Scraped Content

Title: Generative AI Ethics: 8 Biggest Concerns and Risks

Home AI business strategies Tech Accelerator What is generative AI? Everything you need to know Prev Next 7 generative AI challenges that businesses should consider Bard vs. ChatGPT: What's the difference? Download this guide 1 X Free Download What is generative AI? Everything you need to know The potential of AI technology has been percolating in the background for years. But when ChatGPT, the AI chatbot, began grabbing headlines in early 2023, it put generative AI in the spotlight.

This guide is your go-to manual for generative AI, covering its benefits, limits, use cases, prospects and much more. By George Lawton Published: 01 Nov 2023 Like other forms of AI, generative AI can influence a number of ethical issues and risks surrounding data privacy, security, policies and workforces. Generative AI technology can also potentially produce a series of new business risks like misinformation, plagiarism, copyright infringements and harmful content. Lack of transparency and the potential for worker displacement are additional issues that enterprises may need to address . "Many of the risks posed by generative AI ? are enhanced and more concerning than others," said Tad Roselund, managing director and senior partner at consultancy BCG. Those risks require a comprehensive approach, including a clearly defined strategy, good governance and a commitment to responsible AI . A corporate culture that embraces generative AI ethics must consider eight important issues. 1. Distribution of harmful content AI systems can create content automatically based on text prompts by humans. "These systems can generate enormous productivity improvements, but they can also be used for harm , either intentional or unintentional," explained Bret Greenstein, partner, cloud and digital analytics insights, at professional services consultancy PwC. An AI-generated email sent on behalf of the company , for example, could inadvertently contain offensive language or issue harmful guidance to employees. Generative AI should be used to augment, not replace humans or processes, Greenstein advised, to ensure content meets the company's ethical expectations and supports its brand values. 2. Copyright and

Scraped Content

legal exposure Popular generative AI tools are trained on massive image and text databases from multiple sources, including the internet. When these tools create images or generate lines of code, the data's source could be unknown, which can be problematic for a bank handling financial transactions or pharmaceutical company relying on a formula for a complex molecule in a drug. Reputational and financial risks could also be massive if one company's product is based on another company's intellectual property. "Companies must look to validate outputs from the models," Roselund advised, "until legal precedents provide clarity around IP and copyright challenges." This article is part of What is generative AI? Everything you need to know Which also includes: 8 top generative AI tool categories for 2024 Will AI replace jobs? 9 job types that might be affected 19 of the best large language models in 2024 Businesses are scrambling to maximize the benefits of today's generative AI while wrestling with inherent ethical issues. 3. Data privacy violations Generative AI large language models (LLMs) are trained on data sets that sometimes include personally identifiable information (PII) about individuals. This data can sometimes be elicited with a simple text prompt, noted Abhishek Gupta, founder and principal researcher at the Montreal AI Ethics Institute. And compared to traditional search engines, it can be more difficult for a consumer to locate and request removal of the information. Companies that build or fine-tune LLMs must ensure that PII isn't embedded in the language models and that it's easy to remove PII from these models in compliance with privacy laws. 4. Sensitive information disclosure Generative AI is democratizing AI capabilities and making them more accessible. This combination of democratization and accessibility, Roselund said, could potentially lead to a medical researcher inadvertently disclosing sensitive patient information or a consumer brand unwittingly exposing its product strategy to a third party. The consequences of unintended incidents like these could irrevocably breach patient or customer trust and carry legal ramifications. Roselund recommended that companies institute clear guidelines, governance and effective communication from the top down, emphasizing shared responsibility for safeguarding sensitive information, protected data and IP. 5. Amplification of existing bias Generative AI can potentially amplify existing biases -- for

example, bias can be found in data used for training LLMs outside the control of companies that use these language models for specific applications . It's important for companies working on AI to have diverse leaders and subject matter experts to help identify unconscious bias in data and models, Greenstein said. 6. Workforce roles and morale AI can do a lot more of the daily tasks that knowledge workers do, including writing, coding, content creation, summarization and analysis, said Greenstein. Although worker displacement and replacement have been ongoing since the first AI and automation tools were deployed, the pace has accelerated as a result of the innovations in generative AI technologies . "The future of work itself is changing," Greenstein added, "and the most ethical companies are investing in this [change]." Ethical responses have included investments in preparing certain parts of the workforce for the new roles created by generative AI applications . Businesses, for example, will need to help employees develop generative AI skills such as prompt engineering. "The truly existential ethical challenge for adoption of generative AI is its impact on organizational design, work and ultimately on individual workers," said Nick Kramer, vice president of applied solutions at consultancy SSA & Company. "This will not only minimize the negative impacts, but it will also prepare the companies for growth." 7. Data provenance Generative AI systems consume tremendous volumes of data that could be inadequately governed, of questionable origin, used without consent or contain bias. Additional levels of inaccuracy can be amplified by social influencers or the AI systems themselves. "The accuracy of a generative AI system depends on the corpus of data it uses and its provenance," explained Scott Zoldi, chief analytics officer at credit scoring services company FICO. "ChatGPT-4 is mining the internet for data, and a lot of it is truly garbage, presenting a basic accuracy problem on answers to questions to which we don't know the answer." FICO, according to Zoldi, has been using generative AI for more than a decade to simulate edge cases in training fraud detection algorithms. The generated data is always labeled as synthetic data so Zoldi's team knows where the data is allowed to be used. "We treat it as walled-off data for the purposes of test and simulation only," he said. "Synthetic data produced by generative AI does not inform the model going forward in the future. We contain this

generative asset and do not allow it 'out in the wild.'" 8. Lack of explainability and interpretability Many generative AI systems group facts together probabilistically, going back to the way AI has learned to associate data elements with one another, Zoldi explained. But these details aren't always revealed when using applications like ChatGPT . Consequently, data trustworthiness is called into question. When interrogating generative AI, analysts expect to arrive at a causal explanation for outcomes. But machine learning models and generative AI search for correlations, not causality. "That's where we humans need to insist on model interpretability -- the reason why the model gave the answer it did," Zoldi said. "And truly understand if an answer is a plausible explanation versus taking the outcome at face value." Until that level of trustworthiness can be achieved, generative AI systems should not be relied upon to provide answers that could significantly affect lives and livelihoods. Next Steps Skills needed to become a prompt engineer How to prevent deepfakes in the era of generative AI What is an AI hallucination? What is reinforcement learning from human feedback? Generative models: VAEs, GANs, diffusion, transformers, NeRFs Dig Deeper on AI business strategies ChatGPT and human resources: What HR leaders should know By: Carolyn Heinze 5 benefits of using process mining By: George Lawton How AI is shaping the future of ERP By: George Lawton 13 types of business risks for companies to manage By: Mary Pratt Business Analytics The top augmented analytics tools vendors are offering Augmented analytics tools play a crucial role in providing users of all technical skills levels access to AI and analytics daily.... Understand the BI lifecycle to build an effective architecture Following and understanding the five steps of the business intelligence lifecycle is the foundation of a successful and effective... Teradata integrates with DataRobot to aid AI development The partnership provides the data management and analytics vendor's users with more choice so that they can develop advanced ... CIO How to overcome digital transformation fatigue For many employees, poorly executed digital transformation projects lead to fatigue. Learn effective methods for maintaining ... Where 2024 U.S. presidential candidates stand on tech issues The next U.S. president will set the tone on tech issues such as AI regulation, data privacy and climate tech. This guide breaks ... South

Scraped Content

Korea, Taiwan, Japan monitor EU's Digital Markets Act Countries including South Korea and Taiwan are taking a cautious approach to tech regulation and assessing the impact of existing... Data Management IoT data management extends best practices to the edge IoT devices generate and collect data from points all over the network. Organizations must apply general data management best ... Top 6 data fabric tools providers Organizations that want a data fabric architecture should consider how six top vendors use their suite of tools to support a ... Qlik launches Talend Cloud, aims to ensure data is trusted The longtime BI vendor's new integration platform aims to deliver trust in the information users work with to inform analytics ... ERP 8 ERP security best practices to implement now As supply chain attacks continue, ERP security is more important than ever. Learn some best practices to get started protecting ... Companies reassess generative AI in supply chain tech GenAI and other tech is being used to tackle supply chain challenges such as high costs and disruption, but organizations are ... Guide to supply chain management Wars and pandemics offer stark evidence of how global supply chains can break apart. Effective supply chain management can ... Close