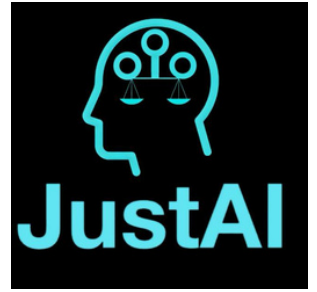


BI-MONTHLY NEWSLETTER



ISSUE 005

1-15 September, 2024

REDEFINING LAW AND JUSTICE IN THE AGE OF AI

OUR SERVICES

Networking Opportunities: We proudly host the JUSTAI Community, a vibrant network of industry specialists, academicians, and student members focused on the intersection of law and artificial intelligence.

Training & Skill Enhancement: At JustAI, we organise various events like Seminars, Guest Lectures and Workshops that features field experts who share valuable insights on AI's impact on society and provide expert. Additionally, we provide courses led by industry experts and powered by AI technology, allowing you to learn at your own pace.

JUSTAI Information Hub: JustAI offers curated resources to stay updated on the latest developments and opportunities in Artificial Intelligence (AI). It serves as a centralized source for AI news, regulatory updates, governance frameworks, and more, designed to support professionals, researchers, and enthusiasts in traversing the rapidly evolving AI landscape.

TOP NEWS

- AI MODELS BY OPENAI AND ANTHROPIC TO UNDERGO SAFETY TESTING BEFORE PUBLIC RELEASE
- NETHERLANDS DATA PROTECTION AUTHORITY IMPOSES FINE FOR ILLEGAL COLLECTION OF BIOMETRIC DATA OF EU CITIZENS
- META TO INFORM BRAZILIANS ON PERSONAL DATA USAGE FOR AI TRAINING
- MANDATORY GUARDRAILS FOR AI DEVELOPMENT PROPOSED BY AUSTRALIAN GOVERNMENT
- WHY AI MAY NOT BE SUITABLE FOR WRITING JUDGMENTS: INSIGHTS FROM PATNA HIGH COURT CHIEF JUSTICE – K. VINOD CHANDRAN
- SANJEEV SANYAL CRITICISES AI REGULATORY APPROACHES FOLLOWED BY EUROPE AND CHINA
- EU INVESTIGATES GOOGLE'S AI MODEL FOR PRIVACY COMPLIANCE
- OPEN AI UNVEILS ITS SECRET PROJECT 'STRAWBERRY': CAN IT THINK BEFORE RESPONDING?

[CLICK HERE TO JOIN JUSTAI COMMUNITY](#)

[CLICK HERE TO SUBSCRIBE JUSTAI NEWSLETTER](#)

BI-Monthly Newsletter

REDEFINING LAW AND JUSTICE IN THE AGE OF AI



AI MODELS BY OPENAI AND ANTHROPIC TO UNDERGO SAFETY TESTING BEFORE PUBLIC RELEASE (3 September, 2024)

OpenAI and Anthropic have partnered with the U.S. AI Safety Institute to subject their upcoming AI models to rigorous safety testing before public release, marking a noteworthy shift toward prioritizing safety and ethical standards in AI development. The U.S. AI Safety Institute, operating under the National Institute of Standards and Technology (NIST), will gain early access to AI models from both companies, allowing for thorough evaluation and feedback to ensure safer AI technologies. This collaboration could influence global AI regulation, encouraging other countries to adopt similar safety assessments for AI models, as seen with recent legislative efforts in California and earlier advisories in India. [Read More...](#)

ANTHROPIC



OPENAI'S APPROACH TO AI SAFETY



[CLICK HERE TO JOIN JUSTAI COMMUNITY](#)

[CLICK HERE TO SUBSCRIBE
JUSTAI NEWSLETTER](#)

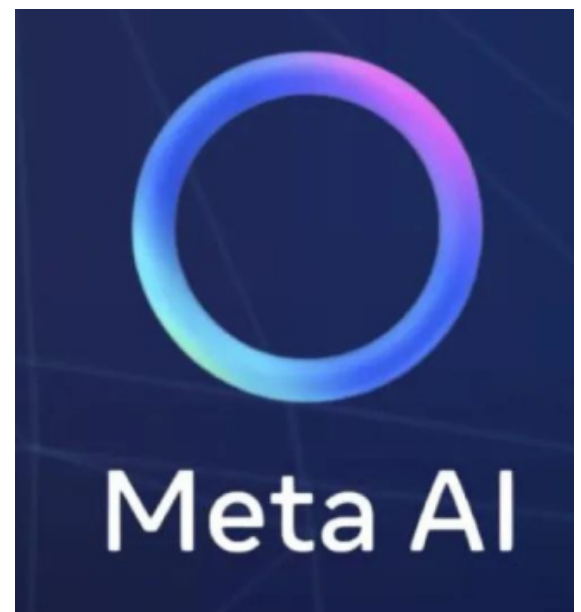
NETHERLANDS DATA PROTECTION AUTHORITY IMPOSES FINE FOR ILLEGAL COLLECTION OF BIOMETRIC DATA OF EU CITIZENS (4 September, 2024)



The Netherlands Data Protection Authority fined Clearview AI €30.5 million for violating GDPR by illegally collecting and processing biometric data of EU citizens. Clearview's facial recognition technology scrapes billions of images from the internet without consent, and the company has not complied with GDPR data access requests. Despite the fines, Clearview continues operations, disputing the legality of the penalties. This case underscores global concerns about facial recognition technology and privacy rights. [Read More...](#)

META TO INFORM BRAZILIANS ON PERSONAL DATA USAGE FOR AI TRAINING (5 September, 2024)

Meta has agreed to notify Brazilian users about how their personal data is used for training AI models. Starting September 2024, users will receive notifications via email and social media, allowing them to opt out of AI training. This follows a suspension by Brazil's National Data Protection Authority (ANPD) over privacy concerns, which has since been lifted after Meta agreed to greater transparency. The suspension of certain AI tools, including WhatsApp AI stickers, remains, with no confirmation of when they will resume. [Read More..](#)



MANDATORY GUARDRAILS FOR AI DEVELOPMENT PROPOSED BY AUSTRALIAN GOVERNMENT (6 September, 2024)

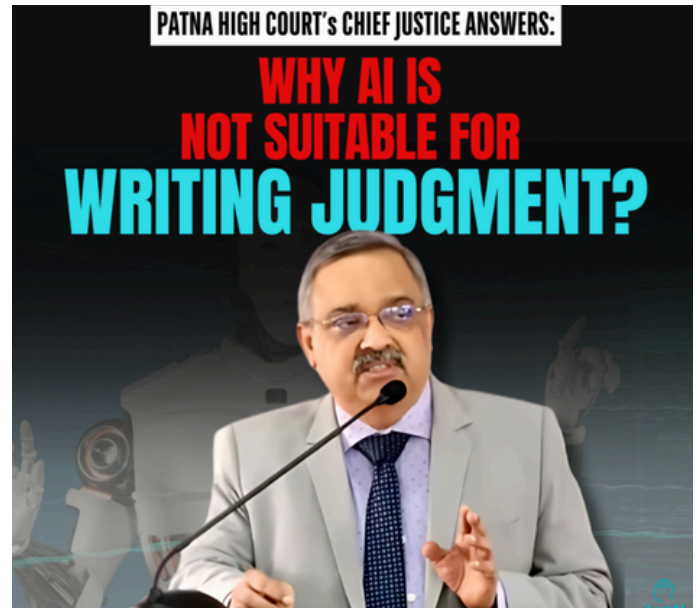
The Australian government has proposed 10 mandatory guardrails for AI development, focusing on high-risk environments. Businesses are encouraged to follow a Voluntary AI Safety Standard while the guardrails undergo public consultation until October 4, 2024. The government aims to prevent AI-related harm while fostering innovation, potentially leading to a new Australian AI Act or updates to existing laws. 10 Guardrails are:



- **Accountability:** Organizations will need to implement and publish an accountability process for regulatory compliance, including policies for data management and risk assessment.
- **Risk Management:** This involves creating processes to identify and mitigate AI-related risks, considering not just technical risks but also impacts on society, specific communities, and individuals.
- **Data Protection:** Organizations must ensure that AI systems protect privacy through robust cybersecurity measures and quality data governance.
- **Testing:** AI systems must undergo rigorous testing and continuous monitoring to ensure they perform as expected without causing unintended harm.
- **Human Control:** Ensures meaningful human oversight throughout the AI lifecycle, allowing for intervention when necessary.
- **User Information:** Requires organizations to clearly inform end-users when they are interacting with AI or when AI is being used to make decisions about them.
- **Challenging AI Decisions:** Provides people negatively impacted by AI systems the right to challenge decisions or outcomes.
- **Transparency:** Requires organizations to maintain transparency regarding their data, models, and systems.
- **AI Records:** Organizations must maintain comprehensive records of their AI systems throughout their lifecycle, including technical documentation.
- **AI Assessments:** AI systems will be subject to conformity assessments to ensure adherence to the guardrails. [Read More...](#)

WHY AI MAY NOT BE SUITABLE FOR WRITING JUDGMENTS: INSIGHTS FROM PATNA HIGH COURT CHIEF JUSTICE – K. VINOD CHANDRAN (8 September, 2024)

Patna High Court Chief Justice K. Vinod Chandran highlighted that AI is not yet suitable for writing judgments. He emphasized that judgment writing requires human insight, emotional intelligence, and moral reasoning—areas where AI falls short. While technology aids the legal process, AI lacks the nuanced understanding needed for interpreting law and making ethical decisions. Chandran underscored the irreplaceable value of human effort in legal work, such as translating judgments, which AI cannot replicate. [Read More...](#)



SANJEEV SANYAL CRITICISES AI REGULATORY APPROACHES FOLLOWED BY EUROPE AND CHINA (11 September, 2024)



Sanjeev Sanyal critiques the AI regulatory approaches of Europe and China, arguing that Europe's stringent regulations may hinder innovation and economic growth, while China's focus on centralized control and state surveillance could compromise individual freedoms. Sanyal suggests that India's AI regulatory framework should seek a balance between fostering technological advancement and ensuring ethical standards and privacy protections, drawing lessons from both approaches while avoiding their drawbacks. [Read More...](#)

EU INVESTIGATES GOOGLE'S AI MODEL FOR PRIVACY COMPLIANCE (13 SEPTEMBER, 2024)



The European Union regulators have launched a detailed investigation into Google's AI model, PaLM2, to determine its compliance with the General Data Protection Regulation (GDPR). This probe is led by Ireland's Data Protection Commission (DPC), Google's lead privacy regulator in the EU via a press release on 12th September, 2024. This investigation is part of a broader trend in the EU, where tech giants like Google, Meta, and X (formerly Twitter) are facing increasing regulatory scrutiny over their AI and data management practices. Recently, X agreed to cease using user data from European Union citizens to train its AI systems following court action by the Irish DPC. The outcome of the investigation into Google's PaLM2 could set a critical precedent for the future of AI development in the EU. It may lead to more stringent regulations regarding how AI models handle personal data and the measures companies must take to protect user privacy.

The European Union is investigating Google's AI model, PaLM2, for potential violations of the General Data Protection Regulation (GDPR). Ireland's Data Protection Commission (DPC) is leading the inquiry to assess whether Google complies with data privacy rules, particularly regarding how personal data is handled in AI development. This reflects broader concerns about the use of large datasets in AI and whether individuals' rights are respected in the process. The outcome may influence how AI systems manage personal data in the future.

[Read More...](#)

OPEN AI UNVEILS ITS SECRET PROJECT 'STRAWBERRY': CAN IT THINK BEFORE RESPONDING? (14 SEPTEMBER, 2024)



On 12th September, 2024, OpenAI officially unveiled its latest AI model, OpenAI o1, previously codenamed “Strawberry” and this marks a significant step forward in the world of artificial intelligence, particularly for tasks requiring advanced reasoning and problem-solving. Unlike the more versatile GPT-4o, which is widely known for handling a broad range of general tasks, OpenAI o1 targets specific domains, excelling in mathematics, coding, and science.

OpenAI o1 introduces a “chain-of-thought” process that allows the model to reason through problems step by step, a feature absent in previous models and this is crucial for tackling complex, multistep problems. For example, in coding or solving high-level math problems, breaking down the task into smaller, manageable steps increases accuracy and reduces errors. The chain-of-thought process significantly improves o1’s reasoning abilities, especially in scenarios requiring logic and deduction. In benchmarks, o1 has demonstrated superior performance in coding tasks, placing in the 89th percentile on Codeforces, a competitive programming platform and it also outperformed human-level accuracy in PhD-level science questions across fields like physics, biology, and chemistry. These results highlight how o1 surpasses general-purpose models when it comes to in-depth reasoning. [Read More...](#)