Explained, yet misunderstood: How AI Literacy shapes HR Managers' interpretation of User Interfaces in Recruiting Recommender Systems

Yannick Kalff[1,*,†], Katharina Simbeck[1,†]

[1]HTW Berlin University of Applied Sciences, Treskowallee 8, 10318 Berlin, Germany

Abstract

AI-based recommender systems increasingly influence recruitment decisions. Thus, transparency and responsible adoption in Human Resource Management (HRM) are critical. This study examines how HR managers' AI literacy influences their subjective perception and objective understanding of explainable AI (XAI) elements in recruiting recommender dashboards. In an online experiment, 410 German-based HR managers compared baseline dashboards to versions enriched with three XAI styles: important features, counterfactuals, and model criteria. Our results show that the dashboards used in practice do not explain AI results and even keep AI elements opaque. However, while adding XAI features improves subjective perceptions of helpfulness and trust among users with moderate or high AI literacy, it does not increase their objective understanding. It may even reduce accurate understanding, especially with complex explanations. Only overlays of important features significantly aided the interpretations of high-literacy users. Our findings highlight that the benefits of XAI in recruitment depend on users' AI literacy, emphasizing the need for tailored explanation strategies and targeted literacy training in HRM to ensure fair, transparent, and effective adoption of AI.

Keywords

AI Literacy, Explainable AI, Recommender Systems, Human Resource Management, Recruitment, HR Analytics, People Analytics

1. Introduction

Artificial intelligence (AI)-based recommender systems have become widespread in recruitment [1, 2]. Recommender systems are software applications that use artificial intelligence techniques to analyze data and provide specific suggestions or predictions to users. AI-based systems typically assist in discovering promising talents for development, identifying the most suitable candidates for a job opening, or assigning the right employees to projects based on their skill sets. In human resource management (HRM), these tools promise to accelerate processes, reduce human bias, and ground decisions in objective data [3, 4]. Current trends, such as HR Analytics and People Analytics, integrate AI to offer a broader promise of analytical rigor, predictive opportunities, and prescriptive recommendations for informed decision-making and actions [5], alongside technological modes of control [6]. In recruiting, AI recommender systems directly influence decisions about individuals, making them the subject of regulations, such as the EU AI Act [7] and the GDPR [8]. Moreover, HR systems have faced severe criticism for fairness issues and biased recommendations [9, 10, 11, 12].

To mitigate potential biases and, equally important, to make optimal decisions, HR managers must understand the underlying data models and the mechanisms by which individual recommendations are generated. Explainable AI (XAI) techniques aim to make "black-box" models interpretable when they are opaque—due to complexity or proprietary constraints [13, 14].

XAI methods offer interpretable, context-specific explanations for model decisions [15, 16]. In recruitment, these explanations can clarify why a candidate's application is ranked highly or why specific competencies are flagged during the CV parsing process. This transparency is essential for HR managers who often have non-technical backgrounds and are responsible for legally and ethically sound decisions that comply with anti-discrimination laws. At the same time, from a human resources management perspective, their decisions must be economically sensible and strategically appropriate for the company. A lack of transparency in AI elements or data, combined with unrecognized distortions, can lead users to incorrect conclusions. The issue is amplified by providers and developers, as transparency and explanations of AI interfaces remain the exception in practice. Lacking transparency often seems to be a deliberate UI design decision (three exemplary dashboards can be found in the appendix Figure 4–6).

However, attaching explanation widgets to a recruitment dashboard does not guarantee impact. For XAI to be effective, HR managers must decode and critically evaluate the provided information [17]. We contend that AI literacy—a combination of knowledge, skills, and attitudes that enables individuals to understand and assess AI systems [18, 19]—directly affects both the subjective and objective effectiveness of XAI. Subjectively, AI literacy influences how helpful, trustworthy, and accessible explanations appear. Objectively, it affects accurate factual understanding that HR managers present when they interpret and act on the information provided by AI dashboards.

To investigate this effect, we conducted an experiment with 410 German-based HR managers, who compared a baseline AI dashboard with versions enriched by three explanation styles: important features (a simplified feature importance approach), counterfactual explanations, and global model criteria summaries [20, 21]. Drawing on a genuine recruitment ranking tool that exists on the market, we measured participants' perceived trust, usability, and assessment quality for each existing dashboard variant. Further, the participants assessed statements on the dashboards with correct or false answer possibilities. We address two guiding research questions:

RQ1 How do HR managers' subjective perceptions of a re-

RecSys in HR'25: The 5th Workshop on Recommender Systems for Human Resources, in conjunction with the 19th ACM Conference on Recommender Systems, September 22–26, 2025, Prague, Czech Republic.

*Corresponding author.

†These authors contributed equally.

yannick.kalff@htw-berlin.de (Y. Kalff);

katharina.simbeck@htw-berlin.de (K. Simbeck)

https://yannickkalff.de (Y. Kalff); https://iug.htw-berlin.de (K. Simbeck)

0000-0003-1595-175X (Y. Kalff); 0000-0001-6792-461X (K. Simbeck)

cruiting recommender system change when adding explainable AI elements, and does this effect differ across different levels of AI literacy?

RQ2 How does HR managers' objective understanding of a recruiting recommender system change when adding explainable AI elements, and does this effect differ across different levels of AI literacy?

Our results show that higher AI literacy is associated with greater perceived usefulness and transparency of XAI-enhanced dashboards. Paradoxically, higher AI literacy also corresponds to a lower objective understanding of the user interfaces. These findings suggest that the benefits of XAI depend critically on users' AI literacy levels, though self-assessed AI literacy may be prone to overconfidence.

The article is structured as follows: First, we review the literature on AI literacy and XAI, with a focus on their intersection in recruitment contexts (2). Next, we present our methodological approach (3) and empirical findings (4). We discuss the theoretical and practical implications for responsibly embedding explainable recommender systems in HRM that arise from AI literacy and its impact on subjective perception and objective understanding (5). We conclude with an outlook on future research topics that can be derived from our findings (6).

2. State of Research

The scholarly discourses on AI literacy and explainable AI (XAI) have so far evolved independently (with few exceptions [22]). AI literacy research emphasizes the competencies required to comprehend, evaluate, and interact with AI-enabled tools [23]. Several authors have developed scales to assess individuals' abilities to recognize, understand, apply, and critically or ethically evaluate AI systems [18, 19, 24]. Empirical studies underscore the need for contextualized training that embeds domain-relevant examples and ethical deliberation, arguing that generic digital skills programs fall short of preparing professionals for AI-mediated work [25, 26]. In HRM, such contextualization is particularly vital, as recruitment decisions carry significant strategic, legal, or ethical weight with impact on companies' success, diversity, equity, and the organization's reputation. A high level of AI literacy would include a foundational understanding of the technical principles underlying AI systems—such as training procedures, or the critical role of data quality—and the ability to use AI tools effectively in appropriate contexts. Proficiency in AI literacy would further indicate awareness of AI's limitations and boundaries, including ethical considerations and potential grey areas, and the necessity of human oversight. Moreover, high levels of AI literacy involve the ability to recognize AI-driven processes, critically assess the outputs generated by such systems, and accurately identify their capabilities and limitations.

XAI research, by contrast, focuses on designing algorithms, systems, and user interfaces that render AI decisions and recommendations transparent and understandable [27, 28]. This approach treats users as accountable agents who must comprehend, evaluate, and, if necessary, correct AI outputs [29]. Initially, XAI addressed the needs of ML engineers and developers seeking to understand and debug complex AI models [14]. Recent developments in XAI have extended the audience for explanations and established that explanations need to account for users' roles, backgrounds, and prior technical knowledge [30]. Researchers distinguish between global explanations, which clarify an entire model's logic, and local explanations, which justify individual decisions [20]. Adequate XAI should also draw on domain-specific knowledge—for example, highlighting key CV attributes or motivational letter elements that influenced an AI recommendation [31].

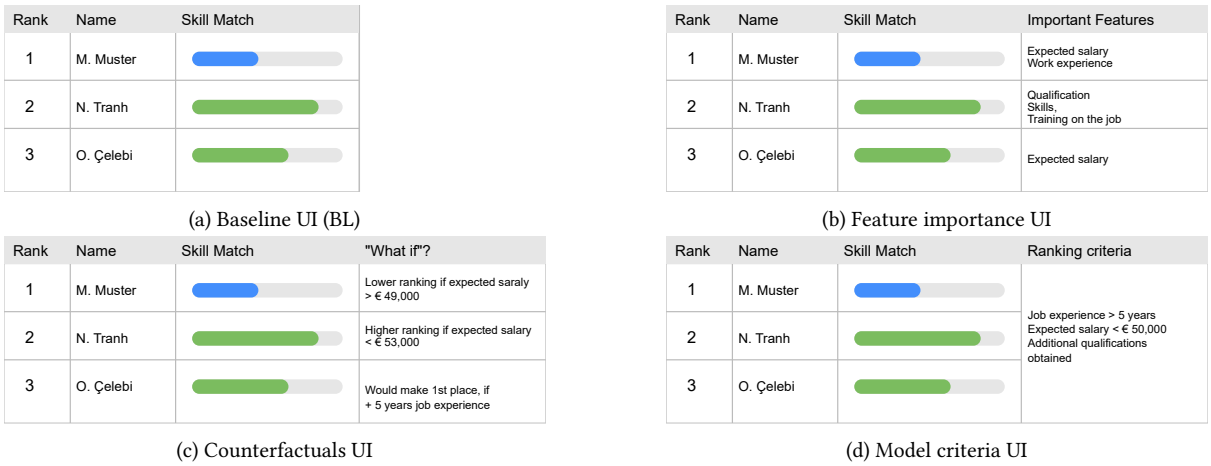
Although contextualization and audience adaptation have been emphasized, little attention has been paid to how users' AI literacy affects the efficacy of XAI elements. Explanations are meaningful only if recipients possess the cognitive and critical frameworks to interpret them: Global explanations of feature weights presuppose familiarity with model training and evaluation metrics, whereas local explanations of ranking positions require understanding how feature contributions differ across cases. Without critical literacy, HR managers may overlook XAI elements or succumb to confirmation bias, disregarding explanations that challenge their prior assumptions. Similarly, deficient practical AI literacy can lead users to fail to recognize when they are interacting with AI, thereby undermining the necessary critical scrutiny. Consequently, even well-designed explanations may fail to foster appropriate trust or may inadvertently reinforce erroneous mental models.

A significant research gap is the lack of systematic insight into how non-technical experts, such as HR managers, perceive XAI subjectively (for example, in terms of perceived usefulness or trustworthiness), how XAI contributes to their objective understanding (such as the accurate interpretation of AI outputs), and how the effectiveness of XAI varies according to different levels of AI literacy among HR managers. Addressing this gap is crucial for three reasons. First, without insight into user comprehension, organizations risk deploying XAI that engenders misplaced trust or unwarranted skepticism. Second, regulatory frameworks increasingly mandate transparency, but compliance depends on decision-makers' ability to understand the provided explanations [32]. Third, investments in AI systems for HR—especially recruiting recommender systems—must not only incorporate explainability features but also ensure that HR professionals receive the AI literacy training necessary to operate these tools responsibly and sustainably.

3. Research design

In an experiment, we queried 427 HR managers in Germany. After excluding implausible cases, the study retained 410 valid responses. We assessed the HR managers' AI literacy using the "scale for the assessment of non-experts' AI literacy" (SNAIL) [18]. The scale comprises three dimensions—*technical understanding* (TU), *critical appraisal* (CA), and *practical application* (PA)—each measured with ten Likert-scaled items from which we picked five. We selected the items from the full 30-item SNAIL scale based on their relevance to the HR domain (cf. Table 4 for an overview of items and individual statistics).

The scale demonstrated high reliability, with strong Cronbach's α values across all three dimensions (Table 1). The dimensions exhibited strong collinearity, indicating that participants who scored low/high on one dimension tended to do so on the other dimensions as well. For further analyses, we classified participants—low, medium, and high AI literacy—by dividing the scale into three equal intervals (Table 2).



(a) Baseline UI (BL)

(b) Feature importance UI

(c) Counterfactuals UI

(d) Model criteria UI

Figure 1: Recreated parts of the interface of Figure 4. Without explanations (a), with feature importance (important features) (b), counterfactuals (what-if?) (c), and model criteria (ranking criteria) (d).

We conducted an experiment to assess the effect of XAI elements on users. We researched interfaces and dashboards of existing HR tool vendors that advertise AI functions (for example, Figure 4-6). Those vendors present their dashboards as advertisements and use cases, usually on the companies' websites. Figure 4 shows an application recommender system with several active applications, their grading, skill match, and additional personal information. On closer examination, the criteria for ranking grade and skill match, and consequently, the resulting recommendation, seem ambiguous. For example, it is unclear why the recommended first-place candidate receives an A grade, despite having a substantially lower skill match than subsequent candidates. There is no explanation of how the sorting was conducted or why the ranking, which initially appears implausible, could nevertheless be justified. Moreover, it is uncertain whether the results reflect possible errors in the underlying AI system.

Table 1
Overview of the index AI Literacy

Scale	Items	Cronbach's α	Mean (SD)	*mean* ($r_{it.corr}$)
TU	5	0.92	2.85 (1.28)	0.79
CA	5	0.90	3.29 (1.22)	0.75
PA	5	0.89	3.10 (1.24)	0.73

Table 2
Overview of the AI Literacy groups

AI Literacy	N	Mean (SD)	Median
Low	91	0.32 (0.09)	0.31
Medium	185	0.61 (0.07)	0.60
High	134	0.83 (0.08)	0.81

We recreated the dashboard and experimental derivations using MOQUPS and added XAI elements to explain its AI results. Overall, the majority of (advertised as) AI dashboards contained no explicit information or warning about the results being AI-generated. Furthermore, the proposed metrics to assess, for example, performance, retention chances, or churn risks, lack further explanation. If the AI-based recommender systems in the tools used in practice resemble the illustrative materials, ambiguities are bound to occur. AI elements are utilized without further explanation of their core function, data sources, operations, or results, making the need for appropriate, targeted, and reliable XAI even more urgent.

The experiment focused on a ranking system that provides a recommendation (ranking) of incoming applications. From the advertisement material, the AI's ranking decision remains in need of explanation. Our experiment addressed this issue: we provided three different XAI-enhanced versions of the interface that each contained a different type of explanation—feature importance to assess the influencing factors on the results, counterfactuals to assess the decision boundaries of the model, and general model criteria to understand the meta-reasoning of the model (Figure 1). To facilitate understanding among non-technical professionals, we referred to the XAI elements used in our experiments using more accessible terms: "Important features" (FI), "What if?" (CF), and "Ranking Criteria" (MC).

We measured *subjective perception* using five Likert-scaled items to assess the perceived trustworthiness, transparency, comprehensibility, usefulness, and practical capability to act on the information provided by the participants. The items constituted a highly reliable indicator for subjective perception with consistently high Cronbach's α values for the baseline dashboard and all XAI enhanced dashboards (cf. Table 3).

Table 3
Overview of the index subjective perception

	N	Cronbach's α	Mean (SD)	*mean* ($r_{it.corr}$)
BL	410	0.92	2.92 (1.21)	0.79
FI	136	0.92	3.31 (1.11)	0.80
CF	138	0.90	3.10 (1.14)	0.75
MC	136	0.91	3.11 (1.10)	0.78

Objective understanding was operationalized as the number of correct responses to five factual statements derived from the information displayed on the dashboards (for example, "The person in second place has more suitable skills," or "The data foundation is known."). These questions could be answered with "yes," "no," or "you can't tell." The last option indicated that the information presented on the dash-

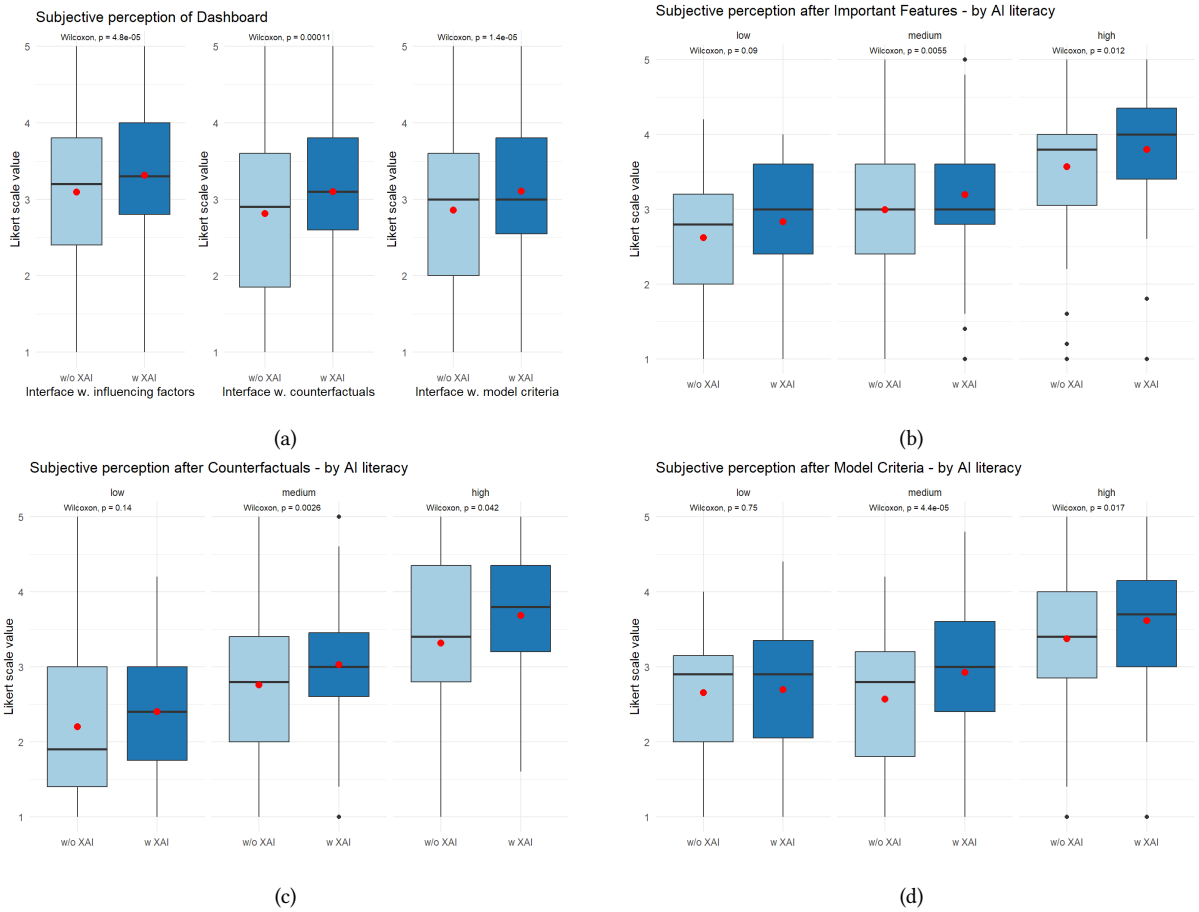


Figure 2: Subjective perception of different XAI interfaces. Without differentiation by AI literacy (a). XAI elements differentiated by AI literacy (b-d). Red dot: *mean* value.

board did not support a definitive yes or no answer. This approach enabled us to award points for correct answers and thereby assess whether individuals could correctly interpret AI dashboards. Using this design, we examined how AI literacy influences objective understanding. By incorporating XAI elements, we evaluated their effectiveness by comparing the number of correct answers and drawing conclusions about XAI's impact across different levels of AI literacy.

We randomly assigned participants to three groups. All participants first evaluated the baseline user interface without explanations. Subsequently, each group was randomly assigned to assess a second interface with a specific explanation type. This randomization was implemented to prevent selection bias and systematic differences between groups.

4. Findings

4.1. Subjective perceptions of the interfaces

The dashboards were first evaluated based on subjective perception by HR managers (Figure 2a). To compare the subjective perception with and without XAI element for each literacy group, we conducted paired-sample Wilcoxon tests, because the results of the Likert scales for baseline and XAI-enhanced interfaces were non-normally distributed (Shapiro-Wilk test: Baseline $W = 0.972$, $p < .001$; XAI $W = 0.974$, $p < .001$).

Adding any of the three XAI components to the interface yielded a consistent, though modest, upward shift in users' subjective perception. Mean ratings for the "important features" interface rose from roughly 3.10 without XAI to 3.31 with XAI, for "counterfactuals" from 2.81 to 3.10, and for "model criteria" from 2.85 to 3.11. Paired-samples Wilcoxon tests confirmed that all three increases were statistically robust ($p = 4.8 \times 10^{-5}$, $p = 1.1 \times 10^{-4}$, and $p = 1.4 \times 10^{-5}$, respectively), indicating that the addition of explanatory information produced a small-to-moderate positive effect on perceived dashboard quality across the board.

When we segmented participants by AI literacy (low, medium, high), however, the effect of XAI explanations was concentrated among users with at least moderate scores on the SNAIL scale (Figure 2b-d). Low literacy users saw slight mean increases of 0.2–0.3 points in all three interfaces, but none of these changes reached significance (influencing factors $p = .09$; counterfactuals $p = .14$; model criteria $p = .75$). Medium literacy users exhibited clear gains across every condition: the influencing-factors interface increased by about 0.3 points ($p = .0055$), counterfactuals by 0.4 points ($p = .0026$), and model criteria by 0.5 points ($p = 4.4 \times 10^{-5}$). High-literacy users displayed a comparable pattern, with all three effects reaching significance (influencing factors, $p = .012$; counterfactuals, $p = .042$; model criteria, $p = .017$), and mean improvements of roughly 0.3–0.4 points.

Taken together, these results suggest that explanatory interfaces—in the form of influencing factors, counterfactu-

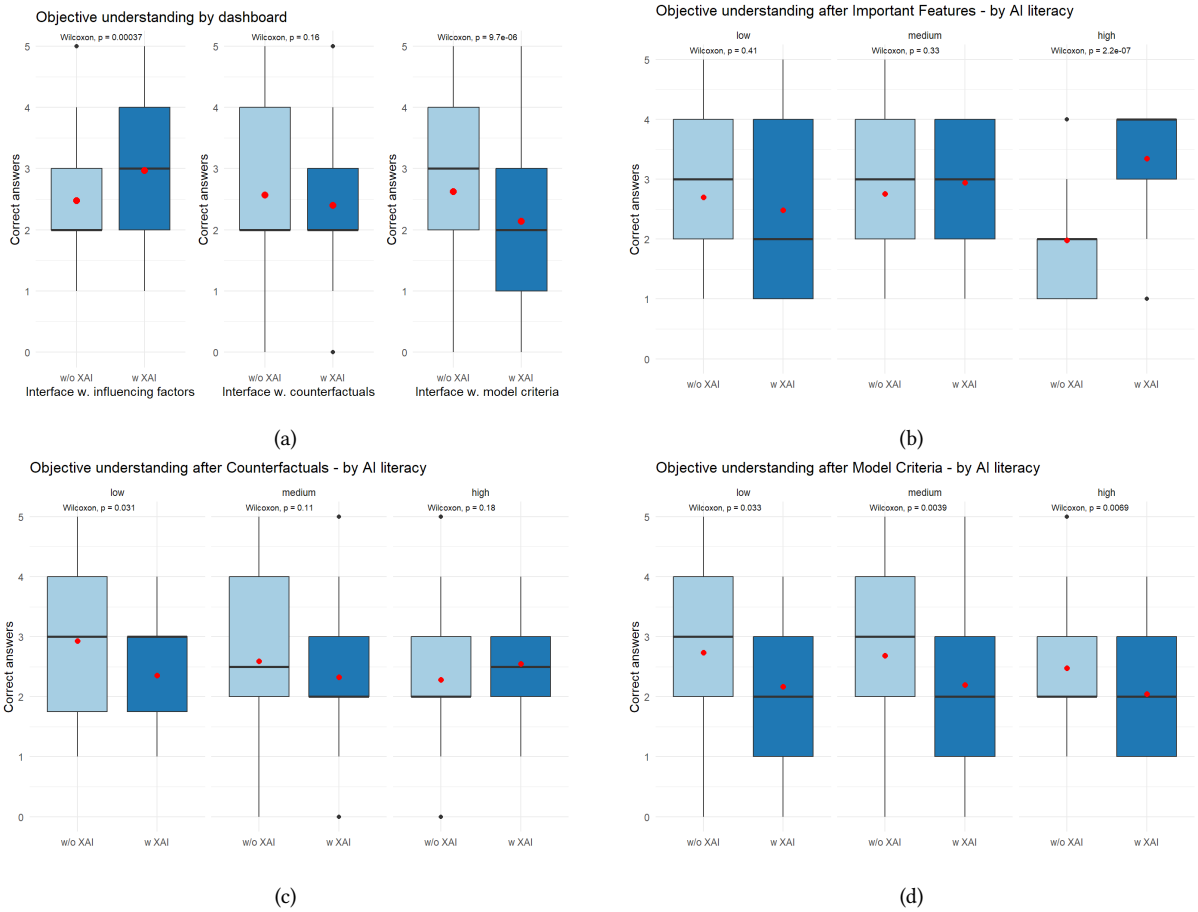


Figure 3: Objective understanding of XAI interfaces. Without differentiation by AI literacy (a). XAI elements differentiated by AI literacy (b-d). Red dot: *mean* value.

als, or explicit model criteria—systematically elevate users' subjective perceptions of a user interface, if the recipient's AI literacy is at least moderate. Crucially, the benefit is most pronounced among individuals who already possess medium or high levels of AI literacy, whereas novices derive less measurable perceived usefulness. XAI elements do not compensate for low AI literacy and do not raise the group's subjective perception of the user interface's perceived quality. This suggests that users require prior knowledge to achieve subjective improvements in any explanation—and that explanation types for low AI literacy levels must be constructed differently.

4.2. Objective understanding of the interfaces

In the second part of the experiment, we examined how the three XAI elements influenced participants' objective understanding of dashboard outputs. We again stratified these results by AI literacy (see Figure 3). To compare performance with and without each XAI overlay within each literacy group, we conducted paired-sample Wilcoxon tests, because the scoring for baseline and each XAI-enhanced dashboard was non-normally distributed (Shapiro-Wilk test: Baseline $W = 0.910$, $p < .001$; XAI $W = 0.904$, $p < .001$).

The XAI elements that show influencing factors on AI results overall performed best, since it was the only one that could increase the scores to the baseline dashboards. When participants were presented with enriched interfaces that displayed relevant features for the assessment scores, low- and medium-literacy users exhibited negligible changes in their interpretation scores (low: mean 2.70 vs. 2.48, $p = .41$; medium: mean 2.75 vs. 2.95, $p = .33$), indicating that additionally represented features neither aided nor hindered their objective understanding. By contrast, high literacy users showed a significant improvement, with mean values rising from 1.98 to 3.35 ($p = 2.2 \times 10^{-07}$). This large, highly significant effect suggests that only those with profound knowledge of AI were able to translate feature-importance annotations into more accurate data interpretations.

The results are astonishing and indicate that the (randomized) group of high AI literacy might have been subjected to a systematic bias: the group performed worst in comparison to all other groups, especially to other randomized groups of high AI literacy, where equal scores would have been expected. The randomization of the experimental groups was successful with respect to the control variable AI literacy (Kruskal-Wallis test $p = .73$). However, significant differences were observed between the experimental groups in the baseline measurement of the dependent variables for subjective perception ($p = .03$), but non-significant differences for objective understanding ($p = .54$). Such discrepancies may arise despite randomization due to random variation.

Introducing counterfactuals as "if-then" instances led to mixed outcomes. Low literacy participants performed worse when counterfactuals were present (mean 2.93 vs.

2.36, $p = .031$), a small but statistically significant decline. Medium literacy users showed a non-significant downward trend (mean 2.59 vs. 2.32, $p = .11$), and high literacy users remained essentially unchanged (mean 2.29 vs. 2.55, $p = .18$). These results suggest that novices may find counterfactual information distracting or confusing, while more experienced individuals neither consistently benefit nor suffer.

Finally, overlaying explicit model criteria as information on what the model deems essential for the presented recommendations and decisions proved to be counterproductive across the board. Low literacy users' interpretation scores fell from a mean of 2.73 to 2.17 ($p = .03$), medium literacy from 2.68 to 2.20 ($p = .004$), and high literacy from 2.48 to 2.04 ($p = .007$). All three decreases were statistically significant and of moderate effect size, indicating that detailed model criteria overwhelmed users regardless of their AI background, leading to poorer objective understanding.

Taken together, these findings demonstrate that XAI elements do not uniformly enhance users' comprehension of dashboard information. While influencing factors can significantly boost accurate interpretations for technically savvy users, counterfactuals and explicit model criteria may impair or fail to improve objective understanding—particularly among those with limited AI literacy.

5. Discussion

Our study provides novel empirical evidence on how AI literacy shapes HR managers' interpretations of AI-based recruitment recommender systems. Addressing our first research question, we found that XAI elements can increase the quality of AI interfaces, depending on users' AI literacy levels. However, enhanced subjective perceptions of informativeness, trustworthiness, and interpretability were statistically significant only for participants with at least moderate AI literacy. This suggests that users must possess foundational conceptual and critical skills to experience benefits from explainability elements in recruitment interfaces [18]. Low-literacy users gained little subjective value from XAI, in line with research that emphasizes the need for contextualized, domain-specific AI literacy interventions, and potentially confusing and misleading complexity [23, 25].

Turning to our second research question, we observed a paradox: XAI explanations did not uniformly improve, and in some cases, impaired objective understanding. Only the "important features (feature importance)" element improved performance, but primarily for high literacy users. Counterfactual and model-criteria explanations either had no effect or reduced objective understanding, particularly among participants with low and medium literacy. Notably, high literacy users—while more likely to benefit subjectively from explanations—exhibited lower absolute understanding, which could be a sign of overconfidence in their AI literacy. This finding complicates earlier claims about the universal efficacy of XAI in fostering trust and better decisions [14, 15], and supports concerns about information overload and cognitive miscalibration of non-technical users [17].

Our results align with and extend the literature on the intersection of AI literacy and XAI. The state of research has emphasized tailoring explanations to users' backgrounds and roles [30, 29], but empirical investigations of explanation efficacy across literacy gradients in HR remain sparse. Our results indicate that designers should carefully tailor explanation formats to the target audience's expertise level, avoiding overly complex explanations or too granular details. Our data suggest that XAI elements, when implemented in HR dashboards, may reinforce subjective confidence without reliably supporting accurate and responsible decision-making. This risk is amplified by regulatory demands for explainability and transparency in high-stakes contexts, such as those imposed by the EU AI Act and GDPR [7, 8]. These findings underline the importance of a nuanced approach to deploying AI tools in sensitive contexts, one that carefully balances explanation design with the end user's AI knowledge. First, our baseline interface shows no clues to AI origins, data quality, or assumptions, leaving users to judge recommendations without deeper insights into underlying models or AI design decisions. Second, while XAI annotations enhance perceived information quality, they do not automatically translate into improved decision-making based on the recommender system. By contrasting subjective perceptions with objective understanding across literacy levels, we uncover the complex interplay between user expertise and explanation efficacy.

Upon further examination, we discovered that the type of explanation is crucial. Only additional important features—inspired by feature importance models without considering effect strengths—yielded an overall improvement in objective understanding performance. In contrast, counterfactual explanations and model-criteria summaries hindered users' evaluations. Counterfactuals—"What would have to change for the AI's decision to differ?"—might impose substantial cognitive demands because they frame reasoning through a negated, hypothetical scenario rather than presenting a direct rationale. This result challenges prior claims about the universal effectiveness of counterfactuals [33]. Likewise, model-criteria explanations—which merely translate the AI's decision rules into human-like justification—systematically failed, regardless of participants' literacy.

Finally, our data reveal an overconfidence effect among high-literacy users: their strong self-assurance contradicts their performance in accurately showing objective understanding. This suggests that relying on self-reported AI literacy may obscure critical performance gaps; future research should incorporate objective measures of AI knowledge. This mirrors work in metacognition and digital literacy, suggesting that self-assessment may not be a reliable proxy for AI literacy, i.e., proper comprehension, critical capacity, or practical application [24, 23]. For HRM practice, this implies that both AI literacy training and explanation design must be empirically validated for effectiveness.

We conclude that XAI is not a one-size-fits-all remedy in riskful settings like HR. Explanations must be carefully tailored to users' expertise and meet the specific demands of their domain. However, providing several explanations at once to satisfy different AI literacy needs could raise the complexity of HR interfaces even further. Our study challenges the assumption that XAI elements are universally beneficial in HR recommender systems. Their impact is conditional on users' AI literacy, the type of explanation, and the context of use. Effective, responsible deployment of XAI in HR, therefore, requires an integrated approach: robust, context-sensitive literacy training, careful user-centered explanation design, and ongoing evaluation of both subjective perception and objective understanding outcomes.

6. Conclusion and Future Work

This research showed that HR managers' AI literacy fundamentally affects the effectiveness of explainable AI elements in recruiting recommender systems. While XAI features enhance perceived transparency and trust among more literate users, they do not guarantee improved objective understanding, sometimes even undermining it. These findings call for a more nuanced, empirically grounded approach to the design and implementation of XAI in HRM. Specifically, we find that *XAI is not a universal remedy* to make AI, its functions, and results accessible to non-technical professionals. Explanation strategies must be tailored to users' actual literacy and cognitive needs. *AI literacy is critical* and investments in AI for HR should be coupled with targeted, domain-specific literacy initiatives that inform and teach on specific tools and utilities and their mechanisms, like recommender systems. Finally, the design of explanations should *address diverse interpretative needs* with flexible formats to serve users with heterogeneous backgrounds, while not overwhelming or misleading them, or giving them a false sense of understanding.

Future research should address several open questions: How can AI literacy interventions best be integrated into HR training programs, and what pedagogical approaches are most effective for non-technical professionals? What hybrid or adaptive explanation strategies (for example, layered explanations, user-driven customization) can accommodate different literacy levels without increasing interface complexity? How do explanation effects evolve with repeated exposure, feedback, or organizational learning? What are the organizational and regulatory implications of overconfidence or miscalibration in AI-literate HR professionals? By pursuing these avenues, we can move toward HR recommender systems that are not only technically robust and legally compliant but also meaningfully transparent, fair, and supportive of human expertise. All in all, these results and the future outlook on open research emphasize that applied AI—in any domain—needs explanations. Explainability is a transdisciplinary process that involves technical interfaces, pedagogical and psychological learning capabilities, and social, vocational, or professional standards of specific groups, like HR managers.

Acknowledgments

The research was part of the project "TRANKI – Standards for transparent AI", funded by the Hans Böckler Foundation (Grant no.: 2022-797-2).

Declaration on Generative AI

During the preparation of this work, the author(s) used OpenAI's GPT-4.1 and o4-mini, as well as Grammarly, to check grammar and spelling and to enhance the writing style. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] Y. Zhai, L. Zhang, M. Yu, AI in Human Resource Management: Literature Review and Research Implications, Journal of the Knowledge Economy (2024). doi:10.1007/s13132-023-01631-z.

[2] A. Malik, P. Budhwar, B. A. Kazmi, Artificial intelligence (AI)-assisted HRM: Towards an extended strategic framework, Human Resource Management Review 33 (2023) 100940. doi:10.1016/j.hrmr.2022.100940.

[3] E. Drage, K. Mackereth, Does AI Debias Recruitment? Race, Gender, and AI's "Eradication of Difference", Philosophy & Technology 35 (2022) 1–25. doi:10.1007/s13347-022-00543-1.

[4] E. K. Kelan, Algorithmic inclusion: Shaping the predictive algorithms of artificial intelligence in hiring, Human Resource Management Journal 34 (2024) 694–707. doi:10.1111/1748-8583.12511.

[5] M. R. Edwards, A. Charlwood, N. Guenole, J. Marler, HR analytics: An emerging field finding its place in the world alongside simmering ethical challenges, Human Resource Management Journal 34 (2024) 326–336. doi:10.1111/1748-8583.12435.

[6] M. Klöpper, S. Köhne, Shifting Structures – a systematic Literature Review on People Analytics and the Future of Work, ECIS 2023 Research Papers (2023) 1–19.

[7] European Commission, Regulation of the European Parliament and of the Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, Technical Report 2021/0106 (COD), Brussels, 2021.

[8] European Parliament, General Data Protection Regulation: GDPR, 2016.

[9] J. Du, Exploring Gender Bias and Algorithm Transparency: Ethical Considerations of AI in HRM, Journal of Theory and Practice of Management Science 4 (2024) 36–43. doi:10.53469/jtpms.2024.04(03).06.

[10] A. Fabris, N. Baranowska, M. J. Dennis, D. Graus, P. Hacker, J. Saldivar, F. Zuiderveen Borgesius, A. J. Biega, Fairness and Bias in Algorithmic Hiring: A Multidisciplinary Survey, ACM Transactions on Intelligent Systems and Technology (2024). doi:10.1145/3696457.

[11] K. Simbeck, HR analytics and Ethics, IBM Journal of Research and Development 63 (2019) 9:1–9:12. doi:10.1147/JRD.2019.2915067.

[12] A. Köchling, S. Riazy, M. C. Wehner, K. Simbeck, Highly Accurate, But Still Discriminatory: A Fairness Evaluation of Algorithmic Video Analysis in the Recruitment Context, Business & Information Systems Engineering 63 (2021) 39–54. doi:10.1007/s12599-020-00673-w.

[13] C. Molnar, Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, 2 ed., Self-publishing, Munich, 2022.

[14] U. Bhatt, A. Xiang, S. Sharma, A. Weller, A. Taly, Y. Jia, J. Ghosh, R. Puri, J. M. F. Moura, P. Eckersley, Explainable machine learning in deployment, in: M. Hildebrandt (Ed.), Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Association for Computing Machinery, New York, 2020, pp. 648–657. doi:10.1145/3351095.3375624.

[15] T. Speith, A Review of Taxonomies of Explainable Artificial Intelligence (XAI) Methods, in: Association for Computing Machinery (Ed.), FAccT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, Association for Com-

puting Machinery, New York, 2022, pp. 2239–2250. doi:10.1145/3531146.3534639.

[16] M. Khalili, Against the opacity, and for a qualitative understanding, of artificially intelligent technologies, AI and Ethics (2023). doi:10.1007/s43681-023-00332-2.

[17] K. Bauer, M. Zahn, O. Hinz, Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing, Information Systems Research (2023). doi:10.1287/isre.2023.1199.

[18] M. C. Laupichler, A. Aster, N. Haverkamp, T. Raupach, Development of the "Scale for the assessment of non-experts' AI literacy" – An exploratory factor analysis, Computers in Human Behavior Reports 12 (2023) 100338. doi:10.1016/j.chbr.2023.100338.

[19] A. Carolus, M. J. Koch, S. Straka, M. E. Latoschik, C. Wienrich, MAILS - Meta AI literacy scale: Development and testing of an AI literacy Questionnaire based on well-founded Competency Models and psychological Change- and Meta-Competencies, Computers in Human Behavior: Artificial Humans 1 (2023) 100014. doi:10.1016/j.chbah.2023.100014.

[20] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A Survey of Methods for Explaining Black Box Models, ACM Computing Surveys 51 (2019) 1–42. doi:10.1145/3236009.

[21] F. Bodria, F. Giannotti, R. Guidotti, F. Naretto, D. Pedreschi, S. Rinzivillo, Benchmarking and survey of explanation methods for black box models, Data Mining and Knowledge Discovery 37 (2023) 1719–1778. doi:10.1007/s10618-023-00933-9.

[22] M. Bhat, D. Long, Designing Interactive Explainable AI Tools for Algorithmic Literacy and Transparency, in: Designing Interactive Systems Conference, ACM, Copenhagen Denmark, 2024, pp. 939–957. doi:10.1145/3643834.3660722.

[23] T. Lintner, A systematic review of AI literacy scales, NPJ science of learning 9 (2024) 50. doi:10.1038/s41539-024-00264-4.

[24] D. T. K. Ng, W. Wu, J. K. L. Leung, T. K. F. Chiu, S. K. W. Chu, Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach, British Journal of Educational Technology 55 (2024) 1082–1104. doi:10.1111/bjet.13411.

[25] M. Pinski, T. Hofmann, A. Benlian, AI Literacy for the top management: An upper echelons perspective on corporate AI orientation and implementation ability, Electronic Markets 34 (2024) 24. doi:10.1007/s12525-024-00707-1.

[26] G. Bassellier, I. Benbasat, B. H. Reich, The Influence of Business Managers' IT Competence on Championing IT, Information Systems Research 14 (2003) 317–336. doi:10.1287/isre.14.4.317.24899.

[27] M.-A. Clinciu, H. Hastie, A Survey of Explainable AI Terminology, Proceedings of the 1st Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence (NL4XAI 2019) (2019) 8–13. doi:10.18653/v1/W19-8403.

[28] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, G.-Z. Yang, XAI-Explainable artificial intelligence, Science robotics 4 (2019). doi:10.1126/scirobotics.aay7120.

[29] S. Chowdhury, S. Joel-Edgar, P. K. Dey, S. Bhattacharya, A. Kharlamov, Embedding transparency in artificial intelligence machine learning models: Managerial implications on predicting and explaining employee turnover, The International Journal of Human Resource Management 34 (2023) 2732–2764. doi:10.1080/09585192.2022.2066981.

[30] E. Cambria, L. Malandri, F. Mercorio, M. Mezzanzanica, N. Nobani, A survey on XAI and natural language explanations, Information Processing & Management 60 (2023) 103111. doi:10.1016/j.ipm.2022.103111.

[31] A. Holzinger, A. Saranti, C. Molnar, P. Biecek, W. Samek, Explainable AI Methods - A Brief Overview, in: A. Holzinger, R. Goebel, R. Fong, T. Moon, K.-R. Müller, W. Samek (Eds.), xxAI - Beyond Explainable AI: International Workshop, Held in Conjunction with ICML 2020, Springer, Cham, 2022, pp. 13–38. doi:10.1007/978-3-031-04083-2_2.

[32] N. Kühl, C. Meske, M. Nitsche, J. Lobana, Investigating the Role of Explainability and AI Literacy in User Compliance, SSRN Electronic Journal (2023). doi:10.2139/ssrn.4558966.

[33] R. Mazzine Barbosa de Oliveira, S. Goethals, D. Brughmans, D. Martens, Unveiling the Potential of Counterfactuals Explanations in Employability, Technical Report, arXiv, 2023. doi:10.48550/arXiv.2305.10069.

[34] M. C. Laupichler, A. Aster, J.-O. Perschewski, J. Schleiss, Evaluating AI Courses: A Valid and Reliable Instrument for Assessing Artificial-Intelligence Learning through Comparative Self-Assessment, Education Sciences 13 (2023) 978. doi:10.3390/educsci13100978.

A. Additional tables of indices statistics

Table 4
Selected items from [18, 34] and individual item statistics

Item	Statement: *I can ...*	Mean (SD)	Item total *r*	α if dropped
TU 1	explain how AI applications make decisions.	2.92 (1.28)	0.80	0.90
TU 2	explain the difference between general (or strong) and narrow (or weak) artificial intelligence.	2.72 (1.31)	0.80	0.90
TU 3	explain how machine learning works at a general level.	2.85 (1.29)	0.81	0.90
TU 4	describe the concept of explainable AI.	2.88 (1.29)	0.79	0.90
TU 5	evaluate whether media representations of AI (for example, in movies or video games) go beyond the current capabilities of AI technologies.	2.89 (1.22)	0.76	0.91
CA 1	explain why data privacy must be considered when developing and using artificial intelligence applications.	3.46 (1.23)	0.72	0.88
CA 2	identify ethical issues surrounding artificial intelligence.	3.28 (1.19)	0.74	0.88
CA 3	name weaknesses of artificial intelligence.	3.23 (1.20)	0.74	0.88
CA 4	describe potential legal problems that may arise when using artificial intelligence.	3.16 (1.25)	0.75	0.87
CA 5	explain why data plays an important role in the development and application of artificial intelligence.	3.33 (1.22)	0.78	0.87
PA 1	give examples from my daily life (personal or professional) where I might be in contact with artificial intelligence.	3.31 (1.23)	0.71	0.87
PA 2	tell if the technologies I use are supported by artificial intelligence.	3.04 (1.24)	0.75	0.86
PA 3	assess if a problem in my field can and should be solved with artificial intelligence methods.	3.10 (1.22)	0.76	0.86
PA 4	name applications in which AI-assisted natural language processing/understanding is used.	2.80 (1.31)	0.72	0.87
PA 5	critically evaluate the implications of artificial intelligence applications in at least one subject area.	3.25 (1.21)	0.72	0.87

B. Screenshots of AI Dashboards

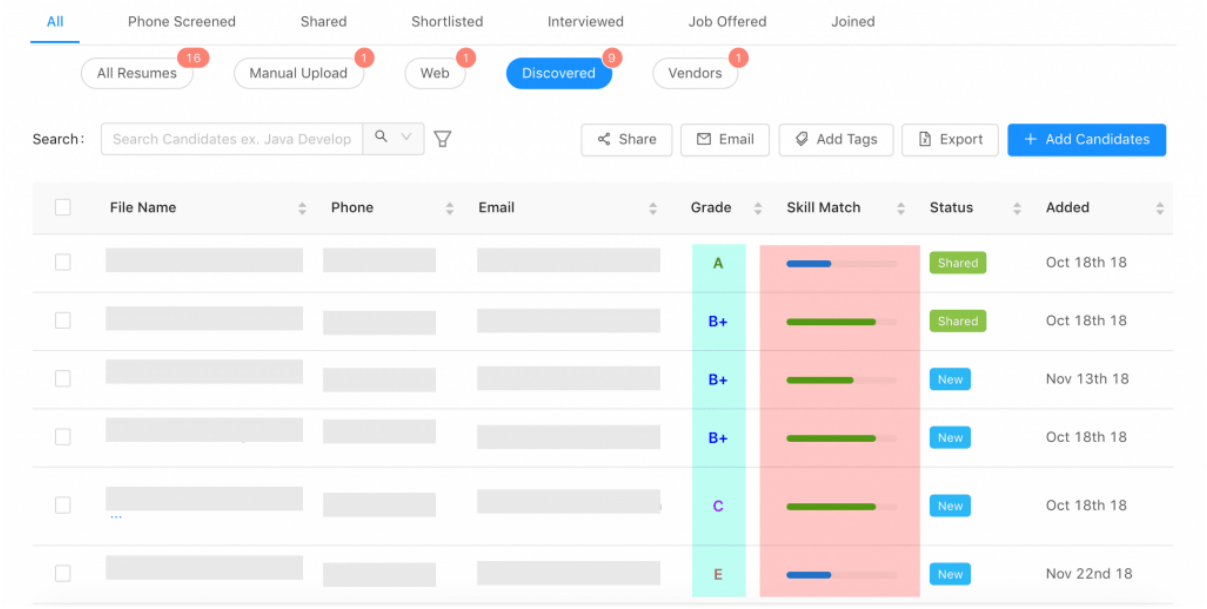


Figure 4: Resume screening and recommender system. Source: https://cvviz.com/product/resume-screening/.

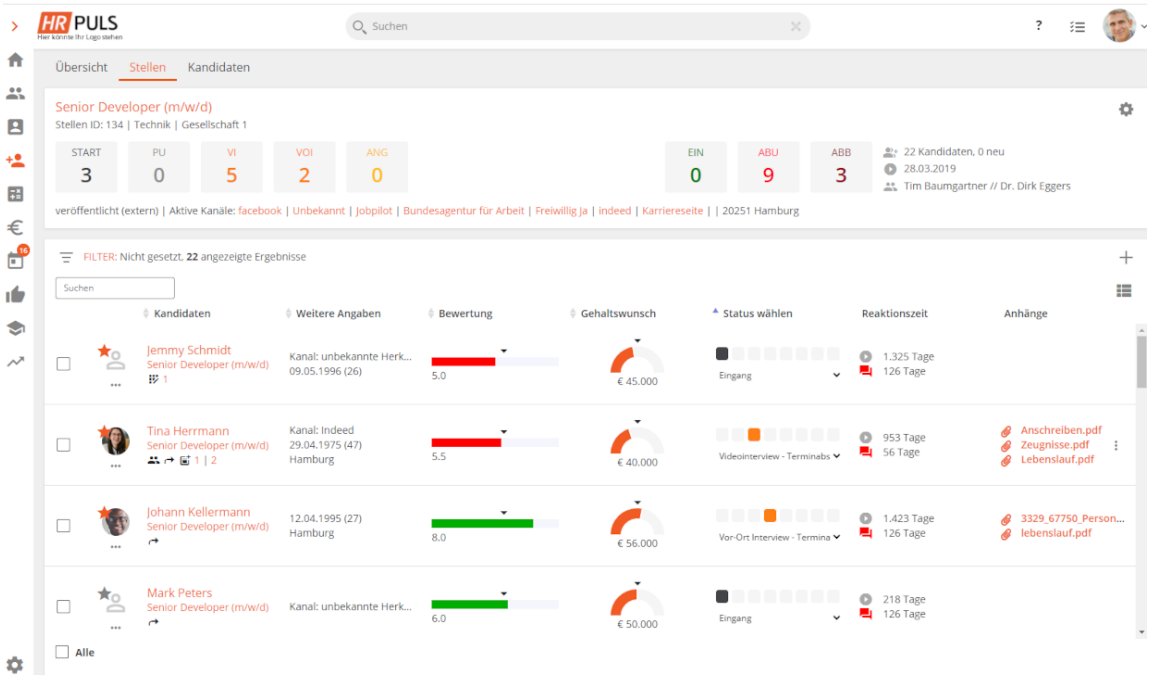


Figure 5: Resume screening and recommender system. Source: https://www.hrpuls.de/bewerbermanagement.html.

Figure 6: Talent management and development recommender system. Source: https://www.softwareadvice.com/bi/edligo-talent-analytics-profile/.