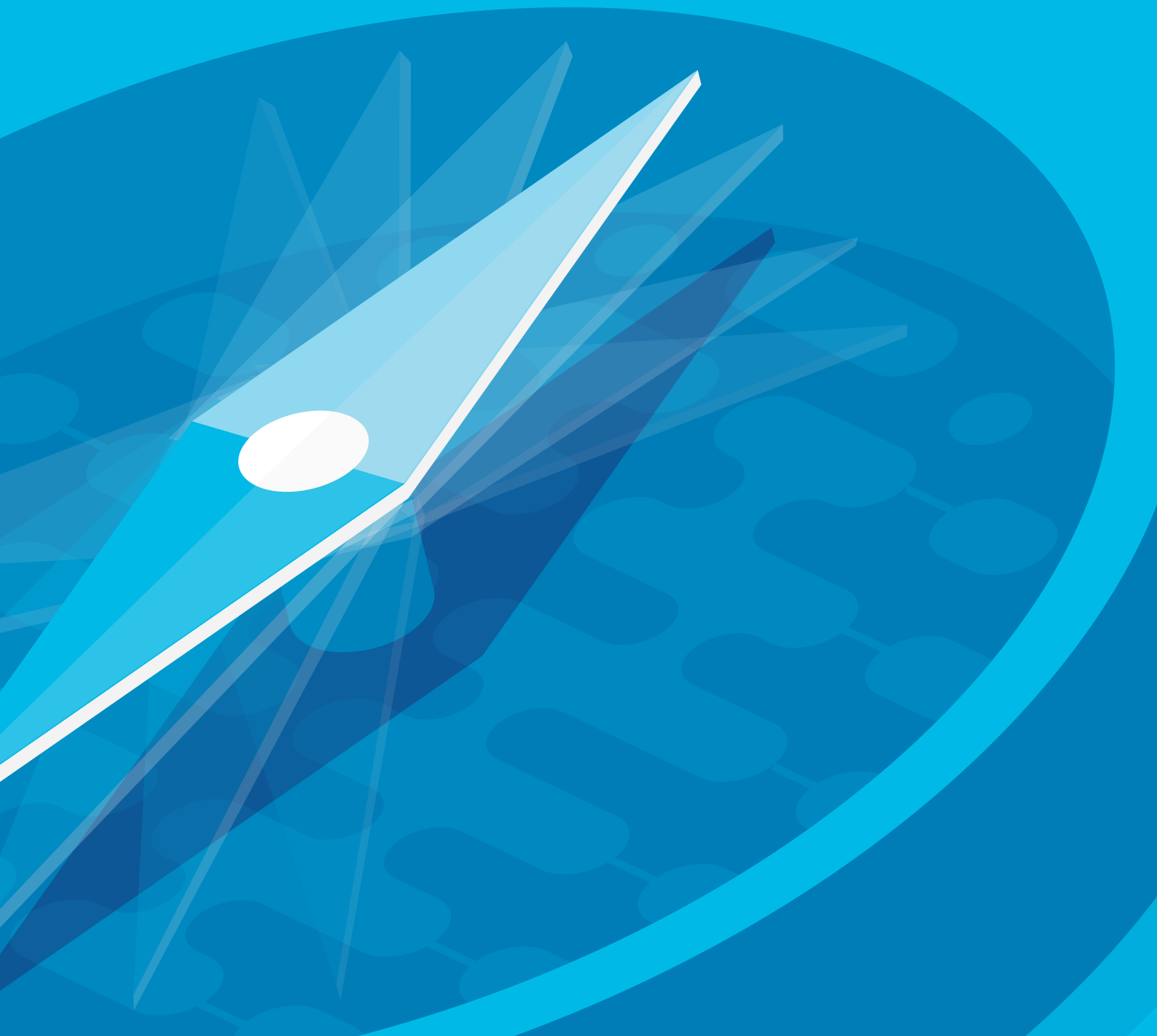
Capgemini
RESEARCH INSTITUTE

Why addressing **ethical** questions in **AI** will benefit organizations



Executive Summary

Every wave of technology development raises its own set of ethical questions, from unintended and harmful consequences for users to concerns about technologies being weaponized. The pace of technological advancement tends to outstrip the pace of regulatory and ethical frameworks – a position that AI finds itself in today. AI has unleashed a range of ethical questions, from concerns over autonomous vehicles to what constitutes end-user consent.

To probe these issues, we have conducted a comprehensive survey of both business executives and individuals to understand ethics and the transparency of AI-enabled interactions and what organizations are doing to address any concerns. We surveyed 1,580 executives in 510 organizations and over 4,400 consumers across countries such as the US, UK, China, Germany, and France. We found that:

Getting ethics in AI right will benefit organizations and mitigate risks

- **AI interactions that consumers and citizens perceive as ethical build trust and satisfaction:** When consumers and citizens feel an organization offers ethical AI interactions, over half said that they would place higher trust in it, share their positive experience, be more loyal, purchase more, and advocate for it. Organizations whose AI systems consumers see as interacting ethically enjoy a 44-point Net Promoter Score (NPS®)# advantage compared to those seen as not using AI ethically.
- **AI interactions perceived as unethical can damage brand reputation:** Nearly two in five consumers would complain to the company and demand an explanation if they experienced an unethical interaction. In the worst case, a third would stop interacting with that company.

Most organizations have encountered ethical issues in AI over the last two to three years. Consumers seem to corroborate this finding

- **Most executives (77%) are uncertain about the ethics and the transparency of their AI systems**
- **Executives in nine out of ten organizations are aware of at least one instance of the use of AI systems that resulted in ethical issues** (see insert below for examples of ethical issues)
- **Close to half of consumers believe they have felt the impact of an ethical issue** (to probe this matter, we gave survey respondents specific instances of unethical practices detailed in the appendix at the end of the report). This is leading a majority of consumers (76%) to advocate regulations on the use of AI: they expect governments, independent bodies (e.g. IEEE) and regulators to lay down the principles that define the just use of AI
- **The pressure to implement AI is fueling ethical issues:** Executives identify the pressure to urgently implement AI as the top reason why ethical issues arise from the use of AI**.**

Executives are starting to realize the importance of ethical AI and are taking action when ethical issues are raised

- 51% of executives believe that it is important to ensure that AI systems are ethical and transparent
- 41% of senior executives report that they have abandoned an AI system altogether when ethics concerns were raised; 55% implemented a "watered-down" version of the system.

First steps to proactively address ethics in AI

Our previous research has established that AI has significant potential across industries and functions.[1] **By addressing ethics issues upfront, we believe that organizations stand to gain additional benefits** as well as avoid regulatory, legal, and financial risks that may result from a market or public backlash on AI. We have developed a comprehensive approach to help organizations proactively address ethics questions in AI. This approach targets three key groups in the organization:

- General management (CXOs, and business leaders and those who have questions of ethics and trust in their remit), who will be responsible for laying the foundational practices and processes for ethical AI
- Customer-and employee-facing teams (e.g., HR, marketing, communications, and customer service) that are responsible for deploying AI ethically for users, and
- AI, data, operations, and IT teams that design, develop, and implement AI systems

For CXOs, business leaders and those with a remit for trust and ethics

- Lay down a strong foundation with a strategy and code of conduct for ethical AI
- Develop policies that define acceptable practices for the workforce
- Build awareness of ethical issues across the organization
- Create ethics governance structures and ensure accountability for AI systems
- Build diverse teams to inculcate sensitivity to ethical issues

For customer-and employee-facing teams such as HR and marketing

- Ensure that ethics questions are debated from the design of AI-infused applications by clearly defining the outcomes and intent of these applications, and ensuring accountability for them
- Educate and inform users on what building trust in AI applications means
- Empower users with more control and ability to seek recourse

For AI, data, and IT teams

- Making AI systems transparent and understandable
- Practice good data management in compliance with regulations where they exist, and address and monitor potential biases in data
- Use technology tools to actually build ethics in AI (bias detection, transparency, explainability, continuous monitoring of accuracy)

What kinds of ethical issues emerge from the use of AI?

By ethical issues arising from the use of AI, we mean the interactions that result in outcomes that are unexplainable, unfair, not transparent and/or biased against a certain group of users. Examples of AI interactions that may result in ethical issues include:

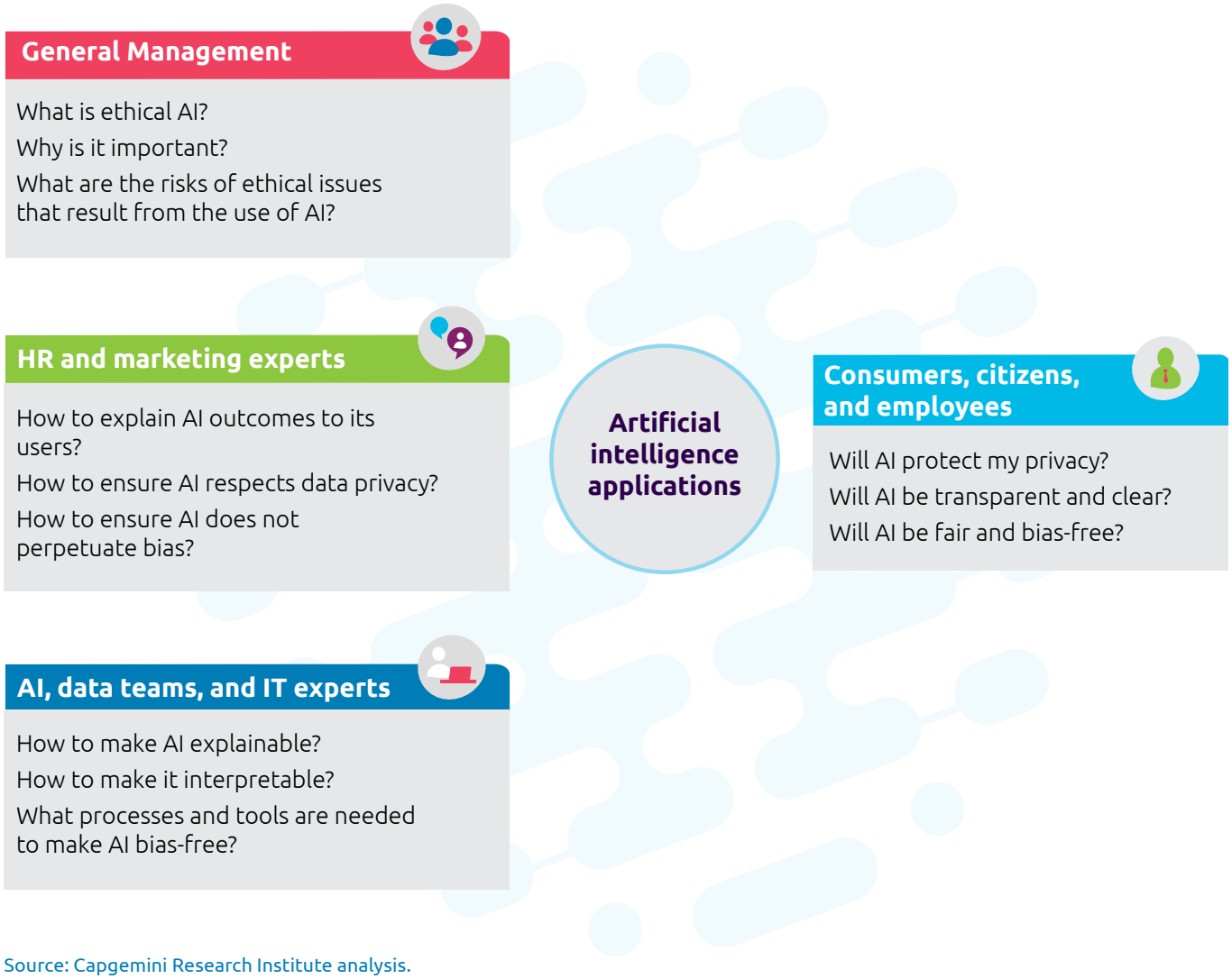
- Using AI to screen job applicants that results in a disproportionate selection of candidates across gender, ethnicity, age, or other factors
- Using AI for surveillance at the workplace without the consent of employees
- Using an AI system that is not able to explain how it arrived at a decision to deny credit or an insurance claim – or cannot be audited on its decision making
- Collecting and processing healthcare personal data in AI algorithms without consent
- Discriminatory pricing or availability of products/ services resulting from automated AI decisions.

#Net Promoter, NPS®, and the NPS®-related emoticons are registered service marks, and Net Promoter Score and Net Promoter System are service marks, of Bain & Company, Inc., Satmetrix Systems, Inc. and Fred Reichheld.

Introduction

Artificial intelligence offers a huge opportunity for businesses and the economy, but significant questions are being raised about the ethical issues surrounding this technology. To examine these questions more closely, we have undertaken this research to understand the current relationship between AI applications and their users. As Figure 1 shows, we examined both sides of the debate – from a business and end-user perspective.

Figure 1. Scope of our research on ethics in AI



Source: Capgemini Research Institute analysis.

We surveyed over 1,500 industry professionals from 500 organizations; over 4,400 consumers; and conducted in-depth interviews with over 20 industry executives, academics, and start-up entrepreneurs (see the research methodology at the end of the paper for more details).

Through this paper, we want to demonstrate:

1. Why it is important to pursue ethics in AI from a business perspective. Our analysis shows that individuals (both consumers and employees) trust organizations that they perceive to be using AI ethically and are willing to advocate for them.
2. Why most organizations have encountered ethical issues in AI over the last two to three years.
3. How organizations can start to address ethics in AI more proactively. We therefore identified clear and actionable first steps while designing, developing, and using AI applications.

What is AI?

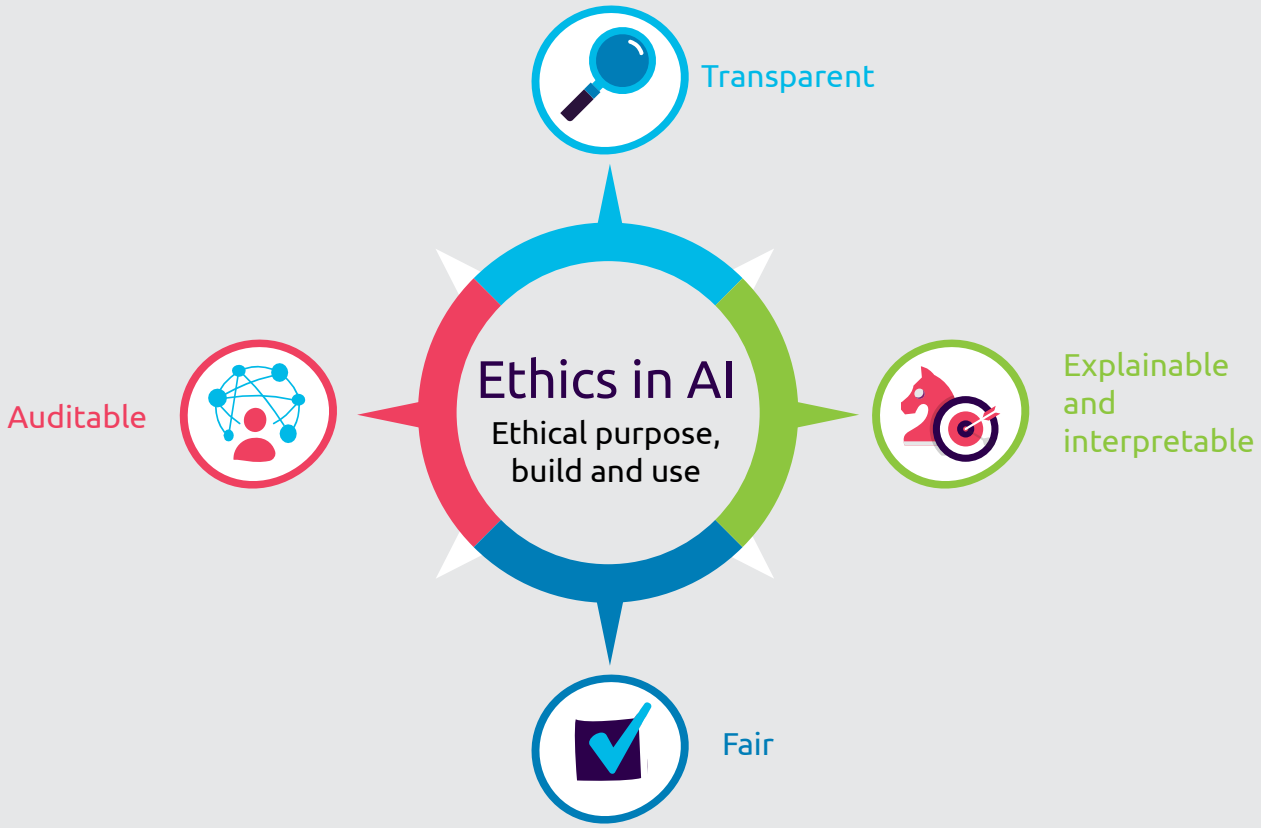
Artificial intelligence (AI) is a collective term for the capabilities shown by learning systems that are perceived by humans as representing intelligence. Today, typical AI capabilities include speech, image and video recognition, autonomous objects, natural language processing, conversational agents, prescriptive modeling, augmented creativity, smart automation, advanced simulation, as well as complex analytics and predictions.

What do we mean by ethics in AI?

According to the European Commission, the ethics of AI is a sub-field of applied ethics and technology that focuses on the ethical issues raised by the design, development, implementation, and use of AI.[2]

Key components of ethical AI include:

- Being ethical in its purpose, design, development, and use
- Transparent AI: AI where it is clear, consistent, and understandable in how it works
- Explainable AI: AI where you can explain how it works in language people can understand
- Interpretable AI: AI where people can see how its results can vary with changing inputs
- Fair AI: AI that eliminates or reduces the impact of bias against certain users
- Auditable AI: AI that can be audited, allowing third-parties to assess data inputs and provide assurance that the outputs can be trusted.



Source: Capgemini Research Institute.

Getting ethics in AI right will benefit organizations

"Trust is something very difficult to gain and very easy to lose. But a classic way of gaining trust, with AI interactions in particular, can be summarized in three words: transparency, accountability, and empowerment. That means transparency so that people can see what you are doing; accountability because you take responsibility for what you are doing; and empowerment because you put people in charge to tell you if something you did was not right or not good."

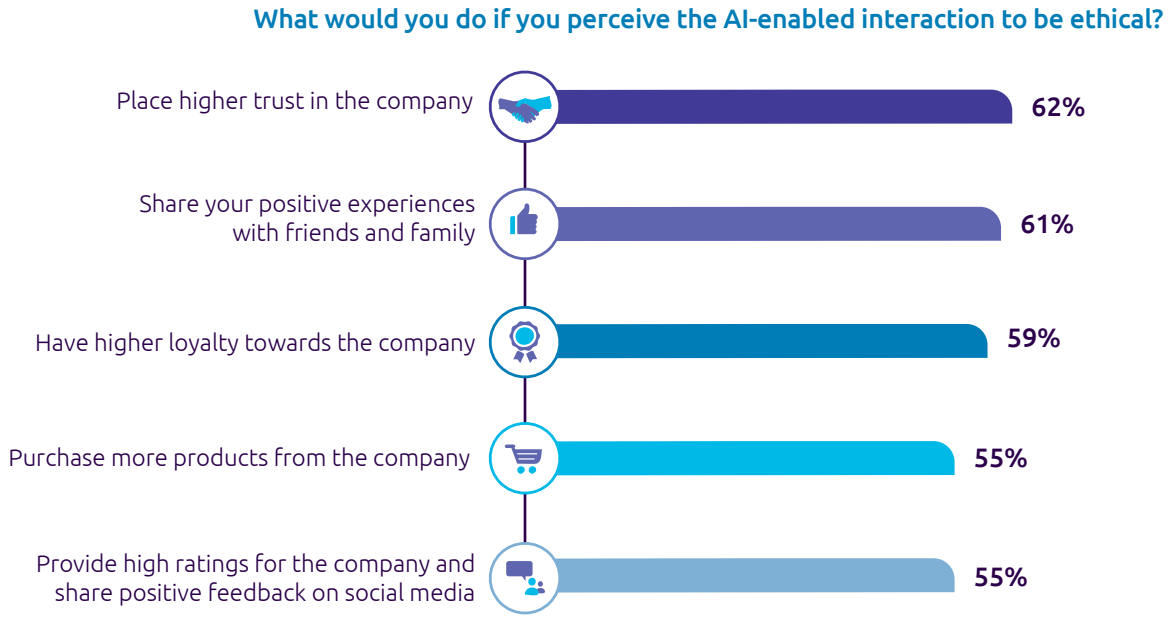
- Luciano Floridi, professor of Philosophy and Ethics of Information and director of Digital Ethics Lab, Oxford Internet Institute, University of Oxford[3]

Our previous research on AI's role in the customer experience established that nearly three-quarters of consumers say that they are aware of having interactions enabled by AI.[4] They see great benefits in these interactions – greater control, 24/7 availability, and convenience. With our current research, we see that organizations can build on these benefits if consumers perceive AI interactions to be ethical.

Ethical AI interactions drive consumer trust and satisfaction

Ethical AI interactions earn consumer trust and build satisfaction. Three in five consumers who perceive their AI interactions to be ethical[5] place higher trust in the company, spread positive word of mouth, and are more loyal (see Figure 2). A positive perception can also have a tangible impact on the top line as well. Over half of the consumers we surveyed said that they would purchase more from a company whose AI interactions are deemed ethical.

Figure 2. Consumers become advocates when they perceive AI interactions as ethical



Source: Capgemini Research Institute, Ethics in AI consumer survey, N = 4,447 consumers.

62% *of consumers will place higher trust in the company if they perceive AI-enabled interactions as ethical*

The Net Promoter Score (NPS®) provides a measure of the positive impact of ethics in AI. Our research found that organizations that are seen as using AI ethically have a 44-point NPS® advantage over those seen as not.

AI interactions that are perceived as unethical can harm business and damage brand reputation

While there is an ethical responsibility on teams to put in place the checks and balances for fairness and transparency, there is also a significant business case for it. *"If a data science team is working on a machine learning project that will be affecting humans, I think that they have both ethical and commercial responsibility to do basic disparate impact analysis,"* says Patrick Hall, senior director of Product, H20.ai, an open source machine learning and artificial intelligence platform. *"Even if you personally don't care about fairness – and you think it's some kind of far-out liberal cause – you'll do financial and reputational harm to your employer if your model is discriminatory."*

If consumers decide to take action when their AI interaction results in ethical issues, (see insert in executive summary for examples of ethical issues in AI) as Figure 3 shows, nearly two in five consumers would complain to the company and demand explanations, a third of them (34%) can even stop interacting with the company – potentially causing loss of business and negative word of mouth.

Figure 3. AI interactions resulting in ethical issues can backfire on organizations

Source: Capgemini Research Institute, Ethics in AI consumer survey, N = 4,447 consumers.

Most organizations have encountered ethical issues in AI over the last 2-3 years

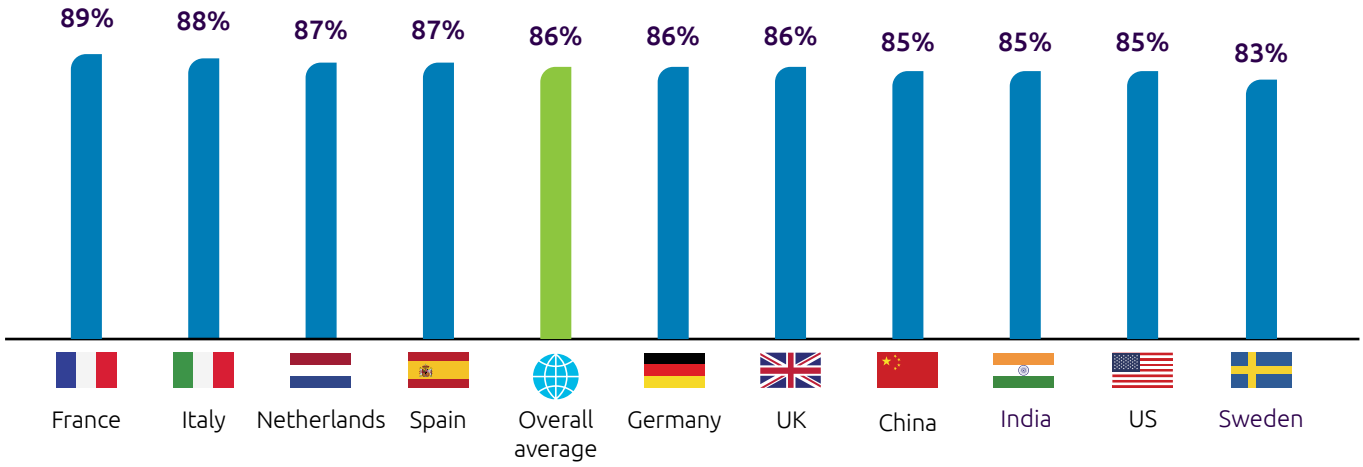
Executives in nine out of ten organizations believe that ethical issues have resulted from the use of AI systems over the last 2-3 years

Our research shows that executives have witnessed at least one instance of the use of AI systems where ethical issues have resulted and close to half of consumers say they have experienced the impact of an ethical issue:

- 86% of executives say they are aware of instances where AI has resulted in ethical issues
- 47% of consumers say they have experienced the impact of an ethical issue
- 77% of executives are uncertain about the ethics and the transparency of their AI systems.

Figure 4. Nearly nine in ten organizations across countries have encountered ethical issues resulting from the use of AI

In the last 2-3 years, have the below issues resulting from the use and implementation of AI systems, been brought to your attention? (percentage of executives, by country)



We presented over 40 cases where ethical issues could arise from the use of AI, to executives across sectors. We asked them whether they encountered these issues in the last 2-3 years.
Source: Capgemini Research Institute, Ethics in AI executive survey, N = 1,580 executives, 510 organizations.

Many of these issues are not intentional but are the result of not having the right checks and balances during the development and deployment of these systems. According to the chief digital officer of a large European consumer products firm we interviewed, organizations currently lack the processes to check for unintended impact of using AI: *"More often than not, AI bias does not come from the people who program the algorithm – they have honest intentions. I think that the bias often comes from the data you feed into the system. This is because the data is basically historical data and historical data is not devoid of bias – it just shows you what consumers did in the past. Therefore, it is best to not fully depend on a historical view of data, but also factor in the socio-economic context."*

As Table 1 shows, the most common issue from executives' perspective is in the banking sector, where they think that banks use machine-led decisions without disclosure.

Table 1. Top ten ethical issues resulting from use of AI: by executive awareness

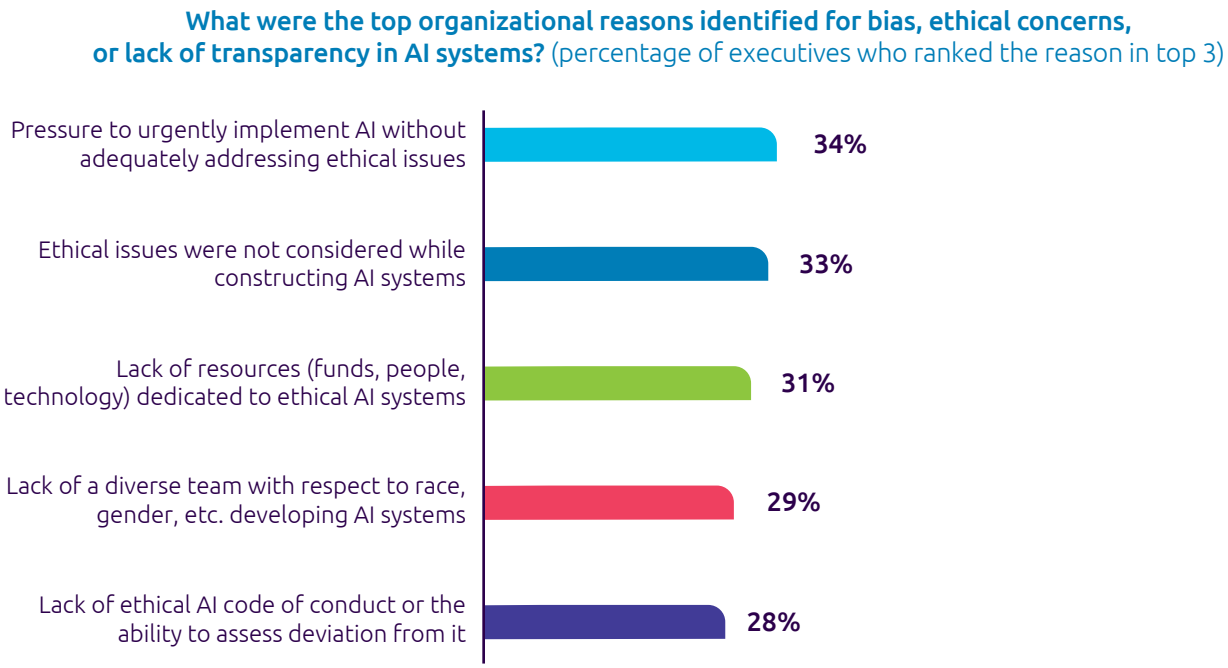
	Top ten issues across all sectors (in decreasing order of share of executives who have encountered these issues)
1	Over reliance on machine-led decisions without disclosure in the banking sector
2	Collecting and processing patients' personal data in AI algorithms without consent
3	Biased/unclear recommendations from an AI-based system for diagnosis/care/treatment
4	Over reliance on machine-led decisions without disclosure in the insurance sector
5	Citizens objecting to use of facial recognition technology by police force for mass surveillance
6	Limiting access/discriminatory pricing of services/products due to consumers' race/gender, etc. (originally part of the company's target group)
7	Discriminatory pricing of insurance policies due to consumers' demographic profiles
8	Processing patients' personal data in AI algorithms for purposes other than for which it was collected
9	Customers demanding reasoning/clarity behind a decision taken by an AI algorithm to deny credit
10	Citizens objecting to the collection and use of their personal data such as biometrics by an AI system

Source: Capgemini Research Institute, Ethics in AI executive and consumer survey, N = 1,580 executives, 510 organizations; and 4,447 consumers; sample size for each sector is different – banks: 124, healthcare: 54, insurance: 127, public sector: 74.

The pressure to implement AI is fueling ethical issues

When we asked executives why ethical issues resulting from AI are an increasing problem, the top-ranked reason was the pressure to implement AI (see Figure 5). This pressure could stem from the urgency to gain a first-mover advantage, acquiring an edge over competitors in an innovative application of AI, or simply to harness benefits that AI has to offer.

Figure 5: Top reasons behind ethical issues in AI



Source: Capgemini Research Institute, Ethics in AI executive survey, N = 1,580 executives, 510 organizations.

34% *of executives identified pressure to urgently implement AI without adequately addressing ethical issues as one of the top organizational reasons for bias, ethical concerns, or lack of transparency in AI systems*

We probed the second-ranked reason, that ethical issues were not considered when building AI, in a separate question. The finding is consistent – about one in three organizations (37%) report to focus significant attention on ethical issues when implementing AI systems, and only about four in ten organizations (44%) are prepared to mitigate ethics issues in AI.

Table 2: Top organizational reasons identified for bias, ethical concerns, or lack of transparency in AI systems – by function (in decreasing order of importance)

	 General management and ethics professionals	 HR and marketing professionals	 AI, data, and IT professionals
1	Pressure to urgently implement AI without adequately addressing ethical issues	Pressure to urgently implement AI without adequately addressing ethical issues	Lack of ethical AI code of conduct or ability to assess deviation from it
2	Ethical issues were not considered while constructing AI systems	Lack of resources (funds, people, technology) dedicated to ethical AI systems	Lack of relevant training for developers building AI systems
3	Lack of a diverse team with respect to race, gender, etc. developing AI systems	Lack of relevant training to developers building AI systems	Ethical issues were not considered when constructing AI systems
4	Lack of ethical AI code of conduct or a deviation from it	Ethical issues were never considered when constructing AI systems	Pressure to urgently implement AI without adequately addressing ethical issues
5	Lack of relevant training for developers building AI systems	Lack of ethical AI code of conduct or to assess a deviation from it	Lack of resources (funds, people, technology) dedicated to ethical AI systems

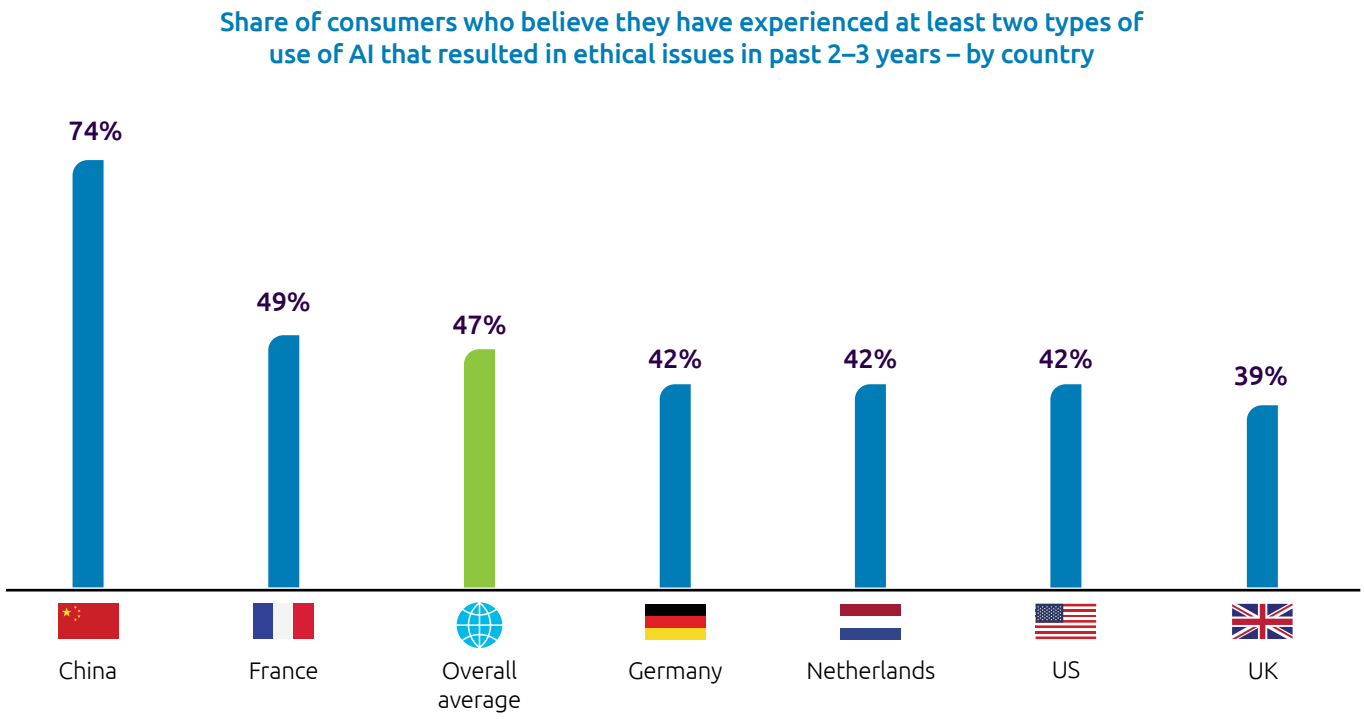
Source: Capgemini Research Institute, Ethics in AI executive survey, N = 1,580 executives, 510 organizations.

47% *of consumers believe they have experienced at least two types of use of AI that resulted in ethical issues in past 2–3 years*

Close to half of consumers feel they have been exposed to ethical issues related to AI

As we have seen, close to half of consumers (47%) feel they have been exposed to more than two instances of ethical concerns resulting from the use of AI in the last two to three years (see Figure 6). To probe this matter, we gave survey respondents specific instances of unethical practices, e.g. reliance on machine-led decisions without disclosure (see appendix at the end of the report for more details).

Figure 6. Close to half of the consumers believe to have been exposed to some use of AI that resulted in ethical issues

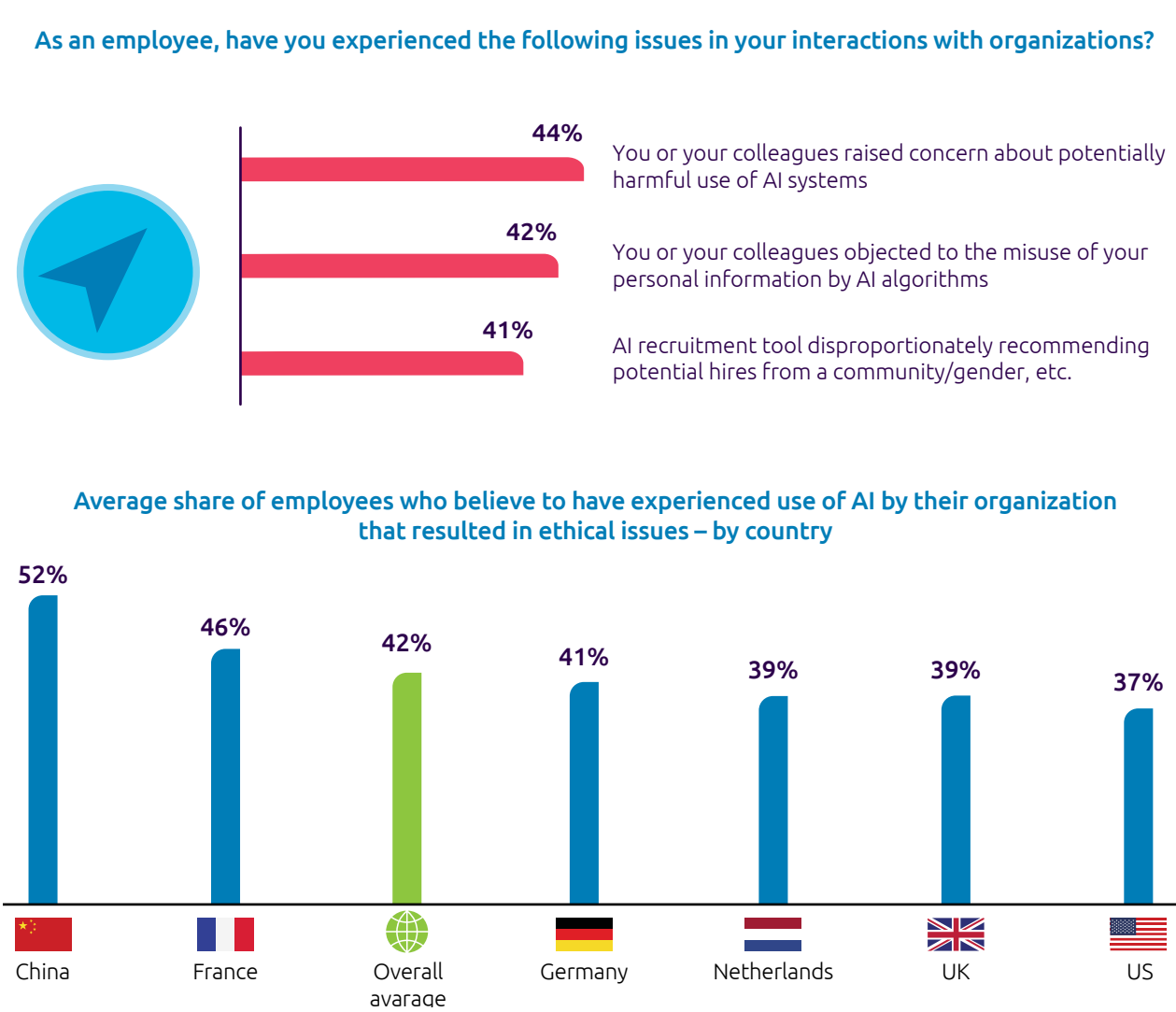


Source: Capgemini Research Institute, Ethics in AI consumer survey, N = 4,447 consumers.

Two in five employees have experienced ethical issues themselves or seen it with the general public

At least 40% of employees have come across some form of AI use that resulted in ethical issues (see Figure 7). Not only are employees aware about ethical issues in AI, they are also raising concerns about the use of such systems. As we found in our research, 44% of employees have raised concerns about the potentially harmful use of AI systems and 42% of employees have objected to the misuse of personal information by the AI systems.

Figure 7. Employees are raising concerns about potential ethical issues resulting from the use of AI systems



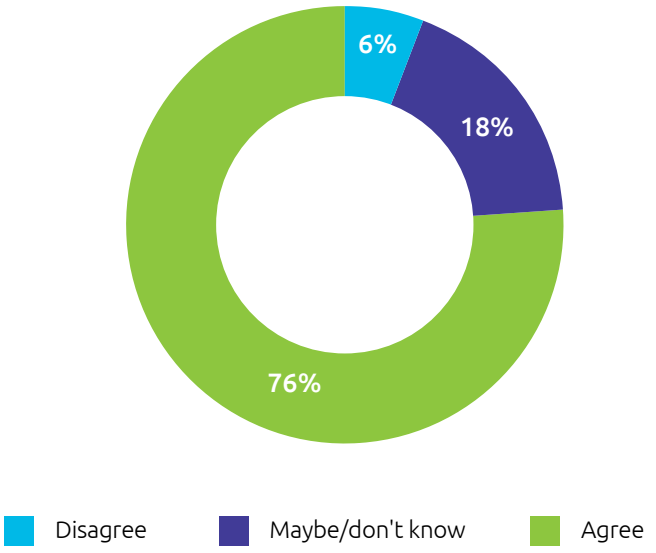
Source: Capgemini Research Institute, Ethics in AI consumer survey, N = 3,288 employees.

Consumers want regulation on the use of AI

Given that many consumers feel they have experienced ethical issues, it is not surprising that over three-quarters (76%) expect new regulations on the use of AI (see Figure 8). This is fueled in part by rising awareness of ethical issues in AI, as well as the positive perception of recent data privacy regulations, such as GDPR. *"Today, we don't really have a way of evaluating the ethical impact of an AI product or service,"* says Marija Slavkovik, associate professor at University of Bregen. *"But it doesn't mean that regulation or law will not catch up with bad actors eventually. Organizations or individuals may not be inherently evil, but we as a society need to develop a way in which we can systematically evaluate the ethical impact of AI to account for negligence."*

Figure 8. Consumers want regulations on the use of AI

Do you think there should be a new law or regulation to regulate the use of AI by organizations?
(percentage of consumers)



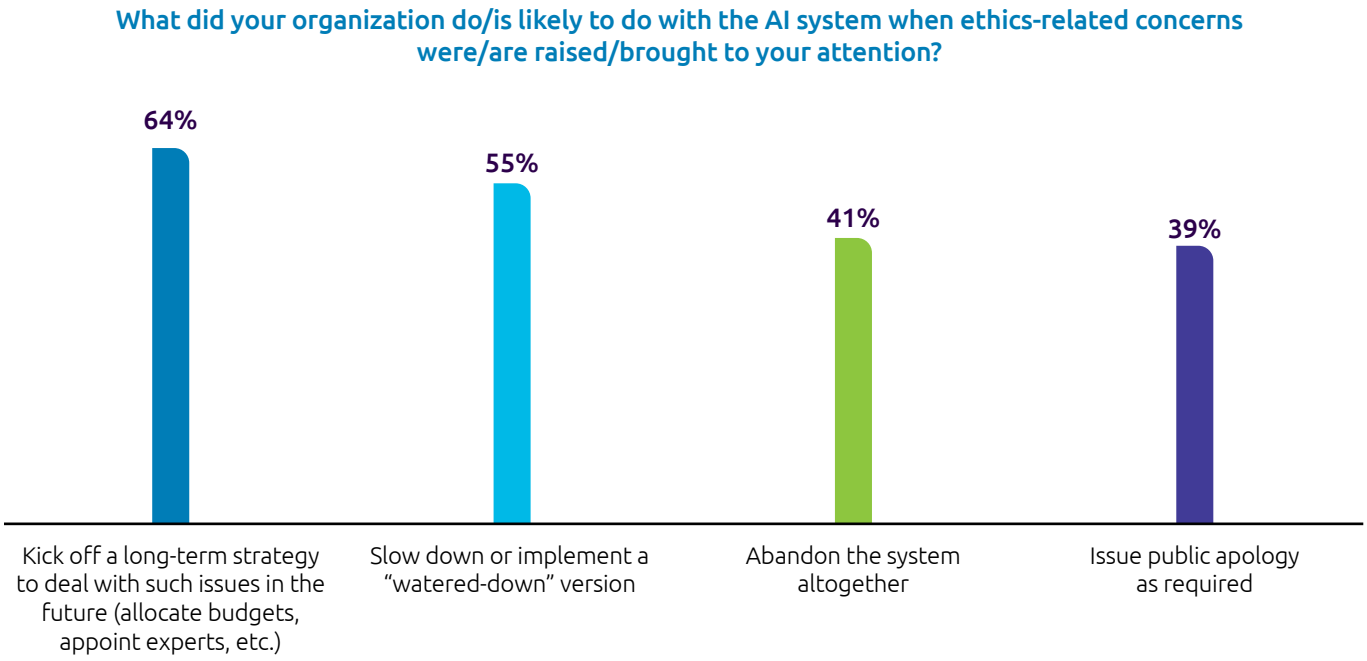
Source: Capgemini Research Institute, Ethics in AI consumer survey, N = 4,447 consumers.

76% *of consumers expect new regulations on the use of AI*

Executives are starting to realize the importance of ethical AI and are taking action when ethical issues are raised

Our research shows that 51% of executives consider that it is important to ensure that AI systems are ethical and transparent. Organizations are also taking concrete actions when ethical issues are raised. As Figure 9 shows, more than two in five executives report to have abandoned an AI system altogether when an ethical issue had been raised.

Figure 9. When ethical issues are raised, organizations are taking action to address the concerns



Source: Capgemini Research Institute, Ethics in AI executive survey, N = 1,580 executives, 510 organizations.

51% *of executives consider that it is important to ensure that AI systems are ethical and transparent*

First steps to proactively addressing ethics in AI

Ethics is a challenging and complex topic that does not stand still, and requires that organizations constantly update their approach. As a start point, organizations need to build a long-term strategy that involves all departments focused on AI planning, development, and deployment, with the goal of building ethics-by-design principles into their AI systems. This requires:

- Understanding the ethical vision of the company and trade-offs against other aspects of company strategy.
- Putting in place policy and governance to give these goals weight within the organization.
- Operationalizing the strategy across the systems, people and processes involved in developing and using AI systems.

Scotiabank, for example, has set a vision for its interactive AI systems – they need to improve outcomes for customers, society, and the bank. The bank also monitors systems for unacceptable outcomes and to ensure there is accountability for any mistakes, misuse, or unfair results.[6]

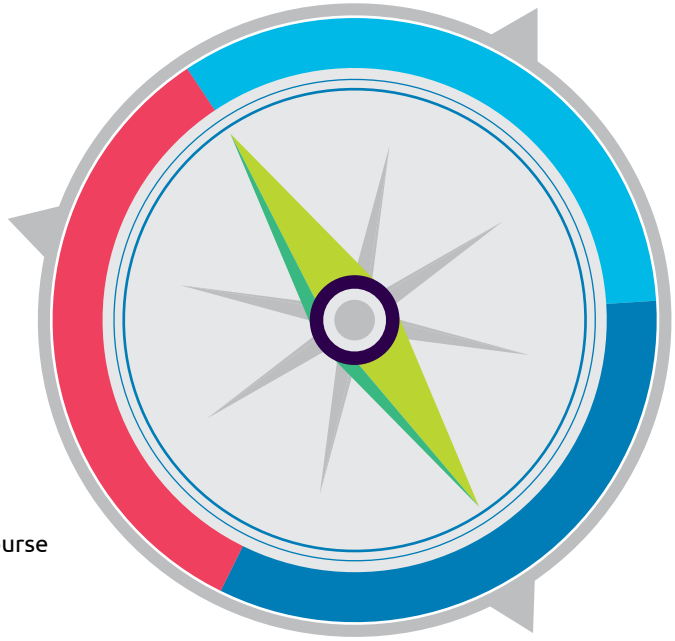
On the basis of our extensive primary research and our conversations with industry experts, startup executives and leading academics in this field, we suggest a three-pronged approach to building a strategy for ethics in AI that embraces all key stakeholders:

1. **General management:** CXOs, business leaders, and those with trust and ethics in their remit who will be responsible for laying the foundational practices and processes for ethical AI, including defining an ethical purpose for using AI
2. The customer and employee-facing teams such as **HR, marketing, communications, and customer service** – who are responsible for designing the finality and intent of the use of AI in their processes and tasks, and are accountable of deploying AI ethically for users
3. **AI, data, and IT** leaders and their teams, who will be responsible for the ethical technology design, development, deployment, and monitoring of AI systems (see Figure 10).

Figure 10. Formulating an ethical AI strategy

For CXOs, ethics, and business leaders

- Lay down a strong foundation with a strategy and code of conduct for ethical AI
- Develop policies that define acceptable practices for the workforce
- Build awareness of ethical issues across the organization
- Create ethics governance structures and ensure accountability for AI systems
- Build diverse teams to inculcate sensitivity to ethical issues.



For customer- and employee-facing teams such as HR and Marketing

- Ensure ethical usage of AI systems
- Educate and inform users to build trust in AI systems
- Empower users with more control and ability to seek recourse
- Proactively communicate on AI issues internally and externally to build trust.

For AI, data, and IT teams

- Make AI systems transparent and understandable to gain users' trust
- Practice good data management and mitigate potential biases in data
- Use technology tools to build ethics in AI

Source: Capgemini Research Institute analysis

1. For CXOs, business leaders, and those with a remit for trust and ethics

Establish a strong foundation with a strategy and code of conduct for ethical AI: Leadership teams need to start by developing a long-term strategy and code of conduct for ethical AI that gives the rest of the organization a roadmap to follow and boundaries to respect. This requires financial investment in building ethical AI – for example, through research and development, hiring external experts, investing in external events and committees and technological investments. But it also means giving departments sufficient time so that data and customer-facing teams do not feel rushed to deploy AI systems without checking the ethical implications of data, data models, algorithms, and the AI systems as a whole.

The code of conduct is a statement on AI ethics that recognizes the trade-offs involved and guides the organization on how to prioritize ethical questions against other business objectives.

"The first step is to define a process," says Nicolas Economou, co-chair, Law Committee of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, and CEO of H5, a Silicon Valley-based consultancy and technology firm. *"What does it mean to implement digital and AI ethics? Beyond what is legal or not –and therefore what you must comply with – you need to determine what you stand for as an organization – what are your brand values? These values should exist whether you use AI or not. You can then define your ethical AI code on the basis of what you stand for; what that implies for how you think of the impact of your decisions on your company, employees, customers, and society at large; and, as a result, determine what kind of AI practices you can deem conformant to your ethics."*

- Organizations can develop a code by drawing on widely acknowledged frameworks, such as the "Ethics Guidelines for Trustworthy AI," from the European Commission's High-Level Expert Group (HLEG) on Artificial Intelligence. "The EU HLEG has taken a first step towards the development of good normative rules around ethical AI that organizations can adapt." says Luciano Floridi, professor of Philosophy

"What does it mean to implement digital and AI ethics? Beyond what is legal or not – and therefore what you must comply with – you need to determine what you stand for as an organization – what are your brand values? These values should exist whether you use AI or not. You can then define your ethical AI code on the basis of what you stand for; what that implies for how you think of the impact of your decisions on your company, employees, customers, and society at large; and, as a result, determine what kind of AI practices you can deem conformant to your ethics."

- Nicolas Economou,
co-chair, Law Committee of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, and CEO of H5

and Ethics of Information and director of Digital Ethics Lab, Oxford Internet Institute, University of Oxford. In our research, only 16% of organizations were "highly influenced" by external benchmarks when designing or implementing ethical and transparent AI systems.

- To build a code of conduct, organizations also need to collaborate with others, including academic institutions and government and regulatory bodies. However, they also need to involve diverse stakeholder groups, encompassing employees, customers and the wider society within which they operate.
- Some organizations have already developed AI principles and ethical guidelines or are in the process of doing so (see Table 3). These organizations are also taking active steps to ensure that the code of conduct is translated into ground-level action. Telefónica, for example, which published its ethical guidelines for AI applications last year, says that it will assess all projects that include AI in accordance with its guidelines.[7] It will apply these principles as rules in all markets in which it operates, throughout its value chain, and across partners and providers. In addition, Standard Chartered Bank is developing a framework to ensure fairness, ethics, accountability and transparency (FEAT) in the Group's use of AI.[8]

Develop policies that define acceptable practices for the workforce and users of AI applications: An organization's ethical code of conduct must be translated into practice through policies that combine the goals of the ethics statement along with applicable regulations and industry best practice. These policies must define boundaries for the workforce, giving them a framework in which to operate and ensuring they know what is acceptable practice. For instance, a large, US-based healthcare organization that we spoke to has designed strict policies to restrict access to customer data, which is classified based on its sensitivity. As part of their approach, it has a rigorous approval process for when personally identifiable and protected customer health information can be accessed for use in AI applications. Those requesting access must have a strong justification for why they need to use a data set.

Build awareness of ethical questions across the organization: Organizations must build awareness across functions and organizational layers – of ethics, transparency, explainability, interpretability, and bias in AI systems. In particular, teams building and deploying AI systems need to be fully aware of these issues if they are to mitigate any ethical risks or weaknesses.

Danya Glabau, faculty member at the Brooklyn Institute for Social Research, points out that organizations will need to build their employees' skills and understanding in what will be a new field for many. *"Thinking about employee education, organizations need to rethink what skills, knowledge, and experiences they expect employees to bring in,"* she says. *"These skills may not follow the typical engineering or executive pathway. Organizations need to think how they can build the teaching of these skills into their employee learning programs, so that there are resources available for employees who are thinking about ethics in AI and other such issues."*

Create ethics governance structures and ensure accountability for AI systems: Leadership teams also need to create clear roles and structures, assign ethical AI accountability to key people and teams and empower them. Key steps can include:

- Adapting existing governance structures to build accountability within certain teams. For example, the existing ethics lead (e.g., the Chief Ethics Officer) in the organization could be entrusted with the responsibility of also looking into ethical issues in AI
- Creating new roles, such as AI ethicists – potentially with a background in business ethics, compliance, and also with an understanding of how that applies to AI – and AI leads who can be held accountable for good AI practices
- Assigning senior leaders who would be held accountable for ethical questions in AI
- Building internal/external committees responsible for deploying AI ethically, which are independent and therefore under no pressure to rush to AI deployment.

"Organizations need to ask the question as to how will they ensure that AI is accountable," says Ryan Budish, assistant director of Research at Berkman Klein Center for Internet & Society at Harvard University. *"One way to do it is through a top-down, system-wide approach where the organization thinks about the sort of standards needed to hold these systems accountable. These could be ethical standards or normative standards or political standards. There could be any number of perspectives that shape the standards used to hold these systems accountable. Such standards can also emerge bottom-up, in a more organic, iterative fashion."*

Build diverse teams to ensure sensitivity towards the full spectrum of ethical issues: To ensure that algorithms are bias-free, and AI systems are as ethical as possible, it is important to involve diverse teams. For example, organizations not only need to build more diverse data teams (in terms of gender or ethnicity), but also actively create inter-disciplinary teams of sociologists, behavioral scientists and UI/UX designers who can provide additional perspectives during AI design. *"I think this is very important because many of these systems will be implemented in different areas for people with different backgrounds,"* says Christoph Luetge, director of the TUM Institute for Ethics in Artificial Intelligence at Technical University of Munich.

"You cannot just assume that you are dealing with some specific group of people only. And, as we know from many famous examples, there might be a problem where systems, for example, assume that people only have a certain skin type. It's very important to involve people with many different backgrounds right from the start." GE Healthcare, for instance, has committed to employing a diverse workforce in teams working on AI. Eighty percent of their data science team have a minority background, 44% who sit outside the US, and 26% are women.[9] They also have mature data practices – all data is clearly sourced and how it can be used is clearly specified.

"You cannot just assume that you are dealing with some specific group of people only. And, as we know from many famous examples, there might be a problem where systems, for example, assume that people only have a certain skin type. It's very important to involve people with many different backgrounds right from the start."

- Christoph Luetge,
director of the TUM Institute for Ethics in Artificial Intelligence at Technical University of Munich

2. For customer- and employee-facing teams such as HR and Marketing

Ensure ethical usage of AI systems: Customer- and employee-facing teams, such as marketing, communications and HR, must ensure AI systems are transparent, explainable, and free of bias for end users.

Working in collaboration with the AI, data and IT teams, they must be empowered, from the first design of the AI application, to define finality and intent of an AI application very clearly, and the corresponding transparency towards end users. The finality and intent as defined by business users would serve as the cornerstone for the proper design, development, and testing phase of the AI application, including any possible impact on users.

End-user testing, drawing on a small set of pilot users, can help weed out adverse effects before the system goes into operation. Up-front action can ensure that AI systems do not result in major ethical problems later. For example:

1. Human-resources professionals must put in place measures for AI systems used in the following areas:
   - Recruitment: that they are fair and non-discriminatory in selecting potential hires from all communities, gender, age, or race/ethnicity
   - Performance reviews: that systems show no bias towards any particular community, gender, and race/ethnicity
   - Workplace surveillance: that systems have the full consent of employees and their purpose is made clear to employees before deployment
   - Collecting and processing employee data: that systems operate with the consent of employees

2. Marketing and communication professionals must ensure that AI applications:
   - Explain their workings and any outcomes when end-users request that explanation
   - Inform end-users that they are interacting with human-like chatbots and not humans
   - Are not limiting access to, or not pricing services/products differently, because of the demographics of customers
   - Do not use biased or sexist words or phrases (e.g., by chatbots/voice bots)
   - Make legally compliant use of personally identifiable data, such as medical records and biometrics, and ensuring that people understand why the data is being collected and how it will be used.

Educate and inform users to build trust in AI systems: The customer- and employee-facing teams have the task of building user trust. Organizations can build trust with consumers by communicating certain principles: the use of AI

for good, valuing human autonomy, and respecting the end consumer's rights. Organizations should also seek to inform users every time they are interacting with AI systems. In other words, the system should not pose as a human when interacting with users. They should also educate users about potential misuse and the impact and risks of using AI systems in an environment where ethical questions are not addressed proactively.

Empower users with more control and the ability to seek recourse: Organizations should empower individuals with the ability to access, share, and seek clarity on the use of their data. This means building policies and processes where users can ask for explanations of AI-based decisions. Consumers want more transparency to when a service is powered by AI (75% in our survey) and know if an AI is treating them fairly (73%).

Effective feedback channels for users will also help them to reach out to organizations for:

- sharing feedback or grievances
- reviewing AI decisions
- seeking explanations.

All of these will help in building greater consumer trust. *"Organizations deploying AI systems are accountable for not abusing the information that they have or the trust that they have been given,"* says Marija Slavkovik, associate professor at University of Bergen. *"If that happens, users should have channels to raise alarm and communicate that they are not happy. For example, consider a user who struggles with infertility that keeps getting advertisements for diapers. This could be a problem for the user, and she should have instruments to object to being in this target group. Users have the right to demand this and organizations have a responsibility to provide such channels."*

Proactively communicate on AI issues internally and externally to build trust: Communication is a bedrock of transparency and trust and can be a useful vehicle to build trust among all stakeholders, especially consumers, employees, and citizens. The communications and marketing team must adopt a well defined strategy to drive internal and external communications to convey the concrete steps the organization is taking to build ethics and transparency into AI applications. Currently organizations adopting AI across their business operations consider it a competitive advantage to be "early movers." To maintain this competitive edge, organizations often do not necessarily communicate openly on their use of AI and its impact on end users or employees. We consider that "smart movers" could turn being transparent in their use of AI into a competitive advantage by thinking "people-first," and getting end users on their side by being as transparent as possible about their use of AI.

"Organizations deploying AI systems are accountable for not abusing the information that they have or the trust that they have been given. If that happens, users should have channels to raise alarm and communicate that they are not happy. For example, consider a user who struggles with infertility that keeps getting advertisements for diapers. This could be a problem for the user, and she should have instruments to object to being in this target group. Users have the right to demand this and organizations have a responsibility to provide such channels."

- Marija Slavkovik, associate professor at University of Bergen.

3. For AI, data, and IT teams

Make AI systems transparent and understandable to gain users' trust: Systems need to be transparent and intuitive for users and business teams. The teams developing the systems should provide the documentation and information to explain, in simple terms, how certain AI-based decisions are reached and how they affect an individual. These teams also need to document processes for data sets as well as the decision-making systems. When we asked consumers what long-term actions would convince them that companies are using AI ethically, close to eight out of ten opted for "Providing explanations for AI decisions in case I request it." Close to an equal number opted for "Informing me about the ways in which AI decisions might affect me."

Practice good data management and mitigate potential biases in data: While general management will be responsible for setting good data management practices, it falls on the data engineering and data science and AI teams to ensure those practices are followed through. These teams should incorporate "privacy-by-design" principles in the design and build phase and ensure robustness, repeatability, and auditability of the entire data cycle (raw data, training data, test data, etc.). The AI practitioners need to:

- ensure that data is sourced ethically and in line with what regulation permits
- check for accuracy, quality, robustness, and potential biases, including detection of under-represented minorities or events/patterns
- build adequate data labelling practices and review periodically
- store responsibly, so that it is made available for audits and repeatability assessments
- constantly monitor results produced by models as well as precision and accuracy, and test for biases or accuracy degradation.

Good data management must also involve creating checks and balances to mitigate AI bias. Teams need to particularly focus on ensuring that existing datasets do not create or reinforce existing biases. For example:

- Identifying existing biases in the dataset through use of existing AI tools or through specific checks in statistical patterns of datasets
- Being mindful of not creating a selection bias on the data when developing algorithms
- Exploring and deploying systems to check for and correct existing biases in the data set before developing algorithms
- Conducting sufficient pre-release trials and post-release monitoring to identify, regulate, and mitigate any existing biases.

Use technology tools to build ethics in AI: One of the problems faced by those implementing AI is the black-box nature of deep learning and neural networks. This makes it difficult to build transparency and check for biases. Increasingly, some companies are deploying tech and building platforms which help tackle this. IBM's AI OpenScale, for instance, gives explanations on how AI models make decisions, and detects and mitigates against bias in the datasets.[10] There are many other open source tools that use AI to detect existing biases in algorithms and check the decisions and recommendations that AI systems provide. These AI tools mean companies can check their data sets and algorithms and make corrections as necessary. For example, ZestFinance which helps lenders use machine learning to deploy transparent credit risk models, developed its "ZAML Fair" tool to help reduce the disparity that affects minority applicants for credit.[11] Some startups are also building AI-based tools that are able to look into AI systems to make them more explainable and interpretable.[12] Organizations can use these tools to check their AI practices, mitigate biases, and build transparency into their AI systems.

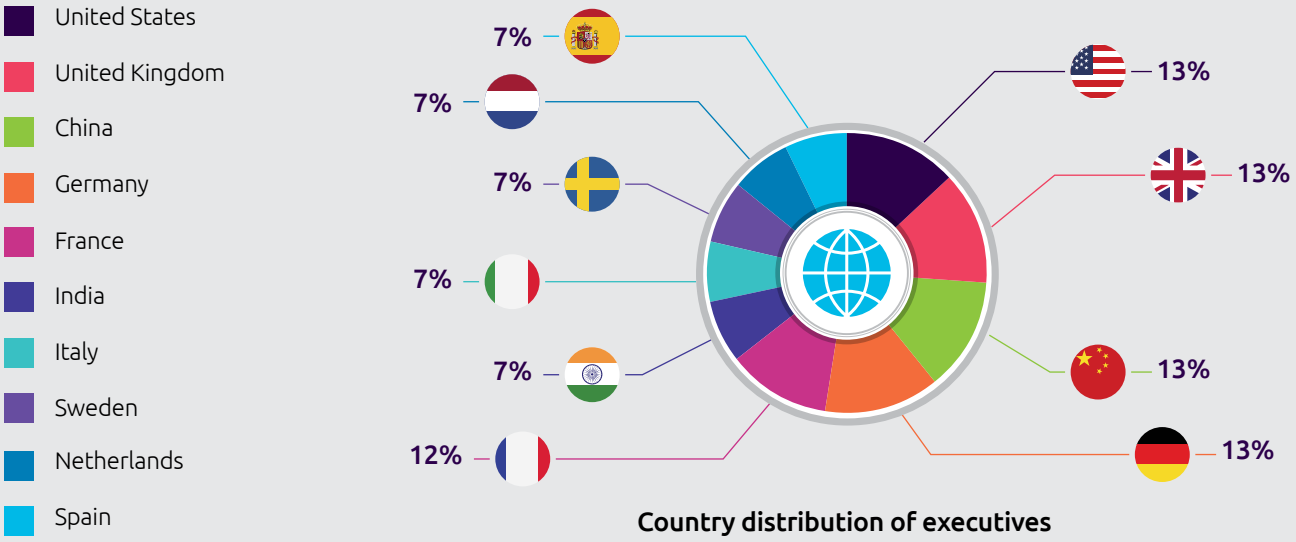
Conclusion

AI offers significant benefits for organizations with the right vision, planning, and approach to implementation. Our research shows that proactively addressing ethical questions in AI from the start is now a critical step to ensuring the adoption of artificial intelligence at scale in organizations. Organizations adopting an "ethics-by-design" approach for AI will earn people's trust and loyalty and greater market share compared to their peers. At the same time, they will stand to gain by preemptively averting significant risks from a compliance, privacy, security, and reputational perspective. Today, however, ethics in AI often does not get the attention it deserves even though some organizations are starting to take action. It is critical for organizations to establish strong foundations for why AI applications are built and how they are used. Organizations need to build accountability – for all teams involved – to infuse ethics into AI applications, by design, from day one. "Getting it right" with ethics in AI is still to be fully defined, and will evolve with the progressive adoption of AI across businesses and organizations, as well as with technology innovation. The first steps in this report will help organizations kick-start this journey towards ethical AI systems and practices.

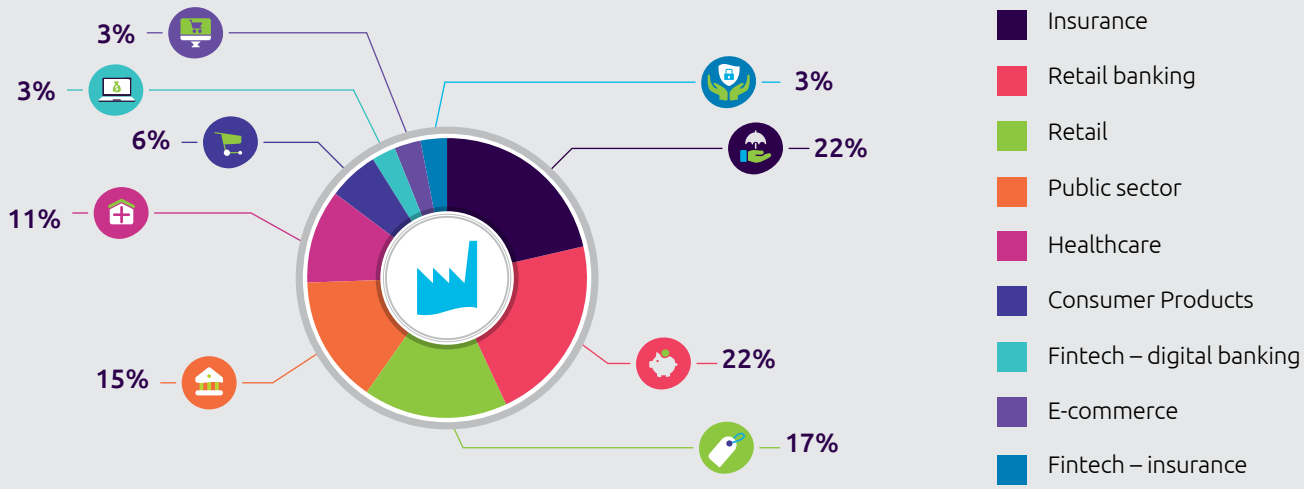
Research Methodology

Executive Survey: We surveyed over 1,580 executives from large organizations (with more than $1 billion in annual revenues each) in ten countries between April 2018 and June 2019. The executives were drawn from three broad groups:

1. General management/strategy/corporate
2. AI, data, and IT
3. HR/Marketing, with one executive from each group per organization.

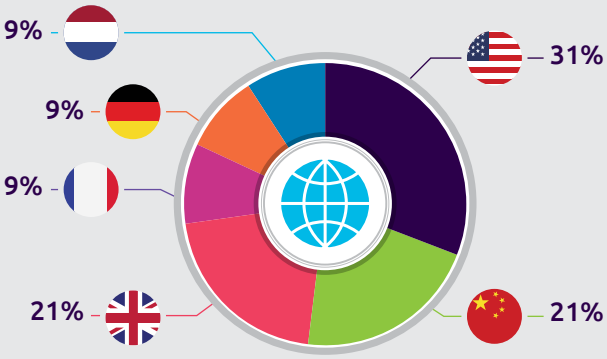


Country distribution of executives



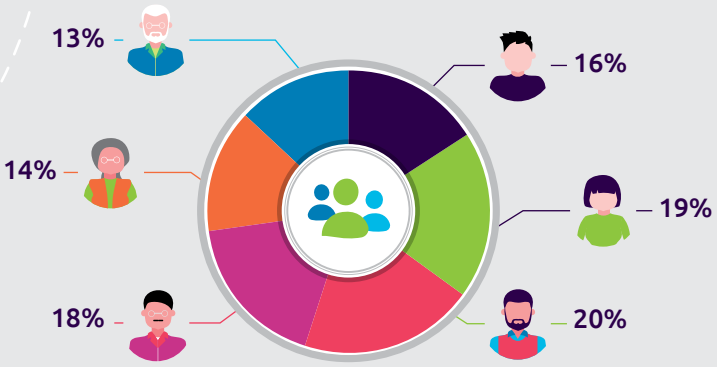
Industry distribution of executives

Consumer Survey: We surveyed 4,400 consumers from six countries. All of these consumers had had some form of AI interaction in the past so that they could relate to ethics-, trust-, and transparency- related issues in AI.



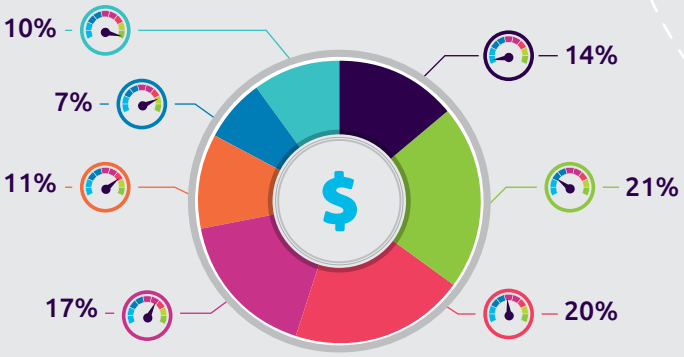
Country distribution of consumers

- United States
- China
- United Kingdom
- France
- Germany
- Netherlands



Age distribution of consumers

- 18–24
- 25–34
- 35–44
- 45–54
- 55–64
- 65+



Annual income of

- Less than $20,000
- $20,000–$39,000
- $40,000–$59,000
- $60,000–$79,000
- $80,000–$99,000
- $100,000–$119,000
- $120,000 or more

We also conducted in-depth interviews with more than 20 academics, industry experts, and entrepreneurs.

Appendix

Top three AI use cases that caused ethical concerns among consumers – by country

Country	AI use case
China	Healthcare – Processing personal data in AI algorithms for purposes other than for which it was collected
	Public sector – Collection and use of personal data (such as biometrics) by an AI system without consent
	Insurance – Reliance on machine-led decisions without disclosure
France	Healthcare – Collecting and processing personal data in AI algorithms without consent
	Healthcare – Processing personal data in AI algorithms for purposes other than for which it was collected
	Insurance – Reliance on machine-led decisions without disclosure
Germany	Public sector – Collection and use of personal data (such as biometrics) by an AI system, without consent
	Healthcare – Biased/unclear recommendations from an AI-based system for diagnosis/care/treatment
	Insurance – Reliance on machine-led decisions without disclosure
Netherlands	Public sector – Denied, without any explanation, aid/public benefits based on an AI algorithm's decision
	Insurance – Premium was set by an AI system based on race/ethnicity/income without any explanation
	Healthcare – Collecting and processing personal data in AI algorithms without consent
UK	Healthcare – Collecting and processing personal data in AI algorithms without consent
	Healthcare – Processing personal data in AI algorithms for purposes other than for which it was collected
	Insurance – Reliance on machine-led decisions without disclosure
US	Insurance – Reliance on machine-led decisions without disclosure
	Public sector – Collection and use of personal data (such as biometrics) by an AI system without consent
	Healthcare – Processing personal data in AI algorithms for purposes other than for which it was collected

References

1. Capgemini Research Institute, "Building the Retail Superstar: How unleashing AI across functions offers a multi-billion dollar opportunity," December 2018; Capgemini Research Institute, "Conversational Commerce: Why Consumers Are Embracing Voice Assistants in Their Lives," January 2018.
2. European Commission High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI," April 2019.
3. All quotes in this paper that are not attributed to a public source come from interviews directly conducted by the Institute.
4. Capgemini Research Institute, "The Secret to Winning Customers' Hearts with Artificial Intelligence: Add Human Intelligence," July 2018.
5. We explained to the individuals we surveyed what ethical AI interactions mean (for example, ethical in purpose, transparent, explainable, fair, etc.) and gave them examples of use of AI that may result in ethical issues to guage their perception.
6. The Globe and Mail, "The AI revolution needs a rulebook. Here's a beginning," December 2018.
7. Telefónica press release, "Telefónica is committed to using Artificial Intelligence with integrity and transparency," October 2018.
8. Standard Chartered, "Annual Report 2018," March 2019.
9. GE Healthcare, "Ethics in healthcare aren't new, but their application has never been more important," October 2018.
10. Forbes, "IBM Wants To Make Artificial Intelligence Fair And Transparent With AI OpenScale," October 2018.
11. Forbes, "ZestFinance Using AI To Bring Fairness To Mortgage Lending," March 2019.
12. H2O.ai website.

About the Authors

Anne-Laure Thieullent
Managing Director, Artificial Intelligence & Analytics Group Offer Leader
annelaure.thieullent@capgemini.com

Anne-Laure leads the Artificial Intelligence & Analytics Capgemini Group Offer (Perform AI), one of Capgemini's seven Group portfolio priorities. She advises Capgemini clients on how they should put AI technologies to work for their organization, with trusted AI at scale services for business transformation and innovation. She has over 19 years of experience in massive data, analytics, and AI systems.

Claudia Crummenerl
Managing Director, People & Organization practice, Capgemini Invent
claudia.crummenerl@capgemini.com

Claudia has more than 15 years of experience in change management and people transformation projects in large organizations in Europe, Asia, and the US. Claudia leads the People & Organization Practice for Capgemini Invent and is responsible for leading the thought leadership on the Future of Work. Besides her interest in leadership in the digital age and shaping the future people agenda, her focus lies on how AI is reshaping society and work and how Capgemini can contribute to architecting positive futures. Claudia has been involved in and has co-authored multiple Capgemini publications since 2011.

Amrita Sengupta
Manager, Capgemini Research Institute
amrita.sengupta@capgemini.com

Amrita is a manager at the Capgemini Research Institute. She tracks the patterns of digital disruptions across industries and their impact on businesses.

Ron Tolido
Executive Vice President and CTIO, Capgemini Insights and Data
ron.tolido@capgemini.com

Ron is the lead-author of Capgemini's TechnoVision trend series and responsible for Artificial Intelligence within Capgemini's Chief Technology & Innovation Network. With global clients, he focuses on innovation, agile architecture, digital strategy, and AI.

Jerome Buvat
Global Head of Research and Head of Capgemini Research Institute
jerome.buvat@capgemini.com

Jerome is head of the Capgemini Research Institute. He works closely with industry leaders and academics to help organizations understand the nature and impact of digital disruption.

Amol Khadikar
Manager, Capgemini Research Institute
amol.khadikar@capgemini.com

Amol is a manager at the Capgemini Research Institute. He leads research projects on key digital frontiers such as artificial intelligence, data privacy, and the future of work to help clients devise and implement data-driven strategies.

Gaurav Aggarwal
Manager, Capgemini Research Institute
gaurav.aggarwal@capgemini.com

Gaurav is a manager at the Capgemini Research Institute. He likes to assess how technology impacts businesses and understand how they respond to it. He is eager to learn about emerging business models, technologies, and trends across sectors.

The authors would like to especially thank Subrahmanyam KVJ, Matthew Newman, Yashwardhan Khemka, and Ranjit Gangadharan for their contribution to this report.

The authors would also like to thank Marie-Caroline Baerd, Michiel Boreel, Menno van Doorn, Moez Draief, Christoph Euler, Allan Frank, Rajesh S Iyer, Priscilla Li, Patrick Nicolet, Anne-Violaine Monnie-Agazzi, Rosemary Stark, Virginie Regis, Fabian Schladitz, Bob Schwartz, Mark Taylor, Charlie Weibel, Manuel Wiegand, Ashwin Yardi, Olivier Auliard, Valérie Perhirin, Isabelle Budor, Marijn Markus, Nicolas Claudon, Reinoud Kaasschieter, Thibault Lafont, David de Magalhaes, and Anika Zeidler for their contribution to this research.

About the Capgemini Research Institute

The Capgemini Research Institute is Capgemini's in-house research center. The Institute publishes research on the impact of digital technologies on large traditional businesses. The team draws on the worldwide network of Capgemini experts and works closely with academic and technology partners. The Institute has dedicated research centers in India, the United Kingdom, and the United States. It was recently ranked Top 1 in the world for the quality of its research by independent analysts.

Visit us at: www.capgemini.com/research-institute

Discover more about our recent research

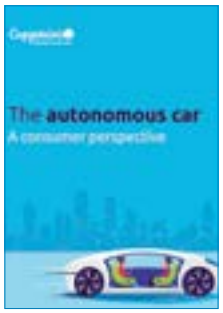
Accelerating Automotive's AI Transformation

How driving AI enterprise-wide can turbo-charge organization value

Building the Retail Superstar

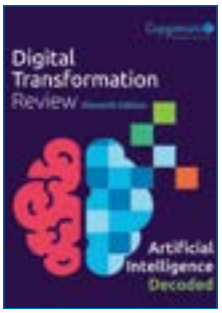
How unleashing AI across functions offers a multi-billion dollar opportunity

The Secret to Winning Customers' Hearts With Artificial Intelligence



The Autonomous Car
A Consumer Perspective

Digital Transformation Review 12 – Taking Digital Transformation to the Next Level: Lessons from the Leaders

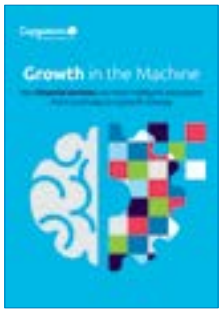


Digital Transformation Review 11 – Artificial Intelligence Decoded

Conversational Commerce: Why consumers are embracing voice assistants in their lives

Reshaping the Future: Unlocking Automation's Untapped Value

Turning AI into concrete value: the successful implementers' toolkit



Intelligent Automation in Financial Services

Intelligent Automation in Energy and Utilities: The Next Digital Wave

Upskilling your people for the age of the machine

For more information, please contact:

Global

Anne-Laure Thieullent
annelaure.thieullent@capgemini.com

Lanny Cohen
lanny.cohen@capgemini.com

Ron Tolido
ron.tolido@capgemini.com

Philippe Christelle
philippe.christelle@capgemini.com

Germany
Fabian Schladitz
fabian.schladitz@capgemini.com

France
Sébastien Guibert
sebastien.guibert@capgemini.com

Netherlands
Pieter Nieuweboer
pieter.nieuweboer@capgemini.com

UK
Joanne Peplow
joanne.peplow@capgemini.com

India
Balaji Thiruvenkatachari
balaji.thiruvenkatachari@capgemini.com

Scandinavia
Robert Engels
robert.engels@capgemini.com

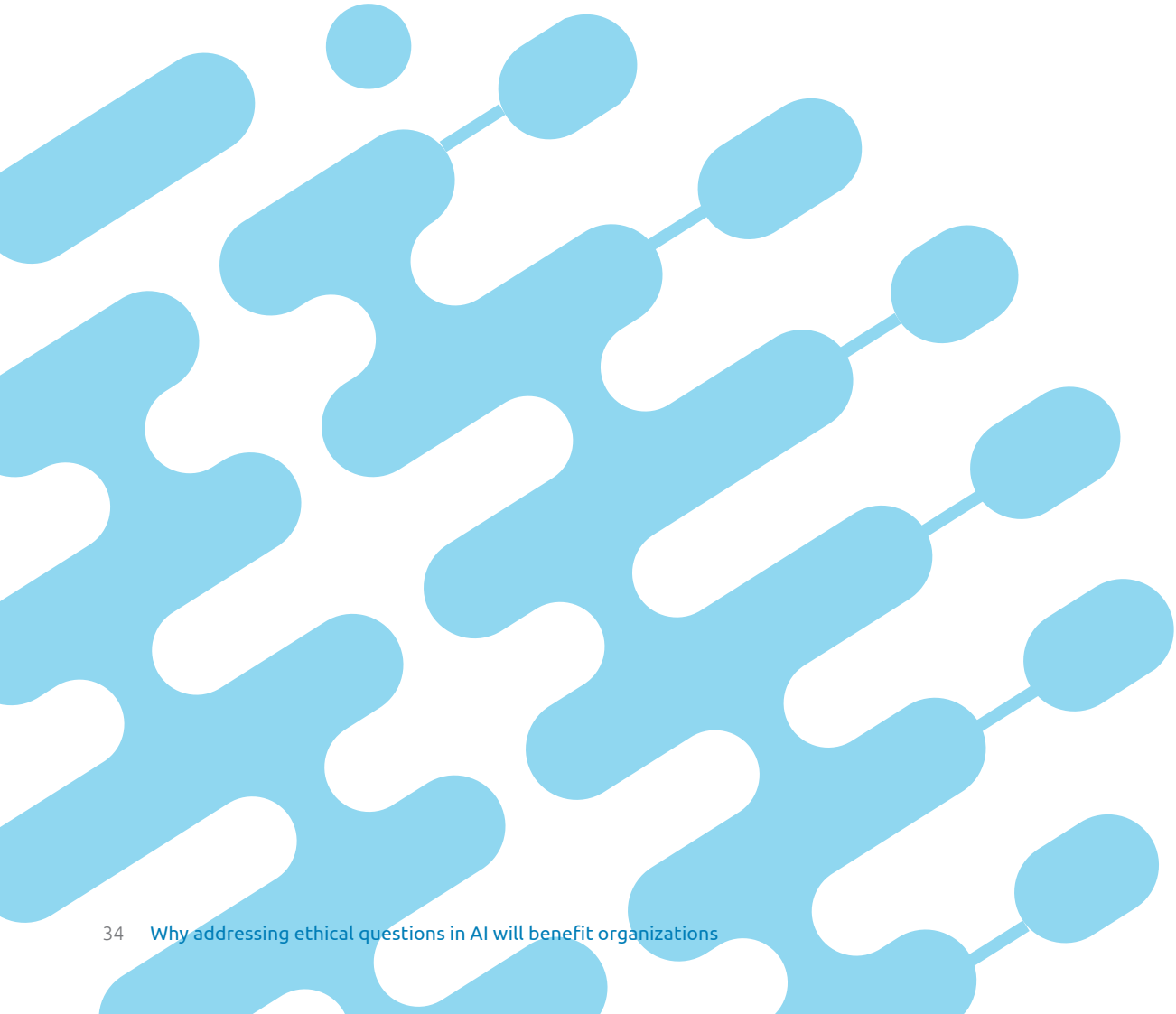
North America
Mark Kirby
mark.kirby@capgemini.com

Australia
Matthew Newman
matthew.newman@capgemini.com

Subscribe to latest research from Capgemini Research Institute

Receive advance copies of our reports by scanning the QR code or visiting

https://www.capgemini.com/Research-Institute/



About Capgemini

A global leader in consulting, technology services and digital transformation, Capgemini is at the forefront of innovation to address the entire breadth of clients' opportunities in the evolving world of cloud, digital and platforms. Building on its strong 50-year heritage and deep industry-specific expertise, Capgemini enables organizations to realize their business ambitions through an array of services from strategy to operations. Capgemini is driven by the conviction that the business value of technology comes from and through people. It is a multicultural company of over 200,000 team members in more than 40 countries. The Group reported 2018 global revenues of EUR 13.2 billion.

Visit us at

www.capgemini.com

People matter, results count.

MACS_04072019_KB