Ethics in Data Science: A Reflection on Simon Blackburn's Being Good

Brandon Nguyen

University of Houston Downtown

Data Science Ethics Final Presentation

Dr. Ryan Pepper

April 20, 2025

Table of Contents

Abstract

Simon Blackburn’s Being Good (2001) offers foundational insights into moral philosophy relevant to data science. This paper applies Blackburn’s treatment of utilitarianism, deontology, and virtue ethics to AI/ML contexts, examines algorithmic bias as false consciousness, and evaluates privacy, trolley style dilemmas, and fairness through universal human rights. Reconstructing categorical imperatives and social contracts, it proposes ethics by design, interdisciplinary collaboration, and continuous oversight to align innovation with human dignity.

Introduction

Data science and machine learning increasingly power decisions that affect individuals, communities, and societies. From predictive policing to medical diagnostics, these systems promise efficiency and objectivity. Yet technical performance metrics—accuracy, precision, recall, tell only part of the story. They rarely capture the broader ethical dimensions: whose data is collected, how biases propagate, who bears the risks, and who reaps the benefits. Blackburn's Being Good (2001) reminds us that ethics is not a peripheral concern but the very environment shaping our choices. His portrayal of the "death of God" (Chapter 1) and the subsequent need to find human foundations for morality resonates with data science's imperative to find non-theistic yet robust guides to right action. This paper adopts Blackburn's structure, surveying major ethical theories, uncovering threats to moral coherence, and rebuilding foundations, to articulate a comprehensive ethics framework tailored to data driven practices.

Ethical Theory

Moral philosophy offers several enduring frameworks for evaluating right and wrong, each with its own view on how ethical decisions should be made. Understanding these theories is crucial for data scientists, who increasingly find themselves shaping technologies with profound social implications.

Utilitarianism, advanced by Jeremy Bentham and John Stuart Mill, judge's actions by their consequences, specifically, the extent to which they promote overall happiness or reduce suffering (Blackburn, 2001, Ch. 12). This consequentialist logic dominates many areas of applied machine learning: algorithms are typically optimized to maximize accuracy, efficiency, or throughput. In fraud detection, for example, a model that minimizes false negatives may be considered ideal, even if it disproportionately flags transactions from specific demographic groups. However, utilitarian metrics like overall accuracy can obscure the experiences of minority populations. An AI diagnostic tool might achieve 95% accuracy by performing well in common cases but severely underperforming on rare or marginalized subgroups. This "aggregate utility" approach, while efficient, risks enacting harm through statistical neglect. In response, critics call for "distributional justice" ensuring that the benefits and burdens of algorithmic systems are shared equitably, not merely averaged. As such, utilitarianism's strengths in scalability and clarity must be tempered with fairness aware metrics like group fairness or equalized odds.

Deontology, most famously associated with Immanuel Kant, focuses not on outcomes but on moral duties and principles. According to Kant, ethical actions are those that respect persons as ends in themselves, never merely as means to an end (Blackburn, 2001, Ch. 18). In the data ethics domain, this translates into respect for individual rights, such as informed consent, data ownership, and autonomy in automated decisions. European legislation such as the General Data Protection Regulation (GDPR) embodies deontological reasoning: individuals have the right to know when they're subject to algorithmic decisions, and to request human review. A facial recognition system that scans unaware pedestrians in public spaces may be technically effective, but it violates the Kantian imperative if it bypasses informed consent. While deontology offers a strong ethical spine, it may be rigid in crisis scenarios. For example, withholding personal data in the name of consent might hamper timely public health interventions during a pandemic. Thus, the challenge for data scientists is to develop generalizable rules—like transparency and reversibility—that honor deontological commitments while remaining pragmatic.

Virtue Ethics, rooted in Aristotle and revived in modern moral thought, shifts focus from rules and consequences to character. It asks: What kind of person should I be? rather than What should I do? (Blackburn, 2001, Ch. 17). For data scientists, this means cultivating habits like intellectual humility, empathy, and courage. A virtuous practitioner resists overconfidence in predictive outputs, questions unjust assumptions in data labeling, and proactively raises concerns about algorithmic harm. Organizations benefit when virtue is embedded in culture, through ethical leadership, inclusive feedback structures, and time for reflection. However, virtue ethics can be subjective and difficult to institutionalize. Without external accountability, appeals to “good intentions” may mask systemic bias or negligence. Thus, while virtue ethics offers a compelling framework for ethical character, it must be supported by structural safeguards and a community of moral practice.

Together, these three ethical theories form a robust lens through which to evaluate AI systems. Utilitarianism offers efficiency and scalability, deontology secures fundamental rights, and virtue ethics centers moral character and judgment. None is sufficient on their own, but in combination they help practitioners navigate the complexity and responsibility of ethical data science.

Real-World Problems in Data Science

Algorithmic Bias and False Consciousness.

In Chapter 7 of Being Good, Blackburn introduces the idea of “false consciousness,” a condition where individuals accept unjust systems as normal or even beneficial. This concept is reflected in the phenomenon of algorithmic bias. Predictive policing systems, for instance, often rely on arrest data that reflects historical discrimination. As a result, these systems perpetuate over-policing in certain communities by flagging them as high-risk. This increases police presence and arrests, reinforcing the system’s flawed assumptions.

The problem is worsened when people trust the algorithm simply because it is data-driven. Blackburn’s discussion warns against this kind of blind trust. To address the issue, data scientists must go beyond performance metrics and look at the broader social implications of their models. This includes conducting bias audits, involving community voices in system design, and applying fairness techniques that reduce discriminatory impacts. Recognizing that bias stems from deeper societal patterns, not just technical shortcomings—is a necessary step toward ethical reform.

Privacy, Consent, and Human Rights

Blackburn explores the concept of rights and paternalism in Chapters 14 and 15. These ideas relate closely to digital privacy in modern data science. Article 12 of the Universal Declaration of Human Rights protects individuals from arbitrary interference in their personal lives. Yet, this right is often undermined in practice. A well-known case is the

Cambridge Analytica scandal, where Facebook data was harvested without clear consent and used for political profiling. This incident highlights the dangers of companies collecting personal data under vague terms of service and repurpose it for profit.

Ethical systems must return control to users. Privacy-centered designs require minimal data collection, clear disclosures, and user-friendly consent tools. In high-stakes cases, such as public health data or biometric surveillance, additional layers of oversight are needed to balance individual rights with collective benefit. Blackburn’s critique reminds us that protecting autonomy is not optional but fundamental to ethical decision-making.

Trolley Dilemmas in Autonomous Systems. In Chapter 6, Blackburn discusses moral dilemmas that stretch our ethical reasoning. One such dilemma has gained real-world relevance through autonomous vehicles. In critical situations, these vehicles may need to choose between different outcomes, such as protecting passengers or bystanders. The trolley problem is no longer hypothetical.

A utilitarian view would seek to reduce total harm, while a deontological stance would object to harming innocents under any condition. Cultural views on these issues vary, making it difficult to design a universal response. To address this, engineers and policymakers are exploring ways to clearly define the ethical priorities of automated systems. These include setting default decision policies and giving users limited customization. Transparency is essential. People must understand how a system might behave in difficult situations to ensure public trust.

Fairness, Representation, and Universal Rights. In Chapter 21, Blackburn argues against moral relativism and supports the idea of universal ethical principles. This is directly relevant to fairness in artificial intelligence. Many machine learning models have shown performance gaps across race and gender groups. A common example is facial recognition technology, which has been less accurate for individuals with darker skin tones.

These issues violate the spirit of Article 7 of the Universal Declaration of Human Rights, which affirms equal protection for all. Addressing these disparities requires more than better data. It requires ethical reflection on how systems define fairness and whose voices are included in that process. Practical steps include balancing datasets, applying fairness-aware algorithms, and assessing impact across diverse user groups. Fairness in AI should be a shared moral goal, not just a technical target.

Responsibility, Accountability, and Social Contracts

Chapter 19 of Being Good explores ethics through the lens of shared agreements. In data science, this raises the question of who is responsible when technology causes harm. It might be tempting to blame individual developers or users, but responsibility often extends across an entire system.

Ethical governance tools can help. Model documentation practices like model cards and dataset reports clarify a system’s purpose and limitations. Logging systems provide traceability, making it easier to investigate errors or misuse. Public ethics boards, composed of experts and community members, can offer additional reviews before deployment. As AI systems become more influential, shared responsibility must become part of their design. Institutions need to embed accountability into every phase of development, ensuring that ethical considerations are not left to chance.

Solution and Reflection

Rebuilding Foundations and Restoring Confidence.

In the final chapters of Being Good, Blackburn turns his attention to the broader challenges facing modern ethical discourse. He explores how moral skepticism, relativism, and cynicism can undermine public trust in any shared sense of right and wrong. This diagnosis is especially relevant to data science, where public confidence in algorithms is increasingly strained by revelations of bias, opacity, and misuse. To counter this erosion of trust, Blackburn calls for a return to ethical reasoning grounded in open dialogue, mutual respect, and a commitment to human dignity.

Applying this to data science requires more than individual awareness. It demands a structural shift in how we design, develop, and deploy technological systems. Ethics must be embedded from the beginning of a project and maintained throughout its life cycle. This includes implementing ethical design principles at the earliest stages of model development, where decisions about data collection, labeling, and feature selection can have profound downstream effects. Additionally, interdisciplinary collaboration is essential. No single discipline holds all the answers to complex ethical problems. Data scientists should work alongside ethicists, sociologists, legal scholars, and members of affected communities to identify risks, surface blind spots, and define acceptable trade-offs.

Ongoing education is another pillar of ethical practice. Just as professionals in law or medicine receive regular ethics training, those who build and maintain AI systems should continuously engage with emerging ethical frameworks and case studies. Institutions must support this with formal training programs, ethical review boards, and open forums for critical discussion. These structures help normalize ethical reflection as part of technical excellence, rather than an afterthought.

Blackburn's Being Good equips data scientists with a practical and philosophical foundation. His work offers not a strict rulebook, but a flexible toolkit: the consequential reasoning of utilitarianism, the principled commitment of deontology, and the introspective discipline of virtue ethics. These perspectives are invaluable when confronting the moral complexity of modern AI systems. Issues like algorithmic bias, data privacy, autonomous decision-making, and representational fairness cannot be resolved through code alone. They require clear ethical reasoning and the courage to act on it.

As someone entering the field of data science, I see ethical responsibility not as a limitation, but as a source of purpose. It provides a guiding compass when trade-offs arise, when business incentives clash with social values, or when technical uncertainty makes decisions difficult. I pledge to approach each project with humility, to remain vigilant against systemic harm, and to seek collaboration across boundaries. My goal is not only to build systems that function, but to ensure they reflect the values and respect the dignity of the people they affect. Ethics in data science must be a continuous, shared practice—an ongoing effort to align innovation with justice, inclusion, and trust.

References

Blackburn, S. (2001). Being Good: A Short Introduction to Ethics. Oxford University Press.

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86–92.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019). Model cards for model reporting. In Proceedings of the Conference on Fairness, Accountability, and Transparency (pp. 220–229).

United Nations. (1948). Universal Declaration of Human Rights. https://www.un.org/en/about us/universal declaration of human rights