

## AI for Good - or is it bad news? Why Governance is Key

Artificial intelligence technologies have the power to transform our world for better or worse. As AI becomes more advanced and integrated into our daily lives, ensuring it is developed and applied responsibly will be critical.

AI is already being used in high-stakes fields like healthcare, criminal justice and finance. The potential for harm if left unchecked is considerable. But with proper safeguards and governance, AI can have hugely positive impacts - from diagnosing diseases more accurately to reducing recidivism in the criminal system.

"The future of AI depends on how we handle its development and rollout," says AI ethics expert Daniela Amodei. "Having principles and procedures in place to promote its beneficial use can help maximize AI's upsides while mitigating risks."

### Establishing Meaningful Oversight

Effective AI governance requires engagement from policymakers, companies, researchers and the public. Key considerations policymakers are now weighing include:

- Defining core values like fairness, safety, transparency and accountability that should underpin all AI systems.
- Regulating high-risk AI applications and setting baseline standards for things like algorithmic bias testing, security protocols, and impact assessments.
- Incentivizing "value-aligned" AI through measures like public funding and voluntary certification programs.
- Investigating and enforcing against harmful, deceptive or dangerous uses through new or adapted regulations.
- Facilitating multistakeholder discussions to build consensus around governance best practices.

Long-term oversight will also mean adapting policies as AI technologies continue innovating rapidly. "Regulation cannot be static," says Meredith Whittaker of the AI Now Institute. "It needs to be continually re-examined to ensure it keeps pace with technical changes."

### Upholding Principles in Practice

For policy frameworks to make a real difference, they must also be upheld in design, development and deployment. Companies building AI systems have a key role to play by:

- Integrating third-party audits, documentation of models, and procedures for investigating issues proactively.
- Prioritizing interpretability to enable monitoring of algorithmic decision-making.
- Amplifying diverse voices through techniques like value sensitive design and multi-stakeholder governance boards.

- Partnering with researchers, civil society groups and regulators to self-regulate responsibly and strengthen oversight collaboratively over time.

With so much riding on AI's development, governance must be an ongoing, inclusive process. By balancing innovation with safeguarding values, the future of AI could help societies solve old problems in new ways - or potentially create all new ones if left unchecked. Oversight will be key to navigating this pivotal moment and steering AI toward benefit.

Here are some practical things directors, boards and management can do to help ensure effective governance around AI usage:

- Understand AI risks and opportunities relevant to their industry/sector. Directors must educate themselves on technology basics and implications.
- Establish clear oversight responsibilities. Appoint an AI governance committee or assign oversight to existing roles like CTO, CIO or CRO.
- Develop an AI strategy and principles. Define how AI will create value responsibly aligned with organizational missions and priorities.
- Implement controls for model development. Require documentation, testing, impact assessments and change management processes.
- Monitor for bias, security and incidents regularly. Conduct ongoing model monitoring, audits and issue response planning.
- Align incentives around values. Ensure executive compensation considers adherence to AI principles not just profitability.
- Communicate openly about AI usage. Disclose model purposes, available explanations and rights/recourse for customers.
- Foster responsible partnerships. Vet vendors/partners and ensure contractual protections for sensitive data and oversight access.
- Advocate for responsible innovation. Champion regulation promoting safe, fair and transparent technology development.
- Obtain independent assessments. Seek third party reviews of high-risk programs to identify gaps and areas for improvement.

Directors should emphasize AI as a strategic asset to manage carefully for long term sustainability and stakeholder trust.

You can find more articles on our website, at [Phundex Knowledge Hub](#), on LinkedIn at [Phundex LinkedIn](#), or for other questions, please email us at: [hello@phundex.com](mailto:hello@phundex.com).

To book a demo or do a trial, you can either use the link on our website or email [support@phundex.com](mailto:support@phundex.com), and they will be happy to set it up for you.