

Responsible AI Starts with Ethics: A Step-by-Step Primer

**Accountability is one of the
core Ethical principles of AI**

What is Accountability in AI Ethics?

Accountability means ensuring that humans remain responsible for AI systems

- their design, outcomes, and societal impact.

AI Accountability Answers

Decision Tracking



How are decisions tracked and explained?

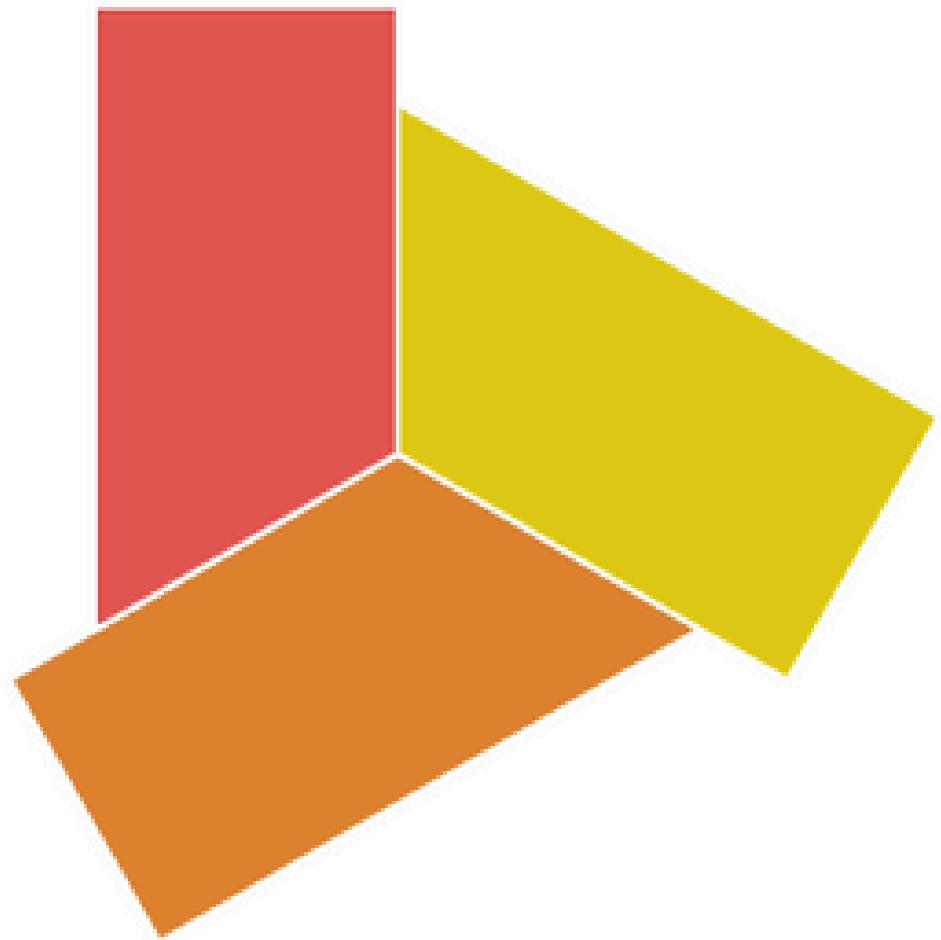
Explaining how AI decisions are monitored and justified.

Harm Mitigation



What steps are taken when harm occurs?

Outlining the procedures followed when AI causes harm.



Responsibility

Who is responsible when AI fails?

Determining who is accountable when AI systems fail.

Who is responsible if AI fails?



Key Legal Frameworks

Framework	Description & Application
Tort Law	Governs civil wrongs and remedies for harms caused by wrongful actions. In the U.S., tort law is largely state-based and can apply to AI-caused injuries. Plaintiffs must prove causation and that the harm was a foreseeable consequence of the defendant's conduct.
Product Liability	Injured parties can claim that an AI product was “defective” or failed to meet industry standards. This applies if, for example, a self-driving car malfunctions or a chatbot causes reputational harm. The plaintiff need not prove the exact technical cause, only that the defendant failed to meet a reasonable standard of care.
Negligence	Requires showing that a party (developer, manufacturer, operator) owed a duty of care and breached it, leading to harm. This is complicated by AI’s opacity and the difficulty of proving causation, especially when multiple actors are involved.
Vicarious Liability	In some cases, organizations may be held responsible for the actions of employees or agents deploying or managing AI, though this is complicated by the autonomous nature of AI systems.
EU AI Liability Directive (proposed)	The European Union has proposed shifting the burden of proof in high-risk AI cases, making it easier for victims to claim damages. These regulations are still emerging and not yet globally harmonized.

Organizational Framework Comparison

Characteristic



Leadership

NIST AI RMF

Govern

US GAO Framework

Governance

IEAI Practical Framework

Clarifies roles



Risk Identification

Map

Data

Ensures transparency



Risk Evaluation

Measure

Performance

Integrates stakeholders



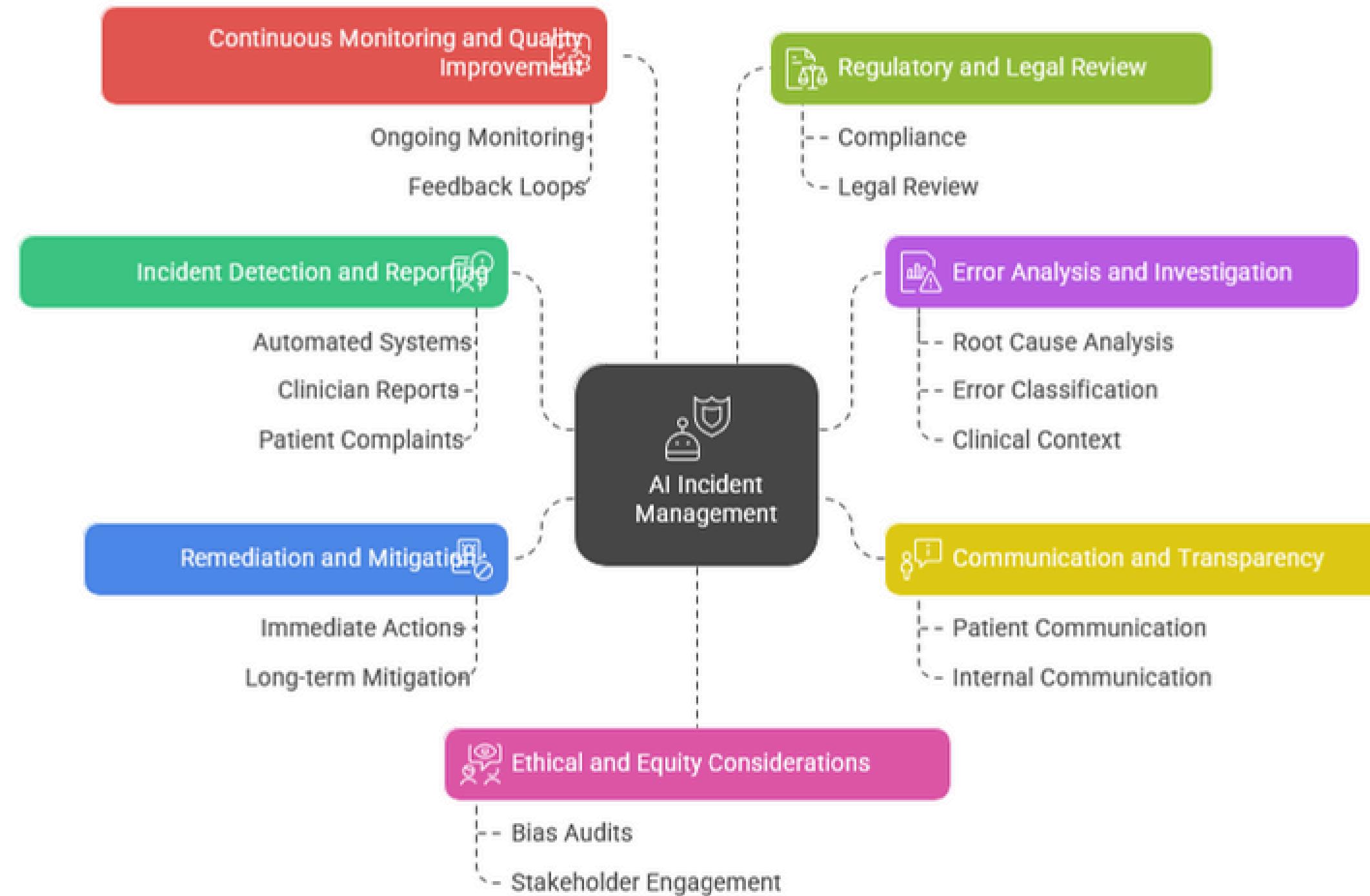
Risk Mitigation

Manage

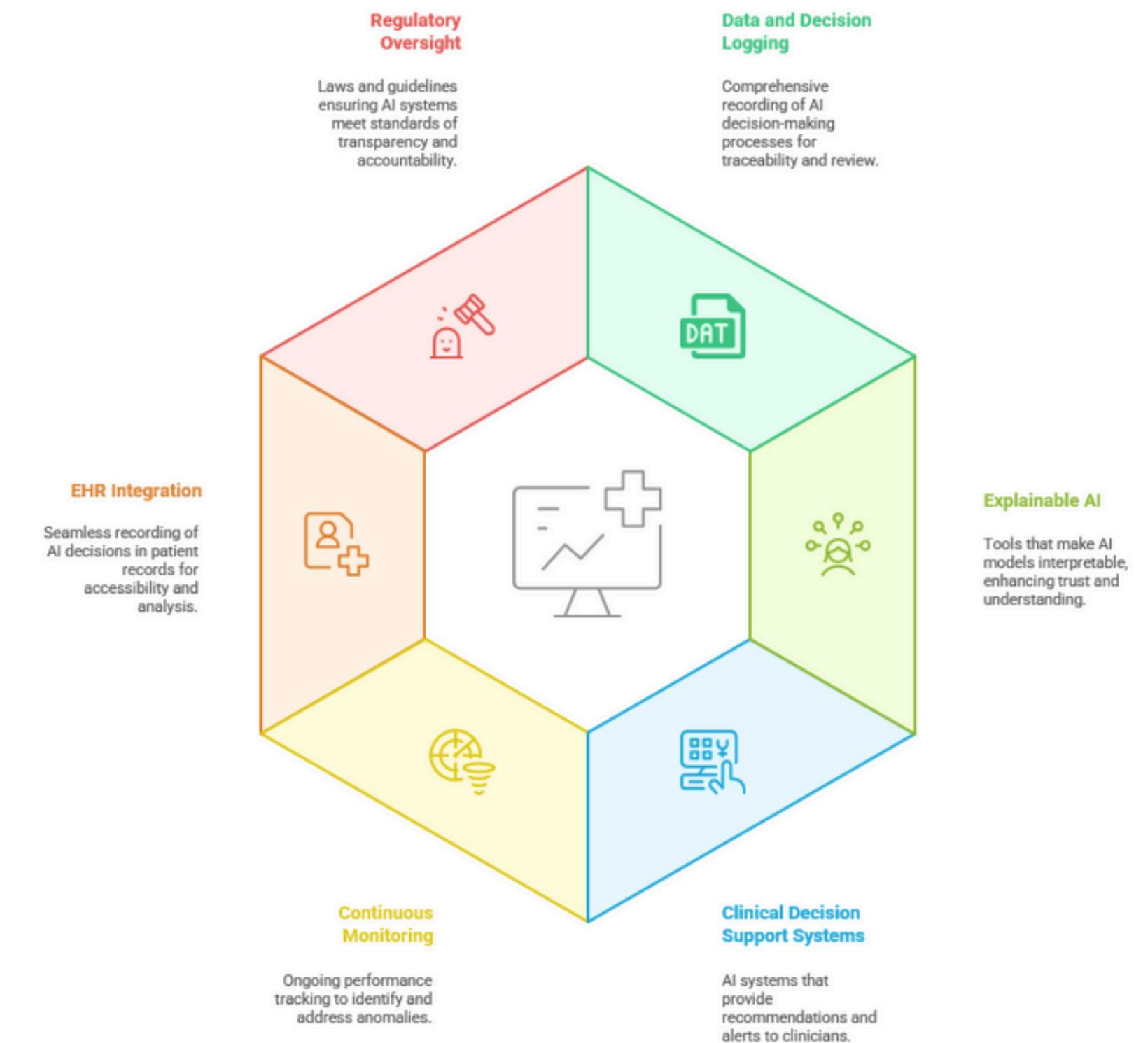
Monitoring

Integrates stakeholders

What steps are taken when harm occurs?



How are decisions tracked and explained?



Risks in AI Accountability

Risks include fairness, bias, security, safety, transparency, and unintended societal impacts.

Risk monitoring involves tracking relevant metrics and conducting regular audits to detect emerging issues.

Proactive risk management (e.g., prevention, bias detection tools, XAI) helps clarify responsibilities and prevent harm.

Key Principles of AI Accountability

Transparency	Make AI decision-making processes, data sources, and assumptions clear and understandable.
Responsibility	Assign clear roles for who is responsible at each stage (developers, providers, users).
Justification	Require explanations for decisions and actions taken by or with AI.
Mitigation	Identify, prevent, and address risks, such as bias, discrimination, or safety failures.
Continuous Monitoring	Regularly evaluate AI systems to ensure ongoing compliance and safety

Real world examples

Sector	Example of AI Accountability Challenge	Risk Addressed	Accountability Approach
Finance	Automated loan approvals causing biased outcomes	Fairness, discrimination	Algorithmic audits, transparency mandates
Customer Service	Chatbots giving misleading or offensive responses	Reputational, customer harm	Incident response, stakeholder mapping
Surveillance	Facial recognition misidentifying individuals	Privacy, wrongful identification	Oversight committees, impact assessments
Auditing	AI systems introducing bias in financial audits	Bias, compliance	Risk taxonomy, data quality

References and further reading

-

NIST AI Risk Management Framework (AI RMF)

U.S. government's gold standard for AI risk, trust, and accountability.

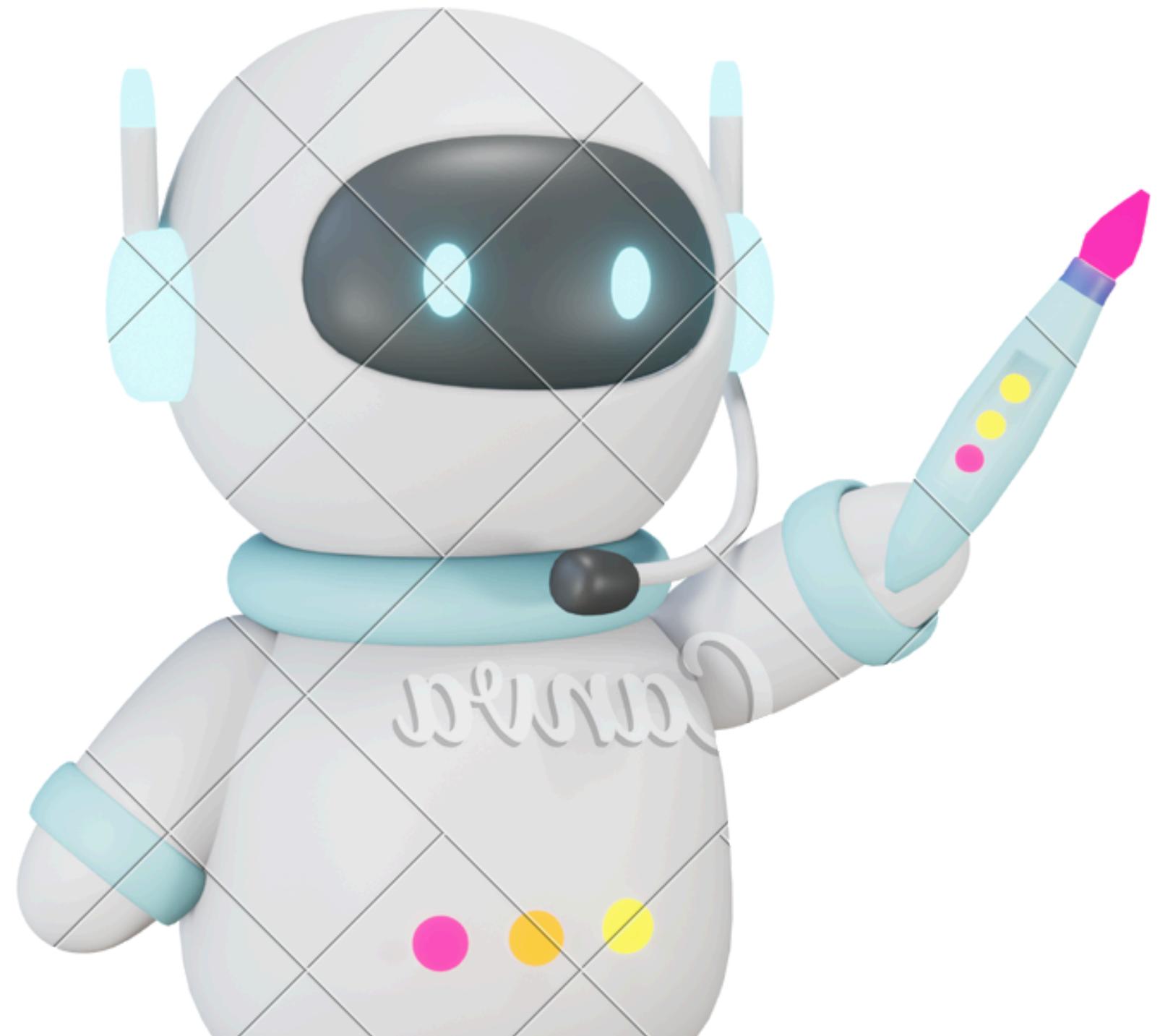
U.S. GAO – “AI Accountability Framework for Federal Agencies”

Operational guide to AI governance in public sector.

OECD AI Policy Observatory – “Advancing Accountability in AI”

Global policies for transparency, fairness, and lifecycle governance.

OECD AI Policy Observatory (Platform)



Thank You

Build powerful AI.
Build with ethics.
Build together.