



# UCL

# THE FIVE KEY COMPONENTS OF AI ETHICS TRANSFORMATION

## Authors

Paul Walton (Capgemini)

Jasmin Booth (Capgemini)

Jamie Rich (Capgemini)

Dino Mariutti (UCL MBA student)

Tom Weston (UCL MBA student)

UCL  
SCHOOL OF  
MANAGEMENT

---

ANALYTICS  
LAB

Capgemini 

---

## ABSTRACT

AI ENABLES RADICAL CHANGES TO BUSINESS PROCESSES BUT BRINGS WITH IT A SWATHE OF ETHICAL ISSUES. THESE ISSUES ARE A MAJOR CONCERN.

“THERE’S NO BUSINESSPERSON ON THE PLANET AT AN ENTERPRISE OF ANY SIZE THAT ISN’T CONCERNED ABOUT [...] WHAT’S GOING TO BE [...] ETHICALLY ACCEPTABLE.”

But how should organisations tackle these issues and respond to the specific ethical challenges that AI brings? This article highlights the five key components of AI ethics transformation. Organisations need:

- An ethical approach to **AI governance**;
- **Analytics** to ensure that they can understand the impact of ethical issues and how to resolve them;
- Measurement of the degree to which AI ethics principles are delivered in the **implementation of AI**;
- Measurement of the maturity of the **enablers** (or inhibitors) of ethical AI;
- An incremental **transformation** approach to improvement.

# AI ETHICS

The scope of ethical issues raised by AI is wide-ranging. And the responses to it encompass everything from the philosophical (for example, as expressed by Michael Sandel) to the more pragmatic.

Numerous sets of principles for ethical AI have been documented (over 150 sets of guidelines are documented in the global inventory). But nevertheless, there is clear common ground captured, for example, in the EU principles shown in Figure 1.



Figure 1: EU principles for ethical AI

---

The principles are important because many ethical questions are likely to arise when machine learning is implemented (machine learning has driven the recent increase in AI). Machine learning uses data that, in itself, reflects the biases and ethical issues of society. So, "bias is machine learning's original sin" and comes in many forms.

An example, from the motor insurance claim process, shows the nature of the problem. The columns in Table 1 show elements of the process, potential uses of AI and the ethical issues that might arise. The table shows highlights the scope and complexity of the issues.

**Table 1: Ethical issues in motor insurance claims**

Process element	Potential AI role	Ethical questions and considerations
Prevent accidents	Understand driver behaviour and influence drivers to improve	<ul style="list-style-type: none"> <li>How far can this go without violating the “human agency” principle? What degree of psychological manipulation is acceptable?</li> <li>Will people agree, knowingly, to how their data is used? What level of consent is needed for different uses? Does the need for consent imply any form of discrimination?</li> <li>When does understanding driver behaviour violate the need for privacy? Many driver behaviours (e.g. driving at night) may not be a choice (e.g. for shift workers).</li> <li>Is there sufficient evidence that the influence techniques work well enough? When is a statistical relationship sufficiently robust?</li> </ul>
Acquire and update information about the claim	Understand the complexity and risk (including for fraud and criminal activity) associated with the incident through analysis of images, documents and interaction with participants (including witnesses, breakdown services, medical facilities, garages, lawyers, ...)	Image, document and interaction analysis and their relationship with risk are all subject to bias and the impact of previous risk thresholds
Acquire and update information about the driver	Understand the propensity for poor driving behaviour, fraud or criminal activity from information about the driver and associates to compare with industry sources	<ul style="list-style-type: none"> <li>Increasing the level of profiling risks violation of privacy and fundamental rights and could lead to a form of social credit</li> <li>How much data is it fair to collect (e.g. from social media) and to what extent is the collection context-specific (e.g. with respect to potential criminality)? To what extent does the transparency principle imply that the customer can see what data has been used? How much consent is or should be needed for access to this data?</li> <li>Profiling requires very robust systems and robust rationale to ensure that the results are fair (how reliable is historical data about people as they change?)</li> <li>Profiling may violate the diversity and fairness principle because of bias in the data</li> </ul>
Interact with parties involved	Manage activities, communications and relationships with participants	<ul style="list-style-type: none"> <li>To what extent should AI demonstrate transparency in the interactions? How transparent should the use of data be and to what extent should it be available to challenge?</li> <li>How can the AI <u>explain decisions</u>, as well as the use of data well enough to different parties?</li> <li>What is an acceptable level of emotion sensing and psychological manipulation in interactions (e.g., Microsoft have withdrawn their emotion sensing technology)?</li> <li>How much will human agency be diminished if users cannot speak to anyone (because of AI-driven interactions, extending current chatbot use)? What is the impact of the diversity principle in establishing fair interaction?</li> </ul>
Decide the next step in the case	Profile the level of risk based on updated information about the incident, driver and other risks. Decide the process to follow and the need for specialist intervention	<ul style="list-style-type: none"> <li>What is an ethical threshold for AI (versus human) decision-making?</li> <li>Can decisions be explained to customers and other stakeholders well enough?</li> <li>As more decisions are made by AI, what impact will this have on human decision-making?</li> </ul>
Litigate	Provide evidence in any claim taken to court	Are the decisions made by AI sufficiently robust and <u>explainable</u> ? To what extent is <u>regulation or accreditation</u> required to provide a level playing field?
Improve	Analyse the process followed and recommend improvements	If AI is used to recommend improvements to the claim process, will those recommendations encompass ethical considerations?

---

# A BURNING PLATFORM

Ethical issues are not just theoretical. The following four factors are creating a perfect storm:

1. **Growth of AI.** The use of AI in organisations is growing fast. It will become pervasive and embedded in multiple forms in business processes. For example, in the insurance example above, McKinsey forecast that as much as 60% of routine insurance claims and customer interactions will be suitable for automated digital resolution by 2030.
2. **Widespread availability.** At the same time, AI is expanding beyond those with the skills to control it. It isn't just that there is a shortage of data scientists – AI is no longer the exclusive domain of knowledgeable data science teams. It is readily available to all developers through free libraries (e.g. Python) or cloud services (e.g. AWS, Azure, Google, IBM) and is being widely included in software products. But it can also be accessed directly by business users in a similar way to Excel and Access (through “no/low code” products like the Microsoft Power Platform and many others).
3. **Risks.** Deliberate or inadvertent misuse of AI (for example, through “artificial stupidity or other causes”) risks triggering reputational, compliance or financial damage. At a wider scale, the proliferation of AI, and increased energy use, may inhibit sustainability.
4. **Urgency.** The problem cannot be deferred. Ethics needs to be designed in and cannot be retrofitted.

Without a clear approach to AI ethics, AI will be in the hands of business users and technical teams driven by financial and delivery pressures without a clear understanding of the risks or the means of mitigating them.

# CHALLENGES

How can organisations address this burning platform? First, it is important to understand the scale of the challenge. Even though the principles are understood, implementing AI ethically is hard: “Organizations can start with an admission: Getting serious about ethical AI is no easy task”. Figure 2 shows that the challenges are multi-faceted, ranging from deep questions about the underlying concepts—touching moral philosophy—to the more pragmatic.



Figure 2: Challenges in implementing AI ethically

References in the diagram:

1. [Ethical Concerns Mount](#)
2. [How Machine Learning Pushes Us](#)
3. [This is How AI Bias Really Happens](#)
4. [A Guide to Explainable AI](#)
5. [Ethical Frameworks for AI Aren't Enough](#)
6. [The Secret to AI is People](#)

---

These challenges mean that the root causes of ethical issues can be very diverse. AI will only be ethical if the organisation as a whole enables it. For example:

- the organisational leadership may not have sufficient understanding of AI, AI ethics and its potential impacts, or the consequent scale and nature of the challenges shown in Figure 2;
- there may not be adequate understanding throughout the organisation, not only of AI, but of the importance of data – good quality data is essential for AI and the existing data architecture, governance or processes may inhibit the production of sufficient quality data and identification of ethical issues;
- people may not have the skills needed, either to implement AI or, more widely, to understand how to work it; and worse still, they may not trust the AI because of the way in which it has been implemented;
- the culture may be at odds with adherence to ethical policies or may inhibit change;
- the technology architecture may not include the tools or integration needed – scaling the use of AI reliably requires an appropriate architecture.

Therefore, resolving these challenges requires transformation:

*“Too many business leaders still believe that AI is just another ‘plug and play’ incremental technological investment. In reality, gaining a competitive advantage through AI requires organizational transformation of the kind exemplified by companies leading in this era... These companies don’t just have better technology — they have transformed the way they do business so that human resources can be augmented with machine powers.”*

The nature of the transformation depends on whether the leadership, people and culture, technology, data, processes and governance of the organisation enables or inhibits an ethical approach to AI. The maturity of these ethics enablers will determine the scale of transformation required in each case. For example, if the quality of data used by AI is poor, then it may make sense to prioritise data governance and data engineering.

As with digital transformation in general (see, for example, “Leading Digital” by Westerman, Bonnet and McAfee), there will always be a tradeoff between developing specific AI capabilities and putting in place the enablers.

With all these difficulties and in a changing environment, how should such a transformation be designed? The key is to acknowledge the uncertainties, learn and improve.

# THE FIVE COMPONENTS OF ETHICAL TRANSFORMATION

Many elements of the solution are understood (including, for example, an ethics committee and ethical governance, ethical infrastructure supplementing existing AI, data and product strategies and organisational change, the need for AI measures).

But these are not enough. They need to be connected in a way that can address the questions raised above. Organisations need to be able to understand:

- What ethical issues are created by the implementation of AI (current or planned)?
- What causes the ethical issues?
- What is the actual or potential impact of the issues on business outcomes? What could resolve or sufficiently mitigate the issues?
- What changes are needed? How should they be prioritised?
- What is the plan to make the changes? How effectively are the changes being made?

Providing answers to these questions indicates the capabilities and how they should be connected. This is shown in Figure 3.

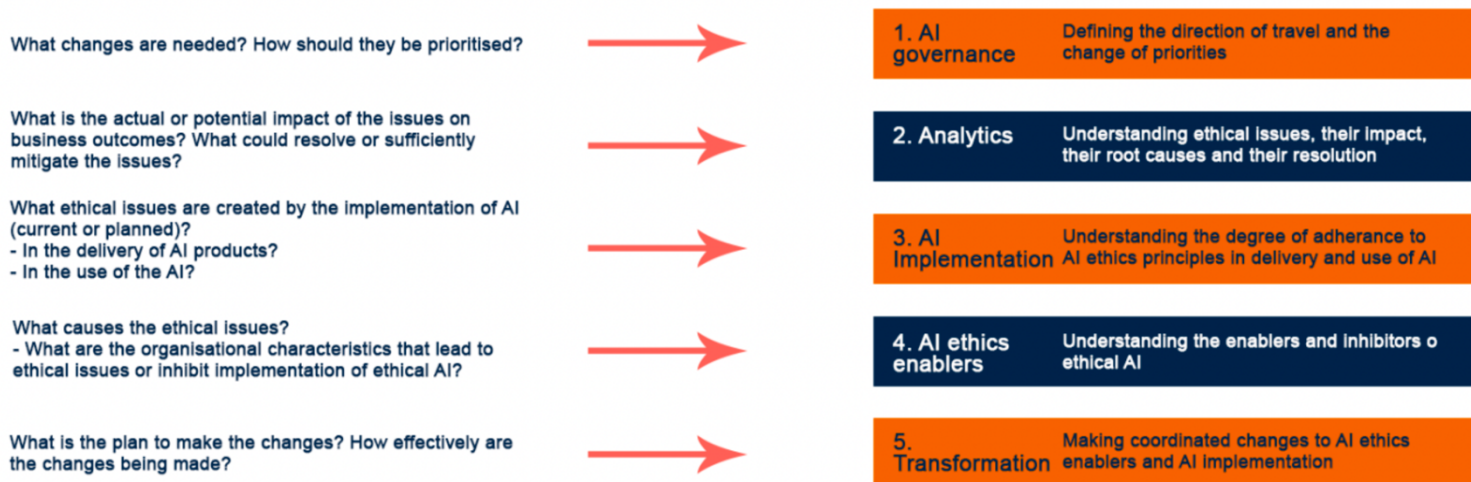


Figure 3: A framework for ethical AI transformation

The governance of AI needs an understanding, not just of the ethical issues, but of the tradeoffs between implementing new AI, resolving immediate ethical issues and tackling the underlying enablers. This understanding requires detailed analysis of the issues, their (actual or potential) impact on business outcomes, their root causes and the changes required to resolve or mitigate them.

---

Ethical issues may arise from both the AI in itself and also its use. So, an understanding of ethical issues needs to take into account both:

- Measures associated with AI product development or integration (for example, whether an ethical approach is built in);
- Measures of the ethics relating to the use of the AI (for example, the skills of its users and its ability to provide explanations).

But this isn't enough. Many causes of ethical issues may lie with the underlying issues discussed above. Leadership, architecture, skills, culture and governance may limit the ability to implement AI ethically. So it is important to measure and understand the maturity of the enablers (with respect to AI ethics).

Any change initiatives need to be coordinated and integrated (perhaps, for example, using McKinsey's "mission control" concept) and propagated through changes to the management of digital products incorporating AI. And this coordination needs to be supported by rigorous and holistic organisational change management to tackle the people, culture and organisational changes needed.

So, it is not enough just to measure adherence to AI ethics principles in the implementation of AI—the other four capabilities shown in Figure 3 are also needed.

---

## CONCLUSION

Implementing AI ethically is necessary and urgent. There isn't yet a mature, agreed approach, so organisations need to learn and improve. To do so, they need five components of ethical transformation:

1. **AI governance:** to define the direction of travel, the priorities and the changes required;
2. **Analytics:** to understand the impact of AI ethics challenges and how to resolve them;
3. **AI implementation:** to measure adherence to AI ethics principles in the development and use of AI;
4. **AI ethics enablers:** to measure the maturity of the enablers (or inhibitors) of ethical AI;
5. **Transformation:** an incremental approach to learning and improvement.

---

## ABOUT US

### CAPGEMINI

**Capgemini** is a global leader in consulting, digital transformation, technology and engineering services. The Group is at the forefront of innovation to address the entire breadth of clients' opportunities in the evolving world of cloud, digital and platforms.

### THE UCL SCHOOL OF MANAGEMENT

The **UCL School of Management** is the business school of University College London, one of the world's leading universities, consistently ranked in the global top 20 for its academic excellence and research. The School offers innovative undergraduate, postgraduate, PhD and executive programmes in Management, Entrepreneurship, Business Analytics, Business Information Systems, and Finance, designed to prepare students for leadership roles in the next generation of innovation-intensive organisations.

### THE ANALYTICS LAB

**The Analytics Lab** is an enrichment module for business students where they are able to explore topical questions in the domain of business analytics and digital economy via hands-on experience. Students are offered the opportunity to conduct research and work on projects with leading technology service and consulting companies.

It aspires to help UCL business students and alumni to be in the heart of fundamental changes and digital transformations in the business environment. Students enhance their practical abilities to manage and operate business activities effectively in view of rapidly developing digital and technological advancements in data analytics.