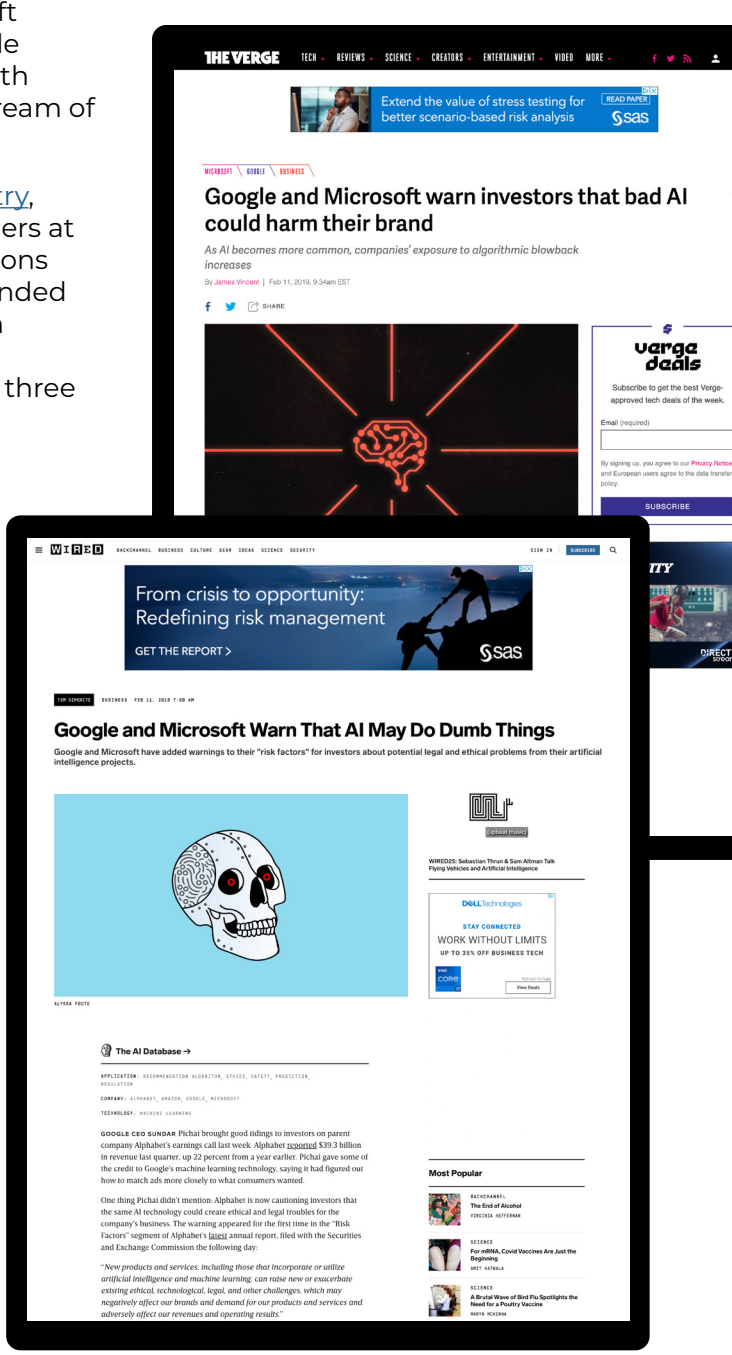
# The Rise of **AI Risk Disclosure**

*How enterprises are disclosing AI risk—and what CIOs and other executives should consider when assessing an AI risk management and disclosure strategy*

arize

# Introduction

Three years ago, Alphabet and Microsoft made waves when they disclosed their use of AI as a risk factor in their annual financial reports. The headlines at the time were unforgiving, from "Google and Microsoft Warn AI May Do Dumb Things" to "Google and Microsoft Warn That Bad AI Can Harm Their Brand." While the controversy has long since passed, it is worth considering the broader ramifications downstream of these early disclosures.




AI is now being deployed in nearly every industry, informing everything from holiday apparel orders at top retailers to potentially life-saving interventions in cancer care. As a result, any risks and unintended consequences from using AI to automate even seemingly benign tasks—such as screening resumes—are likely far more widespread than three years ago.

How many top companies are communicating this to investors and stakeholders? To answer this question, Arize reviewed the two most recent (as of January 1st, 2022) annual financial reports (10-Ks) of U.S.-based Fortune 500 companies as well as the companies' most recent corporate social responsibility or environmental, social and governance (ESG) reports. Since fiscal years vary, comparisons refer to "most recent" and "one year prior" as of the date of collection.

The addition of ESG reports reflects the fact that stakeholder capitalism is now in the mainstream in most boardrooms. In the past three years, a majority of Fortune 500 CEOs signed on to either the Business Roundtable's "Statement of Purpose of a Corporation" or the World Economic Forum's stakeholder capitalism pledge. As part of those commitments, companies are now reporting on ethics and their broader impact to the community and environment in both existing annual financial reports and bespoke corporate responsibility or ESG reports.

# Highlights

**↑ 46.4%**

YoY **increase** in the number of mentions of AI and machine learning in Fortune 500 company annual financial reports

**↑ 20.5%**

YoY increase in the number of Fortune 500 companies citing AI as a risk factor in their annual financial reports

**80**

Fortune 500 companies mention AI and machine learning in their most recent annual financial reports but do not cite AI as a risk factor

**47**

companies—around one in ten (9.4%) in the Fortune 500—cite AI and machine learning as a risk factor in their most recent annual financial reports

**25**

companies specifically mention AI fairness (i.e. algorithmic bias) or AI ethics efforts in their most recent ESG reports

# Recommendations

### Assess Current AI Systems & Implement Machine Learning Observability

Despite the maturation of AI across industries, there is still a wide gap between the technical teams who build and deploy machine learning models and the business executives that rely on them. According to a recent survey, 87.6% of machine learning teams say that their business counterparts cannot consistently quantify AI ROI—and nearly one in five (19.7%) technical executives say they often struggle with accessing and sharing data on AI efforts.

As a result, many enterprises likely lack perfect knowledge about where AI is being used and the various ways it presents business risks. As Cerner notes in its recent annual report, "the lack of standards for measuring the accuracy and effectiveness of ML and AI can raise new or exacerbate existing technological, legal or other challenges." The fact that machine learning operations (MLOps) is still relatively new as a discipline likely amplifies this confusion.

To begin addressing informational asymmetries between teams, organizations should ensure data scientists and ML engineers have access to tools that tie model metrics to business risks and results, making relevant dashboards available to business executives. Bridging the worlds between business intelligence and data science is essential for the ongoing vitality of both.

However, knowing there is a problem is only half the battle; teams also need to figure out why, and therein lies the rub. 84.3% of data scientists and ML engineers cite the time it takes to detect and fix issues with their machine learning models as a pain point today, with over one in four saying it takes them a week or more. Full-stack machine learning observability with ML performance tracing can help close this gap by helping teams automatically pinpoint the source of model performance problems. Leveraging a platform like Arize, teams can make multidimensional comparisons to surface the features and cohorts contributing to algorithmic bias or performance issues and adjust accordingly.

### Consider Carefully Whether Disclosure Is Warranted

When it comes to deployed AI, it is a matter of when – not if – models will encounter issues in production. Unlike the largely rules-based system of software development, successful outcomes in machine learning are dependent on not just system health but also the various complexities of models and underlying data layers. Concept and feature drift, training-production skew, cascading model failures and outliers challenge even the most sophisticated machine learning teams deploying models that perform flawlessly in training.

The decision of whether to disclose the resulting risks on an annual report is likely unique to every company, though the upward trajectory of disclosures suggests more companies are considering doing so given legal uncertainty and real-world challenges. It is worth noting that

software and technology companies, which have the largest data science teams on average, lead all other sectors in disclosing AI risk and in discussing AI ethics in ESG reports.

As legal researcher Sylvia Lu argues in the Vanderbilt Law Review, “the current disclosure framework does not specify what forms of risk are subject to disclosure, leading to firms’ omissions of substantial risks posed by algorithms.”

### Develop An AI Risk Management and AI Ethics Governance Framework and Practice

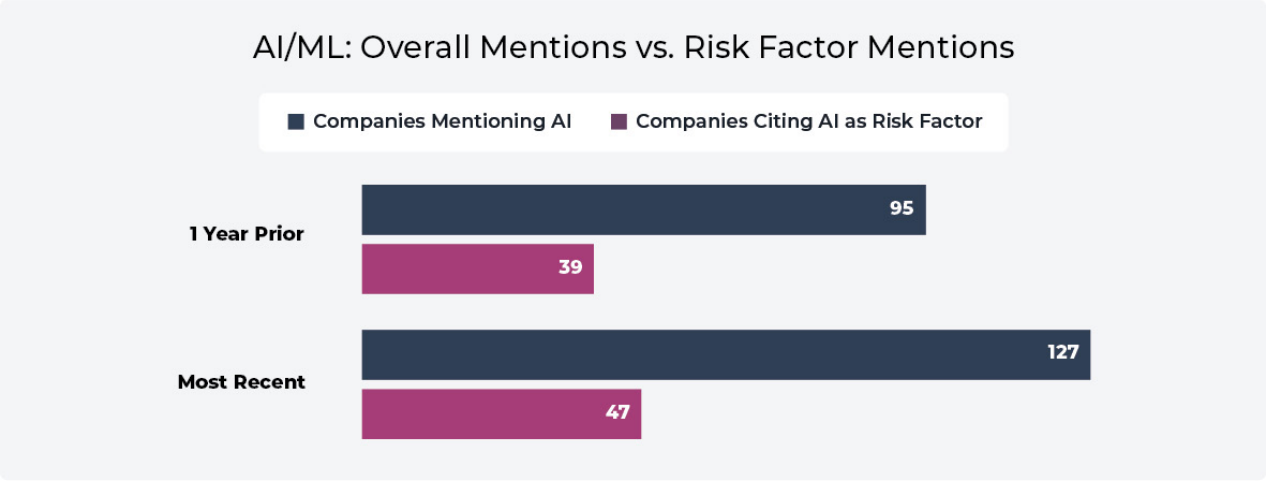
A growing number of enterprises are developing frameworks to both define potential risks upfront and guide ethical AI product development and risk management. By defining what can go wrong and implementing systems to identify both business and ethical risks, organizations can begin to anticipate problems before they happen and incentivize employees to act.

For technical teams, machine learning observability and proactive model monitoring to automatically surface problems at a cohort-level can help achieve that vision. Of course, technology alone is insufficient; a multi-pronged approach that combines purpose-built infrastructure, governance, and dedicated teams is key. Intel, for instance, combines “Ethical Principles for AI Development” with an "Ethical AI impact assessment process” and a “Responsible AI Advisory Council” to guide and provide oversight, according to the company’s Corporate Responsibility Report. “This multi-pillar framework helps address potential issues such as maintaining integrity and protecting privacy while collecting and using data to train AI systems, reducing the risk of harmful bias in AI systems, and building trust in machine learning applications by helping people who interact with these technologies better understand them,” the company says.

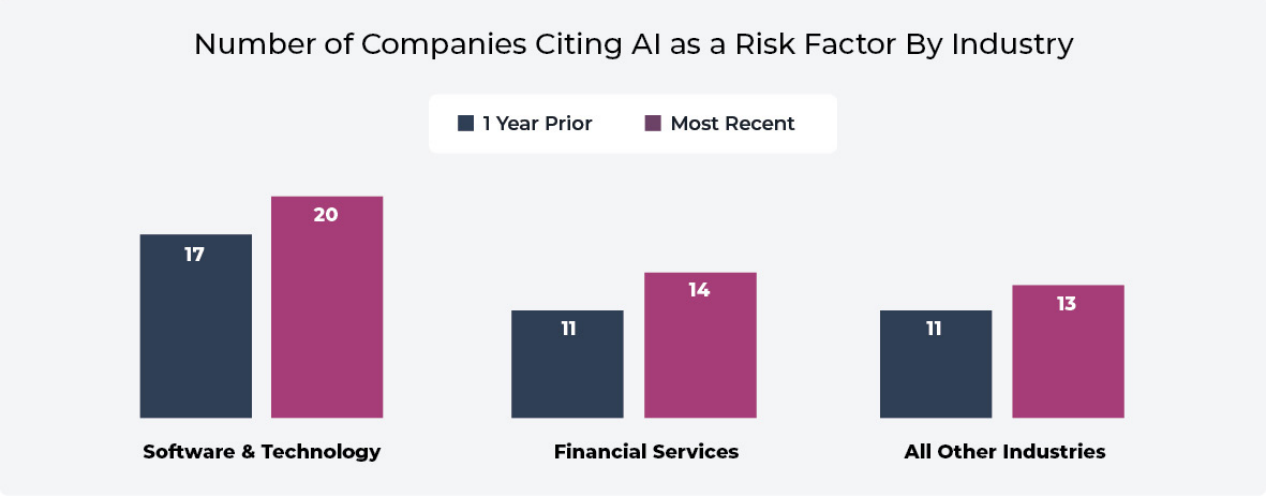
Several larger enterprise clients of Arize also have dedicated model risk officer teams whose job is to put guardrails on AI, quantifying and mitigating any risks to the business. While this practice is not universal yet, there are good reasons to believe it will become more common.

# Annual Financial Reports

One in five Fortune 500 companies now mentions AI or machine learning in their annual financial reports, an increase of 33.7% year-over-year. Under half of these companies (37.0%) cite AI as a risk factor, however.




Among the companies that do cite AI as a risk factor, most are either in the software and technology or financial services industries. The other fields where risk disclosures are on the rise year-over-year are energy and healthcare.




The reasons AI is cited as a risk factor generally fall into three categories, though they often overlap and are rarely standardized in their wording.

## Categories of AI Risk: **Competitive Risks**

### Broad Definition

Failing to keep pace with competitors or achieve AI goals

### Real-World Examples

"Other technological changes also present competitive risks. For example, our competitive position could be impacted if we are unable to deploy, in a cost effective and competitive manner, technology such as artificial intelligence and machine learning that collects and analyzes a wide variety of data points (so-called "big data" analysis) to make underwriting or other decisions, or if our competitors collect and use data which we do not have the ability to access or use."

**—The Travelers Companies [1]**

"The development of new technologies (including artificial intelligence technologies) may impact the healthcare industry, and the development of new, more cost-effective solutions that can be performed by our customers or by patients, and the continued internalization of testing by hospitals or clinicians, could negatively impact our testing volume and revenues."

—**Quest Diagnostics [2]**

## Categories of AI Risk: **General Harms**

### Broad Definition

Physical, reputational, or other harms to company or its stakeholders from AI -- often touching on AI fairness or responsible use of AI

### Real-World Examples

"Social and ethical issues, including the use of AI in our offerings, may result in reputational harm and liability. We are increasingly building AI into many of our offerings. As with many innovations, AI and our Customer 360 platform present additional risks and challenges that could affect their adoption and therefore our business. For example, the development of AI and Customer 360, the latter of which provides information regarding our customers' customers, presents emerging ethical issues and if we enable or offer solutions that draw controversy due to their perceived or actual impact on human rights, privacy, employment, or in other social contexts, we may experience brand or reputational harm, competitive harm or legal liability. Data practices by us or others that result in controversy could impair the acceptance of artificial intelligence solutions. This in turn could undermine the decisions, predictions or analysis AI applications produce, subjecting us to competitive harm, legal liability, and brand or reputational harm."

**—Salesforce [3]**

[1] Source: https://www.sec.gov/ix?doc=/Archives/edgar/data/0000086312/000008631221000011/trv-20201231.htm
[2] Source: https://www.sec.gov/ix?doc=/Archives/edgar/data/0001022079/000102207921000029/dgx-20201231.htm
[3] Source: https://s23.q4cdn.com/574569502/files/doc_financials/2021/ar/Salesforce-FY-2021-Annual-Report.pdf

## Categories of AI Risk: **General Harms (cont'd)**

### Broad Definition

Physical, reputational, or other harms to company or its stakeholders from AI -- often touching on AI fairness or responsible use of AI

### Real-World Examples

"Our use of artificial intelligence and machine learning is subject to risks related to flaws in our algorithms and datasets that may be insufficient or contain biased information."

**—American Express** [4]

"We envision a future in which artificial intelligence operating in our products and services will help our public safety and private sector customers build safer communities with stronger communication platforms. Artificial intelligence may be flawed and datasets may be insufficient or contain biased information. As we work to responsibly meet our customers' needs for products and services that use artificial intelligence, we could suffer reputational damage as a result of any inconsistencies in the application of the technology or ethical concerns both of which may generate negative publicity."

**—Motorola Solutions** [5]

## Categories of AI Risk: **Regulatory Risk**

### Broad Definition

Regulation of AI upending data pipelines or business lines relying on ML models

### Real-World Examples

*"We believe that providing insights from data, including artificial intelligence and machine learning, will become increasingly important to the value that our solutions and services deliver to our customers. However, the ability to provide data-driven insights may be constrained by current or future regulatory requirements or ethical considerations that could restrict or impose burdensome and costly requirements on our ability to leverage data in innovative ways."*

**—Automatic Data Processing** [6]

"The legislative and regulatory environment is beyond our control, may change rapidly and unpredictably and may negatively influence our revenue, costs, earnings, growth, liquidity and capital levels. In addition, some rules and regulations may be subject to litigation or other challenges that delay or modify their implementation and impact on us. Adoption of new technologies, such as distributed ledger technologies, artificial intelligence and machine learning technologies, can present unforeseen challenges in applying and relying on existing compliance systems."

**—Capital One** [7]

[4] Source https://www.sec.gov/ix?doc=/Archives/edgar/data/0000004962/000000496221000013/axp-20201231.htm

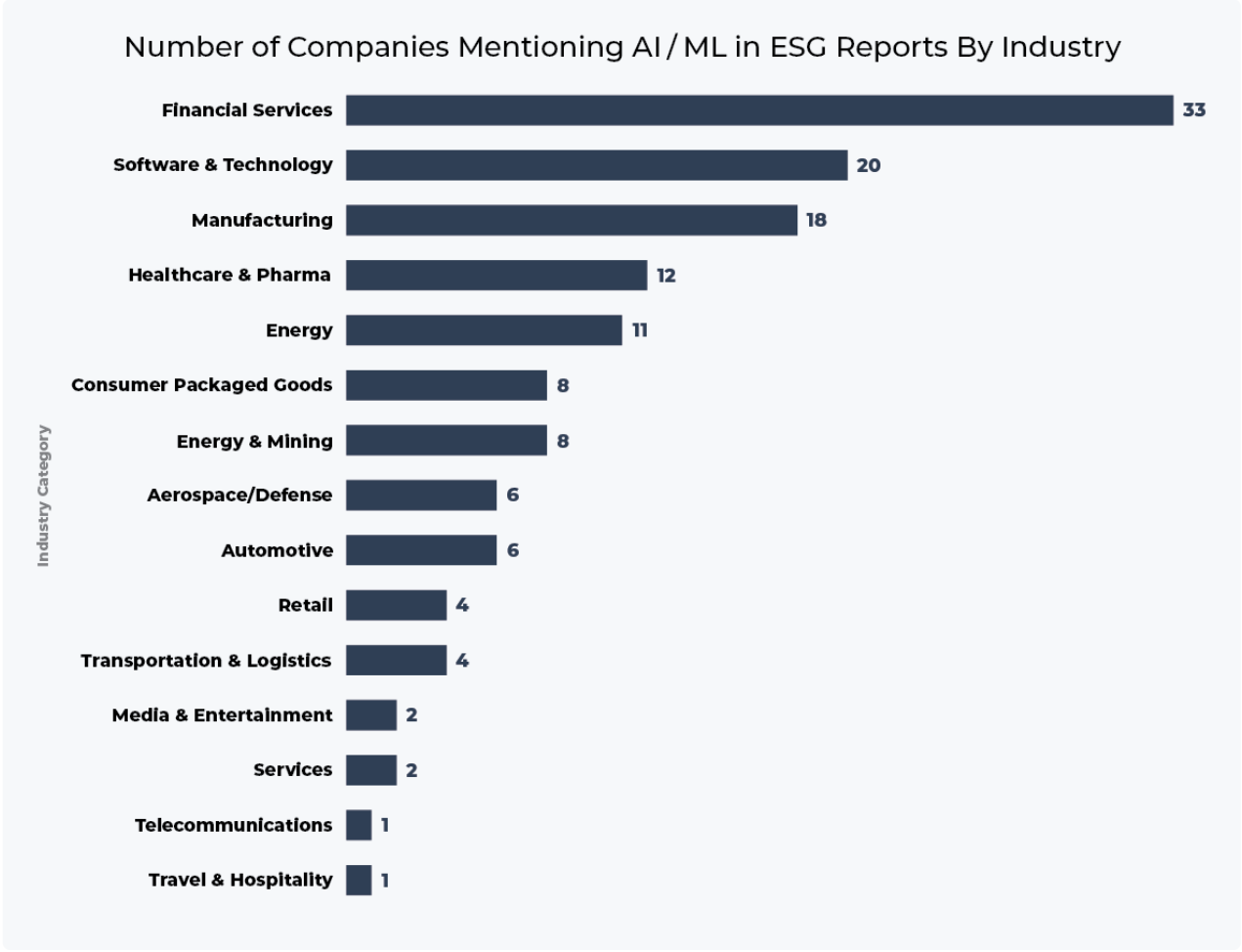
[5] Source https://www.sec.gov/ix?doc=/Archives/edgar/data/0000068505/000006850521000008/msi-20201231.htm

[6] Source https://www.sec.gov/ix?doc=/Archives/edgar/data/0000008670/000000867021000027/adp-20210630.htm

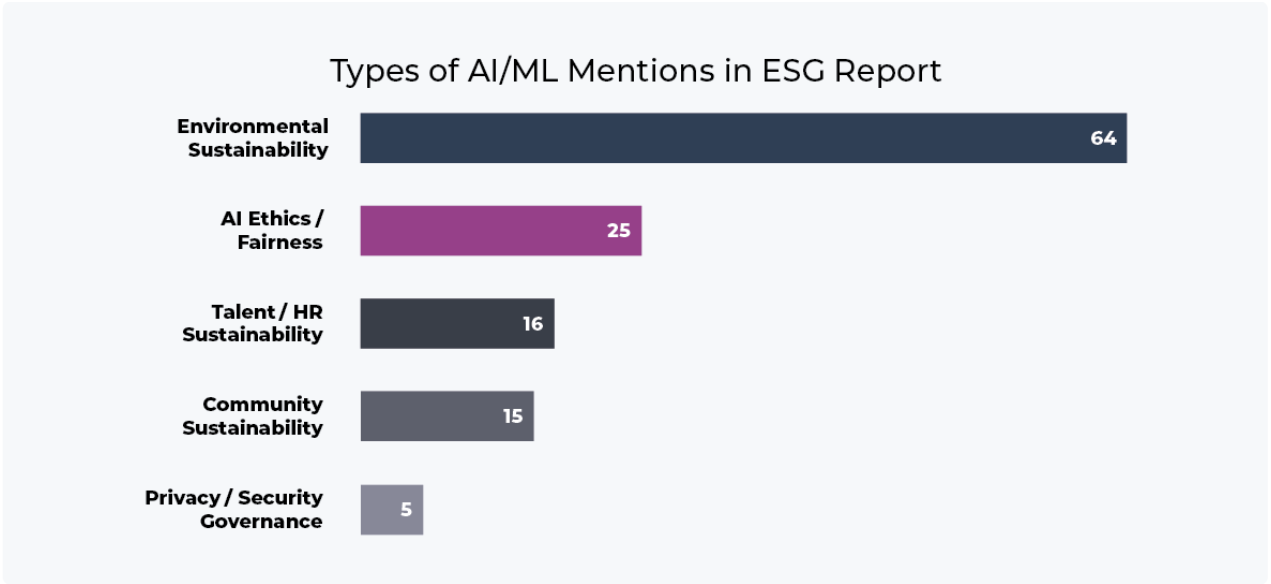
[7] Source https://www.sec.gov/ix?doc=/Archives/edgar/data/0000927628/000092762821000094/cof-20201231.htm

# Annual ESG Reports

ESG reports tell a similar story, with many modern enterprises boasting of the benefits of AI in achieving environmental and broader sustainability goals while relatively fewer acknowledge some of broader ethical and fairness concerns involved.




76% of the mentions of AI in ESG reports discuss the myriad benefits of applied AI rather than referencing ethical governance. In all, 25 companies discuss responsible AI in some form. Only four companies mention model transparency, explainability or investments in machine learning infrastructure designed to ensure AI fairness.




Examples of each category appear below. Similar to financial reports, these often overlap and are rarely uniform in wording.

## Categories of AI Mention: **Environmental Sustainability**

### Broad Definition

Applied AI used to reach environmental goals

### Real-World Examples

*"The Big Data department is actively involved in Enterprise's effort to reduce emissions and improve our operational efficiency. Using realtime data and analytics, we are able to detect and alert flare emission anomalies in record time. This data enables immediate reductions in flare emissions, improving our overall environmental impact. We have an adaptable approach to Machine Learning and Artificial Intelligence, and we are developing additional programs that we expect to further aid in the prediction and prevention of emissions."*

**— Enterprise Product Partners**[8]

[8] Source: http://www.enterpriseproducts.com/documents?download=2019-2020-Sustainability-Report&type=PDF

## Categories of AI Mention: **AI Ethics / Fairness**

| Broad Definition | Real-World Examples |
| --- | --- |
| Broader ethical implications of AI as it's deployed into the real-world | *"While we leverage the power of artificial intelligence to help our colleagues grow in their careers and better serve our customers, we understand the need for safeguards to mitigate the risks associated with its use. In 2020, we established an Artificial Intelligence Governance Forum to review and deliberate on key ethical considerations around the use of artificial intelligence/ machine learning technologies including bias management, transparency, explainability, fairness, quality, privacy and overall soundness. The forum meets regularly and includes key stakeholders from across Citizens to ensure it has appropriate visibility and risk identification to distinguish between intended and unintended consequences in implementation."*<br>—**Citizens Financial Group** [9] |

## Categories of AI Mention: **Talent / HR Sustainability**

| Broad Definition | Real-World Examples |
| --- | --- |
| Applied AI to develop, retain talent and ensure an inclusive workforce | *"Our talent development initiatives provide employees with self-assessment tools, learning experiences, formal and informal education, mentoring opportunities and rotations to grow their skills and careers. These programs offer curated and experience-based learning opportunities that leverage artificial intelligence and machine learning to identify and provide individualized development opportunities. We offer robust career development through our Center for Clinical Advancement, Optum Tech University, Technical Leadership Development, and investments in the training and development of our team members."*<br>—**UnitedHealth Group** [10] |

## Categories of AI Mention: **Community Sustainability**

| Broad Definition | Real-World Examples |
| --- | --- |
| Applied AI to benefit community at large or a broader set of company stakeholders than customers/investors | *"To improve antibiotic use, IQVIA is committed to building on an existing framework of antibiotic prescribing reporting and analysis and using innovative methods such as machine learning to better understand the data. We are providing state-level antimicrobial prescription reports in the U.S. for all 50 states."*<br>—**IQVIA Holdings** [11] |

[9] Source https://www.citizensbank.com/assets/pdf/2020-corporate_responsibility.pdf

[10] Source https://sustainability.uhg.com/content/dam/UHG/PDF/sustainability/final/2020_SustainabilityReport.pdf

[11] Source https://www.iqvia.com/-/media/iqvia/pdfs/about-us/2019-sustainability-report.pdf

Categories of AI Mention: **Privacy/Security Governance**

| Broad Definition | Real-World Examples |
| --- | --- |
| Applied AI to aid in privacy, security and data governance | *"We employ a wide array of policies, procedures and technologies designed to protect client and employee data against unauthorized disclosure, modification or misuse and meet regulatory requirements. Our cybersecurity, infrastructure and risk professionals respond to disruptive events with sophisticated data analytics and threat intelligence. Dedicated teams protect computer systems firmwide in coordination with the Fusion Resilience Center. Our defenses include artificial intelligence and machine-learning technologies that can identify malicious behavior and vulnerabilities in our networks."*<br>—**Morgan Stanley** [12] |

[12] Source https://www.morganstanley.com/content/dam/msdotcom/sustainability/Morgan-Stanley_2020-Sustainability-Report_Final.pdf

# Conclusion

If annual reports from a broad array of companies make one thing clear, it's that the rise of AI-powered systems is leading to incredible breakthroughs in improving productivity and even saving lives across many industries. However, clear-eyed assessments of the associated risks and ethical issues inherent with deploying machine learning systems into a flawed real world are needed. Increasingly, many enterprises are leveraging annual reporting to do just that.