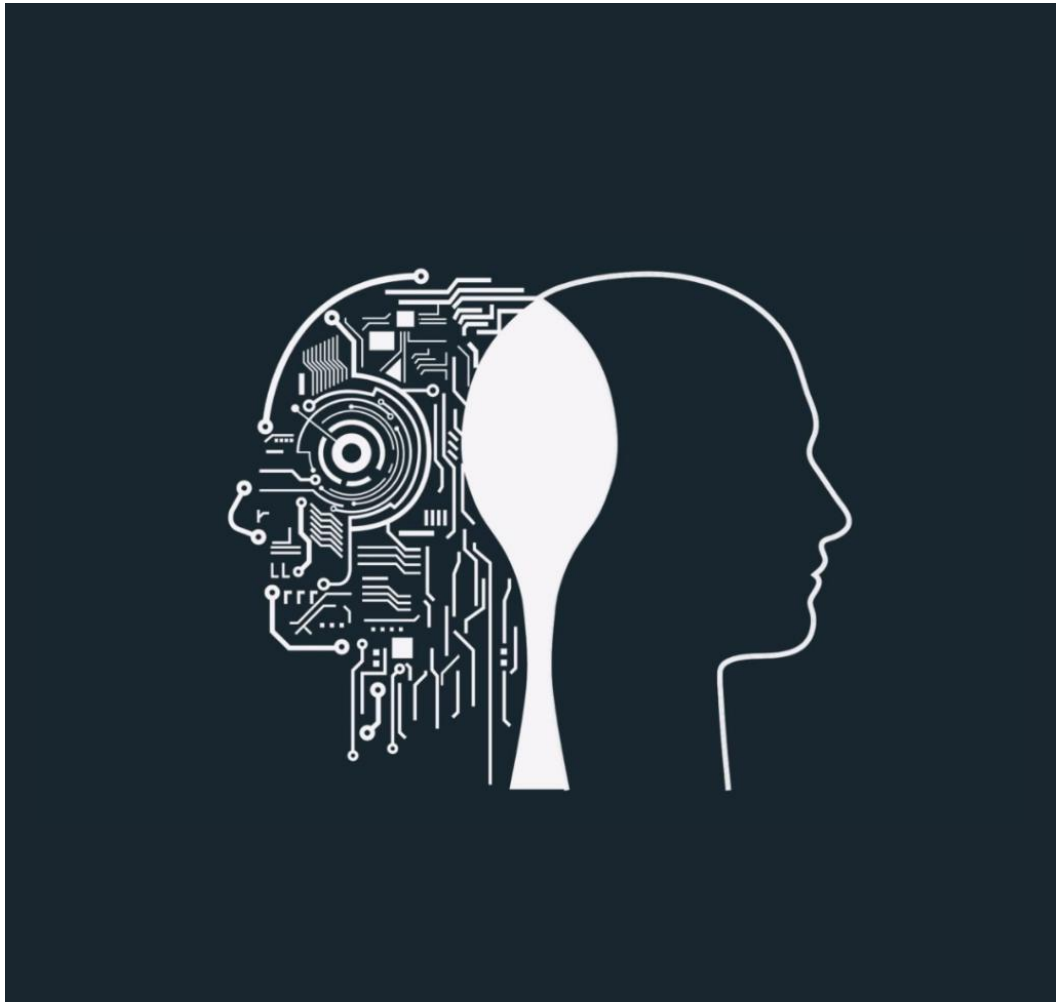


Ethical Use of AI

By:

Sammak Shinde

FS-Prime AI Group,

CDAC Pune

Agenda

- Bias and Fairness
- Governance and Regulations by Indian Govt.
- Privacy and Data Protection
- Accountability matrix
- Ensuring inclusivity
- Autonomous systems and safety
- Unethical use of AI in software applications
- Introduction to Ethical AI Tools/ services

AI BIAS AND FAIRNESS

1. Identify potential biases
2. Define fairness criteria
3. Audit training data
4. Mitigate bias in data
5. Evaluate algorithmic fairness

Credits:

Credits:
IDreamPost

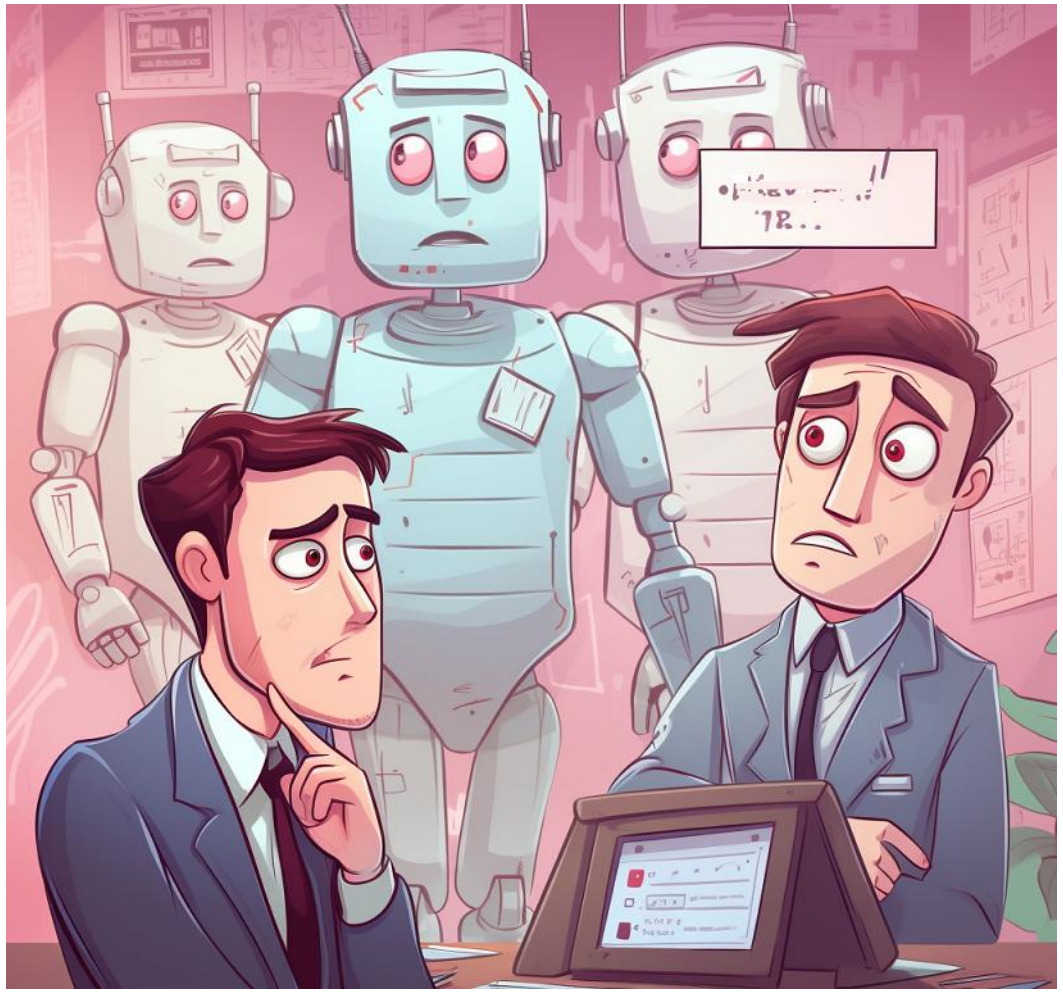
Governance and Regulations by the Indian Government

- IT Act, 2000
- DPDP Act, 2023
- NITI Ayog’s “National Strategy for Artificial Intelligence” , 2018
- India’s AI Act drafted by MEITY

Privacy and Data Protection

Credits: Secure Data Management

- AI models rely on Data.
- Ensures trust
- E.g. DPDP Act India, GPDR Act EU etc



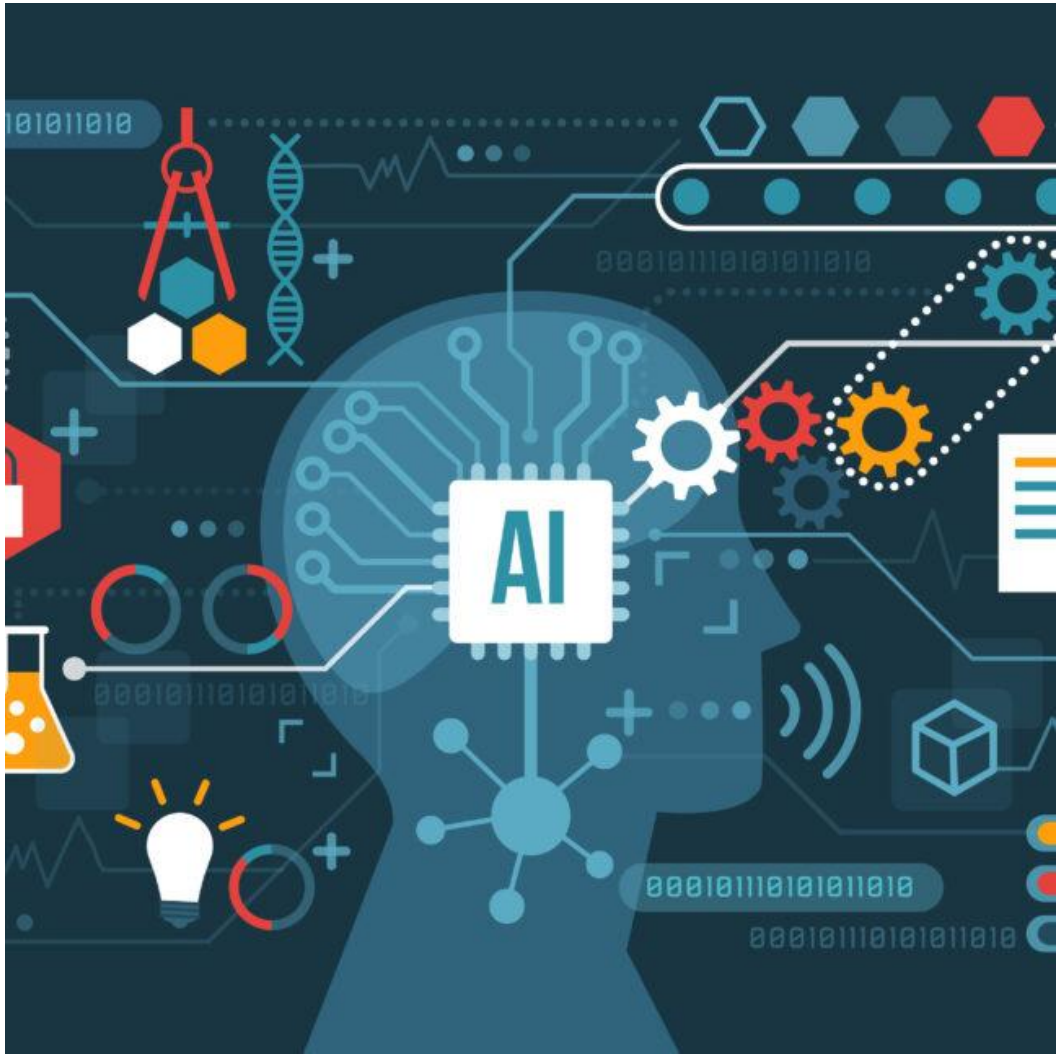
Credits: ESG HOLIST

Accountability Matrix

- AI Developers
- Different Organizations developing AI
- Government

Ensuring Inclusivity

- Need: Avoiding any marginalized group.
- All groups/stakeholders during Data Collection



Credits: STAT News

Autonomous Systems and Safety

Credits: Udacity

- Reliability & Robustness
- Decision Transparency
- Ethical Decision-Making
- Human Oversight & Trust

Unethical Use of AI in Software Applications

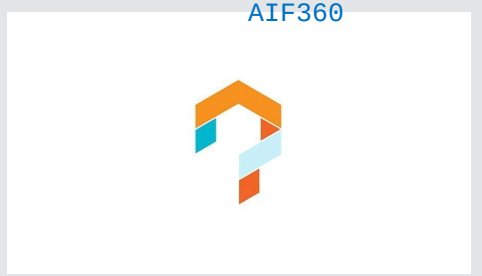
- E.g. Deepfakes, Manipulative Algorithm etc.
- Harm to Individual & Society

Credits: ChatGPT

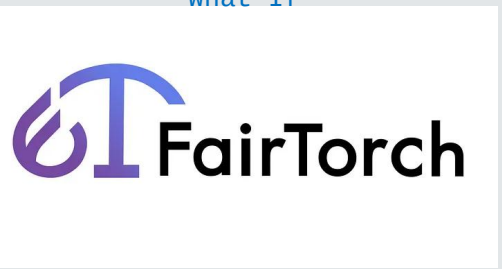
Introduction to Ethical AI Tools/Services

- AIF360 by IBM
- Google’s What If Tool
- FairTorch: A Pytorch Responsible AI Framework
- Microsoft Responsible AI ToolBox
- UNESCO’s AI Ethics Framework

Credits: AIF360



Credits: Google What-If



Credits: FairTorch

Credits: UNESCO

Thank You

References

- ChatGPT
- AIF360: https://aif360.res.ibm.com/
- Google What-If Tool: https://pair-code.github.io/what-if-tool/explore/