



Chartered
Insurance
Institute

Standards. Professionalism. Trust.

Research report

Addressing Gender Bias in Artificial Intelligence



Contents

- 5. Key findings
- 7. CII introduction
- 8. Artificial intelligence and the rise of machine learning
- 9. How insurers are using artificial intelligence
- 10. Trends
- 13. The types of impacts that could result
- 14. What about Test Achats?
- 15. What should insurance do?
- 16. The role of professionals
- 17. Conclusions
- 18. Appendix

About us

The Chartered Insurance Institute (CII), is a professional body dedicated to building public trust in the insurance and financial planning professions. Our 127,000 members commit to high professional standards.

We work with businesses to develop bespoke, company-wide solutions that ensure competitive advantage by enhancing employees' technical and professional competence.

Success in CII qualifications is universally recognised as evidence of knowledge and technical expertise.

Our members are able to drive their personal development and maintain their professional standing by adhering to our Code of Ethics and by accessing an unrivalled range of learning services.

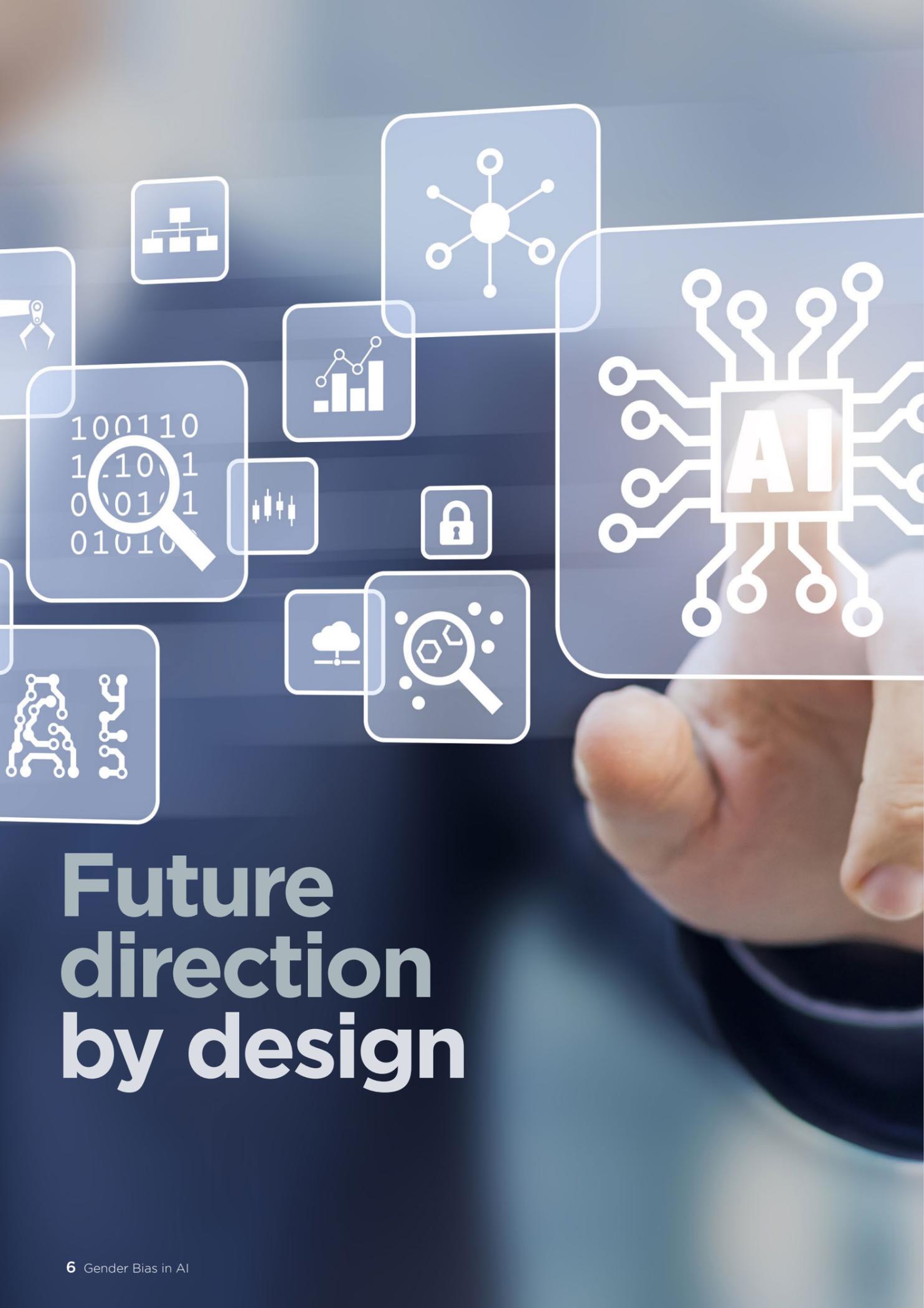
Find out more at: cii.co.uk

Addressing Gender Bias in Artificial Intelligence



Key findings

- Artificial intelligence is transforming insurance, changing the way in which underwriting, claims and marketing is carried out. AI tools such as algorithms and machine learning are increasingly tapping into the 'data exhaust' of our daily lives.
- Algorithms learn by being trained on historic data. More and more of that data is now unstructured, coming from text, audio, video and sensors. Yet engrained in that historic data are decisions based upon historic biases, particularly around gender.
- Recent research confirms widespread gender bias in datasets used to train algorithms. And techniques used by algorithms to generate new insight mean decisions influenced by gender will continue, despite compliance measures for Test Achats. This will create gender based detriment in underwriting, claims and marketing decisions.
- Other forms of bias, such as in relation to race, have been identified as well, but the focus of this Thinkpiece is on gender bias.
- Insurance firms need to prepare a structured response to this issue, starting with visible leadership on the questions it raises. Such a response needs to begin with partner evaluation, go into detail in the planning and testing stages, and then address monitoring and oversight.
- Individuals need to rise to their professional obligations and translate the fifth core duty in the CII's code of ethics into tangible actions to address the risk of algorithmic bias in their line of responsibility. Their technical knowledge and experience should be used to address the tough questions that this issues raises.



Future direction by design

CII introduction

Artificial intelligence is often talked about as the future direction of insurance. Yet it is not free from controversy, with accusations of gender bias often raised. New research has now shed light on the extent of this key challenge.

For insurance firms introducing exciting new techniques around artificial intelligence, the challenge is to both reach for an innovative future and not fall short on their ethical responsibilities. This timely research report explores the issue and shows individual professionals and their firms how they can address it.

The Chartered Insurance Institute is the world's largest professional body for insurance and financial services and is the leader in awarding qualifications to industry practitioners. With over 125,000 members in more than 150 countries, we are committed to protecting the public interest by guiding professionals towards higher technical and ethical standards. Our research reports are a key part of our ongoing commitment to promoting innovative thinking and debate within the profession as well as among wider stakeholders. cii.co.uk

The views expressed and phraseology used within this research report are those of the author and should not be interpreted as reflecting the views or policies of the Chartered Insurance Institute. The author has asserted his right under the Copyright, Designs and Patents Act 1988 to be identified as the author and copyright owner of the text of this work, and has granted the CII worldwide perpetual licence to reproduce and distribute it in whole and in part.

Artificial intelligence and the rise of machine learning

Insurance is starting to be transformed. What we understand to be risks, how underwriting is carried out, how covers are distributed and how claims are handled: each of these things is being questioned and redefined.

This transformation is being powered by the massive amounts of data being produced as we go about our daily lives. The way in which we shop, work, travel, and socialise are all producing this 'data exhaust' of our lives.

Data is now 'big data' and it's getting bigger, but that on its own achieves little. Insurance people want to be able to generate meaningful, actionable insight from it. And for that, they are turning to a variety of tools that fall under the general heading of what is called 'artificial intelligence'.

"Artificial intelligence means... computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision making and translation between languages."

Artificial intelligence gives insurers the capacity to make sense of all that data. And in particular, it allows insurers to keep on top of a key trend. Traditionally, insurance has relied on structured data for underwriting and claims insight, such as age and address. Now however, most of that 'big data' is being produced in the form of unstructured data, through text, audio, video and sensors.

Unstructured data will soon dominate the data that insurers rely on. And those who draw most insight from all that unstructured data will go on to be the insurers who could dominate the market in the future. This quote neatly sums this up:

"Information is the oil of the 21st century and analytics is the combustion engine."

Artificial intelligence encompasses a range of tools and one of the more significant is something called machine learning. As an ever greater volume and variety of unstructured data is exploited, so the capacity of humans to create so called 'hand written' algorithms to analyse this data is slowly but surely reached. The solution lies in the algorithms being taught how to analyse this unstructured data for themselves. The machines must learn.

"An algorithm is a structure of mathematical formulae for identifying relationships within data."

So how do you teach an algorithm? How do you move from algorithms written by humans, to algorithms that learn for themselves? The answer lies in the historic data you use to train them.

If you want an algorithm to learn what a kitten looks like, you have it analyse thousands of pictures of kittens. And if it's still not sure, you just give it many more thousands of pictures of kittens to analyse, until it gets it right. Patterns are then recognised, interpreted and remembered for next time.

This creates an interesting reliance. In order to look forward, machine learning has to look back at all that historic data. Think of those recommendations you often see on retail websites. The algorithms behind those recommendations have learnt something about you from the choices you've made when visiting that and other websites in the past.

"Artificial intelligence means... computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision making and translation between languages."

"Information is the oil of the 21st century and analytics is the combustion engine."

"An algorithm is a structure of mathematical formulae for identifying relationships within data."

How insurers are using artificial intelligence

Insurance firms are starting to put artificial intelligence to use in a wide variety of ways, covering underwriting, claims and marketing. Let's look at each of these in turn, some examples from underwriting.

Underwriting people are using AI to...

- expand the range of risk factors being used to underwrite risks (some motor insurers use over a 1,000 rating factors now)
- reduce the number of questions needed to provide a quote to only a handful (for example, a US insurer will provide you with a health insurance quote in just 42 seconds, by asking only 5 questions)
- reduce the through put time from quotation to policy fulfilment down to a matter of seconds
- streamline the assessment of overall portfolios in order to identify exceptions and predict emerging opportunities and risks
- deliver new types of products in new forms of relationship with consumers, often basing these around advice on risk reduction and mitigation.

Innovations like these will help to remove a lot of the friction that consumers often experience when buying insurance. The result could be increased satisfaction ratings for a sector that has often struggled to win public approval. Remember though that this is often being achieved through complete or significant automation of decision making processes.

Claims people are using artificial intelligence to:

- reduce the time to settle claims that are small, simple and standard
- analyse images to produce quick and accurate estimates of repair costs
- identify with greater confidence those claims likely to be fraudulent
- analyse process flows to identify bottlenecks and inefficiencies
- test and improve settlement strategies for different categories of own and third party claims
- improve reserving strategies and expense management.

Again, increased automation presents opportunities to make the policyholder's claims journey shorter, more responsive and informative.

Marketing people are using artificial intelligence to:

- explore social data to understand consumer identities and their potential insurance needs
- use that same social data to understand customer sentiments about different products and processes
- personalise the interactions with consumers in order to offer them highly tailored products
- predict consumer behaviours so that sales opportunities are identified and exploited
- introduce interactive behaviours that stimulate those sales related opportunities.

This more personalised marketing relies on the automated decision making at the heart of many artificial intelligence tools.

Trends

Three Key Trends

Three key trends are worth noting. Firstly, that artificial intelligence introduces a vastly more automated level of decision making. Secondly, artificial intelligence is paying attention not only to what we say and do, but also to how we say it and the feelings and sentiments that drive those opinions and actions. It's getting underneath our emotive skin.

And thirdly, while artificial intelligence is turning a lot of this unstructured data into structured data (for example, those social media comments that point to you being born in 1971), it is also assembling a lot of that unstructured data into virtual identities for each of us. Characteristics are inferred from clusters of correlations that the algorithms have found in all that unstructured data.

Take the now famous case of the women who shopped in the US retailer Target and then received coupons for pregnancy and baby products. She had only just found out that she was pregnant and had told no one about it. Target didn't have a field in its database to indicate whether the customer was pregnant or not: its marketing algorithms inferred this from how her shopping decisions created a cluster of correlations about that personal characteristic.

And insurers are now going beyond data about what we buy, to include data about what we say on social media. UK insurer Admiral wants to use algorithms that identify the personal characteristics of consumers from what they say, and how they say it, on Facebook.

Using Historic Data to Make Decisions

Let's bring in again that earlier point about how algorithms are evolving from being hand-written by humans, to being self taught through machine learning. The algorithms are being self taught through being trained on huge amounts of historic data. And a lot of that historic data will be unstructured data.

That huge amount of unstructured historic data is in effect a picture of how we as a digital society have been living our lives over the past 10 to 15 years. This covers our decisions, preferences, choices, opinions and actions. And while most of these will be our conscious decisions, the cleverness of artificial intelligence lies in being able to identify the unconscious element that often underlies those conscious decisions. And the power of artificial intelligence allows it to track such decisions and preferences in the billions, rather than the hundreds or thousands.

So, for example, an algorithm that looks at who is and who isn't recruited for a job will be able to identify in the huge number of recruitment decisions not only who did get the job, but the unconscious factors that were influencing those overt decisions. And it does this by picking up the correlations woven into those many millions of recruitment decisions. It creates a form of digital couch.

You can see then the immense potential of artificial intelligence to explore the depths of all those past decisions and opinions making up that historic data. Yet this is a double-edged sword, for if you train an algorithm on historic data, it will learn not only the good decisions we have made, but also the bad ones as well. It will learn the biases in society.

Recent Research on Gender Bias

Research published in the April 2017 edition of the journal 'Science' illustrates this¹. Researchers looked at a machine learning tool known as 'word embedding' that is transforming the way computers interpret speech and text. Word embedding works by building a mathematical representation of language, in which the meaning of a word is distilled into a series of numbers based on which other words most frequently appear alongside it.

The researchers found that words for flowers were clustered close to words linked to pleasantness, while words for insects were closer to words linked to unpleasantness, reflecting common views on the relative merits of flowers and insects. This relative merit would then be learnt by the algorithm and remembered for use in later decisions.

They then looked at words like 'female' and 'woman' and found them being more closely associated with arts and humanities occupations and with the home, while 'male' and 'man' were more closely associated with maths and engineering occupations.

The obvious question from this was of course the extent to which such word clusters reflected outcomes that have actually occurred in the real world. The researchers found a strong correlation with the percentage of woman in 50 occupations in the USA in 2015.

Now you may be thinking that this is just about words, when insurance is much more about numbers. Well, think of those algorithms crawling over social media looking for words that could influence underwriting and claims decisions. And remember that data is intertwined with words, such as in the labels attached to the categories it is segmented into, and how the significance of data is interpreted and expressed.

The Lessons that Algorithms Learn

What does this research tell us? It tells us that artificial intelligence, with its algorithms taught on historic data, is going to learn the gender biases ingrained in so much of our historic data and then propagate this into the automated decisions that insurers will be making in underwriting, claims and marketing. The algorithmic machines will learn societal biases.

Examples of such specifically algorithmic bias are hard to come by in insurance. They are however surfacing in interesting ways elsewhere. Consider the use of online translation software.

The researchers found that AI powered translations to English from gender-neutral languages such as Turkish led to gender stereotyped sentences. For example, Google Translate converted these Turkish sentences with gender neutral pronouns "O bir doktor; O bir hemsire" into these English sentences: "He is a doctor; she is a nurse".

Replacing Turkish with Finnish, Estonian, Hungarian and Persian produced the same result. Similarly, translating the above two Turkish sentences into several of the most commonly spoken languages (Spanish, English, Portuguese, Russian, German and French) resulted in gender stereotyped pronouns in every case. This is an issue for both female doctors and male nurses.

Searches on Google for images of "working women" have been found to turn up lower rates of female executives and higher rates of women in telemarketing, in contrast to the relative numbers of women who actually hold such jobs.

I said earlier that we are not yet seeing direct evidence of algorithmic gender bias in insurance, but what we are seeing is evidence of gender bias amongst decision makers on AI in insurance. A recent article on virtual agents and chatbots in insurance referred to four customer facing 'agents' as Sofia, Iris, Allie and the authors' own Cathy. Such naming patterns reinforce gender stereotypes.

¹"Semantics derived automatically from language corpora contain human-like biases" by Aylin Caliskan, Joanna Bryson and Arvind Narayanan. In 'Science', 14 April 2017. Vol 356, Issue 6334, pp 183-186.





Redefining not repeating

The types of impacts that could result

Let's foresee the types of impacts this could have.

Research published through the CII's 'Insuring Women's Futures'² programme shows that women tend to feel marginally more risk averse, and less financially secure, than men. That 'marginally more risk averse' aspect could lead to women paying more in premiums, through the now established practice of price optimisation.

And that 'less financially secure' aspect could lead to women receiving less in claims settlement, through the emerging practice of claims optimisation. Both of these optimisation practices rely on the systematic analysis of vast amounts of data.

Women are also more likely to report that they lack knowledge relating to financial decisions and to want information and advice. Such information and advice is becoming increasingly automated through artificial intelligence, using tools such as chatbots. A chatbot will tap into historic data and social chat data and from it, learn how to guide the consumer to achieve the goals that have been set for it. Many sectors use chatbots to maximise sales, tailoring the guidance being given to the susceptibility of the consumer to respond to certain emotional prompts. Many females already experience such techniques, through for example, online advertising for cosmetics peaking at those times that social chat data indicates women are most susceptible to such messages.

Can we expect insurers not to follow such examples? Their track record on selling to those less knowledgeable about their products is not a great one, so it is a possibility that shouldn't be dismissed.

That aforementioned research into AI's 'word embedding' feature has two implications for insurers. Firstly, the increasing use of AI in recruitment could result in gender distortions in employment decisions. And in core operations such as underwriting and claims, females in roles that algorithms associate more often with males could find their policies being underwritten differently, or their claims looked at more closely.

You may wonder though about just how significant some of these impacts will be. They will of course vary in scale from insignificant to significant, and in volume from occasional to widespread. That's missing the point though. Gender bias in analogue insurance decisions is illegal; gender bias in digital insurance decisions is also illegal.

The digital transformation of insurance that artificial intelligence is powering to new levels has to work within the same legal framework as everyone else. As the UK's Information Commissioner once told insurers: "big data is not a game played by different rules".

The perspective you take also matters. The impact of gender bias on any one individual may seem small, but on a group basis, it is significant. Insurance firms should be addressing discrimination not just on a per person basis, but on a group and societal basis as well.

² "Risk, exposure and resilience to risk in Britain today - Women's Risks in Life - an interim report" by Chartered Insurance Institute. December 2016.

What about Test Achats?

The word springing to the minds of many of you at this point will be Test Achats, the Belgium consumer group whose test case on gender in underwriting moved the tectonic plates of insurance practice. Surely insurers wouldn't go against the law in again taking account of gender in their underwriting?

Let's be clear: it would be surprising if any insurer in the UK was directly and systematically using gender in their underwriting or claims decisions in ways that were illegal. Is that the end of the story then? Not at all.

Think back to the case mentioned earlier, about the US retailer Target and how it used shopping patterns to detect consumers to whom it could market pregnancy and baby products. Target never collected data specifically showing that any particular customer could be pregnant. Instead, their predictive algorithms had learnt from historic purchase data that the customer was likely to be pregnant and had targeted her for marketing. Her purchase decisions had in effect manufactured a new piece of information about her and learnt what to do with it.

Artificial intelligence is often talked about in terms of its capacity to reveal new and startling insights about the relationships between different parameters. One tool behind this is called correlation clustering, where the focus is more on the relationship between the data objects, rather than the actual representation of the objects themselves. Out of such analysis comes a piece of 'manufactured information': in this case, that there was a high probability that the customer was pregnant.

Imagine then how the algorithms within a firm's artificial intelligence might then 'manufacture' a variety of attributes about a customer from a pattern of activity found in disparate sources of data. Those manufactured attributes wouldn't always be stored as another piece of data, but 'understood' by the algorithm. It is quite possible that from the alignment of those patterns of activity with relatively generic profiles of risk, loss or fraud, there could emerge decision patterns that could be discriminatory.

So there is a real possibility that gender bias could become ingrained in the underwriting and claims systems of the future. And it could be ingrained in ways that are far less discernible than at present. As algorithms that machine learn from vast lakes of unstructured data become the norm, then the basis upon which underwriting and claims decisions are made becomes more and more opaque. And as the focus of those decisions moves from the data objects themselves to the relationship between those objects, then the manufactured information that is thus brought into being starts to become increasingly significant to the outcomes that consumers experience.

The challenge that this then presents is of 'black box' processes learning to fulfil the expectations set for them, without sufficient consideration being given to the ethical considerations that the insurance industry and the public take for granted.

The detriment this causes then never emerges onto the busy radars of business. It blends into the normality of busy firms using complex systems. Complexity may indeed be the new norm, but that doesn't mean that gender bias should be so too.

So what should insurance do?

Let's be honest: insurers should not be surprised to learn that gender bias exists in the historic data making up the 'data exhaust' of our lives. It's been known about for years. What is changing is the increasing amounts of research into how this gender bias might influence the algorithmic decision making that insurers are introducing.

And just as significantly, such research is also producing similar findings in relation to other equally serious concerns, such in relation to race and disability. Gender is but one bias in historic data.

Insurance has not so far been on the radar of such 'algorithmic bias' studies, although there are signs of attention being given to how robo-advice operates. The insurance sector therefore has a window of opportunity to learn and respond to the dangers of gender bias in the algorithmic systems increasingly being used for underwriting, claims and underwriting. How specifically might firms respond then? Here are ten suggestions...

1. adopt a three level response, looking at the data itself, at the algorithms being used, and the practices for managing and overseeing it. This is the emerging approach for addressing complex questions of data ethics.
2. be clear about where you stand, again at three levels: that of the individual manager or executive, that of the firm and that of the sector overall. And then make that stand visible, with colleagues, throughout your firm and across the sector. The risk is that no one moves on this issue for fear of having 'first mover disadvantage'.
3. find a means of building a consensus for action. This could be done around the fifth core duty in the CII's Code of Ethics and then developed into the adoption of some common principles. The 'Principles for Algorithmic Transparency and Accountability' adopted by US and EU sector policymakers earlier this year might serve as one model (see appendix).
4. debate the issues and the significance of the challenge they represent, both for insurers and for consumers. This will require some visible leadership coming forward from the sector on this issue. And such leadership needs to have a pretty powerful voice, given the rapt attention that insurers are giving to the various tools of artificial intelligence.
5. look at the processes your firm is using to assess and implement artificial intelligence into the decisions it makes. Introduce algorithmic impact assessments from the planning stage onwards, within which gender issues could be a standing component.

What should insurance do?

6. introduce a monitoring framework to ensure that legal and regulatory obligations relating to gender, race and disability are respected. Include an independent pair of eyes into that monitoring and ensure that it is familiar with current and emerging trends in artificial intelligence.
7. adopt tools for testing historic data for evidence of any discriminatory features. And use techniques such as fairness data mining to check that the algorithmic outcomes are non-discriminatory.
8. ensure that those managing and overseeing artificial intelligence in your firm receive regulator training in the ethical issues and developments associated with it.
9. remember to see the forest as well as the trees. Gender discrimination needs to be addressed not just on an individual basis, but on a group basis as well.
10. make sure that your business partners and suppliers join with you in this initiative. This means refusing to include any form of 'hold harmless' clause in agreements for data or artificial intelligence. And it also means carrying out thorough due diligence on those you work with.

These steps might seem daunting to some firms, but remember that they fit closely with the requirements of the Senior Managers and Certification Regime (SMCR) that the Financial Conduct Authority will be extending to the insurance sector in 2018. Algorithmic trading has already been identified with a 'significant harm function' in the SMCR regime under which banks now operate. Can insurers assume that their use of algorithms in core operating systems will be treated differently? I would say not.

And the last of those ten points above, about business partners and suppliers, also has a SMCR dimension. The SMCR allows for responsibility to be delegated, but not accountability. Those with responsibility for the functions that artificial intelligence is transforming, and who are identified as such on the firm's responsibility map, will need to remember this. After all, the standards that are to be upheld are still those of the regulated insurer, even when the business is being transformed.

The role of professionals

Insurance is a sector built upon a strong sense of professionalism. So how might professionals, both individually and collectively, respond to the issues raised by algorithms and gender bias?

As indicated in point 3 ethical standards to which insurance professionals in the UK are held are set out in the CII's Code of Ethics, the fifth core duty of which says:

"Treat people fairly regardless of: age, disability, marriage and civil partnership, pregnancy and maternity, race, religion and belief, sex, sexual orientation and transgender."

And as indicated in point 9, this core duty and the public interest obligation within which it sits, applies to 'people' on both an individual and group basis. In other words, the issue should be addressed at the level of both the tree and the forest.

So, against this backdrop, what should insurance professionals do? Here are some suggestions...

- first and foremost, they need to be clear about what is expected of them. One way to do this is to take that core duty relating to equality and fair treatment and express it in more specific terms in relation to the dangers of gender bias in the algorithms their firm is using. This would help set the expectations of what is to be achieved.
- collectively, the professionals within any one insurance firm need to give their active and visible support to initiatives that respond to that risk of gender bias in their algorithms. They need to demonstrate a collective commitment on a defining issue for their profession.
- they need to pressure for positive change. For example, it would be ironic if insurers in the UK were to meet their regulatory requirements on gender pay reporting for their firms (due in 2018), without at the same time having something to say on what they're doing to tackle gender bias within their firm's algorithms. What must be avoided at all costs is the 'black box disconnect', whereupon the gender pay gap in insurance is eliminated, while a gender bias in underwriting, claims and marketing decisions spreads uncontrolled, perhaps even unnoticed.
- professionals need to show personal leadership on this issue, explaining the issue to colleagues, asking questions at meetings, calling out vague or non-existent commitments, and setting a personal example. And by professionals, I mean all gender identities, not just female.
- Look at the business partners and suppliers with whom your firm is working on artificial intelligence projects. How diverse are those firms, or the team you're working with? How does that compare with your firm/team? Ask them what their firm is doing to reduce the risk of gender bias in the type of work they'll be doing for your firm. And ask to see the results of that work. It's the actuality of evidence that counts.

Conclusions

It is vital that the ambitious transformation of insurance that artificial intelligence is bringing about is built upon foundations that are both solid and trustworthy. Should algorithms and machine learning simply absorb the mistaken views of the past and perpetuate them into the future, then the validity, value and veracity of that transformation will rightly be questioned.

Gender bias exists in historic data and insurance firms' use of artificial intelligence must actively address it. There are tools to do so, there are leaders in a position to address it, and there are institutions through which this can be coordinated and monitored.

The main hurdle is most likely to be 'first mover disadvantage', perhaps through being later to the market than some other competitors. Does that mean the 'ends' might be imperilled at the cost of the 'means'? It's happened before with other issues. Let's hope lessons have been learnt.

This transformation of insurance will create losers as well as winners. Expect many more of the former than of the latter. So what will differentiate them? One will be customer perceptions. Can I trust this firm? It wants to get closer to me, but do I want to get closer to it?

In early 2017, the advisory firm PwC published the results of a survey it had conducted across the insurance market. It found that 28% of insurance CEOs were "extremely concerned" that trust would affect their firm's growth, while 72% of insurance CEOs thought it will be harder to sustain trust in a digitised market³. Trust matters because it delivers growth.

Tackling gender bias is both a fundamentally right thing to do, as well as a sensible, trust building move. The long-term winners of this insurance transformation will be those who consumers trust. Tackling bias should be woven into every insurer's trust agenda.

Learning Objectives

Having read this Thinkpiece, readers should be able to:

- understand the role that artificial intelligence is starting to play in the digital transformation of insurance
- recognise the means by which gender bias can be introduced into artificial intelligence tools like algorithms and machine learning
- identify key steps that individuals and firms can take to address the issue of gender bias in artificial intelligence.

CPD Reflective Questions

1. where within an insurance firm should responsibility for issues like gender bias lie, in relation to its artificial intelligence programme?
2. should individuals and firms wait until the regulator issues guidance on how they should deal with issues like gender bias?
3. should individual professionals focus on gender bias in particular, or should they consider biases around race and disability as well?

³ PwC survey, "20th CEO Survey" by PricewaterhouseCoopers LLP, 2017.

Appendix

The following principles were adopted by the US Public Policy Council and the Europe Policy Committee of the Association for Computing Machinery in early 2017, as part of their policy statement on algorithmic transparency and accountability.

Principles for algorithmic transparency and accountability

1. Awareness: Owners, designers, builders, users, and other stakeholders of analytic systems should be aware of the possible biases involved in their design, implementation, and use and the potential harm that biases can cause to individuals and society.

2. Access and redress: Regulators should encourage the adoption of mechanisms that enable questioning and redress for individuals and groups that are adversely affected by algorithmically informed decisions.

3. Accountability: Institutions should be held responsible for decisions made by the algorithms that they use, even if it is not feasible to explain in detail how the algorithms produce their results.

4. Explanation: Systems and institutions that use algorithmic decision-making are encouraged to produce explanations regarding both the procedures followed by the algorithm and the specific decisions that are made. This is particularly important in public policy contexts.

5. Data Provenance: A description of the way in which the training data was collected should be maintained by the builders of the algorithms,

accompanied by an exploration of the potential biases induced by the human or algorithmic data-gathering process. Public scrutiny of the data provides maximum opportunity for corrections. However, concerns over privacy, protecting trade secrets, or revelation of analytics that might allow malicious actors to game the system can justify restricting access to qualified and authorized individuals.

6. Auditability: Models, algorithms, data, and decisions should be recorded so that they can be audited in cases where harm is suspected.

7. Validation and Testing: Institutions should use rigorous methods to validate their models and document those methods and results. In particular, they should routinely perform tests to assess and determine whether the model generates discriminatory harm. Institutions are encouraged to make the results of such tests public.

The Chartered Insurance Institute
42-48 High Road, South Woodford,
London E18 2JP

tel: +44 (0)20 8989 8464

customer.serv@cii.co.uk
cii.co.uk

 Chartered Insurance Institute

 @CIIGroup

© The Chartered Insurance Institute 2017

COH_J012575