

Resume-ing Control: (Mis)Perceptions of Agency Around GenAI Use in Recruiting Workflows

SAJEL SURATI, New York University, USA

ROSANNA BELLINI*, New York University, USA

EMILY BLACK*, New York University, USA

When generative AI (genAI) systems are used in high-stakes decision-making, its recommended role is to aid, rather than replace, human decision-making. However, there is little empirical exploration of how professionals making high-stakes decisions, such as those related to employment, perceive their agency and level of control when working with genAI systems. Through interviews with 22 recruiting professionals, we investigate how genAI subtly influences control over everyday workflows and even individual hiring decisions. Our findings highlight a pressing conflict: while recruiters believe they have final authority across the recruiting pipeline, genAI has become an invisible architect that shapes the foundational building blocks of information used for evaluation, from defining a job to determining good interview performances. The decision of whether or not to adopt was also often outside recruiters' control, with many feeling compelled to adopt genAI due to calls to integrate AI from higher-ups in their business, to combat applicant use of AI, and the individual need to boost productivity. Despite a seemingly seismic shift in how recruiting happens, participants only reported marginal efficiency gains. Such gains came at the high cost of recruiter deskilling, a trend that jeopardizes the meaningful oversight of decision-making. We conclude by discussing the implications of such findings for responsible and perceptible genAI use in hiring contexts.

CCS Concepts: • **Social and professional topics** → *Computing / technology policy*.

Additional Key Words and Phrases: Algorithmic Decision-making, Generative AI, Recruitment, Human-Computer Interaction, Behavioral Agency

ACM Reference Format:

Sajel Surati, Rosanna Bellini, and Emily Black. 2026. Resume-ing Control: (Mis)Perceptions of Agency Around GenAI Use in Recruiting Workflows. In *The 2026 ACM Conference on Fairness, Accountability, and Transparency (FAccT '26)*, June 25–28, 2026, Montreal, QC, Canada. ACM, New York, NY, USA, 22 pages. <https://doi.org/10.1145/3805689.3806498>

1 Introduction

Generative AI systems (genAI) are transforming the workplace, with studies reporting adoption rates up to 80% [22, 59]. Recruitment is an exemplar of this shift: major platforms such as LinkedIn, alongside numerous startups, now integrate genAI across the recruitment funnel: from sourcing and screening to automating the interviewing process itself.

In high-stakes decision-making settings like hiring, scholars and policymakers emphasize the need for “AI-assisted” decision-making, asserting that AI should support, not replace human judgment [33, 53]. But how realistic is maintenance of human agency in practice? While prior work has both problematized and praised

*These authors contributed equally.

Authors' Contact Information: Sajel Surati, New York University, New York, New York, USA, sajel.s@nyu.edu; Rosanna Bellini, New York University, New York, New York, USA, bellini@nyu.edu; Emily Black, New York University, New York, New York, USA, emilyblack@nyu.edu.



This work is licensed under a Creative Commons Attribution 4.0 International License.

FAccT '26, Montreal, QC, Canada

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2596-8/2026/06

<https://doi.org/10.1145/3805689.3806498>

“human in the loop” or “AI assisted” decision-making through controlled experiments [28, 38–40, 75], there is little empirical work investigating how practitioners actually navigate incorporating AI—especially generative AI—into their decision-making on the ground. We use the field of recruitment and hiring as a case study to examine how agency is experienced and redistributed in AI-mediated hiring processes, both on the level of making individual hiring decisions, and systemically, regarding the adoption of or resistance to using genAI in their workflows. We focus primarily on the use of text-generative AI models with multimodal inputs, from ChatGPT to specialized tools that analyze audio and video from interviews. Specifically, we investigate the following:

- **RQ1:** *To what extent do recruiters feel they maintain agency in their decision-making?*
- **RQ2:** *To what extent do recruiters perceive AI adoption as self-directed choice versus a necessity?*
- **RQ3:** *How has shared human-AI decision-making changed the nature of recruiting decision-making processes, and what are the implications of these changes?*

We investigate these questions through interviews with 22 recruiters across eight industry sectors, identifying a widening gap between participants’ perceived professional autonomy and the reality of AI-driven workflows. The mechanisms and implications of this phenomenon form our main contributions.

First, on an individual level (Section 4.1), we find that while recruiters claim ultimate authority—often asserting that either AI does not make decisions or is relegated to low-level tasks—this presents an incomplete picture. Instead, our findings characterize genAI as an *invisible architect*. By defining job qualifications, creating interview rubrics, and summarizing candidate interactions, genAI structures the decision space and exerts a profound, unacknowledged influence on hiring—often before humans realize a decision is being made. Second, on a systemic level (Section 4.2), we find that recruiters often lack agency regarding whether to include this invisible architect in their workflows. Many describe being boxed into using AI due to pressures to maximize personal efficiency, top-down mandates framing genAI as a business necessity, and an escalating arms race where candidate and competitor use of genAI requires automated countermeasures.

Finally, we highlight that the integration of genAI may lead to a subtle deterioration of human agency with three alarming repercussions (Section 4.3). First, despite isolated instances of recruiters upskilling through AI use, we identify a broader risk of workforce deskilling in both domain expertise and the capacity to make deliberate, reasoned decisions about when and how to deploy AI, threatening professional oversight. This trend was particularly severe among junior recruiters. Second, the AI arms race erodes recruiter trust in job-seekers’ applications, pushing recruiters toward heuristic ‘vibe checks’ that may intensify systemic bias, even as some recruiters perceive genAI as a more objective alternative. Overall, despite massive investment, according to our participants’ reports, genAI has failed to transform applicant placement quality. Instead, the industry is inadvertently ‘boiling the ocean’—investing exhaustive effort for seemingly marginal gains while introducing new vulnerabilities like high turnover from poorly matched candidates. Beyond these contributions, to map how genAI systems enter hiring workflows, we perform a functionality analysis of genAI hiring tools mentioned in interviews (Appendix A), and discuss how our findings inform paths towards reclaiming human agency in recruiting workflows through AI governance and future research in Section 5.

2 Background and Related Work

Recruiting refers to the process of identifying, attracting, and selecting individuals to fill specific job roles within a team or a company. This may be performed by recruiters who are either in-house (often located within human resource departments), contracted out to an external agency, or performed by a current employee (a *hiring manager*). A recruitment pipeline consists of five chronological stages: job design, sourcing, screening, interviewing, and offering. First, a job is designed or ‘themed,’ whereby a person (typically a hiring manager) defines the responsibilities for a position. Recruiters can either advertise the position to attract candidates (*inbound*), or approach viable candidates (*outbound*) with a proposition to apply. Applicants are then screened,

where their application materials are evaluated against the position's requirements and organizational factors; those who pass may be invited to an interview to evaluate their qualifications and 'fit' across different teams or increasing staff seniority level. A hiring manager may then select a candidate to extend a job offer. Bias is well-documented in traditional hiring decisions, from résumé screening to culture-fit assessments [12, 25, 30, 49, 63, 67]. Indeed, algorithmic hiring tools are often marketed to reduce such bias [8, 13, 16, 42, 65]—yet as we discuss below, prior work suggests they introduce their own bias.

Algorithmic Hiring Decisions. Our work studies the impacts of AI and genAI tools on an individual's perception of their agency in recruitment workflows. Prior literature highlights potential ethical concerns regarding AI in hiring, largely centered around the potential for discrimination. AI audits, following methodologies like those proposed by Costanza-Chock et al. [26], enable investigators to find evidence of bias previously overlooked. Several papers provide technical audits that have quantified rates of racial and gender-based discrimination in both traditional AI [10, 77] and generative AI [1, 2, 4, 36, 62, 78], while sociotechnical audits have scrutinized claims made by companies developing recruitment tools [65, 68, 72]. Closely related to our work, Li et al. [54] interviewed recruiting professionals regarding traditional AI in screening processes, though their focus remains on barriers to adoption. Similarly, Wilson et al. [79] performed a user study on how humans are influenced by AI-generated candidate recommendations and demonstrate that humans are indeed influenced by AI-based recommendations. However, this study was not interview-based and did not include human resource professionals, instead focusing on isolated recommendations rather than the broader recruitment pipeline. As prior work has focused on traditional AI or did not involve practitioners, this is, to our knowledge, the first study to interview recruiters directly about their use of *generative* AI across the hiring pipeline.

Behavioral Agency in Human-Machine Interactions. Our work builds upon prior work regarding human agency in AI-assisted decision-making teams. In this work, we focus on *behavioral agency* (derived from *agere* “to set in motion, drive forward; to do, perform” [29]), which focuses on an individual's or group's capacity to shape decision-making. Behavioral agency has been explored across a wealth of complex sociotechnical systems [20, 65, 73, 85], being so central to well-designed systems that developers have long been encouraged to “support internal locus of control” [71], despite the tension between human and machine agency [73]. Human reliance on AI is driven by internal factors, such as trust and affinity for technology [20, 52], alongside system traits, like perceived fairness or task simplicity [74]. Thus, investigations into human-AI interactions suggest that decision-making is characterized by relational agency, where a human's capacity to act is constrained by an AI's architecture [57, 60].

GenAI complicates behavioral agency through introducing problems into the heart of decision-making. For instance, LLMs can display “*social sycophancy*”: human-like reasoning and mimic authoritative tones, whereby a model prioritizes saving face over providing critical or truthful feedback [23, 58]. Bućinca [18] and Lai [53] thus caution that the true locus of decision-making authority often resides in pre-computational stages: how an AI frames, filters, and prioritizes data before a practitioner intervenes. Put simply, claims of choice can be done so within a “*curated environment*” where the most cognitively taxing labor, the intentional, initial separation of data, has been performed by a technical blueprint [18]. Prior work has highlighted further issues around navigating agency in human-AI collaboration [45], including governance [9], high-stakes errors [83], and deskilling [24, 51, 84]. Given genAI's potential to affect a decision-maker's agency, it is imperative to learn how decision-makers navigate such effects. We do so in this work through interviewing recruiters about their use of AI in hiring workflows.

3 Methodological Approach

Our study adopts a mixed methods approach, whereby we conducted a series of semi-structured interviews with 22 recruiters across the space of four months (Sept–Dec 2025) on their acceptance and use of such tools in their

work practices (Section 4). We then paired this with a functionality analysis of the genAI hiring tools mentioned in their accounts (see Table 2, Appendix A).

Semi-structured interviews. Given the rapidly evolving nature of genAI integration into recruitment [56], we chose to use semi-structured interviews for both epistemological and practical reasons. As genAI is best characterized by idiosyncratic workflows that vary significantly between individual practitioners, we identified that a semi-structured format enabled the discovery of unanticipated insights: the ‘how’ and ‘why’ behind AI adoption. In addition, recruiters, particularly in large organizational contexts often have to navigate a tension between company policy and practical ‘workarounds’ [69], or unofficial tool use (e.g., Shadow AI [21]). Thus, we hoped that the conversational atmosphere of a semi-structured interview could help build rapport with participants, leading to candid disclosures.

Recruitment Strategy. Participants were recruited using a multi-pronged sampling strategy. First, we conducted purposive sampling via LinkedIn ($N=19$), targeting employees at organizations via the active job postings search feature to increase the likelihood of reaching potential interviewees who were engaged in hiring. An active advert constitutes a listing being discoverable by verified LinkedIn users and remains posted on the site for less than six months¹. We chose to prioritize identifying potential applicants via adverts over searching for specific people to mitigate the algorithmic bias that is baked into LinkedIn’s recommender engines, which prioritize high-follower organizations and ‘popular’ content tailored to a researcher’s personal browsing habits, professional connections, or job history [10].

We sent an initial outreach message to 560 individuals based in the US who met the following inclusion criteria: (a) worked in human resources, recruitment, or run or directly play a significant role in a micro or small business, and (b) reported using genAI in their hiring processes of their current role, or a past role of no later than eight months ago. We chose a purposive sampling approach to cover a range of different states, company sizes, and industries, which included: technology, manufacturing, biotechnology, healthcare, entertainment, publishing, marketing, and nonprofit sectors. While a total of 41 potential participants expressed interest via this method (7.3%), our yield was substantively lower at only 19 participants despite follow-ups (3.4%). To supplement this approach, we also used convenience sampling ($N=1$) by contacting professional acquaintances who had previously described using genAI in their recruitment workflows. Finally, the study also attempted to use referral-based (i.e., snowball) sampling ($N=2$) by encouraging participants to disseminate our study details within their professional networks, which yielded a low conversion rate.

Interview Process. Our interview protocol was designed around five topic areas: introductory conversation, grounding the context, identifying how genAI has been used in hiring, their perception of agency over choices regarding their work, and the future potential of AI. For our second phase, we also leveraged the Critical Incident Technique [34], in which participants were encouraged to reconstruct their most recent hiring event to minimize recall bias and capture relevant data. Additionally, we leveraged a normalization framing inspired by the results of Catherine et al. [21], which describe hesitancy to admit the use of AI tooling, for two questions pertaining to deskilling and the perceived benefits of AI in our protocol to gently challenge our participants’ internal logic. Our interview protocol underwent three iterations and refinement. We refined the protocol’s timing and language through an internal trial with a member of the research team, followed by a pilot interview with a recruitment professional that revealed additional insights. Due to the richness of the pilot data, we chose to include their account in the final analysis (P_1). Interested readers can find a copy of our interview protocol in Appendix B.

Data were collected via synchronous, semi-structured interviews conducted over Zoom video-conferencing software. Each session lasted approximately 60 minutes (46–64 minutes) and was facilitated by a moderator and a designated notetaker to ensure that all data were captured effectively. The research team also informed

¹LinkedIn Help Answers report “Job posts will stay open until you close the job manually or until the job automatically closes after 6 months. You’ll continue to be billed until the job is closed.”

participants of future plans for the data, their timeline for concluding the interviews, and the option to receive a preprint of the research prior to publication for their personal use. Data that were collected were automatically transcribed using Zoom auto-transcribe before being manually verified for spelling and turn-taking errors by the first author of this work for discrepancies. Personal identifiable information, such as names, company organizations, and specific locations were replaced with placeholders to mitigate the unnecessary sharing of identifying data outside of collection. De-identified transcripts were then stored and analyzed in a secure Box folder only accessible to members of the research team.

Analysis. Our interview data were analyzed using a reflexive thematic analysis approach [14] using the qualitative analysis software *Atlas.ti* [37]. Following familiarization by conducting a close reading of the transcripts, we used a hybrid approach of inductive and deductive coding. We drafted an initial set of deductive codes ($N=14$) in accordance with our research questions (Section 1), drawing from prior work on human agency in AI-assisted decisions [85]. These codes were then supplemented with a round of inductive open coding from two transcripts, resulting in 67 additional codes, for a total of 81 initial codes. The research team then independently coded a subset of the transcripts to test the codebook’s applicability and refine definitions. The codebook was refined further through two subsequent cycles with synchronous meetings for coordinating disagreements for a final codebook of 54 codes. Inter-rater reliability of Krippendorff’s Alpha were calculated as a mechanism for consensus-building, at 0.80 agreement (“*Very Good*” [50]). See Figure 3 in Appendix C for a copy of our final codebook.

To further investigate the unique tools mentioned by our interviews, we first confirmed that such software were regularly used by other industry professionals (i.e., non-bespoke) with popular aggregators [64]. Then, we identified AI features across each system, then categorized them according to two qualities: what stage of the recruitment pipeline the tool is intended to be used for (*recruitment phase*), and, what type of action being performed (*functionality*). Our results can be shown in Table 2 in Appendix A.

Ethics and Limitations. This study was reviewed and judged to be exempt by the New York University Institutional Review Board (IRB). Participants were sent an information sheet and a consent form prior to the interview, and written informed consent were taken prior to the interview with participants being asked to reconfirm their consent verbally at the start of the session. Audio recording was voluntary and de-identified note taking was offered as an alternative. Participants were provided a \$50 USD honorarium as remuneration for their time and contributions.

Given the potential for social desirability bias [14], whereby participants could feel pressure to report policy-compliant or ‘ethical’ use of genAI, we designed our protocol to utilize a normalization framing to reduce feelings of discomfort. To further protect our participants from potential retaliation by employers, we chose to not report company names or sizes, report agnostic job titles, and report participant demographics in aggregate where possible (e.g. age), only including information willingly provided by participants. Please refer to Table 1 for participant details. We also report our findings as aggregated themes whereby idiosyncratic details (e.g., software, team structure) are withheld [11].

Our study also has the following limitations. As our participants report their genAI use retrospectively, this may have limited their ability to accurately recall specific prompts or actions. We also encountered a minority of participants ($N=3$) who omitted specific details of their work practice citing company policies. Finally, only interviewing participants in the US means that the international perspective is underexplored, which we look forward to investigating in future work.

4 Findings

We present our findings: how genAI subtly erodes human agency through acting as an *invisible architect* in recruitment decision-making; how the very choice of whether to use genAI systems is not under recruiter control;

<i>ID</i>	<i>Age</i>	<i>Gender</i>	<i>Most Recent Job Title</i>	<i>Exp.</i>	<i>Tools Discussed</i>
<i>P</i> ₁	–	–	Chief Executive Officer	–	ChatGPT, n8n
<i>P</i> ₂	35-44	Man	Sr. Recruitment Specialist	17	Internal system
<i>P</i> ₃	35-44	Man	Director of Recruitment	15	BrightHire, Ashby, Gemini, Gem
<i>P</i> ₄	–	–	Talent Acquisition Specialist	–	Indeed, ChatGPT
<i>P</i> ₅	–	–	Director of Recruitment	–	Metaview, Internal tool, Ashby, Granola, ChatGPT
<i>P</i> ₆	–	–	Chief Operating Officer	–	LinkedIn, ChatGPT
<i>P</i> ₇	35-44	Woman	Director of Recruitment	22	LinkedIn, ChatGPT
<i>P</i> ₈	35-44	Woman	Sr. Recruitment Specialist	10	Undisclosed AI agent service, Greenhouse, Mettl
<i>P</i> ₉	45-54	Man	Director of Recruitment	27	Workday, Gemini
<i>P</i> ₁₀	35-44	Woman	Talent Acquisition Specialist	13	Paraform, LinkedIn, Ashby, Juicebox, Metaview
<i>P</i> ₁₁	25-34	Woman	Sr. Human Resources Manager	8	ChatGPT, LinkedIn, Zoom
<i>P</i> ₁₂	25-34	Man	Talent Acquisition Specialist	5	Willo, Greenhouse, Crisp
<i>P</i> ₁₃	–	–	Talent Acquisition Specialist	–	ChatGPT, Gemini
<i>P</i> ₁₄	35-44	Woman	Talent Acquisition Specialist	14	LinkedIn, Internal LLM, ChatGPT, Google Meet
<i>P</i> ₁₅	35-44	Woman	Sr. Recruitment Specialist	15	ChatGPT, Greenhouse, LinkedIn
<i>P</i> ₁₆	35-44	Woman	Talent Acquisition Specialist	6	ChatGPT, LinkedIn
<i>P</i> ₁₇	25-34	Man	Talent Acquisition Specialist	4	ChatGPT, LinkedIn
<i>P</i> ₁₈	–	–	Talent Acquisition Specialist	–	ChatGPT, LinkedIn
<i>P</i> ₁₉	35-44	Woman	Talent Acquisition Specialist	11	ChatGPT, LinkedIn, Workday
<i>P</i> ₂₀	35-44	Woman	Sr. Recruitment Specialist	15	HireVue, iCIMS, Copilot
<i>P</i> ₂₁	18-24	Woman	Talent Acquisition Specialist	1	Copilot, Internal system, LinkedIn
<i>P</i> ₂₂	18-24	Woman	Talent Acquisition Specialist	2	Bullhorn, ChatGPT, LinkedIn

Table 1. Interview participant demographic information. “Most Recent Job Title” refers to a general class of job titles to retain anonymization. “Tools Discussed” only includes tools that were actively used by the participant, not merely mentioned in their interview.

and how this forced subversion of human agency may lead to bias, recruiter deskilling, and ultimately, little return on investment.

4.1 RQ1: To what extent do recruiters feel they maintain agency in their decision-making?

Recruiters overwhelmingly reported feeling that they alone were in control of their hiring decisions. However, our interviews showed that genAI systems could potentially have more influence over these choices than many participants gave them credit for. We identified that genAI systems were predominantly used to *frame* decisions: our interviewees shared many descriptions of genAI being used to both create assessment schemes of candidate quality, from job descriptions, interview questions, and rubrics. GenAI systems also *processed* the information that went into presenting and evaluating candidates, such as summarizing resumes, interview performance, and more. Thus, our analysis identified that genAI acted as an *invisible architect*—an authority on organizing the framework in which people make decisions, and the information used in decisions, without the necessary visibility that could make participants aware of how their behavior was being shaped. As *P*₃ shared,

“[AI] helps us to really identify the people we’ve been talking to, and boil down to the different patterns that exist... And then using that, it creates, like, an ideal profile, and then from there, you can create a job description. And then it also says, hey, based on this, here is an ideal interview plan. And so it kind of helps you develop the entire operations that way, too, but also make decisions, because you are going to lean on it for some ... Some data, some information, so on and so forth.” (P₃)

Here, P_3 acknowledges that by developing most of the materials used to assess candidates and processing all the information used to go into these assessments, genAI does actually influence the decision-making process itself because you must “*lean on it for... information.*”

4.1.1 Recruiters feel in control and that AI only helps with low-level tasks. More broadly, we observed this dichotomy of perceived control and subtle AI influence through two co-occurring narratives. On the one hand, participants commonly reported feeling in control of hiring decisions and chalked up the work of AI assistance, especially generative AI assistance, to low-level tasks. For example, regarding agency, P_{10} stated that “*I’m the deciding factor, right?*”, and P_9 stated that “*Bottom line, we don’t use AI to make any, like, real decisions in hiring.*” On the other hand, as we demonstrate in Section 4.1.2, a considerable amount of decision-making power was passed on to AI systems via inherent trust of and deference to AI outputs, such as using AI-generated framing and AI-summarized information to make decisions.

Part of participants’ sense of control appeared to rest on framing AI as handling only low-level work: for example, P_{18} stated that the internal generative AI tool in her office that she used to do the initial screening on candidates was “*basically like a filter. You know, it’s filtering it so that it doesn’t waste my time.*” P_{21} stated that the percentage match between a candidate’s resume and job description that she prompts ChatGPT to come up with is “*A helpful tip*”. However, some of the descriptions of AI use had internal contradictions about its role and power: For example, P_5 said their AI interview tool “*Doesn’t make decisions... It’s completely just transcript only*,” despite noting that they used it to organize and highlight information from the interview:

“If... I wanted to know... everything we talked about around compensation, for example, I can actually prompt the tool to tell me everything we talked about that involved a compensation.” (P_5)

Another part of the gap between participants’ assertions that AI does not make final decisions and their deference to it in practice may stem from participants’ belief that AI should not make final decisions about candidates. As P_9 shared, “*I don’t want to make any decisions with AI.*” However, in practice, many recruiters ceded control by rarely questioning or verifying AI output. As P_{18} said about her AI filtering system discussed above, she “*[doesn’t] have to necessarily double-check*” the choices it makes about who it filters out because she trusts the systems’ output. This feeling of control even extended to recruiters using agentic AI systems that completely automate entire sections of the recruiting workflow, such as resume screening: As P_8 states, “*I can almost blindly move those 15% [that an AI agent has approved autonomously]... to talent interview, because I know that they have passed the bot.*” Despite “*blindly*” turning over decisions to the bot, P_8 said “*I still feel I’m in control... because... I am still the responsible one, so the bot is only gonna do what I tell the bot to do.*”

4.1.2 AI as the invisible architect. Despite participants’ descriptions of how little AI and genAI tools influenced their decision process, our interviews demonstrated that genAI systems were invisible architects of the recruitment process. Everything from determining the relevant qualifications for a position, to generating interview questions and rubrics for assessing candidates, to summarizing and highlighting the most relevant pieces of information from an interview to make a final decision, were generated or summarized by an AI system. This fact, plus the common habit of deferring to AI output, means that while the ‘decision’ about whether to select a candidate is often made by a human, genAI, subtly and incrementally, can significantly shape the hiring process.

For example, P_9 shared how genAI systems were essentially responsible for shaping the framework through which recruiters understand what skills the job requires and how they should evaluate and reason about applicants:

“Themeing, interview themes, ... like talking with a hiring manager about a job... We’ll... transcribe that [conversation] and theme it and get all the information so then we can build the concepts for us, the questions for us, the rubrics based upon the job description, the panel interview questions ... We’ll use AI to help, you know, organize the information and then generate things that are useful for us ... so Gemini will transcribe and theme, when we get it ... we don’t have to do too much beyond that” (P_9)

Importantly, P_9 states here that the genAI outputs which frame decision-making don't need "*too much*" done to them, i.e. little editing or oversight, and can be simply used in the recruiting process. Further, despite saying they "*use AI*" to create "*things that help us assess the candidates in the process*," P_9 insists that "*In terms of, like, assessing applications, we don't use it as much as... some other organizations do*"—downplaying AI's influence, thus making it invisible.

Beyond AI designing decision-making frameworks, our interviews also showed that the processing of information relevant to assess candidate qualifications was often also left to AI. Several participants shared how they would use AI summaries to process and highlight information gleaned from candidate interviews, and often these summaries got moved straight along to higher ups. For example, P_{12} stated:

"I use Crisp AI for note-taking in all of my interviews. It does a transcript and an audio recording and has an AI chatbot that you can ask 'summarize all of their answers to the interview questions, and list all the questions they ask me verbatim' ... then I can stick that in there for the hiring manager. (P_{12})"

Again, P_{12} "*sticks*" the AI summaries in the candidate's file for the hiring manager with seemingly little oversight. They went on to explain that while the transcripts exist as a ground truth for what was said during the interview, "*it's very uncommon that someone will actually go and review them*." P_8 seconded this practice, stating

"I kind of tell the bot what... to summarize out of this talk. So, for example, I would say, Summarize... the candidate's experience with enterprise-level clients and then the AI bot would understand that, and I can then fill that into the [candidate's] scorecard." (P_8)

Adding to this pattern, P_{10} states that thanks to the AI summarization, "*I don't have to do anything*" to further process the interview. Evidenced by these examples, AI and genAI systems have huge influence over recruiting decision-making processes by designing assessment frameworks and processing relevant applicant information—but, recruiters do not recognize its influence, despite handing over control to AI.

4.2 RQII: To what extent do recruiters perceive AI adoption as self-directed choice versus a necessity?

Our findings answer this question across three primary sectors of recruiting: the individual-level pressures to be more productive; the business-level pressures to survive; and, finally, how genAI is shaping the field towards pressures that resemble an arms race.

4.2.1 Pressure to Optimize, Adapt, or Perish.

Recruiters justified the adoption of genAI through a complex hierarchy of pressures, ranging from the pursuit of incremental "*small wins*" (P_6) to a perceived existential necessity for professional survival. While many participants initially framed the use of genAI as a personal choice to improve their own efficiency, our data suggest such a choice was often a reaction to a saturated technological landscape where AI is perceived as ubiquitous and non-negotiable. For many, the decision to integrate AI was a response to a psychological nudge. As P_{16} reflected, the integration of genAI had become so normalized that it is often subconscious: "*Yeah, probably, I use it [genAI] more than I realize... because I feel like AI is everywhere... if I'm feeling a little stuck... I'll use that.*"

This "*stuckness*" (P_{16}) in creative work makes AI an easier path to choose, and one of least resistance, leading to an incremental surrender of agency over time. This pressure could also be further compounded by peer comparison; practitioners noted colleagues "*closing*" cases faster, leading to a cost-benefit analysis of time. As P_{18} noted: "*If I can figure out how to use this ethically... It's actually gonna save me a lot of time.*" However, such a drive for efficiency in completing cases faster was not merely a personal preference but is often driven by broader business pressures beyond themselves and their team. Many participants described situations where they framed genAI as a requirement (rather than a choice) for occupational survival. As P_3 shared, "*Every company's gonna be an AI company... if you're not using it, you're gonna die*," with P_1 also echoing this sentiment, noting that avoiding AI could immediately make their company "*less competitive*." Such pressures were also codified by top-down

goals described by our participants, irrespective of the organization, that mold personal experimentation with genAI into compulsory corporate mandates. P_{15} reflected on the shift toward embedding AI in annual objectives:

“One of our corporate goals for next year is going to be utilizing AI to be more productive ... it’s including all departments ... [leader] has assigned each of us ... to look at how we can be utilizing AI more.” (P_{15})

As such, leadership also acts as a motivator for adoption, mandating that staff find ways to justify the technology within existing workflows. Such a top-down directive was often simplified in the daily experience of the recruiter, with P_{20} succinctly noting: *“The company is asking us to use it [genAI]”*. Nevertheless, such findings indicate a tension between the marketing promise of genAI—that it will seamlessly increase productivity—and the actual promised labor that organizations have to provide to implement it effectively. Interviewees shared that in order to fill such corporate mandates (*“use more AI,”* P_1), some organizations were investing heavily in onboarding and training, two areas that our practitioners mentioned were resource heavy. As P_8 shares, *“success”* with using such tools necessitated an adjustment of a rigorous training regime in onboarding new employees:

“That’s why we’ve decided to adjust and really train everybody that joins our team on this system, so that they know how to use the bots, because we have seen great success.” (P_8)

While we cannot comment if P_8 ’s adjustment is or is not positive, our interviewees pointed to the significant human labor required to *“make genAI behave”* (P_{13}) in ways that were useful. Furthermore, we identified that this could have an interesting ripple effect across team collaboration, where the perceived lack of being provided access to genAI was seen as putting their teams at a disadvantage. P_1 reflected on a situation where employees on other teams were actively pushing for more access to genAI, even when the existing infrastructure around training or the justification for them gaining access to it was incomplete or unclear: *“... they [a coworker] asked me last week, like, when is this [enterprise ChatGPT account] going live? Like, we need it... and I have to ask, do you?”*

4.2.2 Field-Driven Escalatory Pressures. An enormous factor that underlined all decisions around agency revealed a cycle of sociotechnical escalation whereby both job seekers and hiring practitioners felt compelled to use genAI based on the behavior of the other. We characterized this by a feedback loop: as more candidates make use of genAI to optimize their applications, thus increasing the volume of applications for recruiters to review, recruiters increasingly turn to ‘gut check’ counter-measures to weed out perceived inorganic and AI-manipulated applications, as well as to reclaim their own agency in decision-making.

Interviewees cited their first adoption of genAI came in response to the perceived increased ease for applicants to apply for specific job roles. Practitioners noted that genAI tooling enabled candidates to bypass a traditional ‘self-filtering’ process, where a human may typically assess their own eligibility before applying. This resulted in a surge of ‘noise’ in the recruitment pipeline, as P_{14} reflected when asked to reflect on genAI and applicant behavior:

“People are using AI to, you know, apply to some postings. Because [as] an individual [applicant], if he’s reviewing it, he’ll see that, okay, I might not be eligible for it, so no point of applying... It’s completely irrelevant ... Why [would] a salesperson apply to a content editor position?!” (P_{14})

Practitioners cited that this flood of irrelevant data was an incentive for them to rely on automated filtering systems to maintain a level of control. However, such a use of AI-to-AI systems creates a paradox as noted by many of our interviewees: as more candidates are perceived to use genAI to bypass filters, the more recruiters shared feeling the need to use genAI to filter them out. This was best evidenced in how recruiters believed that candidates used AI to ‘mirror’ job descriptions with such precision that it triggered suspicion rather than capturing attention. Thus, the perfect resume for the role, once the self-cited goal of recruitment, became a *“yellow flag”* (P_{22}) for inauthenticity. P_{18} describes this reversal of trust, from being excited at a perfect candidate to being suspicious:

“There’s no such thing as a perfect candidate, so when you see a perfect candidate, you go, okay, I’m gonna have to really dig into their experience ... they’ll drop their resume into ... any type of AI tool and say, okay, here’s the job description, like, basically fix my resume for it to present as this.” (P₁₈)

When the data points offered by a candidate felt *“too polished”* (P₁₉), recruiters reported perceiving them as *“a nice way to have something that sounds very polished ... but it’s just so vapid”* (P₁₆) and lacking in genuine substance. P₁₃ shared that this creates a state of noise and distrust where *“exact resume matches”* were more likely to be treated with skepticism or suspicion because they often lacked the nuances around how they perceived applicants to pursue careers. To reclaim such agency in a system that they perceive is being manipulated, recruiters described using ad-hoc detection methods to identify genAI involvement. This was a shared strategy, where recruiters look for subtle linguistic or formatting cues that signal a perceived lack of human effort:

“Whenever candidates have those bolded words, I’m like, okay, that’s another, like, yellow flag... Is it a real candidate, like, tweaking their own resume? Is it something, like, an AI bot?” (P₂₂)

Thus, in reaction to the application overflow and the manipulation of previously reliable information, recruiters returned to *“gut instinct”* (P₂₀), potentially leading to the dismissal of candidates due to perceived genAI use.

4.3 RQIII: Implications of Changes to Decision-making Process and Boiling the Ocean

Our findings present a complex picture of genAI adoption and its impacts on recruiter agency in the realm of recruiting. While many of our participants described narratives that genAI provided a unique *“secret thing”* (P₇) aiding efficiency and data-driven precision to find better candidates, the lived reality of practitioners suggests the presence of what we identify as a *boiling the ocean* effect. We observe this via three mechanisms. First, despite hopes for more objective and data-based decision-making, the use of genAI tools elevates the risk of bias through recruiters reacting to data overload by reclaiming agency through gut and culture-fit based hiring decisions. Second, despite the hopes of increasing recruiter skills through genAI use, we observed evidence of recruiter deskilling, threatening the maintenance of human agency in hiring in the long term. Finally, while AI use is marketed as a time-saver, recruiter reliance on AI tools, despite their reportedly sub-par output, can lead to more work or lower quality outputs overall. This, plus the increased use of AI by candidates increasing workload, suggests that AI use may not be delivering on its promises, as recruiters are ending up with lower-quality candidates, as evidenced by reported higher employee turnover.

4.3.1 Gut- and Culture-fit based Decision-Making. Much prior work has noted that while AI often promises efficient and data-driven, and therefore unbiased, decision-making, automated systems can often further entrench bias [2, 62, 78]. Our findings here add to this conversation by showing that, while participants did believe that using AI tools would reduce bias, the use of genAI tools in recruitment introduced a less-discussed potential mechanism of bias: a retreat to gut-based decision-making that often relies heavily on assessing “culture fit” [25].

Perception of Reduced Bias. Many participants clearly stated that they believed AI tools were a method of reducing bias in recruiting, as P₉ stated that he wanted to get *“past the biases of my recruiters.”* He continued:

“If you take bias out, you may identify a bunch of candidates that we hadn’t thought about previously, who may be able to deliver on that job... I’m really interested in how can I use AI to identify those people.” (P₉)

Some participants also shared examples of when using AI systems helped them become less biased decision-makers. P₁ shared a story of how she did not want to hire a certain applicant because someone on her team already knew the person. But, *“when she was one of the, like, the shortlisted [applicants] by ChatGPT,”* and P₁ thought her performance was *“very good”* on take-home assessments, P₁ hired her—she concluded that it was *“very important”* that ChatGPT was *“not biased.”* While there certainly may be scenarios where the use of AI tools

reduces bias, despite recruiters' good intentions, we observed a different mechanism whereby AI and genAI systems may be increasing bias in recruiting.

Gut- and Culture-fit based Decision-Making. Due to the information overload that comes as a result of the onslaught of ever-rising numbers of AI-polished applications with little signal, and a need to remain useful as employees beyond AI tools, recruiters often expressed that one of the only meaningful choices they can make regarding candidates was based on gut instinct to assess what they believed AI could not replace: an understanding of business culture fit.

Many recruiters spoke directly and indirectly about a retreat from data-driven decision-making. Recruiters no longer trust digital artefacts (e.g., resumes, cover letters) that may have been fabricated with genAI (Section 4.2.2); many participants spoke to reverting back to their “*gut instinct*” (P_{20}) and a reliance on “*vibe checks*” (P_8) as the final sources of truth in the recruiting pipeline. Several recruiters spoke to their belief that what made them irreplaceable in a world of AI-based hiring was understanding business culture fit, which many have argued is a mechanism for unintentional bias in hiring [25, 63, 67]. While undoubtedly solely human-based hiring decisions also include reliance on culture fit, this increased emphasis on it as a reaction to AI may increase bias. P_{18} stated that while she worried about “*At what point am I still going to be relevant, right? At what point are maybe interviews going to be taken over by an avatar completely, you know?*” she pushed against the spectre of being replaceable by “*Learning how to ask more questions around, like, the human experience, which is what actually places people at the end of the day. And that’s what AI can’t do.*” P_{15} concurred:

“So much of it is, about our specific culture...I know sort of the microculture of this team, ... So I know, kind of, what’s going to be a fit and what isn’t, and AI doesn’t know that.” (P_{15})

Such a reliance on a recruiter’s gut assessment of culture fit to combat both the fear of being replaced by AI and cut through perceived AI manipulation on the candidate side, may inadvertently reintroduce the same human biases that many participants believed genAI systems could screen out. P_{16} acknowledges this tension outright, highlighting the intrinsically messy reality of reclaiming agency in a new recruitment landscape:

“What’s my feel on this candidate? Does it seem like they’re aligned with what the manager’s looking for? That’s not a... like, AI’s not gonna do that... So, there’s so much room for bias, right? It’s like, my bias is so imbued in all of our hiring and I have mixed feelings about that.” (P_{16})

4.3.2 Deskilling. Another common consequence of increased reliance on genAI that our interviews surfaced was evidence of recruiter deskilling around core elements of applicant evaluation. Importantly, however, there were also accounts of AI systems being used to acquire new skills.

GenAI used to skill up. Some participants hypothesized that although AI may lead to some currently used skills being underdeveloped, it may lead to the acquisition of others, for example “*[being] more creative*” (P_{10}). P_{12} describes this succinctly:

“I think maybe people that... just got into recruitment, like, this year, I doubt... [they] take as many notes... But they might have an entirely new book of skills at the same time that, you know, people who grew up before it [AI] didn’t.” (P_{12})

This hypothesis was supported by several participants, who described that genAI systems empowered them with efficiently learning about their hiring contexts. As P_{18} shared, “*there’s so many new technologies coming out*” relevant to jobs she places:

“If I see, like, a job description, I also drop that into AI, ... okay, break this down in layman’s terms, like, who am I looking for, candidate-wise? ... they’ll really break down what is this technology.” (P_{18})

P_{18} continued that this research support helped her sound more educated on the area she was hiring in, which helped forge stronger social connections with candidates, reflecting that genAI tools help to “*communicate things*,

because people [applicants] will hang up on you if ... you mispronounce technologies.” Some participants, such as P₁₅, even reported stories of directly using ChatGPT to skill up on the job:

“ChatGPT has helped me a ton with learning Excel ... it’ll write me a super complicated formula, but then also tell me why it doesn’t work. I am actually learning, even though it’s doing a lot of the work for me. Now I can do a VLOOKUP by myself.” (P₁₅)

Evidence of Deskilling. Even when AI was used to empower recruiters with new information or save time, it often simultaneously deskilled workers. Beyond simply atrophying skills, we also highlight deskilling as a lack of intentionality in the choice of tasks some participants assign to genAI, a skill in and of itself.

Sometimes deskilling was subtle, in a slight resistance to doing tasks which were previously a part of day-to-day job operations; other times, especially among more junior recruiters, the loss of agency seemed to result in a fear of or perceived inability to do parts of the job. Many participants described, alongside the ease of using AI, a reluctance or difficulty with performing tasks manually: as P₁ stated regarding reviewing resumes by hand, *“I’m not going to read that. It’s like you’re reading a book.”* P₄ found that a lot of resumes are *“Pretty wordy, or a lot of information,”* so it was *“nice to use that tool to summarize.”* This reluctance persisted even if they acknowledged the product would be better or they would learn more. Both P₁₇ and P₁₈ talked about using ChatGPT to research their candidates’ area of expertise and develop interview questions, but shared they were conscious of depriving themselves of learning skills by using genAI systems to research:

“I would have so much more knowledge if I sat down and had someone do a training on this technology instead of me just dropping it in [to AI] and having it populate screening questions, because really, I’m just reading from a prompter, and at that point, you might as well put me out of a job, you know?” (P₁₈)

Other junior participants expressed fear of being unable to perform aspects of their job without AI assistance. For example, P₂₁ shared that she was working on developing capabilities independent of ChatGPT, in case she ever lost access to her account:

“I’m just trying to get better about using it, and just, like, being more independent on it. ... I’m sure there’ll become a day where ...I don’t have access to it for X amount of time. I don’t know what that would look like, but, like, I don’t ever want to...be in a position where...I cannot do my job without ChatGPT.” (P₂₁)

P₁₃ stated that they like using genAI for aspects of communication they could not figure out on their own: *“I like using it for... sentences where I’m like, this makes no sense, and my brain is broken, but, like, I know I need to communicate this.”* More senior recruiters also expressed frustrations with employee errors that were introduced by not checking over ChatGPT’s outputs, suggesting negligence to the point of deskilling: P₉ shared where his team directly used an unreviewed AI-generated candidate evaluation rubric and jeopardized the quality of the hiring process through inappropriate candidate rankings:

“They’re like... ‘should we have asked questions about it [the rubric item]?’ Like, no, that shouldn’t be in there. Right? And so, like, and that’s the hard thing, like, you know, it wasn’t the first question, it wasn’t the last question, it was, you know, it was the rubric item number, you know, seven. So, [they] didn’t notice it... so that’s the kind of stuff that I really want them looking at.” (P₉)

We identified that when AI systems structure how decisions are made—such as by introducing changes to evaluation rubrics that participants “didn’t notice” (P₉)—they can reduce opportunities for deliberate human reflection on those decisions. This, in turn, can erode the professional skill of intentionally deciding when to delegate to AI versus make a decision independently. Though some recruiters were very cognizant of what tasks they believed should and should not be allocated to genAI, others were not. For example, P₁₀ had a strong reaction to using AI in screening, saying that *“I don’t trust it at all in that aspect,”* while P₁₄ maintained that there should be *“no AI”* in the interview process. On the other hand, when asked about whether there were tasks that should not be allocated to genAI, P₂₁ said *“I don’t really think there’s anything ... Yet, I’m sure there is something I probably*

shouldn't be doing with it, but I cannot think of it." This absence of a clear boundary can make it easier for genAI to creep into more aspects of the hiring pipeline without clear reasoning.

For participants who claimed productivity boosts (Section 4.3.3) from AI use, they had trouble justifying their use of genAI outside of those gains. This is clear in the thinking of P_9 : *"I think any efficiencies we can create using AI, we try to use,"* where efficiency comes first before the consequences of using it *"in all [of] the supporting functions around the picking, and then just a little bit in the picking"* (P_9). This gap in clear intentionality in when to refer to AI in the decision-making process is echoed in P_{18} 's call for *"rules and boundaries"* around AI-use.

In the long term, such deskilling may erode the continued possibility of meaningful human agency over AI-assisted decisions and the ability to repair a pipeline that overrelies on AI.

4.3.3 (Lack of) Return on Investment. Participants cited improved efficiency or profitability as their primary justification for genAI adoption, particularly weighed against the risk of falling behind competitors who did. This can be understood as a stand-in for a rough social calculation of return on investment (ROI); participants regularly cited the marketing promise of genAI—the ability to transform organizations and reap massive efficiency gains in daily work. However, we identified that despite the investment in AI and the change in day-to-day operations, participants often described how genAI seemingly created as much labor as it purported to save, and ultimately did not result in higher quality candidates—ostensibly, the entire purpose of recruiting.

Hidden Labor of AI. Many participants described the ways in which AI saved time in their daily work. These time-saving gains were non-negligible, where some claimed that AI saved *"60% of your day"* (P_5) and helped them *"massively"* (P_8). However, while participants shared that AI tools saved time in the short run, they also gave indications that the use of AI may create more work overall. As a basic matter, participants expressed frustration with the administrative and sometimes cognitive overhead required to make AI useful. For example, P_{21} shared that the level of precision needed for a successful ChatGPT output often rendered using the tool as an inefficient use of time: *"I'm just wasting my time... typing up this whole two-paragraph-long thing to then... I would have read that and figured it out myself at this point."*

Further, participants shared that often, AI created worse outputs than they could on their own, but they would often use it anyway, potentially leading to low-quality work with implications downstream. For example, P_{10} reflected that despite using genAI on the initial volume of applicants, *"manual Boolean searches"* were still required to ensure quality: *"I still need to do it myself to produce the highest amount of quality [candidates]"* (P_{10}). However, crucially, many participants did not invest the manual work to fix AI outputs: when P_{19} talked about her frustration with generating boolean searches, stating that sometimes when using ChatGPT she feels *"this isn't working."* Yet when asked if she would try to fix it on her own or continue to use the tool, P_{19} stated, *"some of both,"* but often *"I just try it again later"* with the tool. P_{18} shared, *"Sometimes it's helpful, sometimes it's not, but I find myself doing it [using genAI] all the time."*

Beyond this loop of using AI tools even if they were unhelpful, the automated nature of these tools leads, in some cases, to professional embarrassment or *"bitten-in-the-ass"* moments (P_{17}) where recruiters rely on saved templates or AI-written messages without a cursory review before posting or sending. Beyond the examples of this discussed above in deskilling (Section 4.3.2), whereby workers that move too fast lose the ability to catch these errors, leading to problems downstream, agency recruiters also spoke of their colleagues *"cutting corners"* with AI to meet their numbers in the short-term to the detriment of long-term gains:

"I've also seen, you know, recruiters cutting corners... they're like, oh, I can just send this candidate over, but it backfires, because you end up using AI editing their resume to make them look more presentable ... and ... when they show up in an interview, it doesn't present the same way, right?" (P_{18})

As P_{18} describes, though editing resumes with AI to increase perceived job fit may improve short-term efficiency for recruiters incentivized to maximize candidate-job matches, it creates downstream inefficiencies for hiring managers, or can even lead to hiring an unqualified employee, potentially incurring high costs long-term.

Finally, as discussed in Section 4.2.2, candidate use of AI created an environment for many of our participants where they had to shift the focus of their agency from “*finding talent*” to “*detecting fraud*” (P_{18}). As P_7 shared, this can actually result in the advent of AI being a time suck overall, as recruiters have more and more irrelevant applications to scan: “[*it’s*] just such a time suck. Like, trying to investigate if someone is real.”

Low perceived ROI. While participants were keen to emphasize that genAI may “*help the day-to-day a little bit*” (P_{18}), this observation was coupled with the reflection that it had failed to deliver the perceived “*insanely beneficial*” (P_{18}) transformation at the organizational or field level. Many participants noted that the core metrics of what was considered success in recruiting (e.g., placement, sales) have not seen a corresponding leap forward:

“No one’s [case] numbers have gone through the roof, and... no company has, like, quadrupled in their placements and sales based off of it... I don’t think it’s moving things faster.” (P_{18})

Overall, between recruiters being forced to use AI, leading them to eventually lose agency or cut corners, and candidates using AI to generate falsely “perfect” profiles, in the end, recruiters painted a world where they are doing more work daily—going through more profiles—but ending up with worse outcomes overall, i.e. not finding high quality candidates. As P_{14} shares, while “*every day these ChatGPTs, they are applying to 10 jobs*” on applicants’ behalf, increasing the load of applications, P_{18} explains that even with recruiter use of AI in response, this leads to more work in screening, and worse candidates overall:

“I’m more screening more now, are you actually going to be able to show up and do the job and not just look like it on paper? ... That’s where we’re having a lot of turnover, because they’re using AI too much.” (P_{18})

Thus, AI’s efficiency gains are eroded. Instead of finding higher quality candidates—the goal of introducing AI—participants report a greater chance of hiring someone who must be replaced, “*wasting everyone’s time*” (P_{18}).

5 Discussion and Concluding Thoughts

Our findings highlight that genAI systems have become “invisible architects” of the hiring process—they a large influence on hiring outcomes through generating decision-making frameworks (rubrics, job descriptions) and processing the data that goes into these frameworks (interviews, resumes), yet this influence is nearly imperceptible as each piece is seen as low-level work (Section 4.1.2). Further, we find that while genAI offered participants some perceivable productivity benefits in tasks such as information gathering (Section 4.1), many reported struggling under the pressure to use such tools (Section 4.2), even when participants deemed such use unsuitable. GenAI’s touted benefits may also be leading to a boiling the ocean effect (Section 4.3), whereby the efforts invested to make genAI ‘work’ for their workflow was often described as contributing more work, unintentional de-skilling, or shaping subpar outputs that needed correcting. We believe these discoveries have implications for AI policy, organizational practice, and areas where further research is needed.

Intentional Decision-Making Policies. Much law surrounding AI accountability focuses on decision-making outputs: for example, in the EU, the GDPR emphasizes the right not to be subject to *decisions “based solely on automated processing”* [33, 82]. However, our findings show that the introduction of genAI systems to high-stakes decision-making pipelines significantly broadens the sphere of influence that AI systems have over these decisions. Generative AI systems create the entire *infrastructure* (see [60, 69]) within which that decision rests: organizational policy documents, rubrics, in addition to *processing* the information that is considered, which may not be perceivable to its users. This creates a tension: recruiters, and other AI-assisted decision-makers, can claim that a final decision was human-authored, even as the decision space itself has been extensively structured by AI.

As a result, legal protections around automated decisions may fail to capture the ways in which agency is altered by genAI-assisted processes. Whether or not this is a problem is an open question, however, our work shows that if regulation aims to limit the extent to which AI makes high-stakes decisions affecting individuals' lives, our current regulatory frameworks may be insufficient.

Relatedly, increased integration of human-AI decision-making may make it more difficult to comply with existing laws regulating high-stakes decisions in practice. For example, in the U.S., Title VII prevents disparities in employment-related assessment outputs that cannot be justified by “business necessity” [32]. In addition, several new state laws require mitigation of disparities of AI models used in high-stakes decisions such as hiring [6, 7]. Thus, AI systems to be used for hiring and other high-stakes decisions need to be tested for bias before deployment—but when human and AI decision-making is so intertwined, it is hard to develop tests that accurately portray how bias may arise during deployment. More generally, the suite of current methodologies for discrimination testing, and other forms of legal compliance, are ill-suited to the level of human-AI collaboration happening in high-stakes decisions, creating a new set of problems for researchers to tackle.

Organizational Practices to Combat Deskilling. Many resources for employees in the workforce are focused on prompt engineering: how to prompt genAI better to produce improved results, promoting an implicit encouragement that genAI adoption is both legitimate and essential [44, 70]. This is a tempting response to presumed participant need; if recruiters just figure out how to use genAI, then this will result in time being saved (Section 4.2.2). Our findings suggest that organizations should also focus on critical oversight of genAI use to prevent the rapid professional deskilling that may arise when recruiters “*move too fast*” (P_{18} , Section 4.3.3) and question the role of genAI in occupational survival (P_3 , Section 4.2.1). To promote the opportunity to learn and develop intentionality, despite the allure of summarization and saving minutes but losing hours, we suggest that future onboarding should incorporate a form of intentional friction to any genAI use. For example, pedagogical interventions that force recruiting professionals to articulate their professional logic behind any gut instinct (P_{16} , Section 4.3.1) could hold promise. Such suggestions mirror Bućinca et al.’s [18] recommendations to incorporate disruptive design elements that force a user to explain their own professional logic, prior to seeing automated content. By forcing recruiters to explain why they have a gut feeling about a candidate or process, this helps to create a reflection-in-action moment, moving from simply checking over genAI work to an independent evaluator. Should such an intervention be incorporated into a workflow, fostering an agency maintenance evaluation metric—such as explaining why certain candidates are a good fit and rewarding meaningful checks—could lead to incentivizing more intentional use of AI assistance, and reduce the long-term costs of poor candidate matching.

Outstanding Research Directions. Several papers have pointed to the fact that genAI systems have fundamentally changed the way that bias testing and mitigation [55] functions in AI systems, and call into question whether it is possible [3]. First, our work shows an empirical observation of a more technical point made by many [82, 86]: How does one mitigate bias in a textual output with an unclear mapping to the final decision? In the case of recruitment, how does one understand the possibility of AI bias in a generated rubric, or set of interview questions? More generally, as researchers, this points to a growing need to build tools to understand the impacts of intermediate model outputs in a decision-making process, not just those providing a concrete recommendation. Our work also highlights a sobering potential outcome: AI bias is not only becoming harder to find because of the structure of its output, but due to incremental over-reliance, practitioners may be less and less equipped to meaningfully watch for signs of bias (Section 4.3.2). This leads to a natural question: How can we help practitioners retain, or achieve, awareness of AI’s influence as its role becomes more and more subtle and intertwined in every step of the human thought process towards decision-making?

Generative AI Usage Statement

The authors of this paper did not use any generative AI assistance when writing this manuscript.

Acknowledgments

First, the authors would like to thank their participants for sharing their time and insights. Additionally, they thank Stevie Chancellor and the reviewers for their generous feedback, which helped to improve this final manuscript.

Positionality Statement

The authors acknowledge their experiences as researchers in relation to the topic of recruitment. As a researcher within the field of Security & Privacy, Rosanna Bellini has experience in analyzing public tools and the effects of their usage. Emily Black, whose background is in Responsible AI, has experience in auditing AI tools, as well as analyzing the requirements around lawful use of such tools within the US regulatory framework. Though the authors have had minimal exposure to industry recruitment themselves, this work focuses on and presents the experience of the participants, who are all recruiting or HR professionals.

References

- [1] Haozhe An, Christabel Acquaye, Colin Wang, Zongxia Li, and Rachel Rudinger. 2024. Do Large Language Models Discriminate in Hiring Decisions on the Basis of Race, Ethnicity, and Gender? *arXiv preprint arXiv:2406.10486* (2024).
- [2] Jiafu An, Difang Huang, Chen Lin, and Mingzhu Tai. 2025. Measuring gender and racial biases in large language models: Intersectional evidence from automated resume evaluation. *PNAS nexus* 4, 3 (2025), pgaf089.
- [3] Jacy Reese Anthis, Kristian Lum, Michael Ekstrand, Avi Feller, and Chenhao Tan. 2025. The impossibility of fair LLMs. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 105–120.
- [4] Lena Armstrong, Abbey Liu, Stephen MacNeil, and Danaë Metaxa. 2024. The Silicon Ceiling: Auditing GPT’s Race and Gender Biases in Hiring. In *Proceedings of the 4th ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (San Luis Potosi, Mexico) (EAAMO ’24). Association for Computing Machinery, New York, NY, USA, Article 2, 18 pages. doi:10.1145/3689904.3694699
- [5] Ashby. 2025. Ashby. <https://www.ashbyhq.com/>
- [6] Colorado General Assembly. 2024. Consumer Protections for Artificial Intelligence (SB24-205).
- [7] Illinois General Assembly. 2024. Artificial Intelligence in Employment (Public Act 103-0804).
- [8] Avature. 2025. Applicant Tracking System. <https://www.avature.net/applicant-tracking-system/>
- [9] Alessio Azzutti. 2024. Artificial Intelligence and Machine Learning in Finance: Key Concepts, Applications, and Regulatory Considerations. In *The Emerald Handbook of Fintech: Reshaping Finance*. Emerald Publishing Limited, 315–339.
- [10] Tina Behzad, Siddhartha Devic, Vatsal Sharan, Aleksandra Korolova, and David Kempe. 2025. An External Fairness Evaluation of LinkedIn Talent Search. arXiv:2511.10752 [cs.CY] <https://arxiv.org/abs/2511.10752>
- [11] Rosanna Bellini, Emily Tseng, Noel Warford, Alaa Daffalla, Tara Matthews, Sunny Consolvo, Jill Palzkill Woelfer, Patrick Gage Kelley, Michelle L. Mazurek, Dana Cuomo, Nicola Dell, and Thomas Ristenpart. 2024. SoK: Safer Digital-Safety Research Involving At-Risk Users. In *2024 IEEE Symposium on Security and Privacy (SP)*. 635–654. doi:10.1109/SP54263.2024.00071
- [12] Marianne Bertrand and Sendhil Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American economic review* 94, 4 (2004), 991–1013.
- [13] Miranda Bogen and Aaron Rieke. 2018. Help wanted: An examination of hiring algorithms, equity, and bias. *Upturn*, December 7 (2018).
- [14] Virginia Braun and Victoria Clarke. 2021. Thematic analysis: A practical guide. (2021).
- [15] Breezy. 2025. Breezy. <https://breezy.hr/>
- [16] BrightHire. 2025. Equitable Hiring. <https://brighthouse.com/solutions/equitable-hiring/>
- [17] BrightHire. 2025. BrightHire. <https://brighthouse.com/>
- [18] Zana Bućinca, Maja Barbara Malaya, and Krzysztof Z. Gajos. 2021. To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-assisted Decision-making. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 188 (April 2021), 21 pages. doi:10.1145/3449287
- [19] Bullhorn. 2025. Bullhorn. <https://www.bullhorn.com/>
- [20] Shiye Cao and Chien-Ming Huang. 2022. Understanding User Reliance on AI in Assisted Decision-Making. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 471 (Nov. 2022), 23 pages. doi:10.1145/3555572
- [21] S Catherine, NV Suresh, T Mangaiyarkarasi, and Leena Jenefa. 2025. Unveiling the Enigma of Shadow: Ethical Difficulties in the Field of AI. In *Navigating Data Science: Unleashing the Creative Potential of Artificial Intelligence*. Emerald Publishing Limited, 57–67.
- [22] Challapally, Aditya, Pease, Chris, Raskar, Ramesh, and Chari, Pradyumna. 2025. *The GenAI Divide*. Technical Report. MIT.

- [23] Myra Cheng, Sunny Yu, Cinoo Lee, Pranav Khadpe, Lujain Ibrahim, and Dan Jurafsky. 2025. Social Sycophancy: A Broader Understanding of LLM Sycophancy. doi:10.48550/arXiv.2505.13995 arXiv:2505.13995 [cs].
- [24] Avishek Choudhury and Zaira Chaudhry. 2024. Large language models and user trust: consequence of self-referential learning loop and the deskilling of health care professionals. *Journal of Medical Internet Research* 26 (2024), e56764.
- [25] Phoebe K Chua and Melissa Mazmanian. 2020. Are you one of us? Current hiring practices suggest the potential for class biases in large tech companies. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–20.
- [26] Sasha Costanza-Chock, Inioluwa Deborah Raji, and Joy Buolamwini. 2022. Who Audits the Auditors? Recommendations from a field scan of the algorithmic auditing ecosystem. In *Proceedings of the 2022 ACM conference on fairness, accountability, and transparency*. 1571–1583.
- [27] Covey. 2024. Covey. <https://covey.framer.website/>
- [28] Maria De-Arteaga, Riccardo Fogliato, and Alexandra Chouldechova. 2020. A Case for Humans-in-the-Loop: Decisions in the Presence of Erroneous Algorithmic Scores. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3313831.3376638
- [29] Online Etymology Dictionary. 2026. Origin and history of agency. <https://www.etymonline.com/word/agency>
- [30] Colin Duncan and Wendy Loretto. 2004. Never the right age? Gender and age-based discrimination in employment. *Gender, Work & Organization* 11, 1 (2004), 95–115.
- [31] Eightfold. 2025. AI talent acquisition & recruiting platform. <https://eightfold.ai/>
- [32] Equal Employment Opportunity Commission; U.S. Department of Labor; U.S. Department of Justice; U.S. Office of Personnel Management; U.S. Department of the Treasury. 1978. Uniform Guidelines on Employee Selection Procedures. 43 Fed. Reg. 38295 (Aug. 25, 1978); codified at 29 C.F.R. Part 1607. Guidance establishing uniform federal standards for employment testing and selection procedures to prohibit discrimination on the basis of race, color, religion, sex, or national origin; adopted by EEOC and related agencies..
- [33] European Parliament and Council of the European Union. 2016. General Data Protection Regulation. *Official Journal of the European Union* L 119 (2016), 1–88. Article 22.
- [34] John C Flanagan. 1954. The Critical Incident Technique. (1954), 1–33.
- [35] Gem. 2025. Gem. <https://www.gem.com/>
- [36] Kate Glazko, Yusuf Mohammed, Ben Kosa, Venkatesh Potluri, and Jennifer Mankoff. 2024. Identifying and Improving Disability Bias in GPT-Based Resume Screening. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (Rio de Janeiro, Brazil) (FAccT '24). Association for Computing Machinery, New York, NY, USA, 687–700. doi:10.1145/3630106.3658933
- [37] ATLAS.ti Scientific Software Development GmbH. 2025. ATLAS.ti Mac. <https://atlasti.com>
- [38] Ben Green and Yiling Chen. 2019. Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments. In *Proceedings of the conference on fairness, accountability, and transparency*. 90–99.
- [39] Ben Green and Yiling Chen. 2019. The Principles and Limits of Algorithm-in-the-Loop Decision Making. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 50 (Nov. 2019), 24 pages. doi:10.1145/3359152
- [40] Ben Green and Yiling Chen. 2021. Algorithmic Risk Assessments Can Alter Human Decision-Making Processes in High-Stakes Government Contexts. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 418 (Oct. 2021), 33 pages. doi:10.1145/3479562
- [41] Greenhouse. 2025. Greenhouse. <https://www.greenhouse.com>
- [42] Hirevue. 2026. Our Science. <https://www.hirevue.com/our-science>
- [43] Hirevue. 2025. HireVue. <https://www.hirevue.com/>
- [44] IBM. 2026. Generative AI for Human Resources (HR) Professionals. <https://www.coursera.org/specializations/generative-ai-for-human-resources/>
- [45] Lujain Ibrahim, Katherine M Collins, Sunnie SY Kim, Anka Reuel, Max Lamparth, Kevin Feng, Lama Ahmad, Prajna Soni, Alia El Kattan, Merlin Stein, et al. 2025. Measuring and mitigating overreliance is necessary for building human-compatible AI. *arXiv preprint arXiv:2509.08010* (2025).
- [46] iCIMS. 2025. iCIMS. <https://www.icims.com/>
- [47] Indeed. 2025. Indeed. <https://www.indeed.com/hire>
- [48] Juicebox. 2025. Juicebox (PeopleGPT) - The Leading AI Recruiting Platform. <https://juicebox.ai/>
- [49] Patrick Kline, Evan K Rose, and Christopher R Walters. 2022. Systemic discrimination among large US employers. *The Quarterly Journal of Economics* 137, 4 (2022), 1963–2036.
- [50] Klaus Krippendorff. 2011. Computing Krippendorff’s alpha-reliability. (2011).
- [51] Olya Kudina and Bas de Boer. 2025. Large language models, politics, and the functionalization of language. *AI and Ethics* 5, 3 (2025), 2367–2379.
- [52] Alisa Küper and Nicole Krämer. 2025. Psychological traits and appropriate reliance: Factors shaping trust in AI. *International Journal of Human-Computer Interaction* 41, 7 (2025), 4115–4131.
- [53] Vivian Lai, Chacha Chen, Q Vera Liao, Alison Smith-Renner, and Chenhao Tan. 2021. Towards a science of human-ai decision making: a survey of empirical studies. *arXiv preprint arXiv:2112.11471* (2021).

- [54] Lan Li, Tina Lassiter, Joohee Oh, and Min Kyung Lee. 2021. Algorithmic Hiring in Practice: Recruiter and HR Professional’s Perspectives on AI Use in Hiring. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) (AIES ’21). Association for Computing Machinery, New York, NY, USA, 166–176. doi:10.1145/3461702.3462531
- [55] Yingji Li, Mengnan Du, Rui Song, Xin Wang, and Ying Wang. 2023. A survey on fairness in large language models. *arXiv preprint arXiv:2308.10149* (2023).
- [56] LinkedIn. 2025. *The Future of Recruiting 2025*. Technical Report. <https://business.linkedin.com/talent-solutions/resources/future-of-recruiting>
- [57] Takuya Maeda and Anabel Quan-Haase. 2024. When Human-AI Interactions Become Parasocial: Agency and Anthropomorphism in Affective Design. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (FAccT ’24). Association for Computing Machinery, New York, NY, USA, 1068–1077. doi:10.1145/3630106.3658956
- [58] Lars Malmqvist. 2025. Sycophancy in Large Language Models: Causes and Mitigations. In *Intelligent Computing*, Kohei Arai (Ed.). Springer Nature Switzerland, Cham, 61–74. doi:10.1007/978-3-031-92611-2_5
- [59] Nestor Maslej, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki, Emily Capstick, Anka Reuel, Erik Brynjolfsson, John Etchemendy, et al. 2025. Artificial intelligence index report 2025. *arXiv preprint arXiv:2504.07139* (2025).
- [60] Stephanie Mertens, Mario Herberz, Ulf J. J. Hahnel, and Tobias Brosch. 2022. The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences* 119, 1 (Jan. 2022), e2107346118. doi:10.1073/pnas.2107346118 Publisher: Proceedings of the National Academy of Sciences.
- [61] Metaview. 2025. Metaview. <https://www.metaview.ai/>
- [62] Huy Nghiem, John Prindle, Jieyu Zhao, and Hal Daumé Iii. 2024. "You Gotta be a Doctor, Lin": An Investigation of Name-Based Bias of Large Language Models in Employment Recommendations. *arXiv preprint arXiv:2406.12232* (2024).
- [63] Bethany J Nichols, David S Pedulla, and Jeff T Sheng. 2025. More than a match: “Fit” as a tool in hiring decisions. *Work and Occupations* 52, 2 (2025), 175–203.
- [64] Phil Strazzulla. 2025. 22 Best Applicant Tracking Systems (ATS): Full Comparison 2025. <https://www.selectsoftwarereviews.com/buyer-guide/applicant-tracking-systems>
- [65] Manish Raghavan, Solon Barocas, Jon Kleinberg, and Karen Levy. 2020. Mitigating bias in algorithmic hiring: evaluating claims and practices. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) (FAT* ’20). Association for Computing Machinery, New York, NY, USA, 469–481. doi:10.1145/3351095.3372828
- [66] LinkedIn Recruiter. 2025. LinkedIn Recruiter. <https://business.linkedin.com/talent-solutions/recruiter>
- [67] Lauren A Rivera. 2012. Hiring as cultural matching: The case of elite professional service firms. *American sociological review* 77, 6 (2012), 999–1022.
- [68] Javier Sánchez-Monedero, Lina Dencik, and Lilian Edwards. 2020. What does it mean to ‘solve’ the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*. 458–468.
- [69] Devansh Saxena and Shion Guha. 2024. Algorithmic Harms in Child Welfare: Uncertainties in Practice, Organization, and Street-level Decision-making. *ACM Journal on Responsible Computing* (March 2024). doi:10.1145/3616473 Publisher: ACM/PUB27 New York, NY.
- [70] Trevor Schachner. Febr. Prompt Engineering for HR. <https://www.shrm.org/labs/resources/prompt-engineering-for-hr>
- [71] Ben Shneiderman, Catherine Plaisant, Maxine Cohen, Steven Jacobs, Niklas Elmqvist, and Nicholas Diakopoulos. 2016. *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (6th ed.). Pearson.
- [72] Mona Sloane, Emanuel Moss, and Rumman Chowdhury. 2022. A Silicon Valley love triangle: Hiring algorithms, pseudo-science, and the quest for auditability. *Patterns* 3, 2 (2022).
- [73] S Shyam Sundar. 2020. Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAII). *Journal of Computer-Mediated Communication* 25, 1 (Jan. 2020), 74–88. doi:10.1093/jcmc/zmz026 _eprint: <https://academic.oup.com/jcmc/article-pdf/25/1/74/32961171/zmz026.pdf>.
- [74] Helena Vasconcelos, Matthew Jörke, Madeleine Grunde-McLaughlin, Tobias Gerstenberg, Michael S Bernstein, and Ranjay Krishna. 2023. Explanations can reduce overreliance on ai systems during decision-making. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–38.
- [75] Bryan Wilder, Eric Horvitz, and Ece Kamar. 2020. Learning to complement humans. *arXiv preprint arXiv:2005.00582* (2020).
- [76] Willo. 2025. Willo. <https://www.willo.video>
- [77] Christo Wilson, Avijit Ghosh, Shan Jiang, Alan Mislove, Lewis Baker, Janelle Szary, Kelly Trindel, and Frida Polli. 2021. Building and auditing fair algorithms: A case study in candidate screening. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 666–677.
- [78] Kyra Wilson and Aylin Caliskan. 2024. Gender, Race, and Intersectional Bias in Resume Screening via Language Model Retrieval. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 7, 1 (Oct. 2024), 1578–1590. doi:10.1609/aies.v7i1.31748
- [79] Kyra Wilson, Mattea Sim, Anna-Maria Gueorguieva, and Aylin Caliskan. 2025. No Thoughts Just AI: Biased LLM Hiring Recommendations Alter Human Decision Making and Limit Human Autonomy. arXiv:2509.04404 [cs.CY] <https://arxiv.org/abs/2509.04404>

	Job description generation	Database filtering	Candidate summary	AI-led screening	Candidate filtering	Qualification grouping	Candidate fraud detection	Match ranking/scores	Interview summary	Interview question generation	Interview transcription	Interview scoring	Interview feedback	AI-led interview
Applicant Tracking Systems														
Workable [80]	•	•	•				•		•					
Ashby† [5]	•	•		•		•				•				
Gem†* [35]		•		•		•	•							
Breezy† [15]						•	•							
Bullhorn† [19]	•		•	•			•	•	•					
iCIMS†* [46]	•	•					•		•					
Greenhouse† [41]	•			•				•	•					
Workday† [81]	•	•		•	•									
Supplemental Tooling														
Indeed† [47]	•	•	•				•							
LinkedIn Recruiter†* [66]	•	•			•									
Metaview†* [61]	•	•						•		•				
Hirevue† [43]								•	•	•	•			
BrightHire† [17]	•			•				•	•	•		•		
Willo† [76]				•				•	•	•	•			
Juicebox†* [48]		•	•											
Agentic Software														
Eightfold [31]	•	•		•			•	•		•	•			•
Covey [27]		•					•	•						

Table 2. Software is organized the AI functionalities it supports and ordered by recruitment stage (Section 2). Applicant tracking systems refer to platforms that support the full-cycle recruitment pipeline, while supplemental tooling offers tooling that focus on a single pipeline phase. Agentic software or agents refer to the distinct functionality of autonomous action that can operate with minimal human intervention (identified by *). For clarity, we denote software that participants reported to actively use with †

[80] Workable. 2025. Workable. <https://www.workable.com/>

[81] Workday. 2025. Workday Talent Management. <https://www.workday.com/en-us/products/talent-management/overview.html>

[82] Alice Xiang. 2024. Fairness & privacy in an age of generative ai. *Science and Technology Law Review* 25, 2 (2024).

[83] John R Zech, Marcus A Badgeley, Manway Liu, Anthony B Costa, Joseph J Titano, and Eric Karl Oermann. 2018. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. *PLoS medicine* 15, 11 (2018), e1002683.

[84] Chunpeng Zhai, Santoso Wibowo, and Lily D Li. 2024. The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments* 11, 1 (2024), 28.

[85] Shuning Zhang, Hui Wang, and Xin Yi. 2025. Exploring Collaboration Patterns and Strategies in Human-AI Co-creation through the Lens of Agency: A Scoping Review of the Top-tier HCI Literature. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article CSCW413 (Oct. 2025), 43 pages. doi:10.1145/3757594

[86] Thomas Zollo, Nikita Rajaneesh, Richard Zemel, Talia Gillis, and Emily Black. 2025. Towards Effective Discrimination Testing for Generative AI. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 1028–1047.

A Taxonomy of AI-Assisted Recruitment Software

In parallel to our interview analyses, we performed an analysis of commonly discussed HR-specific hiring tools. Our taxonomy can be found in Table 2. Through this analysis, we found that there exist AI, and specifically genAI, applications throughout the first four stages of the recruitment process (Section 2), excluding the extension of the final offer. The most common application of AI is ‘Database filtering,’ which facilitates sourcing by searching across candidate databases, though this functionality existed prior to 2022 and the introduction of ChatGPT. ‘Job Description Generation,’ which requires genAI technology, is the second most prevalent application of AI. The stage with the most genAI-specific applications is ‘Interviewing,’ where genAI has introduced new functionalities, such as summarizing interviews and generating interview questions. The AI functionalities that participants did

not explicitly mention were used include ‘AI-led screening,’ ‘Candidate filtering,’ ‘Interview scoring,’ ‘Interview feedback,’ and ‘AI-led interview.’

B Interview Protocol

We provide our complete semi-structured interview protocol below. Due to the nature of semi-structured interviews, we asked follow-up questions and omitted other questions in the protocol as needed.

Grounding Questions

Q1 Tell me about your work and your organization.

- (1) How long have you been in this role?

Q2 Can you tell me about how you are involved with company hiring processes, and a bit about what that process looks like (think of a recent hire)?

- (1) Are there steps you normally go through? Frequency? Number of people involved?
- (2) Is there a rubric you use to evaluate people? What factors do you often consider?

General Use of GenAI in Hiring

Q1 Have you used an LLM to support you with your work related to hiring? If so, how so and where in the process do you use it?

- (1) ChatGPT, Claude are popular tools, what model do you use?
- (2) Is this a general model or personalised to your company?
- (3) Is this a personal or business account?
- (4) What motivated you to use it?

Q2 What do you ask the LLM to do <in the hiring process stated previously>?

- (1) Are there specific tasks you assign to this?
- (2) What information do you give it?
- (3) What do you do with the output?
- (a) Is this useful to you? If yes/no, why?

Honing in on One Application of GenAI

Q1 What are the inputs/outputs at this part of the process?

- (1) How do you use the outputs to reach the next stage of the decision process?

Q2 Has your use of AI matched with your expectations?

Q3 You mention <AI-assisted task>—have you ever looked at the <AI output, e.g. summary of an interview> and felt that it changed your impression of <the input, e.g. a candidate/the position>?

- (1) Has it brought things to your attention you wouldn’t have noticed on your own?

Q4 Do you find yourself deferring to the <AI output>?

- (1) Do you think you do so just because it’s convenient?
- (2) Do you think these tools are improving the selection process?
- (3) What factors make you trust or question its judgment even when your professional instinct disagrees?

Q5 Do you ever check for errors on the outputs of the LLM? What do you look for?

Perception of Agency over Choices

Q1 Is there a reason you haven't used LLMs in other parts of your hiring process?

- (1) Is there something you would never ask an LLM to do for you?
- (2) Do you have any concerns about how it is currently used or could be used in your company?

Q2 Do you feel like you understand why it produces the candidates, scores, summaries etc. that it does?

- (1) Has there been anything that challenges this perspective?

Q3 How much do you feel like you are still in control of the final decision, versus the system guiding or constraining your choices?

Q4 Do you feel like you've lost any skills as a result of relying on AI?

- (1) If the AI suddenly stopped working tomorrow, would you know how to tackle something?

Q5 How do you explain your hiring choices to others (e.g., managers) when the AI system has played a role in shaping those choices?

Q6 <If participant uses genAI in applicant evaluation> Can you give me an example of a prompt you've used to evaluate applicants with an LLM? This prompt can be generic.

C Codebook

See Table 3 for the codebook resulting from our process of thematic analysis.

Received 24 January 2026

Company-focused	Candidate-focused
<i>Business Culture</i> : What is the company's culture like? <i>Business Size</i> : How large is the company <i>Pressure to Use AI in Business</i> : Participant reporting pressure to use AI from their employer <i>Stage of AI Adoption</i> : How much AI the employer is incorporating company-wide <i>Prohibitive Cost of AI</i> : Statement that AI tools are out of the price range for a participant's company	<i>Candidate Use of AI</i> : How candidates use AI <i>Candidate Effort stands out</i> : The manual effort of a candidate is obvious when reviewing their application <i>Desired Candidate Skillsets</i> : Skills the participant looks for <i>Applicant-recruiter Arms Race</i> : Participant's use of AI to combat candidate's use <i>Fake Candidates</i> : Discussion of reportedly "fake" candidates <i>High Volume of Applications</i> : Capturing the great demand on recruiters for reviewing applications
Agency-focused	Impacts/Future
<i>Auto/AI-Resistant Work</i> : Assertion that a specific task cannot be performed by AI <i>Deference to AI</i> : Instance where the participant trusts AI <i>Employer AI Red Line</i> : Determination of when AI should not be used in the recruitment process <i>Human Agency Maintained</i> : Belief that the participant maintains decision-making agency <i>Lack of Trust in AI/Skepticism</i> : Skepticism toward the use of AI in general or a specific task <i>Shared Decision-making</i> : Belief that AI has some impact on the decision-making process <i>Perceived Human Manual</i> : Claim that participant is manually performing a certain task	<i>AI Future Potential</i> : Looking forward to AI's future impact <i>Legal and Privacy Risk</i> : Reported fear of risk or performing actions that constitute risk <i>Loss of Skill/Deskilling</i> : Observed behavior that reflects deskilling <i>Overpromising of AI</i> : Belief that AI marketing has overpromised on the utility of AI <i>AI Used Differently by Generations</i> : Pattern that different age groups use AI in disparate manners <i>Two-tiered AI use</i> : Different usage of AI depending on qualities of the role or candidate
Participant/Role Information	Recruitment Stage
<i>Job Level Type</i> : The level of position that the participant hires <i>Participant Background</i> : Information about participant's role and experience <i>Self-Assessment of AI Proficiency</i> : A participant's confidence in their use of AI <i>Time in Role</i> : How long has the participant been in this role	<i>Assessment</i> : Process of forming a judgment on a candidate <i>Conceptualizing/Themeing</i> : Generation of a job description <i>Final Decision-making</i> : Selection of candidate(s) to send an offer <i>Interviewing</i> : Conversations with candidates <i>Logistics</i> : Planning interviews, sending communication, etc. <i>Sourcing</i> : Finding candidates to apply to a job opening <i>Screening</i> : Comparing a candidate's skills and qualifications to the job requirements
Task Type	Performance of AI Systems
<i>Agentic AI</i> : Using AI agents to complete a task <i>AI Content Generation</i> : Using genAI to produce content <i>Analogue-Technical Processes</i> : A process combining both technical and manual methods <i>Grouping</i> : Classifying candidates in a fixed number of groups <i>Keyword Searching</i> : Looking for specific keywords <i>Manual Work</i> : Human-only completion of a recruitment task <i>Transcription</i> : Translation from a recording to text <i>Summarizing</i> : Reduction of an input into a shortened form <i>Ranking</i> : Comparing candidates to one another individually <i>Research/Information Gathering</i> : Process of learning about a certain domain <i>System Indicators of Job Compatibility</i> : How is job compatibility reflected to the participant	<i>AI Evaluates Candidates Badly</i> : Inconsistent or incorrect evaluation of candidates <i>AI is Biased</i> : Relaying of concerns about biased outputs from AI <i>AI Lacks Nuance</i> : Discusses that AI systems may be brittle or do not understand humans <i>AI Makes More Work</i> : Claim that the participant's use of AI increases their workload <i>AI Reveals Overlooked Details</i> : AI provides the participant with a more detailed interpretation of the input <i>AI Saves Time</i> : The use of AI reduces time spent on a task <i>AI Tool Performance</i> : Reflects how well the AI tool is seen to perform in general <i>AI Use is Essential for Efficiency</i> : Belief that one's use of AI is essential for the efficient completion of tasks

Table 3. The codebook derived from our thematic analysis, grouped by the codes' focus.