

Navigating the privacy paradox in a digital age: balancing innovation, data collection and ethical responsibility

Martin Jones

*College of Business and Entrepreneurship, North Greenville University,
Tigerville, South Carolina, USA*

Abstract

Purpose – The purpose of this paper is to explore the intersection of generative artificial intelligence (AI), data collection and consumer privacy, highlighting ethical tensions in AI-driven advertising. This study examines key challenges, including data mining in smart devices and the implications of high-profile cases like Amazon's proposed acquisition of iRobot.

Design/methodology/approach – This paper uses a conceptual exploration of case studies, regulatory developments and the current legislative responses in the USA and the European Union. This study further proposes ethical self-regulation, drawing parallels to historical precedents like the Defense Industry Initiative, while emphasizing transparency, privacy-by-default and consumer-centric AI design.

Findings – The findings of this study reveal significant gaps in existing regulatory frameworks, particularly in the USA, and underscore the need for proactive industry leadership in self-regulation. This paper identifies practical solutions, such as opt-in data collection models and the incorporation of moral reasoning into AI training, to enhance consumer trust and privacy protections.

Practical implications – Businesses can leverage this paper's proposal to navigate the privacy paradox, enhance consumer trust and mitigate privacy risks, thus fostering ethical innovation.

Social implications – Addressing privacy concerns proactively can alleviate societal apprehensions about AI technologies, contributing to broader acceptance, consumer trust and ethical integration of AI-enabled data collection in daily life.

Originality/value – This study bridges the gap between theoretical discussions on AI ethics and practical, implementable solutions. By advocating for self-regulation alongside robust legislative measures, this study provides a novel pathway to balance technological innovation with ethical responsibility in data-driven advertising.

Keywords Data collection, Artificial intelligence, AI ethics, Generative AI, Consumer privacy, Ethical AI, Targeted advertising, Self-regulation, Moral AI, Innovation

Paper type Conceptual paper



Advances in artificial intelligence-driven data collection

We are all presently witnessing the fastest ever expansion of technology in the history of the human record. Recent research indicates that in less than two years since the public release of

© Martin Jones. Published in *Journal of Ethics in Entrepreneurship and Technology*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

ChatGPT, generative artificial intelligence (AI) tools have achieved a 40% household adoption rate, which far surpasses previous adoption rates for the internet, iPhone and home computers (Bick *et al.*, 2024). As generative AI technologies increasingly become everyday tools, businesses now have greater access to consumer data which they can easily convert into a myriad of targeted advertising opportunities (Davenport and Mittal, 2023). This is not, however, in and of itself, merely a post-ChatGPT phenomenon. Even before November 30th, 2022, it was widely acknowledged that advancements in AI were rapidly changing the way consumer information was not only being collected but also how it was being used to curate highly personalized marketing strategies (Shankar, 2018). Now that businesses are able to leverage an ever-growing plethora of generative AI tools, they are finding it much easier to analyze consumer behavior, as well as predict trends, and subsequently respond with personalized spending recommendations on a scale not previously witnessed. AI tools are not only able to seamlessly interact with consumers but are also capable of easily tracking spending habits, as well as providing brands with precise insights and patterns, quickly implementing hyper-targeted marketing and personalized content creation back to the user. Most consumers are quite unaware of the extent to which their personal information is being scraped by everyday AI-enabled technologies, apps and smart devices (Blaney, 2024). These capabilities enable today's businesses to anticipate consumer preferences with remarkable speed and accuracy. Given the pace with which advancements in AI technologies and data collection are now being made, as well as the present lack of regulatory oversight in the USA, there is a growing need to examine how to balance data collection practices while implementing safeguards to protect consumers from privacy infringements. As brands intensify data mining efforts, using AI to extract and analyze preferences from extensive digital footprints, the need for clear regulation and ethical compliance becomes ever more pressing. Devising an appropriate and ethical framework that recognizes the immense potential of generative AI with respect for personal autonomy is essential for maintaining trust, integrity and sustainable growth in a digitally accelerated economy.

Consumer data and privacy concerns in the age of smart homes

On August 5th, 2022, Amazon announced plans for a \$1.7bn acquisition of iRobot, the maker of Roomba robotic vacuums. At the time, this was heralded as a strategic decision that would further enhance and expand Amazon's existing smart home technology capabilities (Amazon.com, Inc., 2024b). However, within a short period of time, concerns began to surface about the potential ethical implications of the use to which Amazon might put data captured by the Roomba cameras, especially given the fact that the vacuums are equipped with advanced technology that not only maps the layout of the space to be vacuumed but also captures and stores images of consumers' homes. Questions about whether that data would potentially then be analyzed and used to make assumptions about users and create targeted advertising started causing concerns (Blaney, 2024). If the proposed acquisition went ahead as planned, then would Amazon have access to data collected from consumers' homes, and if so, then to what use would that data be put? Blaney observed, "The size of your house is a pretty good proxy for your wealth. A floor covered in toys means you have kids. A household without much furniture is a household to which you can try to sell more furniture" (para. 5). All this information could be an asset to Amazon, who after all, are a company in the business of selling products and making its platform more attractive to both smart home enthusiasts and consumers who are just getting acquainted with using home-related digital devices.

This primary concern was soon echoed by secondary questions about how user data would be stored and who else would be able to access it. Consumer advocates began to sound alarm bells and expressed concern that the mining of data and images from everyday smart

appliances might evolve into the normalization of surveillance of homes and private spaces. The Digital Rights Watch Group raised concerns about the potential for Amazon to combine data from Roomba with data from other Amazon products, such as Ring security cameras and Echo smart speakers. They argued that this could give Amazon too much information about people's lives and could be used to target them with advertising or even adjust their insurance premiums ([Digital Rights Watch, 2022](#)). James Clark, Executive Director of Digital Rights Watch asserted, "At its core, Amazon is a surveillance company. Amazon protects its market power and profitability by using its scale to build detailed profiles on millions of people and uses that to predict market trends and manipulate user behavior" (para 7). Given Amazon's size as an online retailer and the extent to which they already made use of consumer-driven algorithms, the concern was that data acquired by Roomba vacuums could be used for personalized adverts. In addition, advocates warned of the need to address potential privacy violations, given that virtually all of the data generated would be from private dwellings ([The Robot Report, 2022](#)). Furthermore, what impact might this acquisition have on consumer trust, especially if data was found to have been leaked through a breach? Broader concerns were also articulated about an inherent risk of a backlash against smart home technologies. Privacy advocates argued that without stringent safeguards, the potential acquisition might result in the exploitation of intimate details about consumers' daily lives, habits and home environments.

Regulatory bodies took these concerns seriously. The acquisition underwent scrutiny from the Federal Trade Commission (FTC) in the USA, the European Commission and the UK's Competition and Markets Authority. Investigations focused on not only antitrust implications but also the potential impact on consumer data privacy. In relation to the latter, some of the major concerns cited were that the planned acquisition would provide Amazon with access to iRobot's users' data, including not only the data voluntarily provided by the consumer when activating their device but also, and perhaps more importantly from a privacy concern, information actually collected by the Roomba mapping cameras as the devices cleaned the user's home. The European Commission in particular highlighted this had the potential to, "Allow Amazon to better rank organic results and advertisements on its own marketplace and/or to better personalize and target the advertisements" ([European Commission, 2023](#), para 7). Interestingly, while the European Commission eventually unconditionally approved the deal, the discussion around consumer privacy implications remained active and the FTC continued to express concern at the proposal.

Ultimately, the FTC did not have to make a final decision, because in January 2024, both Amazon and iRobot mutually agreed to terminate the acquisition because of regulatory challenges, particularly from the FTC ([Amazon.com, Inc., 2024a](#)). Consequently, Amazon did not complete the purchase of iRobot, and the FTC, in response to the terminated merger, issued statement indicating they were pleased the transaction was abandoned, citing concerns from their own investigation that it would have had negative effects not only on innovation, but specifically for consumer privacy ([Federal Trade Commission, 2024](#)).

Concerns regarding the misuse of consumer data captured from iRobot vacuums are well founded. Several years earlier in 2020, images surfaced on social media of a woman sitting on the toilet in her home. The images were taken by the onboard camera of iRobot's Roomba J7 series vacuum and transmitted via the cloud to human operatives who had been contracted to label audio, photo and video data to train AI ([Guo, 2022](#)). Those human operatives posted the images online, along with a number of others showing users in their homes. Such examples are not in isolation. Concerns regarding the security of consumer data were again highlighted in May 2024, when, in multiple incidents across the USA, another iteration of robotic vacuums called the Ecovacs Deebot X2 were compromised by hackers. These

unauthorized individuals remotely accessed the devices, enabling them to control the vacuums, view live camera feeds and broadcast offensive language through the onboard speakers. In one notable case, a Minnesota lawyer discovered his vacuum emitting racial slurs, while another incident involved a device in Los Angeles chasing a family's dog (Fell, 2024). The breaches were attributed to vulnerabilities in the Deebot X2's security protocols. Researchers had previously identified flaws, including a Bluetooth vulnerability that allowed remote access from over 100 meters away and a PIN system that could be easily bypassed. Despite prior warnings, these issues remained unaddressed, leaving the devices susceptible to exploitation (Naprys, 2024). In response to these incidents, Ecovacs acknowledged the security lapses and announced plans to release a firmware update to enhance the security of the X2 series. The company also advised users to implement strong, unique passwords and to update their devices regularly to mitigate potential risks and protect their privacy. However, these events yet again caused wider concern about the risk posed to consumer data by smart devices as well as the susceptibility of data collected and stored on them.

The Ecovacs breach, coupled with the Amazon-iRobot case, exemplifies the complex challenges at the intersection of AI-driven innovation and data privacy for consumers. As AI technologies become increasingly embedded in everyday consumer products, the potential for unintended or unregulated use of personal information grows. Businesses venturing into this new terrain must navigate not only the opportunities presented by advanced data collection and analysis but also the obligations to respect and protect consumer privacy. In the evolving landscape of AI-powered entrepreneurship, addressing these concerns is essential. King and Meinhardt (2024), advocating for consumer data privacy considerations in their White Paper, assert, "Largely unrestrained data collection poses unique risks to privacy that extend beyond the individual level—they aggregate to pose societal-level harms that cannot be addressed through the exercise of individual data rights alone" (p. 4).

The integration of generative AI tools into brand ecosystems not only offers significant benefits but also necessitates a reevaluation of data handling practices. Transparent policies, robust data protection measures and ongoing dialogue with stakeholders are crucial steps in ensuring that the march of innovation does not come at the expense of consumer trust and autonomy. As AI technologies continue to make seismic advances, brands now have unprecedented capabilities to gather extensive consumer data, which in turn can be used to create and deliver highly personalized promotions. Given the speed with which advancements in AI technologies and data collection are being made, as well as the present lack of regulatory oversight in the USA, there is a growing need to examine how to balance data collection practices alongside safeguards to protect consumers from privacy infringements.

The current legislative and regulatory framework

The rapid advancement of AI technologies and their integration into consumer products has not been met by a similar pace in the development of policy, legislation and compliance. In spite of high-profile cases such as Amazon's proposed acquisition of iRobot and the hacking incident involving Ecovacs robot vacuums, the potential for misuse of personal information gathered by AI-enabled devices continues with little-to-no regulatory framework. Consumer advocates and policymakers are, however, now increasingly recognizing that traditional privacy laws may be insufficient to address the challenges posed by generative AI technologies. For example, Ellucian's AI Survey of Higher Education Professionals represents one of the more recent surveys investigating solutions to the tension between maintaining privacy and AI advances. Ellucian CEO Ipsen observed, "Our survey findings show that while adoption is increasing, leaders are rightfully weighing AI's benefits with

careful consideration of privacy” (eCampus News, 2024, para 4). The Ellucian survey reported that the percentage of respondents worried about data privacy and security concerns increased from 50% in 2023 to 59% in 2024 (Ellucian, 2024).

The ability of AI systems to collect vast amounts of data, analyze it in complex ways and make autonomous decisions raises questions about consent, transparency and accountability that existing regulations were not designed to manage. Moreover, the potential for AI technologies to infringe upon individual privacy rights, whether through intentional misuse or inadvertent security lapses, necessitates a reevaluation of how laws can keep pace with technological innovation. This evolving landscape reflects an urgent need to address the complexities introduced by AI in the realm of data privacy to ensure that the ongoing integration of AI into consumer products proceeds responsibly, with adequate protections against the abuse of personal data. In their White Paper, King and Meinhardt articulate a set of conclusions and predictions about how existing and future privacy and data protection regulations in the USA and the EU will impact the development and deployment of AI systems (King and Meinhardt, 2024). They assert that, “Policymakers must expand the scope of how we approach privacy and data protection to address these weaknesses and bolster data privacy in an increasingly AI dominant world” (p. 6).

The USA has a complex history of privacy protections, starting with the fact that neither the US Constitution nor the Bill of Rights, explicitly either make mention of nor safeguard privacy rights. However, the Supreme Court has consistently recognized a right to privacy contained in several constitutional amendments, including the 1st, 4th and 14th Amendments (Griswold v. Connecticut, 1965). However, these protections primarily guard against governmental intrusion rather than regulating private entities’ use of personal data. Where the Constitution lacks explicit Federal oversight, several individual states, in response to concerns surrounding consumer privacy concerns, have begun to enact their own data privacy laws. California has been at the forefront with the California Consumer Privacy Act of 2018, which grants consumers rights over their personal information held by businesses (California Civil Code §§ 1798.100–1798.199, 2018). The California Consumer Privacy Act was further strengthened by the California Privacy Rights Act, enhancing consumer rights and establishing the California Privacy Protection Agency to enforce the regulations (California Civil Code §§ 1798.100–1798.199.100, 2023). Other states, such as Virginia with the Consumer Data Protection Act and Colorado with the Colorado Privacy Act, have enacted similar legislation, signaling a growing recognition of the need for robust data privacy protections at the state level (Va. Code Ann. § 59.1–575 et seq.; Colo. Rev. Stat. § 6–1 – 1301 et seq.).

The reality, however, is that the rapid advancement of generative AI technologies has outpaced legislative and compliance efforts, creating gaps that continue to leave consumer data and privacy at risk. Although no Federal Legislation has yet been enacted, in October 2022, the White House Office of Science and Technology Policy, recognizing the increasing need for regulation, released a Blueprint for an AI Bill of Rights, outlining principles to guide the design and use of automated systems to protect the American public in the age of AI (The White House, 2022). While not legally binding, the blueprint emphasized the importance of data privacy, algorithmic discrimination protections and transparency. Furthermore, on October 30, 2023, President Biden signed an executive order on the Safe, Secure and Trustworthy Development and Use of Artificial Intelligence, emphasizing the need for responsible AI development to harness its benefits while mitigating associated risks (The White House, 2023). The order outlined a comprehensive approach to ensure AI technologies are developed and deployed in a manner that upholds safety, security and public trust. It specifically addresses the protection of consumer data and privacy, by directing agencies to establish guidelines and standards that safeguard personal information in AI

applications. For instance, it directs the National Institute of Standards and Technology to develop guidelines for AI systems to ensure they are safe and secure. Additionally, the order requires that developers of advanced AI systems share the results of safety tests and other critical information with the US Government before public release, aiming to prevent misuse and protect personal data.

Internationally, where privacy is frequently viewed as a fundamental human right, lawmakers have demonstrated they are willing to proactively lead efforts for better data protection for consumers (Layne, 2024). The European Union has taken a proactive stance with the proposed EU AI Act, aiming to establish a comprehensive regulatory framework for AI (European Commission, 2021). Introduced in April 2021, the Act classifies AI systems based on risk levels and imposes obligations accordingly, focusing on transparency, safety and fundamental rights. The EU AI Act represents a significant effort to regulate AI comprehensively, balancing innovation with ethical considerations. It seeks to mitigate risks associated with AI, such as privacy violations and biased outcomes, by enforcing strict compliance requirements for high-risk AI systems. The EU's approach potentially has global implications beyond their European borders, including US-based companies developing generative AI tools, and may well set a precedent to influence future US legislation, especially as concerns over AI's impact on privacy and consumer rights grow (Engler, 2022). King and Meinhardt (2024) however, conclude that the issue lies beyond legislative reach, observing:

While existing and proposed privacy legislation, grounded in globally accepted Fair Information Practices that implicitly regulate AI development, they are not sufficient to address the data acquisition race as well as the resulting individual and systemic privacy harms. Even legislation that contains explicit provisions on algorithmic decision-making and other forms of AI does not provide the data governance measures needed to meaningfully regulate the data used in AI systems (p. 4).

One thing is certain, as AI continues to permeate consumer products and services, the need for effective regulatory and compliance frameworks becomes increasingly urgent. The USA, while demonstrating some limited regulatory advancements at the state level, lacks a cohesive Federal strategy to address the privacy challenges posed by AI technologies. In contrast, the European Union's comprehensive approach offers a potential model for balancing innovation with consumer protection. However, the need for a legislative framework brings with it the tension that such compliance structures must navigate the delicate tension between securing data protection and privacy rights for consumers, while also simultaneously creating flexible regulations that do not stifle technological advancement (Center for Democracy and Technology, 2022).

Beyond legislation: a proposal for ethical artificial intelligence-enabled data collection

Given the speed with which generative AI tools are being developed and implemented, coupled with the slow pace of uniform legislative solutions, the need for self-regulation and voluntary adoption of ethical solutions becomes paramount. Businesses must not merely focus on the race to design and implement generative AI technologies but also need to adopt a consumer-centric approach, ensuring that data collection and privacy considerations are paramount in the design and development of AI technologies from the outset (Granados, 2024).

In response to these challenges, we are now beginning to see the start of a growing movement advocating for the development of a new mode of compliance aimed at protecting consumers from data abuse while simultaneously not stifling technological progression. This

not only includes the establishment of clear ethical guidelines for AI development but also revisits how LLMs are trained to collect data, and what they do with it once it is obtained. The ultimate goal is to create a balanced system that safeguards consumer rights and privacy without hindering the beneficial advancements that AI technologies can offer. If the marketplace is able to navigate this tension and propose meaningful solutions, then the need for Federal oversight potentially becomes redundant. In other words, if the industry is able to self-regulate, then the need for government legislation dissipates.

Self-regulation via codes of ethical conduct

Developing a model of compliance that safeguards consumer privacy without hindering technological progress requires collaborative efforts among industry leaders, and consumer advocates. Transparent practices, ethical guidelines and enforceable regulations are essential to ensure that AI's integration into society respects individual rights and fosters public trust. While a proposed solution of industry self-regulation may be met with cynicism from some, we do have precedent that the marketplace is able to devise widely accepted ethical principles on a voluntary basis. For example, in 1986, widespread concerns regarding irregularities in defense contracting prompted the creation of the Defense Industry Initiative on Business Ethics and Conduct, a voluntary agreement signed by 50 major defense contractors (Kurland, 1993). This initiative was informed by recommendations from a special commission report that highlighted the need for improved ethical practices within the defense sector. Each signatory voluntarily committed to several key measures, including the implementation of a written code of ethics, the establishment of ethics training programs for employees, the development of monitoring mechanisms to detect and prevent misconduct and a pledge to maintain accountability to the public. The Defense Industry Initiative not only marked a critical turning point for ethical reform and compliance within the defense industry but also provided the foundation for a broader governmental influence on business ethics as a whole. Its principles informed the 1991 US Federal Sentencing Guidelines for Corporations, widely regarded as a cornerstone in the evolution of the modern business ethics movement (LeClair, 1997).

Such a code of ethics would need to include transparency, so that brands proactively disclose to consumers what is happening with their personal data. This could, among other things, adopt principles similar to the FTC's four-pronged test for evaluating the suitability of fine print in advertisements (Federal Trade Commission, 2013). This test emphasizes the importance of prominence, ensuring that critical information is displayed in a size and format that users can easily notice and read. It also stresses presentation, requiring that the wording and format are clear and straightforward. Additionally, it considers placement, which requires disclosures to be located where consumers are likely to look. Finally, it takes into account proximity, emphasizing the need to place it close to the claims it qualifies. Voluntarily integrating these principles into a code of ethics would not only align with existing regulatory best practices but also enhance trust and accountability in the use of AI technologies. In an interview for the Harvard Business School, Layne (2024) concluded, "It's more about the private sector setting the pace. If your company can get ahead and help define the privacy rules, that puts you at a huge advantage. Essentially, you're setting the standard, and everyone else will have to catch up" (para. 15).

One of the more recent, and potentially most effective, suggestions for protecting consumer data, is that AI tools ought to require consumers to specifically *opt in* to allow their data to be mined or retained, rather than the present system of forcing them to *opt out*. King and Meinhardt (2024) advocate for such a system, arguing that the solution lies in a shift away from the historical opt-out practice, which frequently mandates users to follow a

convoluted and ambiguous pathway through a myriad of terms and conditions to discover how to notify an organization that they do not wish their information to be stored or used for future marketing purposes. Instead, they propose opt-in data collection, in which generative AI tools must operate through privacy by default strategies rather than the present reverse system, concluding, “An industry default of not collating or scraping for user data would be a much more appropriate, understandable, and it is argued, ethical practice” (p. 34).

McClain *et al.* (2023) observe that when considering how to restore trust in data collection and storage, the challenge is that:

In an era where every click, tap or keystroke leaves a digital trail, Americans remain uneasy and uncertain about their personal data and feel they have little control over how it's used [...] The public largely feels as if they have no control over the data that is collected about them online (Paras. 1 and 6).

Users not only ought to regain control over whether their data is mined but also need to be educated as to how their personal information might be used to train AI tools. Most users have little-to-no comprehension how LLMs frequently operate in the background to scrape data that is frequently entered online or shared on social media (Li *et al.*, 2024). AI tools need to be designed in such a way that humans are able to easily specify their privacy preferences, when interacting with LLM-powered systems, and this ought to be in a usable, convenient, efficient and effective way. Many consumers are simply unaware how their data is being shared. Efforts to balance these interests are already evident in current legislative debates. For instance, the proposed American Data Privacy and Protection Act includes provisions for data minimization, requiring companies to limit data collection to what is necessary for specific purposes, thus addressing privacy concerns without unduly hindering business operations (American Data Privacy and Protection Act, H.R. 8152, 117th Cong. § 101, 2022).

Incorporate moral reasoning into artificial intelligence development and training

As AI capabilities continue to improve and become more complex, there is an increased burden on developers to ensure that ethical protection for consumers also increases at a similar level. Google DeepMind Senior Staff Research Scientist Iason Gabriel highlighted the importance of building ethical compliance alongside AI tools, noting, “Building better AI means building more ethical AI” (Gabriel, 2024). If AI tools and agents are going to increase in capability to the stage where they are beginning to be semi-autonomous or even fully autonomous agents (AGI), then the AI needs to be trained to think morally, and to understand, as best as a machine can understand, that it is required to act in the best interests of the individual it is serving. In other words, as AI capabilities develop there is a greater need for LLMs and AI tools to be trained on not only data and web-based content but also, and perhaps even more particularly, moral values. This is going to be especially important as advances toward AGI are made. As important as constitutional boundaries and parameters are, the reality is that we are already discovering these are inadequate. Digital guardrails and codes of ethics are only part of the solution moving forward. An AI tool should act as an agent on the consumer's behalf, and as an agent ought to comply with and understand the role of fiduciary duties, including things such as confidentiality and acting in the best interest of the principal. Moral intelligence, in other words, is going to be an important factor in AI. One of the ways this could be accomplished is to train AI tools to have an aversion to behaving in any way that would be morally inconsistent with the best interests of the human principle. For example, moral intelligence in relation to data privacy could include a trained

aversion to the abuse or misuse of consumer data. Gorny, writing for Forbes about the need to protect consumer data from AI tools (2024), asserts:

AI tools should be programmed carefully to avoid wading unnecessarily into private terrain. These tools can be taught questions to never ask, including personally identifiable information, and learn to understand what types of information must be redacted immediately (para. 13).

When it comes to AI privacy, one of the responsibilities of generative AI tools ought to be to protect user data and privacy in the same way and to the same level that the user would protect himself (Gabriel, 2024). The White House recognized that AI tools inherently reflect the values of those who create and train them. “AI reflects the principles of the people who build it, the people who use it, and the data upon which it is built” (The White House, 2023).

In addition, inverse reinforcement learning (IRL) might also be considered as part of training AI to understand human moral reasoning. This is a technique where AI learns from human feedback to infer and adopt the underlying preferences and values that guide human interactions. IRL provides a framework for training AI to infer human moral values by observing behavior and deducing the underlying reward function guiding those actions (Oliveira *et al* (2023)). Unlike traditional reinforcement learning, where the reward function is predefined, IRL allows AI to work backward from observed behaviors to determine what people value implicitly. This approach is particularly useful for teaching AI to handle complex and nuanced moral principles, as human ethical reasoning often cannot be easily codified. By analyzing decisions in ethically charged scenarios, such as balancing fairness and utility, the AI infers the implicit values or rewards that drive human decision-making (Jara-Ettinger (2019)).

The application of IRL to moral training of AI tools faces challenges, including ambiguity in human behavior, cultural and individual diversity in moral frameworks and the need to align the inferred values with ethical priorities. For example, observed human actions may reflect biases or incomplete information rather than pure moral principles, complicating the AI’s task of inferring accurate reward functions. Additionally, care must be taken to ensure the AI does not overgeneralize or misapply learned values. When designed effectively, however, IRL allows AI systems to adapt to new ethical contexts, navigate value trade-offs and make decisions that reflect human ethical norms, such as prioritizing safety in autonomous vehicles or fairness in social algorithms (Ng and Russell, 2000). By using IRL, AI systems can move beyond rigid, rule-based morality to better emulate human ethical reasoning in dynamic, real-world contexts. This approach requires robust observation data, carefully designed algorithms and human oversight to validate the learned values and ensure alignment with accepted ethical principles. As a result, IRL has significant potential to contribute to the development of ethical AI systems capable of addressing complex moral challenges across industries.

Conclusion

As generative AI tools revolutionize consumer data collection, the tension between innovation and privacy protection grows increasingly urgent. Balancing the benefits of AI with the ethical imperatives of data privacy and consumer protection is one of the most pressing challenges of our time. AI technologies have enabled unprecedented levels of data mining, often resulting in highly targeted advertising, while simultaneously heightening concerns over surveillance, consent and systemic harm. While the European Union has taken the lead with comprehensive AI regulatory initiatives, the USA lacks a cohesive federal strategy to safeguard consumer data. Addressing these concerns requires a multifaceted approach, combining proactive industry self-regulation with the integration of moral

reasoning into AI training systems. Ultimately, success will depend on voluntary collaboration between businesses and AI developers to design systems that prioritize consumer-centric values, transparency, fairness and trust. By aligning innovation with the safeguarding of consumer rights, this approach ensures that technological progress and ethical responsibility can coexist.

References

- Amazon.com, Inc (2024a), “Amazon to acquire iRobot”, [Press release], available at: <https://press.aboutamazon.com/news-releases/news-release-details/amazon-and-irobot-sign-agreement-amazon-acquire-irobot>
- Amazon.com, Inc (2024b), “Amazon and iRobot agree to terminate pending acquisition [press release]”, available at: www.aboutamazon.com/news/company-news/amazon-irobot-terminate-acquisition
- Bick, A., Blandin, A. and Deming, D. (2024), “The rapid adoption of generative AI”, Federal Reserve Bank of St. Louis, available at: www.stlouisfed.org/on-the-economy/2024/sep/rapid-adoption-generative-ai
- Blaney, R.P. (2024), “Amazon’s recent acquisitions highlight the value of consumer data and evolving privacy concerns”, The National Law Review, available at: www.natlawreview.com/article/amazon-s-recent-acquisitions-highlight-value-consumer-data-and-evolving-privacy
- California Civil Code §§ 1798.100–1798.199 (2018).
- California Civil Code §§ 1798.100–1798.199.100 (2023).
- Center for Democracy and Technology (2022), “AI policy and governance”, available at: <https://cdt.org/area-of-focus/ai-policy-governance>
- Davenport, T.H. and Mittal, N. (2023), “How generative AI is changing creative work”, Harvard Business Review, available at: <https://hbr.org/2023/07/how-generative-ai-is-changing-creative-work>
- Digital Rights Watch (2022), “Amazon’s acquisition of Roomba raises privacy concerns”, available at: <https://digitalrightswatch.org.au/2022/08/08/media-release-amazons-acquisition-of-roomba-raises-privacy-concerns>
- Ellucian (2024), “Ellucian’s AI survey of higher education professionals reveals surge in AI adoption despite concerns around privacy and bias”, Ellucian, available at: www.ellucian.com/news/ellucians-ai-survey-higher-education-professionals-reveals-surge-ai-adoption-despite-concerns
- Engler, A. (2022), “The EU AI act will have global impact, but a limited Brussels effect”, Brookings Institution, available at: www.brookings.edu/articles/the-eu-ai-act-will-have-global-impact-but-a-limited-brussels-effect
- European Commission (2023), “Mergers: commission clears unconditionally the acquisition of iRobot by amazon”, available at: https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_23_3702/IP_23_3702_EN.pdf
- Federal Trade Commission (2013), “Disclosures: how to make effective disclosures in digital advertising”, available at: www.ftc.gov/sites/default/files/attachments/press-releases/ftc-staff-revises-online-advertising-disclosure-guidelines/130312dotcomdisclosures.pdf
- Federal Trade Commission (2024), “Statement regarding the termination of amazon’s proposed acquisition of iRobot”, available at: www.ftc.gov/news-events/news/press-releases/2024/01/statement-regarding-termination-amazons-proposed-acquisition-irobot
- Fell, J. (2024), “Hackers take control of robot vacuums in multiple cities, yell racial slurs”, ABC News, available at: www.abc.net.au/news/2024-10-11/robot-vacuum-yells-racial-slurs-at-family-after-being-hacked/104445408
- Gabriel, I. (2024), “The ethics of AI assistants [audio podcast episode]”, In H. Fry (Producer), Google DeepMind: The Podcast. Spotify, available at: https://open.spotify.com/episode/0jLw5WApj1qW1AQwQkwc9?si=AaNBjIL_SfmpafGnoFx6HQ

- Granados, A. (2024), "AI and personal data: Balancing convenience and privacy risks", Velaro, available at: <https://velaro.com/blog/the-privacy-paradox-of-ai-emerging-challenges-on-personal-data>
- Griswold v. Connecticut (1965), 381 U.S. 479.
- Guo, E. (2022), "A Roomba recorded a woman on the toilet", How did screenshots end up on Facebook? MIT Technology Review, available at: www.technologyreview.com/2022/12/19/1065306/roomba-irobot-robot-vacuums-artificial-intelligence-training-data-privacy
- Jara-Ettinger, J. (2019), "Theory of mind as inverse reinforcement learning", *Current Opinion in Behavioral Sciences*, Vol. 29, pp. 105-110, doi: [10.1016/j.cobeha.2019.04.010](https://doi.org/10.1016/j.cobeha.2019.04.010).
- King, J. and Meinhardt, C. (2024), "Rethinking privacy in the AI era: Policy provocations for a data-centric world [white paper]", The Stanford University Center for Human and Artificial Intelligence, available at: <https://hai.stanford.edu/sites/default/files/2024-02/White-Paper-Rethinking-Privacy-AI-Era.pdf>
- Kurland, N.B. (1993), "The defense industry initiative: ethics, self-regulation, and accountability", *Journal of Business Ethics*, Vol. 12 No. 2, pp. 137-145, doi: [10.1007/BF00871933](https://doi.org/10.1007/BF00871933).
- Layne, R. (2024), "Navigating consumer data privacy in an AI world", Harvard Business School Working Knowledge, available at: <https://hbswk.hbs.edu/item/navigating-consumer-data-privacy-in-an-ai-world>
- LeClair, D.T., Ferrell, O.C. and Ferrell, L. (1997), "Federal sentencing guidelines for organizations: legal, ethical, and public policy issues for international marketing", *Journal of Public Policy and Marketing*, Vol. 16 No. 1, pp. 26-37, available at: www.jstor.org/stable/30000419
- Li, T., Das, S., Lee, H.-P., Wang, D., Yao, B. and Zhang, Z. (2024), "Human-centered privacy research in the age of large language models", arXiv, doi: [10.48550/arXiv.2402.01994](https://doi.org/10.48550/arXiv.2402.01994).
- McClain, C., Faverio, M., Anderson, M. and Park, E. (2023), "How Americans view data privacy", Pew Research Center, available at: www.pewresearch.org/internet/2023/10/18/how-americans-view-data-privacy
- Naprys, E. (2024), "Hacked Ecovacs robot vacuums go berserk yelling racial slurs and chasing dogs", CyberNews, available at: <https://cybernews.com/news/hacked-ecovacs-robot-vacuums-yelling-racial-slurs>
- Ng, A.Y. and Russell, S.J. (2000), "Algorithms for inverse reinforcement learning", Proceedings of the Seventeenth International Conference on Machine Learning, 663–670.
- Oliveira, N., Li, J., Khalvati, K., Barragan, R.C., Reinecke, K., Meltzoff, A.N. and Rao, R.P.N. (2023), "Culturally-attuned moral machines: implicit learning of human value systems by AI through inverse reinforcement learning", arXiv, doi: [10.48550/arXiv.2312.17479](https://doi.org/10.48550/arXiv.2312.17479).
- Shankar, V. (2018), "AI and the future of marketing", *Journal of the Academy of Marketing Science*, Vol. 46 No. 5, pp. 1-8.
- The Robot Report (2022), "iRobot addresses privacy concerns amid pending amazon deal", available at: www.therobotreport.com/irobot-addresses-privacy-concerns-pending-amazon-deal
- The White House (2022), "Blueprint for an AI bill of rights", available at: www.whitehouse.gov/ostp/ai-bill-of-rights
- The White House (2023), "Executive order on the safe, secure, and trustworthy development and use of artificial intelligence", available at: www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence

Further reading

Colorado Privacy Act (CPA) (2023), Colo. Rev. Stat. § 6-1-1301 et seq.

- eCampus News Staff (2024), "Despite privacy concerns, higher ed's AI adoption surges", eCampus News, available at: www.ecampusnews.com/ai-in-education/2024/11/08/higher-education-ai-adoption-surges
- Gorny, T. (2024), "The key to protecting customer privacy while embracing AI-powered experiences", Forbes, available at: www.forbes.com/sites/tomasgorny/2024/09/25/the-key-to-protecting-customer-privacy-while-embracing-ai-powered-experiences
- Kietzmann, J., Paschen, J. and Treen, E. (2022), "Artificial intelligence in advertising: how marketers can leverage AI to enhance customer relationships", *International Journal of Advertising*, Vol. 41 No. 1, pp. 10-20. available at: [a7f00cf28745c60cf9a49f24086f1c5445d7.pdf](https://doi.org/10.1086/ijad.41.1.20220001)
- U.S. House of Representatives (2022), "American data privacy and protection act", H.R. 8152, 117th Cong, available at: [www.congress.gov/bill/117th-congress/house-bill/8152](https://www.congress.gov/bills/117/congress/house/bills/8152)
- Virginia Consumer Data Protection Act (CDPA) (2023), Va. Code Ann. § 59.1-575 et seq.

Corresponding author

Martin Jones can be contacted at: martin.jones@ngu.edu