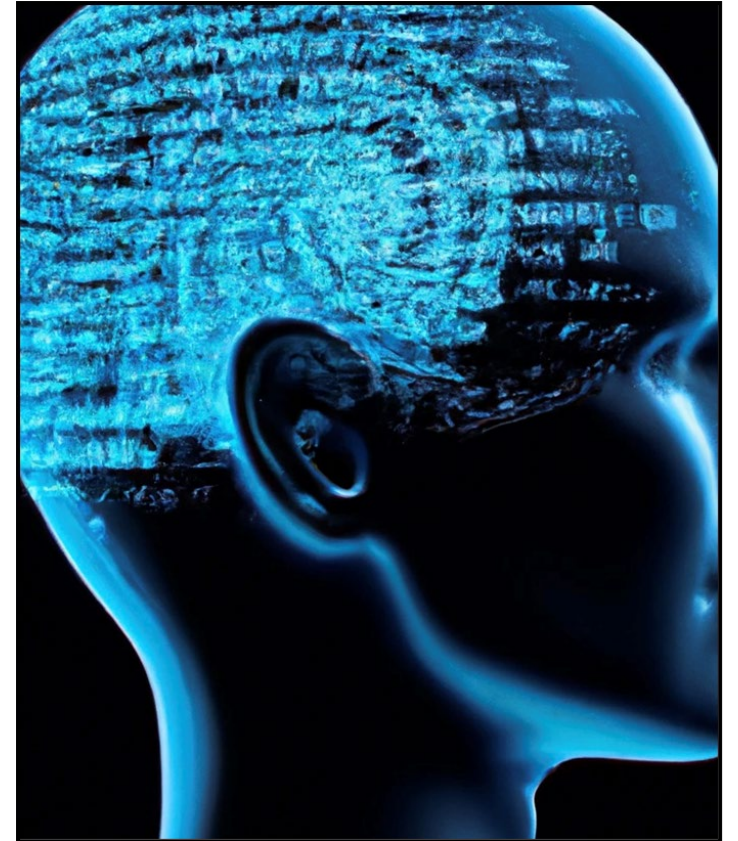




Challenges and Issues With Generative AI

Christiana Till and Art Robbins



Scope of the Document



This document aims to shed light on the **top challenges** and **issues** related to deploying Generative AI



AI possesses **transformative power**, positively impacting business and **driving innovation**



However, real **risks** and **challenges** accompany AI's growth, requiring **attention** and **proactive measures**



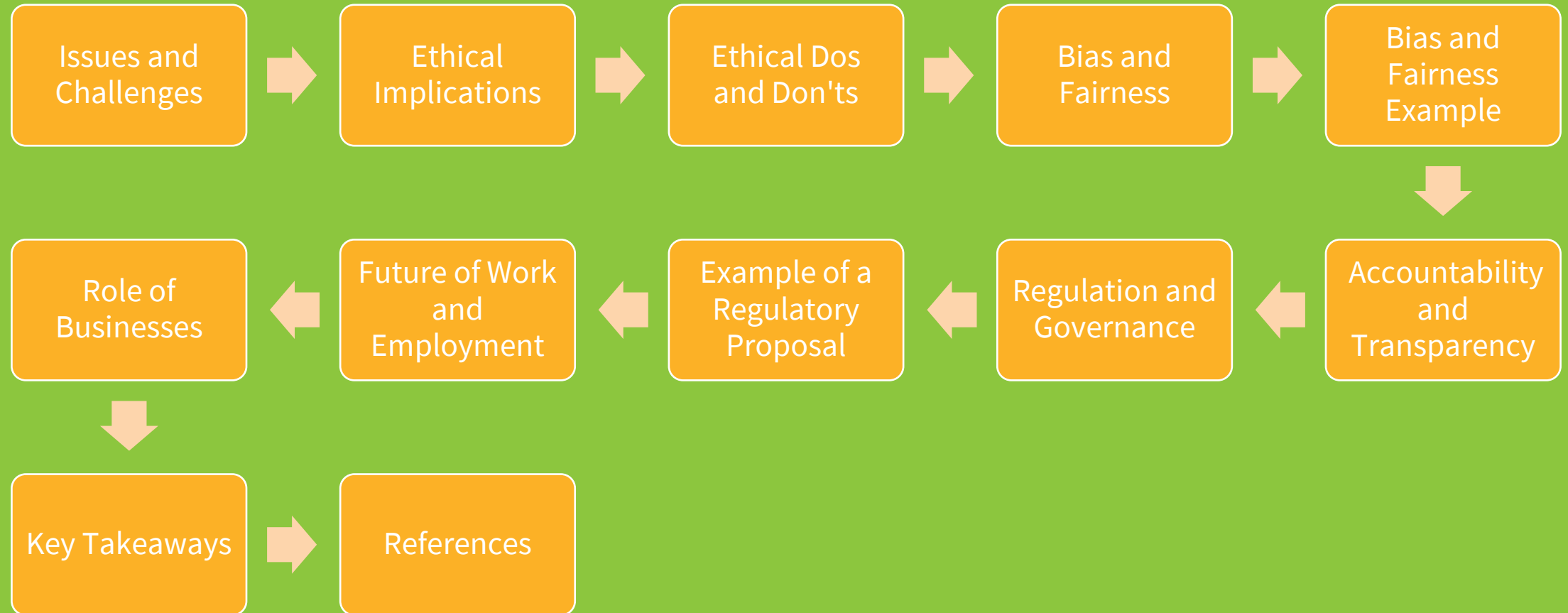
It is necessary to foster a **responsible AI landscape** that prioritizes **human well-being** and **societal benefits**



Adopt AI with a mindful approach, by recognizing both its potential challenges and risks, then acting responsibly to follow best practices to fully leverage its benefits.



Agenda



Issues and Challenges

SITUATION:

- Generative AI is an emerging technology that has rapidly gained popularity.
- Due to its capabilities, generative AI poses unique and extensive risks compared to traditional technologies.
- These risks have raised genuine concerns among individuals and institutions about its responsible deployment.
- Ethical, security, and societal impacts are at the forefront of concerns policymakers, experts, and institutions have raised.

EXAMPLES:

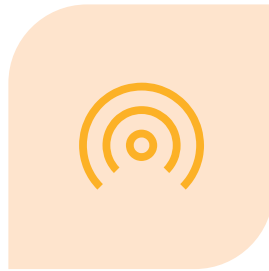
- Lack of truthful results that the audience can trust
- Misuse for malicious purposes if in the wrong hands
- Incorrect or misleading information leading to harmful decisions and outcomes
- Lack of transparency and accountability to garner confidence in the results

To mitigate these risks, it is essential to implement responsible AI development practices, transparency measures, and appropriate regulations.

Prevalent Issues



ETHICAL
IMPLICATIONS



BIAS AND FAIRNESS



ACCOUNTABILITY
AND TRANSPARENCY



REGULATION AND
GOVERNANCE

Ethical Implications

DEFINITION

AI applications and algorithms must be developed based on key ethical principles. These ethical guardrails result in truthful information and outcomes that benefit society and engender trust and confidence in the AI solutions

EXAMPLES OF ISSUES

Human Rights and Dignity

- AI-generated content can be used to create fake news and harmful misinformation, amplifying the spread of false information

Privacy and Security

- Generative AI requires vast amounts of data, including copyrighted material and private information. Its misuse can lead to legal challenges

Peace and Safety

- Unchecked AI systems have the potential to interfere with basic functions that promote the safe, peaceful existence of individuals and communities

Environment and Sustainability

- Unchecked AI systems have the potential to not comply with environmental laws and regulations impairing climate and quality of life

Ethical Solutions - Dos and Don'ts

DO

- **Adopt and Enforce Ethical Standards** – Follow the principles of Human Rights and Dignity, Privacy and Security, Peace and Safety, and Environment and Sustainability in all AI solutions.
- **Prioritize Transparency:** Ensure AI systems are transparent, explainable, and provide insights into their decision-making process. Continually adjust to eliminate false information and hallucinations.
- **Enforce Human Oversight:** Integrate human mechanisms to review and validate AI-generated outputs, especially when the results impact people and society
- **Enforce compliance:** Ensure all AI algorithms comply with regulations and laws related to human rights and the environment
- **Promote Privacy:** Respect individuals privacy rights and protect their personal data throughout the AI lifecycle
- **Consider social impact:** Assess the potential consequences of AI applications and take steps to minimize negative impacts

DON'T

- **Ignore Accountability:** Hold developers, organizations, and users accountable for the ethical use of AI systems
- **Operate in a Black Box:** Avoid using AI algorithms that lack transparency
- **Avoid Generalizations:** Treat each AI situation individually
- **Assume Accuracy:** Question AI-Generated Content
- **Prompt Harmful Actions:** Rigidly follow the ethical AI principles
- **In summary, since Generative AI is in its infancy, do not relax Quality Control processes, to ensure you are building responsible AI solutions.**

UNESCO's Recommendation of the Ethics of Artificial Intelligence – November 2021, outlines key principles and considerations in implementing ethical, responsible AI solutions

<https://unesdoc.unesco.org/ark:/48223/pf0000381137>

Bias and Fairness

DEFINITION

AI solutions and algorithms must be developed to eliminate both intentional and unintentional biases that compromise the principles of diversity, inclusivity, and non-discrimination. AI systems must result in fair outcomes.

CONSIDERATIONS

Identifying and Mitigating Bias

- Bias might be subtle and emerge in unexpected ways, making it difficult to detect
- Developing tools and methods to measure and quantify bias is an ongoing challenge

Efforts to Develop Fair Systems

- Techniques such as fairness-awareness training and adversarial testing are being explored
- Transparent AI design and interpretability efforts aim to assess decision-making processes

Bias in Algorithms and Data Sources Leads to Discrimination

- AI systems trained on biased data can perpetuate and amplify existing societal biases
- Discriminatory outcomes in areas such as hiring, lending, and law enforcement can result

Continual Monitoring and Improvement

- Enforce regular audits and bias assessments with feedback loops
- Adjust machine learning processes and algorithms to improve outcomes resulting in fairer systems, over time

Bias and Fairness

Example: Hiring Systems

- AI hiring systems may unintentionally favor certain demographics while discriminating against others
- Algorithms may tend to be biased towards candidates who attended prestigious universities, come from specific geographical locations, or have certain language patterns commonly associated with specific groups
- This can be attributed to biased training data and the use of historical hiring data that reflects past discriminatory practices
- If the data predominantly consists of candidates from a particular demographic, the AI system may learn to prioritize those characteristics
- Biased AI hiring systems can perpetuate inequalities in the job market, limiting opportunities for underrepresented groups
- This can lead to a lack of diversity in the workforce
- Biased hiring decisions can lead to implications such as harming an organization's reputation and exposing them to legal and ethical problems

Accountability and Transparency

DEFINITION

AI systems must be developed as a “white box” to support understanding of the algorithms’ decision-making process, providing clear explanations and audit trails in order to build trust and confidence with the audience.

Concept of Hallucinations

- AI can experience “hallucinations” where it applies data incorrectly, leading to erroneous outcomes, and the AI model remains unaware of its mistakes.

Distinguishing Bias and Hallucinations

- Bias in AI is often a result of intentional or careless application of data, while hallucinations arise from the model’s lack of awareness of its errors.

Identifying Unintentional Misuse

- Hallucinations highlight the need to address the unintentional misuse of AI and understand the factors contributing to such inaccuracies.

Shared Responsibility

- Stakeholders, including developers, companies, and regulatory bodies, must share the responsibility of preventing and addressing AI misinformation.

Ensuring Model Responsibility

- AI developers must establish mechanisms (e.g. audit trails, logs) with clear explanations to hold AI models accountable for their decisions and actions.

Human Oversight and Intervention

- Implementing human oversight (quality reviews) becomes essential to identify and correct AI model misinformation.

Liability Frameworks

- Developing liability frameworks and legal guidelines can help attribute responsibility and address potential harms caused by AI misinformation.

Transparent Auditing

- Regular audits and reviews of AI systems can help detect misinformation and assess their impact, allowing for necessary corrections and improvements.

Regulation and Governance

DEFINITION

New standards, regulations and laws must be enacted to create an AI ecosystem that minimizes the risks and enhances AI's value to society and business.

Importance of Government Regulation

- Governments must recognize the urgency of implementing regulatory and legal frameworks to address potential abuse and criminal AI applications.
- Government intervention is crucial to safeguard individuals and society from AI-related risks.
- Collaboration between governments, industry stakeholders, and experts to craft effective AI regulations and laws will result in practical guidelines.

Industry AI Standards and Guardrails

- Advocating for standards that require AI systems to be transparent and accountable for their decisions fosters trust and understanding among users and stakeholders.
- Establishing standards for AI applications to be responsible, safe, and beneficial adds real business and societal value while preventing harm from malicious or false information.
- Balancing innovation with the establishment of necessary safeguards is essential in order to mitigate risks.

Corporate Governance

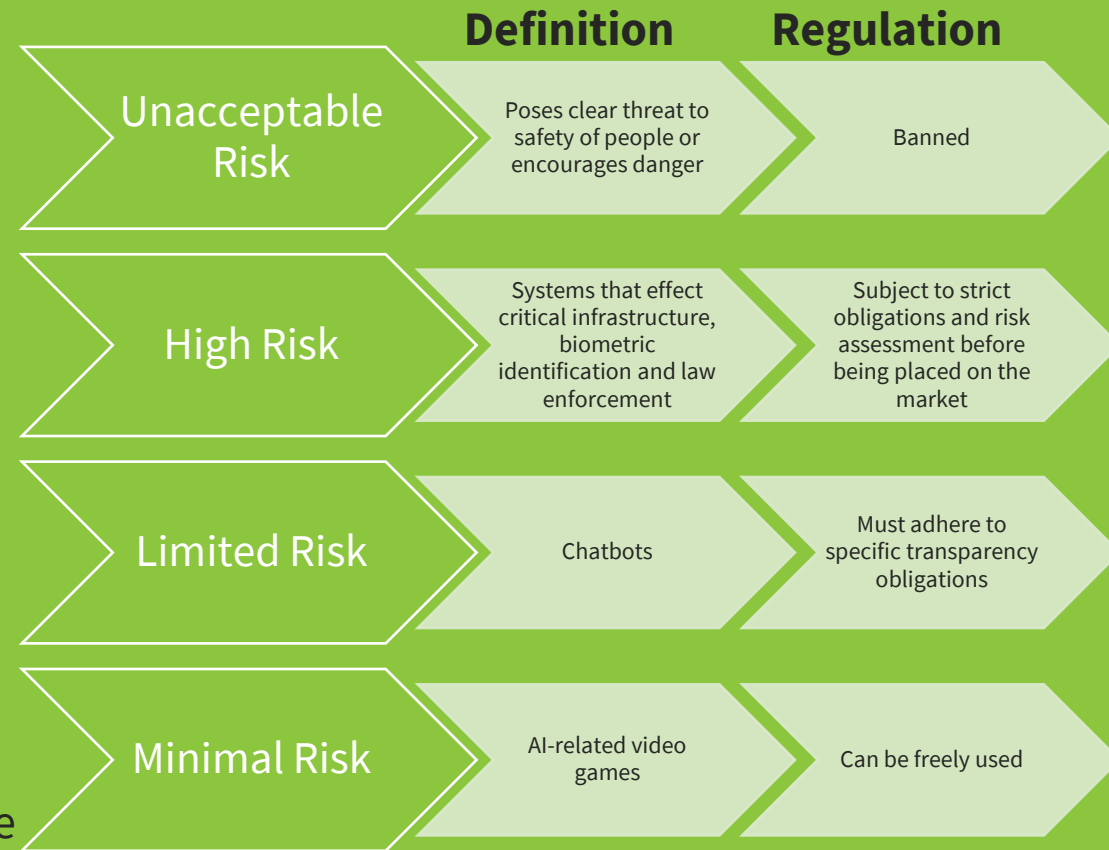
- Companies must respond internally to comply with the regulations, such as create corporate standards where company policies align with AI ethics.
- In many countries, it is going to take a long time for the government to pass regulations and laws.
- During that gap period it will be the companies' responsibility to self govern by creating internal AI Councils and specific jobs such as Chief AI Officer.

Example Regulatory Solution: The Proposed EU AI Act

- Europe is a leader putting AI regulations into place as the EU is working on a Regulatory Framework
- The Regulatory Framework aims to establish clear requirements and obligations for AI developers, deployers, and users, ensuring AI trust and safety
- It addresses risks associated with AI systems that may create undesirable outcomes
- It defines four levels of risk: unacceptable risk, high risk, limited risk, and minimal or no risk -- each level contains specific requirements

High Risk Assessment Process

1. A high risk system is developed
2. It undergoes a conformity assessment
3. It is registered in an EU database of AI systems
4. A declaration of conformity must be signed, then the system can be placed on the market



Being the first, this could serve as a framework for future regulations in other countries

Future of Work and Employment

How will AI impact the workforce – will it eliminate jobs or create jobs?

- AI will lead to increased productivity and specialization in job roles emphasizing “human skills” like creativity, problem solving, and quantitative abilities
- Economic growth from AI will not be evenly distributed, with low-skilled demographics facing additional strain due to job vulnerability and insecurity
- AI will benefit labor in high-growth fields like healthcare, complementing job roles that require highly-skilled practitioners
- Some industries relying on standard routines may face job threats from AI-based automation
- Existing jobs are going to change to incorporate AI and new AI-only jobs will also emerge
- Enhancing corporate training and transitioning workers to new skills will be imperative during this time

Example: Impact on Healthcare Industry

- Specialized AI and machine learning are being utilized in the healthcare industry to improve diagnoses and treatment options
- Precision medicine uses AI algorithms to assess previous patient information and recommend the most effective treatment options
- Deep learning algorithms help identify potentially cancerous lesions and anomalies in radiology pictures, aiding in early diagnosis and preventing the spread of malignancies and chronic diseases.
- AI tools in healthcare enable professionals to allocate their time more efficiently, leading to better patient care and potentially creating more specialized job roles

This illustrates how AI can positively impact certain industries such as healthcare, complementing the expertise of human professionals and improving overall outcomes

Role of Business

Driving principles that businesses must follow when developing and deploying AI solutions:

Human Intervention

Adopt a policy of incorporating human intervention in critical AI decisions, human review helps identify potential biases and misinformation, thereby preventing blind acceptance of AI-generated outcomes

Select and Audit Vendors

Create criteria for selecting vendors with robust security measures and ethical guidelines
Conduct rigorous audits ensuring compliance with established controls and regulatory standards

Institute Internal Procedures, Tools, and Training

Establish internal policies that guide AI development and deployment, these policies outline ethical principles, responsible use, and adherence to applicable regulations and laws
Develop internal AI policies and governance outlining responsible AI use and guidelines allowing for development and oversight controls
Evaluate job enhancements due to AI and provide employees with comprehensive training on AI usage, development and ethics

By focusing on these areas, business leaders can play a significant role in ensuring the prudent control of AI, protecting their businesses and maximizing the value AI delivers, responsibly

Key Takeaways

- The world's **ability to address these AI issues** is going to have a serious impact on the **level and speed of the adoption of this technology**
- Right now, Generative AI is at a **nascent stage, with unbridled, uncontrolled growth** – like the Wild West
- The challenge of potential **misuse and abuse** in AI is akin to cybersecurity concerns, necessitating the **implementation of ongoing controls and new solutions to mitigate such risks effectively**
- A **specialized segment of the security industry will emerge** to ensure that AI adheres to **established standards, regulations, and societal expectations**, as it cannot evolve naturally without deliberate oversight
- The evolution of AI and the **measures taken to address its challenges** are part of the **natural lifecycle** observed in the development of every technology
- **Companies must play a key role driving internal risk mitigation**, ahead of the pace of government intervention
- By adapting these **proactive measures and solutions**, AI can flourish and **unlock new opportunities and benefits**, fostering growth for businesses and society as a whole

References

- 1) “Artificial Intelligence: Examples of Ethical Dilemmas.” *UNESCO.Org*, www.unesco.org/en/artificial-intelligence/recommendation-ethics/cases#:~:text=But%20there%20are%20many%20ethical,an d%20privacy%20of%20court%20users. Accessed 30 July 2023.
- 2) Guinness, Harry. “Ai Ethics: The Ethical Issues of Artificial Intelligence.” *Zapier*, 22 Mar. 2023, zapier.com/blog/ai-ethics/.
- 3) ain, Pragma. “Ai and the Future of Work in the United States.” *American University*, 18 Feb. 2021, www.american.edu/sis/centers/security-technology/ai-and-the-future-of-work-in-the-united-states.cfm.
- 4) Novelli, Claudio, et al. “Accountability in Artificial Intelligence: What It Is and How It Works - AI & Society.” *SpringerLink*, 7 Feb. 2023, link.springer.com/article/10.1007/s00146-023-01635-y.
- 5) “Regulatory Framework Proposal on Artificial Intelligence.” *Shaping Europe’s Digital Future*, digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai. Accessed 30 July 2023.
- 6) Tom Wheeler, Blair Levin, et al. “The Role of Corporations in Addressing AI’s Ethical Dilemmas.” *Brookings*, 9 Mar. 2022, www.brookings.edu/articles/how-to-address-ai-ethical-dilemmas/.
- 7) “What Business Leaders Must Do to Avoid Extreme AI Risks.” *I by IMD*, 5 June 2023, www.imd.org/ibyimd/innovation/what-business-leaders-must-do-to-avoid-extreme-ai-risks/.
- 8) *Unesdoc.Unesco.Org*, unesdoc.unesco.org/ark:/48223/pf0000381137. Accessed 10 Aug. 2023.