Educating for Adaptive AI Awareness: Enabling Users to Recognize and Resist Algorithmic Influence

Alexandru Mateescu[1]

[1]*Université Paris 1 Panthéon-Sorbonne, École Doctorale 280, Paris, France*

Abstract

Adaptive AI systems—such as recommender platforms and personalized learning environments—continuously adjust their outputs in response to user behavior. This adaptivity enhances personalization but also shapes user beliefs, preferences, and actions through opaque feedback loops. While current approaches to trustworthy AI stress transparency, explainability, or fairness, they often treat users as passive recipients rather than active participants in co-adaptive processes.

This paper proposes that awareness in such systems can be understood as a form of epistemic agency: the capacity to recognize, reflect on, and respond to algorithmic influence. To support this agency, we introduce the notion of *educational requirements*—user-facing design principles that embed epistemic support into system interfaces.

Rather than prescribing a fixed framework, we outline three orientations—intelligibility, deliberation, and control—and show how they can foster critical reflection in recommender systems. We also discuss implications for AI education, emphasizing the importance of designing interactions that support not only efficiency, but also user understanding and autonomy.

1. Introduction

Adaptive AI systems—such as recommender platforms and personalized learning environments—shape user experience by continuously adjusting outputs in response to individual behavior. While these systems are often praised for improving relevance or efficiency, they also exert influence through subtle feedback loops that are rarely visible to users. As algorithmic mediation becomes more pervasive, a growing concern has emerged: how can individuals meaningfully understand and respond to the effects of such adaptive systems?

Existing approaches to trustworthy AI focus on transparency, explainability, or fairness, often emphasizing system-level metrics or compliance standards. While valuable, these efforts tend to treat users as passive recipients of information. In contrast, this paper approaches AI awareness from the perspective of *epistemic agency*: the capacity of users to recognize, reflect on, and respond to the influence exerted by adaptive systems over time.

We introduce the notion of *educational requirements*—user-facing design principles aimed at supporting epistemic engagement within interaction. Rather than framing education as an external intervention, we suggest that awareness can emerge through interaction itself, provided that systems are designed to scaffold observation, interpretation, and correction [1, 2].

Our goal is not to prescribe a definitive framework, but to explore how adaptive AI systems might support user autonomy by embedding critical engagement directly into interface design. Using recommender systems as a case study, we outline how educational requirements could foster intelligibility, deliberation, and interactive control. We then examine broader implications for AI education and suggest that fostering epistemic resilience may be as essential as technical robustness in the development of trustworthy systems.

2nd Workshop on Education for Artificial Intelligence (edu4AI 2025, https://edu4ai.di.unito.it/), Co-located with ECAI 2025, the 28th European Conference on Artificial Intelligence which will take place on October 26, 2025 in Bologna, Italy
alexandru.mateescu@etu.univ-paris1.fr (A. Mateescu)
0009-0007-7084-8284 (A. Mateescu)

2. Related Work and Conceptual Background

A wide range of initiatives across computer science, education, and HCI seek to promote AI awareness and trust. These efforts typically follow three main directions: (1) fostering user understanding through AI literacy, (2) enhancing system transparency and explainability, and (3) managing the persuasive or behavioral effects of AI systems. Each contributes important insights to the problem of user awareness—but also reveals certain limitations that motivate the notion of educational requirements we develop in this paper.

Research on AI literacy has made significant progress in helping users grasp the principles behind machine learning systems, data usage, and algorithmic decision-making [3]. These interventions often take the form of curricular modules, explainer platforms, or interactive tools designed to demystify AI for non-experts. While essential, these approaches tend to frame awareness as a static form of conceptual knowledge, often delivered externally to the system. In contrast, adaptive AI systems—such as recommender platforms or personalized learning environments—affect users through ongoing interactions and feedback loops. This dynamic aspect can be difficult to address through conventional literacy methods alone.

Another key area of work concerns explainability and trustworthiness. Here, the goal is not necessarily to educate the user, but to make the system more interpretable. Methods range from model interpretability techniques to transparency dashboards and accountability frameworks [4, 5]. These developments have influenced both technical research and regulatory guidelines. However, they often rely on abstract explanations of model behavior, rather than concrete support for users navigating evolving algorithmic environments. Explanations may be difficult to interpret, poorly contextualized, or ignored altogether if they do not align with user goals or experiences.

A third strand addresses the persuasive and behavioral effects of adaptive systems. Recommender platforms in particular can steer attention, reinforce engagement loops, or shape belief formation over time—sometimes without users realizing it [6, 7]. This literature raises questions not only about manipulation and autonomy, but also about the erosion of epistemic agency: the capacity to assess and regulate one's own beliefs and information sources. While some proposals focus on defensive measures—such as friction mechanisms or interface redesigns—we argue that these issues also pose pedagogical challenges. Users need support not only to resist influence, but to recognize its dynamics and implications.

Finally, our approach is informed by the tradition of reflective interaction design, which aims to foster user awareness of habits, assumptions, and decision paths [8]. In this view, awareness is not merely a matter of information access, but of developing the ability to pause, reflect, and redirect attention. We position educational requirements within this line of work, as design principles that embed epistemic support into interaction. Rather than replacing existing strategies, they aim to complement them by fostering user capacities that are critical in co-adaptive environments.

3. Educational Requirements

In adaptive AI systems, awareness is not a fixed state but a process—something users cultivate over time, through repeated exposure, trial and error, and reflective engagement. Traditional design strategies aimed at fairness, robustness, or explainability often assume a static or transactional view of the user: someone who must understand a model, trust it, or audit its outcomes. But when system behavior evolves with user behavior, this view becomes insufficient. The user is not just interpreting outputs—they are shaping them, being shaped in return, and participating in a dynamic relationship. We propose that this relationship demands a different kind of support: educational requirements.

Educational requirements are user-facing design features that embed epistemic support directly into interaction with adaptive systems. They help users recognize patterns of influence, understand how the system is adapting to them, and develop strategies to steer or resist that adaptation when appropriate. Unlike formal transparency obligations or static explanations, educational requirements are contextual, situated, and oriented toward long-term understanding. They treat the user not as a passive recipient of information, but as a participant in the system's logic of personalization.

3.1. Defining educational requirements

We define educational requirements as system-level design principles that aim to cultivate user reflection, awareness, and epistemic agency in the context of adaptive AI. They are requirements in the engineering sense—criteria to guide design decisions—but their goal is pedagogical: to create interactional conditions in which learning becomes possible.

This approach draws on research in reflective design, critical pedagogy, and HCI [8]. It also aligns with recent proposals to integrate socio-technical perspectives into trustworthy AI frameworks. Yet educational requirements are distinct in that they do not seek to deliver content or teach concepts in isolation. Rather, they embed opportunities for reflection within the system's affordances themselves. For instance:

- **Feedback visualizations** can show users how their past behaviors have influenced recommendations, revealing behavioral loops and filter effects.
- **What-if tools** allow users to explore how different actions might shift system outputs, supporting counterfactual reasoning.
- **Diversity prompts** or **exposure meters** can signal when content homogeneity is increasing, nudging users to recalibrate.
- **Traceability mechanisms** let users inspect how specific content was selected or adapted, giving insight into personalization paths.

These are not explanations in the technical sense, but tools for epistemic exploration: ways for users to observe the system observing them.

3.2. Why adaptation matters

Adaptation adds a layer of complexity that most existing AI literacy or explainability tools do not address. A classifier returns the same output for the same input. A recommender system does not. Once a system begins adapting to behavior over time, static snapshots of its logic become insufficient. The user must learn to reason about trajectories, feedback, and co-evolution.

This is why we argue that adaptive systems create educational obligations. If a system can learn from users, it should also be designed so that users can learn from it. This mutual intelligibility is not simply a matter of ethics or usability—it is a precondition for meaningful agency.

3.3. From feature to requirement

Treating these supports as requirements—not optional features—has several implications. First, it brings user education into the scope of system design and evaluation. Educational supports become part of what it means for a system to be trustworthy—not as external add-ons, but as integrated affordances [5]. Second, it allows for accountability: requirements can be tested, iterated, and assessed over time. Finally, it offers a way to scale epistemic resilience. In contexts like recommender systems, where millions of users interact with personalized outputs daily, interface-level interventions may be more feasible and effective than external education campaigns.

In the next section, we demonstrate how these principles can be applied to recommender systems specifically, and how educational requirements might support users in recognizing and resisting influence across various stages of algorithmic personalization.

4. Application to Recommender Systems

Recommender systems provide a relevant domain for applying educational requirements. They are widely used, socially influential, and inherently adaptive: recommendations evolve with user behavior and shape attention, preferences, and beliefs. Yet the mechanisms behind this evolution often remain opaque. Most interfaces conceal the logic of personalization and offer limited opportunities for reflection or correction. We propose a functional model structured around three moments: observation, interpretation, and correction.

4.1. Observation: surfacing adaptive dynamics

The first step is to help users perceive that adaptation is occurring. Many assume that recommendations are either fixed or directly tied to explicit choices. Interfaces seldom make the learning process visible [9]. Educational requirements begin by treating adaptation as a learnable structure.

Interfaces can support this through features such as:

- Timeline views showing how content changed over time;
- Scroll histories or heatmaps revealing engagement patterns;
- Notifications such as "We're showing you more of X because you clicked on Y."

Such elements support *algorithmic legibility*: recognizing system behavior as historically shaped rather than static. They encourage users to ask not only "why this item?" but "why this trend?"

4.2. Interpretation: scaffolding epistemic reasoning

Observation alone is not enough if users lack resources to interpret system behavior. Educational requirements call for scaffolding—tools that help users relate observations to their own informational goals.

Examples include:

- Prompts comparing current recommendations with past activity;
- Counterfactual tools simulating alternative choices;
- Indicators clarifying the type of signal used (e.g., "Based on watch time").

These interventions encourage active inquiry and support users in reasoning about their position within the system—whether they are entering a filter bubble, for example, or being nudged toward certain views.

4.3. Correction: enabling agency and recalibration

The final moment is correction: allowing users to intervene when needed. This does not mean rejecting personalization altogether, but providing ways to recalibrate the system's inferences.

Examples include:

- Interfaces to inspect and edit inferred preferences;
- "Forget" or undo buttons for unwanted interactions;
- Controls to toggle between novelty, diversity, or relevance.

While some systems offer such options [10], they are often hidden or underused. Reframing them as epistemic supports emphasizes their role in helping users monitor and shape their own informational environments.

4.4. Implications for design

Embedding educational requirements in recommender systems repositions interface design as a form of epistemic support. It suggests that systems should be evaluated not only by accuracy or usability, but by how they foster user understanding and reflection.

This orientation complements ongoing work on fairness and explainability. Rather than treating personalization solely as a technical process, it highlights its cognitive and social dimensions—and invites pedagogical attention to how users learn to navigate it.

5. Implications for AI Education

Educational requirements invite us to reconsider how AI awareness can be supported—not only through external instruction, but through interaction itself. Traditional approaches to AI education often emphasize conceptual knowledge: how algorithms work, what risks they entail, or how to interpret their outputs. While foundational, such approaches may fall short when users face adaptive systems whose influence evolves over time and through feedback.

In this context, awareness cannot be limited to prior training or static explanations. It must be sustained, situated, and responsive. By embedding epistemic support mechanisms—such as traceability, counterfactuals, or recalibration tools—into user interfaces, designers can foster reflection within the very process of interaction. This aligns with constructivist learning theories, which emphasize inquiry, feedback, and learning-by-doing.

Moreover, integrating education into system design may help address known limitations of standalone literacy efforts, such as limited scalability or user engagement. As Verbeek has argued, "If ethics is about how to act and designers help to shape how technologies mediate action, designing should be considered a material form of doing ethics."[1] From this perspective, educational requirements do not replace pedagogical initiatives, but complement them—by making adaptive systems themselves part of the pedagogical landscape.

6. Conclusion

As adaptive AI systems become increasingly embedded in everyday life, awareness of their influence becomes essential—not only for informed use, but for preserving epistemic autonomy. This paper has argued that awareness in such contexts must go beyond surface-level understanding or static explanations. It must be cultivated through interaction, reflection, and agency. We proposed the concept of *educational requirements* as a response to this need: system-level design principles that embed support for user observation, interpretation, and correction within adaptive systems.

Rather than treating education as an external add-on, educational requirements integrate pedagogical aims into the system interface itself. This reframing aligns with recent shifts in AI ethics and human-centered design, but it adds a crucial dimension: education is not merely a means of compliance or literacy-building—it is a condition for meaningful interaction with adaptive AI.

By applying this framework to recommender systems, we have illustrated how personalization can be made legible and negotiable, allowing users to track influence and recalibrate engagement. This approach supports a deeper form of AI awareness: one that includes the ability to recognize when influence occurs and to resist it when necessary.

This paper builds on prior work [1, 2] by refining the concept of educational requirements and clarifying its implications for adaptive system design. Embedding educational support into interaction is not just an opportunity—it is a condition for cultivating awareness, autonomy, and responsibility in an algorithmically mediated world.

7. Declaration on Generative AI

During the preparation of this work, the author used ChatGPT, Grammarly in order to: Grammar and spelling check, Paraphrase and reword. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] A. Mateescu, Educational requirements for trustworthy recommender systems, in: Proceedings of the IEEE RE-TRAI Workshop 2025, 2025. doi:10.1109/RETRAI59676.2025.00014.

[1] Verbeek, *Moralizing Technology* (University of Chicago Press, 2011), p. 91.

[2] A. Mateescu, Supporting epistemic agency in personalized learning environments: Educational requirements for adaptive systems, in: Book of Abstracts - HELMeTO 2025, Studium s.r.l., Naples, Italy, 2025. URL: https://www.helmeto.it/wp-content/uploads/2025/09/BOA_HELMETO_2025_con_ISBN.pdf, doctoral Consortium Abstract. ISBN verified via external source: 978-88-99978-68-6.

[3] D. Long, B. Magerko, What is ai literacy?competencies and design considerations, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (2020) 1–16.

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

[5] N. Diakopoulos, Accountability in algorithmic decision making, Communications of the ACM 59 (2016) 56–62. doi:10.1145/2844110.

[6] A. J. B. Chaney, B. M. Stewart, B. E. Engelhardt, How algorithmic confounding in recommendation systems increases homogeneity and decreases utility, in: Proceedings of the 12th ACM Conference on Recommender Systems, ACM, 2018, pp. 224–232.

[7] P.-P. Verbeek, Moralizing Technology: Understanding and Designing the Morality of Things, University of Chicago Press, 2011.

[8] P. Sengers, K. Boehner, S. David, J. Kaye, Reflective design, in: Proceedings of the 4th Decennial Conference on Critical Computing, ACM, 2005, pp. 49–58.

[9] E. Rader, R. Gray, Understanding user beliefs about algorithmic curation in the facebook news feed, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI), ACM, 2015, pp. 173–182.

[10] P. Kouki, J. Schaffer, J. Pujara, J. O'Donovan, L. Getoor, Personalized explanations for hybrid recommender systems, in: Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI), ACM, 2019, pp. 379–390. doi:10.1145/3301275.3302306.