the operational context, which includes planned maintenance operations on the grid, and provides action to ensure grid security if needed.</p> <p><i>Note: a sub-scenario could address the case where the AI assistant can't provide any relevant preventive action and make this clear to the human operator, see UC2.Sim2Real.</i></p>	<p>There is a chance that the system security is not ensured at the forecasted horizon in an N or N-1 situation (for a specific case that could arise) if no action is performed.</p> <p>Thus, the AI assistant proposes actions to the operator.</p>	<p>The AI assistant continuously checks that the transmission grid security is ensured at the appropriate horizons (e.g., from a few hours ahead down to 30 minutes ahead) when considering a list of contingencies defined in the operational policies of the TSO.</p> <p>The transmission grid state (and corresponding security assessment) is forecasted. The Grid system is in a normal situation; there is no contingency (unexpected event on the grid), and N/N-1 situations are secured.</p>	<p>The human operator chooses one of the recommendations provided by the AI assistant. The transmission grid goes into the state as predicted by the AI assistant, which informs the human operator about the transmission grid state following the action performed.</p>
2	AI assistant learns from human operator	<p>The AI assistant updates its list of recommendations with actions that were performed by the human operator.</p>	<p>Decisions of human operators are used to improve the learning of AI assistants in new contexts.</p>	<p>The AI assistant is acting on new episodes that were not seen during training</p>	<p>All new episodes are rerun with an AI assistant trained on these new episodes. The result is compared with AI assistants not trained in these new episodes.</p>

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
3	(Nice to have scenario) Human operator learns from AI assistant	The AI assistant provides feedback to the human operators on his/her actions.	The AI assistant provides feedback on actions performed by the human operator with KPIs comparing the initially recommended action and the action chosen by the operator.	Run scenario 1 from the use case Power Grid Assistant	The human operator wants to replay the scenario to get detailed feedback. The AI assistant provides feedback to the human operator on his/her actions.

## 4.2 Steps of scenario 1

Note: For each step, an example of operational business context is given; this will be further detailed during the definition of scenario data. Here, the scenario starts when handling a planned maintenance operation on the grid at the beginning of an operator's shift.

Step no.	Event	Name of process/ activity	Description of process/ activity Service	Information producer (actor)	Information receiver (actor)	Information Exchanged	Requirement
1	Start	The human operator prepares his/her shift	<p><b>Example of context:</b></p> <p>At 08:00 AM, the previous operator ended his/her shift.</p> <p>The planned outage on line L0 beginning at 09.00 AM requires 2 actions:</p> <ul style="list-style-type: none"> <li>• P1: Change topology in an adjacent substation</li> <li>• P2: Coordinate and validate a transit limitation with a DSO</li> </ul>	(empty)	(empty)	(empty)	(empty)
2	Overload forecasted	The AI assistant raises an alert	<p><b>Example of context:</b></p> <p>At 08:10, the AI assistant raises an alert for a potential overload that could occur starting at 10:00 AM on line L1 (after the N-1 situation): with the current hypothesis and forecasts, the load flow performed on the 10:00 AM situation would result in an overload.</p> <p>This overload, if confirmed, needs remedial action (else operational limits would be violated)</p> <p><i>Note: The time horizon of the scenario might need to be adjusted because for the moment all Grid2Op environment data are 1h horizon</i></p>	AI assistant	Human operator	AIAL	(left empty)



Step no.	Event	Name of process/ activity	Description of process/ activity Service	Information producer (actor)	Information receiver (actor)	Information Exchanged	Requirement
3	Action recommendations	The human operator processes the recommendations	<p><b>Example of context:</b> The AI assistant proposes different possible remedial actions:</p> <ul style="list-style-type: none"> <li>• A.R1: load transfer from DSO (time limit 08:15 AM)</li> <li>• A.R2: change of topology in substation S1 (time limit 09:40 AM)</li> <li>• A.R3: limitation of RES generation (costly, time limit 09:50 AM)</li> </ul> <p>AI assistant indicates A.R2 seems the best option.</p> <p><i>Note: it is more interesting to have both preventive and curative remedial actions</i></p>	AI assistant	Human operator	AIR	(left empty)
4	Time limit for remedial action R1 is reached	The AI assistant raises an alert	<p><b>Example of context:</b> At 08:15 AM, the AI assistant indicates that 1.R1's time limit is reached</p>	AI assistant	Human operator	AIAL	(left empty)
5	Operator's decision	The operator decides to ignore the recommendation R1	<p><b>Example of context:</b> The operator decides to ignore A.R1 and wait</p>	Human operator	AI assistant	D	(left empty)
6	Time limit for preparing the planned outage is reached	The AI assistant raises an alert	<p><b>Example of context:</b> The 2 actions required for planned outage beginning at 09.00 AM have to be done in time</p>	AI assistant	Human operator	AIAL	(left empty)
7	Operator's action	The operator prepares planned outage	<p><b>Example of context:</b> The operator implements action P1:</p> <ul style="list-style-type: none"> <li>• Simulation of flows with changed topology</li> <li>• Action list to change the topology</li> </ul>	Human operator	Environment	HA	(left empty)

Step no.	Event	Name of process/ activity	Description of process/ activity Service	Information producer (actor)	Information receiver (actor)	Information Exchanged	Requirement
8	Unplanned event	An unplanned outage is needed	<p><b>Example of context:</b>            At 08:45 AM, the operator receives a call from the maintenance team.            The risk of an explosion of measuring equipment requires an urgent (ASAP) and unplanned outage.            The operator stops the ongoing actions for the planned outage to deal with the urgent outage and calls the maintenance team in charge of the planned outage to indicate that he has to stop due to another urgent outage.</p> <p><i>Note: unplanned outage could concern either:</i></p> <ul style="list-style-type: none"> <li>• a busbar: the interest is that this outage could impact in turn the list of possible remedial actions, but it might not be realistic to implement it effectively,</li> <li>• or a line, which is a simpler case.</li> </ul>	Environment	AI assistant	E	(left empty)
9	Action recommendations	The human operator processes the recommendations	<p><b>Example of context:</b>            According to the current hypothesis, the outage would result in overload in the N-1 situation at 08:50 (due to the new topology following the urgent outage).            The AI assistant proposes different possible remedial actions:</p> <ul style="list-style-type: none"> <li>• B.R1: change of topology in substation S1</li> <li>• B.R2: change of topology in substation S2</li> </ul> AI assistant indicates B.R2 would make A.R2 remedial action unavailable	AI assistant	Human operator	AIR	(left empty)
10	Operator's decision	The operator decides to implement an action	<p><b>Example of context:</b>            The operator goes for action B.R1</p> <p><i>Note: other combinations of cross-impacts could be imagined, for example, cases where the only possibility is that A.R2 remedial action becomes unavailable and the only possible choice is A.R1</i></p>	Human operator	AI assistant	D	(left empty)

Step no.	Event	Name of process/ activity	Description of process/ activity Service	Information producer (actor)	Information receiver (actor)	Information Exchanged	Requirement
11	Operator's action	The operator prepares unplanned outage	<p><b>Example of context:</b> The operator performs the urgent outage and implements remedial action B.R1 The operator calls the maintenance team in charge of the unplanned outage so that the urgent work can start.</p> <p><i>Note: to be detailed according to what type of grid element is concerned by the outage</i></p>	Human operator	Environment	HA	(left empty)
12	Action recommendations	The human operator processes the recommendations	<p><b>Example of context:</b> At 09:00 AM, the AI assistant proposes to continue with the remaining actions to prepare for the planned outage of line L0</p>	AI assistant	Human operator	AIR	(left empty)
13	Operator's decision	The operator decides to implement an action	<p><b>Example of context:</b> The operator decides to continue with the remaining actions to prepare planned outage of line L0</p>	Human operator	AI assistant	D	(left empty)
14	Operator's action	The operator prepares planned outage	<p><b>Example of context:</b> The operator confirms with DSO that action P2 can be performed The operator implements action P2:</p> <ul style="list-style-type: none"> <li>• Topology with the simulation of agreed load transfer from DSO</li> <li>• DSO contact information</li> </ul> <p>The operator fully disconnects line L0 At 09:20 AM, the operator confirms to the maintenance team that the maintenance work can start.</p>	Human operator	Environment	HA	(left empty)
15	Time limit for remedial action R2 is reached	The AI assistant raises an alert	<p><b>Example of context:</b> At 09:40 AM, overload is still forecasted and A.R2's time limit is reached</p>	AI assistant	Human operator	AIAL	(left empty)
16	Operator's decision	The operator decides to implement an action	<p><b>Example of context:</b> Given that A.R2 is the only available action, the operator decides to perform A.R2</p>	Human operator	AI assistant	D	(left empty)
17	Operator's action	The operator implements an action	<p><b>Example of context:</b> The operator implements A.R2</p>	Human operator	Environment	HA	(left empty)

### 4.3 Steps of scenario 2

<b>Step no.</b>	<b>Event</b>	<b>Name of process/ activity</b>	<b>Description of process/ activity Service</b>	<b>Information producer (actor)</b>	<b>Information receiver (actor)</b>	<b>Information Exchanged</b>	<b>Requirement</b>
1	Start	Run episodes where the AI assistant provides recommendations	The AI assistant is acting on new episodes that were not seen during training	(empty)	(empty)	(empty)	(empty)
2	Action recommendations	The human operator processes the recommendations	(per episode) The AI assistant proposes action recommendations to the operator	AI assistant	Human operator	AIR	(left empty)
3	Operator's decision	The operator decides to implement an action	(per episode) The operator decides to take remedial action.	Human operator	AI assistant	D	(left empty)
4	Operator's preference learning	The AI assistant logs human operator's preferences	(per episode) All operator's decisions are logged for the AI assistant's learning	Human operator	AI assistant	D	(left empty)
5	Evaluation	The AI assistant's learning is evaluated	All new episodes are rerun with an AI assistant trained on these new episodes. The result is compared with the AI assistant not trained in these new episodes.	(empty)	(empty)	(empty)	(empty)

#### 4.4 Steps of scenario 3

<b>Step no.</b>	<b>Event</b>	<b>Name of process/ activity</b>	<b>Description of process/ activity Service</b>	<b>Information producer (actor)</b>	<b>Information receiver (actor)</b>	<b>Information Exchanged</b>	<b>Requirement</b>
1	Start	Run a scenario where the AI assistant provides recommendations	Use scenario 1 from the use case Power Grid Assistant	(empty)	(empty)	(empty)	(empty)
2	Operator's decision	The operator decides to implement an action	The human operator doesn't choose the remedial action recommended by the AI assistant.	Human operator	AI assistant	D	(left empty)
3	Operator's action	The operator implements an action	The operator implements the remedial action	Human operator	Environment	HA	(left empty)
4	AI assistant's instant analysis	The AI assistant provides feedback on actions performed	The AI assistant provides feedback on actions performed by the human operator with KPIs comparing the initially recommended action and the action chosen by the operator.	AI assistant	Human operator	AIAN	(left empty)
5	Replay of scenario	Go back to step #1	The human operator wants to replay the scenario	(empty)	(empty)	(empty)	(empty)
6	Action recommendations	The human operator processes the recommendations	The AI assistant provides recommendations	AI assistant	Human operator	AIR	(left empty)
7	Recommendation simulation	The human operator asks for action simulation	The human operator chooses the recommended action to see its effects / or another recommendation. The AI assistant provides simulated results of the recommended action	AI assistant	Human operator	AS	(left empty)

## 5 Information exchanged

<b>Information exchanged (ID)</b>	<b>Name of information</b>	<b>Description of information exchanged</b>
HA	Action implemented by human operator	Action (e.g., topology) implemented by human operator.
AIAL	AI assistant alert	AI assistant alerts for an overload occurring on one or several grid elements. AI assistant alert for reached time limit of a given action.
AIAN	AI assistant analysis	The AI assistant provides feedback on actions performed to the human operator.
AIR	AI assistant recommendations	List of remedial action recommended by the AI assistant
D	Decision from human operator	Human operator's choice
E	Environment information	Information on the environment, e.g., outages. <i>In case an adversarial agent is used to model unplanned events, this information would be replaced by an "adversarial attack".</i>
NRA	New remedial action	Remedial action that is not known by the AI assistant

## 6 Requirements

Requirements		
Categories ID	Category name for requirements	Category description
Ro	Robustness	It encompasses both its technical robustness (the ability of a system to maintain its level of performance under a variety of circumstances) as well as its robustness from a social perspective (ensuring that the AI system duly takes into account the context and environment in which the system operates). This is crucial to ensure that, even with good intentions, no unintentional harm can occur. <i>Source: EU-U.S. Terminology and Taxonomy for Artificial Intelligence. First Edition</i>
E	Efficiency	The ability of an AI system to achieve its goals or perform its tasks with optimal use of resources, including time, computational power, and data.
I	Interpretability	Make the behavior and predictions of AI systems understandable to humans, i.e., the degree to which a human can understand the cause of a decision. <i>Source: Molnar, Christoph. Interpretable machine learning. Lulu.com, 2020.</i>
Re	Regulatory and legal	The AI system's capacity to meet its objectives while complying with relevant laws, regulations, and ethical standards.
HAO	Human Agency and Oversight	The design phase involves including mechanisms for human intervention and ensuring that people can easily understand and monitor AI systems. During deployment, it means continuous monitoring and evaluation to ensure that the systems act within their ethical boundaries.
DG	Data governance	Rules, processes, and responsibilities to drive maximum value from data-centric products by ensuring applicable, streamlined, and ethical AI practices that mitigate risk and protect privacy.
FAIR	Non-discrimination and fairness	This means that AI systems are developed and used in a way that includes diverse actors and promotes equal access, gender equality, and cultural diversity while avoiding discriminatory impacts and unfair biases that are prohibited by Union or national law. <i>Source: EU AI Act</i>
Acc	Accountability	Relates to an allocated responsibility. The responsibility can be based on regulation or agreement or through assignment as part of delegation. In a systems context, accountability refers to systems and/or actions that can be traced uniquely to a given entity. In a governance context, accountability refers to the obligation of an individual or organization to account for its activities, to complete a deliverable or task, to accept the responsibility for those activities, deliverables, or tasks, and to disclose the results transparently. <i>Source: EU-U.S. Terminology and Taxonomy for Artificial Intelligence. First Edition</i>
Requirement R-ID	Requirement name	Requirement description
Ro-1	Keep electrical grid security	The AI assistant monitors all the contingencies in the list and recommends valid actions that consider all relevant operational constraints to keep the electrical grid operating in a secure state. Thus, the physical constraints and operational limits of the electrical network should be passed to the AI system.
Ro-2	AI informs the human operator about its confidence in the output recommendation ( <i>self-awareness</i> )	Confidence of the recommendation is given by the AI assistant: Is the event really “well known” by the model thanks to its training? or is it out of distribution, and then few or no relevant recommendations can be given? The AI assistant shall indicate its confidence in the effectiveness of its recommendations with clear information, such as green, orange, or red indicators.

Ro-3	Fault tolerance	The AI system must maintain seamless grid operation despite potential failures or malfunctions within the AI infrastructure. This requires establishing robust, thoroughly tested, and efficient fallback mechanisms to ensure uninterrupted functionality.
Ro-4	Reproducibility and traceability of recommendations for <i>post-mortem</i> analysis	All recommendations made by the AI system must be reproducible at a later point, given the same input or specific context/conditions. While the actions recommended by the system do not need to be identical in a strict mathematical sense - acknowledging the variability inherent in distributed computing environments - they should be closely aligned and functionally equivalent, ensuring reliable and predictable outcomes under similar conditions. Moreover, it should be possible to trace back which AI model or rules led to the decision(s) or recommendation(s) of the AI system, which is very relevant for audits from the Energy Regulator.
Ro-5	Adaptability to different operating conditions	The system should be able to adapt to different scenarios or operational conditions without significant degradation in performance (i.e., maintain appropriate levels of stability). The scenarios considered are related to the training examples but are particularly challenging.
Ro-6	Do not increase cybersecurity risk	The AI assistant should not increase the system's overall cybersecurity risk level. It must be closed to adversarial attacks from external parties so that no control is taken over the information provided to the human operator. It must also be designed to prevent any communication with commands of grid components (e.g., opening of circuit breakers).
Ro-7	Keep acceptable performance levels under natural or adversarial perturbations during operation	The training of the AI system should include scenarios with natural or adversarial perturbations in its input/state vector, which can originate from missing or erroneous values from the environment ( <i>or adversarial attacks from agents</i> ).
Ro-8	Robustness to attacks targeting model space and reward function	Reward functions and models should be stored and operated in highly cyber-secure Information Technology (IT) systems. In the event of an attack, the previously trained model could be quickly restored. Model training should be done in a secure and controlled digital environment, and model retraining is possible.
Ro-9	Detect changes in AI behavior	Changes in the AI system should be auditable and controlled by humans. Nevertheless, several supervised and reinforcement learning algorithms have online learning, and it might be difficult to evaluate or detect changes in the AI system. Thus, automatic mechanisms are required to detect data and model shifts.
E-1	Relevance of the recommendations	The AI assistant often becomes confident in its ability to propose relevant recommendations to solve situations and limits its number of warnings to the human operator to help him focus his/her attention.
E-2	Computational efficiency	The AI system must be designed to ensure efficient training and inference capabilities on various computer hardware, from small-scale development setups with limited processing power to configurations involving multiple servers and GPUs.
E-3	Scalability	The AI system's training and inference methodology and algorithms must be designed to scale up for applications in large and realistic electrical networks.
E-4	Adequate training environment	AI-friendly digital environments should be used to train the AI system, which generates high-quality representative data of the environment where the system will be deployed. However, the transfer of knowledge from simulation to the real environment should be carefully designed – see UC2.Power Grid “Sim2Real, transfer AI-assistant from simulation to real-world operation”.



I-1	Action rating	Frame recommendations into different scenarios/strategies, and rate these scenarios based on their consequences, e.g., identify a “robust” strategy that could work in all cases or a “no regret” strategy.
I-2	Transparency during system training	The AI system must exhibit high transparency in its decision-making processes. This necessitates that documentation on the system’s training data, training methods, and scenarios is available and understandable to relevant stakeholders.
I-3	Capacity to explain recommendation(s) to the human operator (and other stakeholders)	Depending on the type of AI model used, different options are possible, such as (non-exhaustive list): a) empirically compare the outcomes of various strategies and evaluate the proposed recommendations against predefined KPIs; b) relate the recommendations with features importance of the state/input vector; c) use inherently interpretable models and/or knowledge distillation to explain the decisions of a more complex/large model. A trade-off between accuracy and interpretability needs to be evaluated.
I-4	Adaptability to different levels of interaction and human operator preferences and experience	Each operator has its own preferences (e.g., one operator can be more risk averse than others): ideally, the AI assistant interacting with a given operator could provide decision support that fits the preferences of this operator but is not necessary of another, especially given the type of situation that can require more attention. Thus, the AI system shall be able to interact with the human operator according to his/her preferences and experience, such as a) fully manual, b) get notified every time an overload is detected, and c) only get notified when the AI assistant is not confident enough.
Re-1	Compliance with existing operational policies	The AI assistant's recommendations comply with operational policies and network codes for power grids.
Re-2	European AI Act	The AI system must be prepared to comply with the regulations and standards stipulated in the European AI Act. This compliance involves adhering to the defined transparency, safety, data governance, and accountability requirements.
Re-3	Transparency to humans in terms of interaction with an AI system	The human operator should be aware of their interaction with an AI or another human. In this case, operators are advised of the AI assistant and, hence, not be confused about whether they interact with a human or AI system.
Acc-1	Allow audits for the AI recommendations and human operator actions	Audits are to be expected, though no formal assessment process is available for software in the power grid domain. The regulator will look at the case if a grid user or electricity market agent has a complaint. This is strongly related to requirements Ro-4 and I-3.
Acc-2	Reporting of potential vulnerabilities, risks, or biases	A database with vulnerabilities, risks, and biases, similar to <a href="#">AI Vulnerability Database</a> should be created. However, the vulnerabilities and risks of other systems, e.g., SCADA, should be evaluated together due to interdependencies with the AI system (e.g., source of input data).
HAO-1	Mitigate addictive behavior from humans	The AI system should operate as a recommender (i.e., one more additional tool to support the human operator's decisions), and all the decisions should be solely taken by the human operator (human-in-command approach). The AI assistant shall not create a craving among the operators to use it. On the other hand, we should maintain credibility and intimacy between the operator and the AI system.
HAO-2	Mitigate de-skilling in the human operators	The usage of the AI system must not lead to de-skilling in the human operators. This requires new metrics that monitor workers' skill levels and provisions for actions to compensate workers' de-skilling. Furthermore, a higher knowledge of the fundamentals behind the AI system can help human operators understand the decision-support process.

DG-1	Processing of human operator data	The AI system can use historical data about human operator actions, employing techniques such as imitation learning. However, it is imperative that this data undergoes complete anonymization, as the identification of individual operators is unnecessary. Including action timestamps is mandatory, ensuring compatibility with a table of operator shifts. Consequently, even when cross-referenced, it should remain impossible to discern the operator's identity or correlate specific actions with individuals (including performance metrics). Additionally, the knowledge database must exclude any actions characterized by poor performance.
FAIR-1	Avoid creating or reinforcing unfair bias in the AI system	The system must not unfairly favor specific producers or consumers of electrical energy. A level playing field in the electricity market, as well as fair competition, must be provisioned. Measures must be implemented to ensure these fairness constraints are observed. Note that: 1) Occurring bias may very well originate from technical or physical limitations of electrical grid operations and hence may (in part or wholly) not be avoidable. 2) Requiring the AI system to adhere to fairness standards that are not required from existing alternative techniques may put it at a disadvantage, especially if those originate from the source of the previous issue.
FAIR-2	Regular monitoring of fairness	Using the physical equations of the power grid, it is possible to compare the decisions made by the AI system and the impact that other grid users would have in solving the technical problem. For instance, <i>ex-post</i> , it is possible to run an optimal power flow with the redispatch costs and compare its solution with the AI system. Having a least-cost solution is the primary goal. Metrics such as Jain's fairness index have been used to evaluate fairness in load shedding <sup>8</sup> and fairness in renewables' curtailment <sup>9</sup> .

<sup>8</sup> F. Moret and P. Pinson, "Energy Collectives: A Community and Fairness Based Approach to Future Electricity Markets," IEEE Trans. Power Syst., vol. 34, no. 5, pp. 3994–4004, Sep. 2019.

<sup>9</sup> M. Z. Liu Liu, A. T. Procopiou, K. Petrou, L. F. Ochoa, T. Langstaff, J. Harding, and J. Theunissen, "On the Fairness of PV Curtailment Schemes in Residential Distribution Networks," IEEE Trans. Smart Grid, vol. 11, no. 5, pp. 4502–4512, 2020.

## 7 Common Terms and Definitions

Common Terms and Definitions	
Term	Definition
TSO – Transmission System Operator	A natural or legal person is responsible for operating, ensuring the maintenance of, and, if necessary, developing the transmission system in a given area and, where applicable, its interconnections with other systems and for ensuring the long-term ability of the system to meet reasonable demands for the transmission of electricity. Source: Directive 2009/72/EC and ENTSOE glossary
SCADA - Supervisory Control And Data Acquisition	A system of different hardware and software elements that come together enables a power grid operator to monitor and control various components of a power system in real time, such as generators, transformers, and transmission lines.
EMS – Energy Management System	Optimal control center solution enables secure, efficient, and optimized electric power system operation.
Nominal grid (“N” situation)	Network operating condition where all grid elements are available
Contingency (“N-1” situation)	Electric system’s state after the loss of one grid element, and possibly several grid elements, depending on the TSO’s policy
Load (or power) flow calculation	Calculations are used to determine the voltage, current, and real and reactive power at various points in a power system under steady-state conditions.
Line’s load	It is defined as the observed current flow divided by the thermal limit of each powerline (no unit): the value is within [0; 1] interval. A line’s load is associated with a line for a given state: it is therefore referred to as “N line’s load” or “N-x line’s load”. <i>Note: this measure is referred to as “rho” in Grid2Op digital environment</i>