

Who will watch the robots? Ethics, Insurance and Artificial Intelligence

More knowledge, more ethics?

More information should help us become more ethical. Ethics is (at least partially) a matter of recognising and predicting how our actions affect others, and then pursuing those actions that result in the least harm or the greatest good.

With more information and better prediction our ethical decision-making should improve. At least, this would be the case if one assumes, as Socrates did, that “knowledge is virtue”. Socrates believed that when we truly grasp what is good, and why it is good, such knowledge is compelling – it effectively guides behaviour. The inverse, then, is that bad behaviour is ultimately a form of ignorance.

Artificial Intelligence (AI) for good

It stands to reason, then, that the advent of Big

Data and the predictive power of AI opens new moral opportunities. And there are uses of Big Data and AI that do exactly that. Consider, for instance, the personal behavioural improvements made possible by the myriad of tracking apps available to individuals. Receiving alerts and summaries about the driving “mistakes” you make allows you to consciously adjust your driving, learn new habits, and consequently reduce the risk of accident, injury or death. In a Socratic sense, these applications help the moral agent to “know thyself”¹.

On a bigger scale, organisations like “AI for Good” are pursuing moral good through the use of AI. One example is rAlnbow (or “Bo” for short) – an AI-powered conversational bot that provides a safe and non-judgemental space to identify and prevent abuse and gender-based violence². Another example, from the field of psychiatry, is the use of machine learning to identify the risk of suicide from an analysis of social media posts, creating algorithms that will aid clinical decision-making in future³. Through technologies like these, new avenues open up for the effective pursuit of health, mental and physical wellbeing, community and social justice.

The vices of Big Data

But while there are examples of AI being applied for moral purposes, it seems that the age of Big Data has

brought with it more vice than virtue. Instead of using data to prevent moral harm, many of our applications of data tend instead to undermine our most valued goods, including privacy, autonomy, justice and democracy. The recent phenomenon of “election meddling” illustrates the point. In both the 2016 and the 2020 US elections it was alleged that data applications were used in an attempt to manipulate the electoral outcome. Reportedly 19% of tweets relating to the 2016 US election were created by so-called “bots” (or automated Twitter accounts that share content)⁴. The use of AI makes these bots more difficult to distinguish from humans, and the use of these bots often has malicious intent – sowing division and ill will intended to drive voting behaviour.

¹ Aristotle, the father of virtue, provided three practical rules for good conduct. The second rule reads “We must notice the errors into which we ourselves are liable to fall, and we must drag ourselves in the opposite direction. (The Nichomachean Ethics, Book Two, ix)

² For more information, visit <https://www.aiforgood.co.uk/>

³ Roy, Nikolitch, McGinn, Jinah, Klement & Kaminsky. 2020. “A machine learning approach predicts future risk to suicidal ideation from social media data” in npj Digital Medicine. Available at: <https://www.nature.com/articles/s41746-020-0287-6> [Accessed: 2 July 2021]

⁴ Guglielmi, G. 2020. “The next-generation bots interfering with US elections”. Nature 582 (21). Available at <https://www.nature.com/articles/d41586-020-03034-5#ref-CR1>. [Accessed: 4 July 2021.]



Schalk Engelbrecht

Chief Ethics Officer
Risk Management

Tel: +27 82 713 7656

Email: schalk.engelbrecht@kpmg.co.za

Regarding the effects of bots and fake news, a recent study found that fake news/stories may only have a small influence on people's behaviour, but that these small effects could still cause serious harm. Think, for instance, of the potential harm that can result from only a small percentage of people opting against vaccination⁵.

What is clear, therefore, is that corporations and nations alike are not necessarily using data and AI to promote moral good. Instead, information gathering and prediction are used to get to "know" people better, with the express purpose of guiding them (sometimes by means of targeted false information) toward other-serving actions. The aim of knowledge becomes power, rather than virtue.

AI and insurance

How does this relate to insurance? Insurance is fundamentally about information and prediction. Information and prediction drives decisions regarding underwriting, claims, pricing, marketing and products. Among other uses, data helps insurance companies to determine risk more accurately; to make decisions about whether or not to accept risk; and, to judge whether to invest time or resources in the investigation of a claim. Big Data and AI are consequently being used to make insurance more efficient, to make decisions faster, and to help clients better.

While the increase in knowledge can be put to virtuous use – matching the right clients with the right products; lowering the cost of insurance; increasing the speed and efficiency with which clients are helped – it can also lead to less ethical outcomes, intentional or unintentional. The following related ethics risk stands

out: the erroneous use of data, bias, and opacity.

As AI and algorithms are increasingly used to make insurance decisions, we may rightly ask whether we can trust these decisions. When we employ the help of AI to decide how risky a policyholder is, or whether a claim is suspicious or not, can we be sure that a learning machine is not incorporating factors like gender or ethnicity into its determinations? This could lead to erroneous decisions, but may also lead to a violation of rights. It is possible that we intentionally or unwittingly allow bias into the insurance process.

Towards the responsible use of data in insurance

To counter the ethics risks associated with data and AI, the insurance industry must focus on increased transparency and on the cultivation of what philosopher of technology Shannon Vallor calls "technomoral virtue".

One of the problems of AI is that the algorithms used and the calculations performed happen behind the magician's curtain. We are impressed and astounded by the results, but we are seldom sure how the trick works. When the trick is not meant to entertain, but to determine whether I am insurable, what the cost of my insurance should be, or whether I am attempting to commit insurance fraud, then a good magician always reveals his secrets.

This is one of the key principles for the responsible use of data in the insurance industry, namely of human involvement, monitoring and transparency. Before an application is used in the insurance industry, companies must consider how best to explain AI outcomes

to customers. This kind of transparency lies in the extension of such treasured financial services values of *fairness, informed consent* and *autonomy*.

Secondly, human involvement is crucial in the use of data-driven technologies. When these technologies throw up red flags relating to underwriting or fraud risk, human governance is required to understand the basis on which these flags were raised. This also allows companies to improve these technologies – correcting them when flags are false.

If on the institutional and technological level transparency and human involvement are important, then on the individual level, design, operational and commercial decisions should be informed by so-called technomoral virtue.

We often ask which technical skills students will need to prepare themselves for future professions and employability. What we don't consider is that new industries and employment opportunities may require new moral skills. To steer technology towards moral purposes, for instance, data scientists working in financial services will need to develop new moral skills, tailored for the world of technological opacity, and the risks associated with data. Shannon Vallor suggests that the virtues that will be required include, among others, technomoral wisdom and empathy.

As an illustration, think of the use of AI to identify and counter claims fraud. To solve this problem, data scientists break up the problem in a myriad of smaller problems, that AI addresses individually. It requires wisdom and empathy for a data scientist to steer this process towards an accurate and fair outcome.

According to data scientist Eric Sibony:

It's crucial to understand that AI should not be trying to determine whether or not a claim is fraudulent. AI should be making decisions about various aspects of the claim to determine if they are suspicious or not. As such, algorithms must be designed to identify questionable behaviors and determine the extent of suspiciousness. Fundamentally, the claim is not suspect simply because of a policyholder's ethnicity or gender, or the neighbourhood in which they live. Algorithms must represent this fact⁶.

It is this kind of practical wisdom that insurance companies need to develop to prevent bias, and to ensure that the use of AI builds confidence in the industry, treats customers fairly, and benefits the insured.

Who will watch the robots?

Increased information and predictive power allow us to better know ourselves and our customers. It also enables us to pursue human and financial wellbeing more effectively. Through negligence or malice, however, Big Data and AI can also result in more opaque forms of discrimination and injustice.

One often gets the feeling of being watched by our technologies, especially when a targeted advertisement pops up pushing a product you remember mentioning casually to a friend. Data researchers dispel the conspiracy that Facebook is recording your conversations. Instead, AI is just getting better at predicting our wants. Adding a human element to data-driven technologies – cultivating ethics focused on the realities of AI applications – is a way for humans to "watch back".

⁶ Sibony, E. 2020. "Keeping the robots trustworthy: The ethics of artificial intelligence". Available at: <https://www.propertycasualty360.com/2020/11/03/keeping-the-robots-trustworthy-the-ethics-of-artificial-intelligence/?sreturn=20210528223245>. [Accessed: 5 July 2021]

⁵ Greene, C. 2021. "COVID-19: the first study to look at whether fake news actually changes people's behaviour" in The Conversation. Available at <https://theconversation.com/covid-19-the-first-study-to-look-at-whether-fake-news-actually-changes-peoples-behaviour-144819>. [Accessed: 5 July 2021]

