Building Ethical AI solutions: using the ethics funnel and a trusted framework

Based on the University of Waterloo's Center for Accounting Ethics and CPA Canada's 4th Ethics Symposium, "The Impact of Technology on Ethics, Professionalism and Judgement in Accounting" held April 25-26, 2019 in Toronto, Ontario

Automation, Artificial Intelligence (AI), robotics, blockchain, among other technologies could eliminate 30% of all jobs by the mid 2030s, **according to PwC**. This includes many accounting functions. For good reason: From automating rote tasks to reviewing complex contracts to using modelling and analysis to make judgments, AI is increasing efficiency, improving quality and in the process freeing up time for professionals to focus on roles that bring greater value to clients.

Perhaps the greatest role of the chartered professional accountant (CPA) is that of trusted advisor. But how do you build and maintain trust when a machine is undertaking more and more of the work? What are the threats and the advantages of relying on technology to create financial reports, assess and enhance business performance, provide tax services and audit financial statements? How will the public interest fare when Intelligent Agents and not humans are making the decisions? Who (or what) is accountable when mistakes are made?

CPA CHARTERED PROFESSIONAL ACCOUNTANTS CANADA

On April 25 and 26, 2019 in Toronto, the University of Waterloo's Centre for Accounting Ethics and Chartered Professional Accountants of Canada (CPA Canada) brought together thought leaders from academia and the profession to explore the intersection of technology and ethics and what it means for the way we do business, run our organizations and manage risk. Realizing the potential and benefits of AI will require creating a framework to ensure human-computer interactions actually help people. CPAs are in a unique position to pave the way forward by advancing an ethics by design approach to developing and deploying AI systems that demonstrate integrity, transparency and accountability.

The following is a summary of the symposium presentations and key messages.

Ethics and Applied Innovation: How can AI and machine learning act as ethical moral agents?

Imagine a world where everyone has a self-driving car. You go home at the end of the day, park in a garage and everyone's travel history gets uploaded into one centralized server. That is what we're working towards. "That's the power of AI and machine learning," said presenter Jeff Lui, CPA, Director in Deloitte's Artificial Intelligence Practice.

While self-driving cars may best encapsulate that power, a host of new AI use-cases are taking hold across industries and sectors. For example, digital music service Spotify is using machine learning and algorithms to gain a much deeper understanding of why we enjoy certain songs in order to make its recommendations. This goes well beyond genre and is much more personal.

Evolutions in natural language processing, computer vision and prediction/recommendation systems are advancing the march to singularity – the term for when computers surpass the human brain. It's anticipated we have 30 years to understand how to coexist with robotics and AI before singularity takes hold.

This is a critical time and it's important that we be thoughtful in moving forward and understand the implications of building machines that are better than humans in almost every way. In other words, just because you can build something, it doesn't mean you should. And even if you should, are you building it in a way that mitigates or eliminates bias? That protects privacy and the best interest of the public?

When exploring the use of AI, people should be considering what Lui describes as the Ethics funnel:

Justification: What do we build?
Not every new solution/product needs to be AI based. For example, The United Nations is currently grappling with the push from countries such as the U.S. and China to develop weaponized AI. While the U.S. wants weaponized AI ultimately controlled by humans, China wants fully autonomous weapons. The UN is tasked with creating a global framework that captures the spirit of the different approaches.

Ethos: Should we build it?
In 2015, toy maker Mattel released Hello Barbie, the first fashion doll that can have a two-way conversation with its users. By pressing a button, a child can speak to Barbie, Mattel uploads what the child is saying to its cloud server, processes it, and then Barbie issues an appropriate response. The company received immediate backlash, with concerns around who has access to the server, what Mattel is doing with the data and how it determines what an appropriate response should be. It appears no one at Mattel considered the implications.

Fairness: How do we build it?
It's important to understand what goes into datasets as there has been a history of data bias in our past. For example, since the early 1970s, every new car has had to undergo crash test dummy safety testing. What wasn't stipulated was the physiology of the dummy. As a result, a replica of a 200 pound man was used as the model. It wasn't until 2012 that a female physiology was also required for testing. The impact of this bias: if you are a female driving a car made before 2012 you are 50% more likely to be seriously injured or die in a car accident compared to a man in the same situation. While this example predates AI, bias continues to be a cause for concern. For example, if you do a Google image search for CEO, dozens of pictures of Caucasian males dressed in suits appear. At some point, the first female image was that of a CEO Barbie doll. How did this happen? Because of how Google designs its machine learning deployments. It scours the web for any image labeled CEO. Since the web represents the culture of society and the majority of CEO images are white men in suits, that is what appears. Imagine a child who doesn't know what a CEO is, goes onto the web to find out and this is the first image they see. We have to be conscious of this.

Safety: Are we building it safely?
Research has proved that commands can be hidden in high frequency sounds not detectable by humans but easily picked up by digital assistants such as Google Voice, Siri and Alexa. Known as Dolphin Attacks, these commands can activate these devices, telling them to open a malicious web page or pull private information.

"For the first time, we are asking computer scientists to solve much bigger problems than they are used to solving and we need everyone at the table – economists, accountants, anthropologists, lawyers – to have a say in how AI is designed and employed," said Lui.

How do you teach AI the Value of Trust?

A few key facts:

- 95 percent of companies have or are planning to adopt AI[1]
- 85 percent of AI projects will have erroneous outcomes due to bias[2]

1 www.ey.com/Publication/vwLUAssets/EY-Growth-Barometer-2018-Canada-Highlights/$FILE/EY-Growth-Barometer-2018-Canada-Highlights.pdf

2 www.ey.com/Publication/vwLUAssets/EY-Growth-Barometer-2018-Canada-Highlights/$FILE/EY-Growth-Barometer-2018-Canada-Highlights.pdf

- 85 percent of CEOs trust AI[3]
- 33 percent of employees trust AI[4]

AI is a prediction system and the reason we use it is because we have complex, incomplete information to develop accurate models. Still, no prediction is ever 100% accurate. "You should not expect AI to be perfect. What you should do is have the right measuring and monitoring systems in place to identify erroneous outcomes and mitigate the risk," said presenter Cathy Cobey, CPA, CISA, Ernst & Young, Global Trusted AI Advisory Leader.

This will require creating a trusted AI ecosystem, one that is based on:

- Ethics (Does it comply with ethical/social norms including corporate values?)
- Accountability (Is there a clear line of accountability to an individual and clarity on how it operates, the data that it uses and the decision framework that is applied?)
- Social responsibility (Did the AI design consider local and macro social impact, including the impact on the financial, physical and mental well-being of humans and our natural environment?)
- Reliability (Can it be trusted to perform as intended, not just during the initial training or modelling but also throughout its ongoing 'learning' and evolution?)

When creating a trusted AI ecosystem, it's important to consider that AI is designed to give the best answer based on the information it has; not necessarily the correct answer. And, perhaps more significantly, AI is only as good as its teacher. Its outcomes can be affected by poor training, bad data and bias. While the use of AI is poised to grow fast it will take time to mature and its processing speed and decision-making capabilities are quickly outpacing the monitoring and validation tools available.

EY has created a Trusted AI Framework[5] that identifies the five attributes necessary to sustain trust:

- Performance: AI outcomes must be aligned with and deliver stakeholder expectations.
- Transparent: End users must be notified they are interacting with AI.
- Explainable: Human operators must understand and have access to the AI's training methods and decision criteria.
- Resilient: The data used and the algorithm itself must be secure from attack.
- Unbiased: Inherent biases must be identified and addressed and the AI system designed to promote positive stakeholder impact.

In addition to following a trusted AI framework, businesses should also consider implementing leading governance practices when developing AI capabilities that may include:

3 assets.ey.com/content/dam/ey-sites/ey-com/en_gl/news/2019/05/ey-the-ai-race-barriers-opportunities-and-optimism.pdf
4 assets.ey.com/content/dam/ey-sites/ey-com/en_gl/news/2019/05/ey-the-ai-race-barriers-opportunities-and-optimism.pdf
5 www.ey.com/en_gl/digital/how-do-you-teach-ai-the-value-of-trust

- Forming a multidisciplinary, independent AI advisory board comprised of members from ethics, law, philosophy, technology, privacy, regulations and science who report to the Board
- Establishing AI ethical design standards
- Documenting an inventory of AI projects and corresponding impact assessments
- Creating AI validation tools
- Providing AI awareness training to employees
- Performing independent audits. IEEE is developing a certification system for AI focused on transparency, bias, and more.

The IESBA Code and Technology Initiative: An Update

In June 2019, the International Ethics Standards Board of Accountants (IESBA) revised and restructured code of ethics will come into effect, impacting some 3 million accountants worldwide. (Note: While the IESBA Code has not been adopted in Canada, ethics standards adopted by provincial bodies are created in alignment with the IESBA Code.) Renamed the International Code of Ethics for Professional Accountants (including international independence standards), the intent was to make the Code more clear and easier to apply and enforce. Among the key changes are new guidance on professional judgment and professional skepticism, the standard for non-compliance with laws and regulations (NOCLAR), and new provisions addressing pressure to breach Fundamental Principles (professional competence and due care, integrity, objectivity, confidentiality, professional behavior).

IESBA has also launched a Technology Initiative. A report with recommendations is to be presented to the IESBA Board in December 2019.

To date, five themes have emerged:

- Common ethical principles are applicable to AI
- Professional accountants need to have a growth mindset (one that adapts to new technologies)
- Bias in AI can be a significant risk
- Should we address fairness, transparency and accountability within the Code?
- As companies and organizations adopt new technologies, do professional accountants have an ethical responsibility to make clients aware of the risks and benefits? More broadly and significantly, should professional accountants promote ethical behavior in society above and beyond their professional activities?

Each of these themes has significant practical implications around the world and will lead to rule changes from the national standard setters, and ultimately change behavior of professional accountants.

How the Ethics of Privacy Relates to Accounting and the Audit: A Panel discussion with Jerrard Gaertner, President, Managed Analytic Services and adjunct professor at Ryerson University and Eric Cohen of Cohen Computer Consulting and co-founder of XBRL

Unlike Europe, which treats privacy as a fundamental human right, made explicit and protected by the General Data Protection Regulation, in Canada and the U.S., privacy is a moving target, mandated by statute that can be overturned. What's more, privacy, which, unlike concepts such as "confidential" that can be reasonably well defined without context, lives within situations, jurisdictions, cultures and social norms. How does any organization, then, come up with ethical standards with respect to privacy?

At the same time, the value of data is tied to how personal it is. The more anonymized and protective of privacy, the less attractive it is from a commercial standpoint. For the profession, the question becomes how might emerging technologies impact our ethical responsibility to ensure privacy and how do we build electronic trust? Realtime reporting and continuous auditing are the way forward but how do today's rules and codes of conduct – which currently do not discuss privacy – work in an environment of blockchain and the decentralized audit? As accountants, we are being challenged more and more.

The more we progress in computational power, the more difficult it is to anonymize something.

It's not linear. The increase in value of data is not linear, it goes up rapidly and because of that there is strong commercial incentive to accumulate data.

Throughout the symposium, PhD students from universities in Canada and the U.S. presented the following research papers:

- Are Individuals More Willing to Lie to a Computer or a Human? Evidence from a Tax Compliance Setting
- Black Box Analytics and Ethical Decision Making
- Complex Estimates and Auditor Reliance on Artificial Intelligence
- How Do Intelligent Agents Impact Managers' Aggressive Reporting and Ownership of Responsibility?

Call to action: AI is increasingly becoming the technology of choice across all sectors and industries to solve a wide range of problems and to create competitive advantage. This will only continue as the technologies evolve and advance. That doesn't mean everyone is on board or that AI is always necessary or safe. Critical to building trust and moving forward is an ethical approach. CPAs are in a unique position to lead and help define ethical principles and frameworks to guide AI design, development and deployment in a way that helps and protects the public interest.