

CATEGORIA

Notícias

CITAÇÃO

Rev. *Ciência Elem.*, V12(02):022
doi.org/10.24927/rce2024.022

EDITOR

João Nuno Tavares
Universidade do Porto

COPYRIGHT

© Casa das Ciências 2024.
Este artigo é de acesso livre,
distribuído sob licença Creative
Commons com a designação
[CC-BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/), que permite
a utilização e a partilha para fins
não comerciais, desde que citado
o autor e a fonte original do artigo.

rce.casadasciencias.org



Funcionários de grandes empresas de IA apela por supervisão mais rigorosa e transparência

Um grupo formado por atuais e ex-funcionários de algumas das principais empresas de inteligência artificial (IA) do mundo afirmou acreditar fortemente no potencial das tecnologias de IA para trazer benefícios significativos à humanidade. No entanto, expressam preocupações profundas sobre os riscos que essas tecnologias podem representar, que variam desde a ampliação das desigualdades existentes e a manipulação da informação, até a perda de controle sobre sistemas autónomos de IA, levando até mesmo a ameaças de extinção da humanidade. Entre eles contam-se alguns dos principais inventores da IA, tais como Jacob Hilton, William Saunders, Carroll Wainwright, Daniel Kokotajlo, Daniel Ziegler (OpenAI), Ramana Kumare Neel Nanda (Google DeepMind), e ainda Yoshua Bengio, Geoffrey Hinton e Stuart Russell, bem conhecidos nos meios académicos ligados à investigação em IA.

As advertências não são levianas, uma vez que as próprias empresas de IA, governos ao redor do mundo e especialistas independentes já reconheceram esses riscos. “Temos esperança de que estes riscos possam ser adequadamente mitigados com orientação suficiente da comunidade científica, dos decisores políticos e do público”.

Entretanto, um obstáculo significativo permanece: as empresas de IA têm fortes incentivos financeiros para evitar uma supervisão eficaz. Os denunciantes argumentam que as

estruturas personalizadas de governação corporativa são insuficientes para enfrentar essa realidade. As empresas possuem informações críticas e não públicas sobre as suas capacidades tecnológicas, medidas de proteção e níveis de risco, mas atualmente têm pouca obrigação de compartilhar tais informações com os governos e nenhuma com a sociedade civil.

O grupo de funcionários observa que, sem uma supervisão governamental eficaz, são poucos os que podem responsabilizar estas empresas publicamente. No entanto, grandes acordos de confidencialidade impedem de expressar as suas preocupações, exceto internamente, onde essas questões podem ser ignoradas. Eles destacam que as proteções usuais para denunciante são insuficientes, uma vez que se concentram nas atividades ilegais, e muitos dos riscos atuais ainda não estão regulamentados.

Com base nisso, os funcionários apelam a que as empresas de IA acelerem seu comprometimento com quatro princípios fundamentais:

1. Abertura à Crítica: as empresas não devem celebrar nem fazer cumprir acordos que proíbam críticas relacionadas a riscos e não devem retaliar por tais críticas.
2. Processos Anónimos de Denúncia: facilitação de processos verificados onde funcionários atuais e antigos possam levantar preocupações de risco anonimamente ao conselho, reguladores e organizações independentes.
3. Cultura de Crítica Aberta: apoio a uma cultura que permita críticas abertas e elevando preocupações sobre riscos publicamente ou a organizações apropriadas, protegendo segredos comerciais e outros interesses de propriedade intelectual.

4. Proteção Contra Retaliação: nenhuma retaliação contra funcionários que compartilhem publicamente informações relacionadas a riscos após a falha de outros processos.

A declaração reflete a urgência da necessidade de uma supervisão mais robusta e enfatiza a importância de permitir que essas vozes sejam ouvidas em prol da segurança e continuidade do desenvolvimento humano.

Especialistas alertam para desvantagens e riscos das provas de avaliação em formato digital

Com a crescente digitalização do ensino, a implementação de provas de avaliação em formato digital tem gerado intensas discussões entre professores, estudantes e especialistas em tecnologia. Embora a mudança prometa modernizar o sistema educativo, várias desvantagens e riscos têm sido apontados, levantando preocupações significativas.

Falta de Infraestrutura Adequada e Problemas Técnicos Frequentes

Muitas escolas ainda não possuem a infraestrutura tecnológica necessária para realizar provas digitais de forma eficiente. A ausência de uma rede de internet robusta e a insuficiência de dispositivos adequados podem prejudicar o desempenho dos alunos na execução das provas. Por outro lado, os testes em formato digital estão sujeitos a falhas técnicas, que não apenas atrasam o processo de avaliação, como também podem resultar na perda de dados importantes, forçando os alunos a reiniciar as suas provas ou até mesmo refazê-las.

Desigualdade de Acesso

A digitalização das provas pode intensificar as desigualdades já existentes no sistema educativo. Alunos de escolas com menos recursos, ou aqueles que não possuem acesso a computadores e internet em casa, encontram-se em desvantagem em comparação com os seus colegas mais favorecidos. Isso pode agravar ainda mais a disparidade no desempenho escolar entre diferentes grupos socioeconômicos.

Segurança de Dados e Privacidade

A segurança dos dados dos alunos é outra preocupação crítica. Provas digitais envolvem o armazenamento e o processamento de informações pessoais e acadêmicas, o que pode ser alvo de ciberataques. A adequação das medidas de proteção de dados ainda é uma área de incerteza e preocupação para pais e educadores.

Impacto no Desempenho dos Alunos

Estudos indicam que a transição para avaliações digitais pode inicialmente impactar negativamente o desempenho dos alunos. A necessidade de adaptação a um novo formato de prova pode causar ansiedade adicional, além de exigir capacidades técnicas que nem todos os alunos dominam.

Conclusão

Enquanto a digitalização das provas de avaliação oferece inúmeras vantagens potenciais, como a rapidez na correção e o acesso a uma gama mais diversificada de perguntas, os numerosos riscos e desvantagens não podem ser ignorados. A implementação bem-sucedida de provas digitais exige um planejamento cuidadoso, investimento em infraestrutura e um compromisso com a equidade educacional para garantir que todos os alunos tenham as mesmas oportunidades de sucesso.