

Towards an AI assistant for human grid operators*

Marot Antoine Rozier Alexandre Dussartre Matthieu
AI Lab AI Lab R&D
RTE France RTE France RTE France
 Paris, France Paris, France Paris, France

Crochepierre Laure Donnot Benjamin
R&D RTE France AI Lab
 Paris, France *RTE France*
Université de Lorraine, CNRS, LORIA Paris, France
 Metz, France

Abstract—Power systems are becoming more complex to operate in the digital age. As a result, real-time decision-making is getting more challenging as the human operator has to deal with more information, more uncertainty, more applications and more coordination. While supervision has been primarily used to help them make decisions over the last decades, it cannot reasonably scale up anymore. There is a great need for rethinking the human-machine interface under more unified and interactive frameworks. Taking advantage of the latest developments in Human-machine Interactions and Artificial intelligence, we share the vision of a new assistant framework relying on an hypervision interface and greater bidirectional interactions. We review the known principles of decision-making that drives the assistant design and supporting assistance functions we present. We finally share some guidelines to make progress towards the development of such an assistant.

Index Terms—assistant, artificial intelligence, human-machine interaction, hypervision

I. INTRODUCTION

From the beginning, power systems have been complex artificial systems to operate. Over time, complexity has been rising and control room operators have been using more and more applications setting side by side on multiple screens to manage the system. While this simple approach to incrementally grow a monitoring alarm-driven management system has been relatively effective until now, we are now reaching a glass ceiling to add in any new application. This is true in terms of physical space, but most importantly, in terms of manageable cognitive load for any human operator. In that regard, human-machine interfaces have indeed been identified as a risk factor to human error [1] and should be considered more closely. In addition, latest research in psychology [2] and neuroscience [3] have shed light on human decision-making process and its limits, which could be a lead to improve in turn decision-making.

Nowadays, the system complexity keeps rising given the advent of intermittent renewable energies on the production side and of prosumers on the demand side, coupled with the globalization of energy markets over a more and more interconnected European grid. Grids are also aging, and grid developments become more limited nowadays. Operators will hence need to operate a system closer to its limits while dealing with additional automaton on the grid inducing cyber-physical dynamics [4]. While there could eventually be a temptation to develop a fully autonomous grid to cope with

that complexity, it fails short for such large critical system operations. Indeed coordination, responsibility, accountability, and explainability are a must when operating such a system and can only be reasonably achieved by humans today: human operators remain key players. Lately, several works have proposed situation awareness frameworks [5] [6] to help augment the operator’s comprehension of safety-critical situations especially. In addition to better information processing, it is also urgent to rethink the operator human-machine interface [5] and interactions to assist the operator’s regular real-time decision making. Rather than having operators adapt to the machine through a technology-centered system engineering design, machine and operator’s could co-adapt following human centered-design [7].

In terms of interfaces, we have seen tremendous innovations in other domains, especially in consumer products such as smartphones, connected homes, social media and social networks, search engines, recommendation systems. They have been well-adopted for displaying the most relevant information to the user on single screens through homogeneous format, while dealing in the background with vast diverse amounts of information and user interactions. The development of such interfaces has also been made possible thanks to the latest developments of Artificial intelligence (AI) in the last decade. These developments enabled more advanced and practical large scale real-time information processing, such as in computer vision [8], image understanding [9], natural language processing [10], recommendations [11]. This now translates into an even more advanced form of interface: an assistant. Assistants were found useful to both improve user’s performance on tasks and enhance group collaboration on a common task [12]. In power systems, the notion of an AI assistant was used lately in [13] [14].

In this paper, we present the design of an AI-infused assistant for the control room operator based on the latest developments in human-computer interaction design (HCID), AI, and decision-making science. We first review the effectiveness and limits of humans and AI in decision making, highlighting their complementarities. We further define an assistant. We then discuss the specificities of an assistant for grid operators. Finally, we devise guidelines for how researchers can develop and test such an assistant.

II. EFFECTIVENESS AND LIMITS/BIAS OF HUMAN AND AI IN TAKING DECISIONS

A. Human decision making

Human Decision-making is primarily a matter of attention and executive control [3]. Taking proper decisions first involves paying attention to the right information in the environment and making sense of it. It further implies selecting relevant actions while inhibiting inappropriate ones, and eventually executing one in a timely-manner. Following dual process theory [2], we can describe 2 underlying fictitious operating and cooperative agents, called System 1 and System 2. System 1 is the fast intuitive and heuristic agent while System 2 is the slow and reasoning agent. System 2 is the one responsible for decisions assisted by System 1 which continuously provides him predictions for action. Most of the time, System 2 just lazily accepts System 1 proposal in usual situations without much more thinking, resulting in successful, quick and cognitively effortless decisions. When confronted to unusual situations however, System 2 can develop more explicit conscious thinking, beyond System 1 predictions, to deliberate and come up with novel and acceptable decisions while cognitively costly.

Young operators will rely more heavily on System 2 and can hence struggle taking any good decision on time for several situations that still appear complex and unusual for them. As they are very focused on trying to make sense of it, they have a narrow attention and might miss important new information. As they learn overtime through appropriate training and feedback becoming expert, their intuitive System 1 grows for that field of expertise, enabling them taking good and quick decisions even more often with ease. For an expert operator, it has become a lot easier to operate a system intuitively, being able to take more decisions as well as decisions in more difficult situations. However, the downside can be overconfidence, overlooking unusual information that would require more deliberation from System 2.

B. Human biases and desirable assistance

Because it relies on fast heuristics and mostly jump to conclusions, System 1 indeed introduces several potential biases which can lead to human errors, hence limiting the effectiveness of human decision-making. Cognitive biases are beautifully summarized in the cognitive bias codex [15] and classified through 4 problems they are trying to circumvent: a limited memory, the need to act fast, the information overload and a lack of meaning. Among possibly damaging biases, we can more specifically list:

- anchoring bias: be over-reliant on the first piece of information we see.
- confirmation bias: tend to pay attention only to information that confirms our preconceptions.
- overconfidence bias: too confident about one's abilities which causes to take greater risks.

- information bias: tendency to seek information when it does not affect action (more information is not always better).
- availability heuristic: overestimate the importance of information that is available.
- ostrich effect: ignore dangerous or negative information.
- outcome bias: judge a decision based on the outcome rather than how exactly the decision was made.

An assistant should hence help the operator avoid falling into those biases by addressing the following needs:

- augment its memory, knowledge retrieval and keep track of latest events
- better filter or highlight information, enhancing attention
- make sense of a situation and give feedback
- make recommendations, possibly handle some tasks or alert on some undesirable expected consequences

Let's now consider what AI in its latest developments could bring in that regard.

C. AI potential for assistance

The latest deep learning revolution demonstrated some impressive practical abilities of AI, being able to digest lots of information, memorize large historical datasets, and learn by imitation to infer quickly effective actions in context. Turing-award Yoshua Bengio recently described current deep-learning AI as a System 1 kind of intelligence [16], while missing System 2 type. It is indeed presented as advanced pattern matching and recognition machines like System 1 [2], being coined as artificial intuition [17]. It however lacks the ability to reason about causality [18], hence lacking understanding and some common sense. Yet Human and AI can be seen as complementary heterogeneous intelligences that could achieve a superior outcome when combined by co-evolving and learning from each other [19]. This is best exemplified by Centaur's chess [20], having humans play with machines, and not against, to reach a superior performance. Used through an assistant, AI seems capable of addressing the needs mentioned previously, providing a complimentary and hence enhanced System 1 to the operator whose System 2 remains in charge of decisions overall.

To be used effectively by any human however, the AI assistant will need to work along with a proper interface. Indeed, while human Systems 1 and 2 are fully integrated into a cognitive system as a whole, AI and human are clearly separated entities at first. Interactions, communication and shared representations [21] need to be defined. Interpretable [22], explainable [23] and trustworthy [24] AI are in that regard attributes that should get integrated into the assistant. The AI should eventually be provably beneficial to humans, pursuing preferably not a fixed pre-determined objective but pursuing operator's fuzzy objectives, by continuously learning its preferences under uncertainty [25]. Moving away from these different maturing fields of AI that should prove useful to create an assistant, let's now define it more precisely.

III. DEFINING AN ARTIFICIAL ASSISTANT

An assistant is an agent that helps in someone's job, supporting him and taking over on agreed tasks when possible.

A. Assistant: balancing assistance, user control and Automation

To make things clearer, we should distinguish an assistant from single assistance functions and from a whole automaton. An assistance function helps the user gather or alert him about some new relevant information. It can take in a user request in some expected format and compute a result to be interpreted. Other ones can also monitor the user state and warn him about risks it does not seem aware of. Situation awareness offers, in that sense, advanced assistance functions. An assistant relies on assistance functions at its core. But importantly, it also engages actively with the user. It offers a unified interface and allows for dynamic bidirectional interactions with the user to cooperate efficiently on task completion. However, in that configuration, the user remains responsible for the proper operations of the system. While some tasks could eventually become automated if they always get delegated to the assistant, there is no explicit goal to automate any particular task in the first place: this is mostly left to the user's choice over time.

Other industrial sectors have also defined different autonomy levels that we can reflect on, like the one from the International Association of Public Transport :

- GoA0: Manual operation with no automatic protection
- GoA1: Manual operation with automatic protection
- GoA2: Semi-automatic operation
- GoA3: Driverless operation
- GoA4: Unattended operation

Many such fields, however, aims at a fully autonomous system without operators. They diverge on that point from our goal of obtaining an augmented operator through an assistant, which is closer to GoA2 level. Beyond that point, GoA3 and GoA4 are targeting automation, and thus ends the comparison to our problem here. GoA1 and GoA2 offers assistance functions discussed previously, but without too much considerations of interface and interactions. An assistant, as we illustrate on Figure1 and later discuss, is yet a subsequent level not described there whose goal is to offer the right balance between user control and autonomy [21] for enhanced decision-making. We will now focus on the interface and interactions that more uniquely defines an assistant.

B. Hypervision: smart interface & information management

Today's supervision leaves the user the cognitive load to prioritize, organize, and link every displayed information and alarms before considering any decision. It can be regarded as a fragmented ecosystem from an operator viewpoint. While it has been manageable for up to ten applications, it becomes impractical with always more information and uncoordinated applications to control under heterogeneous formats. Supervision gives access to the user to every information available without much more processing. However, it does not help deal with the information overload and lack of meaning problems

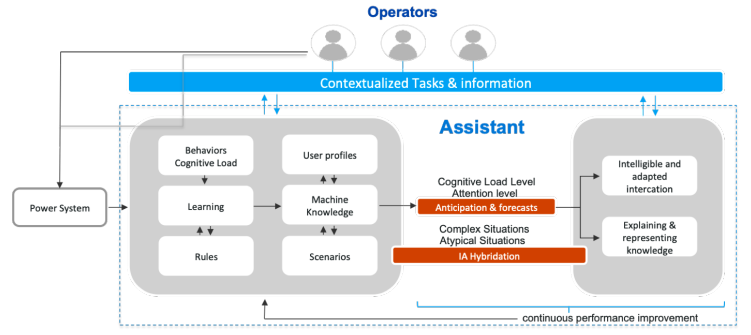


Fig. 1. The AI operator's assistant: Hypervision interface, bidirectional interactions and AI components running altogether in a coordinated and modular fashion. Continuous revision is important to up to date shared representations

that need to be tackled for improved decision making: it dilutes the operator's attention. Let's recall that humans can only take sequential decisions one after another, with a limited working memory space of 4 information to manipulate at a time.

To be effective at continuous decision-making, it is important to focus on one task at a time, with the highest priority, and present only the most relevant information to it. In that regard, we propose an "hypervision" framework to bring the right information at the right time to the right person. It helps overcome multiple biases, such as both information bias and anchoring bias, by taking advantage of them rather than being influenced by them. Hypervision relies on the definition of tasks created by processing and synthesizing the necessary information. Those tasks do not have to be only-real time. However, they are still preferably the ones anticipated to be completed or configured ahead of time thanks to forecast, hence defining an expected trajectory that might be adapted along the way. Reaching this higher level of information enables the assistant to establish a simplified but relevant dialogue with the operator, eventually providing him with diagnostics or even recommendations on solutions. Hence, hypervision's goal is to refocus the operator on task completion rather than alarm monitoring, as illustrated in figure 2. It creates the basis for more advance and effective bidirectional interactions under shared representations of tasks.

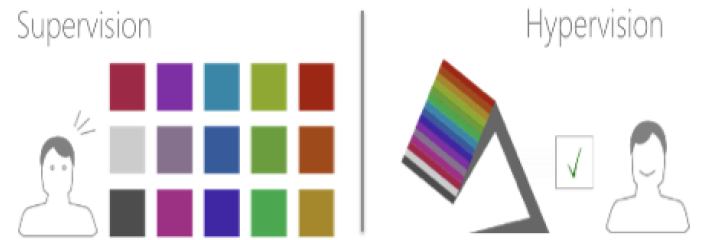


Fig. 2. Alarm monitoring with Supervision - Task completion with Hypervision

C. Bidirectional interactions

While the choice of the interaction modalities (visual, audio, haptic, etc.), as well as the form of the assistant avatar matters for enhanced interactions, we will leave it open here and focus on the importance of bidirectional interactions.

Human-Machine interactions have become a new scientific discipline in the 80's, especially thanks to Lucy Suchman [26]. At a time AI was mostly centered around expert systems with pre-defined rules, she shed light on the ineffectiveness of such systems, mostly attributed to the lack of well-designed interactions and learning loops beyond knowledge retrieval. She noted that plans, similarly to predefined rules, are not prescriptive and not something to follow exactly, because everything eventually depends on circumstances and contingencies. Plans should rather be seen as heuristic and available resources for actions that help focus one's attention while abstracting the details, but that should get updated through interactions to take appropriate decisions. In the end, interfaces should not draw a dry delimitation with its user but rather re-configures itself and conforms with it.

Research [27] has shown an increased efficiency in Human-AI coupling when both agents were able to initiate and respond to interactions. These were historically mostly unidirectional, the assistant either asking a predefined set of questions to build its context representation, or the user asking to perform some predefined tasks. In a bidirectional relationship, the interaction is collaborative, with neither the system nor the user in control of the whole interaction. The assistant is capable of interacting with the latter to refine its context representation (e.g. ask for a clarification when ambiguities arise), thus improving its efficiency when asked to perform a specific task. A good example of such bidirectional interaction is found in [28], where when asked to find the shortest path to evacuate wounded people, the assistant will for instance first ask which vehicles are available, then react accordingly.

While new approaches let an assistant learn how and when to defer to an expert [29], creating true human-computer partnerships becomes a practical reality [30], [21] and are the object of study of the "Cockpit and Bidirectional Assistant" project [31]. We should now review the expected functions of such an assistant.

IV. EXPECTED FUNCTIONS FOR AN ASSISTANT

The paper [32] provides a "unified set of design guidelines" to keep in mind when designing AI-infused assistants, helpful for deciding which features should receive a particular focus. We highlight below a relevant subset of these guidelines:

- **Time services based on context** - Grid operators evolve in a time-constrained environment where having the right information at the right time is paramount. An assistant should engage interaction when the context allows it, based on the operator's cognitive load and impact of the interaction. An assistant should logically engage the operator with a new task if more critical, while not disturbing him from the current one if critical.
- **Show contextually relevant information** - While knowledge databases might suggest to the operator some usual curative action for a given issue, the context of a nearby maintenance, for instance, might make it inefficient. This contextual event should be brought to the operator's attention in that case. Inversely, insensitive context should get filtered out.
- **Support efficient invocation and dismissal** - The number of actions an operator can do in a time period is limited, and interacting with an assistant should be as efficient as possible and not a burden. Should the assistant be in charge of a task, it should not invoke subsequent interactions on it if nothing significant has changed. In tense situations, the assistant should be shorter and dismiss its interaction sooner.
- **Support efficient correction / Encourage granular feedback** - Grid operators are well-trained experts, capable of evaluating the assistant's answers and provide feedback. Thus, the latter must be able to learn from them, for instance, by understanding in hindsight that some additional context needs to be considered to select a curative action or remembering that a line is under maintenance during a defined period.
- **Inform the user about uncertainty in services provided** - If there is too much uncertainty yet when considering some preventive actions, it is reasonable to inform the operator about it so that he waits for the last instant to decide, before the action opportunity expires. Also, the operator should be able to assess any additional risks due to its actions in the next hours. Operators should also get informed of possibly missing or bad quality data and hence uncertain observability. The operator should finally know if any result is deterministic or probabilistic.
- **Scope services when in doubt** - When not yet sure about some action implementation because of uncertainty, the operator can first simply indicate his intention of using such or such flexibility and later decide how he would like to implement it: the assistant should be able to deal with different levels of abstraction. This is also true when giving some contextual information: ahead of time, it might be more relevant to only communicate about aggregated loads in some areas, only sharing individual load values near more certain real-time.
- **Learn from user behaviour / Remember recent interactions** - Operators often have specific decision-making processes, some relying on numerous power-flow simulations to assess a situation, others more akin to rely on their expertise of the considered area, and a good assistant should adjust to these user-specific behaviours. For instance, when supporting a user accustomed to simulations, it could recompute them when the user is busy on the phone to integrate the latest grid changes when he hangs up.
- **Convey the consequences of user actions** - As well as assistants should learn over time, so should operators! It is often deplored that the consequences of unary grid

operations are poorly monitored, which in return, prevent operators from valuable feedback. Assistants delivering a detailed report of how the grid evolved after a specific action would tremendously speed up the way operators acquire experience, and yield better grid management. Interpretability and explainability are required here.

The assistant should eventually help the operator prioritize his tasks thanks to these functions and the hypervision interface. We will now propose some initial guidelines to more concretely design, implement and test such an assistant.

V. GUIDELINES FOR DESIGNING, IMPLEMENTING AND TESTING AN ASSISTANT

Designing an assistant in practice might still seem complex beyond the discussed framework and principles. We devise here some pragmatic guidelines to start simple on a common but modular ground, listing some already available building blocks as well.

A. Grounded Design Considerations

1) *Modeling Tasks as shared representations*: It should be noted that in other industrial sectors such as aeronautics, tasks in processes have been codified more precisely at a granular level, which gives the operator a clearer framework to work and coordinate with, as for the assistant. We should aim at such explicit modeling.

A task is first defined by the problem it needs to solve specifically, such as a safety problem - an overload over a line, its priority and the residual time to complete it. It should then contain relevant context to understand the root of the problem, what might be already known about it, recent related events or tasks, as well as the persons involved. It should further come with some suggestions about available actions to the operators, and their expected effectiveness. It should finally retain a decision for completion and meta-attributes about it. Task categories and attributes should be more exhaustively establish through future works.

Also, opposite to traditional approaches in power systems that mainly tries to focus on the most critical situations we ever have to solve, we suggest here to start studying tasks in regular situations and gradually increasing the number of needed bidirectional interactions. To operators it should prove useful to start with the most basic but sometimes time-consuming tasks with often low added value. That way, building trust in the first place should be easier while still helping ease their cognitive load.

2) *Simple situational use case as a sandbox*: We offer a simple interesting use case to highlight key difficulties in daily grid management through the interplay here of preventive and curative decisions under uncertainty. This makes us think about how the operator-agent interaction should take place.

An operator starts monitoring a grid composed of two smaller areas at 7:00am. Forecasts show that an incident should occur around 9:00am in area 1, with three available corrective actions after simulations, each being able to be executed just before 9:00am. Another incident should happen

around 8:30am in area 2, leaving only a couple of minutes to execute the only preventive action. A couple of questions

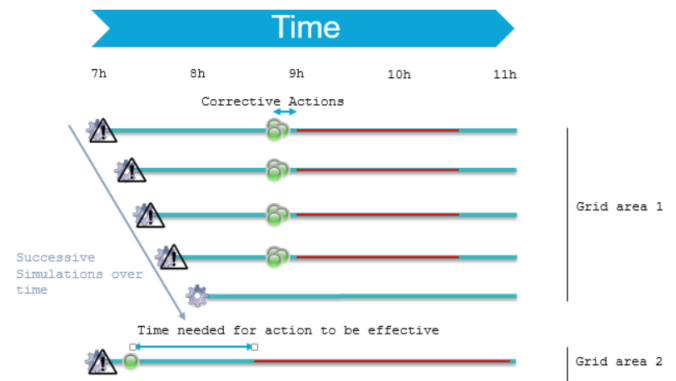


Fig. 3. A simple scenario where incidents are forecasted in multiple parts of a grid, and several corrective actions with different setup durations are possible. Green dots each represent an available action, and the blue arrows the time it needs to be effective. For instance, starting a nuclear power plant takes longer than applying topological grid changes, and would have a longer blue arrow.

arises:

- Which decisions have priority ? It seems that a preventive action on area 2 should be urgently taken, but maybe the forecast isn't that reliable. The assistant should be able to provide the operator with these uncertainties.
- How should the result of the simulations be presented to the operator ? We can see that the last simulation of 8:00am in area 1 shows that the forecasted incident should not happen anymore. Is it a simulation artifact, or has the situation improved with more recent measurements? Maybe the assistant should run successive simulations over time and alert the operator as soon as the forecasted situation evolves.
- How does applying a corrective action on area 1 reflect on area 2 ? Would it lead to a less secure grid state ? What coordination is required ? Maybe there's a new maintenance operation in this area that isn't taken into account by the simulation.
- Which of the three corrective action in area 1 should be taken ? The operator has to mediate between economical, practical and safety arguments, each with a degree of uncertainty. How could the assistant be helpful there ?

Our objective here is not to provide any viable solution, but rather to demonstrate that grid operators are confronted with complex decisions even on apparently simple cases, in which context-dependent trade-offs always have to be made. Future works could build a library of such canonical cases to be studied in the community.

B. Unified Interface & Data collection as an industrial stack

The hypervision framework relies on a generic and single interface that should be able to integrate any kind of tasks, and apply to different industrial systems for instance. While previous supervised applications would still run in the background, hypervision is a master process responsible for

displaying the right information in a proper and standardized format. An example of such an existing framework is the open-source Operator Fabric [33] which could be used both by industrial and researchers as a unified interface for decision-making processes. Such a framework is also a corner-stone to digitalize the decision-making process, centralize every necessary information and hence capitalize on them. This historical data-collection is essential for continuous improvement, experiments, as well as for creating the datasets from which AI can learn recommendations. Data should get labelled and its quality properly monitored. These developments should create a necessary technical stack for an assistant.

C. Power system AI modules for assistant functions

Recent surveys list interesting developments of AI for power systems [34], [35]. For an assistant, AI can today be used to make corrective action recommendations to an operator through adaptive interpretable expert system [14], imitation learning [13] or reinforcement learning [36]. It can learn from user behaviour and help convey the consequences of operator's action by comparison. Exhaustive risk assessment [37] also helps in prioritizing tasks. Further, automatic hierarchical and contextual representations of the grid [38] enable scope services and give greater flexibility to convey the right context and interpret a situation. [39] also lets an AI learn interpretable and physically-consistent contextual indicators associated with a particular operator's task. Finally, [40], [41] and [42] let operators explore interactively and iteratively historical explainable factors across similar situations and decisions for augmenting and keeping up-to-date the system knowledge and proper labels. This is an illustrative sample of today's AI potential to provide effective assistance functions. New developments are needed to augment the assistant functions continuously.

D. Assistant evaluation & development of shared benchmarks

In order to assess the relevance of an assistant in a real-world scenario, and eventually compare multiple assistants, it is necessary to set up repeatable evaluation protocols and define common benchmark tasks.

1) *Evaluation*: As for now, there is not yet a standard testing protocol to evaluate artificial assistants. However, we could draw insights from other domains such as interpretable machine-learning [22] or interactive visualization [43]. As done in [22], we could come to structured and step-by-step experimental practices to evaluate candidate-assistants on incrementally difficult task.

Moreover, several Virtual-Assistant (VA) related studies have also tried to define custom evaluation criteria. For instance in [44], authors compare their VA against both a simpler interactive data-exploration scheme and a non-interactive solution-search. They assess the use of their assistant on three factors :

- **Performance** - Is an operator more efficient with a VA ? Here, task-specific performance metrics are used.

- **Human-learning** - Is a virtually-assisted operator learning as much about the problem and it's underlying model than without the assistant ? This factor is assessed using questions and tests at the end of each task.
- **Usability** - Is a user able to handle the use of the assistant ? This evaluation is performed using a standard System Usability Scale (SUS) [45].

2) *Benchmarks*: Because of confidentiality issues, it is often hard to share real-world data on decision-making problems. Thus, we should aim at developing synthetic but realistic environments from which to extract representative and relevant scenarios. To define these scenarios, we could rely on already existing decision-making assessment frameworks, such as [46] originally designed for road safety. Moreover, synthetic and realistic environments for sequential decision-making have recently been developed for power systems [47] and could be interactively further studied with grid2viz study tool [48].

VI. CONCLUSION

In this paper, we have presented the framework and principles for designing an AI assistant for grid operators, opening new research directions for augmented decision-making. We have provided initial guidelines and already available materials in power systems to start exploring this rich and promising new field of human-machine partnerships.

REFERENCES

- [1] E. Flaspöler, A. Hauke, P. Pappachan, D. Reinert, T. Bleyer, N. Henke, and R. Beeck, "The human machine interface as an emerging risk," *EU-OSHA (European Agency for Safety and Health at Work)*, Luxembourg, 2009.
- [2] D. Kahneman, *Thinking, fast and slow*. Macmillan, 2011.
- [3] L. Naccache, S. Dehaene, L. Cohen, M.-O. Habert, E. Guichart-Gomez, D. Galanaud, and J.-C. Willer, "Effortless control: executive attention and conscious feeling of mental effort are dissociable," *Neuropsychologia*, vol. 43, no. 9, pp. 1318–1328, 2005.
- [4] F. Allgoewer, J. B. de Sousa, J. Kapinski, P. Mosterman, J. Oehlerking, P. Panciatici, M. Prandini, A. Rajhans, P. Tabuada, and P. Wenzelburger, "Position paper on the challenges posed by modern applications to cyber-physical systems theory," *Nonlinear Analysis: Hybrid Systems*, vol. 34, pp. 147–165, 2019.
- [5] M. Naderpour, J. Lu, and G. Zhang, "An intelligent situation awareness support system for safety-critical environments," *Decision Support Systems*, vol. 59, pp. 325–340, 2014.
- [6] M. Panteli and D. S. Kirschen, "Situation awareness in power systems: Theory, challenges and applications," *Electric Power Systems Research*, vol. 122, pp. 140–151, 2015.
- [7] G. A. Boy, "Human-centered design of complex systems: An experience-based approach," *Design Science*, vol. 3, 2017.
- [8] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.
- [9] M. Z. Hossain, F. Sohel, M. F. Shiratuddin, and H. Laga, "A comprehensive survey of deep learning for image captioning," *ACM Computing Surveys (CSUR)*, vol. 51, no. 6, pp. 1–36, 2019.
- [10] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, *et al.*, "Language models are few-shot learners," *arXiv preprint arXiv:2005.14165*, 2020.
- [11] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: A survey and new perspectives," *ACM Computing Surveys (CSUR)*, vol. 52, no. 1, pp. 1–38, 2019.

- [12] R. Winkler, M. Söllner, M. L. Neuweiler, F. Conti Rossini, and J. M. Leimeister, "Alexa, can you help us solve this problem? how conversations with smart personal assistant tutors increase task group outcomes," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2019.
- [13] B. Donnot, I. Guyon, M. Schoenauer, P. Panciatici, and A. Marot, "Introducing machine learning for power system operation support," *arXiv preprint arXiv:1709.09527*, 2017.
- [14] A. Marot, B. Donnot, S. Tazi, and P. Panciatici, "Expert system for topological remedial action discovery in smart grids," 2018.
- [15] "Cognitive bias codex," <https://busterbenson.com/piles/cognitive-biases/>.
- [16] Y. Bengio, "From system 1 deep learning to system 2 deep learning," *Neural Information Processing Systems. December 11th*, 2019.
- [17] C. E. Perez, *Artificial Intuition: The Improbable Deep Learning Revolution*. Carlos E. Perez, 2018.
- [18] J. Pearl, "Theoretical impediments to machine learning with seven sparks from the causal revolution," *arXiv preprint arXiv:1801.04016*, 2018.
- [19] D. Dellermann, P. Ebel, M. Söllner, and J. M. Leimeister, "Hybrid intelligence," *Business & Information Systems Engineering*, vol. 61, no. 5, pp. 637–643, 2019.
- [20] N. Case, "How to become a centaur," *Journal of Design and Science*, 2018.
- [21] J. Heer, "Agency plus automation: Designing artificial intelligence into interactive systems," *Proceedings of the National Academy of Sciences*, vol. 116, no. 6, pp. 1844–1850, 2019.
- [22] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," *arXiv preprint arXiv:1702.08608*, 2017.
- [23] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (xai)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018.
- [24] E. Commission, "White paper on artificial intelligence—a european approach to excellence and trust," 2020.
- [25] S. Russel *et al.*, *Artificial intelligence: a modern approach - 4th Edition*.
- [26] L. A. Suchman, *Plans and situated actions: The problem of human-machine communication*. Cambridge university press, 1987.
- [27] E. Horvitz, "Principles of mixed-initiative user interfaces," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 159–166, 1999.
- [28] J. F. Allen, D. K. Byron, M. Dzikovska, G. Ferguson, L. Galescu, and A. Stent, "Toward conversational human-computer interaction," *AI magazine*, vol. 22, no. 4, pp. 27–27, 2001.
- [29] H. Mozannar and D. Sontag, "Consistent estimators for learning to defer to an expert," *arXiv preprint arXiv:2006.01862*, 2020.
- [30] M. Beaudouin-Lafon and W. E. Mackay, "Rethinking interaction: From instrumental interaction to human-computer partnerships," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–5, 2018.
- [31] "Cockpit and bidirectional assistant project." https://www.irt-systemx.fr/wp-content/uploads/2020/10/SystemX_201015_projetCAB_Final.pdf73.pdf.
- [32] S. Amershi, D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, K. Inkpen, *et al.*, "Guidelines for human-ai interaction," in *Proceedings of the 2019 chi conference on human factors in computing systems*, pp. 1–13, 2019.
- [33] "Operator fabric framework." <https://opfab.github.io/>.
- [34] M. Kezunovic, P. Pinson, Z. Obradovic, S. Grijalva, T. Hong, and R. Bessa, "Big data analytics for future electricity grids," *Electric Power Systems Research*, vol. 189, p. 106788, 2020.
- [35] L. Duchesne, E. Karangelos, and L. Wehenkel, "Recent developments in machine learning for energy systems reliability management," *Proceedings of the IEEE*, 2020.
- [36] A. Marot, B. Donnot, C. Romero, B. Donon, M. Lerousseau, L. Veyrin-Forrer, and I. Guyon, "Learning to run a power network challenge for training topology controllers," *Electric Power Systems Research*, vol. 189, p. 106635, 2020.
- [37] B. Donnot, I. Guyon, A. Marot, M. Schoenauer, and P. Panciatici, "Optimization of computational budget for power system risk assessment," in *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pp. 1–6, IEEE, 2018.
- [38] A. Marot, S. Tazi, B. Donnot, and P. Panciatici, "Guided machine learning for power grid segmentation," in *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pp. 1–6, IEEE, 2018.
- [39] L. Crochepierre, L. Boudjeloud-Assala, and V. Barbesant, "Interpretable dimensionally-consistent feature extraction from electrical network sensors," in *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases ECML/PKDD'20*, 2020.
- [40] A. Marot, A. Rosin, L. Crochepierre, B. Donnot, P. Pinson, and L. Boudjeloud-Assala, "Interpreting atypical conditions in systems with deep conditional autoencoders: the case of electrical consumption," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 638–654, Springer, 2019.
- [41] L. Boudjeloud-Assala, P. Pinheiro, A. Blansch , T. Tamisier, and B. Otjacques, "Interactive and iterative visual clustering," *Information Visualization*, vol. 15, no. 3, pp. 181–197, 2016.
- [42] D. Gkorou, M. Larranaga, A. Ypma, F. Hasibi, and R. J. van Wijk, "Get a human-in-the-loop: Feature engineering via interactive visualizations," 2020.
- [43] R. Borgo, L. Micallef, B. Bach, F. McGee, and B. Lee, "Information visualization evaluation using crowdsourcing," *Computer Graphics Forum*, vol. 37, 2018.
- [44] A. V. i. Martin and D. Selva, "Daphne: A virtual assistant for designing earth observation distributed spacecraft missions," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 30–48, 2020.
- [45] J. Brooke, "Sus: A 'quick and dirty' usability scale," 1996.
- [46] A. Benabbou, D. Lourdeaux, and D. Lenne, "Generation of obligation and prohibition dilemmas using knowledge models," pp. 433–440, 11 2017.
- [47] A. Marot, B. Donnot, C. Romero, L. Veyrin-Forrer, M. Lerousseau, B. Donon, and I. Guyon, "Learning to run a power network challenge for training topology controllers," 12 2019.
- [48] "Grid2viz study tool." <https://github.com/mjothy/grid2viz>.