Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 373–382
https://doi.org/10.54941/ahfe1008270

A Design Methodology for Explainable Supervisory Interfaces Supporting Situation Awareness and Human-AI Collaboration in Augmented SOCs

Kahina Amokrane-Ferka[1], Reda Yaich[1], Emeline Durimel[2], Guillaume Giraud[3], and Christelle Fouquiau[2]

[1]IRT SystemX, 91120, Palaiseau, France
[2]Airebus Protect, 78996 Elancourt, France
[3]RTE, Paris, France

ABSTRACT

The increasing integration of Artificial Intelligence (AI) agents into monitoring environments, such as Security Operations Centers (SOCs), is significantly transforming analysis, decision-making, and incident response activities. While these agents allow the automation of certain processes, they also raise new challenges regarding transparency, oversight, and human–AI coordination. Operators must be able to understand the status and behavior of AI agents, oversee their actions, adjust task delegation, and maintain awareness of their own ability to intervene In this paper, we propose a methodology to design an explainable monitoring dashboard intended to support interaction between a human operator and a set of AI agents. The design principles are based on the concepts of Situation Awareness and Human-AI Teaming, as well as recommendations from research in Human-AI Interaction. The methodology is applied to the context of augmented SOC where an analyst monitors and analyzes security events in real time to identify potential threats, assess alerts, and determine appropriate responses. More specifically, the dashboard is structured around four complementary dimensions: (i) visualizing the operational status of the AI agents, (ii) maintaining a shared understanding of the collaborative activity between humans and the agents, (iii) supporting adjustable task delegation through supervision and reallocation mechanisms, and (iv) taking into account the operator's cognitive and operational state. This approach leads to the definition of an interface that incorporates mechanisms for transparency, explainability, and human control, aimed at improving understanding of AI agents' behavior while keeping the operator at the center of the decision-making loop.

Keywords: Human-AI interaction, Situation awareness, Explainable AI, Security operations center, Supervision dashboard, Human-AI teaming, Adaptive delegation

INTRODUCTION

Human Factors Engineering (HFE) focuses on integrating human cognitive, physical, sensory, and teamwork capabilities into system design throughout its lifecycle. Through task analysis and function allocation, it

Received October 16, 2026; Revised November 21, 2026; Accepted November 27, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

aims to optimize overall system performance, usability, maintenance, and safety while reducing lifecycle costs (Folds et al., 2008). Human Systems Integration (HSI) specialists work within the Systems Engineering process to ensure that human needs are considered from design and development to system retirement. These human-centered approaches have led to significant improvements in workload, maintainability, and personnel safety, avoiding billions of dollars in costs and preventing numerous fatalities and disabling injuries (Booher and Minninger, 2003).

The integration of AI agents into Security Operations Centers (SOCs) transforms analysts' work by automating alert processing and incident response, but also raises concerns about AI opacity and the difficulty of maintaining adequate Situational Awareness (SA). According to Endsley (1995), SA involves perceiving, understanding, and anticipating changes in a dynamic environment, which depends heavily on the quality of information provided by interfaces.

Human–Computer Interaction research, particularly Norman's work (2013), emphasizes that effective interfaces should make available actions discoverable, provide meaningful feedback, and help users develop accurate mental models of system behavior. Similarly, Human-Centered Explainable AI research highlights that explanations should be adapted to users' needs, tasks, and interaction contexts (Doshi-Velez and Kim, 2017; Miller, 2019), using interactive and visual representations to support understanding and decision-making (Ehsan et al., 2021; Mucha et al., 2021).

Based on these principles, this paper proposes a theory-driven interface design framework that translates Situational Awareness requirements for Human–AI Teaming into concrete interface components guided by Human–AI interaction guidelines.

STATE OF THE ART

Situational Awareness and Monitoring of Automated Systems

Situational Awareness (SA) is essential for monitoring complex systems. According to Endsley (1995), it involves three levels: perceiving relevant elements, understanding their meaning, and anticipating their future evolution. In Security Operations Centers (SOCs), maintaining situational awareness is crucial for effective decision-making. However, insufficient visibility into autonomous systems can reduce operator vigilance, impair trust, and negatively affect performance (Sweller, 1988; Parasuraman et al., 2008).

In human–AI collaboration, effective supervision requires a dynamic balance of control between humans and AI, allowing both to take initiative while preserving operator awareness and control (Horvitz, 1999). Interface design plays a key role in supporting this process, as operator performance depends on the quality of the information presented (Wickens et al., 2015).

Research on human–AI teaming (Endsley, 2023) further emphasizes the need to provide visibility into AI agents' states, activities, confidence levels, limitations, behaviors, and intentions, as well as task allocation and

coordination mechanisms. Operators' own cognitive and operational states must also remain visible to support appropriate supervision and timely intervention.

Therefore, interfaces act as cognitive aids that help operators understand AI actions, decisions, and uncertainties. In an augmented SOC, a monitoring dashboard should integrate agent status, task progress, delegation mechanisms, and operator availability to strengthen transparency, situational awareness, and effective human control.

Human-AI Interaction Design Principles

Designing oversight interfaces for AI-augmented environments requires more than traditional usability principles. Interfaces must support system understanding, human oversight, trust calibration, and decision-making by integrating explanations directly into operational workflows.

Human–AI Interaction guidelines, particularly G1, G2, G4, and G11 (Amershi et al., 2019), emphasize making AI capabilities, performance, actions, and reasoning transparent. In SOC environments, this can be achieved through agent status views, confidence indicators, decision justifications, and activity traces. Research on Explainable AI (XAI) further highlights the importance of interactive, contextualized, and user-centered explanations, including natural language, visual highlighting, and multiple explanation formats (Nguyen et al., 2024). These explanations should address users' actual information needs and clarify not only what the AI recommends, but also why, under which assumptions, and with what confidence (Hoffman et al., 2018; Mueller et al., 2021).

Effective supervision also requires interaction history, adaptation to user feedback, and global control mechanisms, as emphasized by guidelines G12, G13, and G17. These principles help operators review recommendations, adjust task allocation, and retain authority over AI-driven processes.

Furthermore, explainable dashboards should provide both a high-level overview and access to detailed information on demand, following the visualization principle of overview first, followed by zooming, filtering, and detailed inspection (Munzner, 2025; Card et al., 1999). Finally, iterative and participatory design, supported by prototyping and stakeholder validation, is essential to ensure that interfaces meet the operational needs and cognitive constraints of SOC operators (Hong and Park, 2025).

PROPOSED DESIGN FRAMEWORK

Problem Statement

The increasing deployment of AI-based agents in interactive systems is profoundly transforming the role of users, which is gradually shifting from direct operators to supervisors of systems. This evolution poses significant challenges for research in human-computer interaction, as users must remain capable of understanding and supervising processes that operate without continuous human intervention. Designing effective human-AI interfaces therefore requires enhancing users' situational awareness,

maintaining an appropriate level of trust in autonomous behaviors, and preserving their ability to intervene quickly when necessary. A key design question is determining what information must remain visible to enable users to form adequate mental models of the system, including the agent's goals, its reasoning process, its confidence levels, its current actions, and its anticipated future behaviors. At the same time, interfaces must define how users can dynamically interact with the system by monitoring, correcting, interrupting, validating, adjusting the level of autonomy and task sharing through appropriate interaction mechanisms.

More generally, the challenge consists of designing collaborative interactions that maintain a balance between system autonomy and human action, while ensuring that increasing automation does not reduce the user's cognitive engagement, control, or ability to participate actively in the decision-making process.

The interface design is based on the assumption that effective human-AI collaboration requires the operator to maintain several complementary forms of situational awareness (SA), as described in (Endsley, 2023).

To translate these cognitive requirements into concrete interaction mechanisms, we based our approach on the guidelines for human-AI interaction recommended by Amershi et al. (2019). Each implemented interface component was therefore designed as a mechanism supporting a specific dimension of SA necessary for effective collaboration between human agents and AI agents, as well as the various recommendations regarding explainable interfaces.

User Case: Adaptive Human-AI Collaboration for Cyber Alert Management

In a cybersecurity operations center (SOC), operators must manage multiple alerts with varying levels of severity, uncertainty, and complexity. AI agents initially process alerts through detection, correlation, and contextual enrichment before displaying them on a centralized dashboard.

An orchestration AI agent then recommends one of four automation modes—manual, collaborative, automated, or autonomous—based on AI confidence, alert uncertainty, and operational criticality. However, operators retain control and can modify the proposed automation level at any time.

In collaborative mode, tasks are dynamically shared between humans and AI according to alert volume, severity, operator workload, and agent availability. In automated mode, AI performs most tasks while requesting human validation for critical actions. Both modes allow AI agents to request human intervention when needed. Operators can also interrupt AI execution and regain control at any time.

Finally, a dedicated interface provides an overview of AI agents' activities, availability, and operational status. This scenario highlights the need for interfaces that support adaptive automation, transparency, dynamic task allocation, human oversight, and controllability, while maintaining operators' Situation Awareness throughout human–AI collaboration.

Supporting Teamwork Situation Awareness (Maintaining Coordination and Human Control)

Effective Human–AI Collaboration Requires **Teamwork Situation Awareness**, which enables operators to monitor team coordination, manage responsibilities, and maintain control over autonomous processes. This awareness ensures that human operators remain actively involved in coordination and can adapt team behavior to changing operational conditions. To support this dimension, guideline G17 (Provide global controls) was implemented.

This guideline motivated the design of the **Alert Delegation Panel** (Figure 1) and the **Dynamic Task Reallocation Panel** (Figure 2). The first allows operators to review and modify the automation level assigned to each alert through drag-and-drop interactions. The second enables real-time adjustments to task allocation between human operators and AI agents according to changing workloads, priorities, and operational constraints.

Additionally, the interface provides explicit control mechanisms to pause or stop AI agents and to validate, modify, or reject AI recommendations requiring human judgment (Figure 3). These features ensure that AI-driven activities remain under human supervision.

Together, these components support Teamwork Situation Awareness by improving coordination, enabling flexible task allocation, and preserving human authority over delegation and AI intervention. This ensures meaningful human control throughout the collaborative process.

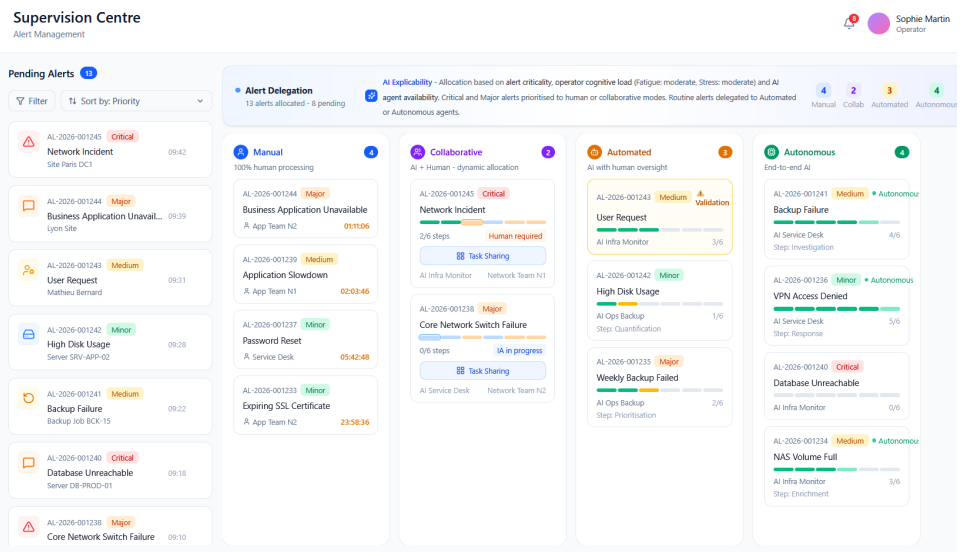


Figure 1: Supervision center dashboard with dynamic alerts delegation panel / Progress in the activity workflow of each alert (overview with Use dark green for completed tasks, light green for tasks in progress, gray for tasks that haven't been started and orange for steps that require operator action).

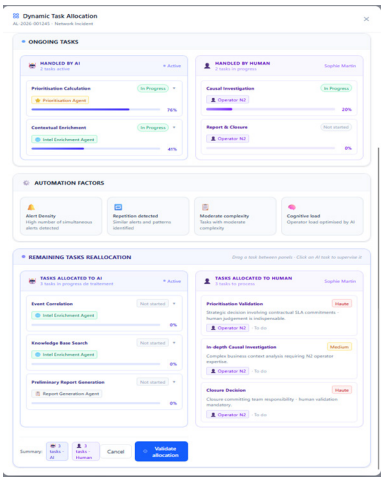


Figure 2: Dynamic task allocation panel in collaborative mode.

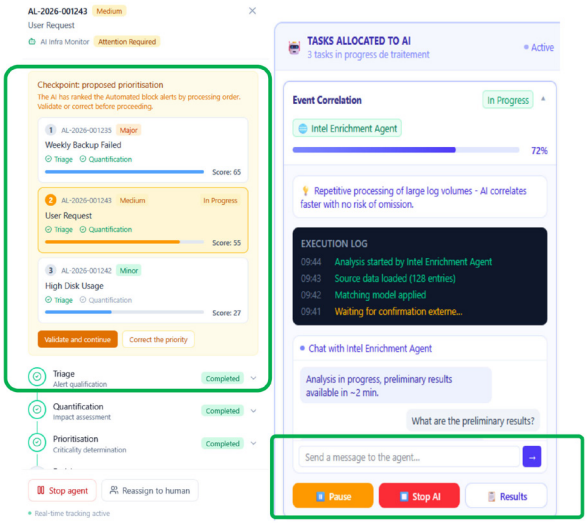


Figure 3: Human control illustration (Left : validate or correct suggestion, Stop agent in automated mode; Right : pause, Stop AI in collaborative mode).

Supporting Taskwork Situation Awareness (Understanding Task Execution)

Effective human–AI collaboration requires Taskwork Situation Awareness, which enables operators to monitor task execution, overall system progress, operational context, and expected performance. This awareness helps operators assess whether activities are progressing as expected, anticipate future states, and intervene when necessary. To support this dimension, guideline G2 (Make clear how well the system can do what it can do) was implemented.

This guideline motivated the integration of workflow monitoring elements (Figure 4), which display the progression of alert management activities and distinguish completed, ongoing, and pending steps. Operators can also access detailed views of individual alerts to monitor their execution status.

In collaborative mode, the interface additionally displays ongoing and remaining tasks, their allocation between human operators and AI agents, and their current status.

By providing continuous visibility into workflow progression and task execution, the interface helps operators evaluate automation performance, anticipate future system behavior, and make informed decisions. This supports Taskwork Situation Awareness and enables timely human intervention when task conditions or system performance require adaptation.

Supporting Agent Situation Awareness (Understanding the AI Agent)

Effective human–AI collaboration requires Agent Situation Awareness, which enables operators to understand the AI agent's capabilities, operational state, limitations, and decision-making process. This understanding helps operators anticipate AI behavior, interpret its actions, and determine when supervision or intervention is necessary. To support this dimension, guidelines G1 (Make clear what the system can do) and G11 (Make clear why the system did what it did) were implemented.

Guideline G1 motivated the design of the AI Agent State and Behavior Panel (Figure 4), which displays the agent's current state, active behavior, intentions, and limitations. In autonomous and automated modes, it also shows the progression of operational steps and subtasks, helping operators understand the agent's activities and capabilities.

Guideline G11 led to the design of Explanation Zones, which make the AI's decision-making process more transparent by presenting the contextual factors underlying its decisions, such as those involved in a triage recommendation (Figure 4). This helps operators interpret AI outputs and assess their appropriateness.

Together, these components strengthen Agent Situation Awareness by clarifying both what the AI can do and why it behaves in a particular way. This supports effective supervision, appropriate trust calibration, and informed human–AI collaboration by helping operators adjust their reliance on AI according to its observed behavior, confidence, and contextual suitability.

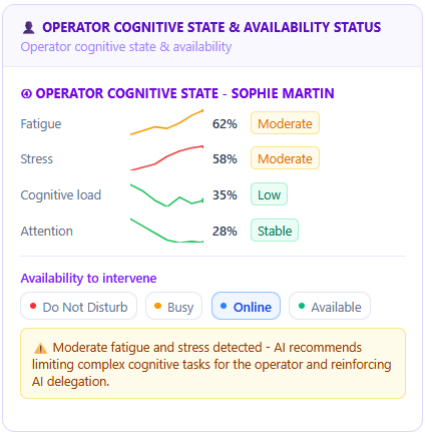


Figure 4: Operator state.

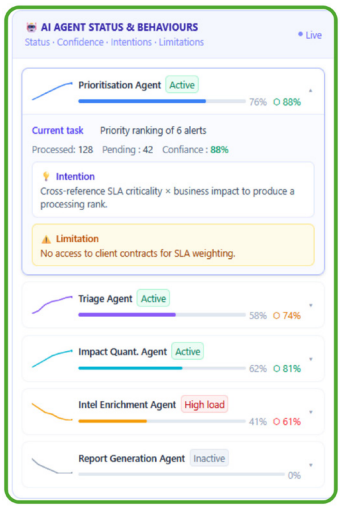


Figure 5: Agent situation awareness.

Supporting Shared Situation Awareness (Human-AI Coordination)

Effective human–AI collaboration requires Shared Situation Awareness, which enables human operators and AI teammates to maintain a common understanding of task objectives, responsibilities, task allocation, and ongoing activities. This shared understanding supports coordination, anticipation of each other's actions, and adaptation to changing task conditions. To support this dimension, guideline G4 (Show contextually relevant information) was implemented.

This guideline led to the design of the Task Delegation Panel (Figure 2), which displays task distribution between the operator and AI agents, including subtasks, completion status, timing, and execution progress. The Remaining Tasks Reallocation panel also proposes allocation strategies for unfinished tasks, which operators can review and modify by transferring tasks between themselves and AI agents through drag-and-drop interactions.

By making task ownership, delegation decisions, execution progress, and future allocation explicit, the interface strengthens shared awareness, improves coordination and transparency, and reduces ambiguity regarding responsibilities. In line with Team Situation Awareness requirements (Endsley, 2023), this approach helps operators understand the overall collaborative process and make informed delegation and intervention decisions.

Supporting Self Situation Awareness and Trust Calibration

In addition to monitoring AI teammates, operators must maintain **Self Situation Awareness** to assess their cognitive and operational state, ensuring their ability to supervise AI agents, make decisions, and intervene when necessary. This is essential for calibrating trust in automation and avoiding both over-reliance and under-reliance on AI systems.

To support this dimension, guidelines G2, G4, and G17 were implemented through the **Operator Cognitive State and Availability Status panel** (Figure 8). This interface displays key indicators such as fatigue, stress,

cognitive workload, attention level, and operational availability, alongside context-aware AI recommendations.

When high fatigue or workload is detected, the system suggests adapting task allocation, for example by reducing cognitively demanding tasks for the operator or increasing AI delegation. By making the operator's cognitive state and adaptation mechanisms explicit, the interface supports informed supervision and delegation decisions, improves trust calibration, and maintains awareness of human responsibility in human-AI collaboration.

CONCLUSION AND PERSPECTIVES

The proposed contribution consists of an operational methodology for designing interfaces that support effective collaboration between humans and AI. It is based on the hypothesis that effective collaboration depends not only on the system's intelligence but also on the user's ability to maintain accurate "situational awareness" regarding task execution, agent behavior, uncertainty, and coordination dynamics.

This methodology translates theoretical concepts from the literature on human-AI collaboration and situational awareness into concrete design requirements and interface mechanisms. It organizes these requirements around four complementary cognitive objectives: understanding the task progression and its environment, understanding the AI agent, and maintaining a shared situation awareness of collaborative goals and responsibilities.

By establishing an explicit link between theoretical concepts and guidelines for human-AI interaction, on the one hand, and the requirements of situational awareness and the corresponding interface mechanisms, on the other, this methodology proposes a systematic approach to the design of transparent and coordination-oriented human-AI interfaces. It thus offers a transferable design framework that can be adapted to various collaborative contexts, with the aim of promoting transparency, effective coordination, the establishment of trust, and adaptive collaboration between human operators and intelligent agents.

As an initial proposal, however, this methodology still needs to be refined and empirically validated. Future work will therefore adopt an iterative, user-centered design approach, involving SOC analysts from our project partners. Their expertise and feedback will be used to progressively refine both the design requirements derived from the methodology and the corresponding interface mechanisms, while ensuring their validity in realistic operational contexts. In the long term, this iterative process will lead to the implementation of user studies aimed at empirically evaluating the methodology and the resulting interface design, particularly with regard to their ability.

ACKNOWLEDGMENT

This work has been supported by the French government under the "France 2030" program, as part of the SystemX Technological Research Institute R&D Program CYBELIA. The CYBELIA Program is co-funded by Airbus Protect and Réseau de Transport d'Electricite (RTE).

REFERENCES

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Horvitz, E. (2019, May). Guidelines for human-AI interaction. In Proceedings of the 2019 chi conference on human factors in computing systems (pp. 1–13).

Card, S. K., Mackinlay, J., & Shneiderman, B. (Eds.). (1999). Readings in information visualization: using vision to think. Morgan Kaufmann.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv prep*rint arXiv:1702.08608*.

Ehsan, U., Liao, Q. V., Muller, M., Riedl, M. O., & Weisz, J. D. (2021, May). Expanding explainability: Towards social transparency in AI systems. In Proceedings of the 2021 CHI conference on human factors in computing systems (pp. 1–19).

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. Human factors, 37(1), 65–84.

Endsley, M. R. (2017). Toward a theory of situation awareness in dynamic systems. In Situational awareness (pp. 9–42). Routledge.

Endsley, M. R. (2023). Supporting Human-AI Teams: Transparency, explainability, and situation awareness. Computers in Human Behavior, 140, 107574.

Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2018). Metrics for explainable AI: Challenges and prospects. arXiv preprint arXiv:1812.04608.

Hong, S., & Park, W. (2025). Developing user-centered system design guidelines for explainable AI: A systematic literature review. Artificial Intelligence Review, 58(12), 1–50.

Horvitz, E. (1999, May). Principles of mixed-initiative user interfaces. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 159–166).

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial intelligence, 267, 1–38.

Mucha, H., Robert, S., Breitschwerdt, R., & Fellmann, M. (2021, May). Interfaces for explanations in human-AI interaction: proposing a design evaluation approach. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1–6).

Mueller, S. T., Veinott, E. S., Hoffman, R. R., Klein, G., Alam, L., Mamun, T., & Clancey, W. J. (2021). Principles of explanation in human-AI systems. arXiv preprint arXiv:2102.04972.

Munzner, T. (2025, August). Visualization analysis and design. In Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Courses (pp. 1–2)

Nguyen, T., Canossa, A., & Zhu, J. (2024). How human-centered explainable ai interface are designed and evaluated: A systematic survey. arXiv preprint arXiv:2403.14496.

Norman, D. (2013). The design of everyday things: Revised and expanded edition. Basic books.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. Journal of cognitive engineering and decision making, 2(2), 140–160.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive science, 12(2), 257–285.

Wickens C. D. HollandsJ. G. Banbury S. Parasuraman R. (2015). Engineering Psychology and Human Performance. Psychology Press. Available online at: https://www.taylorfrancis.com/books/mono/10.4324/9781315665177/engineering-psychology-human-performance-christopher-wickens-justin-hollands-simon-banbury-raja-parasuraman