The Cost of Ethical AI Development for Startups

James Bessen
TPRI
Boston University
Boston, MA USA
jbessen@gmail.com

Stephen Michael Impink
Stern School of Business
New York University
New York, NY USA
impink@outlook.com

Robert Seamans
Stern School of Business
New York University
New York, NY USA
rseamans@stern.nyu.edu

ABSTRACT

Artificial Intelligence startups use training data as direct inputs in product development. These firms must balance numerous tradeoffs between ethical issues and data access without substantive guidance from regulators or existing judicial precedence. We survey these startups to determine what actions they have taken to address these ethical issues and the consequences of those actions. We find that 58% of these startups have established a set of AI principles. Startups with data-sharing relationships with high-technology firms or that have prior experience with privacy regulations are more likely to establish ethical AI principles and are more likely to take costly steps, like dropping training data or turning down business, to adhere to their ethical AI policies. Moreover, startups with ethical AI policies are more likely to invest in unconscious bias training, hire ethnic minorities and female programmers, seek expert advice, and search for more diverse training data. Potential costs associated with data-sharing relationships and the adherence to ethical policies may create tradeoffs between increased AI product competition and more ethical AI production.

CCS CONCEPTS

• Social and professional topics • Computing/technology policy • Commerce Policy • Antitrust and competition.

KEYWORDS

artificial intelligence, ethics, data, startup, competition

ACM Reference format:

James Bessen, Stephen Michael Impink, and Robert Seamans. 2022. The Cost of Ethical AI Development for Startups. In Proceedings of 2022 ACM Conference on Artificial Intelligence, Ethics, and Society (AIES'22), August 1–3, 2022, Oxford, United Kingdom. ACM, 10 pages. https://doi.org/10.1145/3514094.3534195

AIES '22, August 1–3, 2022, Oxford, United Kingdom.

ACM ISBN 978-1-4503-9247-1/22/08...$15.00.

https://doi.org/10.1145/3514094.3534195

1 Introduction

AI is important for future macroeconomic growth, and data is a key input in AI production [15,27,32]. Many future AI products will replace some aspects of human labor, augment human capabilities, and revolutionize data analysis [1,10,12,28]; yet these gains will only be realized through training algorithms supporting AI products with tons of data. The mass collection and use of data, particularly personally identifiable information (PII) or biometric data, raises concerns about firms' ethical obligations to protect individuals and reduce the likelihood of discriminatory outcomes [7,4,14,72]. Moreover, scholars highlight many fairness issues stemming from the development and use of AI and algorithms, such as unrepresentative training data, biased programmers, and an overemphasis on prediction accuracy, which may negatively impact decision-making or harm a demographic subgroup [6,20,31,55].

Over the past several years, AI's ethical use has been a popular topic among academics and journalists alike, with much of the research focusing on issues related to privacy and outcome fairness [65]. Despite the increased usage of AI in organizations, awareness of the potential pitfalls of using AI and certain training data lags. A recent study of Google search trends shows a significant increase in searches related to AI-based recruitment and algorithmic bias since 2016; however, there has been no increase in similar searches focused on ethical recruitment and hiring [53]. Though there have been numerous calls to action, urging governments to regulate AI and provide substantive guidelines for ethical AI development and usage [70], no widescale policies have emerged. The European Union (EU) General Data Protection Regulation (GDPR) and California Consumer Protection Act (CCPA) are the most wide-scale data regulation; however, they only cover one aspect of ethics: privacy. Even though the EU is currently drafting broader AI regulations, it may be years before these regulations are finalized and take effect. Moreover, similar to privacy regulations, these regulations with be concentrated in more developed countries, primarily in Europe[1].

Given the lack of guidance from governments, guidelines focused on AI design, deployment, and usage are usually created and adhered to at the firm level, leaving ample room for ambiguity and providing limited guidance for startups. Ethical AI principles may be substantially different and applied in various ways across firms and industries. Still, it often includes guidance on understanding potential biases in training data and algorithms

[1] 18 ethical AI policies from governments, Jobin et al. 2019

and how they manifest in algorithms and impact the fairness of their AI product's recommended outcomes. These policies signify the importance of ethics to these firms' managers, investors, and stakeholders. Creating more ethical AI products, adopting these policies, and ultimately following these policies have tangible costs. For instance, firms may hire additional minority programmers to code their algorithms, yet the search for and scarcity of minority engineering talent may be expensive. Furthermore, searching for more representative training data or dropping training data deemed to be less fair or possibly lead to biased outcomes may be an insurmountable hurdle for a newly founded startup with limited funding.

For example, these policies may require that data containing PII or information on legally protected statuses like ethnicity, gender, or age, are not used in AI production, limiting access to needed training data resources. Limited training data may reduce algorithmic effectiveness at predicting outcomes. Moreover, the policies may guide managers to build AI products that prioritize prosocial outcomes. For instance, it is not simply that the firm wants to build an algorithm that is the best at predicting an outcome in aggregate; they want an algorithm that best predicts an outcome for all races and genders individually. Furthermore, some startups may even turn down deals that conflict with their ethical guidelines, such as refusing to work with the military or police forces, directly reducing their revenues. Larger high-technology firms have crafted ethical AI policies that align with their business models, reduce liabilities, and limit the chances of a public relations fiasco. There are many examples of these larger firms turning down business that conflicts with their ethical principles [2]. Still, it may be that turning down revenue from potentially unethical customers makes other customers more willing to collaborate with them. However, AI startups may only have a handful of customers. Turning down business from a single firm may threaten their survival

If AI startups lack adequate training data, they may partner with a technology firm to access better or more representative training data. Many high-technology firms have established norms around AI development and share their guidance on potential negative externalities from AI usage [36,64]. As a result, private firms in the US, UK, and EU published a quarter of the available AI guideline documents [40][3]. Intel[4], IBM[5], Microsoft[6], SAP[7], Sony[8], and Google[9] have shared ethical guidelines that provide insight into algorithmic transparency, bias, and fairness, describing many implications for justice, equity, and privacy. These partnerships may make ethical AI production concerns more salient, and in some cases, these startups may be held accountable to the high technologies firm's AI development norms.

Larger technology firms, especially the largest three technology firms (e.g., Amazon, Google, and Microsoft) that are also cloud services providers (CSPs), share many data-related resources with startups on their cloud platforms. They may be more comfortable partnering with AI startups that take ethics seriously. However, these technology firms invest heavily in AI product development and often compete downstream in the AI product market. The information and feedback startups share through the data-sharing partnership may reduce downstream competition [23,46], potentially making these data-sharing relationships very costly. For instance, startups may exchange non-pecuniary resources, like product information, industry insight, or feedback on the larger firm's development platform, products, and training data, in exchange for data through this relationship [56].

Data as inputs in AI production leads to two interesting points of tension in the AI development industry. First, using personal data to train AI directly or indirectly (i.e., through an omitted covariate) may negatively affect consumer welfare by leading to biased outcomes or revealing consumer willingness to pay. Data-sharing partnerships with technology firms may incentivize startups to follow their ethical norms and enable startups to avoid using less appropriate training data. However, this relationship may allow technology firms to collect and aggregate valuable information and feedback about the AI industry from many startups, reducing downstream competition. In turn, this downstream competition could also negatively impact consumer welfare, possibly even muting the gains to consumer welfare from more ethical AI startup product development. Second, the costs associated with developing more ethical AI products may reduce startup entry and resulting innovation. Entrepreneurs generally have limited resources compared with more established firms, leading them to experience more ethical dilemmas and impacting how they respond to these dilemmas [35]. Since more ethical AI production has numerous additional costs, these nascent startups may not have the resources to adhere to ethical AI guidelines or develop products demanded by a more ethically conscious market.

This paper explores the relationship between prior resources and AI startups' actions to address numerous data-related ethical issues. Our research question is whether and how prior resources, such as funding, data sharing relationships, and experience with GDPR, impact the adoption and use of ethical AI principles. We assess these issues by collecting and analyzing novel survey data from 225 AI startups. We find that more than half of responding

[2] Washington Post: Microsoft won't sell police its facial-recognition technology, following similar moves by Amazon and IBM, https://www.washingtonpost.com/technology/2020/06/11/microsoft-facial-recognition/ and Big tech companies back away from selling facial recognition to police, https://www.vox.com/recode/2020/6/10/

[3] Jobin et al. 2019 details 84 ethical guidelines produced worldwide in the last several years.

[4] Intel's AI Privacy Policy White Paper: Protecting Individuals' Privacy and Data In The Artificial Intelligence World (Intel, 2018).

[5] Everyday Ethics for Artificial Intelligence (IBM, 2018). Transparency and trust in the cognitive era. IBM https://www.ibm.com/blogs/think/2017/01/ibm-cognitive-principles/ (2017).

[6] Microsoft AI Principles. Microsoft https://www.microsoft.com/en-us/ai/our-approach-to-ai (2017).

[7] SAP's guiding principles for artificial intelligence (AI). SAP https://www.sap.com/products/leonardo/machine-learning/ai-ethics.html#guiding-principles (2018).

[8] Sony Group AI Ethics Guidelines (Sony, 2018). https://www.sony.net/SonyInfo/csr_report/humanrights /AI_Engagement_within_Sony_Group.pdf

[9] DeepMind, acquired by Google in 2014, issues guidelines in addition to Google. DeepMind ethics and society principles. DeepMind https://deepmind.com/applied/deepmind-ethics-society/principles/ (2017). Artificial intelligence at Google: our principles. Google AI https://ai.google/principles/ (2019).

startups have ethical AI principles. However, many of those have never invoked their ethical AI principles in a way that negatively impacted their business, such as firing an employee, dropping training data, or turning down a sale. Startups with data-sharing relationships with high-technology firms or prior experience with privacy regulations are more likely to establish ethical AI principles and incur adverse business outcomes, like dismissing employees, dropping training data, or turning down business, to adhere to their ethical AI policies. Moreover, startups with ethical AI policies are more likely to invest in unconscious bias training, hire ethnic minorities and female programmers, seek expert advice, and search for more diverse training data. Potential costs associated with data-sharing relationships and the adherence to ethical policies may create tradeoffs between increased AI product competition and more ethical AI production.

On the one hand, having an ethics policy may signal that a startup is more willing to adapt to the industry's broader norms. Investors may be more willing to invest in AI startups with ethical policies and leaders who prioritize developing ethical AI norms. For instance, investors could worry about PR-related issues or reduced exit opportunities for startups with ethical issues or products that facilitate discriminatory behavior. Furthermore, as potential algorithmic bias and resulting discrimination are more universally acknowledged as a risk, there could be increased liability related to product usage once judicial rulings are established. On the other hand, producing more ethical AI has costs startups may not be able to incur, such as turning down business from early customers, hiring niche programmer talent, or dropping data. Furthermore, startups may provide non-pecuniary resources (i.e., information, feedback [57]) to develop a data-sharing relationship, making the cost of these relationships, even though they are related to more ethical decisions, more costly than they realize.

Our paper makes several contributions. First, we contribute to the developing literature on the role of data for AI-producing firms [12,11,13,15,21,22,38,32]. Second, we contribute to a broader literature on digitization [20,35,70]. Third, we contribute to the nascent literature on ethics in AI development by highlighting the case of AI startups, where these ethical issues are often related to the acquisition and use of resources needed to survive [50,67,68,61,63,69].

This paper proceeds as follows. Section 2 introduces the academic literature on ethical development and usage of AI, highlighting gaps in the research on how ethics interplays with data access issues. Next, Section 3 discusses the data collected from our survey of AI startups ending in March 2021, including several data limitations, such as issues of non-response (response rate: 6%) and the cross-sectional nature of our data. Section 4 explains our research design, which relies on Heckman's selection to help overcome potential non-response based on observable characteristics and Coarsened Exact Matching (CEM) to support that our treatment group, startups with AI ethics principles, is demographically similar to those without principles. Our findings (Section 5) highlight correlations between a firm's prior resources and ethics outcomes, focusing on data-sharing collaborations with the high technology firms and prior experience with GRPR. Lastly, Section 6 discusses possible antecedents to AI startups adapting ethical AI principles and concludes.

2 Prior Research on Ethical AI

First, we review research on ethical issues stemming from personal data's mass collection, monitoring individuals, and related privacy implications. This literature highlights tradeoffs between using certain types of data and possible ethical issues that could arise. Next, we focus on more recent issues of algorithmic bias stemming from training data and programmers. In addition to the choice of training data, how programmers code the algorithm could introduce bias, impacting the fairness of outcomes. Lastly, there is nascent literature on the impact of managers within organizations using potentially biased AI outcomes in ways that accounts for these biases, highlighting the importance of ethics education for managers.

2.1 Privacy and Monitoring

Firms must choose what types of data they use to train their AI products' algorithms. Often, firms use data about unique individuals, such as information on past purchases or preferences, which, when with other demographic information, enables the identification of a specific individual. The ability to identify someone from their data, passively collected through normal business activities, leads to numerous concerns about the possible negative externalities [14]. In most cases, consumers willingly share this information to enjoy services or social media platforms they highly value [2]. In response to growing privacy concerns, governments created substantive legislation (e.g., GDPR, CCPA) covering a broad range of ethical issues, including the right to access, delete, and prevent the sale of one's personal data held by a firm and to know what personal data is being collected and stored by one's employer.

Ethical concerns also arise from using AI to monitor individuals by capturing personal information, such as motions and gestures, to identify anomalies or unwanted behavior patterns. AI products focused on monitoring have many socially positive outcomes applications, such as increasing fairness by ensuring students do not cheat on exams [9] or reducing road hazards by buzzing truck drivers that start to dose. Even for socially beneficial outcomes, monitoring technologies raise concerns about the organization's role in forcing members to be subjected to these more intrusive, albeit passive, technologies. Alternately, there are substantial risks associated with these "Big-brother" style surveillance systems. For instance, police or military could use AI in detrimental ways, such as discriminating against or targeting a particular group by searching crowds for individuals based on their faces or appearance.

The literature on the impact of data privacy regulation also gives us insight into how increased adherence to ethical norms impacts smaller, newer firms. For example, increased government regulation, though uniformly guiding firms on handling ethical issues that may harm consumers, may also reduce the amount of training data collected and used in AI development. This tradeoff between increased regulation and data availability could asymmetrically negatively impact smaller firms that need data to develop their products and grow, creating competitive barriers for some startups [13,42,44,53].

2.2 Algorithmic Bias

Issues with an algorithm's training data can introduce bias into outcomes, impacting the fairness of outcomes [8,54]. For example, an unrepresentative sample of data collected from a single race or gender could bias outcomes and is particularly relevant when training AI with biometrics data, such as images of faces, eyes, or fingerprints [59]. A typical example is using photos of human faces to train AI products focused on recognizing an individual. If you only use photos of one demography's faces, the AI product will better identify individuals similar to those in the sample. Models may need to control for sensitive demographic information instead of ignoring it; otherwise, the endogeneity of related variables may confound results, introducing unaccounted-for biases [19].

Programmers are often unaware that their product's resulting outcomes are biased. Even if they know that their results are biased, they often cannot determine that bias's exact source [5]. AI products are often referred to as a "black box," having little ability to explain causal relationships [3,26]. Data regulation has required that certain types of AI products, such as training data and algorithms used in recruitment and hiring, and results are more "explainable." However, even additional transparency may not entirely reduce bias, especially when the source of the bias and relationships among the model's inputs are causally unknown.

Even if the sample is representative, programmers could introduce bias by building algorithms in a certain way. A recent field study has shown that a particular demographic subgroup is not more biased per se; still, there is a benefit (i.e., reduced prediction errors) to having more demographic subgroup diversity [22]. Moreover, how programmers code outcomes could introduce bias into an algorithmic model [20]. For example, including all employees who left a firm in a turnover algorithm may capture individuals who left the firm for poor performance and, additionally, individuals who left because they did not fit with the firm's culture. In many cases, outcomes are endogenous, with other aspects of demography not accounted for in an AI product's underlying algorithms. Moreover, feedback loops, where outcomes from the initial algorithm are used in future iterations, can further exacerbate these initial biases [20]. Lastly, algorithmic bias can also occur when programmers are overly focused on prediction accuracy instead of weighing the benefits of accuracy with other prosocial outcomes [19].

2.3 Impact of Managers

Most algorithms have one goal, prediction accuracy, and in most cases, the algorithm type and the associated training data are chosen as the first step toward reaching this goal. In cases where programmers prioritize algorithmic accuracy, it may be best for humans to make the final decision by weighing the potential source of biases and the organization's prosocial goals. Even if outcomes from AI products are biased, managers can use these outcomes combined with other information to produce a final outcome that is ultimately less biased. Just because the AI is biased does not mean that the resulting decisions are biased; managers can use AI as a tool to complement their constrained decision-making processes [45].

Often biases from algorithms or less representative samples emerge in organizational decision-making processes. Managers benefit from education about recognizing sources of algorithmic bias and other forms of unconscious bias that could impact their decision-making [51,62,66]. Even though managers may benefit from a more metered approach, using AI to complement other information in their decision-making process, they often avoid using AI when made aware of potential biases [21]. Though recent frameworks have been introduced to help identify ethical issues [63] and raise awareness of the tradeoffs between ethics, bias, and the use of big data in machine learning processes [68], theoretical development needs to reflect the speed with which organizations are adopting AI. Some managers are unaware of the possibility that their AI products are biased, producing less fair outcomes; however, given the amount of evidence provided in the current stream of research, managers should be made aware of ethical issues and held accountable for ameliorating these issues [50]. Moreover, augmentation of human-AI interaction and research to better understand this interaction is needed as AI becomes more prevalent in organizations [67].

Lastly, prior research argues specific demographic subgroups may make better decisions than others. For instance, female students [49,60] and female managers are more likely to make ethical business decisions [16,17,24,25,33,29]. This issue is exacerbated as high-paying STEM jobs, designing algorithms and AI, are often advertised more to men than women [48]. High technology firms have made strides in more inclusive hiring, and business school faculty members, aware of the potential issues of working with big data, have created awareness around the need for ethical AI development education [3,20]. Furthermore, those who take courses in ethics are more likely to make more ethical decisions. Over the past decade, the curriculum at many business schools has focused more on business ethics [18,30,52]. However, men are more likely to go to business school, and some argue that men benefit more than women from business education [71]. Additionally, many top-ranked business schools (e.g., Harvard, MIT, NYU) have built out courses that focus on managing AI development and use within organizations.

3 Survey Data and Measures

We use data from a survey of AI startups, including questions regarding the development of ethical AI principles and the impact of these principles on business decisions. We list these questions in Appendix A. We pretested the survey with several academics and practitioners associated with startups and then administered the survey from January 2021 to March 2021 through Qualtrics. We received 225 responses from AI startups in our sample; these firms confirmed they develop AI products in the first survey question.

Respondents to our survey come from several sampling frames, including Crunchbase, Pitchbook, Creative Destruction Labs, and an incubator at Technische Universität München

(TUM). From Crunchbase and Pitchbook, we identify firms associated with the keyword "artificial intelligence" that are in operation yet have not experienced an initial public offering (IPO). Additionally, we received a contact list of AI startups from the Creative Destruction Lab, a startup incubator based in Toronto, and another contact list from Philipp Hartmann and Joachim Henkel at TUM [38]. To develop a more homogenous sample of firms for our analysis, we exclude larger (more than 500 employees) or older (more than ten years old) firms. We sent a digital survey via email to the 3,790 firms in our sample address to the founders, chief technology officer, or executives who know their firm's business model and technologies.

Firms responding to our survey are about four years old and employ, on average, 23 employees, with more than 40% of firms having ten or fewer employees. Even though the survey was administered worldwide, most of our responses are from more developed countries, with almost 80% of responses from the United States, Canada, and Europe. Since our response rate is 6%, we analyze whether certain firms are more or less likely to respond to the survey. In the first stage of our analysis, we use Heckman's selection procedure, described in the section below, to help correct any potential non-response bias based on demographics observables. However, correlated unobservables, variations not captured by our data, may bias our estimates and results.

There is variation by region, with Asia-headquartered firms most likely to have policies and middle eastern and African headquartered firms least likely to have policies. There are, however, only slight differences between policy adoption of firms in North America and Europe, where more than three-quarters of responding firms are located. Similar to ethical AI policy adoption, we find that business outcomes resulting from startups adhering to their ethical AI principles vary by firm size and HQ location. Larger firms are much more likely to dismiss an employee and drop data, and firms in Europe are more likely to turn down business and drop data. Larger startups with more than ten employees are more likely to hire a female or minority programmer. Additionally, firms in the US are more likely to acquire additional, more diverse training data and conduct unconscious bias training than firms in Europe and other parts of the world.

For the selection probit, we include indicator variables for firms that are more than six years old (0.29 SD 0.45), have a minority founder (0.05 SD 0.21), or are located in the United States (0.37 SD 0.34). For the main results, we also include indicator variables for smaller firms with ten or fewer employees (0.41 SD 0.49) and firms located in cities with a higher concentration of venture capital firms (San Francisco, London, New York, Boston, and Hong Kong, 0.24 SD 0.43).

Table 1.A – Firm Summary

Measure	Mean	SD	Min	Max
Age	4.08	1.89	1	10
Young (< 3 years)	0.49	0.50	0	1
Older (>6 years)	0.29	0.45	0	1
Employee (Avg. Count)	23	26	1	175
Small (<11 employees)	0.41	0.49	0	1
Large (>100 employees)	0.24	0.42	0	1
VC Location	0.24	0.43	0	1
Headquarters Location				
United States	34%			
California	3%			
New York	6%			
Massachusetts	2%			
United Kingdom	11%			
Canada	4%			
European Union	27%			
Other Developed	17%			
Other Emerging	7%			

Table 1.A summarizes demographics for the 159 startups included in our analyses.

We pair the survey data with firm-level data from both Crunchbase and Pitchbook to build our measures. Of the 225 survey respondents, 159 responding firms have available demographic and funding information in Crunchbase or Pitchbook[10]. From the survey, we create dummy variables for a) if the firm had established an ethical AI policy (0.58 SD 0.50); b) if the firm acted upon those principles by turning down business (0.21 SD 0.41), dismissing an employee (0.06 SD 0.24), or dropping training data (0.20 SD 0.40); and c) if the firm had taken the following ethics-related actions, including considering diversity when selecting training data (0.45 SD 0.45), hiring a minority or female programmer (0.67 SD 0.48), offering bias training (0.26 SD 0.44), or seeking expert advice (0.35 SD 0.47).

Next, we create indicator variables for if a startup has a data-sharing partnership with a technology firm (0.45 SD 0.50), prior experience with GDPR (0.62 SD 0.48), or received funding a) before the survey was administered in 2021 (0.64 SD 0.47), b) before for the survey was administered from a VC investor (0.28 SD 0.45). Lastly, we create indicator variables for if a startup was founded by a female (0.10 SD 0.30) or Master of Business Administration degree (MBA) graduate (0.11 SD 0.32).

[10] Some startups in our sample are listed in Crunchbase and/or Pitchbook with nothing more than a description. These startups are likely small, nascent ventures that will have additional data paired in the future as they grow. We drop any firms that are public (IPO), have acquired another firm, or are more than 10 years old.

Table 1.B – Measure Summary

Measure	Mean	SD	Min	Max
Dependent Variables				
Ethical AI Principles?	0.58	0.50	0	1
Due to these principles, has your firm:				
Dismissed Employee	0.06	0.24	0	1
Dropped Data	0.20	0.40	0	1
Turned Down Business	0.21	0.41	0	1
Independent Variables & Controls				
Log (Funding through 2020)	8.90	6.80	0	17.7
Funding through 2020	0.64	0.47	0	1
VC Backed	0.28	0.45	0	1
Has your firm done the following:				
Considered Diversity	0.45	0.45	0	1
Hired Minority	0.67	0.48	0	1
Unconscious Bias Training	0.26	0.44	0	1
Sought Expert Advice	0.35	0.47	0	1
Tech Firm Collab.	0.45	0.50	0	1
GDPR (Prior Exp.)	0.62	0.48	0	1
Diversity (Minority) Founder	0.05	0.21	0	1
Female Founder/CEO	0.10	0.30	0	1
Founder with MBA	0.11	0.32	0	1

Table 1.B summarizes measures for the 159 startups included in our analyses.

4 Methods

We use regression models to explore prior resources and ethical AI policies. We use Heckman's selection approach [39,40] and Coarsened Exact Matching (CEM; [41]) to help address selection and endogeneity issues. First, given our lower survey response rate and reliance on cross-sectional data, we analyze if our survey respondents are similar to the broader population of startups in Crunchbase and Pitchbook. From initial t-tests, we find that responses from the United States are overrepresented, but California, where many startups are based, is underrepresented. Moreover, younger firms are less likely to respond to our survey. To confirm this, we use a probit regression model to estimate the likelihood of response (Table 2, selection specification below)[11].

Based on this, we use Heckman's two-step procedure to account for selection issues from possible respondent missingness to support the argument that our sample of respondents does not bias our main results. We include dummy variables for young startups and startups with HQ locations in the United States or California in the first step, below, to obtain estimates of γ.

$$response_i = w_i\gamma + \mu \quad (1)$$

where, $response_i$ takes the value of 1 if a firm in the population responds to the survey, otherwise 0; w_i refers to a vector of firm demographic dummy variables (e.g., US, Older (>6 years old), and minority founder) plausibly correlated with sample response.

Now that we have obtained the estimates of γ from the selection equation, we compute the inverse Mills ratios of each observation.

$$\lambda = \frac{\phi(w_i\gamma)}{\Phi(w_i\gamma)} \quad (2)$$

where, $\phi(w_i\gamma)$ refers to the probability density function, and $\Phi(w_i\gamma)$ refers to the complementary cumulative distribution function.

Next, we use CEM weighting and matching, based on region, startup age, employment size, and VC funding before the survey was administered, to ensure that the firms with an ethical AI policy are observationally similar to those without an ethical AI policy. The match reduces the difference in standardized means across these observable demographic variables between the respondents who have and do not have ethical AI principles. We provide a table comparing these demographic variables of firms before and after the match and weighting in Appendix D.

We use the following regression specification for our analysis:

$$y_i = \beta_0 + \beta_1 prior\_resource_i + \beta_2\lambda_i + \rho_i + \mu \quad (3)$$

where, y_i refers to an indicator variable for if the firm has a set of ethical AI principles (yes, 1; otherwise, 0) in the main analysis; $prior\_resouce_i$ refers to an indicator variable if a firm collaborates with a technology firm to access data, has prior experience with GDPR, or prior funding; ρ are controls for small employment size (<11 employees) and HQ location in a top VC city. When controlling for selection (i.e., representativeness of our sample compared with the population of AI startups that we sourced and contacted) with IMR, the inverse of the Mills ratio, we include the term λ, in the specification.

We control for firm size and if a firm is located in a city with a high concentration of venture capital firms, which may be related to outcomes. For example, smaller firms have fewer employees, so they may be less likely to fire employees. Additionally, proximity to the largest VC may capture the impact of these interactions between large institutional investors and similar firms in the same location; these learnings may be informally shared, impacting our results. Furthermore, cities with more VC firms also have more technology firms. We pay close attention that these control variables are not highly correlated with firm age, founder demographics, or HQ location in the US, as we use those control variables in the first stage of the Heckman selection model. We provide these correlations in Appendix B, Tables B.1-B3.

5 Findings

5.1 Selection

First, we examine aspects of selection through a probit specification (equation 1) to understand which attributes of startups in the sample are related to higher or lower response rates. We find that older startups (+0.22 SD 0.08) and firms with a minority founder (+0.45 SD 0.18) are more likely to respond to the survey. On the other hand, firms based in the US (-0.28 SD 0.08), where the bulk of startups in our sample are located, are less likely to respond. From these estimates, we calculate IMR (equation 2),

[11] We look at response to the survey against the total population of AI startups matching our criteria, currently available in Crunchbase and Pitchbook (4,967 startups)

which can be added as a control to correct estimates for selection issues in the main specification.

Table 2 - Response Probit

	(1)
DV, Dummy:	Response
Old (>6 Years)	0.223***
	(0.079)
USA	-0.282***
	(0.075)
Minority Founder	0.447**
	(0.180)
Firms	3,790

Table 2: * p<0.1, ** p<0.05, *** p<0.01. 225 startups respond to the survey. This probit model is the first stage of the Heckman selection procedure.

5.2 Main Results

Next, we explore the relationship between prior resources and policy adoption through a series of cross-sectional OLS regressions. We find a significant positive relationship between having a technology firm data collaboration and adopting AI principles (Table 3, model (1): +0.25 SD 0.08). In model (2) we use a CEM weighted regression, and in model (3) we include additional controls for size (ten or fewer employees) and location near a city with a greater presence of high-technology and venture capital firms (SF, London, New York, Boston, Hong Kong), finding similar results. Lastly, in model (4) we include an additional control for if a firm received funding before 2021, a proxy for prior fundraising performance, to provide additional support for our argument that prior data collaborations with technology firms are not necessarily related to observable measures of prior performance (model (4): 0.29 SD 0.08). While we make no causal assumption, these technology firms may be more likely to share data with AI startups with ethical AI principles than those without, limiting their exposure to and liabilities from data misuse. At the same time, it could be that these technology firms provide information through their corporate accelerators, marketing programs, or cloud-service relationships about the importance of developing AI more ethically.

Furthermore, using the same build-up of regression models, we show a significant positive relationship between prior experience with GDPR and the increased adoption of AI principles (Table 3, model (8): +0.19 SD 0.08). Data privacy is a critical component of ethical AI development, and many aspects of GDPR reduce ethical issues and make those concerns salient to startups. Moreover, prior experience with GDPR may incentivize firms to think more proactively about other aspects of potential regulations, such as specific AI development-related regulations. These findings support a significant positive relationship between prior resources, such as data relationships and prior regulatory experience, and future AI ethics principles adoption. Lastly, we show that prior funding is not related to adopting ethical AI principles (Appendix C, Table C.1).

Table 3 - AI Ethics Principles (Prior Resources)

DV [0,1]:	Does your firm have Ethical AI Principles?			
	(1)	(2)	(3)	(4)
Tech Firm Data Coll.	0.250***	0.294***	0.291***	0.286***
	(0.076)	(0.079)	(0.079)	(0.080)
Employees (<11)			0.063	0.096
			(0.082)	(0.083)
VC Loc.			0.070	0.066
			(0.089)	(0.088)
Funding bf. 2021				-0.102
				(0.081)
Firms	159	159	159	159
Adj R2	0.057	0.071	0.068	0.071
CEM	No	Yes	Yes	Yes

	(5)	(6)	(7)	(8)
GDPR	0.226***	0.199**	0.198**	0.189**
	(0.081)	(0.084)	(0.084)	(0.084)
Employees (<11)			0.022	0.054
			(0.082)	(0.084)
VC Loc.			0.086	0.082
			(0.091)	(0.091)
Funding bf. 2021				-0.099
				(0.086)
Firms	159	159	159	159
Adj R2	0.043	0.024	0.023	0.025
CEM	No	Yes	Yes	Yes

Table 3: * p<0.1, ** p<0.05, *** p<0.01. All measures are indicator variables. Coefficients are estimated using OLS regression and include robust standard errors, in parentheses below the coefficient. Models (2)- (4) and (6)- (8) include weighting (CEM), based on firm age, employment size, and region.

We explore attributes of founders that could be potential mechanisms moderating the relationship between having a data-sharing relationship with a technology firm. Based on the specification we use for the main results, we include an interaction for two scenarios: first, having a female founder and, second, having a founder with an MBA, reported in Table 4. In Table 4, model (1) we show having a female founder in addition to having a data-sharing relationship with a large technology firm is related to a significant increase in adopting ethical AI principles (model (1): +0.53 SD 0.08). Moreover, in model (2) we show that having a founder with an MBA is related to a significant increase in adopting ethical AI principles (model (2): +0.33 SD 0.16).

Table 4 - Tech Collaboration (Interaction)

	(1)	(2)
DV, Dummy:	AI Principles?	
Tech Firm Data Collab.	0.285***	0.269***
	(0.084)	(0.085)
Female	0.054	
	(0.161)	
Tech x Female	0.531***	
	(0.078)	
Founder MBA		-0.143
		(0.153)
Tech x MBA		0.329**
		(0.160)
Employees (<11)	0.047	0.061
	(0.087)	(0.083)
VC Location	0.058	0.066
	(0.091)	(0.090)
Firms	159	159
Adj R2	0.0614	0.0614
CEM Weighed	Yes	Yes

Table 4: * p<0.1, ** p<0.05, *** p<0.01. All variables included are indicator variables. Coefficients are estimated using OLS regression and include robust standard errors, in parentheses below the coefficient. All models include weighting (CEM) based on firm age, employment size, and region.

5.3 AI Policy Adherence

A set of ethical AI principles in and of itself is not important unless firms adhere to those principles. From the survey, we asked startups with AI policies to provide additional information on how adherence to these policies impacts their business outcomes to determine if ethics policies are followed instead of being signals to investors. More than half of the 90 firms with ethical AI principles experienced at least one negative business outcome because they adhered to their ethical AI principles. Though this is not an exhaustive list of outcomes, these fields provide insight into how being more ethical impacts AI startups' operations, costs, and revenues in data-centric production. Furthermore, though this question asks respondents if they have incurred a negative business outcome due to their AI principles, there is a possibility that these outcomes could have occurred anyways or were misattributed to adopting AI principles.

For the treatment group, firms with an AI ethics policy, we find that startups with a data-sharing collaboration with a technology firm are substantially more likely to dismiss an employee (Table 5, model (1): +0.14 SD 0.08) and turndown business due to their AI principles (model (5): +0.25 SD 0.11). Moreover, startups with prior experience with GDPR are more likely to drop training data when adhering to their AI principles (model (4): +0.20 SD 0.11)[12].

[12] There is no significant difference between how GDPR effects startups in the US versus Europe. Even though it is a European regulation, it is widely followed in most developed countries.

Table 5 – Ethical Outcome (Treatment)

	(1)	(2)
DV, Dummy:	Dismissed Employee	
Tech Firm Data Collaboration	0.136*	
	(0.077)	
GDPR		0.051
		(0.066)
Employees (<11)	0.082	0.045
	(0.081)	(0.068)
VC Location	0.077	0.077
	(0.084)	(0.086)
Adj R2	0.0222	-0.0148
	(3)	**(4)**
	Dropped Training Data	
Tech Firm Data Collaboration	0.077	
	(0.108)	
GDPR		0.204*
		(0.107)
Employees (<11)	-0.119	-0.108
	(0.107)	(0.104)
VC Location	0.056	0.052
	(0.115)	(0.111)
Adj R2	-0.00192	0.0279
	(5)	**(6)**
	Turned Down Business	
Tech Firm Data Collaboration	0.246**	
	(0.109)	
GDPR		-0.047
		(0.114)
Employees (<11)	0.070	-0.024
	(0.110)	(0.104)
VC Location	0.277**	0.280**
	(0.115)	(0.118)
Adj R2	0.0908	0.0349

Table 5: * p<0.1, ** p<0.05, *** p<0.01. All variables included are indicator variables. Coefficients are estimated using OLS regression and include robust standard errors, in parentheses below the coefficient. All models include 90 startups, all with Ethical AI policies, and weigh regression based on firm age, employment size, and region (CEM).

5.4 Ethics Related Actions

Many firms have taken ethics-related actions not necessarily connected to adopting ethical AI principles, such as hiring minorities, using more diverse training data, seeking expert advice, and conducting bias training. We examine the relationship between having an ethical AI policy and these ethics-related actions. Using CEM weighted regressions and controlling for size and location, we find a positive relationship between having an ethical AI policy and all ethics-related actions. In Table 6, we report that having an ethical AI policy is related to sourcing more diverse training data (Table 6, model (1): +0.19 SD 0.08), seeking expert data ethics advice (model (3): +0.30 SD 0.07), hiring minority and female programmers (model (5): +0.16 SD 0.08), and hosting unconscious bias training for employees (model (7): +0.19 SD 0.07).

Table 6 - Ethical Firm Actions

	(1)	(2)	(3)	(4)
DV, Dummy:	Training Data Diversity		Expert Advice	
Ethical AI Policy	0.187**	0.191**	0.300***	0.299***
	(0.080)	(0.081)	(0.070)	(0.070)
IMR		0.302		-0.107
		(0.290)		(0.272)
Employees (<11)	0.000	-0.015	0.016	0.022
	(0.080)	(0.082)	(0.072)	(0.073)
VC Location	0.115	0.110	0.200**	0.201**
	(0.093)	(0.093)	(0.091)	(0.092)
Firms	159	159	159	159
Adj R2	0.029	0.029	0.123	0.118
CEM Weighed	Yes	Yes	Yes	Yes
	(5)	(6)	(7)	(8)
	Minority Hire		Bias Training	
Ethical AI Policy	0.155**	0.150*	0.192***	0.196***
	(0.077)	(0.077)	(0.065)	(0.066)
IMR		-0.356		0.319
		(0.278)		(0.265)
Employees (<11)	-0.173**	-0.154**	0.021	0.005
	(0.077)	(0.078)	(0.068)	(0.069)
VC Location	0.073	0.079	0.152*	0.147*
	(0.080)	(0.082)	(0.085)	(0.088)
Firms	159	159	159	159
Adj R2	0.047	0.052	0.057	0.061
CEM Weighed	Yes	Yes	Yes	Yes

Table 6: * p<0.1, ** p<0.05, *** p<0.01. All variables included are indicator variables. Coefficients are estimated using OLS regression and include robust standard errors, in parentheses below the coefficient. All models include weight regressions based on firm age, employment size, and region (CEM). Additionally, models (2), (4), (6), and (8) include IMR from Heckman's selection procedure

5.5 Robustness

To further support these results, we run these models using probit and report results in Appendix C, Table C.2. In this additional analysis, we find similar results to our previous regressions in Table 3, prior data-sharing relationships with high technology firms and being negatively impacted by GDPR are positively related to having an ethical AI policy.

Next, we re-run the same specifications as in Table 3 with an additional selection control, IMR, from Heckman's selection equation, reported in Table 2. These results are similar to the main analysis (Appendix C, Table C.3). We omit IMR in the main regressions because the coefficient on the term is not significant. We interpret this as such; observable aspects of selection correlated with response do not significantly relate to adopting ethical AI principles. Even though there is some evidence that older firms and firms with a minority founder are more likely to respond and that firms in the USA are less likely to respond, these aspects of selection do not materially impact the ethical AI policy adoption and adherence. We show that our results remain similar regardless of the inclusion of IMR, and even though IMR is not significant, it may still effectively adjust estimates. We do not run the lasso model with the added selection control, IMR, as we need to ensure no substantial correlation between variables used in the first stage of Heckman's selection procedure and controls in the second-stage regressions.

Lastly, in Table C.4 we use a linear regression model with two types of lasso procedures (double selection and partialing out) to select control variables. Even when the machine learning model chooses control variables, results remain significant and similar to the main results.

Our results rely on cross-sectional survey data, which has its limitations. However, we attempt to address these issues by detailing plausible alternative explanations [65] and Coarsened Exact Matching. Though we control for startup size and headquarters location in a city with a large concentration of venture capital firms, we cannot entirely rule out that firms with ethical AI policies are not more likely to partner with high technology firms. Moreover, other unobservable aspects of these relationships may be correlated with policy adoption. For instance, having a supplier relationship with a CSP, a specific type of technology firm, may impact outcomes. We use prior VC funding in our CEM equation and control for prior funding in the main regressions, yet we rely only on observable data.

6 Conclusion

This study provides insight into how AI startups, newly founded firms on the front lines of AI product development, address ethical issues. Combined with our prior surveys, this data is a step towards understanding entrepreneurship around AI and the issues impacting startups relying on big datasets in production. Given the value of AI to the economy more broadly, these firms' innovations are anticipated to be important for labor productivity and future macroeconomic growth. However, similar to the impact of data privacy, there are tradeoffs between many ethical issues and the ability to access and use certain data.

In all scenarios explored, the relationship with a technology firm is related to more ethical startup behavior. These technology firms have created norms, often cited and followed by startups, since there are no extant government policies other than privacy regulations addressing many of these ethical issues. Without regulation, these startups may be more reliant on these technology firms. Many of these technology firms, such as CSPs, have invested heavily in AI and work with many startups. This increased reliance could negatively impact downstream competition in the AI product market. Recent entrepreneurship research has clearly articulated the risks of information sharing and the close relationship between certain startups and the largest technology firms to downstream competition [23,46]. Moreover, information and feedback, non-pecuniary resources provided to the larger firm, though not particularly valuable to the startup, in aggregate, may be highly valuable to the recipient. Without regulation, these prior relationships and regulatory experience may enable startups to navigate complex ethical issues and make more ethical choices, increasing their operations costs by dropping current training data to search for more diverse training data or reducing revenues by refusing to sell to less ethical customers.

The conversation around the cost of ethical AI development is becoming even more important as those costs will likely further increase as startups use more sophisticated algorithms that require even more data. For example, when more startups use neural networks, they may be less willing to drop data as larger data sources are needed to train a functional product. This training data will be even more valuable to startups as, without it, their product will not function properly, raising the cost of deleting data to comply with internal ethics policies.

In addition to reducing the need for startups to rely on technology firms, AI regulation could incentivize firms to develop more ethically. From our analysis, it is apparent that many AI startups are aware of possible ethical issues and have created ethical AI policies yet have not invoked those policies in more costly ways (e.g., dropping data, turning down business, dismissing employees). It is plausible that these startups have not been in situations that conflicted with their ethical AI policies. However, it is just as plausible, if not more so, that startups do not fully adhere to their internal policies as these costs are too much for their nascent startup to bear.

Though we highlight the tradeoff being regulation and data availability, future research must ask whether downstream competition and product innovation are more negatively impacted by increased regulation or by startups sharing valuable information with high technology firms. Both regulation and the en mass sharing of information and feedback with technology firms potentially raise the cost of AI entrepreneurship, possibly reducing entry. Understanding these tradeoffs is critical as the AI development industry grows in size and importance.

REFERENCES

[1] Acemoglu, D., and Restrepo P. Artificial Intelligence, Automation and Work. No. w24196. National Bureau of Economic Research. 2018.

[2] Acquisti, A., Taylor, C., and Wagman, L. The Economics of Privacy. Journal of Economic Literature, 54(2), 442-92. 2016.

[3] Athey, S. Beyond Prediction: Using big data for policy problems. Science 355, no. 6324, 483-485. 2017.

[4] Barocas, S., & Levy, K. (2020). Privacy dependencies. Wash. L. Rev., 95, 555.

[5] Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. Calif. L. Rev., 104, 671.

[6] Barocas, S., and Boyd, D. Engaging the ethics of data science in practice. Communications of the ACM, 60(11), 23-25. 2017.

[7] Barocas, S., and Nissenbaum, H. Big data's end run around anonymity and consent. Privacy, big data, and the public good: Frameworks for engagement, 1, 44-75. 2014

[8] Barocas, S., Hardt, M., and Narayanan, A. Fairness and Machine Learning. URL: http://www. fairmlbook.org. 2018.

[9] Bellamy, R. K., Dey, K., Hind, M., Hoffman, S. C., Houde, S., Kannan, K., ... and Zhang, Y. AI Fairness 360: An extensible toolkit for detecting, understanding, and mitigating unwanted algorithmic bias. arXiv preprint arXiv:1810.01943. 2018.

[10] Bessen, J. E. How Computer Automation affects Occupations: Technology, Jobs, and Skills. Boston Univ. School of Law, Law and Economics Research Paper, 15-49. 2016.

[11] Bessen, J. E., Impink, S. M., Reichensperger, L., and Seamans, R. GDPR and the Importance of Data to AI Startups. Available at SSRN 3576714. 2020.

[12] Bessen, J. E., Impink, S. M., Reichensperger, L., and Seamans, R. The Business of AI Startups. Boston Univ. School of Law, Law and Economics Research Paper, 18-28. 2016.

[13] Bessen, J., Impink, S. M., Reichensperger, L., & Seamans, R. (2022). The role of data for AI startup growth. Research Policy, 51(5), 104513.

[14] boyd, D., and Crawford, K. Critical Questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication and Society 15, no. 5: 662-679. 2012.

[15] Brynjolfsson, E., Rock, D., and Syverson, C. Artificial Intelligence and the Modern Productivity Paradox: A clash of expectations and statistics (No. w24001). National Bureau of Economic Research. 2017.

[16] Cohen, J. R., L. W. Pant and D. J. Sharp: 2001, 'An Examination of Differences in Ethical-Decision Making Between Canadian Business Students and Accounting Professionals', Journal of Business Ethics 30(4), 319–336.

[17] Cole, B. C. and D. L. Smith: 1996, 'Perceptions of Business Ethics: Students vs. Business People', Journal of Business Ethics 15(8), 889–896.

[18] Cornelius, Nelarine, James Wallace, and Rana Tassabehji. "An analysis of corporate social responsibility, corporate identity and ethics teaching in business schools." Journal of Business Ethics 76, no. 1 (2007): 117-135.

[19] Cowgill, B., and Stevenson, M. Algorithmic Social Engineering. AEA Papers and Proceedings, vol. 110, pp. 96-100. 2020.

[20] Cowgill, B., and Tucker, C. E. Economics, Fairness and Algorithmic Bias. Journal of Economic Perspectives. 2019.

[21] Cowgill, B., Dell'Acqua, F., and Matz, S. The Managerial Effects of Algorithmic Fairness Activism. AEA Papers and Proceedings, vol. 110, pp. 85-90. 2020a.

[22] Cowgill, B., Dell'Acqua, F., Deng, S., Hsu, D., Verma, N., and Chaintreau, A. Biased Programmers? or biased data? A field experiment in operationalizing ai ethics. In Proceedings of the 21st ACM Conference on Economics and Computation, pp. 679-681. 2020b.

[23] Cunningham, C., Ederer, F., and Ma, S. (2021). Killer acquisitions. Journal of Political Economy, 129(3), 649-702.

[24] Dalton, Derek, and Marc Ortegren. "Gender differences in ethics research: The importance of controlling for the social desirability response bias." Journal of Business Ethics 103, no. 1 (2011): 73-93.

[25] Dawson, L. M.: 1997, 'Ethical Differences Between Men and Women in the Sales Profession', Journal of Business Ethics 16(11), 1143–1152

[26] Donnelly, R., Ruiz, F. R., Blei, D., and Athey, S. Counterfactual Inference for Consumer Choice Across Many Product Categories. arXiv preprint arXiv:1906.02635. 2019.

[27] Farboodi, M., and Veldkamp, L. A Growth Model of the Data Economy. Working Paper, Columbia Business School, New York, June. 2019.

[28] Felten, E. W., M. Raj, and R. C. Seamans. The Variable Impact of Artificial Intelligence on Labor: The role of complementary skills and technologies. Available at SSRN. 2019.

[29] Fleischman, G. and S. Valentine: 2003, 'Professionals' Tax Liability and Ethical Evaluations in an Equitable Relief Innocent Spouse Case', Journal of Business Ethics 42(1), 27–44.

[30] Floyd, Larry A., Feng Xu, Ryan Atkins, and Cam Caldwell. "Ethical outcomes and business ethics: Toward improving business ethics education." Journal of Business Ethics 117, no. 4 (2013): 753-776.

[31] Friedman, B., and Nissenbaum, H. Bias In Computer Systems. ACM Transactions on Information Systems (TOIS) 14, no. 3, 30-347. 1996.

[32] Furman, J., & Seamans, R. (2019). AI and the Economy. Innovation Policy and the Economy, 19(1), 161-191.

[33] Glover, S. H., M. A. Bumpus, J. E. Logan and J. R. Ciesla: 1997, 'Re-Examining the Influence of Individual Values on Ethical Decision Making', Journal of Business Ethics 16(12), 1319–1329

[34] Goldberg, S., Johnson, G., and Shriver, S. Regulating Privacy Online: The Early Impact of the GDPR on European Web Traffic and E-Commerce Outcomes. Available at SSRN 3421731. 2019.

[35] Goldfarb, A., & Tucker, C. (2019). Digital economics. Journal of Economic Literature, 57(1), 3-43.

[36] Guo, A., Kamar, E., Vaughan, J.W., Wallach, H., and Morris, M. R. "Toward fairness in AI for people with disabilities: A research roadmap." arXiv preprint arXiv:1907.02227. 2019.

[37] Harris, J. D., Sapienza, H. J., and Bowie, N. E. (2009). Ethics and Entrepreneurship. Journal of Business Venturing, 24(5), 407-418.

[38] Hartmann, P., & Henkel, J. (2020). The rise of corporate science in AI: Data as a strategic resource. Academy of Management Discoveries, 6(3), 359-381.

[39] Heckman, J. J. Sample selection bias as a specification error. Econometrica: Journal of the Econometric Society, 153-161. 1979.

[40] Heckman, J. J. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement, v. 5, n. 4 (pp. 475-492). 1976.

[41] Iacus, S. M., King, G., and Porro, G. Causal inference without balance checking: Coarsened exact matching. Political Analysis, 1-24. 2012.

[42] Jia, J., Jin, G. Z., and Wagman, L. The Short-run Effects of GDPR on Technology Venture Investment (No. w25248). National Bureau of Economic Research. 2018.
[43] Jobin, A., Ienca, M., and Vayena, E. The Global Landscape of AI Ethics Guidelines. Nature Machine Intelligence 1, no. 9: 389-399. 2019.
[44] Johnson, G. A., Shriver, S. K., and Du, S. Consumer privacy choice in online advertising: Who opts out and at what cost to industry?. Marketing Science. 2020.
[45] Kahneman, D., and Tversky, A. "Choices, values, and frames." In Handbook of the fundamentals of financial decision making: Part I, pp. 269-278. 2013.
[46] Kamepalli, S. K., Rajan, R., and Zingales, L. (2020). Kill zone (No. w27146). National Bureau of Economic Research.
[47] Kleinberg, J., Ludwig, J., Mullainathan, S. and Rambachan, A. Algorithmic fairness. In AEA Papers and Proceedings (Vol. 108, pp. 22-27). 2018.
[48] Lambrecht, A., and Tucker, C. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. Management Science, 65(7), 2966-2981.
[49] Malinowski, C. and K. A. Berger: 1996, 'Undergraduate Student Attitudes About Hypothetical Marketing Dilemmas', Journal of Business Ethics 15(5), 525–535
[50] Martin, K. E. Designing ethical algorithms. MIS Quarterly Executive June. 2019.
[51] Martin, K., Shilton, K., and Smith, J. Business and the ethical implications of technology: Introduction to the symposium. 2019.
[52] May, Douglas R., Matthew T. Luth, and Catherine E. Schwoerer. "The influence of business ethics education on moral efficacy, moral meaningfulness, and moral courage: A quasi-experimental study." Journal of Business Ethics 124, no. 1 (2014): 67-80.
[53] McSweeny, T., O'Dea, B. Data, Innovation, and Potential Competition in Digital Markets. CPI Antitrust Chronicle, February. 2018.
[54] Mitchell, S., Potash, E., Barocas, S., D'Amour, A., and Lum, K. Prediction-based decisions and fairness: A catalogue of choices, assumptions, and definitions. arXiv preprint arXiv:1811.07867. 2018
[55] Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., and Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 1-16.
[56] Mujtaba, D. F., and Mahapatra, N. R. Ethical considerations in AI-Based recruitment. IEEE International Symposium on Technology and Society (ISTAS), pp. 1-7. IEEE, 2019.
[57] Nagle, F., Seamans, R., & Tadelis, S. (2020). Transaction cost economics in the digital economy: A research agenda. Harvard Business School Strategy Unit Working Paper, (21-009).
[58] Nanda, R. Financing high-potential entrepreneurship. IZA World of Labor. 05530. 2016.
[59] North-Samardzic, A. (2020). Biometric technology and ethics: Beyond security applications. Journal of Business Ethics, 167(3), 433-450.
[60] Peterson, Robert A., Richard F. Beltramini, and George Kozmetsky. "Concerns of college students regarding business ethics: A replication." Journal of Business Ethics 10, no. 10 (1991): 733-738.
[61] Raisch, S., and Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. Academy of Management Review, 46(1), 192-210.
[62] Rambachan, A. and Roth, J., 2019. Bias in, bias out? Evaluating the folk wisdom. arXiv preprint arXiv:1909.08518.
[63] Silva, S., and Kenney, M. Algorithms, platforms, and ethnic bias. Communications of the ACM, 62(11), 37-39. 2019.
[64] Smith, B., and Browne, C.A. Tools and weapons: The promise and the peril of the digital age. Penguin. 2019.
[65] Spector, P. E. (2019). Do not cross me: Optimizing the use of cross-sectional designs. Journal of Business and Psychology, 34(2), 125-137.
[66] Tarafdar, M., Teodorescu, M., Tanriverdi, H., Robert, L. P., and Morse, L. Seeking Ethical use of AI Algorithms: Challenges and Mitigations. 2020.
[67] Teodorescu, M. H., Morse, L., Awwad, Y., and Kane, G. C. (2021). Failures of fairness in automation require a deeper understanding of human-ML augmentation. MIS Quarterly, 45(3), 1-17.
[68] Teodorescu, M. H., Morse, L., Awwad, Y., and Kane, J. A Framework for Fairer Machine Learning in Organizations. In Academy of Management Proceedings (Vol. 2020, No. 1, p. 16889). Briarcliff Manor, NY 10510: Academy of Management. 2020.
[69] Tsamados, A., Aggarwal, N., Cowls, J., Morley, J., Roberts, H., Taddeo, M., and Floridi, L. (2021). The ethics of algorithms: key problems and solutions. AI & SOCIETY, 1-16.
[70] Tucker, C. (2019). Digital data, platforms and the usual [antitrust] suspects: Network effects, switching costs, essential facility. Review of Industrial Organization, 54(4), 683-694.
[71] Wang, L. C., & Calvano, L. (2015). Is business ethics education effective? An analysis of gender, personal ethical perspectives, and moral judgment. Journal of Business Ethics, 126(4), 591-602.
[72] Whittaker, M., Crawford, K., and Dobbe, R. Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Mysers West, Rashida Richardson, Jason Schultz, and Oscar Schwartz. AI Now Report 2018. New York: AI Now Institute at New York University. 2018.

Appendix

Appendix A.1 –Survey Questions

Ethics Questions:

Q50. Has your firm taken any of the following actions?

	Yes (1)	No (2)	I don't know or N/A (3)
Offered unconscious bias training (1)			
Hired an under-represented minority or female programmer (2)			
Considered gender or racial diversity as criteria for selecting training data (3)			
Sought expert advice on navigating ethical issues (4)			

51 Does your firm have a set of ethical AI principles?

Yes
No
I don't know

52 If your firm has ethical AI principles, has your firm ever:

	Yes (1)	No (2)	I don't know or N/A (99)
Turned down business due to a conflict with these ethical AI principles (1)			
Dismissed an employee that did not follow these ethical AI principles (2)			
Stopped using certain training data that did not align with these ethical AI principles (5)			

Data Sharing Question:

30 Do you collaborate with other technology firms to access data?
Yes
No
I don't know

GDPR Questions:

37 Have you created a new position to handle compliance associated with the General Data Protection Regulation ("GDPR")?
Yes
No
I don't know or N/A

38 Have you reallocated resources in your firm to handle GDPR compliance?
Yes
No
I don't know or N/A

39 Have you deleted some data in order to comply with GDPR?
Yes
No
I don't know or N/A

Appendix B – Correlations

Table B.1 – AI Ethics Principles and Ethical Outcomes Correlations

	Ethical AI Principles?	Dismissed Employee	Dropped Data	Turned Down Business	Funding before 2019	Tech Firm Data
Dismissed Employee	0.2263*					
	0.0051					
Dropped Data	0.4317*	0.3267*				
	0	0				
Turned Down Business	0.4578*	0.4307*	0.4336*			
	0	0	0			
Funding before end of 2019	-0.0756	-0.0355	-0.1126	0.0441		
	0.356	0.6656	0.1686	0.5905		
Tech Firm Data Collaboration	0.2313*	0.1882*	0.2013*	0.2791*	0.1349	
	0.0041	0.0202	0.0129	0.0005	0.0985	
GDPR (Prior Capability)	0.2477*	0.0959	0.2234*	0.0897	-0.0087	0.0683
	0.0021	0.2398	0.0057	0.2719	0.9157	0.4029

Table B.2 – Firm Ethical Action Correlations

	Diverse Training Data	Sought Expert	Hired Minority/ Female	Bias Training	Funding before 2019	Tech Firm Data
Sought Expert	0.2103*					
	0.0093					
Hired Minority/ Female Prog.	0.2137*	0.1502				
	0.0082	0.0647				
Bias Training	0.3604*	0.2396*	0.1337			
	0	0.0029	0.1004			
Funding before 2019 (Dummy)	0.0886	0.0733	-0.0758	-0.0541		
	0.2793	0.3713	0.3552	0.5096		
Tech Firm Data Collaboration	0.1774*	0.1191	0.1676*	0.0494	0.1349	
	0.0287	0.1439	0.039	0.5454	0.0985	
GDPR (Prior Capability)	0.1159	0.139	0.1635*	0.1646*	-0.0087	0.0683
	0.1552	0.0876	0.0441	0.0428	0.9157	0.4029

Table B.3 – Firm Demographics Correlations

	US	VC Location	Less than 11 Emps.
VC Location	0.2103*		
	0.0093		
Less than 11 Employees	0.2137*	0.1502	
	0.0082	0.0647	
Less than 3 years old	0.1159	0.139	0.1635*
	0.1552	0.0876	0.0441

* p<0.05

Appendix C – Robustness

Table C.1 - AI Ethics Principles (Funding)

	(1)	(2)	(3)	(4)	(5)	(6)
DV, Dummy:	**Does your firm have Ethical AI Principles?**					
Funding before 2019 (dummy)	-0.051	-0.046	-0.058			
	(0.083)	(0.085)	(0.086)			
Inst. Investors (Series A or later, dummy)				-0.243***	-0.231**	-0.250***
				(0.090)	(0.092)	(0.093)
Employees (<11)		0.005	0.006		-0.019	-0.020
		(0.084)	(0.084)		(0.081)	(0.082)
VC Location			0.154			0.208
			(0.155)			(0.130)
CEM Weighting:	No	Yes	Yes	No	Yes	Yes
Firms	159	159	159	159	159	159
Adj R2	-0.001	-0.019	-0.011	-0.004	-0.020	-0.012

* p<0.1, ** p<0.05, *** p<0.01. Models (1)-(3) show that prior funding is not correlated with ethical AI principles. Standard errors are robust.

Table C.2 – Probit AI Ethics Principles

	(1)	(2)	(3)	(4)
DV, Dummy:	**Does your firm have Ethical AI Principles?**			
CEM:	All Models Weighted			
Funding before 2019 (dummy)	-0.145			
	(0.212)			
Inst. Investors (Series A or later, dummy)		-0.706***		
		(0.243)		
Tech Firm Data Collab.			0.590***	
			(0.209)	
GDPR (Prior Capability)				0.638***
				(0.212)
Employees (<11)	-0.116	-0.195	-0.012	-0.028
	(0.209)	(0.213)	(0.214)	(0.213)
VC Location	0.407	0.614	0.357	0.479
	(0.435)	(0.463)	(0.437)	(0.440)
Firms	159	159	159	159

* p<0.1, ** p<0.05, *** p<0.01. Alternate specification, Probit instead of OLS, for robustness.

Table C.3 - AI Ethics Principles (Prior Resources) with IMR

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DV, Dummy:	**Does your firm have Ethical AI Principles?**							
Tech Firm	0.229***	0.216***	0.225***	0.229***				
Data Collab.	(0.076)	(0.079)	(0.081)	(0.081)				
GDPR					0.243***	0.229***	0.235***	0.238***
					(0.080)	(0.082)	(0.083)	(0.083)
IMR		-0.044	-0.029	-0.317		0.111	0.128	-0.152
		(0.286)	(0.298)	(0.338)		(0.301)	(0.308)	(0.349)
Emp. (<11)			0.048	0.040			0.039	0.032
			(0.086)	(0.087)			(0.086)	(0.086)
VC Location				0.235*				0.231*
				(0.138)				(0.132)
CEM:	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Firms	159	159	159	159	159	159	159	159
Adj R2	0.0467	0.0346	0.0305	0.0318	0.0506	0.0377	0.0330	0.0340

* p<0.1, ** p<0.05, *** p<0.01. Models (2)-(4) and (6)-(8) include IMR to control for selection issues. Standard errors are robust.

Table C.4 – AI Ethics Principles (Lasso Model, Control Selection)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DV, Dummy:	**Does your firm have Ethical AI Principles?**							
CEM:	All Models Weighted							
Funding before 2019 (dummy)	-0.063	-0.074						
	(0.083)	(0.082)						
Inst. Investors (Series A or later, dummy)			-0.235**	-0.246***				
			(0.091)	(0.086)				
Tech Firm Data Collab.					0.247***	0.246***		
					(0.077)	(0.078)		
GDPR (Prior Capability)							0.258***	0.228***
							(0.080)	(0.077)
Lasso Procedure	DS	XPO	DS	XPO	DS	XPO	DS	XPO
Firms	159	159	159	159	159	159	159	159

* p<0.1, ** p<0.05, *** p<0.01.

Appendix D – CEM Matching

Table D.1 - CEM

	Treatment: Established an ethical AI policy?			
	Pre/All		CEM Weighted	
	No	Yes	No	Yes
Young	0.75	0.56	0.56	0.56
US HQ	0.41	0.41	0.41	0.41
Small	0.37	0.28	0.28	0.28
Firms	159	159	159	159