



## Where responsibility lies when it comes to AI Ethics in Banking

 Dr. Giles Cuthbert\*

### Introduction

The world of banking has changed a lot since the Chartered Banker Institute was first established almost 150 years ago in 1875 and in recent times some regard it as changing almost beyond recognition, driven by greater competition, greater globalization, significant socio-demographic changes and of course, by the greater possibilities created by new technologies such as Artificial Intelligence (AI). We at the Chartered Banker Institute are determined to keep abreast of all the latest developments in banking and FinTech, whilst retaining pride in our past history and achievements and are keen to ensure that the banking sector's ongoing drive towards its digital transformation is sustainable, responsible and reflecting on the skills, expertise and professional judgement that will be required by bankers of the future.

We are, therefore, excited to see the opportunities that Artificial Intelligence will bring to the banking sector. However, we think that the use of AI is nuanced and caution should be exercised to ensure it adheres to the same professional standards and values as humans. As a professional body, we will continue to help our members identify how AI can enhance and benefit their roles, rather than focussing on the threats often presented in the media and that is why earlier this year, we launched our 'Certificate in Digital & AI Evolution in Banking' to develop the learner's

knowledge of digital, AI and automated banking and the role of FinTech and which counts towards that individual becoming an Associate Chartered Banker. Through my own research I have also looked in some detail at the ethics of AI in banking and I wanted to share some of my insights, which individuals working across the sector may find useful as practical guidance going forward.

AI - the concept of emulating human-like intelligence in machines or software, enabling them to execute tasks that humans commonly undertake is an area which also continues to focus the minds of global financial regulators. The use of AI, which include some element of machine learning in banking, is not new but the use of new forms of AI, such as Generative AI, is what is leading to heightened risk concerns in banking and financial services around explicability, capacity, accountability and potential bias. When looking at the use and potential uses of AI and considering whether the current regulatory systems are fit for purpose, there is a need for a heightened focus on ensuring there is a broader understanding of "agency" and "responsibility" across the sector and where responsibility lies when it comes to AI ethics.

One of the key features of an ethical situation is if the person actually has the ability to take an ethical decision. For example, a 3-year-old child would not be expected to make ethical decisions about complex matters - or even basic matters - as they do not

---

\*Managing Director, Chartered Banker Institute, United Kingdom.

have the ability to do so. This pinpoints the concept of 'capacity' in ethical decision-making. Capacity is a precursor to responsibility - that is, in order to take responsibility for something, we need to have capacity to make the decision in the first place. There are different forms of incapacity, such as a mental illness, age or a number of other factors.

Once it has been established that a person has capacity in an ethical situation, they then have the ability to take a decision that is seen as valid: this is what we describe as agency. In the context of a professional, 'agency' refers to their capacity to take a decision in a certain situation for which they then have responsibility. So, once we know there is agency, the next step is identifying 'responsibility'.

When identifying responsibility, there are two key features which must be considered: firstly, whether the person has sufficient knowledge of the situation and secondly, whether a person has control of the situation (or certain aspects of the situation). This is important, as a person can be involved into a situation but not have control over the outcome and therefore cannot be considered to be responsible.

It is that first factor of sufficient knowledge where 'professionalism' comes into play, because only professionals can take responsibility in certain situations where professional knowledge is required. For example, an amateur may be involved in an ethical situation, but they could not be seen as responsible as they do not have professional authority or knowledge needed for that situation. It is this factor of responsibility that allows for praise and blame in certain ethical situations.

It is these features of agency, professionalism and responsibility which allow us to understand the distinction between good and bad; this can also be referred to as a value-based (or axiological)

distinction. And our role as a professional body is ensuring that we have individuals that are confident in their role as the responsible human to call on their professionalism, expertise and judgement to look at the distinction between good and bad outcomes i.e. is it better or worse that I do this than an alternative action?

The concept of ethical decision-making in relation to AI ethics, centres on how digital systems and AI take automated decisions throughout the course of their daily operations. The key distinction is, therefore, on the unique ability that humans have to intuitively know all these things they are doing are good on a normal, day-to-day basis. When humans are struck by an ethical dilemma, their identification of that ethical dilemma is triggered by their 'ethical sensitivity' - the human capacity to notice that there is a problem or challenge that requires particular ethical attention or analysis. In other words, it is our ethical sensitivity that causes us to stop and think about some things more than others.

Although this may sound simple on the surface, there is a significant history of ethical dilemmas and scandals caused by organisations or individuals who did not stop long enough to reflect on whether there was an ethical problem with something they were attempting to do. It is this failing of ethical sensitivity which lies at the heart of many ethical problems.

That said, it is very difficult in large complex organisations or large complex societies to see where responsibility lies and you could argue that is a very hard in this situation for an individual to be responsible for everything they do because they have such a requirement for the help and support of others to allow them to take the actions they do. This has led people and financial regulators in particular, to think that the concept of 'individual responsibility', which

is very popular in western culture is the right way to think about responsibility. For example, if you look at a large complex organisation you could think of it in terms of 'complicit responsibility'.

The concept behind 'complicit responsibility' is that we as individuals all work together, we are all inter-reliant on each other and therefore no one individual is fully responsible, there is always a complicity around the responsibility which we have. In legal terms, you encounter this around joint and several liability, which is very similar to responsibility where we all take a share in responsibility or liability for a situation, so it may sound fanciful at first to say we are all complicit in actions taken by an organisation but perhaps there is some truth in that.

This is even more so the case when we start to consider AI - when we consider the complexity of the number of organisations that may be involved in the production, development and design of AI and the organisations that deploy it and then how all those different digital systems and AI interact with each other, we start to see a very complex map of interactions between different systems, different organisations. So perhaps when we start to think about AI, we need to have slightly different forms of responsibility. Maybe – it is more about complicit responsibility rather than individual responsibility as we see all these maps of responsibility draw together. This is a contentious concept because we typically find ourselves in a situation in society where people look to blame or look to praise an individual and if we start saying our responsibility and indeed our accountability is complicit then that perhaps undermines a central tenet of our society.

So, can a machine/AI ever be held responsible? We all know we have been in a situation where AI has taken a decision or a digital system has taken a

decision and we have to stop and consider bearing in mind what we already know that all decisions are inherently ethical, it may be they are good but they are still ethical. Should we have AI or machines taking decisions on that basis? Are we comfortable for that to happen? This brings us back to the requirements for responsibility so a machine can have knowledge of a situation potentially, they may have control over aspects of the situation and they may be able to influence its outcome but the big question is surely do they have capacity or agency? Typically, only humans have agency – for instance, we would never say a dog is responsible for biting someone, we would typically blame the owner and we would not sue the dog if we were bitten by the dog, we would sue the owner! So, bearing in mind we could not sue the AI, can we say that AI has agency? And if does not have that agency, can it take decisions because after all we know that responsibility requires that agency. This is one of the central debates about AI and ethics and it is why I wanted to explore responsibility in the way I have done in this article because it is really critical to understand what role AI can play; can it replace the professional? Can it replace the human decision maker? This is an area of huge debate and a lot of the literature on AI ethics looks at whether a robot can ever have agency. The key issue for us, in answer to that question, is can we as responsible individuals allow a system which is de facto “not responsible” to take decisions on our behalf?

One further point is, if we know AI cannot be responsible but it is taking highly complex decisions and if we have a concept of individual responsibility rather than complicit responsibility surely we are creating a dangerous situation where individuals are given responsibility for something they cannot control. Because we know control is a requirement, perhaps we are getting into the perfect storm here

where AI can never have agency and can, therefore, never be seen as fully responsible, yet humans do not have full control and therefore, cannot be seen as fully responsible. In which case we analyse responsibility, we find a serious lack in the concept of responsibility at this stage and this undermines some of our most basic ethical thinking.

At the moment, regulators seem to be taking the view that their existing model risk management standards has given them enough powers to deal with the risks around these AI models because fundamentally they believe a human is in control of the model and they need someone to be accountable for the model and those individuals will be held to account. They have also tried to remain technology agnostic up to now but it is possible that this latest round of AI developments might need them to supplement what they have done in some areas, but they feel that have not reached that stage yet.

Nonetheless, regulators are clearly in a hurry to keep up to speed with developments in AI, in terms of the frameworks they are presenting and that is why I have taken into account all the ethical considerations raised above to design a model, which can be applied in AI in banking and which is available from the Chartered Banker Institute. One critically important feature of this model is that it is a reminder that ethical, personal and professional values are not a 'nice to have', they are an essential component of what leads

to good ethical outcomes. Digital ethics cannot be approached solely with a good framework or code without the underlying professional competencies, knowledge and ethical commitments that individuals bring to the environment. This model demonstrates how all the different factors at work in the theory and practice can be shown to come together and thus, visualises for individuals a model which they can use to, hopefully, optimise ethical thinking. This is not, of course, a guarantee of positive ethical outcomes, it is a model designed to support ethical decision making.

This last point is a key one, we have a plethora of ethical decision making frameworks, but those frameworks can only support a human to do the right thing. Ultimately, human has to use their intuitions and moral courage to act in the right way. Research in moral psychology has shown that humans have better intuitions, the greater their expertise and we also know that individuals are more likely to have the moral courage to act the right way when placed in a difficult ethical situation. Both these factors, expertise and community, are provided by the professional bodies represented at Asian-Pacific Association of Banking Institutes (APABI), so I can say with some confidence that whatever the future world of AI and digital holds for us, professional banking bodies are an essential part of that future.

