Ethical AI

Building the automated machines of the future.

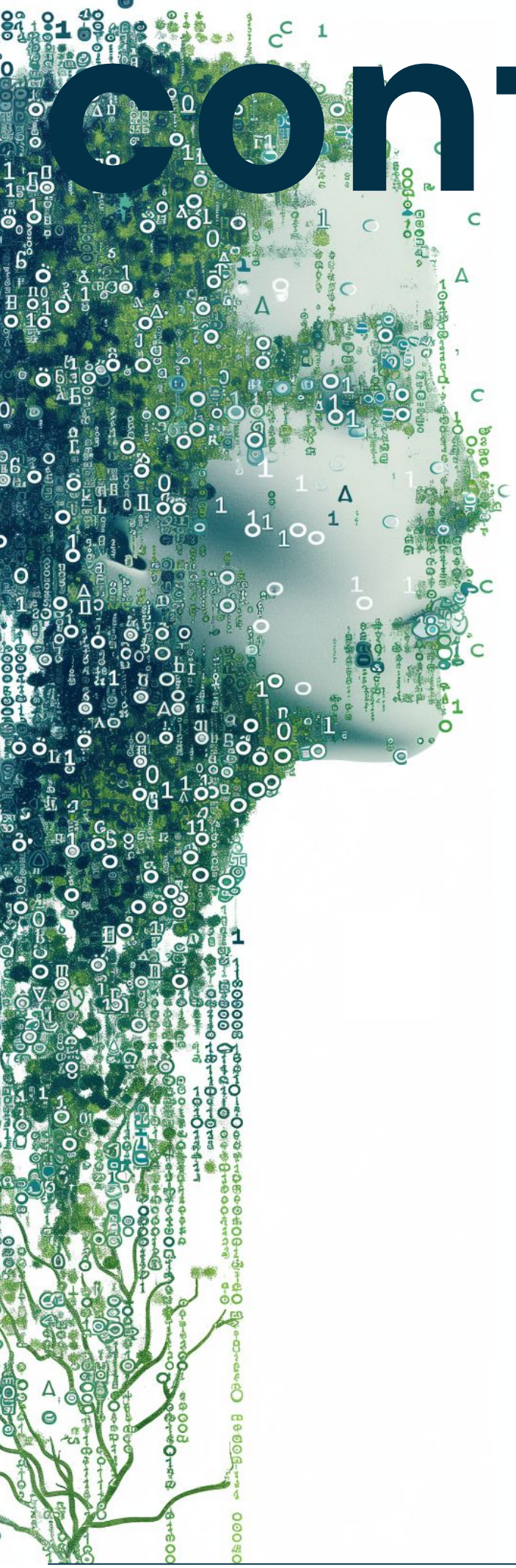
"The ethical deployment of AI in businesses is more than a tech upgrade; it's a fundamental shift towards aligning innovation with human values."

- ChatGPT

EXADEL

info@exadel.com // exadel.com

contents



61%

of companies cannot explain the decisions their AI makes

68%

of companies do not monitor changes in their AI performance and model accuracy

74%

of companies do not ensure their AI reduces bias

Source: IBM

AI made at least one decision today that directly affected your life. It knows all about your shopping habits and preferences. It gives you the next show to watch, navigates those windy country roads, and probably decides your credit score.

We all trade AI some decision making power in return for unprecedented efficiency. But how can we ensure that the decisions it makes reflect the ethical standards of our companies?

What happens when AI gets it wrong?

One corporate study on AI bias used an image generator to create archetypal people based on salary bands. High salary prompts like *lawyer* and *architect* produced pictures of people with lighter skin tones, while inputs like *fast food worker* and *social worker* generated images of people with darker skin tones.

This is not an isolated case of bias. The laundry list of ethical mishaps and subsequent public fallout is long.

AI is as risky as it is profitable.

We're not here to preach to you about the exact moral compass your AI should follow. Let's leave the social debate to our boardrooms and Zoom rooms.

Our goal is to help you understand the current state of AI ethics in plain terms.

We want to help you ask the right questions and get clear answers so you can make smart technology decisions.

We want to help you hedge against regulation, become leaner and more profitable, and also rest easy — knowing that the machines that work for you are programmed to represent your values.

Artificial intelligence is far too beneficial to ignore.

According to a recent survey of business leaders, 91% of executives are now willing to invest in AI.

But, where should you put your money?

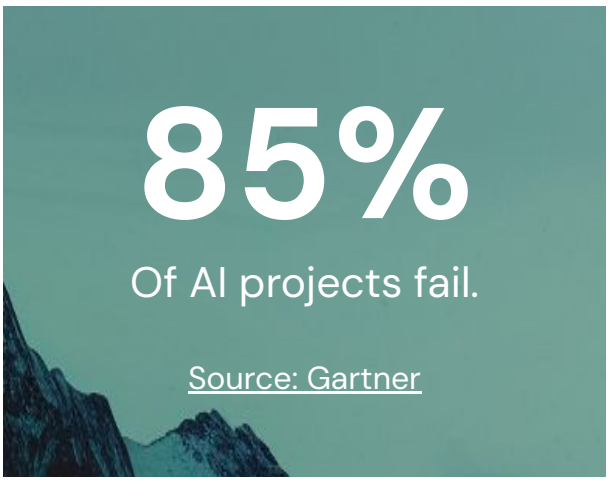
Process Automation
Complete time-consuming and routine tasks faster than ever with process automation. Instead of wasting time on repetitive tasks, your team can focus on new technology strategies that add more value to your business.

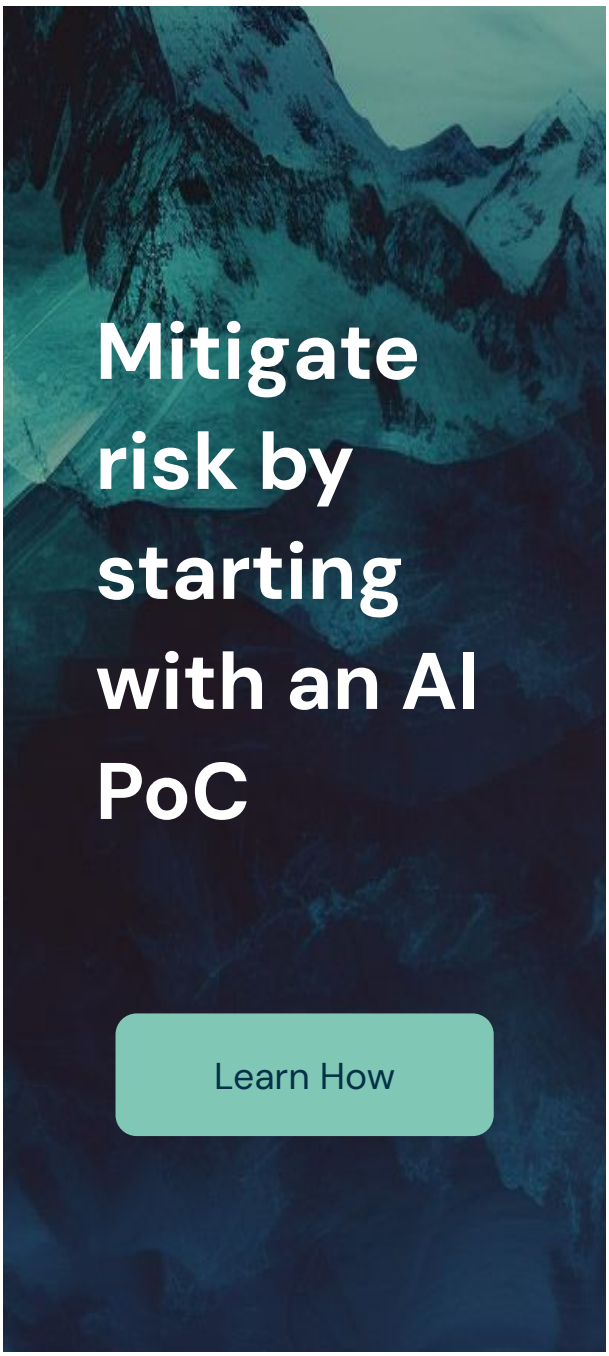
Cost Reduction
Streamline operations, automate tasks, and optimize resource allocation. With predictive analytics and process optimization, AI helps you improve efficiency, minimizes downtime, and identifies areas for savings.

Improved Decision-Making
Uncover patterns and extract actionable insights from extensive datasets. Have AI sift through your historical data and current trends so you can make well-informed decisions. Refine strategies, improve operational efficiency, and achieve optimal outcomes, regardless of your field.

Better Customer Experience
Integrate chatbots and intelligent agents into your services to give superior customer experiences. Chatbots, equipped with NLP capabilities can engage your customers in real-time by addressing queries, providing information, and even assisting with transactions 24/7.

Competitive Advantage
Turn market trends and consumer behavior into a continuous advantage. Use AI algorithms to analyze customer preferences, predict spending patterns, and provide personalized recommendations.





Leaders are concerned about the machines... and for good reason.

AI Bias

AI bias occurs when algorithms are trained on inaccurate or poor-quality data, resulting in fittingly inaccurate results. Biases are typically introduced into a system in two ways:

- **Primary bias** is related to the input of biased data, relying on biased modeling as well as bias creeping into data processing and reporting.
- **Secondary bias,** whereby the above biases affect the equality of opportunities and outcomes. This is reflected in models inadvertently favoring one group over another.

The resulting inequalities can be categorized as:

- **Illegal bias** that contravenes regulatory constraints.
- **Unfair bias,** which, while not illegal, is in conflict with the company's values.
- **Inherent bias**, which stems from flawed data, algorithmic bias and confirmation bias - where a model is trained on observational data drawn from its own predictions.

It's all too easy for human bias, whether conscious or unconscious, to penetrate machine learning algorithms during development.

This can be overcome by implementing thorough data collection processes and continuous monitoring. The result is significantly enhanced performance and improved reliability of the AI model.

Resistance to Change

Not everybody is ecstatic over the all-too-real cliché of the robots coming to steal our jobs.

According to a World Economic Forum report, nearly a quarter of all jobs globally are expected to change in the next five years due to technological trends, including the integration of AI into various industries.

The reality is that employees who do not adapt and retrain will simply not have a place in the modern market.

What responsibility do companies have?

As of now, the law hasn't fully caught up, but that's likely to change soon.

According to a study conducted by the Pew Research Center, 85% of Americans support legislation that would limit the automation of jobs. Although no specific law exists, guidance from entities like the Equal Employment Opportunities Commission (EEOC) aims to help employers navigate existing laws like the Americans with Disabilities Act (ADA) when using AI.

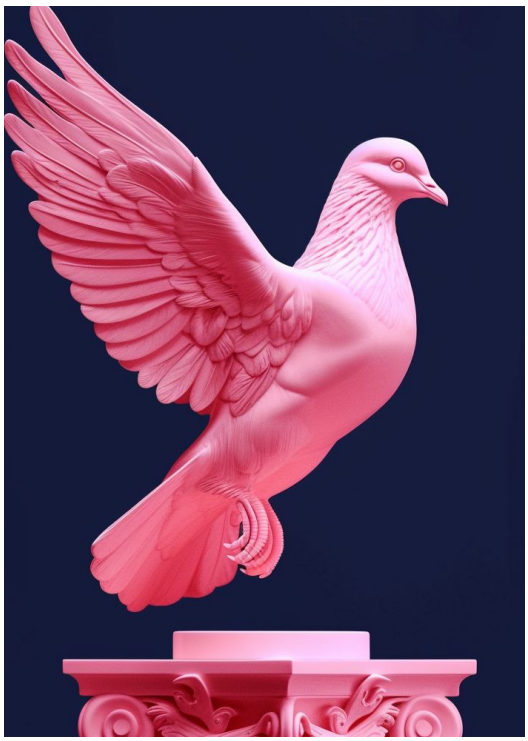
The Black Box

Advanced algorithms frequently struggle to provide explanations for their decision-making processes.

Explainable AI systems, while offering transparency, often fall short in terms of cognitive abilities and performance. They're generally less able to provide the best predictive and prescriptive analytics for businesses.

Non-interpretable AI systems, on the other hand, are not considered reliable in highly-regulated use cases or in high-stakes situations. This makes them unsuitable for practical use.

If business leaders don't understand how their AI works, who then is responsible for its outputs? And if these outputs don't match our moral standards, how do we fix a machine that we can't understand?



Data privacy isn't just a tech trend. It's a human right.

Effective privacy programs match your technology and risk tolerance to ever-changing regulations.

Learn More

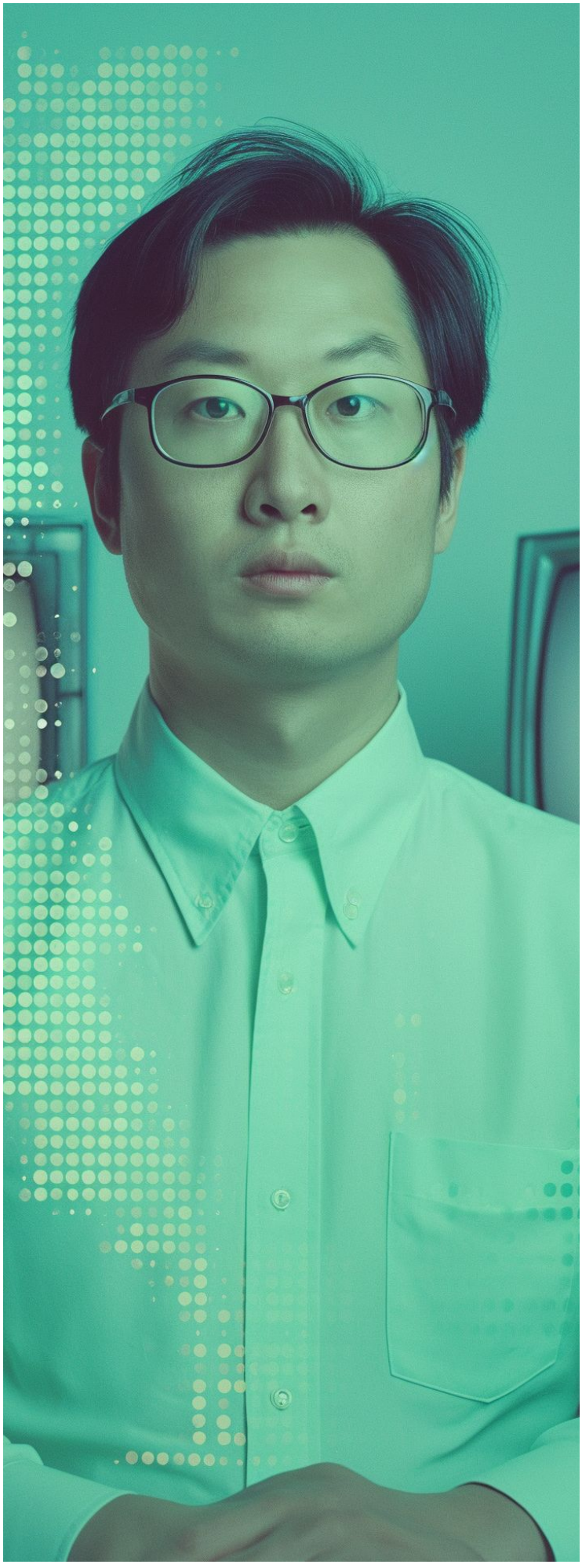
The law slowly fights back

The General Data Protection Regulation (GDPR) regulates the ethics of AI by enforcing principles such as transparency, data protection, and accountability. It mandates that individuals must be informed about the collection and use of their personal data, including when AI is involved. GDPR also requires that AI systems using personal data be designed with privacy in mind (privacy by design) and that data subjects have rights to access, rectify, or delete their data. Additionally, GDPR stipulates that organizations must be able to explain decisions made by AI systems, ensuring accountability and fairness.

The FTC regulates the ethics of AI by enforcing laws against deceptive or unfair business practices, including those involving AI technologies. It emphasizes the importance of transparency, fairness, and accuracy in AI systems, especially those that make decisions affecting consumers. The FTC has issued guidance for companies using AI, stressing the need for explainability, transparency, and fairness, and has the authority to take action against companies that use AI in ways that are deemed deceptive, unfair, or discriminatory.

More recently, the US government introduced an executive order on *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*. This order sets up a regulatory framework for more than 50 federal entities to engage in over 100 specific actions to be implemented across eight overarching policy areas.

The Insitute of Electrical and Electronics Engineers (IEEE) addresses the ethics of AI through guidelines and standards aimed at ensuring ethical design and implementation. It emphasizes principles like transparency, accountability, and privacy. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems offers comprehensive recommendations for ethically aligned design, focusing on prioritizing human well-being in AI systems. This initiative encourages professionals to consider the broader impacts of AI technologies on society and the environment, advocating for responsible innovation. While not legally binding, these standards can help companies prepare for current and upcoming regulation.



A new ethical code emerges.

As the law slowly catches up to societal expectations, a new ethical code emerges that can help companies prepare for the looming backlash against AI.

Microsoft's principles of Ethical AI does an excellent job of capturing the mood and concerns of major corporations. It covers everything from accountability and inclusiveness to reliability, safety, fairness, transparency, privacy and security.

Accountability

Organizations can draw on relevant industry standards to produce accountability norms. These standards help guarantee that AI systems don't have the final say in any decisions that can negatively affect people's lives, and that people retain meaningful control over highly autonomous AI systems. Facebook and X use AI to identify and delete harmful content, like hate speech and misinformation. By taking responsibility for content moderation, they uphold ethical standards and foster a more secure online community for users.

>75%

of companies have not implemented AI ethics

Source: Datamation

Inclusiveness

Inclusiveness is pivotal as it ensures that AI technologies are accessible and beneficial to everyone, regardless of their background, ability, or status. This principle supports the creation of AI systems that reflect and serve the diverse global community, fostering innovation and trust by embracing different perspectives and needs.

Reliability and safety

Reliability and safety ensure that AI systems operate consistently and can be trusted to make decisions that protect human welfare. By prioritizing these principles, companies can build AI technologies that users can depend on for critical tasks, reducing the risk of harm and enhancing user confidence in adopting AI solutions across various domains.

Fairness

The inclusive approach to minimizing bias utilizes training datasets that accurately represent diverse groups. Conversely, factors such as gender, race, or ethnicity can be removed from data sets like resumes before making AI-powered decisions and recommendations. Some companies have been publicly named and shamed for favoring men over women in AI-powered recruitment.

Transparency

Companies can openly share information about how AI algorithms are designed, the data they use, and the decision-making processes they follow, enabling users to trust and verify the fairness and accuracy of these systems. This helps to identify and mitigate biases, understand the rationale behind AI decisions, and safeguard against unethical practices. Emphasizing transparency in AI fosters an environment of trust, encourages responsible innovation, and supports the ethical integration of AI technologies into society, ensuring they serve the public interest.

Privacy and security

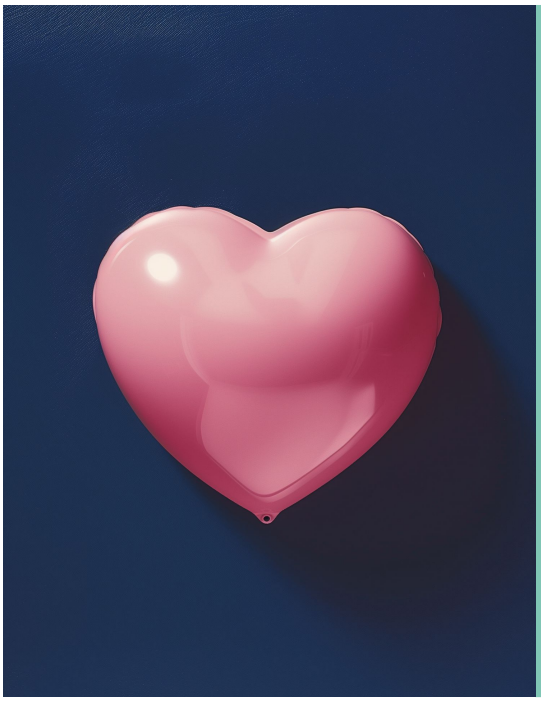
To reduce security risks for AI, companies implement robust data encryption, conduct regular security audits, and use secure AI training environments. Additionally, employing anomaly detection systems to monitor AI behavior and investing in AI-specific cybersecurity measures can mitigate potential threats. Continuously updating AI systems to guard against emerging vulnerabilities and fostering a culture of security awareness among development teams can help greatly.

Facial recognition shouldn't be dystopian.

CompreFace is an advanced, open-source facial recognition AI tool designed for seamless integration into any system. It has all essential features for facial recognition, including face identification and verification, alongside age and gender detection, and landmark detection capabilities.

The development of CompreFace was guided by ethical AI practices to minimize bias and prevent misuse. The fairness and impartiality of the system were rigorously tested using datasets encompassing a wide range of ethnicities and demographics. These datasets are crucial for ensuring the system accurately recognizes individuals from diverse backgrounds without discrimination.

Learn More



AI is diagnosing disease.

Exadel's Dr. Karol Przystalski helps avoid invasive procedures by detecting blocked arteries with computer vision.

Read More

Every company needs ethical AI, but how hard is it to get?

Data Quality and Accessibility Challenges

Training effective algorithms for business solutions requires substantial data. However, acquiring such data is particularly challenging in regulated industries like healthcare, biotech, and financial services due to stringent regulations like HIPAA and GDPR. Moreover, publicly available datasets may suffer from biases or incompleteness, compromising algorithm accuracy. Companies can get around this by generating high-quality synthetic data.

High Implementation Costs

There is a great financial burden adopting AI, which includes hardware and software expenses, alongside costs related to labor, training, and ongoing maintenance. Despite predictions that global spending on AI systems could hit $300 billion within three years, only 11% of companies report significant returns on their AI investments.

Shortage of Skilled Professionals

Crafting and deploying AI solutions demands a team of specialists, including data scientists, machine learning engineers, and software developers. The competition for such talent is fierce, leading many organizations to consider partnerships with technology firms to fill the gap.

Challenges in System Integration

On average, organizations utilize around 400 different data sources for analytics and business intelligence. Integrating AI across such a diverse technological landscape presents significant hurdles, as existing IT systems often feature disparate technologies and architectural frameworks.

The ethical AI checklist.

Ensure your AI programs are impactful and ethical.

Align Your Business Goals

Each of the outcomes mentioned above can be implemented ethically with the right questions:

- What are the ethical risks associated with the initiative?
- What measures do you have in place to mitigate the risks?

Establish Ethical AI Guidelines

A good place to start is with the draft recommendations of the European Union's ethical AI guidelines. These broadly state that AI systems should be accountable, explainable and unbiased. Once ratified, the EU guidelines will also include penalties for failure to adhere to its stipulations.

Explore More Training Data

Adding synthetic data to data sets helps get to the root causes of AI bias.

Adopt an Ethical AI Solution

Technology tools such as IBM Fairness 360 and Google's What-If tool, help assess whether an AI system can be trusted.

Educate Tech Teams on Ethical AI

Once trained up in the principles and themes of Ethical AI Guidelines, teams will be able to promptly flag any violations.

Ask the Right Questions About Your Current Deployments

- Are you putting people first?
- Are you still in step with your organization's core values?
- Do you have a plan to keep your AI systems ethical as they are being developed further?

Turn your data into rapid growth with AI

Read More