



Ethics in Technology

- 1 Introduction
- 2 Artificial Intelligence
- 3 Fraudulent Activities in Tech
- 4 Regulation – Global and Regional
- 5 Other Ethical Considerations



Introduction

Ethics in Technology

Ethics in technology refers to the application of moral principles and values in the design, development, and deployment of technological solutions. It encompasses a wide range of considerations, from data privacy and security to the broader societal impact of technological innovations. At its core, it seeks to ensure that technology serves humanity in a way that is fair, responsible, and beneficial.

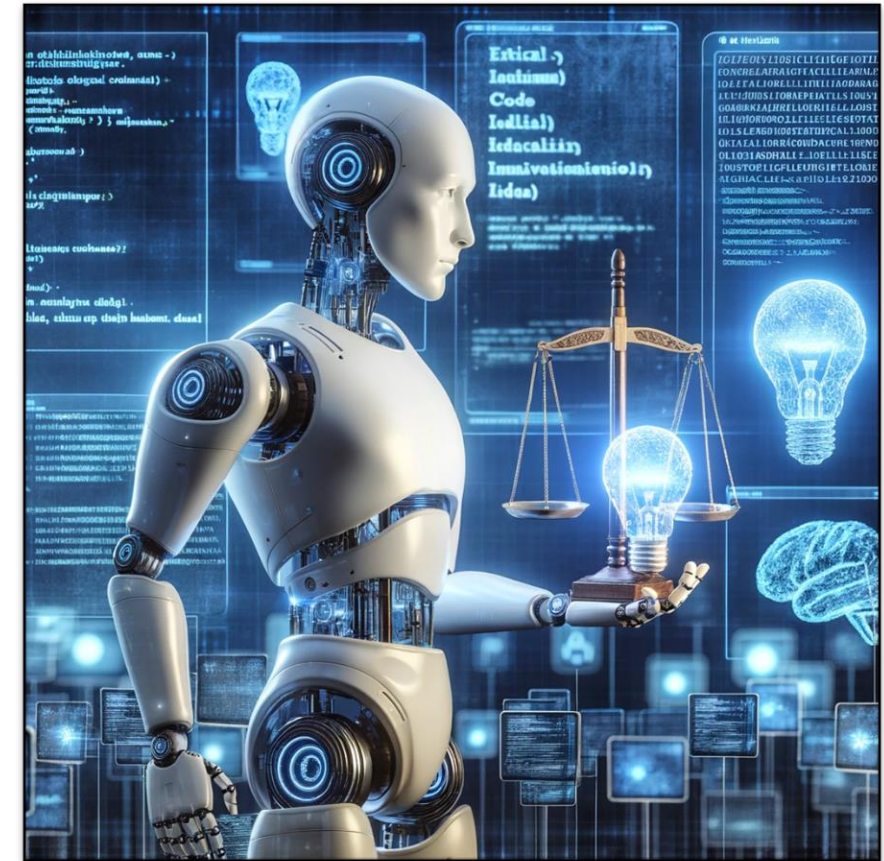


Importance of Ethical Considerations in Tech

As technology becomes increasingly integrated into every aspect of our lives, the importance of ethical considerations grows. Ethical technology aims to:

- Protect Individual Rights
- Foster Trust In Technology
- Prevent Harm

In an era where technology can influence everything from individual choices to global politics, ensuring that these powerful tools are used responsibly is crucial.



Fundamental Ethical Principles in Technology



Principle of
Beneficence (Doing
Good)



Principle of
Nonmaleficence
(Avoiding Harm)

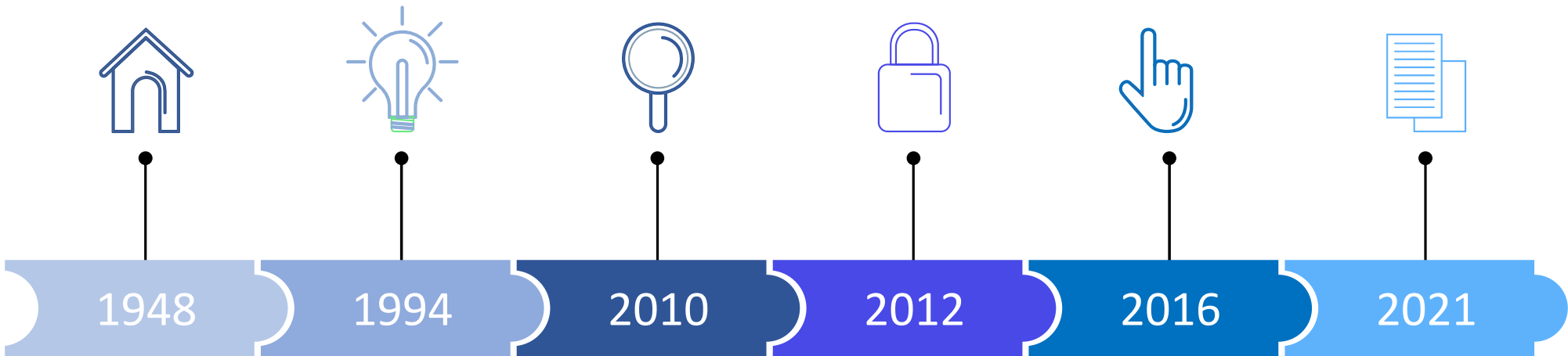


Principle of Autonomy
(Respecting User's
Rights)



Principle of Justice
(Fairness and Equity)

Evolution of Privacy Protection



UN Declaration of Human Rights establishes the right to privacy.

Netscape releases browser that makes online tracking possible for the first time.

Facebook changes the default setting for user profiles from 'private' to 'public'.

A US survey shows 70% customers don't trust social media with their data.

EU enacts the GDPR – biggest change in 20+ years.

Council of the EU agrees on negotiating mandate for ePrivacy regulation.



Artificial Intelligence

What is Artificial Intelligence (AI)?

It is the simulation of human intelligence processes by machines, especially computer systems. These processes include:

- Learning
- Reasoning
- Self-correction

It involves creating algorithms that allow computers to perform tasks that typically require human intelligence.

This can range from simple tasks like recognizing speech patterns to more complex ones like making decisions or solving problems.



Discuss and jot down the pros and cons for using AI

Pros

Cons

Overview of How to Create an AI System

STEPS TO BUILD AN AI SYSTEM



RELEVANT

relevant.software

Key AI Concepts

Algorithm

Machine
Learning

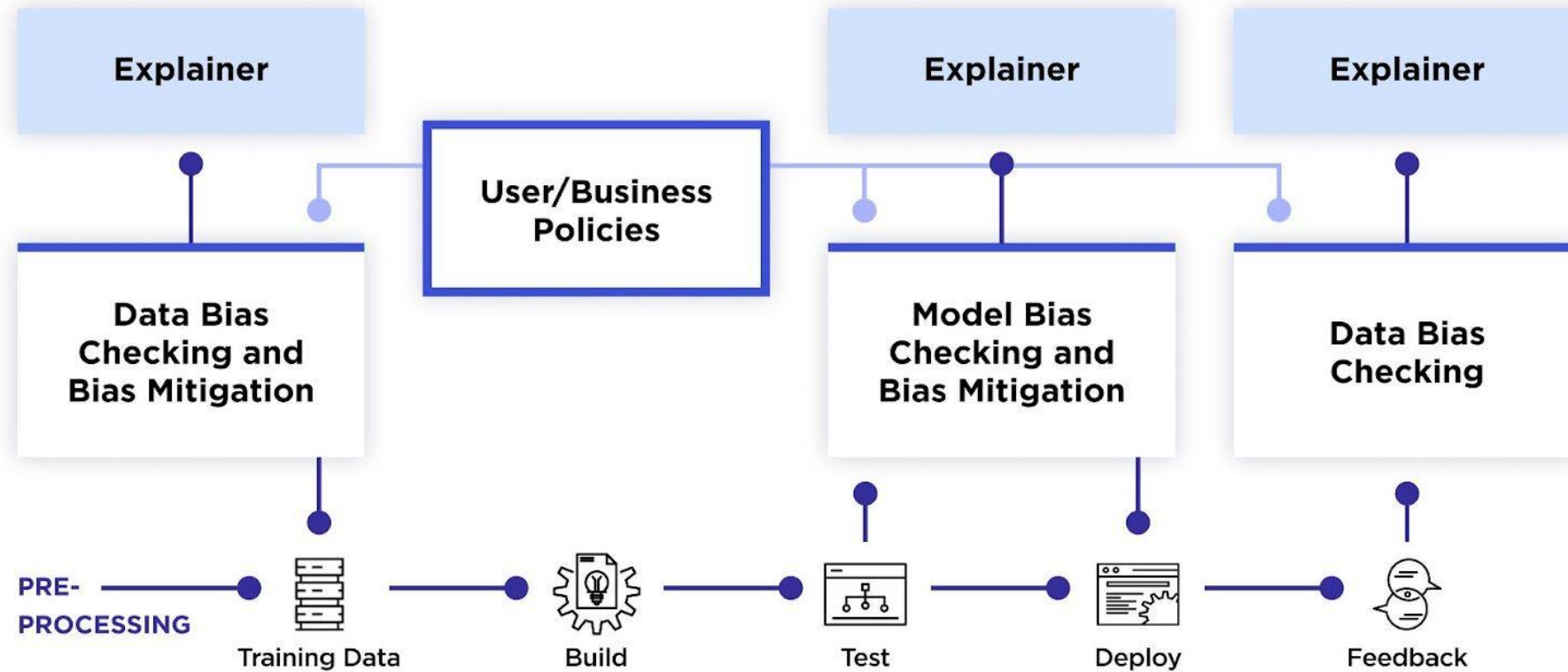
Deep
Learning

Bots

Autonomous
Systems

MITIGATING BIAS THROUGHOUT THE AI LIFECYCLE

Source: IBM

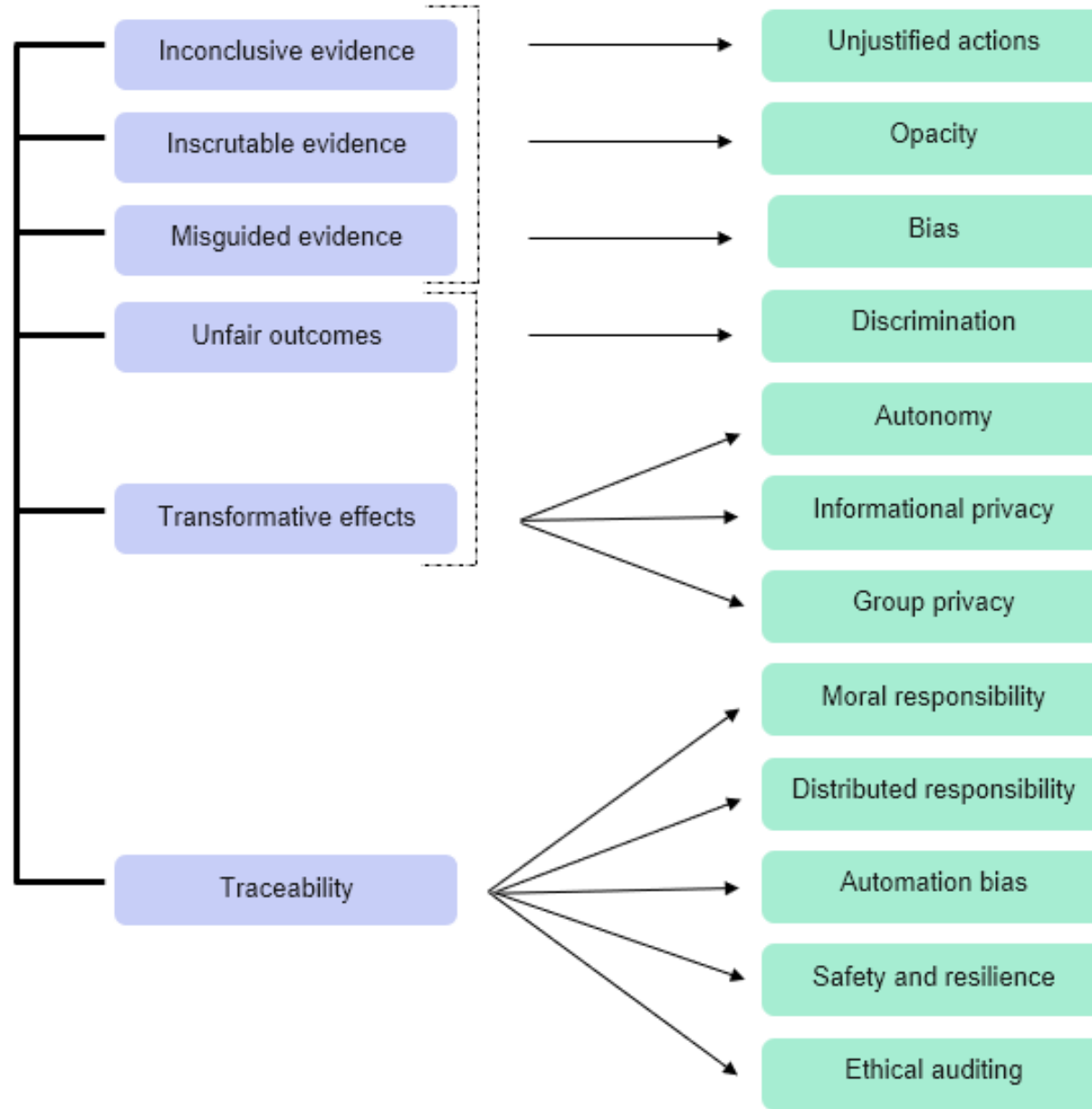


RELEVANT

relevant.software

Common Ethical Challenges In AI

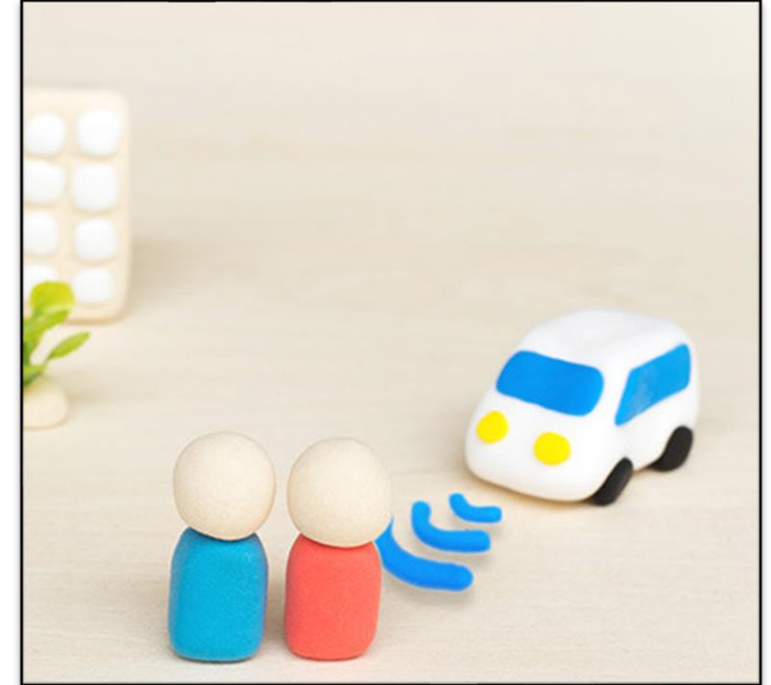
Types of Concerns



Ethical Challenges

Example of Ethical Dilemmas in AI – Self Driving Cars

- The autonomous car must undertake a considerable amount of training in order to understand the data it is collecting and to be able to make the right decision in any imaginable traffic situation.
- Moral decisions are made by everyone. When a driver chooses to slam on the brakes to avoid hitting a jaywalker, they are making the moral decision to shift risk from the pedestrian to the people in the car.
- Imagine an autonomous car with broken brakes going at full speed towards a grand-mother and a child. By deviating a little, one can be saved.
- This time, it is not a human driver who is going to take the decision, but the car's algorithm.



Who would you choose, the grandmother or the child? Do you think there is only one right answer?

Pause Giant AI Experiments: An Open Letter

Future of Life Institute published an Open Letter in May 2023 to immediately pause (for at least 6 months) the training of AI systems which may be more powerful than GPT-4.

Some notable signatories include:

- Elon Musk (CEO of SpaceX, Tesla and X)
- Steve Wozniak (Co-Founder, Apple)
- Chris Larsen (Co-Founder, Ripple)
- Tom Gruber (Siri Designer)
- Craig Peters (CEO, Getty Images)
- Rachel Bronson (President, Bulletin of Atomic Scientists)



Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

[View this open letter online.](#)

Published	PDF created	Signatures
March 22, 2023	May 5, 2023	27565

AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research¹ and acknowledged by top AI labs.² As stated in the widely-endorsed [Asilomar AI Principles](#), *Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources.* Unfortunately, this level of planning and management is not happening, even though recent months have seen AI labs locked in an out-of-control race to develop and deploy ever more powerful digital minds that no one – not even their creators – can understand, predict, or reliably control.

Contemporary AI systems are now becoming human-competitive at general tasks,³ and we must ask ourselves: *Should we let machines flood our information channels with propaganda and untruth? Should we automate away all the jobs, including the fulfilling ones? Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders.*

Key Points from the Open Letter

The development of powerful AI systems poses significant risks to society.

We must ensure that the effects of AI systems are positive, and their risks are manageable.

Powerful AI systems should only be developed once we are confident in their safety.

OpenAI has called for independent review and limits on the rate of growth of compute used for creating new models.

We call on all AI labs to pause for at least 6 months the training of AI systems more powerful than GPT-4.

Key Points from the Open Letter

AI developers must work with policymakers to accelerate the development of robust AI governance systems.

These systems should include new regulatory authorities, oversight and tracking of AI systems, provenance and watermarking systems, auditing and certification, liability for AI-caused harm, and public funding for technical AI safety research.

Humanity can enjoy a flourishing future with AI, but we must engineer these systems for the clear benefit of all and give society a chance to adapt.

Why bestselling author Malcolm Gladwell distrusts Artificial Intelligence

Malcolm Gladwell, in the article, expresses a nuanced view on the use of AI in journalism and writing. Here are the key points reflecting his distrust and caution towards AI:

Preservation of Human Element in Storytelling:

- Gladwell emphasizes the importance of writers and journalists maintaining the human touch and emotional connection in their storytelling. He suggests that AI should aid in focusing on what humans do best - telling compelling stories that capture imagination.

AI as a Tool, Not a Replacement:

- He acknowledges AI's role in facilitating research, making the process more efficient, but firmly believes it should enhance, not replace, journalists. This implies a concern that over-reliance on AI could diminish the role of human journalists.

Distinguishing Machine vs. Human Expression:

- Gladwell stresses the importance of recognizing the qualitative differences between machine-generated content and human expression. This indicates his belief that AI and human writing have distinct qualities that should be discernible.

Risk of Errors Introduced by AI:

- He points out the risk that AI might introduce errors into our understanding of subjects. While acknowledging human fallibility, Gladwell seems more concerned about the potential for greater mistakes by AI.

Why bestselling author Malcolm Gladwell distrusts Artificial Intelligence

Importance of Various Communication Channels:

- Discussing the relevance of journalists, he highlights the importance of effective communication through various channels, indicating a belief that AI should not overshadow traditional forms of storytelling.

Definition of Successful Work:

- Gladwell describes successful work as hard, engaging, and independent. This reflects a view that success is not about external validation (fame, money), which could be seen as a critique of AI's potential focus on efficiency and output over depth.

Value of Mistakes and Learning:

- He emphasizes learning from mistakes, suggesting a preference for the human learning process over AI, which might not have the same capacity for nuanced learning from errors.

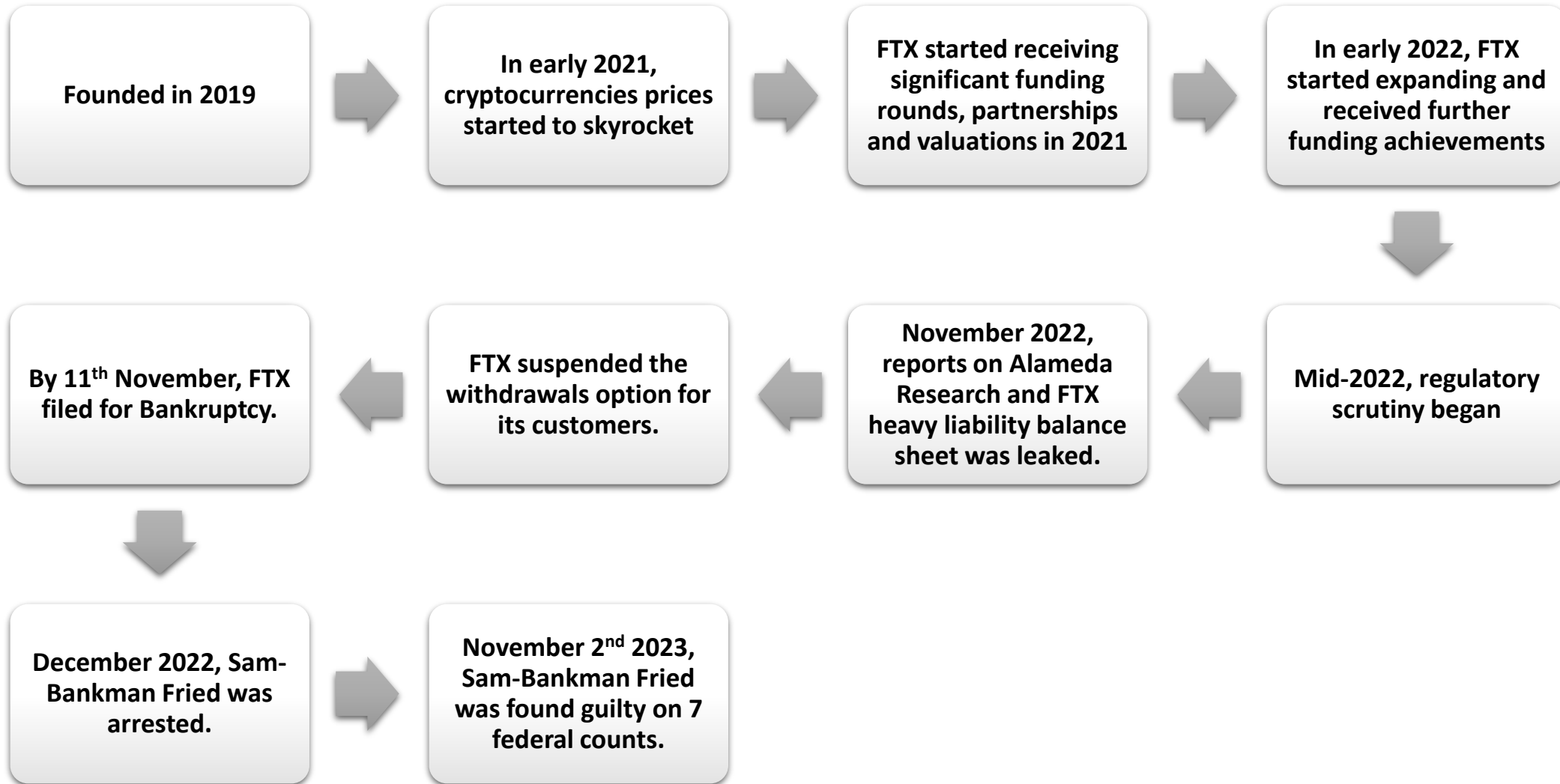
Storytelling and Patience:

- Gladwell distinguishes between anecdotes and stories, emphasizing that great stories require time to develop. This may imply a skepticism towards AI's ability to replicate the depth and complexity of human storytelling.



Fraudulent Activities in Tech

FTX Cryptocurrency Exchange: The Case of Sam-Bankman Fried



Group Activity: The Case of Sam-Bankman Fried

In groups of 5-6, find the following:

1. FTX Debtors released their first report on April 9, 2023. What points were mentioned in that report as reasons of FTX's failure?
2. What role did Sam-Bankman Fried's parents play in this case?
3. How was Binance involved in FTX?
4. What lessons can be learned from FTX's failure?

Group Exercise: Presentation

In groups, research and present at least 2-3 companies that have been in a controversy that led to either huge fines or bankruptcy.

Below are a few examples to help you start:



In your presentation:

- Explain what the company's core operations were.
- What controversy was it involved in?
- How could such a scenario be avoided in the future?
- What ethical lessons did you learn from each company?

ETHICS

Regulation – Global & Regional

Recent regulations and initiatives reflect a growing recognition of AI's impact and the need for comprehensive governance to ensure its responsible and ethical use.

President Biden recently released an *Executive Order on Artificial Intelligence* that include the following crucial points:

- **Strengthening AI Governance:** Involves designating Chief AI Officers, establishing AI Governance Boards, and expanding AI use reporting.
- **Advancing Responsible AI Innovation:** Federal agencies are directed to develop AI strategies, remove barriers to AI use, and explore the use of generative AI with adequate safeguards.
- **Managing Risks from AI Use:** Mandates safeguards for AI uses impacting public rights and safety, including AI impact assessments, testing, and public consultation.

Automated systems and algorithms have been massively useful in helping farmers grow food more efficiently, predicting storm paths and identifying diseases in patients.

But they also have ethical issues such as being unsafe, ineffective or biased.

Algorithms used in hiring and credit decisions have been found to reflect and reproduce existing unwanted inequities or embed new harmful bias and discrimination.

Unchecked social media data collection has been used to threaten people's opportunities, undermine their privacy, or pervasively track their activity—often without their knowledge or consent.

Regulation in Technology – United States

The White House Office of Science and Technology Policy has identified five principles that should guide the design, use, and deployment of automated systems to protect the public in the age of artificial intelligence.



Safe and Effective
Systems



Algorithmic
Discrimination
Protections



Data Privacy



Notice and
Explanation



Human
Alternatives,
Consideration, and
Fallback

Summary of Each Section – Safe & Effective Systems

Automated systems should be developed with diverse input, undergo rigorous testing and monitoring for safety and effectiveness, and adhere to specific standards, with outcomes including the possibility of non-deployment or removal if unsafe.

Systems should avoid endangering safety, protect from unintended harms, and prevent misuse of data, with independent evaluations confirming safety and effectiveness made public.



Summary of Each Section – Algorithmic Discrimination

Systems should be designed to prevent algorithmic discrimination based on legally protected characteristics, with proactive measures like equity assessments and disparity testing.

Independent evaluations, including algorithmic impact assessments, should be performed to confirm protections against discrimination and results should be publicly available.



Summary of Each Section – Data Privacy

Design and deployment of automated systems should prioritize user privacy by default, limit data collection to necessity, and respect user decisions regarding data use.

Enhanced protections should be applied in sensitive domains, with restrictions on surveillance and access to reports confirming respect for privacy and data decisions.



Summary of Each Section – Notice and Explanation

Users should receive clear, plain language information about the use and functioning of automated systems, including updates and explanations of outcomes.

Systems should provide technically valid and meaningful explanations, with public reporting on system clarity and quality of notices and explanations.



Users should have the option to opt out of automated systems in favor of human alternatives, with access to timely human intervention for errors or appeals.

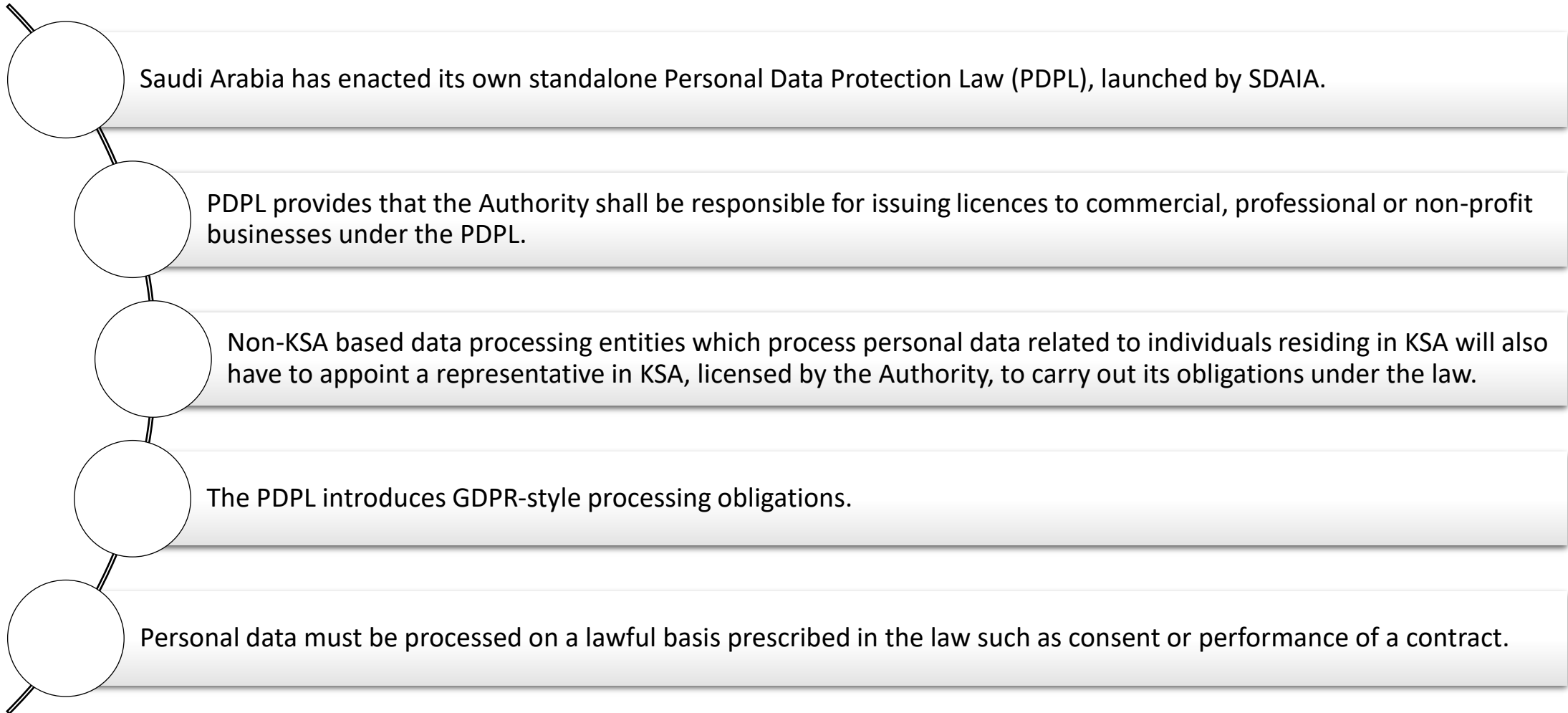
Human governance processes should be equitable and effective, especially in sensitive domains, with public reporting on their outcomes and effectiveness.



Regional Regulations – KSA and UAE

- Saudi Arabia and the United Arab Emirates have both introduced significant regulations regarding data usage and privacy.
- The Saudi Data & AI Authority (SDAIA) is the competent authority in the Kingdom concerned with data and AI including big data.
- The UAE federal data protection law provides greater clarity on what is permissible in terms of the collection, processing, review and transfer of personal data in onshore UAE.

Regulations in KSA



The National Data Regulations sets out principles for dealing with personal data, which include:

The purpose of collection of personal data should be known

The data subject's consent should be sought for collection and processing

Collection of personal data shall be limited to what is necessary for the purpose

Personal data should be used for the agreed purpose only, and data shall be protected against breach

Federal Decree Law No. 45 of 2021

- This law constitutes an integrated framework to ensure the confidentiality of information and protect the privacy of individuals in the UAE.

First Omnibus Personal Data Protection Law

- The law came into force on 2 January 2022. It is the first comprehensive personal data protection law in the UAE, aligning with many of the core principles and concepts found in the GDPR. Executive regulations were expected to be published within six months of its enforcement.

Greater Transparency and Harmonization

- The introduction of this landmark federal data protection law on 20 September 2021 has moved the UAE towards greater transparency and harmonization with other jurisdictions in the region.

Standardization of Data Protection

- The Personal Data Protection Law (PDPL) standardizes data protection for companies operating in and/or serving data subjects who are residents of the UAE, aligning with global privacy regulation best practices

Regulations in UAE

The United Arab Emirates (UAE) has a comprehensive federal Personal Data Protection Law (PDPL) much like GDPR in the EU. Key aspects of the UAE's data protection law include:

Consent Requirement

- The PDPL mandates that businesses and entities must obtain explicit consent from individuals before collecting and processing their personal data.

"Opt-in" Consent

- Consent must be obtained through a specific, clear, and unambiguous positive statement or action by the data subject.

Transparency

- Businesses are required to inform individuals about the purposes for which their data is being collected and how it will be used.

Technology

Fintech firm Wise fined \$360,000 by Abu Dhabi regulator

Abu Dhabi Global Market's Financial Services Regulatory Authority (FSRA) "found that Wise did not establish and maintain adequate AML systems and controls to ensure full compliance with its AML obligations".

ETHICS

Other Ethical Considerations

Doing the right thing for the Environment



Hiring Bias

Amazon's AI-powered hiring tool, implemented in 2018, showed gender bias.

Trained on resumes and employment data predominantly from male candidates, the AI favored male candidates, perpetuating existing gender biases in hiring.

In effect, Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain."

Eventually, Amazon scrapped the secret AI recruiting tool.



Lending and Credit Scoring

There's a significant correlation between FICO credit scores and race, impacting credit and loan decisions.

For instance, white homebuyers have notably higher average credit scores compared to Black and Hispanic homebuyers.

This disparity affects the types of loans and interest rates available to borrowers.

The FICO logo is displayed in a large, bold, blue sans-serif font. The letters are thick and blocky. A registered trademark symbol (®) is located at the bottom right of the letter 'O'.

AI in healthcare has shown bias in treatment recommendations.

For example, the Framingham Heart Study cardiovascular risk score performed well for Caucasian patients but not for African American patients.

A 2019 study highlighted that algorithms used by health systems were racially biased, affecting care recommendations.

Millions of black people affected by racial bias in health-care algorithms.



Facial Recognition Technology

Facial recognition software has struggled with accurately detecting darker skin tones.

This was notably demonstrated in Joy Buolamwini's study at MIT.

Issues with such technology have led to difficulties in various applications, including online proctoring software and law enforcement, with some cases resulting in wrongful arrests.

Examination of facial-analysis software shows error rate of 0.8 percent for light-skinned men, 34.7 percent for dark-skinned women.



**Massachusetts
Institute of
Technology**

You have reached the end of the training. Please use the link and credentials below to complete your final assessment:

Link: www.emergingmarketft.com/elearning

Username: Your email address

Password: *Test123!* or Use the one if you changed it for the pre-course assessment



NBK - Ethics in Technology

