



Ethical Issues in any Automated Decision-Making Model

By Neil Raden

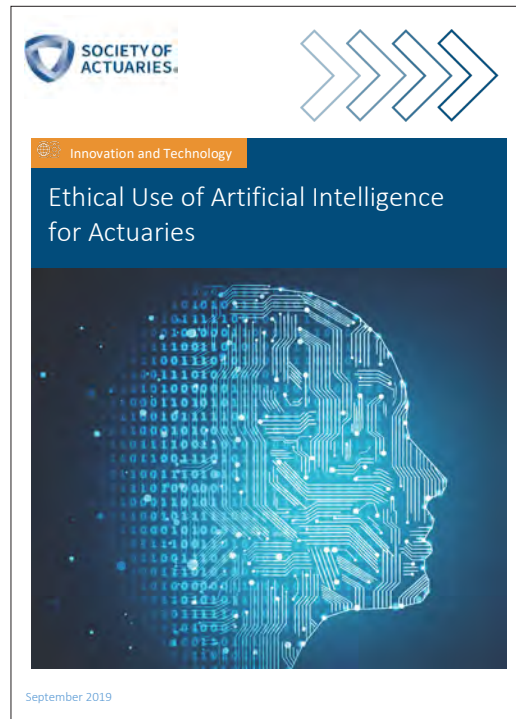
Editor's note: The views expressed in this article are solely those of the author and do not necessarily represent the views of the Society of Actuaries.

The pace of new technology creates difficult ethical questions for insurance companies. The accelerating use of unattended decision-making applications opens the door to risk of reputation and liability. AI and Machine Learning (ML) models can contain bias, provoke discrimination, intrude on privacy, and unwittingly violate regulations. Unlike your other applications, there is no code to follow when something goes wrong; AI models are devilishly difficult to explain. Risk arises from the skill of the team, the quality of the data and issues of placing algorithms into production systems.

There are a multitude of risks from the development process itself such as: unclear goals, improper and unprepared data, data produced by unregulated third parties, overly enthusiastic desire to fit a model, and overconfidence when the first few “proofs of concept succeed,” but plans to scale are incomplete.

In September 2019, the SOA released my research report, “Ethical Use of Artificial Intelligence for Actuaries,” to cover the essential concepts of risk, how AI works, how to form an AI team for actuaries and other topics, such as:

- AI, the social context and the five pillars of ethical AI;
- digital transformation, future technologies and InsurTech;
- ethical risk in general: data, bias and do-It-yourself risk;
- ethical aspects of automation;
- organizing the actuarial department for: skills, team, diversity;
- IT, AI engineer and actuary team;
- government and regulatory initiatives;



- advanced analytics types;
- review boards: discussion and examples; and
- path forward.

We presented our findings at the SOA Annual Meeting in Toronto, in an SOA podcast, at the Enterprise Risk Management Symposium of the Casualty Actuarial Society in Tampa, Fla. (I was not able to attend, but AI Klein of Milliman artfully presented) and in an SOA webinar and newsletter in June.

The report is available to download from our website; we developed the subject matter for actuarial organizations but found a wider audience of insurance organizations such as general insurance, reinsurers, and insurance industry professional organizations, all of whom find the material current and relevant. We know that many of the issues in this paper will resonate with you.

ISSUES THAT DEMAND MORE DISCUSSION

AI is moving so fast, and new ethical issues are apparent. It is time to review the subject, first by commenting on what has materially changed in the last few years, and what ethical issues have arisen.

Specifically, the White House Office of Science and Technology Policy (OSTP) announced **10 Principles for Stewardship of AI Applications**. These are important indications of the federal government's direction. Some states in the U.S. already put a stake the ground, not to mention the EU, and the Organization for Economic Co-operation and Development (OECD), an intergovernmental organization that represents 37 countries for economic concerns and world trade. OSTP probably shouldn't have bothered, as there was nothing new or even provocative in the 10 principles, and it has no teeth. If anything, it most resembled similar doctrines from China.

EIGHT AI ETHICS ISSUES TO CONFRONT

Ethical issue #1

How can organizations follow an ethical path with AI when the central government gives no guidance? The State of Washington just signed into law landmark legislation about facial recognition. According to the *WSJ*, "Washington state adopted a **Microsoft Corp**-backed law enshrining the most detailed regulations of facial recognition in the U.S., **potentially serving as a model for other states** as use of the technology grows." But should we entrust these issues to be addressed at a per-state level?

Ethical issue #2

Should a mega-tech company be writing legislation about a controversial AI application? Not everyone is comfortable with this, according to the *WSJ* article. Some feel the bill gives Washington State too much leeway. One provision allows police to use the technology without a warrant if "exigent circumstances exist."

We covered data bias in a report, "**Ethical Uses for Artificial Intelligence for Actuaries**," but it needs some explaining here. Data is an ethical problem, and always has been. Businesses should take every effort to minimize risks from data, especially when the data is from a third party, even *Data.gov* or *CDC.gov*, because data on its own has no context. How it was recorded and under what logic is missing when looking at a table.

There must be transparency around lineage, acquisition methods, and model assumptions, both initially and on an ongoing basis when the data is changing. There must be mandated security procedures to prevent loss from tampering and introduction of malware—all reinforced by comprehensive rights to audit, seek injunctive relief and terminate. The problem is that **data brokers** are mostly unwilling and unable to provide this.

Ethical issue #3

AI engineers, data scientists and predictive modelers crave new data. There is an aching desire to try all kinds of data to see if they can improve their models. The issue is that many data brokers are unscrupulous, and developers are wittingly or unwittingly poisoning their models with bad data. The problem

is even worse when the data is reliable but the motivations of the modelers are less than pristine.

A good example is the use of credit scores for underwriting personal auto insurance. While there is an undeniable correlation between poor credit and risk, the causal relationship is not the same. Poor credit is a function of many societal factors, it is not the driver. The working poor are forced to carry car auto insurance which is expensive, a regressive tax, and if they fail to make a payment, they are likely to get a ticket or have their vehicle impounded for hundreds of dollars, depriving them of the right to have a vehicle to go to work, to transport their kids to school, to travel to a decent grocery store, all exacerbating their situation.

Some decisions cannot be made by matching against known patterns. According to Vegard Flovik in "**How do You Teach Physics to Machine Learning Models**":

"If there is no direct knowledge available about the behavior of a system it is not possible to formulate any mathematical model to describe it in order to make accurate predictions."

In "**Deep Learning for Physical Processes: Incorporating Prior Scientific Knowledge**," Emmanuel de Bezenac adds that the most prevalent use of AI (outside defense and intelligence, where it is not possible to gauge its breadth) is in targeted selling. The reason is in selling applications en masse produces the cost of being wrong sometimes is almost zero. If you process 100,000 credit requests a day and get 1,000 wrong, you still made 99,000 correctly. There are other kinds of decisions that require a higher percentage of correct responses. Little decisions add up—those judged unfairly will notice at some point in time.

Ethics issue #4

The people building AI are not sophisticated enough to engineer-in domain expertise. Here is the big potato: job loss from automation. One school of thought is that most jobs have unseen complexities that currently require a human in the loop, such as different types of data a machine can't cope with, or the person who remembers birthdays with thoughtful presents. This subtlety and finesse is never described. Automation can only go so far as the AI engineer understands the job.

In many periods of realignment, organizations found that staff made redundant were responsible for many tasks that were never recognized. AI obviously cannot replace things it isn't aware of. However, learning AI watches and learns, and as time goes on, more work is done by the machine than the person—and you get a mix of human agents and cognitive robots working together. But as time goes by, the proportion of work done by the robot could increase and the human part could decrease.

Ethical issue #5

Good intentions are that the AI will augment workers not replace them. This overlooks the learning aspect of AI, and organizations may fail to plan for the situation where the employee actually becomes redundant. And why is all this on the employer?

Because AI as a machine doesn't have an ethical framework; we have to give it an ethical framework.

If you put in enough data at the right level of quality, the AI will eventually become very good at spotting a pattern, and can tell you about it. That may be good for making recommendations to buyers, but for more multiplex problems, the question of what to do next is complex. AI cannot, at this point, tell you what to do next. The only way is through modeling and simulation. Data never speaks for itself. With ML, the action is not learned. It is predetermined: "If you see this pattern, perform this action."

Ethical issue #6

Understand the limits of what the AI can tell you. Conway's Law: organizations which design systems are constrained to produce designs which are copies of the communication structures of these organizations. In implementing AI solutions, developers must be aware that people are diverse and complex and live within groups and cultures. AI is not like coding. In many cases, there is no coding at all. There is nothing to examine for potential bias. As a result, data selection and labeling, feature engineering, model development and review all reflect the attitudes and belief of the group. No diversity in the group leads to insensitiveness to those affected.

Ethical issue #7

It is too easy to be lulled into exposing personal information. In fact, it is too easy for bad actors to snatch personal data when you're not looking. Federated learning is a powerful idea for distributed applications and data, <https://ai.googleblog.com/2017/04/federated-learning-collaborative.html>, but first movers have chosen medical data as a testbed. Tread carefully.



Ethical issue #8

Using AI ethically ought to reflect that diversity is essential. "Fairness" isn't uniform; there are different versions of it. The emergence of Federated Learning, on the one hand, has positive implications for privacy, but on the other, is likely to exacerbate the explainability issue.

This is hardly a complete list—so it will be a recurring series. ■



Neil Raden is an active industry analyst, consultant, speaker and author of "Ethical Use of Artificial Intelligence for Actuaries." He can be contacted at diginomica.com/author/neil-raden.