

# Ethical Considerations in AI Development: Balancing Autonomy and Societal Benefits

Yukta Parag Butala



## Introduction

In an age dominated by technological advancements, the ethical considerations about data usage and autonomy have increased. Consider this scenario about a firm utilizing data from Reddit users to train a Large Language Model (LLM). While this practice has led to significant advancements in artificial intelligence (AI), it has also sparked complaints from Reddit users who argue that they did not consent to their data being used and are not benefiting from the profits generated by the company. At the heart of this issue lies the ethical principle of autonomy – the right of individuals to make informed decisions about the use of their data. This paper aims to address this autonomy ethical issue within the context of data usage from public spaces to train AI models. Through the lens of utilitarianism, the paper will propose a policy that balances the interests of individuals with the societal benefits derived from AI development. The focus will be on maximizing overall happiness and minimizing harm while ensuring that the company's actions align with ethical principles and contribute to the greater good. By employing utilitarianism as the guiding normative theory, this paper will offer a comprehensive analysis of the autonomy ethical issue and provide a clear pathway towards a just and ethical policy framework. Training models using data sourced from public spaces involves gathering information from publicly available sources such as social media platforms, forums, and websites to train AI algorithms to recognize patterns, understand language, and generate responses (Gunning, 2016). For instance, companies like OpenAI have utilized data from platforms like Reddit to train large language models (LLMs) such as ChatGPT, enabling these models to generate human-like text based on the input provided (Radford et al.,

2021). The data collection raises ethical questions surrounding consent and autonomy, particularly when individuals may not be aware of or have explicitly consented to their data being used for AI training purposes.

In the specific case of Reddit, users have voiced complaints regarding the usage of their data without explicit consent. Many argue that they did not provide consent for their posts and comments to be used to train AI models, and they are not benefiting from the profits generated by the companies utilizing their data (Zeng, 2020). This issue underscores the importance of autonomy – the right of individuals to control their personal information and make informed decisions about its use. Reddit users believe that their autonomy is compromised when their data is collected and utilized without their explicit consent, raising concerns about privacy and fair compensation for their contributions.

Furthermore, similar concerns have been raised in other contexts, such as the lawsuits filed by artists against companies that use their data without proper consent or compensation. For example, in the music industry, streaming platforms like Spotify have faced legal challenges from artists who argue that their music is being exploited for profit without adequate compensation (Jennings, 2019). These lawsuits highlight the broader ethical implications of data usage from professionals, including artists, musicians, and other content creators, without their explicit consent. By examining the complaints from Reddit users and the lawsuits from artists, it becomes evident that a balance must be struck between respecting individual autonomy and promoting the advancement of technology for the greater good.

### **Consequence-based Normative Theory/Utilitarianism**

Utilitarianism, as a normative ethical theory, prioritizes the maximization of overall happiness and the minimization of harm in decision-making (Mill, 1863). Rooted in consequentialism, utilitarianism assesses the morality of actions based on their outcomes, to achieve the greatest good for the maximum people. This approach holds that the consequences of an action determine its moral worth, rather than adhering to rigid rules or principles. In the context of public data usage for AI training, utilitarianism provides a framework for evaluating the ethical implications of such practices. By considering the overall societal benefits derived from advancements in AI technology, such as improved healthcare diagnostics,

enhanced customer service, and more efficient resource allocation, utilitarianism guides policymakers in weighing the potential positive outcomes against any negative consequences (Brey, 2012).

Furthermore, utilitarianism offers an approach to address the autonomy concerns raised by individuals whose data is used without explicit consent. While respecting individual autonomy is important, utilitarianism emphasizes the collective welfare of society. Thus, in formulating policy regarding public data usage, a utilitarian perspective would consider not only the rights of individuals but also the broader societal benefits that may come from the development and implementation of AI technologies. By focusing on the consequences of different policy options, utilitarianism provides a rational basis for decision-making that seeks to optimize overall well-being while minimizing harm (Perry, 2006). The complex ethical challenges posed by public data usage, employing a utilitarian framework allows policymakers to strike a balance between individual autonomy and societal welfare, ultimately guiding the formulation of ethical and socially responsible policies.

## **Proposed Policy**

### ***A. Company's Response to Reddit Users' Complaints:***

1. Acknowledgment of Concerns: The company should issue a public statement acknowledging the concerns raised by Reddit users regarding the usage of their data without explicit consent. They should recognize the importance of transparency and accountability in data practices and understand the significance of respecting individual autonomy (Floridi, 2014). This acknowledgment will demonstrate their commitment to ethical principles and willingness to address user concerns promptly and transparently.
2. Explanation of the Benefits of AI Development for Society: In response to Reddit users' concerns, the firm should provide a detailed explanation of the societal benefits derived from AI development. They should highlight specific examples of how AI technologies, powered by data from platforms like Reddit, have led to advancements in areas such as natural language processing, medical research, and disaster response (Russell & Norvig, 2022). By emphasizing the positive impact of AI on society, the company should contextualize the importance of data usage for AI training while also fostering understanding amongst users about the broader implications of their data contributions.

3. Offer of Transparency and Opt-Out Options for Users: To address autonomy concerns, they should enhance transparency and provide opt-out options for Reddit users who do not wish to contribute their data to AI training. They should implement user-friendly mechanisms through which users can access information about how their data is being used and easily opt out of data collection processes if desired (Floridi, 2019). By empowering users to make informed decisions about their data, they should uphold users' autonomy and ensure that their privacy rights are respected.

***B. General Policy Regarding Data Used from Public Spaces:***

1. Implementation of Ethical Guidelines for Data Usage: The firm should establish comprehensive ethical guidelines for the collection and usage of data from public spaces. These guidelines will include principles such as informed consent, data anonymization and encoding, to protect user privacy and autonomy (Bostrom & Yudkowsky, 2014). They should also prioritize accountability in their data practices, implementing measures to mitigate potential misuse and ensure transparency in data usage.

2. Consideration of Compensation for Data Contributors Where Feasible: Recognizing the value of user contributions to AI training data, the company should explore options for providing compensation to data contributors where feasible. While not all data contributions may warrant compensation, they should consider implementing rewards programs or revenue-sharing models to fairly compensate users for their data contributions (Mittelstadt et al., 2016). By acknowledging the importance of user contributions and providing fair compensation where appropriate, they should foster trust and reciprocity between the company and the Reddit community.

Overall, these policies aim to address the autonomy concerns raised by Reddit users while also promoting responsible data practices and maximizing the societal benefits derived from AI development and data usage. By prioritizing transparency, user empowerment, and ethical considerations, the firm should seek to build trust with users and contribute positively to the advancement of AI technologies.

**Implications of Policy on Artists, Musicians, and Professionals**

The utilization of data from artists, musicians, and other professionals in the entertainment industry raises significant ethical considerations, as demonstrated by the recent surge in lawsuits against companies accused of leveraging such data without proper consent or compensation. For instance, streaming

platforms like Spotify have found themselves in legal cases with artists who assert that their music is being exploited for profit without a fair share given to them(Jennings, 2019). These legal disputes highlight the complex ethical issues surrounding data usage within the entertainment sector, making the establishment of clear guidelines and regulations to safeguard the rights of content creators.

In examining the potential benefits and harms of data usage for professionals in the arts and entertainment industry, it becomes evident that the impact is multifaceted. On one hand, data-driven insights have the potential to offer valuable analytical feedback to artists and musicians, helping in understanding audience and content customization. Additionally, data-driven algorithms can provide personalized recommendations for consumers, potentially expanding exposure for emerging artists and diversifying content consumption habits (Cox, 2019). However, concerns persist regarding the commercialization of creativity and the washing away of artistic autonomy. When data is utilized solely for commercial gain without equal compensation or acknowledgment, it risks undermining the integrity and creative autonomy of artists and musicians.

Balancing the interests of professionals in the arts and entertainment industry with broader societal benefits necessitates careful navigation of ethical principles and industry norms. While data-driven technologies hold promises for transforming content creation and distribution, it is imperative to ensure that the rights and autonomy of artists and musicians remain safeguarded. This may be done by the implementation of regulations for the ethical collection and usage of data, as well as the establishment of fair compensation structures for data contributors (IFPI, 2021). Furthermore, boosting transparency and collaboration among content creators, platforms, and consumers can increase trust and encourage responsible data practices within the industry.

Practical changes to industry standards and corporate practices may include the adoption of transparent data usage policies, the establishment of fair compensation mechanisms, and the implementation of regulatory frameworks that prioritize the rights and autonomy of content creators.

## **Objections**

One potential objection is the challenge of quantifying and comparing different forms of happiness or well-being, which are central to utilitarian calculations. Subjective measures of happiness may not adequately

capture the diverse range of values and preferences within society. Prioritizing the maximization of overall happiness may lead to overlooking the rights and autonomy of individuals, particularly when their consent is not explicitly obtained for data usage.

Another counterargument is about the consequentialist nature of utilitarianism, which captures the impact of actions based on their outcomes. This approach overlooks the intrinsic value of ethical principles, such as individual autonomy and privacy. Even if data usage from public spaces has societal benefits, it may be unethical to disregard the autonomy of individuals who have not explicitly consented to the use of their data. Additionally, there could be concerns about utilitarianism to support actions that result in long-term harmful consequences, even if they reflect short-term benefits.

However, it is essential to acknowledge the limitations of utilitarianism while also using it in ethical decision-making. While quantifying happiness and comparing different forms of well-being may pose challenges, consequence-based theory provides a framework for assessing the outcomes of different policy options and prioritizing actions that maximize overall societal welfare. Rather than focusing just on subjective measures of happiness, utilitarianism encourages a comprehensive assessment of the consequences of actions, considering factors such as individual rights, autonomy, and long-term societal impacts.

Moreover, consequence-based theory emphasizes the importance of balancing competing interests and minimizing harm while maximizing overall happiness. In the context of data usage from public spaces, our proposed policy aligns with utilitarian principles by prioritizing transparency, respecting individual autonomy, and promoting responsible data practices. By offering transparency and opt-out options for users, our policy aims to empower individuals to make informed decisions about their data while also maximizing the societal benefits derived from AI development and data usage.

Overall, while utilitarianism has arguments against ethical decision-making, our proposed policy shows a commitment to balancing the interests of individuals with broader societal benefits. By addressing autonomy concerns and promoting ethical data practices, the policy upholds utilitarian consequence-based principles while fostering trust and accountability within the community.

## **Conclusion**

The proposed policy regarding data usage from public spaces, in utilitarian normative theory lens, strikes a balance between respecting individual autonomy and maximizing societal benefits. By acknowledging concerns, promoting transparent practices, and considering opt-out options, the policy aims to address concerns while promoting responsible data practices that enhance overall societal welfare. Utilitarianism provides an amazing framework for justifying this policy, as it emphasizes the importance of maximizing overall happiness and minimizing harm. However, it is crucial to recognize the ongoing importance of balancing autonomy concerns with societal benefits in the realm of AI development and data usage. While advancements in technology hold great potential for societal progress, it is essential to protect individual rights and autonomy, making sure that ethical considerations remain one of the top priorities in the pursuit of innovation and progress.

## References

1. Gunning, D. (2016). Explainable artificial intelligence (XAI). Defense Advanced Research Projects Agency (DARPA).
2. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2021). Language models are unsupervised multitask learners. OpenAI.
3. Zeng, D. D. (2020). From privacy to autonomy in the age of artificial intelligence. *IEEE Transactions on Technology and Society*, 1(1), 1-8.
4. Jennings, D. (2019). Spotify Faces \$1.6 Billion Copyright Lawsuit Over Tom Petty, Weezer, Neil Young Songs. Retrieved from Billboard website: <https://www.billboard.com/articles/business/legal-and-management/8525562/spotify-wixen-copyright-lawsuit>.
5. Mill, J. S. (1863). *Utilitarianism*. Parker, Son, and Bourn.
6. Brey, P. (2012). Anticipating ethical issues in emerging IT. *Ethics and Information Technology*, 14(4), 305-317.
7. Perry, S. (2006). Utilitarianism and the catastrophe of unforeseeable consequences. *Ethical Theory and Moral Practice*, 9(1), 13-32...
8. Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.
9. Russell, S., & Norvig, P. (2022). *Artificial intelligence: A modern approach*. Pearson.
10. Floridi, L. (2019). Soft ethics and the governance of the digital. *Philosophy & Technology*, 32(2), 169-182.
11. Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In *The Cambridge handbook of artificial intelligence* (pp. 316-334). Cambridge University Press.
12. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 2053951716679679.
13. Cox, J. (2019). How Music Data Is Being Used to Personalize Your Playlist. Retrieved from Rolling Stone website: <https://www.rollingstone.com/music/music-features/playlist-algorithms-data-music-839950/>.
14. IFPI. (2021). *Investing in Music*. Retrieved from International Federation of the Phonographic Industry website: <https://www.ifpi.org/wp-content/uploads/2021/10/IFPI-InvestingInMusic-2021.pdf>.