Trends in Ethical Principles in Corporate AI Implementation

OZAWA Yasumasa

Digital Hollywood University, Graduate School

As artificial intelligence (AI) becomes increasingly integral to enhancing corporate competitiveness, attention to ethical principles during its implementation is crucial. This research note examines the evolving trends in ethical principles that companies should prioritize when adopting AI. Failure to address these ethical considerations may not only lead to suboptimal outcomes but also harm the company's reputation and long-term performance. By analyzing current trends, this note aims to provide guidance for companies on how to practically apply ethical principles in AI implementation, effectively balancing technological advancement with social responsibility.

1. Introduction

Artificial intelligence (AI) has great transformative power and significant implications for societies and economies around the world. AI is playing an increasingly important role in shaping the development of economies and financial sectors, and is expected to drive productivity and economic growth through increased efficiency, improved decision-making processes, and the creation of new products and industries. However, companies that proceed without taking into account the complexities of AI ethics and data integrity may harm their reputations in pursuit of short-term gains.

For instance, according to IBM's "The CEO's guide to generative AI"[1], executives understand what is at stake in rushing to adopt AI, including the risk of damaging reputations for short-term gains, and 58% of them believe that the introduction of generative AI will increase significant ethical risks. And these risks will be very difficult to manage without new, or at least more mature, governance structures. However, many companies are struggling to put the principles into practice. 79% of executives say AI ethics are important for their company-wide AI approach, but less than 25% of companies are operationalizing common principles of AI ethics.

To illustrate, according to PwC Japan's "AI Predictions for 2023"[2], when only companies in Japan and the US that have already implemented or are considering implementing AI governance were surveyed, the results were as follows. In Japan, 228 companies were surveyed, and in the US, 840 companies were surveyed.

Among these, 22% of Japanese companies and 40% of US companies responded that they are considering establishing an ethics committee. Additionally, 5% of Japanese companies and 0% of US companies reported that their governance measures are being implemented or progressing without any problems.

These results highlight the fact that even companies that have already implemented AI governance or are considering it are not necessarily doing so well.

Furthermore, according to the Information-Technology Agency, Japan's "Survey Report on Security Threats and Risks When Using AI"[3], only respondents who indicated that they use, authorize, or plan to use AI (n=1,000) were asked whether they have established rules and systems related to security when using generative AI.

The survey results revealed that while approximately 60% of respondents perceive threats related to AI security, less than 20% have established, documented, or systematically considered such rules. Even when including companies that have detailed security rules in place, the proportion remains at only around 40%.

This indicates a significant gap between the perceived risks and the implementation of security measures.

These data show that while companies recognize that AI ethics and governance are important, they are not taking sufficient measures. In this research note, I discuss the ethical benefits of introducing AI from the perspective of effective accelerationism and effective altruism, and present a cycle that satisfies stakeholders without separating ethical behavior from business behavior.

2. Effective Altruism and Effective Accelerationism

2.1 Effective Accelerationism

Effective accelerationism (e/acc) is a philosophy that aims to reform society through the promotion of technology and capitalism. This idea has gained a great deal of support, especially in the technology industry in Silicon Valley. It has become popular through platforms such as Twitter and Substack since around 2022, and many entrepreneurs and investors agree with this idea.

The main tenet of e/acc is to accelerate technological progress indefinitely and embrace the social and economic changes that arise in the process. This idea is at odds with voices that emphasize the safety of AI as research and development of AI progresses rapidly. In some parts of the tech industry, there is a growing argument that "technological progress should be accelerated without regulating technology," and some have said that "AI regulation advocates are pessimists."

Supporters of this idea include prominent venture capitalists and executives of major accelerators. For example, Marc Andreessen of Andreessen Horowitz and Gary Tan, CEO of Y Combinator, have expressed their support for effective accelerationism. They aim to maximize innovation and market-driven growth through this idea.

However, it has also been pointed out that the promotion of this rapid technological progress may give rise to social costs and ethical issues. Effective accelerationism aims to simultaneously advance technological progress and social reform, but there is currently no clear guideline established for how to deal with the various problems that arise in the process.

2.2 Effective Altruism

Effective Altruism (EA), as a social movement, aims to solve the world's problems in the most effective way. The movement has divided employees and executives of AI development companies, especially in Silicon Valley, demonstrating its influence. EA pursues ways to generate the greatest social impact based on reason and scientific evidence. This approach plays an important role, especially in the development of conversational AI.

EA advocates emphasize safety and ethical issues in AI research and development. They argue that a more cautious approach should be taken to the potential risks posed by technological advances. This way of thinking calls for a balance to deal with rapid progress and the ethical challenges that come with it, especially in artificial intelligence research.

The Effective Altruism (EA) movement has also significantly contributed to the development of ethical guidelines for AI. Through the EA framework, discussions have focused on critical issues such as how AI can enhance human welfare and how to ensure fairness in AI decision-making processes. By addressing these considerations, the EA movement aims to advance technological progress while carefully managing its societal impact.

However, there is a major divide between effective accelerationists and effective altruists regarding their approach to AI.

3. Advantages of Effective Altruism for Companies

There are several reasons why companies choose services from companies based on EA when introducing AI.

1. Ethical Considerations and Social Responsibility: EA is a philosophy that aims to maximize the benefit of others in the most effective way. By choosing a service based on this philosophy, companies can ensure that ethical considerations and social responsibility are taken into account when introducing AI. For example, the potential risks and impacts of AI can be evaluated in advance to ensure that it is beneficial to society.

2. Sustainable Growth: Companies that advocate EA do not simply pursue profits, but aim for sustainable growth from a long-term perspective. Even when introducing AI, the use of technology in a way that does not negatively impact the environment or society is prioritized, improving the sustainability of the company.

3. Trustworthiness and Transparency: Companies that practice EA place importance on trust and transparency. This ensures that data handling and decision-making processes are transparent even in AI introduction projects, increasing trust in the partnership. It is expected that companies will be able to appropriately respond to the ethical issues and challenges they face when introducing AI.

4. Reputation Management: A company's reputation as a socially responsible and ethical company is an important factor for consumers and investors. Partnering with a company based on EA can improve a company's social reputation and increase brand value.

5. Innovation and Leadership: Companies that engage in EA are proactive in innovating to maximize the benefits to society as a whole. Even in the introduction of AI, they can be expected to provide innovative technologies and solutions, helping companies establish leadership in their industries.

For these reasons, it can be said that companies considering introducing AI can benefit greatly in terms of ethical considerations, social responsibility, sustainability, improved trust, and reputation and innovation by choosing the services of a company based on EA.

4. Implications of AI Ethical Guidelines

According to "AI Ethics in the Age of Generative AI"[4], as awareness of the risks and potential harms of AI grows, government agencies, AI development companies, and researchers have formulated various ethical guidelines for AI, called ethics codes, guidelines, statements, policies, principles, frameworks, etc., to seek healthy and reliable research, development, and use of AI.

Can various ethical guidelines for AI (AI guidelines) have an impact on society and guide government agencies, companies, academic institutions, and citizens in the healthy use and development of AI? Recent research on AI ethics has harshly criticized AI guidelines in the digital field for being used to deceive consumers, evade regulations, and buy time for lobbying activities to weaken regulations. In the field of environmental ethics, the deceptive and dishonest practices of companies and government agencies that attempt to appear more environmentally friendly, sustainable, and ecologically friendly than they actually are are sometimes called "green washing" . Similarly, in the digital field, companies and government agencies have been criticized as "ethics washing" for pretending to have a greater moral responsibility to citizens and a wide range of stakeholders than they actually do. Building on these arguments, Munn[5] has ignited controversy by arguing that the AI guidelines are "useless" and being used as a tool for ethics washing.

While it may be true that AI guidelines are not fully implemented across all organizations, this does not mean that the guidelines themselves are "useless." Instead, this indicates the need for stricter oversight and enforcement mechanisms to ensure their effectiveness. Ethical guidelines not only encourage organizations to voluntarily fulfill their responsibilities but also highlight the necessity for external pressure and regulation.

The risk of "ethics washing" is indeed real, but dismissing AI guidelines entirely due to this risk is an overreaction. To prevent ethics washing, it is crucial to ensure transparency and establish third-party audits so that guidelines do not remain superficial statements. Additionally, including diverse stakeholder perspectives in the development of guidelines can make them more fair and effective, ensuring that ethical responsibilities lead to concrete actions rather than mere "window dressing."

AI guidelines may not be flawless in the short term, but in the long run, they can play a crucial role in establishing ethical standards for AI development and usage. Historical precedents show that guidelines and norms often evolve over time to influence corporate behavior. Similarly, AI guidelines have the potential to drive sustainable change as societal awareness and technology continue to advance.

The argument that AI guidelines are "useless" underestimates their potential value and capacity for evolution. While the risk of "ethics washing" exists, this should not lead to the outright dismissal of guidelines. Instead, efforts should be directed towards strengthening these guidelines and enhancing their effectiveness. Ethical guidelines are vital tools for promoting responsible AI

practices, and when properly implemented, they can significantly contribute to society.

5. Effective Accelerationism

If a company chooses to adopt services provided by a company based on e/acc instead of EA, especially in the context of the EU Artificial Intelligence Act(AI act), the company's executives could face several significant risks:

1. Legal Risks and Penalties

The EU's AI Act emphasizes the importance of safety and ethics in AI development and usage, with strict regulations, especially for high-risk AI systems. e/acc prioritizes rapid technological advancement, which may lead to the disregard of these regulations. This could result in the following risks:

Hefty Fines: If a company violates the AI Act, the EU can impose substantial fines. Specifically, companies may face penalties of up to 7% of their annual global turnover. Such fines can have a severe impact on the company's financial health.

Legal Liability: If regulatory violations are discovered, company executives may also face personal legal liability. If executives are found to have knowingly allowed these violations, they could be subject to legal actions, including criminal penalties.

2. Failure in Risk Management

Services based on e/acc aim to accelerate technological progress, which might result in inadequate risk management procedures as required by the AI Act. This can expose companies to the following risks:

Lack of Compliance: If AI systems do not meet the necessary risk assessments and audit procedures, regulatory authorities may increase scrutiny, potentially disrupting business operations.

Product Recalls or Suspension Orders: If non-compliant AI systems are released into the market, companies risk facing product recalls or orders to suspend usage. This can lead to a significant loss of market trust.

3. Reputation Risks

While AI services based on e/acc might appear innovative in the short term, neglecting ethical concerns can severely damage a company's reputation. Given the enforcement of the EU's AI Act, the following risks are pertinent:

Loss of Consumer Trust: If regulatory violations come to light, companies may lose the trust of consumers and partners, leading to a substantial decline in brand value.

Pressure from Investors: Investors who prioritize social responsibility are likely to avoid companies that violate regulations. This could make it difficult for the company to secure funding.

4. Restricted Market Access

The EU, through the AI Act, promotes the use of safe and ethical AI systems. Companies that do not comply with these regulations risk restricted access to the EU market.

Market Exclusion: Companies that violate regulations could be effectively excluded from the EU market, resulting in significant economic losses, particularly for those reliant on access to the EU market.

If a company's executives choose to adopt services based on e/acc, they face substantial risks in light of the EU's AI Act, including legal risks, risk management failures, reputation damage, and restricted market access. These risks can have serious implications for the company's financial stability, credibility, and long-term sustainable growth, necessitating careful consideration in decision-making.

6. Conclusion

The choice between Effective Altruism (EA) and Effective Accelerationism (e/acc) is not just a philosophical debate. It has tangible implications for companies, especially in the context of AI governance. While e/acc emphasizes rapid technological progress and market-driven growth, it carries significant risks, particularly in light of stringent regulations like the EU's AI Act. These risks include legal penalties, failure in risk management, reputational damage, and potential exclusion from critical markets.

On the other hand, adopting principles based on EA offers a more balanced approach. Companies that align with EA can benefit from ethical considerations, social responsibility, and sustainable growth. By prioritizing transparency, trustworthiness, and innovation that serves societal well-being, EA-aligned companies are better positioned to navigate the complex landscape of AI ethics and governance.

Moreover, while criticisms like "ethics washing" raise valid concerns about the superficial application of AI guidelines, dismissing these guidelines outright would be shortsighted. Instead, efforts should focus on strengthening these frameworks, ensuring they are robust, enforceable, and inclusive of diverse stakeholder perspectives. Over time, well-implemented guidelines have the potential to shape responsible AI practices that align with both ethical standards and business objectives.

In conclusion, companies should carefully weigh the benefits of sustainable, ethical AI practices against the short-term gains of unchecked accelerationism. By choosing to integrate EA principles into their AI strategies, companies can mitigate risks, enhance their reputation, and contribute to a future where technological advancement and ethical responsibility go hand in hand. This balanced approach not only supports long-term growth but also aligns with the increasing demands for corporate accountability in the digital age.

References

[1] IBM: "The CEO's guide to generative AI" https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/ceo-generative-ai/responsible-ai-ethics. Accessed 2024-8-26

[2] PwC Japan: "AI Predictions for 2023" https://www.pwc.com/jp/ja/knowledge/thoughtleadership/2023-ai-predictions.html. Accessed 2024-8-26

[3] Information-technology Promotion Agency, Japan: "Report on the investigation into security threats and risks when using AI" https://www.ipa.go.jp/security/reports/technicalwatch/eid2eo00000022sn-att/report.pdf. Accessed 2024-8-26

[4] TANAKA Ikko: "AI Ethics in the Age of Generative AI: The Significance and Controversy of AI Guidelines", J. F. Oberlin University Journal of Advanced Research. Humanities, vol 4, p. 117-132, Publication date 2024-03-25 https://obirin.repo.nii.ac.jp/records/2000146

[5] Luke Munn: "The uselessness of AI ethics"

https://link.springer.com/article/10.1007/s43681-022-00209-w

Accessed 2024-10-6

[6] European Union: "The EU Artificial Intelligence Act"

https://digital-strategy.ec.europa.eu/en/policies/regulatory-

framework-ai

Accessed 2024-10-6