# Exploring or Exploiting? Social and Ethical Implications of Autonomous Experimentation in AI

Sarah Bird Solon Barocas Kate Crawford

Fernando Diaz Hanna Wallach

Microsoft Research
{slbird,solon,kate,fdiaz,wallach}@microsoft.com

## ABSTRACT

In the field of computer science, large-scale experimentation on users is not new. However, driven by advances in artificial intelligence, novel autonomous systems for experimentation are emerging that raise complex, unanswered questions for the field. Some of these questions are computational, while others relate to the social and ethical implications of these systems. We see these normative questions as urgent because they pertain to critical infrastructure upon which large populations depend, such as transportation and healthcare. Although experimentation on widely used online platforms like Facebook has stoked controversy in recent years, the unique risks posed by *autonomous* experimentation have not received sufficient attention, even though such techniques are being trialled on a massive scale. In this paper, we identify several questions about the social and ethical implications of autonomous experimentation systems. These questions concern the design of such systems, their effects on users, and their resistance to some common mitigations.

## 1. INTRODUCTION

Many computer scientists see experiments on users as a necessary step toward improving products and services. Indeed, user experimentation, such as classic A/B testing, has become a normalized part of the technology sector. But, thanks to advances in artificial intelligence, we are now at the beginning of a new phase of experimentation involving autonomous systems. As these systems spread throughout our lives, affecting even critical infrastructure, they raise difficult questions about the ethics of autonomous experimentation practices and their wider social implications.

Consider, for example, navigation services that are responsible for providing millions of users with real-time directions. Given the current traffic conditions, these services attempt to suggest optimal routes for drivers. Experimentation is likely a core part of suggesting optimal routes. This is because service providers often lack information about traffic conditions on those routes to which they have purposefully *not* directed drivers. To determine whether a previously slow route is still slow, these services will deliberately send some users along it. Although such experiments may have beneficial effects for the system as a whole, they can be problematic for individual users or groups of users. For some users, taking a slow route might mean that they are slightly late for work; for others, though, it might delay a trip to the hospital. Moreover, users seldom know whether they are part of an experiment, nor do they have any way to convey that one journey is more urgent than another.

Alternatively, consider ad placement systems intended to increase click-through rates and, ultimately, revenue. These systems display combinations of ads to users in order to determine which combinations are most effective. In isolation, each ad may seem completely innocuous, but, together, these combinations and users' responses may reveal privacy-violating information or, worse yet, uniquely identify users.

As these examples illustrate, we are witnessing the advent of new autonomous experimentation systems that are intended to maximize efficiency and seize opportunities to learn, at the cost of providing some users with a sub-optimal experience. In the aggregate, this approach results in sophisticated systems that can rapidly adapt to changing conditions. But it requires that users contribute to the "greater good" of the system—that they ask not what their algorithm can do for them, but what they can do for their algorithm.

In this paper, we set out to identify the complex ethical questions raised by autonomous experimentation. Some of these questions concern privacy, although for novel reasons; other questions are local, concerning the disadvantages experienced by individual users or groups of users; finally, some are more general and relate to informed consent, the reduced agency of users, and increased power asymmetries.

## 2. WHEN MACHINE LEARNING MEETS EXPERIMENTATION

Autonomous experimentation systems draw on both machine learning and experimentation. Machine learning is a subfield of AI that is concerned with modeling observed data, either to uncover meaningful patterns hidden in the data or to make predictions about future, yet-to-be-observed data. "Training" a machine learning system often requires a large amount of data. Experimentation provides system designers with a way to leverage user responses to evaluate different design decisions, settings, and algorithms (even machine learning algorithms). Unfortunately, both machine learning and experimentation can expose users to risks. Machine learning systems can reflect and reinforce any biases that are present in the training data [6], while experimen-

tation can expose users to experimental treatments that are not in their best interests or to which they would not have knowingly consented. Both machine learning and experimentation can threaten users' privacy, but in different ways.

Although researchers have begun to acknowledge and address these kinds of risks, focusing separately on machine learning [4] and experimentation [3], new systems are being deployed that combine machine learning and experimentation via multiworld testing, interactive learning, explore–exploit, and reinforcement learning algorithms. Search engines were early prototypes for these kinds of systems [19, 9], but increasingly they are found even in critical infrastructure, such as transportation [13] and healthcare [11].

These new autonomous experimentation systems are inherently adaptive and learn by conducting experiments without human intervention. Typically, the experiments are relatively crude and involve sequences of potentially sub-optimal actions in order to explore the relationship between actions and rewards. One of the most well-known approaches is reinforcement learning—a mature subfield of machine learning. Reinforcement learning systems learn how software agents should take actions in some changing environment in order to maximize some long-term reward. Reinforcement learning recently played a crucial role in helping a computer defeat a highly ranked human Go player [16]. In this context, the environment is the state of the board, the actions are all possible stone placements, and the reward is winning the game. But reinforcement learning has also been deployed in online platforms. Here, for example, the environment might be user attributes, the actions might be be various ad placements, and the reward might be increasing revenue [5, 17]. Roughly speaking, reinforcement learning systems autonomously adapt in the same way that other systems are iteratively refined via manual A/B testing; however, the number and range of experiments are much larger and the experiments are performed at a considerably faster speed.

## 3. PROBING PRIVACY

Machine learning systems learn patterns and associations that are present in the data on which they are trained. Crucially, they are only able to learn patterns and associations for which there are sufficiently many examples. Standard machine learning systems, involving supervised or unsupervised learning, are therefore limited in their ability to learn privacy-violating patterns and associations (e.g., predicting the sexual orientation of particular social media users [12]) by the number of supporting examples that are present in the training data. In contrast, autonomous experimentation systems are not restricted to learning from historical, already-observed data. These systems perform experiments to obtain new data about users' responses in situations for which there are few or no examples in the existing training data. As a result, they are capable of uncovering user-specific insights that may be sensitive or privacy violating.

Autonomous experimentation overcomes the primary limitation of observational studies—i.e., their inability to identify causal relationships. Although causal inference from observed data is an active area of research in machine learning and related fields, these methods are less powerful than randomized, controlled experiments that can isolate the effects of an experimental treatment. Because autonomous experimentation systems can establish that some treatment *causes* a user to respond in a particular way, they can reveal an enormous amount of information. For example, such systems might discover that there are specific combinations of ads that are much more likely to cause a user to make a purchase. Together, these combinations may uniquely identify that user. Finally, sequential variations in experimental treatments can generate cumulative causal insights that far exceed users' expectations and even their levels of consent.

We argue that autonomous experimentation can be actively *hostile* to users' privacy because it is capable of uncovering user-specific insights in novel, uncertain situations.

## 4. ETHICAL PRINCIPLES FOR EXPERIMENTATION

Moving beyond privacy, autonomous experimentation systems also raise new and difficult questions about broader ethical frameworks, especially as these systems spread throughout our lives. The ongoing debate about existing research ethics regululations and their applicability to data science is especially relevant. The U.S. Department of Health and Human Services recently released a notice of proposed rule making (NPRM) that proposes revisions to the Common Rule that more effectively cover data-intensive research [8]. Although this NPRM characterizes such research as inherently "low risk"—a contentious issue in and of itself [14]—it notes that relationships between researchers, experimental subjects, and data are in flux, with 1) subjects caring more than ever about data management, 2) researchers being able to access data without directly interacting with subjects, and 3) the risk profile of human-subjects data changing unpredictably. Given these observations and their relevance to autonomous experimentation, we argue that it is illustrative to re-examine the basic ethical principles that have constrained research experimentation on humans in the past.

The Common Rule, or Federal Policy for the Protection of Human Subjects, was established in response to a series of breaches to the public trust, and draws upon the Nuremberg Code, the Declaration of Helsinki, and the Belmont Report.

The Nuremberg Code [2] was created after the atrocities of World War II as a means to define ethical norms for human-subjects research. It laid out many standard principles of ethical research, including requiring subjects to give consent and maintaining a balance between potential risks and benefits. Although it was not specifically codified in U.S. law, it was the first document to advocate informed consent.

The Declaration of Helsinki [1] was developed by the World Medical Association as a guide for the medical community. It draws upon the Nuremberg Code, but also stipulates that human experiments must be grounded in animal trials, that researchers must be medically and scientifically qualified, and that research protocols must be independently reviewed.

These principles were enacted in U.S. law by the National Research Act in 1974. The National Resarch Act created the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, which subsequently published the Belmont Report in 1979. This report established three basic ethical principles: respect for persons, beneficence, and justice. Respect for persons protects human autonomy, allows for informed consent, and requires researchers to be truthful. Beneficence embodies the idea of "do no harm," with the aim of minimizing any risks to research subjects while maximizing research benefits. Finally, justice addresses the process of distributing risks and bene-

fits to potential research subjects in a manner that is fair.

We argue that autonomous experimentation challenges these basic ethical principles. Although the Common Rule does not have any legal force outside of government-funded research, it is widely viewed as a baseline for ethical conduct, and we see its violation as a cause for concern. This viewpoint is loosely echoed by companies that have recently drafted review procedures that look to these basic ethical principles to guide their interactions with users [18, 15]. In the remainder of this paper, we therefore examine autonomous experimentation in the context of respect for persons, beneficence, and justice, and outline some of the obstacles to realizing these basic ethical principles in practice.

# 5. VIOLATING ETHICAL PRINCIPLES

As we described in section 2, autonomous experimentation systems conduct experiments that involve sequences of potentially sub-optimal actions in order to explore the relationship between actions and outcomes. By subjecting users to potentially sub-optimal actions, some users will experience inconveniences or risks not experienced by others. For example, by deliberately sending some users along a previously slow route to determine whether it is still slow, navigation services expose these users to risks, even though the goal is to improve the system as a whole. This disparity raises the question of who should serve as an "explorer" and who should "exploit" the information discovered. One obvious answer is to require that explorers be selected uniformly at random. Unfortunately, even though this selection procedure is technically unbiased, it can still violate the ethical principle of respect for persons if some users are unknowingly enlisted into the process of taking potentially sub-optimal actions.

We argue that in the absence of informed consent, such experiments take advantage of users' ignorance and potentially direct them to engage in activities that depart from their goals, preferences, and expectations. A person who will lose their job if they are late to work might decline an invitation to serve as an explorer, even if there were some chance that they would arrive at work earlier than expected. Other users might prefer to take a route with an uncertain duration over one that is certain to take a long time. Respect for persons dictates the importance of allowing users to make such decisions for themselves in an informed manner—that exploration will discover information that improves the system as a whole may not justify the risks for any specific user.

Even systems that do not have a non-experimental default setting—such as those that rely on confidence-based sampling [7]—will subject some users to actions that are less certain to be optimal. These users will serve as more adventurous explorers, and may be exposed to greater risks.

In practice, explorers do not need to be selected uniformly at random. Moreover, because exploration can be driven by uncertainty and uncertainty arises from a lack of information, users who belong to some minority group (about which there is proportionally less information by definition) may be more likely to serve as an explorer. In other words, autonomous experimentation systems can disproportionately target users who do not resemble the majority, in some cases because they belong to a historically disadvantaged group. Although other machine learning systems can also discriminate against minority users [10], the context of experimentation raises questions that relate to justice—i.e., the process of distributing risks and benefits in a manner that is fair.

Unfortunately, the standard notion of justice does not translate cleanly to autonomous experimentation. In a traditional human-subjects experiment, upholding justice means ensuring that one group does not bear most of the risks, while another accrues most of the benefits. However, although autonomous experimentation systems may disproportionately experiment on minority users, these same users are the ones who stand to benefit the most from the experiments. As a result, a more pertinent question is whether the extent or nature of the experiments is necessary—would other experiments, with fewer risks, be equally effective at benefiting those users? This is exactly what beneficence captures. Here, the goal is to ensure that any risks to research subjects are minimized, while maximizing research benefits. Indeed, upholding beneficence sometimes requires researchers to consider using alternative research methods.

Ultimately, any rigorous discussion of fairness and autonomous experimentation will have to address which experiments are justified and which are not, as well as who is most likely to be affected by them. When humans lack *any* intuition about which actions are optimal, a system that selects explorers and actions uniformly at random will not elicit the fairness concerns described above. No user will be subjected to actions that are known to be sub-optimal. However, there are many scenarios where humans—and especially domain experts—have well-honed intuitions about the kinds of actions that are more or less likely to be optimal, even if a system does not. For example, a user may know a near-optimal route to work, even though the navigation service she relies on has not discovered it yet. One possible way to assess a system's fairness is therefore to allow users to ask whether they are part of an experiment, what information that experiment is intended to discover, and whether that information could have been discovered by other means.

# 6. REALIZING ETHICAL PRINCIPLES

In a traditional human-subjects experiment, informed consent means that potential research subjects must be provided with information about the experiment (e.g., the selection procedure for research subjects, the research methods, the potential risks and benefits) and given the opportunity to assent. Unfortunately, it is much harder to obtain informed consent from the users of an autonomous experimentation system. First, the reasons for performing an experiment depend on the internals of the system—its current representation of the environment, its set of possible actions, its strategy for selecting explorers and actions, and its assessment of the potential risks and benefits. Communicating these reasons to a user in a manner that is unambiguous and that accords with the user's own notions of risks and benefits is not easy. Second, the number and range of experiments and the speed at which they are typically performed make enacting the informed consent process extremely difficult.

These obstacles to obtaining informed consent also make it hard to develop procedures for ensuring accountability. For government-funded human-subjects research, the Common Rule requires that institutional review boards (IRBs) review all experiments to evaluate research protocols and methods and to assess potential risks and benefits. There is no clear analog for autonomous experimentation systems. One possibility is an automated "checks and balances" mechanism, perhaps as part of a system's reward signal, for monitoring and regulating a system's actions to avoid or penalize exper-

iments that an IRB would find problematic. In many scenarios, though, it is likely that such a mechanism would have to be accompanied by some form of human review. This hybrid approach would make it possible to share accountability between the system designers and external reviewers, but again raises the challenge of communicating the reasons for performing an autonomous experiment to a human. In addition, it would slow the system down and obfuscate who users should hold accountable if something goes wrong.

Regardless of the procedure for ensuring accountability, the task of assessing potential risks and benefits to users is non-trivial. In autonomous experimentation systems, risks and benefits are typically derived from behavioral data such as clicks or wearable sensors—often by another machine learning system, which may be crude or even vulnerable to approximation errors and bias. If a system's assessments of risks and benefits are inaccurate, then any decisions based on those assessments will necessarily also be inaccurate.

Finally, reviewing an experiment according to IRB standards involves the notion of minimal risk. Minimal risk means that the probability and magnitude of harm or discomfort experienced in an experiment are no greater than those ordinarily encountered in a research subject's day-to-day life. If an experiment qualifies as minimal risk, it is typically allowed to proceed with fewer precautions and a less rigorous review process. In practice, although an autonomous experimentation system may have a sophisticated model of the comparative risks caused by various actions, it is unlikely to have any model of the absolute probability or magnitude of these risks or how these values might compare to those associated with other systems or other activities in a user's life. This makes maintaining IRB standards difficult.

## 7. CONCLUSIONS

We are at the beginning of a new phase of experimentation, with autonomous experimentation systems increasingly found even in critical infrastructure, such as transportation and healthcare. However, despite the ongoing debate about existing research ethics regululations and their applicability to data science, researchers have largely ignored the unique risks posed by autonomous experimentation. As we have shown in this paper, autonomous experimentation systems challenge the ethical principles outlined in the Belmont report and subsequently codified in the Common Rule: respect for persons, beneficence, and justice. We therefore believe that more research is urgently needed in order to fully understand the social and ethical implications of autonomous experimentation systems and to develop mitigating strategies.

## 8. REFERENCES


[1] Human Experimentation: Code of Ethics of W.M.A. *British Medical Journal*, 2(5402):177–177, 07 1964.

[2] The Nuremberg Code (1947). *British Medical Journal*, 313(7070):1448, 12 1996.

[3] *Workshop on Ethics of Online Experimentation at the Eighth International Conference on Web Search and Data Mining*, 2015.

[4] *Workshop on Fairness, Accountability, and Transparency in Machine Learning at the Thirty-Second International Conference on Machine Learning*, 2015.

[5] D. Agarwal, B.-C. Chen, and P. Elango. Explore/exploit schemes for web content optimization. In *Proceedings of the Ninth International Conference on Data Mining*, 2009.

[6] S. Barocas and A. D. Selbst. Big data's disparate impact. *California Law Review*, 104, 2014.

[7] R. Dearden, N. Friedman, and D. Andre. Model-based Bayesian exploration. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, 1999.

[8] Department of Health and Human Services. Notice of Proposed Rulemaking. *Federal Register*, 80(173), September 2015.

[9] F. Diaz. Integration of news content into web results. In *Proceedings of the Second International Conference on Web Search and Data Mining*, 2009.

[10] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel. Fairness through awareness. In *Proceedings of the Third Innovations in Theoretical Computer Science Conference*, 2012.

[11] I. Hochberg, G. Feraru, M. Kozdoba, S. Manor, M. Tennenholtz, and E. Yom-Tov. Encouraging physical activity in diabetes patients through automatic personalized feedback via reinforcement learning improves glycemic control. *Diabetes Care*, 2016.

[12] M. Kosinski, D. Stillwell, and T. Graepel. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805, 2013.

[13] P. Mannion, J. Duggan, and E. Howley. An experimental review of reinforcement learning algorithms for adaptive traffic signal control. In L. T. McCluskey, A. Kotsialos, P. J. Müller, F. Klügl, O. Rana, and R. Schumann, editors, *Autonomic Road Transport Support Systems*, pages 47–66. 2016.

[14] J. Metcalf and K. Crawford. Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society*, 2016.

[15] J. Polonetsky, O. Tene, and J. Jerome. Beyond the common rule: Ethical structures for data research in non-academic settings. *Colorado Technology Law Journal*, 13, 2015.

[16] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 01 2016.

[17] D. Silver, L. Newnham, D. Barker, S. Weller, and J. McFall. Concurrent reinforcement learning from customer interactions. In *Proceedings of the Thirtieth International Conference on Machine Learning*, 2013.

[18] O. Tene and J. Polonetsky. Beyond IRBs: Ethical guidelines for data research. *Washington and Lee Law Review Online*, 72(3):458, 2016.

[19] Y. Yue and T. Joachims. Interactively optimizing information retrieval systems as a dueling bandits problem. In *Proceedings of the Twenty-Sixth International Conference on Machine Learning*, 2009.