1 Description of the use case

1.1 Name of the use case

ID	Application Domain(s)	Name of Use Case
UC2.Power Grid	Power grid	Sim2Real, transfer AI-assistant from simulation to real-world operation

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)	Final review
0.6	24.04.2024	Cyrill Ziegler (FHNW)	Insertion of Human Factors KPI's
1.0	08.07.2024	Ricardo Bessa (INESC TEC)	Final version

1.3 Scope and objectives of use case

Scope and Objectives of Use Case	
Scope	Power grid real-time operation and operational planning (hours-ahead). It integrates the global concept of the assistant framework (developed in UC1.Power Grid) and deepens a specific "real world" complication (in comparison, UC1.Power Grid has a more "theoretical" vision).
Objective(s)	This use case is to assess the capability of an AI assistant to be used for the operation of a "real" transmission grid, in the sense that the "real" environment doesn't exactly behave as the one available to the agent (that is implemented in the AI assistant) during training and simulation procedures, even if they share the same functional properties (same grid components and topology), and operational constraints. Therefore, Sim2Real stands for "Simulation to Reality". The main objectives are: 1. Look at additional technical considerations to succeed at deploying an AI assistant in the real world besides its sole ability to find solutions to simulated situations. 2. Improving human trust when such systems are deployed in real-world environments. 3. Allowing for iterative human-AI refinements with human feedback and insights.
Deployment model	Possible deployment models of AI considered in ISO/IEC TR 24030: cloud services, on-premises systems,

1.4 Narrative of use case

Narrative of Use Case
Short description
The use case outlines two paths for an AI assistant to manage a transmission grid: A) In coping with real-world conditions, the AI assistant monitors grid situations, raises alerts for human intervention, and provides action recommendations, considering uncertainty coming from

bad or low-quality data (e.g., partially missing). The human operator makes decisions based on AI suggestions, with feedback loops to continuously improve interactions and learn from realized actions.

B) When data limitations prevent full autonomy, the AI assistant alerts the human operator due to missing or poor-quality data. The human operator may also choose actions that do not yield expected results due to various factors. In such cases, the operator can provide missing information to aid the AI. Enriched context, including human input and decisions, is logged for continuous learning, enhancing the AI assistant's robustness in making recommendations for grid actions.

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).

Note 1: This use case is linked with the broader notion of "transfer learning", which is the possibility to adapt a pre-trained model to a new environment only with a slight additional training. One of the possible associated research questions is to evaluate the minimum amount of real data that would be needed to align a model with the "real world". In the context of this use case, transfer learning won't be applied, and the model trained in the context of the Power Grid Assistant use case will be used.

Note 2: As for the AI-assistant training, the human operator's decision and perception will rely on "theoretical simulations" (training and simulation tools).

Complete description

The use case can be divided into two paths:

A. The AI assistant copes with real-world conditions

The AI assistant can still carry out its role and provide the human operator with action recommendations, even if data is not of good quality as in training.

1. The AI assistant monitors the transmission grid situation *[same as in UC1.Power Grid]*

2. When anticipating issues requiring intervention, the AI assistant raises alerts for decisions at the appropriate horizon (e.g., a few hours ahead to 30 minutes ahead) to the human operator in time for carrying out corresponding actions *[same as in UC1.Power Grid]*
   The action recommendations from the AI assistant will reflect the additional uncertainty due to bad-quality data and the sensitivity to uncertainty.

3. 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 *[same as in UC1.Power Grid]*

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) *[same as in UC1.Power Grid]*

5. The human operator performs needed actions according to his/her decision *[same as in UC1.Power Grid]*

6. 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) *[same as in UC1.Power Grid]*

B. Real-world conditions require specific interactions between AI assistant and human operator

Available data doesn't allow the AI assistant to provide the human operator with action recommendations in a fully autonomous way and requires the AI assistant to call for additional feedback or information from the human operator: the AI assistant raises an inaccuracy alert.

1. **The first type of situation is where the AI assistant can't evaluate the need for action due to missing and bad-quality** data and thus can't determine any action recommendations. It raises a corresponding alert to the human operator.
   The main reasons can be:
   - Bad or low-quality data:

- Due to uncertainty because the forecasts aren't always accurate or even available, or uncertainty as "epistemic uncertainty", which is the model uncertainty due to sampling (or underrepresentation) problems
    - The state estimator does not directly use the measurement values but first goes through a readjustment. This means that the raw measurement values from the Energy Management System (EMS) can't be directly used to compute the load flow because the needed adaptations (missing or wrong measurement values due to, e.g., measurement device issues) performed by the state estimator will be missing.
- Evolution of the electric system: trends such as higher renewable penetration or consumer behavior change (adaptation) that shift data distribution over the years.

2. **The second type of situation** is where a recommended action doesn't have the expected consequences on the transmission grid's state.
   The main reasons can be:
   - Reproducibility of remedial actions, one or several prerequisites needed to perform an action recommended by the AI-assistant are missing due to:
     - Device failure (e.g., the failure of a circuit breaker might prevent changing the topology as proposed).
     - Unavailability of flexibility (that might prevent performing planned redispatching).
   - Real-time behavior of the transmission grid is significantly different from simulation due to:
     - Different load flow calculation than the one available at training and inference time.
     - Add or upgrade new elements on the grid: substations, lines, etc., even automatic devices.
     - Distributed energy resources (DER) can impact grid congestion and decision-making since they can be a source of additional complexity and difficulty: a model might not be able to analyze or predict the real-world cumulative effect of smaller grid-connected assets.
     - changing grid equipment characteristics (e.g., climate impact or DLR).
     - transient grid dynamics that steady state simulation doesn't capture, for example, in the context of a windstorm.
     - cyber-physical considerations with the integration and modeling of more automatic devices.

3. When the AI assistant can't evaluate the need for action, or a recommended action doesn't have the expected consequences, the human operator can provide the AI assistant with specific missing information to help the AI assistant forecast system state and assess action recommendations.
   This is only possible if the human operator can easily provide missing information to the AI assistant (i.e., it doesn't generate an important additional workload), e.g., the status (open/closed) of a given busbar coupler.

4. The difference between the original context used by the AI assistant and the enriched context is logged to continuously learn from realized actions and improve the robustness and novelty of recommendations for actions by the AI assistant.
   Enriched context includes at least:
   - information given by the human operator.
   - Decisions are made by the human operator (visible as topology changes or other actions on the transmission grid).

Stakeholders
See UC1.Power Grid
Stakeholders' assets, values
See UC1.Power Grid
System's threats and vulnerabilities

Human manipulation: Human operators with malicious intent may attempt to manipulate the AI system by providing misleading feedback or deliberately misusing the AI learning process. It is important to ensure that this co-learning process complies with regulatory requirements and industry standards for power grid management.

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.

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
Technical robustness to real-world imperfections	Describes the ability of the AI system to maintain its performance level under natural or adversarial perturbations, namely bad or low-quality data, or when recommended action does not have the expected impact on the transmission grid's state. This KPI can be quantified by comparing the technical performance of the AI assistant without and with the perturbations, using KPIs from UC1.Power Grid. From those KPIs, the following metrics (or properties) can be computed: 1) The extent to which the output of the AI system or a specific KPI (e.g., total operational cost) varies with the perturbations, e.g., measured with the output/KPI variance and/or average difference. 2) Assess whether a particular decision holds for input variation (data quality issue) in the same context. During the training-time of the AI assistant, the slope of the reward/loss function deterioration can also be used to measure technical robustness.	Objectives: 1,2,3
Resilience to real-world imperfections	Ability to prepare for and adapt to changing conditions and withstand and recover (to a "normal" state) rapidly from natural or adversarial perturbations or unexpected changes. The quantification of this KPI can be made with the magnitude and/or duration of reward/loss function performance degradation compared to an unperturbed system for the same context. It can, for instance, be measured by the area between the reward curves of the unperturbed and perturbed AI system. This can be computed during training or operational testing time.	Objectives: 1,2,3
Transferability across fidelity levels	Measures how effectively a policy or model trained in one environment (low-fidelity simulation) performs when applied to different environments (e.g., high-fidelity simulation or real-world operation). Evaluated by directly applying the policy trained in a low-fidelity simulation to a high-fidelity simulation and measuring its effectiveness by computing the KPIs from UC1.Power Grid.	Objectives: 1,2,3
Generalization to different grids	The ability of a policy to perform well in an unseen grid environment was not part of the training experience. Tested by exposing the previously trained AI system to different environments with changed grid elements and observing how well it adapts and performs by determining the KPIs from UC1.Power Grid.	Objectives: 1,2,3
Assistant disturbance	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: • 0 meaning that the notification was not considered disturbing at all by the human operator, • 5 meaning that the notification was considered as fully disturbing by the human operator.	Objectives: 3

Name	*Description*	*Reference to the mentioned use case objectives*
Workload	It is based on a workload assessment of the AI assistant by the human operators. It shall be determined according to the NASA-TLX[1] methodology or similar[2].	Objectives: 3
Assistant self-awareness	It is based on the number of times the AI assistant agent is right about its ability to perform action recommendations ahead of time. Moreover, a confusion matrix can be calculated to show: • True positive cases: AI assistant raises inaccuracy alert indicating it has insufficient data to estimate the state of the grid and it actually doesn't have the required data, • False positive cases: AI assistant raises inaccuracy alert indicating it has insufficient data to estimate the state of the grid, but it actually does have the required data (i.e., it should be confident, but it isn't) • False negative cases: AI assistant doesn't raise inaccuracy alert, but in reality, it can't properly assess the situation (i.e., is falsely confident) *Note: This KPI is the equivalent of the "Assistant alert accuracy" KPI initially defined for UC1 "Power Grid Assistant"*	Objectives: 3
Trust towards the AI Tool	"(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 & 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.	Objectives: 2,3
Human motivation	"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 & Deci, 2000, p. 54). The human operators perceived internal work motivation can be measured by using the Job Diagnostic Survey (Hackman & Oldham, 1974) or similar. The questionnaire needs to be adapted to the AI context (e.g., problem detection with AI-assistance).	Objectives: 2,3
Human control/ autonomy over the process	"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 & Oldham, 1975, p. 162). It consists of three interrelated aspects centered on freedom in decision-making, work methods and work scheduling (Morgeson & 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 & Humphrey, 2006) or similar. The questionnaire needs to be adapted to the AI context (e.g., problem detection with AI assistance).	Objectives: 2,3

[1] https://humansystems.arc.nasa.gov/groups/tlx/index.php
[2] 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
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: 2,3
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: 2,3
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 an 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: 2,3
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: 2,3

1.6 Features of use case

Task(s)	Planning, prediction, interactivity, recommendation, inference.
Method(s)	Reinforcement learning has been applied to this use case, but other AI approaches are possible.
Platform	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

Classification Information
Relation to existing standards
ISO/IEC 23894:2023, Information technology — Artificial intelligence — Guidance on risk management. 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 *ISO 31000:2018, Risk management – Guidelines*. *ISO/IEC 24029-2:2023, Artificial intelligence (AI) — Assessment of the robustness of neural networks — Part 2: Methodology for using formal methods*. Artificial neural networks are generally a building block of AI assistants for power grid operation (see results from L2RPN competitions); thus, methodologies for using formal methods to assess the robustness properties of neural networks are important. This standard is focused on how to select, apply, and manage formal methods to prove robustness properties. The technical report ISO/IEC TR 24029-1:2021 complements this standard and presents an overview of different methods to assess the robustness of neural networks. *ISO/IEC 42001:2023, Information technology – Artificial intelligence – Management system*. 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. *IEEE 7000-2021, IEEE Standard Model Process for Addressing Ethical Concerns during System Design*. 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.
Standardization requirements
Assessment of AI robustness should go beyond artificial neural networks (ISO/IEC 24029-2:2023) and consider other AI models, as well as the communication of this information to the end-user/decision-maker and the interaction between AI and the environment.

1.9 Societal concerns

Societal concerns
Description
Responsibility: Provide the capacity to evaluate the quality of the AI decisions and their corresponding impacts in case of low-quality decisions. Provide mitigation mechanisms to ensure the security, integrity, validity, and accuracy of the AI assistant. **Explainability and transparency**: Disclose to stakeholders the evaluation methods used to assess robustness, explain AI failures (e.g., the impact of input data contamination, communications failure), and allow them to submit test cases and adversarial examples. **Accountability:** Mitigate, detect, and correct erroneous or harmful AI decisions when operating the model. **Safety and security:** The AI system should perform consistently across different scenarios and consider the complexity of the environment in which the AI system will be used. The key question is to understand if technology is fit for its purpose and real-world operating conditions.
Sustainable Development Goals (SGD) to be achieved
SGD7. Affordable and clean energy / SGD13. Climate action

2 Environment characteristics

See UC1.Power Grid.

3 Technical details

3.1 Actors

Actor Name	*Actor Description*
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

References						
No.	*Type*	*Reference*	*Status*	*Impact on use case*	*Originator / organisation*	*Link*
1	AI competition	Paris Region AI Challenge for Energy Transition, Low-carbon Grid Operations, April 2023	Public	The track "Sim2Real" has inspired the use case	Paris Region, RTE	Paris Region[3]

[3] 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

All scenarios happen in an intraday time horizon, meaning not more than a 24-hour forecast period. Scenario 2 is a “nice to have” scenario, which means that it is of less priority than the other scenarios for the project.

Notes regarding scenario and environment data:

- *Scenario #1 uses Power Grid Assistant data from Use Case 1, scenario 1, which is progressively altered (e.g., replace data points by zero or NaN if possible). For harder cases, the following modifications could be:*
    - *a grid element is added or removed on the zone*
    - *generation changes (e.g., increase of RES generation capacity)*
    - *the AI assistant is used in a different zone*
- *It is specific to scenario #2*

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 the integration/development 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	Adaptation to real-world conditions	The AI assistant’s robustness is tested on bad or low data. The situation can worsen to the point where the transmission grid state can’t be estimated properly by the AI assistant, which can’t propose any action recommendation. *Note: other more difficult cases could be:* • *new grid elements on the zone* • *the AI assistant is used on a different transmission grid than in the training phase (transfer learning)*	Issues and inconsistencies are present in the data, and data are also missing. Forecasting of transmission grid state is challenged or can’t even be performed by the AI assistant because the quality of input data is too low and/or the proportion of missing data is too high.	Run scenario 1 from the use case Power Grid Assistant	The recommendations from the AI assistant make the human operator aware of the sensitivity to the uncertainty of recommended actions. All new episodes are rerun with an AI assistant trained on the episodes with altered perception. The result is compared with AI assistants not trained in these conditions.
2	(Nice-to-have scenario) Additional information from the human operator	The effect of actions recommended by the AI assistant is challenged by unexpected events or dynamics, like the shift of distribution (in this scenario, RES generation). Due to the magnitude of change, specific information is needed from the human operator.	The AI assistant has provided one or several action recommendations. The human operator has assessed that the proposed actions are not feasible or didn’t have the expected	The real-time behavior of the transmission grid is significantly different from the simulation.	The AI assistant proposes new alternative actions with the help of information provided by the human operator.

Scenario conditions					
No.	Scenario name	Scenario description	Triggering event	Pre-condition	Post-condition
		Note 1: a subcase could be added where the human operator is not able to provide information to the AI assistant. *Note 2: other cases could be where one or several prerequisites (e.g., data) needed to perform an action recommended by the AI assistant are missing or have changed.* *Note 3: This scenario shares a lot in common with the first scenario of the use case Power Grid Assistant. However, even if it also includes an unplanned event, the one considered is a shift of the distribution of RES generation pattern, which is not an event monitored in the same way as a list of predefined outages. In addition, this scenario also includes the use of additional information from human operators by the AI assistant.*	consequences for the transmission grid's state.		

4.2 Steps of scenario 1

Step no.	Event	Name of process/ activity	Description of process/ activity Service	Information producer (actor)	Information receiver (actor)	Information Exchanged	Requirement
1	Start	Run episodes from scenario 1 from the use case Power Grid Assistant	AI assistant's perception of the environment is altered. *Harder cases could be:* • *a grid element is added or removed,* • *the AI assistant is used in a different zone from the one used in the training*	(empty)	(empty)	(empty)	(empty)
2	Action recommendations	The human operator processes the recommendations	The AI assistant proposes action recommendations to the operator The recommendations from the AI assistant make the human operator aware of the sensitivity to the uncertainty of recommended actions.	AI assistant	Human operator	AIR	(left empty)
3	Unfeasibility of action recommendation	The AI assistant can't provide recommendations	The AI assistant can't propose action recommendations to the operator and indicate the reasons.	AI assistant	Human operator	AIAL	(left empty)
4	Evaluation	The AI assistant's handling of the real world is evaluated	All episodes are rerun with an AI assistant trained on the episodes with altered perception. The result is compared with AI assistants not trained in these conditions to evaluate especially what will be the reaction of the human operator when working with each assistant. *Note: the following distinction shall be between:* • *False Positives: AI assistant doesn't raise inaccuracy alert, but it can't properly assess the situation,* • *False Negatives: The AI assistant indicates it has insufficient data to estimate the state of the grid, but it does have the required data.*	(empty)	(empty)	(empty)	(left empty)

4.3 Steps of scenario 2

*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** (start as scenario 1 from the Power Grid Assistant UC).*

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	**Example of context:** At 08:00 AM, the previous operator ended his/her shift. A planned outage of line L0 starts at 09:00 AM. For this planned outage, a load limitation has been agreed upon beforehand with DSO on a selected set of substations (100 MW max load), knowing that this load is netted with connected RES generation.	(empty)	(empty)	(empty)	(empty)
2	Operator's action	The operator prepares planned outage	**Example of context:** The operator calls the DSO to confirm that the limitation is implemented before the beginning of the outage The operator fully disconnects line L0 The operator confirms to the maintenance team that the maintenance work can start.	Human operator	Environment	HA	(left empty)
3	Overload forecasted	The AI assistant raises an alert	**Example of context:** A potential overload is foreseen (N situation) starting at 12:00 PM on the line L1. This overload, if confirmed, needs remedial action (else operational limits would be violated)	AI assistant	Human operator	AIAL	(left empty)
4	Action recommendations	The human operator processes the recommendations	**Example of context:** AI assistant proposes one possible curative remedial action (the same as the one foreseen during operational planning preparation of the outage): it consists of opening a line L2	AI assistant	Human operator	AIR	(left empty)
5	Overload alert	The AI assistant raises an alert	**Example of context:** The flow on line L1 is increasing and exceeds the admissible flow in the N situation	AI assistant	Human operator	AIAL	(left empty)
6	Operator's decision	The operator decides to implement an action	**Example of context:** The human operator decides to perform the recommended action and opens the line L2, which brings the flow on line L1 back to admissible level	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***
7	Unplanned event	Change of forecasted flows	**Example of context:** The flow on the line is different from what is forecasted This can correspond to real situations, e.g., sudden gusts of wind (or, on the contrary, sudden drops)	Environment	AI assistant	E	(left empty)
8	Overload alert	The AI assistant raises an alert	**Example of context:** The flow on the line L1 exceeds again the admissible flow in N situation	AI assistant	Human operator	AIAL	(left empty)
9	Action recommendations	The human operator processes the recommendations	**Example of context:** AI assistant proposes only one possible curative remedial action: load shedding	AI assistant	Human operator	AIR	(left empty)
10	New information from human operator to AI Assistant	The human operator provides additional information in the context of constraint-solving	*Note: a subcase could be added where the human operator is not able to provide information to the AI assistant* **Example of context:** After analysis, the human operator realizes that the load on the agreed substations exceeds the agreed volume of 100 MW The human operator checks with the DSO that one transformer can be opened (no risk of load shedding) and adds this as a possible remedial action in the AI assistant	Human operator	AI assistant	NINF	(left empty)
11	Action recommendations	The human operator processes the recommendations	**Example of context:** The AI assistant assesses possible actions and recommends going for opening the transformers	AI assistant	Human operator	AIR	(left empty)
12	Operator's decision	The operator decides to implement an action	**Example of context:** The human operator decides to perform the recommended action	Human operator	AI assistant	D	(left empty)
13	Operator's action	The operator implements an action	**Example of context:** The operator opens the transformer, which brings the flow on line L1 back to admissible level	Human operator	Environment	HA	(left empty)

5 Information exchanged

Information exchanged (ID)	*Name of information*	*Description of information exchanged*
HA	Action implemented by a human operator	Action (e.g., topology) implemented by human operator
AIAL	AI assistant alert	AI assistant alert for an overload occurring on one or several grid elements. AI assistant alert for reached time limit of a given action
AIR	AI assistant recommendations	List of remedial action recommended by the AI assistant
D	The decision from a human operator	Human operator’s choice
E	Environment information	Information on the environment, e.g., outages. *In case an adversarial agent is used to model unplanned events, this information would be replaced by an “adversarial attack”.*
NINF	New information	Information related to the environment context 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) and 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. *Source: EU-U.S. Terminology and Taxonomy for Artificial Intelligence. First Edition*
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. *Source: Molnar, Christoph. Interpretable machine learning. Lulu. com, 2020.*
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. *Source: EU AI Act*
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.
Requirement R-ID	**Requirement name**	**Requirement description**
Ro-1	Adaption to increased uncertainty	The AI system should demonstrate the ability to sustain operational stability and decision performance in diverse and partially unpredictable scenarios, such as increased forecasting errors, missing data, unavailable control actions, and delayed measurements.

Ro-2	Network change responsiveness	The AI system must be able to handle changes within the transmission grid infrastructure, such as introducing new grid elements and modifying the grid topology as the electrical grid evolves.
Ro-3	Cognitive load and stress	The AI system shall not increase the complexity of the situation and the associated level of stress for human operators (due to additional misinformation).
Ro-4	Reproducibility of recommendations for *post-mortem* 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	Increase technical robustness to missing or erroneous input data	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 (*or adversarial attacks from agents*).
Ro-6	Robustness to attacks targeting model space and reward function	Reward functions and models should be stored and operated in highly cyber-secure Information Technology 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.
E-1	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.
I-1	Adaptability to different levels of interaction and human operator preferences	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.
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, *ex-post*, 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[4] and fairness in renewables' curtailment[5].
HAO-1	Additional training about AI for human operators	The type of recommendation from this use case is already known by the human (i.e., the same as traditional tools in power system control rooms), but humans should be trained to understand the rationale behind the AI system (e.g., understand how reinforcement learning works) and its limitations.

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
EMS – Energy Management System	Optimal control center solution to enable secure, efficient, and optimized operation of the electric power system.
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.

[4] 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.

[5] 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.