

Can you trust a robot?

Sophie Falcon, ICAEW's integrity and law manager, explores the ethics of new technology

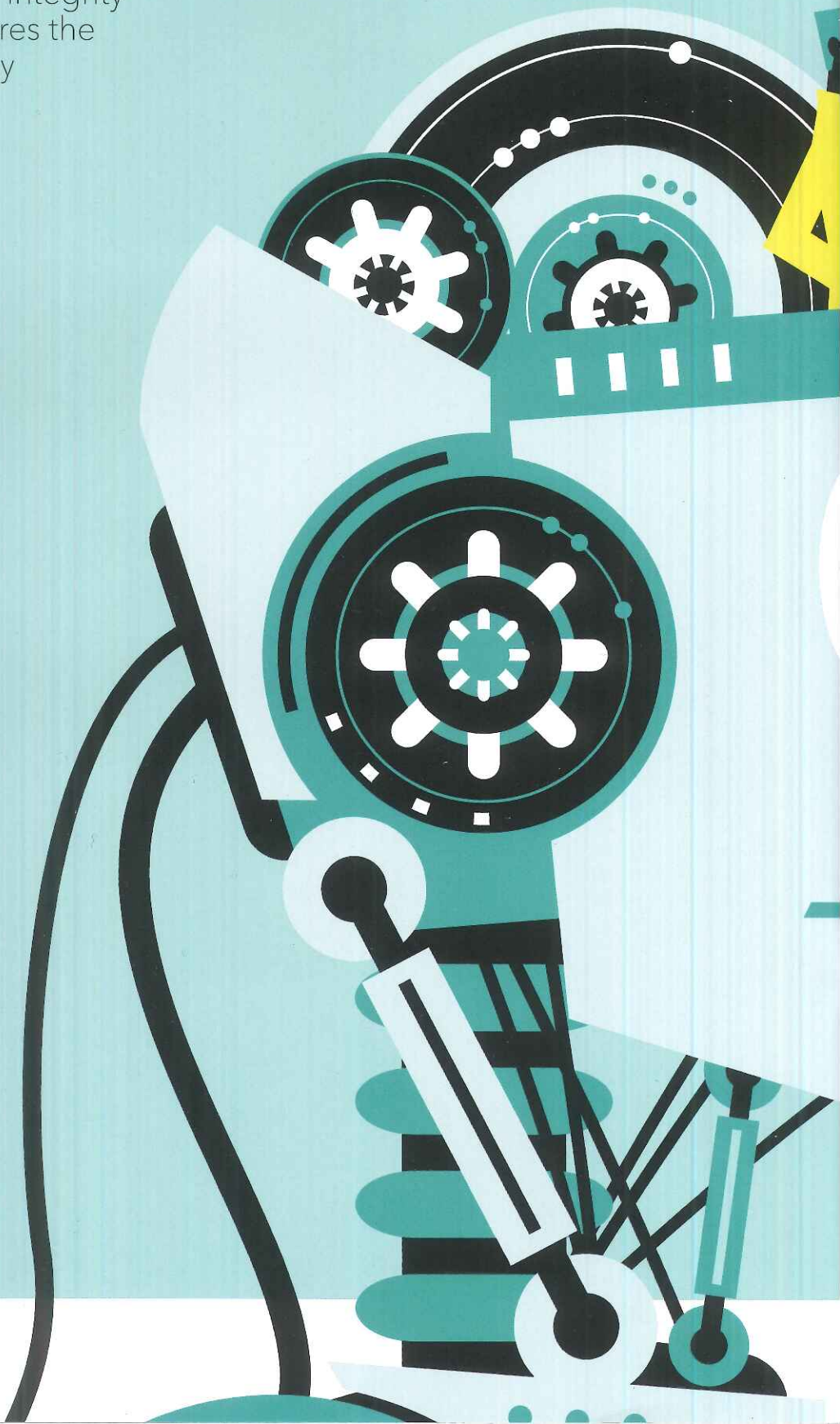
Accountancy is changing. New technology is permeating our daily work, and artificial intelligence (AI) systems are expected to replace or enhance key parts of traditional accountancy roles. But what does this mean for the ethics of the profession?

Members and students of ICAEW must comply with the Fundamental Principles of the Code of Ethics which require us to act with integrity and objectivity; to work competently and diligently; to respect confidentiality and behave professionally. But what if an AI machine becomes the auditor? How can we trust it to observe these ethical principles and what can we do to control its actions?

We need to train AI how to think ethically. So what makes a person ethical? There may be different answers to this question depending on where you are in the world, and the sector you operate in. Humans are imperfect and they have bias, whether conscious or unconscious. If you use data from previous human decisions to train AI, it's fairly likely that you will create a machine with those same biases. If the machine then starts making decisions using the same values, it could reinforce undesirable outcomes in terms of gender, race or disability, among others.

EMOTIONAL CUES

Another ethical concern with letting machines make decisions is whether the decision will be fair. A human making a decision that involves, for example, vulnerable people, will be influenced by their emotions. The human's reasoning won't always be explicitly explained, it will just be a feeling of whether something is right or not. What we might think is



acceptable business practice may vary if the client is suffering from mental health issues. Will a machine identify that an individual is vulnerable, and will it properly take such factors into account?

Humans will be the ones to educate AI, so maybe we need to think of AI as a child, and that we are teaching it right from wrong. If AI learns from every situation it encounters, then perhaps you can create an ethical being. Properly trained AI, that understands and applies ethical principles, and is free from human bias, could potentially become fairer and more consistent in decision-making than a human. The challenge is the complexity of technology underlying AI, and the extra learning we will all need to undertake to interact with and teach AI effectively.

"If you use data from previous human decisions to train AI, it's fairly likely that you will create a machine with the same biases"

Machines, like humans, make mistakes. But if AI comes out with an unexpected answer, how do we know if it's wrong? Will the human users feel able to override a machine decision? If we don't know how AI has reached its conclusion, how can we criticise its output?

There are two issues here. The first is professional scepticism. Just because a machine, albeit one designed to be infallible, has produced the unexpected result, doesn't mean that as accountants or auditors we should just accept what it gives us. The same critical review should be made as if a human had calculated the result. We also need AI to show its workings. It remains to be seen whether the AI that emerges will provide this rationale, or whether it will work on a 'black box' basis. In some organisations, such as the US military, there is a call for all AI systems to be explainable AI or 'XAI'.

The way that an audit is performed may fundamentally change if you can get a machine to check every transaction and every balance, rather than a selected sample. This is a real opportunity for audits to provide a higher level of comfort on the financial statements.

What about the auditor independence implications of AI? Under the auditing

ethical standards, the audit partner must be independent of their audit clients, and safeguards must be implemented to prevent any threats to independence. One of those threats is familiarity, so audit partners of listed and public interest entities are required to rotate off the audit after five years of acting, to safeguard against any real or perceived lack of independence from the client.

REPLICATION ERROR

So what about AI taking the place of the audit partner? How do ethical standards apply? Can a machine develop a familiarity bias and need to be rotated to preserve independence? That raises the tricky question of how you rotate one intelligent machine for another. If multiple machines have all been trained in the same way it is debatable what you would achieve by swapping one machine for another that thinks identically.

Perhaps the competitive advantage between firms will come down to the quality of their firm-specific AI, and by rotating one audit firm for another, you apply a distinctly different AI system's thought processes, rather like a different human audit partner. This raises the issue of whether there is one single best way to do something? If so then convergence between the best AI machines should be inevitable, which takes you back to square one again.

In an imperfect world, we can't prevent all failures. But who is responsible when the AI system goes wrong? It is hard to imagine that an accountant could justifiably blame AI for the mistakes, and

how far does that go? Would the programmer become liable? It remains to be seen whether the human desire to blame someone would extend to AI, and whether that would be a valid defence. Until AI holds the same legal and professional status as an individual, humans will still be responsible for AI's error.

New technology and AI present huge opportunities for the accountancy profession but we need to take care with what we teach intelligent machines, and retain responsibility for their actions and the decisions they take on our behalf. Professional scepticism may mean something new in the age of AI, but it will take on even greater importance so that we can navigate the ethical landscape of new technology. ● For more on ethics and new technology visit icaew.com/ethicsandtech

