From Profit to People: Reimagining Data Governance Beyond the Corporate Model

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Trishal Muthan

Spring 2025

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

William J Davis, Department of Engineering and Society

"Data is the new oil" - Clive Humby, 2006

Introduction

Modern society has become deeply and likely irreversibly intertwined with technology, both to our collective advantage and disadvantage. One would be hard-pressed to find people without laptops, without YouTube logins, or without Facebook accounts (*Pew Social Media Fact Sheet,* 2024). It is unlikely people have conversed with others who have never ordered a package via Amazon or used a Maps app on their mobile devices. With all that reliance and usage of technology comes an unbelievably powerful resource: data.

Individuals generate exorbitant amounts of data, whether it be through their phones, computers, TVs, or any plethora of other common technologies present in modern society. In fact, in 2016 there were 2.5 quintillion bytes of data being produced every day, with that number having risen dramatically in the modern day (Joubert, 2016). Rarely, however, do we ever think about how this data is being collected and how it is being used. For example, take location data. People are quick to share their locations with an app if it means better weather forecasts or their dinner being delivered quicker. A *New York Times* study found 75 companies receiving anonymous, precise location data from users who had opted to share their locations, several of which had been tracking up to 200 million smartphone devices in the United States alone. In some cases, this data was accurate to within yards and had been updated 14,000 times *a day* (Valentino-DeVries, 2018).

This data is not being collected for no reason. Similar to how oil served as a foundation of the Industrial Revolution, we are all part of an increasingly data-driven world where data is being extensively used by powerful entities such as companies and governments to shape

decision-making in essentially every aspect of human life, from grocery purchasing to healthcare (Martinez-Martin, 2020, p. 79). As advancements continue in this field, significant concerns emerge regarding privacy, inequality, agency, and ethical responsibility (Martin, 2024). Raw data, much like crude oil, gains its value through extraction and refinement and often in ways that generate immense power and raise serious ethical concerns. Data is no longer just a passive outcome of technological progress; it has become a potent tool capable of shaping societal dynamics, reinforcing existing inequalities, and impacting individual freedoms. The personal data collected through social media, devices, government surveillance, and countless other sources can be harnessed to predict and influence behavior, challenging traditional concepts of privacy and personal autonomy (Quay, 2024, p. 708). This is not even to mention the ascendance of artificial intelligence (AI), which has few boundaries in place to contain and mitigate its potentially harmful effects (Pant et al., 2024, p. 2).

The collection and use of data by corporations and governments are not neutral processes but potentially deliberate exercises of power that reinforce existing inequalities. This paper argues that data governance must shift from a corporate-driven model to one that prioritizes individual agency and collective democratic control. This paper presents a series of case studies through which the issues of data privacy, bias, and governance can be analyzed. The paper then discusses these data issues through the lens of moral theories including utilitarianism, deontology, and virtue ethics to identify common problems that arise. Additionally, the paper aims to put forth possible regulatory approaches that can be used to ensure the responsible collection and usage of data going forward.

Background and Significance

As of 2023, 81% of Americans say they are concerned about how companies are using the data regarding them that is collected. 67% of Americans say they understand little to nothing about what companies are doing with their data, while 73% of Americans believe they have little to no control over what companies or the government do with their data (McClain). Even though technology use is drastically increasing, concerns over data largely seem to be growing, with nearly all associated metrics regarding these data concerns increasing.

Data has become an unavoidable component of our society, shaping our lives in ways that are difficult to understand (Pentland, 2013). Despite the aforementioned concerns over data privacy, individuals continue to share data, often unknowingly, through frequent, ordinary digital interactions. The increasing pervasiveness of smart devices, location tracking, social media, and algorithmic decision-making has created a social environment where privacy is often sacrificed for convenience and innovation. This tension between technical progress and individual rights lies at the core of the ethical problems regarding data practices. While data collection is not inherently problematic, the lack of transparency regarding how data is gathered, stored, and utilized creates an environment ripe for exploitation by powerful entities.

These problems can manifest themselves in many ways in the real world. For example, many social media platforms and online services bury critical information about data practices in lengthy and complicated terms of service agreements. A study in 2023 found that of the 20 most popular websites in 19 countries, the average length of a terms-of-service agreement was 6938 words long while the majority of agreements had a Flesch Reading Ease Score (a metric for reading difficulty) of “difficult” (college-level), with many only readable by graduates (Šlekytė). This sort of practice makes it extremely challenging for the general public to fully grasp what exactly they are consenting to when using these popular platforms. Additionally, it is not

uncommon for companies to employ vague language regarding things like third-party sharing, allowing them to sell or distribute user information to other entities without explicit user awareness.

Companies have also been known to mislead consumers regarding their privacy protections. In 2020, Google was sued for allegedly tracking users in "Incognito Mode" despite its implications of private browsing via advertising tools on websites (Allyn, 2024). Google then used this information to measure web traffic and sell advertisements. When the majority of people use a service told to be secure and private, they do not expect for their data to be extracted and used. This sort of situation is not uncommon and highlights how a lack of transparency can enable corporations to exploit user data while maintaining illusions of privacy. The consequences here can go far beyond the individual level. In 2018, it was revealed that Cambridge Analytica, a political consulting firm, had been harvesting the Facebook data of tens of millions of users without their consent (Confessore). The firm, which had been involved in analytically assessing the campaigns of Donald Trump and Ted Cruz in 2016, had used the data to build psychological profiles of the users to aid various political entities in their elections. Ultimately, Facebook was forced to pay 5 billion dollars in fines and Cambridge Analytica eventually filed for Chapter 7 bankruptcy.

Despite the eventual apprehension of those involved, what is particularly alarming about this situation is the extent of the potential consequences. While often data collection by companies is associated with things like advertising on Google and Amazon or customer analysis, in this case, it was shown to be capable of influencing entire elections. This has become prevalent in increasing capacity as things like disinformation on social media, deepfakes, and targeted propaganda become more and more utilized and impact the public's very understanding

of the truth (Barthel et al., 2016). It is difficult to think about futures in which entities such as corporations are capable of affecting our livelihoods and societies on such a grand scale but this is increasingly becoming the case and, thus, those futures become necessary to consider.

Another destructive outcome lies in the realm of predictive policing. Predictive policing involves the use of algorithmic tools to try and predict those who might commit crimes before they occur. In 2016, an analysis of the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) tool, essentially an algorithm used to determine the risk of arrested individuals committing a crime again in the future, revealed that black people were being given higher likelihoods of risk at a disproportionate rate compared to white people, even when taking into account the types of crimes being committed (Angwin). These risk scores had direct effects on the types of sentences individuals were receiving, therefore leading to black people tending to get longer sentences than others. This sort of bias can very easily arise when there is a lack of transparency regarding how data is being used and the black-box nature of algorithms makes them particularly prone to ethical issues.

To assess the larger ethical problems present in corporate-driven data practices, these sorts of instances of data misuse can be analyzed through the lens of moral theory. Utilitarianism, initially brought forth by philosophers such as Jeremy Bentham and John Stuart Mill posits that the morality of an action is determined by its consequences, aiming to maximize overall happiness while minimizing overall suffering (Mill & Crisp, 1998). Such a theory can be applied as something of a cost-benefit analysis, where the positives and negatives of some particular action are considered and the appropriate action is done which, on a large scale, achieves the greatest good for the most people. In contrast to the outcome-focused analysis of utilitarianism, deontological ethics, most notably associated with Immanuel Kant, centers on duties, principles,

and the inherent rights of individuals (Alexander & Moore, 2007). It emphasizes that certain actions are right and wrong, regardless of outcome, and focuses on following moral rules. A common idea associated with deontology is that people should be treated never as a means to an end but rather as ends in themselves. Another moral theory that can be applied to data practices is virtue ethics. Virtue ethics focuses on the character and virtues of the moral agent rather than worrying about rules or consequences (Hursthouse & Pettigrove, 2003). It advocates for the development of moral virtues such as honesty, integrity, and fairness and guides entities to act in ways that reflect a virtuous character. These ethical frameworks can serve as tools by which we can analyze the actions taken by companies, governments, and other powerful entities and assess the moral problems within those actions. Each theory offers a different lens through which the actions can be examined and will allow us to derive some common issues that persist across the data landscape.

Despite growing awareness of data-related risks, regulatory frameworks have struggled to keep pace with technical advancements and the evolving data landscape. The European Union's General Data Protection Regulation (GDPR) law and the California Customer Privacy Act (CCPA) are examples of significant steps in the right direction, however, there remains much to be desired when it comes to enforcement and extent (Hilliard, 2020, p. 1288). Additionally, data practices tend to largely be approached from technical perspectives in existing literature, attempting to provide specialized solutions to existing issues, yet failing to integrate these methods within the broader societal and ethical considerations at play.

Addressing the challenges of data privacy, bias, and governance requires going a step beyond simply identifying unethical practices. It necessitates a more structured analytical approach to understand why these issues persist despite growing public awareness and how the

issues can be mitigated. By examining these data concerns through various moral frameworks, it is possible to better grasp the broader implications, find common themes and problems among the current corporate-driven environment, and propose meaningful solutions.

Methods

To analyze the ethical implications surrounding data collection, usage, and governance, this paper employs a framework rooted in the aforementioned classical moral theories: utilitarianism, deontology, and virtue ethics. These frameworks offer distinct yet complementary lenses through which to examine the power dynamics, individual rights, and societal consequences of corporate-driven data practices. This ethical approach was selected to move beyond purely technical or policy-based solutions, instead engaging with the deeper questions of responsibility, justice, and human well-being that underlie the current data landscape. The application of moral theory allows for a more holistic, value-driven critique of existing data practices and the formulation of more ethically sound regulatory models.

Taking these moral theories one at a time, it can be shown how each of them provides a unique viewpoint into existing problems with data usage. The utilitarian approach involves assessing data practices based on the balance of benefits and harms they provide for society at large (Mill & Crisp, 1998). For instance, the widespread collection of user data by companies or governments might lead to better technology and more security for individuals which would enhance societal welfare (*Pros and Cons of Big Data*, 2024). However, there exist potential harms such as breaches of privacy, misuse of information, and bias/inequality. A utilitarian viewpoint would critically assess whether these advantages outweigh the downsides and what that means for how data should be allowed to be used. The deontological approach in the context

of data would mandate that entities respect the fundamental rights of users such as their privacy and autonomy (Alexander & Moore, 2007). In this theory, collecting the data of users without their consent would likely be seen as unethical, even if that data were used for some beneficial reason, as doing so would be treating users as simply a means. Virtue ethics might encourage powerful entities to embody characteristics like transparency and responsibility when it comes to the way they manage their collected data and aims to mitigate instances in which things like financial or political incentives are prioritized over ethical behavior (Hursthouse & Pettigrove, 2003). All these theories have potentially differing assessments of what the correct course of action is depending on the specifics of the situation and analyzing data practices through these different ideologies can give valuable insight into how to come up with a more encompassing framework moving forward.

Another critical approach integrated into this analysis is the surveillance model, which provides a framework for understanding how data collection and usage operate as instruments of control. The surveillance model (Lyon, 2014) emphasizes that data collection is not a neutral process but rather one that transforms individuals into observable data points, often without their explicit consent or awareness. This idea of surveillance views data as a rational reflection of reality, passively acquired to monitor behavior, while alternative models like the capture model suggest that the act of data collection itself shapes how individuals interact with technology and society. This paper also draws on the "datafication" model proposed by Mai (2016), which shifts attention from the collection stage to the processing and analysis of data, underscoring how new knowledge is generated from existing data. Incorporating the surveillance model into this research allows for a deeper analysis of how data practices facilitate power imbalances between corporations, governments, and individuals. It highlights how data is often commodified and

used to influence behavior and complements ethical analysis by situating individual rights within larger systems.

The analytical process undertaken follows several key steps to ensure a proper analysis of the issues at hand. First, an extensive literature review was conducted to obtain a better understanding of existing scholarly work on data ethics, surveillance techniques, and regulatory models. This entailed examining academic papers, legal policies, and other scholarly resources to establish a foundational understanding of past analyses in the field. It provided not only background on specific existing ethical concerns surrounding data but also helped identify gaps in past research. Following this, a variety of case studies were analyzed to identify patterns in data governance failures and ethical issues. These case studies were selected based on their relevance to current debates on privacy, algorithmic bias, and moral misuse of data. By examining these cases, the study was able to find broader trends in how data is improperly used for commercial, political, and social purposes. Lastly, the findings were evaluated through the theoretical frameworks mentioned before to draw conclusions about the implications of data usage and its governance. This final stage of analysis ensured that the research was not simply descriptive of past issues but also interpretive and offered insights into how unethical actions can occur in environments where data is involved and what can be done to challenge these past practices and make them more democratic.

This structured approach is appropriate for several reasons. First, by looking at existing problems through the lens of moral theories and surveillance models, the study offers a more human-oriented analysis of data ethics rather than a purely technical assessment. While much of the existing literature on data practices focuses on legal or technical solutions, this research emphasizes the deeper societal consequences of these practices, framing them as ethical concerns

that affect real people. Additionally, the incorporation of multiple different types of cases allows for a more nuanced critique that accounts for both structural and individual agency in data practices. Rather than positioning data collection as an inevitable consequence of increased reliance on technology, this paper investigates the motivations and consequences behind these practices, revealing the deliberate choices made by corporations and governments.

Ultimately, this analytical approach provides a way to interrogate the power dynamics of data practices, making it possible to propose meaningful solutions. By understanding how data functions within social, economic, and political contexts, we can better advocate for policies that prioritize individual rights while balancing the benefits of technological progress (Panai, 2023). Furthermore, this research acknowledges that data ethics is not a single-issue problem but one that intersects with broader societal concerns such as inequality, discrimination, and corporate accountability. Recognizing this is essential for developing effective strategies to combat the exploitation of data and ensure a future in which digital technologies serve humanity rather than control it.

Analysis

One of the most striking failures of corporate data governance is its inability to balance individual benefit with collective benefit. Utilitarian analysis suggests that ethical actions should aim to maximize overall well-being (Mill & Crisp, 1998). There are many potential collective benefits that could theoretically come with corporate data collection. Taking the COMPAS case as an example, in an ideal world such a tool might be useful. If such an algorithm was truly capable of accurate, unbiased predictions of recidivism, there are many potential upsides with respect to safety and security for the general population (Angwin, 2016). Even in general,

data-driven AI models and applications have significantly enhanced fields like healthcare and public services. Additionally, supporters argue that the increased collection of data by large entities allows for greater economic efficiency, improved user experiences, and more innovation. In theory, if these benefits outweigh the harms, corporate data collection could be considered ethically sound. However, reality suggests that the harms tend to outweigh the supposed benefits. COMPAS disproportionately targeted marginalized communities, reinforcing systemic racial and socioeconomic biases rather than increasing any sense of safety or security. The Google Incognito lawsuit revealed how even when companies claim to respect user privacy, they often engage in deceptive practices in their efforts to what they say is innovation but what is actually using the data for their own monetary benefit (Allyn, 2024). Additionally, the Cambridge Analytica scandal showed how large-scale data misuse can go so far as undermining democratic integrity by allowing bad actors to manipulate voter behavior (Confessore, 2018). These examples indicate that corporate data governance, rather than maximizing well-being, often amplifies social harms while prioritizing economic profit. From a utilitarian perspective, this suggests that the current data governance model is fundamentally unethical.

Beyond utilitarianism, corporate data practices also largely fail to align with deontological ethics. Again, according to Kantian ethics, individuals should always be treated as ends in themselves, not merely as means to an end (Alexander & Moore, 2007). In the context of data governance, this means that individuals should have a clear understanding of how their data is collected, processed, and used. However, the reality of modern data collection reveals a fundamental lack of transparency and meaningful consent (Panai, 2023). Many companies employ deceptive interface designs that manipulate users into agreeing to extensive data tracking without full awareness. Moreover, tech corporations often structure their services in ways that

make it very difficult to opt out of data collection (Marcovitch & Rancourt, 2022, p. 1162). The Google Incognito lawsuit underscores this issue, as users were led to believe they were browsing privately when, in reality, Google continued to track their online activity. From a deontological perspective, these practices constitute an ethical failure, as they violate individual autonomy and fail to respect users as rational decision-makers.

In addition to utilitarian and deontological critiques, virtue ethics also provides a lens through which corporate data governance falls short. Ethical governance ought to be guided by values such as integrity, fairness, and responsibility. Yet corporations consistently prioritize profit-driven decision-making over ethical responsibilities to their consumers and the public. These virtues are not arbitrary ideals as they are often explicitly invoked in mission statements and commitments of tech companies. For instance, while Google's mission and branding emphasize user trust and ethical behavior (even once operating under the motto "Don't be evil"), its location tracking practices violated user expectations and transparency norms, revealing a gap between corporate language and action (Nakashima, 2018). When such stated commitments are contradicted by practice, the ethical failure is not just theoretical but a failure to uphold the very values corporations publicly align themselves with. Additionally, rather than fostering trust, these companies often make use of exploitative practices such as misleading privacy policies and evasion of accountability through legal loopholes (Özçetin & Wiltse, 2023, p. 191). Consider the previously mentioned data on terms-of-service agreements. These types of documents are typically where critical information is written regarding data privacy and usage yet they remain incredibly challenging to parse through due to their length and reading difficulty. Companies know that, due to these challenges, few people really take the time to read through all this content. This type of practice is deceptive and is not in line with the sort of behavior one would

expect based on virtue ethics. In another instance, the case of predictive policing algorithms highlights how AI-based systems can reinforce discrimination while remaining largely unregulated and opaque (Angwin, 2016). If ethical leadership in the corporate tech sector were truly a priority, firms would adopt more transparent, fair, and accountable data practices. However, the repeated ethical failures of companies like Facebook and Google indicate an absence of moral responsibility in corporate-driven data governance.

What becomes obvious when viewing corporate data governance through the lenses of utilitarianism, deontology, and virtue ethics is that the problems are not isolated but rather intrinsic and systemic to the corporate model itself. Across all three moral frameworks, the common thread that becomes apparent is the prioritization of profit and success over ethical principles. Whether the goal is to maximize shareholder value, optimize user engagement, or avoid legal entanglements, the central incentive structure of for-profit entities encourages practices that disregard collective welfare, individual autonomy, and moral integrity. In each ethical approach, the failure is almost always due to consumers (or data producers generally) being treated not as stakeholders but as resources. They, therefore, are subject to the same treatment as the most corporate resources; they can be quantified, abstracted away, and commodified (Canellopoulou-Bottis & Bouchagiar, 2018, p. 206). This perspective on people inherently is in conflict with any ethical model because it fails to take seriously both the dignity of individuals and the collective good. Whether it be the utilitarian harms of inequality, the deontological violations of individual agency, or the absence of corporate virtue, the problems come from a logic that treats data as a private asset rather than a shared social good. The previously identified failings are not merely a result of poor implementation or isolated incidents but structural outcomes of a system designed to extract rather than serve (Crain, 2016). Thus, it

is not enough to focus on individual unethical practices and it is the corporate model of data governance itself that necessitates reevaluation.

Rather than treating data as a private asset controlled by corporations, democratic governance models frame data as a collective resource to be managed in the public interest (Viljoen, 2021, p. 583). These models aim to reintroduce ethical responsibility by incorporating values like consent, fairness, and accountability into the infrastructure of data management itself. One example of such a model is a data trust, a legal structure that holds data on behalf of a group and makes decisions based on fiduciary responsibility to the actual members of that group (Artyushina, 2021). By design, this shifts decision-making away from corporate entities and towards more transparent, community-based oversight. Another approach is the data commons, systems that use participatory rule-setting and monitoring to ensure that data use is ethical and equitable (Jeong et al., 2022). Initiatives like Barcelona's DECODE project illustrate how communities can directly govern how their data is shared and used, reinforcing democratic control and public trust (*DECODE*, 2023). Another similar alternative is the data cooperative, where individuals voluntarily pool their data and vote on its use, sharing in the value it creates (Ungureanu et al., 2024). This model challenges the centralized power of tech companies by redistributing both control and benefits directly back to the people generating the data. Early examples in health and agriculture suggest that such cooperatives can produce innovation while respecting autonomy and ensuring fairness. Unlike profit-driven systems that treat users as means to an end, democratic governance approaches attempt to institutionalize ethical principles in how data is handled, used, and valued. In doing so, they offer a pathway toward a data ecosystem where human dignity and social good are central, not incidental. As described before, although regulations such as GDPR and CCPA exist, current laws largely fail to address the more

systemic issue at play when misconduct occurs, often targeting symptoms of the issue rather than the cause. Moving forward, it is critical that more focus is placed on the underlying power differences between corporations and individuals, and on building governance structures that prioritize democratic approach in the management of data.

Conclusion

Corporate data governance today fails to meet core ethical standards across utilitarian, deontological, and virtue ethics frameworks. While data is often promoted as a driver of innovation, its collection and use under corporate control tend to prioritize profit over public welfare. As seen in cases like COMPAS, Google's Incognito dilemma, and Cambridge Analytica, individuals are routinely treated as resources rather than real people. The result is a governance model that amplifies social harm, undermines autonomy, and lacks integrity. These failures are not incidental but structural, revealing a misalignment between corporate incentives and ethical responsibilities.

To address this, data must be reframed not as a commodity but as a collective resource governed in the public interest. Democratic models such as data trusts, commons, and cooperatives offer frameworks for redistributing control and embedding ethical values like transparency, fairness, and consent into data infrastructures. These alternatives promote shared oversight and emphasize the role of communities in determining how data is used and for whose benefit.

Still, implementing such models is not without challenges. Legal, institutional, and cultural barriers including corporate resistance and public disengagement make it difficult to quickly implement these things. Additionally, ethical data systems must still confront issues like

bias, surveillance, and digital inequality. Yet despite these limitations, the need for structural change is clear. Democratizing data governance is not just a technical task but a moral one, and it demands that we prioritize accountability and human dignity in our increasingly datafied world.

References

Alexander, L., & Moore, M. (2007, November 21). *Deontological ethics*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/ethics-deontological/

Allyn, B. (2024, April 1). Google to delete search data of millions who used "incognito" mode. *NPR*. https://www.npr.org/2024/04/01/1242019127/google-incognito-mode-settlement-search-history

Angwin, J., Larson, J., Kirchner, L., & Mattu, S. (2016, May 23). Machine bias. *ProPublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Artyushina, A. (2021, June 10). *The future of data trusts and the global race to dominate AI*. Bennett Institute for Public Policy. https://www.bennettinstitute.cam.ac.uk/blog/data-trusts1/

Barthel, M., Mitchell, A., & Holcomb, J. (2016, December 15). *Many Americans believe fake news is sowing confusion.* Pew Research Center. https://www.pewresearch.org/journalism/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/

Canellopoulou-Bottis, M., & Bouchagiar, G. (2018). Personal data v. big data: Challenges of commodification of personal data. *Open Journal of Philosophy*, *08*, 206–215. https://doi.org/10.4236/ojpp.2018.83015

Confessore, N. (2018, April 4). Cambridge Analytica and Facebook: The scandal and the fallout so far. *The New York Times*. https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

Crain, M. (2016). The limits of transparency: Data brokers and commodification. *New Media & Society*, *20*(1), 88–104. https://doi.org/10.1177/1461444816657096

DECODE. (2023, November 14). Barcelona Digital City. https://ajuntament.barcelona.cat/digital/en/technology-accessible-everyone/accessible-and-participatory/decode

Duhigg, C. (2012, February 16). How companies learn your secrets. *The New York Times*. https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html

Hilliard, E. (2020). The GDPR: A retrospective and prospective look at the first two years. *Berkeley Technology Law Journal, 35*(4), 1245–1289. https://doi.org/10.15779/Z384J09Z3K

Holder, A. E. (2020). What we don't know they know: What to do about inferences in European and California data protection law. *Berkeley Technology Law Journal, 35*(4), 1331–1364. https://doi.org/10.15779/Z38MP4VP1V

Hursthouse, R., & Pettigrove, G. (2003, July 18). *Virtue ethics*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/ethics-virtue/

Jeong, J. C., Hands, I., Kolesar, J. M., Rao, M., Davis, B., Dobyns, Y., Hurt-Mueller, J., Levens, J., Gregory, J., Williams, J., Witt, L., Kim, E. M., Burton, C., Elbiheary, A. A., Chang,

M., & Durbin, E. B. (2022). Local data commons: the sleeping beauty in the community of data commons. *BMC Bioinformatics, 23*, 1–20. https://doi.org/10.1186/s12859-022-04922-5

Joubert, S. (2016, May 13). *How much data is produced every day?*. Northeastern University Graduate Programs. https://graduate.northeastern.edu/knowledge-hub/how-much-data-produced-every-day

Martin, B. (2024). [Review of the book *Data cartels: The companies that control and monopolize our information*, by S. Lamdan]. *Prometheus, 40*(1), 56–59. https://www.jstor.org/stable/48813346

Lyon, D. (2014). Surveillance, Snowden, and big data: Capacities, consequences, critique. *Big Data & Society, 1(*2). https://doi.org/10.1177/2053951714541861

Mai, J. E. (2016). Big data privacy: The datafication of personal information. *Information Society, 32*(3), 192–199. https://doi.org/10.1080/01972243.2016.1153010

Marcovitch, I., & Rancourt, E. (2022). A data ethics framework for responsible responsive organizations in the digital world. *Statistical Journal of the IAOS*, *38*(4), 1161–1172. https://doi.org/10.3233/sji-220067

Martinez-Martin, N. (2020). Big data, corporate surveillance and public health. *The American Journal of Bioethics, 20*(10), 79–81. https://doi.org/10.1080/15265161.2020.1806394

McClain, C., Faverio, M., Anderson, M., & Park, E. (2023, October 18). *How Americans view data privacy*. Pew Research Center. https://www.pewresearch.org/internet/2023/10/18/how-americans-view-data-privacy/

Mill, J. S., & Crisp, R. (1998). *Utilitarianism*. Oxford University Press.

Nakashima, R. (2021, April 20). Google tracks your movements, like it or not. *AP News*. https://apnews.com/article/828aefab64d4411bac257a07c1af0ecb

Özçetin, S., & Wiltse, H. (2023). Terms of entanglement: A posthumanist reading of terms of service. *Human–Computer Interaction*, *40*(1–4), 171–194. https://doi.org/10.1080/07370024.2023.2281928

Panai, E. (2023). The latent space of data ethics. *AI & Society, 39*, 2647–2665. https://doi.org/10.1007/s00146-023-01757-3

Pant, A., Hoda, R., Spiegler, S., Tantithamthavorn, C., & Turhan, B. (2024). Raising AI ethics awareness through an AI ethics quiz for software practitioners. *ACM Transactions on Software Engineering & Methodology, 33*(3), 1–35. https://doi.org/ 10.1145/3635715

Pentland, A. (2013). The data-driven society. *Scientific American*, *309*(4), 78–83.

Pew Research Center. (2024, November 13). *Social media fact sheet*. https://www.pewresearch.org/internet/fact-sheet/social-media/

Pros and cons of big data. (2024, February 28). Harvard Online. https://www.harvardonline.harvard.edu/blog/pros-cons-big-data

Quay, A. (2024). Desperation for legislation: The need for the American data privacy and protection act. *Wisconsin International Law Journal, 41*(4), 707–729. https://doi.org/10.59015/wilj.rrzv1592

Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 42–47. https://doi.org/10.1109/CTS.2013.6567202

Šlekytė, I. (2023, October 23). *NORDVPN study shows: Nine hours to read the privacy policies of the 20 most visited websites in the US*. NordVPN. https://nordvpn.com/blog/privacy-policy-study-us/

Ungureanu, C. T., Gheorghiu, A., & Greavu-Șerban, V. (2024). A 3D approach on European data cooperative as an intermediation service. *Eastern Journal of European Studies*, *15*(2), 48–68. https://doi.org/10.47743/ejes-2024-0204

Valentino-DeVries, J., Singer, N., Keller, M. H., & Krolik, A. (2018, December 10). Your apps know where you were last night, and they're not keeping it secret. *The New York Times*. https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html

Viljoen, S. (2021). A relational theory of data governance. *The Yale Law Journal*, *131*(2), 573–654. https://www.jstor.org/stable/45400961

Zwitter, A. (2014). Big data ethics. *Big Data & Society, 1*(2). https://doi.org/10.1177/2053951714559253