

From transparent to translucent decisions in Human-AI teams

Maëlle François^{1,3}, Elodie Bouzekri², Gilles Coppin¹ and Thierry Coye de Brunélis³

¹IMT Atlantique, Lab-STICC, UMR CNRS 6285, France

²Univ Brest, Lab-STICC, CNRS, France

³Thales Defense Mission Systems, Valbonne, France

Abstract

As sensors evolve, they enable the sensing of increasing amounts of data to support operators of critical systems, such as Sound Navigation and Ranging (SONAR) operators. AI-based classification systems are investigated as new team members to support decision-making based on large amounts of data. Transparency of Human-AI teams enables understanding of behavior and accuracy of team members by displaying more or less information. Transparency models define different levels and different types of information to convey within the interaction within the team. Focusing on another aspect to achieve the same goal of team decision performance, we propose the translucency model, which proposes enabling in-depth exploration of information that leads to the team's decision and their relationships.

Keywords

Translucency, Transparency, Decision-making, Classification

1. Introduction

In critical complex systems, such as those used in the military domain, the need for user assistance in performing specific tasks has become essential. Advances in sensor quality and increased deployment have significantly expanded the amount of data collected from the environment, resulting in a larger volume of information available to the users. Even though data pre-processing serves as an initial filter, the amount of displayed data remains substantial. Therefore, it is necessary to support users in processing information effectively. In this paper, we focus on user AI-based assistance for a classification task. This AI-based assistance refers to a classification system for decision support which aims to enhance the decision-making process of users by providing classification recommendations that help them make the most satisfying decision within a specific context [1]. More specifically, from a technical standpoint, we define a classification system as an algorithmic or rule-based system designed to assign input data to predefined categories or classes based on specific features or patterns.

The user may distrust and disuse the system (i.e., underestimating the system's capacities) or over-trust and misuse it (i.e., overestimating the system's capacities) [2], both of which impair the effectiveness of decision-making support. The user may reject a correct recommendation or accept an incorrect one. To avoid these pitfalls, the system must provide the user with an appropriate level of trust. One factor that supports informed decision-making is 'transparency.' Transparency involves conveying information about the behavior and accuracy of both the individuals and the team formed by the user and the decision-support system [3, 4]. However, we argue that a transparency mechanism that conveys or withholds certain types and amounts of information to support different levels of transparency is limited.

We propose to explore a new approach called 'translucency' that aims to support individuals' understanding of information with contextualization elements such as decision-making context, teamwork, history, system capacities, or accuracy. Transparency aims to present this information as it is, without the possibility of accessing alternatives or actual relationships between different information, while remaining sober in the information display. In the translucency model, we propose enabling in-depth exploration of alternatives and their relationships with other information.

AutomationXP25: Hybrid Automation Experiences, April 27, 2025, Yokohama, Japan. In conjunction with ACM CHI'25.



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In the remainder of the paper, we present the translucency model, building on the limitations of existing transparency models. This model is illustrated using a future decision-support classification system for Sound Navigation and Ranging (SONAR) operations.

2. Transparent classification systems for decision support

The term ‘transparency’ is a metaphor borrowed from the field of physics, referring to the property of a material of “being easy to see through” (Cambridge Dictionary) [5]. While this physical definition of the concept of transparency is widely accepted in this field, it appears to be more difficult to grasp in the domain of human-machine interaction, specifically when characterizing an AI-based decision support system. Indeed, in this field, transparency remains a debated concept, with no clear consensus on its definition or design. Transparency is a characteristic of the communication between an AI agent and a human agent. Models of transparency describe how an AI agent should cooperate with a human, and especially define what information should be conveyed. Additionally, some research has explored how this information can be represented and delivered through the interface. The following section focuses on what it means for a system to be transparent through three perspectives. Haresamudram and al. [6] formerly proposed three types of transparency, which are “algorithmic transparency”, “interaction transparency,” and “social transparency,” that work together to make AI-based systems transparent.

Haresamudram and al. [6] formerly proposed three types of transparency, which are “algorithmic transparency”, “interaction transparency,” and “social transparency,” that work together to make AI-based systems transparent. They define “algorithmic transparency” as the ability of machine learning algorithms to offer information about how they interact with the data they process, and to provide explanations for the decisions they make, even when operating on data volumes and through mechanisms that are beyond human capacity to manage or fully understand. Thus, this type of transparency is implemented in the field of Explainable AI (XAI). More specifically, with regard to XAI, Richard et al. [7] propose that classification systems are deemed transparent when they are understandable (i.e., conveyed information and concepts are part of users’ knowledge), interpretable (i.e., user can make sense of the results and of the underlying process), traceable (i.e., absence of a stochastic process both in the learning system and in the classifier), and revisable (i.e., the user can provide feedback to the classification system, which is taken into account by the system to improve the results). Interaction transparency refers to the mutual exchange of information between the AI-based system and the user. The example provided by Haresamudram et al. [6] involves embodied systems, such as smartwatches equipped with sensors that collect physiological data from users and provide them with recommendations. This form of transparency aligns with the definition of bidirectional transparency proposed by Chen et al. [3], and Lyon et al. [4]. For them, the agent (in our case, the AI-based classification system) must provide what they refer to as ‘transparency information’—including the agent’s intent, reasoning, future plans, uncertainties, as well as the humans’ intents, constraints, objectives, and their shared tasks and interactions- and where the system is able to collect data on the user’s cognitive state, including stress and fatigue levels. Finally, social transparency refers to the ethical obligation regarding data privacy.

It should be noted that the terms transparency and explainability are used in many subfields of artificial intelligence, robotics, and autonomous systems with slightly different meanings. As Patidar et al. [8] suggest, we contend that an explainable AI system significantly enhances the transparency of this framework.

In this paper, we focus on the transparency of the system for user understanding and interpretability. Therefore, in the following sections, the term ‘transparency’ refers to a system characteristic that allows users to access data (i.e., perceiving the information on the display or being able to perceive it) that explains the process supporting its decisions and is represented in a manner easily understandable by humans [9]. In contrast, the term explainability denotes the extent to which the information made transparently available to users can be readily interpreted by them. It requires describing the causality behind a system’s decision [9].

If a consensus on the definition of transparency has yet to be reached in the literature, the same holds

for its design, particularly in terms of how it should be concretely applied to classification systems. It remains unclear what information should be provided and how it should be presented. Most researchers have focused on the representation of a specific type of information from the classification systems, which are explanations regarding XAI. As mentioned earlier, XAI aims to address the “black box” problem (i.e., “the challenge of understanding and interpreting how complex machine learning models, particularly deep learning algorithms, make decisions” [10]) in complex AI models by making their decision-making processes understandable to humans. Various explainability methods, such as post-hoc explainers, have been developed, offering different approaches for visualizing explanations. Nevertheless, there is limited empirical evidence on whether these “interpretable models” and explanation representation are actually understandable and usable by users [11, 12, 13].

Most empirical studies [14, 15] focus on designing static explanations, typically by comparing user interfaces that present varying ‘levels of explanation.’ These levels refer to the granularity of the explanations. It refers to the level of information detail provided to users to justify or clarify an AI system’s decision-making process. Granularity can range from high-level (coarse-grained) explanations, which offer general insights into model behavior, to low-level (fine-grained) explanations, which provide detailed, instance-specific justifications, such as feature contributions or intermediate processing steps. These studies compare interfaces that provide varying levels of explanations about the decisions made by the classification system, focusing on which level of granularity has the greatest impact on factors such as performance, trust, and workload. It is also worth noting that some studies indicate that performance declines when transparency is high (i.e., the amount of displayed information exceeds a certain threshold). However, researchers have yet to provide a clear explanation for this effect—specifically, whether the drop is due to cognitive overload or a lack of training, both of which may result from poor design representation [16]. In this static approach, a fixed level of explanation granularity is imposed on the user without allowing them to explore further. This approach assumes that there is a single level that best aligns with the user’s needs. This approach is unsatisfactory because it overlooks the individual variability of our mental models used to make decisions. As some researchers argue, system transparency must consider user preferences and individual differences. Vered et al. [17] propose a model called the demand-driven transparency model, which aligns with this viewpoint. They argue that users should be able to choose the level of transparency based on their needs. On their interface, the level of transparency is either predetermined and imposed on the user or adjustable through a button click by the user.

To go beyond these binary representations of information, we propose a new approach for exploring decisions made by the classification system through a depth-based layout, which we refer to as the translucent approach.

3. Translucent classification systems for decision support

3.1. A new approach to represent decisions: Translucency

We mentioned that there is a need to design a classification system qualified as transparent to enhance the performance of the human-classification system team. Achieving this transparency requires displaying additional information to the user. However, current design solutions for information representation still do not fully address or resolve the problem of information overload, where the volume of incoming data exceeds the operator’s capacity to process it effectively. To make a decision, the user has to deal with diverse information from various sources, particularly from the classification system behavior (i.e., a final decision and explanations related to it), information from the other teammates, information derived from the user’s operation, and information related to the tactical and environmental context. To mitigate this information overload, it is essential to implement strategies that streamline information presentation and enhance decision-making processes. Although not directly related to the field of decision-making, Harrison et al. [18] explored an approach to information display by organizing information into depth-based layouts. They developed a concept of layered displays which aimed to “better support both focusing attention on a single interface object (without distraction from other

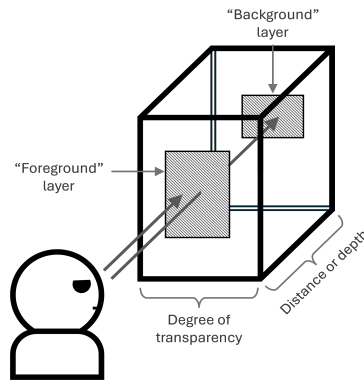


Figure 1: Illustration of Buxton et al.'s concept of layered-displays from [18]

objects) and dividing or time sharing attention between multiple objects (to preserve context or global awareness)". The user interface was composed with what they called 'semi-transparent' windows, menus, dialogue boxes, screens, or other objects where semi-transparency fits into a design space of "layered" interface objects (see Figure 1) [18]. Our concept of translucency aligns with this semi-transparent approach. Our concept seeks to go beyond the mere juxtaposition of graphical objects by treating objects as information and incorporating semantic criteria to establish relationships between them.

We consider translucency as a new approach to represent information to support decisions. The concept of translucency in decision representation can take several forms. It is important to distinguish between representations aimed at assisting users during the decision-making process and those that communicate a decision after it has been made.

The term 'translucency' usually refers to the property of being 'translucent,' which is defined as "allowing some light to pass through" (Cambridge Dictionary) [19]. In the field of optics, translucency is defined as "the property of a specimen by which it transmits light diffusely without allowing a clear view of objects through it" [20]. A translucent material is characterized by the scattering of light within it, alongside the effects of reflection and refraction. This scattering arises from the presence of particles that interact with the light, causing it to diffuse in various directions. Due to their inherent properties of absorption, refraction, and scattering, these particles are capable of diffusing incident light, resulting in a softened or blurred transmission, which is the main defining feature of translucency [21, 22]. Different levels of translucency can be distinguished. The level of translucency depends on the particle density of the material. The higher the material's density, the lower the level of translucency; conversely, the lower the density, the higher the level of translucency [21]. Depending on the degree of translucency, visual perception will vary. When looking at an object through a translucent filter with a low degree of translucency, one can still perceive its colors and contours, but details become difficult to discern. This makes it hard for the observer to recognize the object, though its intrinsic shape remains identifiable. As a result, the viewer is likely to search for details. In contrast, viewing an object through a highly translucent filter causes colors to appear less saturated and contours more blurred. This makes it nearly impossible for the observer to recognize the object. Consequently, attention is drawn to the sharp part of the image, which simply serves to guide the viewer toward the most relevant information [23].

The blur effect has been applied to information representation in the field of human-machine interfaces as a depth and selective cue, helping to guide the user's attention when large amounts of information are displayed [24]. Experimental studies show that, when combined with other depth cues such as transparency and contrast cues like color [25], blur becomes an effective means of directing the user's focus.

We consider a decision-making scenario where a human agent must make a decision and compare it with those of their teammates, which include both AI agents and human agents within the framework of Human-Autonomy Teaming. To converge toward a shared decision, the human agent needs access to sufficient information supporting each option. In their research, Guarino et al. [26] define a list of

information that supports the interpretation of a decision, which they refer to as meta-information. This supporting information provides the human with contextualization, allowing for a better interpretation of the decisions made by their teammates. They define meta-information as "qualifiers" that contextualize information and "therefore can critically influence how a decision-maker will process, understand, and act on that information." For example, uncertainty, reliability, or characteristics of the source of the decision are meta-information.

3.2. Use case: a translucent representation of decisions to assist a SONAR operator in a classification task

Our research is conducted within a military context. We focus on the classification task performed by SONAR operators. They must develop a clear understanding of their surrounding environment, which involves accurately identifying the class of each suspicious noise source. SONAR operators have distinct and complementary roles, with each operator having a unique perspective of the environment based on the sensors they use. In order to classify these noise sources, they need to extract and interpret relevant information through graphical displays (i.e., data extracted from sensors and processed through various methods) and sound analysis, then share this information along with the preliminary classification derived from their combination with the other operators. The classification proposed by the operators is discussed among them and refined if consensus is not achieved. In addition to communication among teammates, it is now essential to consider another form of communication: the collaboration between an operator and a classification system designed to assist in the classification task. This human-machine collaboration aims to enhance the operator's decision-making process. Therefore, it appears that an operator must manage a substantial amount of information from various sources and of different natures. Furthermore, this classification task involves making decisions made under time constraints, with uncertain and incomplete information, and where the risks and consequences of errors can be significant. In addition to these factors, stress, tiredness, and data overload make it challenging to perform classification and prioritize tasks effectively and accurately.

To illustrate our concept of translucency, we consider a simplified classification task in which a human agent, *OP1*, must classify an object, *O1*. *OP1* collaborates with two teammates—another human agent, *OP2*, and an artificial agent, *OP3*—to make the final classification decision. *OP3* refers to the classification system. The object, *O1*, can belong to one of three possible classes: 'friend', 'neutral' or 'suspect'. The interface used by *OP1* displays the decisions made by the teammates regarding *O1* through a translucent representation. We propose a preliminary design to illustrate the dissonance and consonance among these decisions (see Figure 2, Display view). We consider a translucent disk representing the decisions of *OP1*, *OP2*, and *OP3*, who can agree or disagree about the class of the object. In this configuration we decided to represent the class of the object through the color and the shape to refer to North Atlantic Treaty Organization (NATO) formalism, the source of the decision through the value of θ (i.e., $\theta \in [\pi/2; -\pi/2]$ for human agent decisions and $\theta \in [-\pi/2; \pi/2]$) and the decision dissonance or consonance through the value of r (i.e., the distance between the center of the disk and the decision).

In this configuration, we illustrate different cases of disagreement. For instance, the top-right figure shows the *OP1* decision as being consonant with the *OP2* decision and therefore represented as closer together (i.e., with a smaller r value). However the *OP1* decision and *OP3* decision are dissonant and therefore represented as more distant (i.e., with a larger r value). Here, we display the decisions from *OP2* and *OP3* with translucency to indicate to *OP1* the extent to which they support its own decision.

This preliminary design is a conceptual sketch that does not yet demonstrate the relationship between the decisions and their contextual information.

We assume that this model of translucency for information representation could subtly guide the operator's attention and decision-making process by presenting information in a translucent way. It will encourage the operator to engage with the underlying structure when necessary without imposing it upon them. Therefore, this model leverages the 'nudge' effect, potentially enhancing both the efficiency and effectiveness of decision-making. It signals the presence of an internal structure associated with the emerging information without overwhelming the operator with the content of this structure while

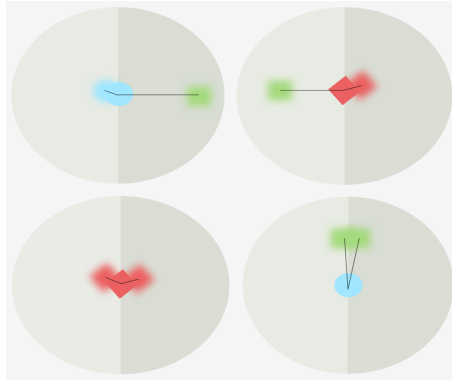


Figure 2: Representation of translucent decisions.

allowing access if the operator deems it necessary.

This model of translucency for information representation is still in its early stages and naturally requires further refinement and conceptual development. We aim to conduct an initial exploratory experiment to test our preliminary hypothesis, primarily comparing the effects of sharp and blurred stimuli on users and testing whether making information translucent can intrigue this ladder and create a nudge effect. We expect that translucent decision representation will impact factors such as cognitive overload, performance, or situation awareness.

4. Conclusion

The growing need for transparency in critical, complex systems to support informed decision-making has highlighted a significant challenge of information representation. The literature shows that this issue has garnered the interest of several researchers and is still under investigation. The existing transparency models and designs in classification systems have generally shown positive effects on trust and human-system classification performance [16, 27]. However, these results should be interpreted with caution, as comparisons are challenging due to inconsistencies in the definition of transparency and significant variability in protocol design. However, one notable observation is that the proposed approaches to designing transparency have certain limitations. Transparency often requires displaying more information, but studies have shown that when too much information is presented, users become overloaded, leading to a decline in the performance of the human-system team [16]. Interestingly, no clear explanations are provided to account for these findings. We believe that the real solution lies not in fixing a specific level of transparency but rather in allowing users the flexibility to explore and access the information they need to interpret decisions effectively. The preliminary approach of translucency we developed is designed to tackle this issue. We have embedded our concept of translucency into the field of decision support for operators of critical and complex systems, where they must manage vast amounts of information, interact with other operators, and engage with a classification system to make informed decisions. The goal is to introduce a new method of presenting information that allows for a detailed and broad exploration of decision options and how they relate to other information.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-4 to check grammar and spelling. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

- [1] H. A. Simon, *Models of man: social and rational; mathematical essays on rational human behavior in society setting*, New York: Wiley, 1957.
- [2] J. D. Lee, K. A. See, *Trust in Automation: Designing for Appropriate Reliance*, Human Factors (2004) 31.
- [3] J. Y. C. Chen, S. G. Lakhmani, K. Stowers, A. R. Selkowitz, J. L. Wright, M. Barnes, Situation awareness-based agent transparency and human-autonomy teaming effectiveness, *Theoretical Issues in Ergonomics Science* 19 (2018) 259–282. doi:10.1080/1463922X.2017.1315750.
- [4] J. B. Lyons, Being transparent about transparency: A model for human-robot interaction, in: 2013 AAAI spring symposium series, 2013.
- [5] Cambridge Dictionary, transparency, in: Cambridge University Press and Assessment, Cambridge Dictionary, 2025. URL: <https://dictionary.cambridge.org/us/dictionary/english/transparency>.
- [6] K. Haresamudram, S. Larsson, H. Heintz, Three levels of AI transparency, *Computer* (2023) 93–100.
- [7] A. Richard, B. Mayag, F. Talbot, A. Tsoukias, Y. Meinard, Transparency of classification systems for clinical decision support, 2020, pp. 99–113.
- [8] N. Patidar, S. Mishra, R. Jain, D. Prajapati, A. Solanki, R. Suthar, K. Patel, H. Patel, Transparency in AI Decision Making: A Survey of Explainable AI Methods and Applications, *Advances of Robotic Technology* (2024).
- [9] I. S. A. a. others, Ieee standard for transparency of autonomous systems, *IEEE Std* (2022) 7001–2021.
- [10] H.-F. Cheng, R. Wang, Z. Zhang, F. O'connell, T. Gray, F. M. Harper, H. Zhu, Explaining decision-making algorithms through UI: Strategies to help non-expert stakeholders, in: *Proceedings of the 2019 chi conference on human factors in computing systems*, 2019, pp. 1–12.
- [11] A. Abdul, J. Vermeulen, D. Wang, B. Y. Lim, M. Kankanalli, Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda, in: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, Association for Computing Machinery, New York, NY, USA, 2018, pp. 1–18. doi:10.1145/3173574.3174156.
- [12] F. Doshi-Velez, B. Kim, A roadmap for a rigorous science of interpretability, *arXiv preprint arXiv:1702.08608* (2017).
- [13] E. S. Vorm, A. D. Miller, Modeling user information needs to enable successful human-machine teams: Designing transparency for autonomous systems, in: *International Conference on Human-Computer Interaction*, 2020, pp. 445–465.
- [14] T. Helldin, U. Ohlander, G. Falkman, M. Riveiro, Transparency of Automated Combat Classification, in: D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics*, Springer International Publishing, Cham, 2014, pp. 22–33. doi:10.1007/978-3-319-07515-0_3.
- [15] C. Westin, C. Borst, B. Hilburn, Automation transparency and personalized decision support: Air traffic controller interaction with a resolution advisory system, *IFAC-PapersOnLine* 49 (2016) 201–206.
- [16] A. Bhaskara, M. Skinner, S. Loft, Agent Transparency: A Review of Current Theory and Evidence, *IEEE Transactions on Human-Machine Systems* 50 (2020) 215–224.
- [17] M. Vered, P. Howe, T. Miller, L. Sonenberg, E. Velloso, Demand-driven transparency for monitoring intelligent agents, *IEEE Transactions on Human-Machine Systems* 50 (2020) 264–275.
- [18] B. L. Harrison, H. Ishii, K. J. Vicente, W. A. Buxton, Transparent layered user interfaces: An evaluation of a display design to enhance focused and divided attention, in: *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1995, pp. 317–324.
- [19] Cambridge Dictionary, translucent, in: Cambridge University Press and Assessment, Cambridge Dictionary, 2025. URL: <https://dictionary.cambridge.org/us/dictionary/english/translucent>.
- [20] A. E284-17, Standard Terminology of Appearance, Standard, ASTM International, 2022. doi:<https://doi.org/10.1520/E0284-17>.
- [21] B. Xiao, B. Walter, T. Zickler, E. Adelson, K. Bala, Looking against the light: How perception of translucency depends on lightning direction, *Journal of Vision* 14 (2014). doi:<https://doi.org/10.1167/14.3.17>.

- [22] D. Gigilashvili, J.-B. Thomas, J. Y. Hardeberg, M. Pedersen, Translucency perception : A review, *Journal of Vision* 21 (2021). doi:<https://doi.org/10.1167/jov.21.8.4>.
- [23] P. Martin, Le flou est-il quantifiable?: étude du flou-net de profondeur en photographie et en cinéma, Ph.D. thesis, Saint-Etienne, 2001.
- [24] G. Colby, L. Scholl, Transparency and blur as selective cues for complex visual information, in: *Image Handling and Reproduction Systems Integration*, volume 1460, SPIE, 1991, pp. 114–125.
- [25] R. Kosara, S. Miksch, H. Hauser, J. Schrammel, V. Giller, M. Tscheligi, Useful properties of semantic depth of field for better f+ c visualization, in: *ACM International Conference Proceeding Series*, volume 22, 2002, pp. 205–210.
- [26] S. L. Guarino, J. D. Pfautz, Z. Cox, E. Roth, Modeling human reasoning about meta-information, *International Journal of Approximate Reasoning* 50 (2009) 437–449.
- [27] F. Rajabiyazdi, G. A. Jamieson, A review of transparency (seeing-into) models, in: *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020, pp. 302–308.