AI AND WORKPLACE DEI INITIATIVES: OPPORTUNITIES AND CHALLENGES

Posted on June 25, 2024

Categories: Insights, Publications

Given their wide-ranging and ever-developing capabilities, artificial intelligence ("**AI**") and AI-supported tools certainly have the potential to support the introduction and advancement of innovative and effective diversity, equity and inclusion ("**DEI**") initiatives in the workplace.

However, while AI can help to identify patterns of unconscious bias and support the development of more inclusive workplaces, among other potential benefits, the use of AI in furtherance of DEI initiatives is not without risks.

Potential Opportunities

By leveraging automation, AI may enable organizations to collect and analyze data that may have previously been too costly or time-consuming to process manually. For instance, AI-enabled tools may have the capacity to aggregate, analyze and draw conclusions from large or complex datasets to spot indicators of bias or inequity in recruitment, advancement, compensation, performance evaluation, work allocation and other workplace processes.

One popular use of AI-enabled tools in the DEI space is to support more equitable recruitment efforts. Certain AI products may promote greater objectivity and consistency when drafting job postings and recruitment materials, screening application materials and generating interview questions based on core job competencies.

AI tools may also be used to support the development and implementation of workplace accommodations, including for employees who do not work in their first language and/or who have disabilities. For example, AI is enabling increasingly sophisticated text-to-speech and speech-to text tools, as well as the ability to translate text or speech nearly instantaneously, both of which may improve communications with employees.

Potential Risks

Notwithstanding the potential benefits of implementing AI-enabled tools in furtherance of DEI initiatives, it is important to be aware that the adoption of such tools may also present challenges and risks that could

undermine the advancement of workplace DEI objectives.

One prevalent concern is the potential for AI to introduce or perpetuate bias in workplace processes.

Bias can arise or be perpetuated because of an AI tool's algorithm itself. An example of algorithmic bias occurs when an AI tool is designed to over-emphasize a particular characteristic (such as time spent out of the paid workforce) when scanning and ranking application materials.

Bias can also arise from the training data used to teach an AI tool. If data of a biased or discriminatory nature is used to train an AI tool, the tool may produce outputs reflecting or even exacerbating that bias. For instance, an employer that wishes to use an AI-enabled tool to scan and rank application materials may be prompted to input data related to the application materials of past successful candidates so that the tool can identify similar traits, skills or experiences. However, if the employer has historically hired candidates that are predominately of a certain race or gender, these traits may be over-represented in the data, which could in turn prompt the AI-driven identification and automated decision making to over-emphasize the importance of these traits.

Where an AI-enabled tool is used to make or inform employment-related predictions or decisions, the introduction of bias can have devastating consequences, including prompting unfair, adverse or inaccurate decisions about individuals who are members of equity-seeking groups. This can seriously compromise a workplace's DEI objectives, and possibly cause an employer to run afoul of human rights legislation by discriminating against certain groups of candidates or employees.

Finally, the collection and storage of sensitive demographic information in furtherance of DEI initiatives, and the processing of such information using AI-enabled tools, introduces unique security risks. Unauthorized access or use may be perpetuated by malicious third parties, including via hacking or ransomware attacks, as well as through employee snooping or misuse in the workplace. Canadian privacy regulators have also identified that there are security threats of particular concern when using generative AI tools, including prompt injection attacks, model inversion attacks and jailbreaking, all of which may jeopardize the confidentiality of sensitive employee personal information.[1]

Action Items

Employers who are considering using AI tools to advance workplace DEI initiatives should take proactive measures to manage risks associated with the use of such tools. Action items to consider implementing include:

- **Carrying out risk assessments to vet vendors and AI tools before procuring or implementing them.** Consider implementing policies and practices – such as standard questionnaires – regarding the selection and onboarding of AI-enabled tools and the engagement of vendors who develop and

implement them. Evaluate what provisions are mandatory and prudent to include in your contract or terms of service with any third parties involved in providing an AI solution.

- **Conducting privacy impact assessments prior to processing personal information for DEI initiatives.** Map and evaluate the intended processing of personal information in connection with the use of the AI-enabled tool – from collection, through to disposal – to understand your workplace's obligations under privacy, employment and human rights law and to formulate a plan to ensure compliance with those obligations. For example, consider any provisions under privacy and human rights laws that govern if and how your organization can collect and use job applicants' and employees' demographic information for your intended purposes, as well as any notice and consent obligations under applicable privacy and employment laws.
- **Understanding how bias will be prevented, monitored for and addressed.** As part of your due diligence process concerning AI vendors and tools, confirm that the AI model is trained on diverse, high-quality data, that any measured characteristics and indicators are job relevant and do not perpetuate biases, and that bias in the algorithm and training data have otherwise been evaluated and effectively mitigated. Make inquiries about how bias will be monitored for and addressed by the vendor on an ongoing basis and what role an employer is expected to play in this process.
- **Remembering the importance of human oversight.** Ensure that there is still appropriate human oversight with respect to the processes undertaken or impacted by the AI-enabled tool. Ultimately, employers are responsible for ensuring adherence to employment, human rights and accessibility laws, and relying on AI tools alone will not discharge this responsibility. Ensure that role-specific training is provided to personnel who will be using AI-enabled tools, including to address when and how such tools can be used and how and when human oversight must be provided.
- **Implementing appropriate security safeguards.** Ensure that your organization, and any relevant vendors, have appropriate physical, organizational and technological safeguards in place to protect sensitive DEI-related information from loss, theft and unauthorized access, use and disclosure, among other risks. Evaluate the safeguards periodically to ensure that they continue to adequately address risks specific to the use of AI-enabled tools.
- **Developing an AI governance structure.** Organizations should develop an internal AI governance structure that takes into account any DEI-related use cases for AI-enabled tools. In addition to the vendor and tool vetting procedure referenced above, it is also important to have appropriate policies to establish guardrails around employees' use of AI-enabled tools in the course of their job functions, including prescriptive limitations on if and how employees are allowed to use such tools.

If you have any questions about implementing these and other practical measures to address risks associated with the use of AI-enabled tools to further DEI initiatives, a member of McMillan's Privacy & Data Protection

Group would be happy to assist you.

[1] Principles for responsible, trustworthy and privacy-protective generative AI technologies

By Kristen Pennington and Laura Kabbabe (Student-at-Law)

A Cautionary Note