



Financial Reporting Council



Government
Actuary's
Department

The use of Artificial Intelligence and Machine Learning in UK actuarial work



October 2023

Authors

Sam Davies FIA

Sam Kinshuck

Chris Paterson FIA

Government Actuary's Department

Contents

1. Executive summary	4
1.1 Method	4
1.2 Extent of AI/ML use	4
1.3 Governance and quality assurance	6
1.4 Risks from use of AI/ML	7
2. Introduction	10
2.1 Background to the research	10
2.2 Research scope	10
2.3 Survey methodology	10
2.4 Interview methodology	12
3. Extent and nature of AI/ML usage among actuaries	14
3.1 Extent of use	14
3.2 Trends in the extent of AI/ML use	16
3.3 Areas in which AI/ML tends to be applied	17
3.4 Variety of techniques used	21
3.5 Expanding use of Large Language Models (LLMs)	22
3.6 Distribution of AI/ML expertise across teams	24
4. Governance and quality assurance	25
4.1 Robustness of governance	25
4.2 Governance and quality assurance practices for complex AI/ML models	27
4.3 Communication with decision makers	28
4.4 Reproducibility	30
4.5 Consideration of regulation in governance and quality assurance process	31
5. Explainability	32
5.1 Importance of explainability when using AI/ML techniques	32
5.2 Ease of explainability for AI/ML models	32
5.3 Different approaches to explainability	33
6. Data privacy	34

7. Increased risks from use of AI/ML techniques	35
7.1 Overview of risks of use of AI/ML techniques	35
7.2 Lack of actuaries with AI/ML skills	36
7.3 Bias and discrimination	36
7.4 Lack of human oversight and automation	37
7.5 Poor data quality	38
7.6 Over-reliance on AI/ML analysis	38
7.7 Insufficient holdout data	40
7.8 Lack of supporting infrastructure	40
7.9 Other risks identified in survey responses	40
Appendix A – Technical terms	41
Appendix B – Use of AI/ML techniques in different areas in actuarial fields	45

Liability

No party accepts any liability for any loss, damage or costs howsoever arising, whether directly or indirectly, whether in contract, tort or otherwise from any action or decision taken (or not taken) as a result of any person relying on or otherwise using this publication or arising from any omission from it.

© The Financial Reporting Council Limited 2023

The Financial Reporting Council Limited is a company limited by guarantee.

Registered in England number 02486368.

Registered Office: 8th Floor, 125 London Wall, London, EC2Y 5AS

1. Executive summary

This report presents the findings from a survey and a series of interviews conducted by the Government Actuary's Department (GAD) on behalf of the Financial Reporting Council (FRC) to explore how Artificial Intelligence and Machine Learning (AI/ML) are used for actuarial work in the UK.

1.1 Method

The research involved an online survey that was completed by 104 respondents across different fields and 20 interviews (37 interviewees) in the first six months of 2023. Interviews with more than one person covered a single organisation, except in one case.

The surveys and interviews were targeted at UK actuaries who use AI/ML in their work, although contributions from non-UK actuaries or individuals who do not use AI/ML were also welcomed and made up a small minority of survey respondents and interviewees. More General Insurance actuaries completed the survey (c44%) and attended the interviews (c38%) than actuaries from any other field. The higher survey response rate from General Insurance actuaries and evidence from the interview stage suggests that there is a concentration of AI/ML work in General Insurance.

1.2 Extent of AI/ML use

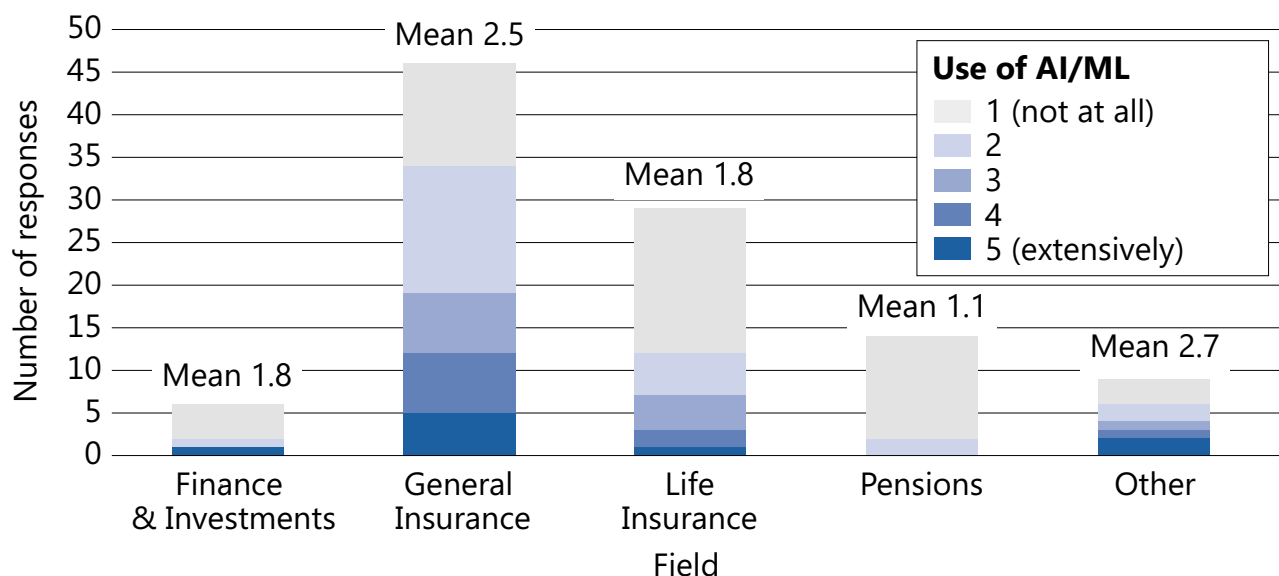
The main use of AI/ML techniques in UK actuarial work relates to insurance pricing, particularly in General Insurance, with use being more limited in other areas. There are, however, signs that the use of AI/ML is expected to increase in future.

The research suggested AI/ML use is currently limited among actuaries across the profession but revealed a concentration of AI/ML activity among General Insurance actuaries, particularly for the purposes of General Insurance pricing. The uses of AI/ML cover a wide range of applications within General Insurance pricing, including determining claims risk for policyholders, forecasting price-elasticity of demand for policyholder groups, as well as informing the 'front- end' process for customers and policyholders.

A minority of respondent actuaries work outside of the more established actuarial fields of Pensions, General Insurance, Life Insurance and Finance & Investment. This group tended to use AI/ML to a greater extent than other respondents, applying more advanced techniques in more diverse use cases. Examples of AI/ML use in the 'Other' fields of actuarial work that were referred to included analysing the impact of public health interventions, assisting pharmaceutical companies with patient profiling, and developing long term economic projections.

Whilst, at the time of conducting the research, the use of AI/ML techniques in actuarial work was somewhat limited, both the survey and interviews found its usage was increasing rapidly. Research participants from organisations that were already using AI/ML extensively were more likely to report stronger intentions to increase use in future than those that were using it less.

Figure 1 – Use of AI/ML by actuarial field



The plot is based on responses to the question: “To what extent do you use Artificial Intelligence and Machine Learning for work carried out by you or teams that you manage?”. Answers were on a scale from 1 (‘not at all’) to 5 (‘extensively’), split by actuarial field. The mean score across all responses and all fields was 2.1. Within each field the mean score ranged from 1.1 in Pensions to 2.5 in General insurance and 2.7 in ‘Other’ fields. Among survey participants AI/ML was therefore used least in Pensions, and most in General Insurance and ‘Other’ fields.

There was a wide range of complexity in the AI/ML techniques being applied. Simpler methods included basic decision tree classification analyses and Generalised Linear Models (GLMs), which some participants classified as AI/ML; more advanced techniques were also mentioned, including deep triangle and extreme gradient boosting. A glossary of technical terms can be found in Appendix A.

The timing of this research meant that ChatGPT (a Large Language Model, or ‘LLM’, developed by OpenAI) grew from having no mention in survey responses (January to April 2023) to being raised by almost all interviewees (March to June 2023) as a relevant topic for extensive discussion. Some interviewees’ organisations were making use of LLMs to aid efficiency of some processes. At the simpler end LLMs were being used to assist programming, horizon-scanning or summarising large volumes of text. At the most advanced end, participants were developing bespoke in-house LLMs trained on internal documents to build entire financial models of their own, based purely on natural language prompts. Interviewees noted that the integration of LLMs into businesses had allowed some organisations to make significant gains in productivity.

1.3 Governance and quality assurance

Governance and quality assurance processes are generally being informally adapted for models using AI/ML techniques.

Most participants said their organisations have not changed their formal governance processes for AI/ML, because these processes are effective, but noted that there were often differences in governance practice to focus on ensuring analysis is explainable to actuarial reviewers, model owners, senior decision makers and regulators. This is widely referred to as “explainability”. Participants noted that many of the additional governance processes applied when using more complex AI/ML techniques were geared towards enhancing explainability in addition to ensuring accuracy of outputs.

There were, however, a small number of participants who described a different governance process being used for AI/ML work. This included requiring higher levels of sign-off and more thorough documentation, as well as having different expectations on data protection and IT security. Participants also gave examples of how they had considered existing regulations in adapting their governance processes.

Participants indicated that actuarial practitioners with the skills and knowledge to review AI/ML work are in short supply, particularly in pensions or finance & investment. Some organisations overcame this by shaping AI/ML model outputs to align with metrics that were more familiar for practitioners with limited data science experience and were developing interactive tools to aid understanding.

Participants shared a range of interpretations of the concept of reproducibility. Regardless of which is applied, reproducibility is, however, often not a high priority in respect of governance and quality assurance.

The research suggests that explainability is a key factor in the choice of modelling techniques and tended to be a greater challenge for AI/ML models than for established modelling techniques. The most referenced increased risk to actuarial work by using AI/ML techniques was a lack of model understanding due to models being treated as ‘black boxes’.

Participants typically thought explainability could be more challenging for most AI/ML analysis compared to traditional techniques. Participants mentioned organisations prioritising explainability to internal and external stakeholders when making modelling choices. Examples given included organisations opting for very simple forms of AI/ML model or building models in a way that ensured the underlying rationale for every stage of the model’s decision-making process was transparent and amenable to statistical interrogation. Participants explained how more engaged senior decision makers can overcome explainability issues. They gave examples where deploying techniques including more advanced data visualisations or SHapley Additive exPlanations (SHAP) analysis – which helps to shed light on how different input variables contribute to a model’s predictive output – were often helpful in enabling decision makers to understand the inner workings of AI/ML models.

Participants highlighted that a key risk with an AI/ML model is its calculations not being readily visible or understandable due to its functioning as a 'black box'. There were a range of views on the ease of achieving explainability for models using AI/ML techniques, depending heavily on the type of AI/ML analysis employed. Some considered decision tree models as easy to explain, and some suggested that neural networks could be inherently unexplainable.

1.4 Risks from use of AI/ML

A key risk highlighted by those using AI/ML techniques was that of bias or potential discrimination, either as a result of the modelling techniques used or bias in the underlying data.

Participants were concerned that the risk of inadvertently discriminating based on protected characteristics was higher when using AI/ML models than was the case with more traditional techniques. Participants also suggested the risk of gender-based pricing discrimination could be higher when using AI/ML models due to the generally lower level of explainability within an AI/ML model and the level of technical information provided to decision makers.

A range of views were offered on how to address this issue. One organisation stated they had a mathematically watertight and legally defensible method for proving that their models did not discriminate on the basis of gender.

Participants also noted the risk of inadvertent bias and potential discrimination due to the potential for inherent bias in datasets which were made more useable by AI/ML techniques. For example, over-representation of certain groups in input datasets could lead to AI/ML models in population health management recommending public health interventions that favour over-represented groups.

Other areas of increased risks to actuarial work were said to potentially arise from using AI/ML techniques including poor quality – or inappropriately sourced – data, insufficient human oversight, over-reliance on results and data privacy.

Interviewees noted that the ability of AI/ML models to use new and more extensive data sources, such as unstructured text, indirectly increases the magnitude of potential data privacy concerns. One interviewee based outside the UK explained that some companies had begun using "alternative data" for insurance pricing, for example using a policyholder's smartphone app usage data or shop visit records to derive a better-informed assessment of their risk profile.

Risks associated with data labelling were mentioned, such as the burden of the data labelling process and the potential for inaccuracies. Interviewees also expressed reservations about potential biases or systematic inaccuracies in the data used to train AI/ML models, leading to flawed patterns of judgement and low generalisability of models to different contexts.

Participants mentioned risks arising from a lack of human oversight in deploying AI/ML models, in particular the full automation of LLMs. While participants saw human involvement as fundamental to the mitigation of mistakes that LLMs can make, there was concern about the effectiveness of human involvement given that humans also make mistakes. Participants explained that AI/ML is sometimes seen as a “magical” solution to analytical problems and that there is a risk of some actuaries being overly optimistic about the kinds of problems it can solve. Participants suggested that actuaries may rely on these techniques excessively without understanding their limitations, or carrying out appropriate validation. This was a particular worry for LLMs, where the cases of errors or inaccuracies, while rare, are very difficult to predict.

Participants took the view that data privacy concerns are of equal relevance for both AI/ML and more traditional analysis, as AI/ML relates primarily to how data is processed, not the type of data that is processed. However, AI/ML techniques greatly expand the content and diversity of the datasets that can be fed into subsequent modelling stages, meaning they indirectly lead to more data privacy concerns.

Lack of consistent language in relation to AI/ML work may hamper ongoing communication and understanding of modelling issues, or the management of risks associated with them.

Within this study a definition was provided for the terms AI and ML. At the time of writing, there is as yet no commonly accepted definition¹, and this has the potential to hamper understanding among individuals working in this area, and the development of appropriate governance or regulatory approaches. For example, there were varying views on whether generalised linear models (GLMs), which are considered to be a ‘traditional’ actuarial technique, would be classified as an ML model.

Equally, participants had varying interpretations of terms such as ‘explainability’ and ‘reproducibility’. With explainability, in particular, being seen as a key issue, participants emphasised the importance of a common understanding of its characteristics among actuaries developing models, as well as those using or relying on the results of such models.

Discrepancies in the understanding of key terms regarding AI/ML not only introduce technical challenges, but also potential governance considerations. For instance, some interviewees noted that their organisation had a distinct governance process for AI/ML analysis that did not apply to more traditional forms of analysis. The decision as to whether to classify a given technique as AI/ML could therefore have practical implications with respect to the level of governance to which associated analysis is subject.

¹ Some definitions have, however, been proposed. See, for example, definitions that appear in the [Turing Institute’s Data Science glossary](#), which was referenced in a [recent IFOA risk alert](#) about the use of AI techniques.

The recent proliferation of Large Language Models (LLMs) has had a rapid and potentially significant impact on actuarial work. This highlights the risk of adopting new technical advances when they may not be fully understood.

Interviewees noted that while LLMs can perform some tasks with a high level of accuracy, it is difficult to predict when, and under what circumstances, they will fail to do so (even if this only happens rarely). This in turn makes it difficult to mitigate against the associated risks. Interviewees also highlighted that many LLMs may create a risk of data leakage because in some cases the LLM owner may have the right to share the prompts it receives with third parties; this increases the risk that sensitive information could fall into the wrong hands.

2. Introduction

2.1 Background to the research

AI and ML have become increasingly integrated into the workings of the global economy, including within the actuarial profession. As well as potentially enormous benefits to productivity and innovation, these technological advances can also bring substantial new risks when applied to actuarial work.

The FRC commissioned GAD to conduct research into the use of AI and ML in UK actuarial work through surveying and interviewing professionals in the field. This report presents the findings from the research, aiming to provide a clearer picture of how AI and ML are used, both now and in the future.

No statement in this report relating to any rule, policy, or regulatory approach should be interpreted as an endorsement or lack thereof, by either GAD or the FRC. Any discussion of regulation in this report is intended solely to present the views of the research participants on the topic. Nothing in the report should be seen as indicative of GAD's or the FRC's opinion on the regulation of AI/ML use among actuaries.

2.2 Research scope

Our working definition of AI and ML refers to techniques that allow computers to learn from data without being explicitly programmed. It involves algorithms that can adapt and improve over time, learning from experiences (in the form of data inputs) to predict outcomes or make decisions. AI/ML algorithms identify patterns and relationships within data, which can then be used to predict future data or inform decision-making processes.

By "actuarial work" we mean work that involves the exercise of judgement, and for which the principles and/or techniques of actuarial science are central. For example, using predictive analytics to estimate future mortality rates would be 'actuarial work', whereas developing a chatbot to interface with customers would not.

Recruitment of survey participants and interviewees was heavily focused on the UK, although there were some contributions from non-UK-based actuaries. The research therefore cannot be used to draw any inferences about practices or market dynamics in relation to actuarial use of AI/ML outside of the UK.

The research addressed the use of AI and ML for actuarial work. This included use of AI and ML at any stage of data collection or analysis that will be an input to actuarial work.

2.3 Survey methodology

Research was first carried out through an online survey targeting UK-based actuaries, engaging a total of 104 participants. While the focus was on actuaries who were already using AI and ML in their work, professionals who did not currently use these technologies were also welcome to respond. Respondents were recruited via the following approaches:

-
- Website and social media announcements by GAD, the FRC and the Institute and Faculty of Actuaries (IFoA).
 - Articles about the research in The Actuary and The Actuarial Post.
 - Emails and posts to LinkedIn groups of actuaries volunteering for the IFoA on relevant working parties, and committees, or those subscribed to relevant IFoA interest groups. This included IFoA members who had completed the IFoA's Certificate in Data Science.
 - Communication to alumni of universities offering actuarial science or similar degrees.
 - Publicising the research at relevant IFoA webinars.
 - Targeted approaches to actuaries with relevant experience, for example if they had published relevant articles or spoken at conferences.
 - Targeted approaches to actuaries employed by the largest UK actuarial employers.
 - Direct approaches to contacts of GAD's staff.

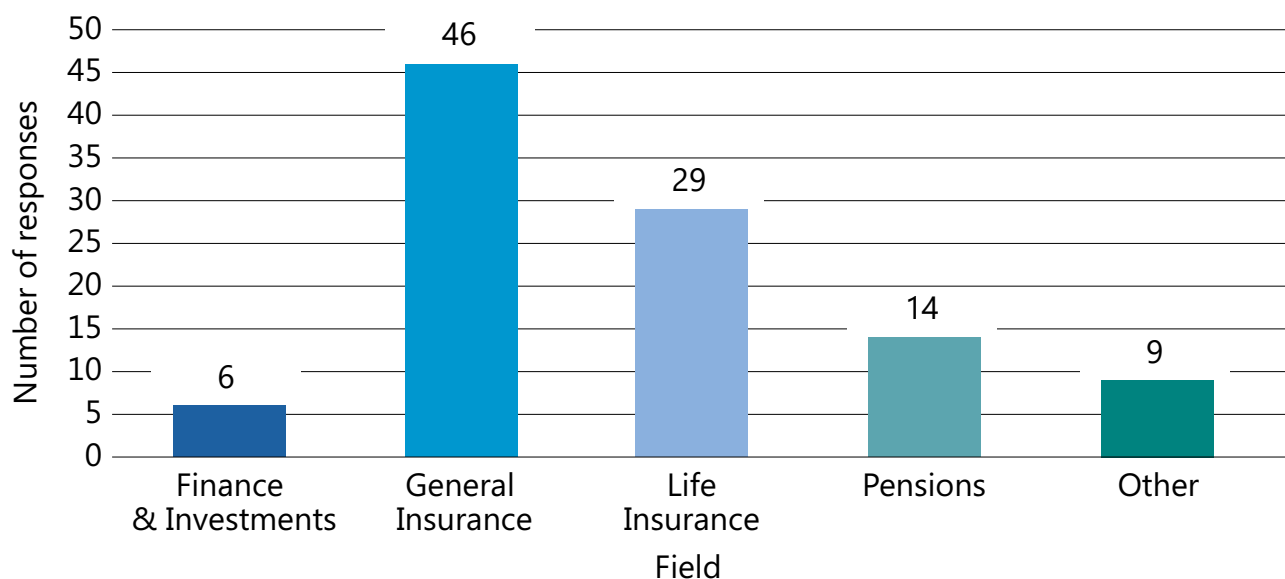
The survey was open from January to April 2023. It comprised a mix of multiple-choice and open-text questions aimed at understanding the extent and way respondents and their organisations use AI and ML. Most questions were identical for all respondents. A small subset of questions contained minor variations depending on the respondent's self-reported actuarial field. The survey questions can be found [here](#).

Figure 2 shows the number of survey respondents from different actuarial fields. Given analysis of the timing of survey responses, we believe that the higher number of respondents in certain fields – particularly General Insurance – reflected the greater extent of AI/ML use in those fields as opposed to a preference for General Insurance actuaries in the recruitment of survey participants. We also verified that the survey provided a good representation of people with knowledge about AI/ML use in the specific sub-areas of these fields. Some actuarial fields employ more actuaries than others, meaning that proportionate representation across fields would not necessarily have translated into equal numbers of respondents in each field. Indicative statistics on the distribution of actuaries by field² show that there is a very small number of actuaries working in Investment, which could partly explain the low response rates in this area.

All figures presented in this report are based on data from the survey unless otherwise stated in the figure caption.

2 <https://www.actuarialcareers.co.uk/profession-overview/areas-of-work/>

Figure 2 – Survey completions by field



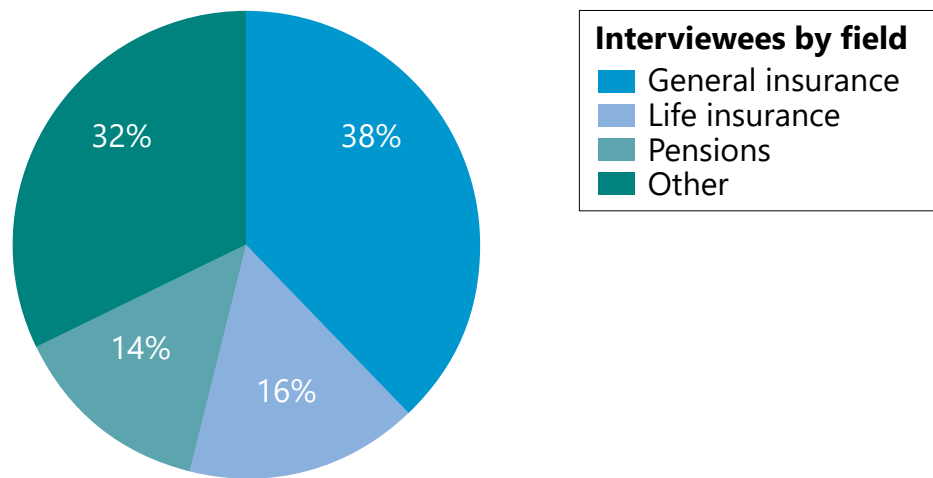
2.4 Interview methodology

Most interviewees were recruited via the survey, which had an option to be contacted for an interview. The remainder were recruited by targeting larger actuarial employers in the UK, with initial contact being made through the survey recruitment process.

To enable a more thorough exploration of some of the key themes emerging from this research, the survey was supplemented by 20 in-depth interviews with actuaries. Each interview involved between one and three interviewees, with 37 interviewees in total, covering 21 organisations. Interviews with more than one person covered a single organisation, except in one case.

In most interviews, at least one attendee had already completed the survey. Interviews were conducted in a semi-structured manner using a defined list of questions. Interviewees were free to steer the discussion towards the topics that they thought were most pertinent. Interviewers were able to prompt for clarification or to request further details when points of particular interest were raised. Interviews were conducted between March and June 2023.

Figure 3 – Distribution of interviewees by field



Note that these fields are used to be consistent with those provided by interviewees in the Survey which most interviewees completed. This chart is based on self-reported categories, and so interviewees working for Reinsurers and Lloyd's of London syndicates may be included under General Insurance or Other. There were no interviewees in the field of Finance & Investment.

3. Extent and nature of AI/ML usage among actuaries

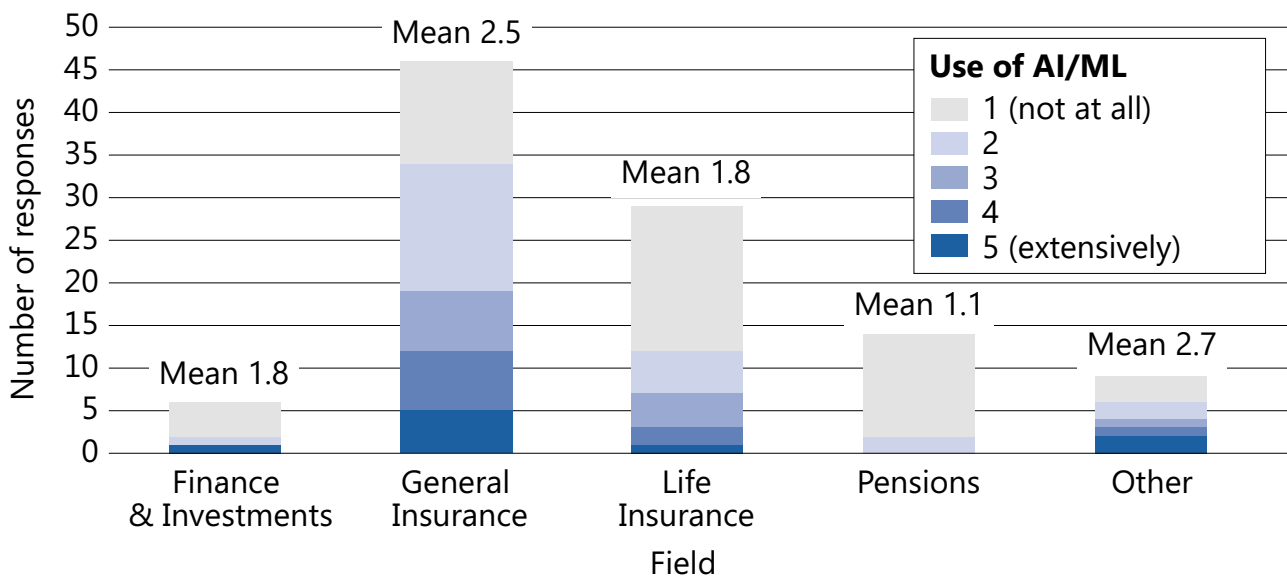
3.1 Extent of use

As shown in Figure 4, the mean rating of the use of AI/ML by survey respondents' organisations on a scale from 1 ('not at all') to 5 ('extensively') was around 2.1, though there were substantial differences across actuarial fields. The mean response to a similar question about organisational plans to increase the use of AI/ML over the coming years was much higher, at 3.6 (see Figure 5).

'General Insurance' actuaries and 'Other' actuaries rated the extent of their organisations' AI/ML usage substantially higher than all the other traditional actuarial fields, with 'Pensions' actuaries rating the extent of their usage lowest.

Despite the low reported usage in Pensions, these respondents said that they use AI/ML to some extent for "data collection and cleaning", even though all but two answered the earlier, more general, question about AI/ML usage by saying their organisation does not use AI/ML at all.

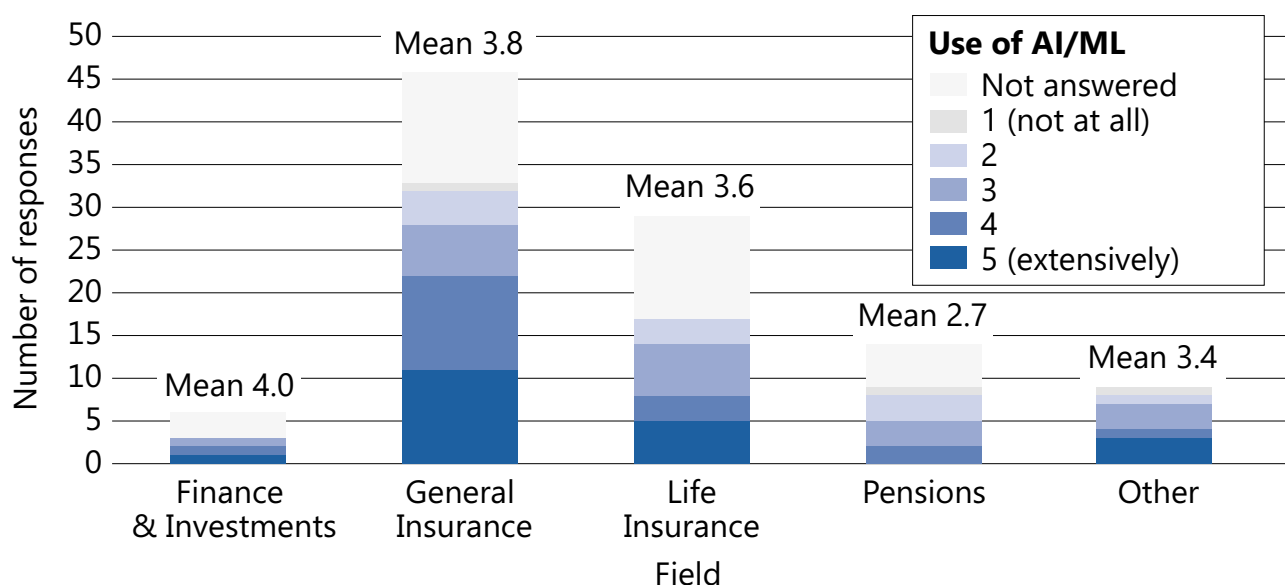
Figure 4 – Use of AI/ML by field



The plot is based on responses to the question: "To what extent do you use Artificial Intelligence and Machine Learning for work carried out by you or teams that you manage?", which was answered on a scale from 1 (not at all) to 5 (extensively). Responding to this question was mandatory.

Figure 5 – Reported plans to increase use of AI/ML

How much will the organisation increase AI/ML based analysis over the next 5 years?



The plot is based on responses to the question:

“To what extent does your organisation plan to increase its use of Artificial Intelligence and Machine Learning-based actuarial analysis within the next 5 years?”

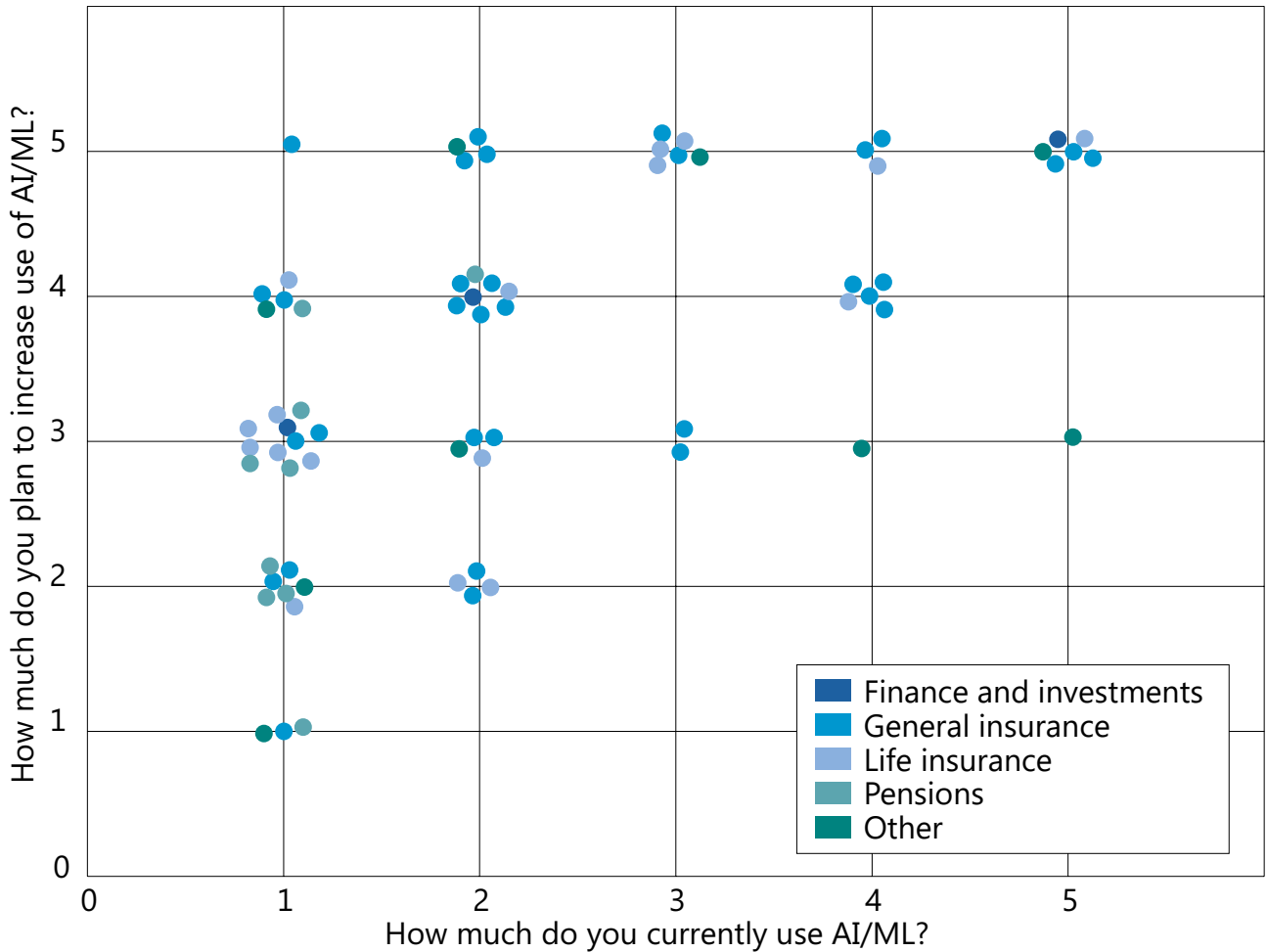
Interviewees’ statements were generally consistent with the survey in terms of the fields and areas in which use of AI/ML was seen as being most common. Interviewees generally echoed the survey findings that AI/ML use among actuaries was limited but undergoing rapid advancement, and that usage was concentrated among General Insurance actuaries and even more so in General Insurance pricing. The participants of ten interviews said the use of AI/ML among actuaries was undergoing a rapid expansion, with the participants of one interview pointing out that this was the source of a key risk because the growing sophistication of techniques being used increasingly threatened to outpace actuaries’ expertise.³ One interviewee added that use of AI/ML was highly advanced in the insurance industry in the (non-UK) country where their organisation was based.

³ This reflects the findings from a 2020 survey of actuaries in which respondents generally agreed that there is increasing pressure on actuaries to upskill to keep abreast of recent data science advances: <https://www.actuaretech.com/insights/the-evolving-role-of-the-actuary>

3.2 Trends in the extent of AI/ML use

Figure 6 below shows the relationship between current levels of AI/ML use and intentions to increase its use over the subsequent five years. The figure shows that respondents who already had a high usage of AI/ML expressed a greater intention on the part of their organisation to increase its use in future.

Figure 6 – Current use of AI/ML compared to plans to increase its use in the future.



The data in the plot are based on the questions: "To what extent do you use Artificial Intelligence and Machine Learning for work carried out by you or teams that you manage?" and "To what extent does your organisation plan to increase its use of Artificial Intelligence and Machine Learning-based actuarial analysis within the next 5 years?" A small amount of random noise has been superimposed on the positions of the dots to prevent them from obscuring each other.

Interviewees who reported a high current level of AI/ML usage also tended to describe more detailed and ambitious plans about how they would expand its applications in the future.

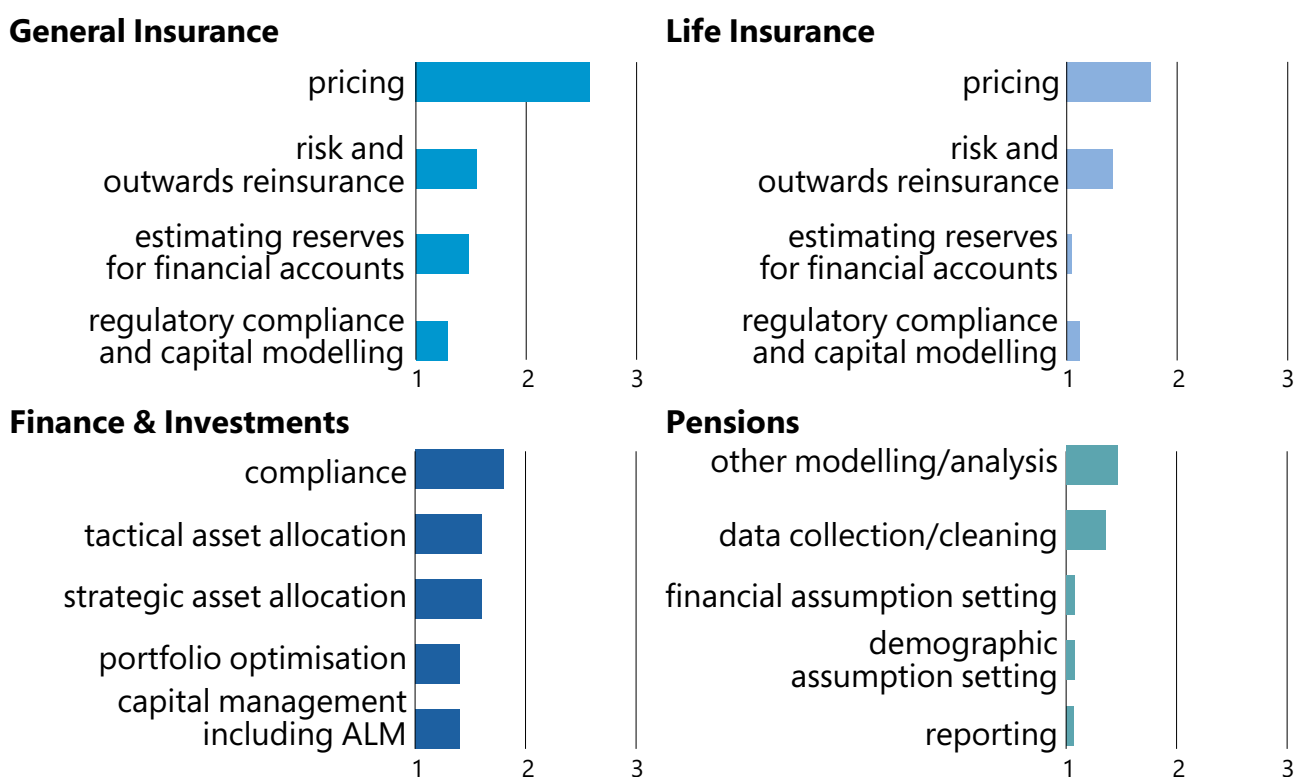
3.3 Areas in which AI/ML tends to be applied

The survey and interviews highlighted patterns in the specific areas in which actuaries, at the time of this research, tended to apply AI/ML techniques, including:

- AI/ML was being primarily used for the purposes of pricing, especially among General Insurance actuaries. Within pricing, the interviews and survey responses revealed multiple different applications for AI/ML that were increasingly being implemented.
- There were several less traditional areas in which actuaries were applying AI/ML techniques, often to a greater extent than in areas of work more closely linked to actuarial science.
- AI/ML was employed slightly more frequently by General Insurance and Life actuaries in risk and reinsurance relative to some other areas.

Figure 7 illustrates some of these findings from the survey – the responses in many areas were very low, with General Insurance pricing being the only area where the mean extent of use was above 2 (on a scale of 1-5 where 1 is 'not at all' and 5 is 'extensively'). A more detailed breakdown of the kinds of techniques that were being used in different areas of different fields can be found in Appendix B.

Figure 7 – Use of AI/ML in different areas within fields



This plot is based on responses to the question: "To what extent is Artificial Intelligence and Machine Learning used in each of the following areas within your organisation?" This question was asked to all respondents except those with a field of 'Other'. Respondents were asked to provide answers on a scale from 1 (not at all) to 5 (extensively) for each area within their field.

When asked to describe how they used AI/ML in each area, the survey responses from General Insurance and Life actuaries were more detailed in relation to applications for pricing than for other areas. 15 General Insurance and 6 Life actuaries discussed the use of AI/ML to determine the risk profiles of different policy holders in order to set 'back-end' prices⁴. Respondents also talked about applications of AI/ML across broader aspects of pricing-related decisions, such as forecasting demand for different insurance products, predicting the likelihood of policy termination following price increases⁵ or predicting the revenue implications of pricing decisions. There was also one mention of using AI/ML to determine market-wide pricing benchmarks.

The reported patterns of usage across different areas matched the findings of prior research in this area which also found that, among Life actuaries, AI/ML was mainly used for the purposes of pricing and claims analytics.⁶

Some of the other ways in which respondents used AI/ML are presented in Table 1.

Table 1 – Reported use cases of AI/ML from the survey. These data are based on the following questions:
 "Please briefly describe the type of analysis that your organisation uses Artificial Intelligence and Machine Learning for in each of the following areas...", which was following by open-text inputs alongside labels denoting all of the areas linked to the respondent's field. This question was presented differently to 'Other' respondents, who were asked about AI/ML use cases in general with no area-based breakdown.
 "Please describe any other areas where Artificial Intelligence and Machine Learning is used for in your organisation that are not included in your actuarial work previous answers?"

Use described	Fields	Areas
Determining policy holder risk profiles to set back-end prices	General Insurance (GI) (15 mentions) Life (6 mentions)	Pricing (21 mentions)
Forecasting price-elasticity of demand for different policyholder groups	GI (3 mentions) Life (2 mentions)	Pricing (5 mentions)
Setting market-wide pricing benchmarks	GI (1 mention)	Pricing (1 mention)
Detecting fraud	GI (4 mentions) Other (1 mention)	Miscellaneous (4 mentions) Risk and Outwards reinsurance (1 mention)

4 Back-end prices in this context was taken to refer to calculating the costs of policies before commercial considerations were applied

5 We note that any AI/ML modelling carried out to test the price-sensitivity of policyholders should comply with the FCA General insurance pricing practices market study PS21/15 and PS21/11.

6 [Surveying skillsets: AI deployment within life insurance, The Actuary](#)

Use described	Fields	Areas
Projecting inflation	GI (2 mentions)	Estimating reserves for financial accounts (1 mention) Regulatory compliance and capital modelling (1 mention)
Predicting 'Incurred But Not Reported' claims	Life (1 mention)	Estimating reserves for financial accounts (1 mention)
Identifying patterns in reserving diagnostics to determine which should be prioritised for actuarial review	GI (1 mention)	Estimating reserves for financial accounts (1 mention)
Reading in and processing unstructured data	GI (3 mentions) Life (1 mention)	Risk and Outwards Reinsurance (3 mentions) Miscellaneous (1 mention)
'Assessing risk' or 'estimating losses' (minimal details provided)	GI (3 mentions) Life (2 mentions)	Risk and Outwards Reinsurance (3 mentions)
Scraping news articles and assessing negative sentiment to serve as an 'early warning signal'	Finance & Investment (1 mention)	Portfolio Optimisation (1 mention)
Forecasting portfolio returns	Finance & Investment (1 mention)	Strategic Asset Allocation (1 mention)
Targeting marketing at individuals deemed to be good prospective clients	Finance & Investment (1 mention)	Tactical Asset Allocation (1 mention)
Scraping and analysing content from news articles to inform horizon scanning exercises	Pensions (1 mention)	One individual mentioned this use case for both Data Collection/Cleaning and for Other modelling/analysis
Increase digitalisation of data (no further detail provided)	Pensions (1 mention)	Data Collection/Cleaning (1 mention)
Analysis of survey responses, assessing credit risk	Pensions (1 mention)	Other modelling/analysis (1 mention)
Performing sentiment analysis on organisation's social media pages	Pensions (1 mention) Other (1 mention)	Other modelling/analysis (1 mention) Miscellaneous (1 mention)
Various uses including using a range of advanced AI/ML techniques	Other (6 mentions)	Miscellaneous (6 mentions)

Use areas for AI/ML in 'Other' fields

The use cases described by respondents in the 'Other' field were more detailed and appeared to be more complex than use cases in more traditional actuarial fields. This was consistent with higher mean reported use and usefulness (relative to traditional actuarial approaches) of AI/ML provided by these respondents.

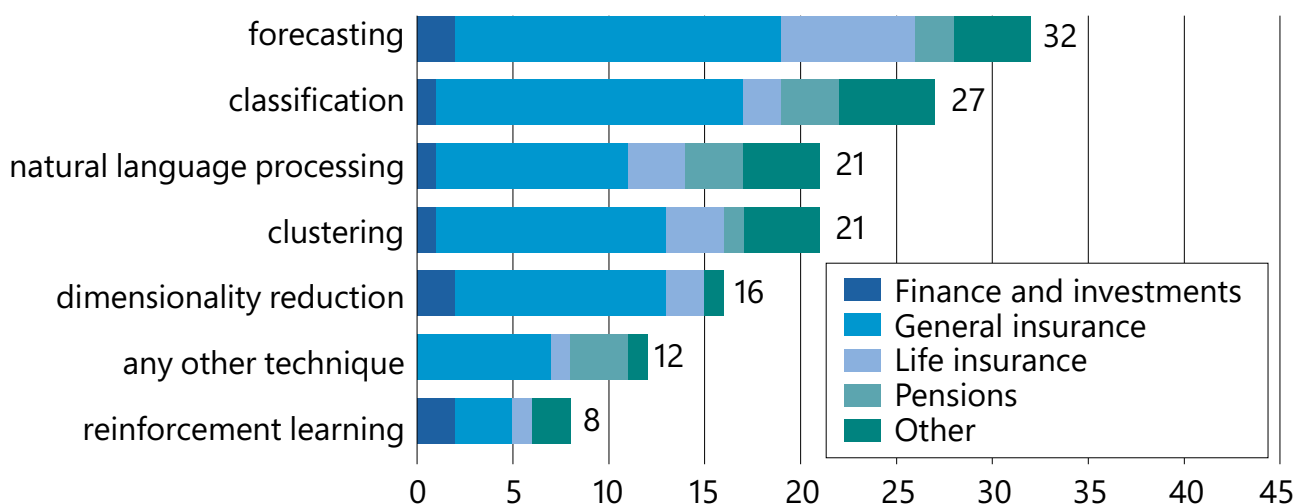
Healthcare was represented by two survey respondents and in two interviews. Use areas mentioned included analysing the impact of public health interventions, biometric experience analysis, and advising pharmaceutical clients on the treatment cost profiles of different patient groups. One of the interviewees stated that actuaries using AI/ML techniques in the healthcare system was presently low but increasing. For example, three interviewees mentioned an increasing trend for health insurers to use AI/ML models to encourage healthier behaviour among their customers to reduce future claims. One interviewee stated that the practice was used extensively in their own company, but that this was only dealt with by data scientists and not by actuaries.

'Other' respondents to the survey also reported using more complicated AI/ML techniques at a higher rate – such as reinforcement learning and natural language processing. Consistent with this, 'Other' interviewees typically gave the impression that the AI/ML approaches they used were more advanced. Examples included participants reporting that their organisation had a dedicated team responsible for researching and facilitating the implementation of new AI/ML methods and describing particularly advanced use cases that included in-house fine-tuning of LLMs that could accurately answer text prompts about the company's internal policies, processes and data. Some 'Other' interviewees also reported that actuaries within their organisations typically worked closely with data scientists, had strong data science expertise, and formed a minority within highly technical teams. However, survey respondents in the 'Other' category did not generate higher average scores in response to the question about plans to increase use of AI/ML over time.

One of the 'Other' interviewees discussed the use by their organisation of extremely advanced LLMs internally for non-actuarial purposes, and questioned whether models of this nature were sufficiently explainable to be applied to actuarial work or other regulated work where explainability was required.

3.4 Variety of techniques used

Figure 8 – number of survey respondents who reported using different techniques in one or more area.



A further breakdown of technique use by areas within fields is shown in Appendix B.

Some participants cited the use of advanced and relatively new approaches, like deep triangle and extreme gradient boosting. Others reported using Generalised Linear Models (GLMs) or other simpler and more well-established AI/ML methods such as basic decision tree classification analyses.

While some interviewees cited the use of GLMs as an example of ML within their organisation, others noted that GLMs were typically treated as a more traditional statistical technique within their organisation, where ML was seen as referring to more advanced methods such as neural networks.

Survey respondents also differed on whether GLMs were presented as an ML technique (three survey responses) or a traditional technique (six survey responses). Two of the respondents who cited GLMs as an example of ML explicitly noted the divergent perceptions of its status as an ML technique, one of whom discussed this matter in detail in the context of a broader point about the fact that ML techniques varied widely in their complexity and that many were very similar to the statistical tools that actuaries are trained to understand and already used extensively.

3.5 Expanding use of Large Language Models (LLMs)

LLMs were mentioned explicitly in 14 of the 20 interviews, mainly with reference to ChatGPT, which interviewees typically presented as an exciting and interesting technological development. There was widespread agreement that LLMs like ChatGPT would have a huge impact across a wide range of economic sectors, including for actuaries.

“Think carefully about how you regulate ChatGPT – it’s either going to change the world or lead us to doom.” An interviewee discussing the risks and opportunities created by ChatGPT

“You can quote me on this: [LLMs] are changing the way we develop models and software and everything that relates to us driving stuff with a computer.” An interviewee discussing the seismic changes being created by LLMs

Participants explained that while the take-up of AI/ML had been slower in actuarial work in comparison to other industries, the emergence of new technology such as LLMs could lead to rapid change, which companies and regulators need to be aware of. In most cases, interviewees emphasised the importance of caution in using these powerful tools because of the various risks involved.

Only two survey respondents mentioned LLMs in relation to their work (or the work of their organisations), and they did so only in contexts not specific to actuarial work. This discrepancy between the prevalence of references to LLMs in the survey and interviews is likely due to timing. The surveys were conducted between January and April 2023, whereas the interviews were conducted from March to June 2023, by which time media coverage of ChatGPT had become far more extensive.

Interviewees reported a diverse range of use cases for LLMs, covering both speculative possibilities and actual applications. The most commonly reported ‘already in use’ example was to leverage LLMs to help to write programming code, or to assist novice programmers with their learning. Other “already-in-use” use cases included asking ChatGPT to describe likely future trends to inform a horizon-scanning exercise for an insurer that focused on a specific demographic group and producing written summaries for large sections of computer code in order to explain the code’s workings to non-technical audiences. One interviewee stated that their organisation had trialled the use of a third party LLM for processing transcripts of customer complaints, reporting that it had saved a lot of manual effort. The interviewee in question did not mention any of the potential risks involved in sharing data with a third-party LLM in this way.

One of the speculative use cases raised was the use of ChatGPT to parse the contents of written reports to inform assessments of the risk associated with insuring certain assets. This approach was eventually abandoned, despite ChatGPT being highly accurate in the way it processed the reports, because it was very difficult to identify a pattern behind its reported (rare) instances of inaccuracy due to the ‘black box’ nature of the model’s internal workings. Some of these inaccuracies could have led to catastrophic consequences if they fed through to subsequent stages of pricing modelling, so the company decided that the risks of using ChatGPT for this were unacceptably high.

At the most advanced end of the spectrum, two interviewees stated that their organisations had started developing their own LLMs by fine-tuning third-party models using proprietary information held within the organisation's own records, enabling the models to respond accurately to prompts about internal policies and information. One of these interviewees stated that these fine-tuned LLMs were already producing ground-breaking productivity gains throughout their organisation to an extent comparable with the transformative changes resulting from the introduction of computers to actuarial work in the 1970s. These interviewees noted that, because the LLM was operated entirely from the company's own servers, there was no risk of data contained in the prompts submitted to the LLM being transmitted to untrustworthy third-parties. They also emphasised that they implemented various safeguards around internal use of LLMs – for example, one interviewee reported that their organisation meticulously ensured that all processes involving AI were subject to human oversight and that no decisions were taken based purely on the output of an AI model.

In this case, the organisation was at the stage where it could feed the LLM a long text document with details about a financial model that they wanted to build – including mathematical formulae characterising the precise mechanisms of different processing stages – and the LLM would then produce the described model from scratch. The LLM was so familiar with the organisation's internal policies that it knew exactly where to read in the relevant data and where to save audit information about how the data had been processed. The interviewee said that only relatively minor human modifications to the LLM's output were required to produce a functioning model that met the specifications provided in the text input to the LLM.

Participants in some interviews outlined other potential use cases for LLMs in their organisations beyond those described above. For example, an interviewee who was using natural language processing to build tailored investment advice based on client responses to an open-ended questionnaire hinted that some colleagues thought that it might be possible for ChatGPT to be used for this purpose.

More broadly, the examples of use cases given in many interviews suggested large variations in familiarity with ChatGPT's functionality and the ways it might be used. Some interviewees were using it to write code and tended to see ChatGPT as a web-based tool available on OpenAI's website. Other interviewees discussed various ways ChatGPT could be accessed, such as through API (Application Programming Interface) calls executed as part of a script in a programming language like R or Python. These interviewees also then discussed the implications of this for how user inputs were shared with the server where the ChatGPT API was exposed.

One interviewee's organisation accessed ChatGPT via its Microsoft Azure account rather than through OpenAI because it had more confidence in Microsoft's data security safeguards than OpenAI's. Another interviewee expressed a similar view but noted that their organisation did not have access to the Microsoft-provided version yet. Those interviewees who had deeper knowledge of how users could interact with ChatGPT tended to work for organisations that had considered more advanced and sophisticated use cases; they also tended to be more informed about some of the risks associated with its use.

One of the interviewees who had been involved in such use cases stated that their organisations' employees were advised against using the online version of ChatGPT, or similar widely accessible LLMs, for work purposes because of the risks involved in sharing data with third parties such as OpenAI; they emphasised that this issue was eliminated when the organisation operated an LLM internally.

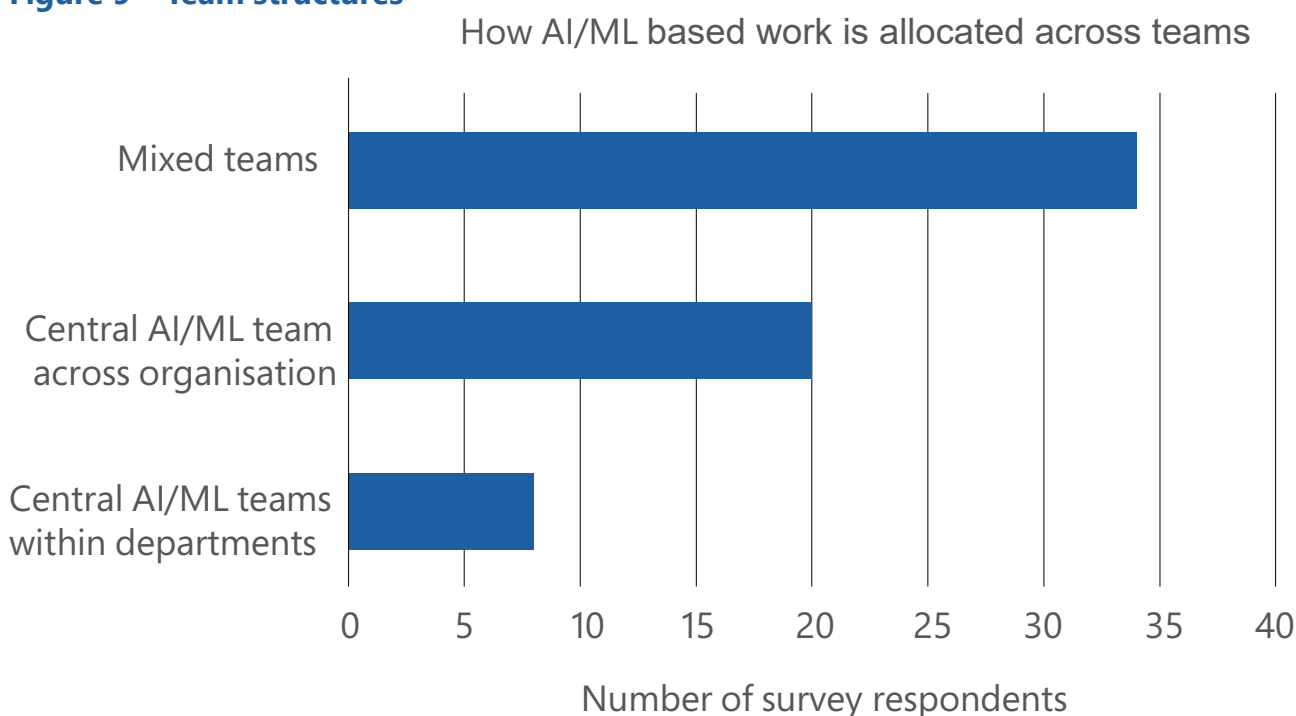
3.6 Distribution of AI/ML expertise across teams

Interviewees described how data science expertise was distributed in their organisations. Participants in seven interviews stated that data science activities were largely under the remit of a central team that may also have featured multi-disciplinary groups. In one case it was stated that the centralised structure was deemed to be optimal for managing the governance of AI/ML work.

The participants of four interviews stated that data scientists tended to be dispersed through their organisation, with one interviewee explaining that there were multiple data science teams in different parts of their organisation. One interviewee described a central team comprised exclusively of data scientists that was responsible for exploring innovative new applications of ML techniques coupled with multiple teams of actuaries and data scientists spread across the organisation.

There was no discernible pattern linking the quantity or nature of the AI/ML analysis undertaken by an organisation to the distribution of data science expertise within. The distribution of AI/ML expertise across teams reported by the survey respondents is shown below in Figure 9:

Figure 9 – Team structures



Data from this plot are based on the question:
"Which of the options below best describes how Artificial Intelligence and Machine Learning-based actuarial work is allocated across roles and teams within your organisation?"

4. Governance and quality assurance

4.1 Robustness of governance

In general, interviewees indicated that quality assurance of AI and ML-based actuarial work was taken seriously in their organisations and that the various governance mechanisms in place ensured the analysis was robust and appropriately validated. Two interviewees explicitly stated that the quality assurance and general governance procedures relating to AI/ML in their organisations probably required improvement; another raised the concern that actuaries may have struggled to keep abreast of the rapid advances in the AI/ML techniques increasingly being applied in their work, calling into question their ability to serve as credible reviewers of analysis that AI/ML was used to produce.

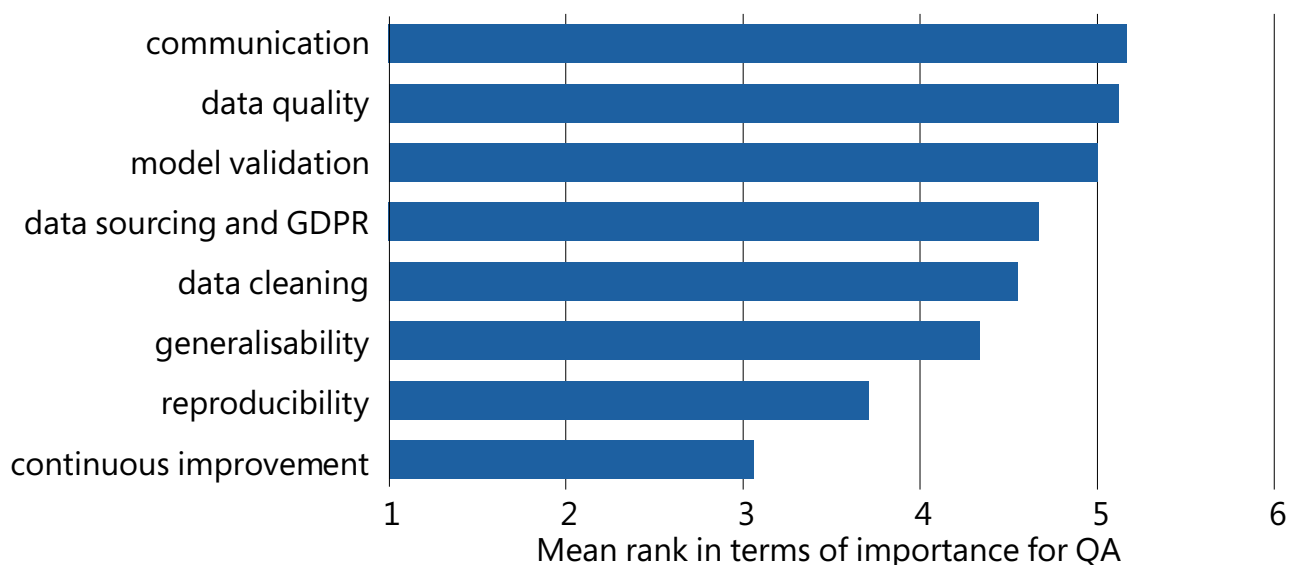
Elaborating on the point about governance processes needing improvement, one interviewee expressed concern that senior executives may have been too distanced from the technical checks and validation processes essential for robust quality assurance and compliance, to the extent that it could have been difficult for them to judge with confidence whether all necessary testing was performed. Another interviewee raised the issue of the uncertainty, within their organisation, about who was most the appropriate person to review AI/ML analysis, given the technical expertise required.

Similarly, two survey respondents, when answering a question about how caveats and uncertainties are communicated to senior stakeholders, described concerns about their organisations' governance processes. One respondent said that AI/ML practitioners in their organisation often failed to communicate uncertainties and caveats to senior stakeholders, and another suggested that senior actuaries may not have the necessary technical skills to adequately review AI/ML analysis, resulting in ambiguity over who was the most appropriate reviewer. However, as with the interviews, most respondents indicated that the governance processes in their organisation were robust.

Some interviewees gave detailed and concrete descriptions of how the governance and quality assurance processes in their organisation applied to AI/ML analysis. For example, they talked about the different levels of sign-off that were required for different types of analysis and the formal procedures that practitioners in their organisation had to complete before an ML technique could be put into operation. In contrast, other interviewees strongly asserted their organisations' commitment to robust quality assurance and sound governance but made less definitive statements about how these ideals were put into practice.

As set out in Figure 10, respondents indicated that they placed greatest importance on communication, data quality and model validation in their QA processes. Reproducibility, and particularly continuous improvement, were considered of lower importance.

Figure 10 – Perceived importance of different aspects of the AI/ML QA process



This plot is based on data from the question:
“Please rank the following in terms of the importance you would ascribe to each when quality assuring Artificial Intelligence and Machine Learning-based actuarial work. You may rank two or more items equally.”

Governance differences compared to non AI/ML analysis

Interviewees from only two organisations confirmed that they had a separate governance process specifically for ML models; interviewees from one of these organisations described an even more specific governance process that applied exclusively to LLMs and other models using generative AI. However, in eight interviews, it was stated explicitly that there were no formal differences between the governance processes for AI/ML-based analysis and those for more traditional analysis.

Three survey respondents described formal differences among the governance processes for AI/ML analysis. These included requiring a higher level of sign-off, more thorough documentation of assumptions and reviews, completion of a ‘Model Risk Register’, and more rigour around data protection and IT security. Two of these three respondents also referenced the greater need to ensure that models were explainable when using AI/ML, compared to the case with traditional analysis, saying that the latter tended to be more transparent and interpretable by nature. This idea of AI/ML models needing to be presented in an explainable way was also echoed by five other respondents in the same question and by 32 respondents in the survey, albeit not always in the context of discussing governance differences between AI/ML and traditional analysis.

Separately, in a different question asking about the risks of AI/ML and associated mitigations, six respondents mentioned the idea of implementing more rigorous quality assurance requirements or requiring higher-level sign-offs as potential mitigations against AI/ML risks.

4.2 Governance and quality assurance practices for complex AI/ML models

Interviewees indicated that more complex AI/ML techniques typically required more thorough quality assurance; these techniques tended to be less well known, and more commonly employed by organisations that were using AI/ML more extensively. One interviewee, describing the use of a random forest model for assessing credit risk, stated the relative simplicity of the technique meant a less stringent governance regime was required, partly because the model's simplicity made it easy to explain. Interviewees using more advanced techniques described far more involved governance procedures running alongside the model, including the interviewees who reported a bespoke governance process specifically for AI/ML.

Another interviewee stated that the importance of ensuring explainability within their organisation's model-governance process depended on the complexity of techniques being employed and the business criticality of these techniques' applications. The respondent's view was that their organisation was more comfortable with the techniques being deployed at a smaller scale on lower-priority tasks that needed explaining.

Participants discussed how many of the additional governance processes (formal or informal) were considered applicable when using more complex ML techniques that were geared towards enhancing explainability in addition to merely ensuring accuracy of outputs. In an exception, one interviewee mentioned a lower relative importance of explainability in their governance process, emphasising instead the need to prioritise model accuracy. They explained that it was viewed as more important to use appropriate assumptions and identify rare but impactful edge cases where the model performed poorly. There were also several examples from interviews of organisations actively avoiding using a more accurate AI/ML model in favour of a corresponding traditional model because the AI/ML model was less transparent and more difficult to explain.

There was variation in the types of technical procedures used by interviewees either to provide insight into a model's behaviour, or to assess its performance. Much of this was attributable to differences in the type and complexity of the AI/ML analysis performed. For example, some interviewees and one survey respondent mentioned the role of SHapley Additive exPlanations (SHAP) analysis and feature importance analysis in their quality assurance process for supervised learning methods.

Interviewees using unsupervised learning methods⁷ mentioned alternative quality assurance techniques that involved checking that key metrics and trends in the data followed the expected statistical patterns. However, most of the AI/ML techniques that interviewees reported were supervised methods, and the techniques used to quality assure them were directed at understanding the relationships between input variables and model predictions.

⁷ See Glossary for definition of unsupervised learning.

4.3 Communication with decision makers

The participants of four interviews mentioned a lack of actuaries' AI/ML skills as putting greater onus on AI/ML practitioners to communicate outputs in a form accessible for actuarial reviewers (and others involved in sign-off). This included making sure that AI/ML models were not perceived as 'black boxes' that converted inputs to outputs in an inexplicable manner.

Many interviewees stated that the need for explainability was an important part of the governance process surrounding AI/ML analysis. They explained that it was necessary to communicate model outputs and behaviour in a way that regulators, actuarial reviewers, and other people responsible for analysis sign-off could understand and relate to actuarial concepts that they were familiar with.

There were examples of interviewees adapting their approach to the needs of the decision makers. Two interviewees stated that they must adapt the way they communicate model outputs to account for the fact that key decision makers responsible for signing-off AI/ML analysis tend to assess models by relying heavily on their intuition and expertise, or on more traditional actuarial metrics for assessing the reasonableness of model behaviour. One interviewee stated that 'black box' issues for AI/ML models were less of a concern in their organisation because senior decision makers have a habit of "unpicking" models until they understand them thoroughly and there is "healthy scepticism" within their team about the extent to which AI/ML models can be trusted. Another response said the issue of explainability was mitigated by the main customer for their model being highly inquisitive and keen to understand how it worked.

One interviewee addressed the challenge of communication to decision makers by attempting to translate metrics extracted from their AI/ML models into concepts more familiar from an actuarial and underwriting perspective. They used interactive dashboards to show how changes to model inputs and parameters affected outputs. One group of interviewees identified the same need to present outputs using language and metrics that were familiar to actuaries. This issue also came up in one survey response.

Another interviewee addressed this challenge by developing simplified 'challenger' models using techniques familiar to the decision makers to demonstrate that these techniques yield similar outputs to the AI/ML approach. Similarly, in the survey question on risks of AI/ML and associated mitigations, seven respondents⁸ suggested comparing AI/ML outputs to traditional models as a mitigation against some of the risks from AI/ML.

A further interviewee briefly stated that the necessity for underwriters to understand ML models often leads to a "blended approach of analytics and expert judgement". Consistent with this, survey responses (in Figure 10 above) showed communication rated as the most important feature of both the governance process and the quality assurance process.

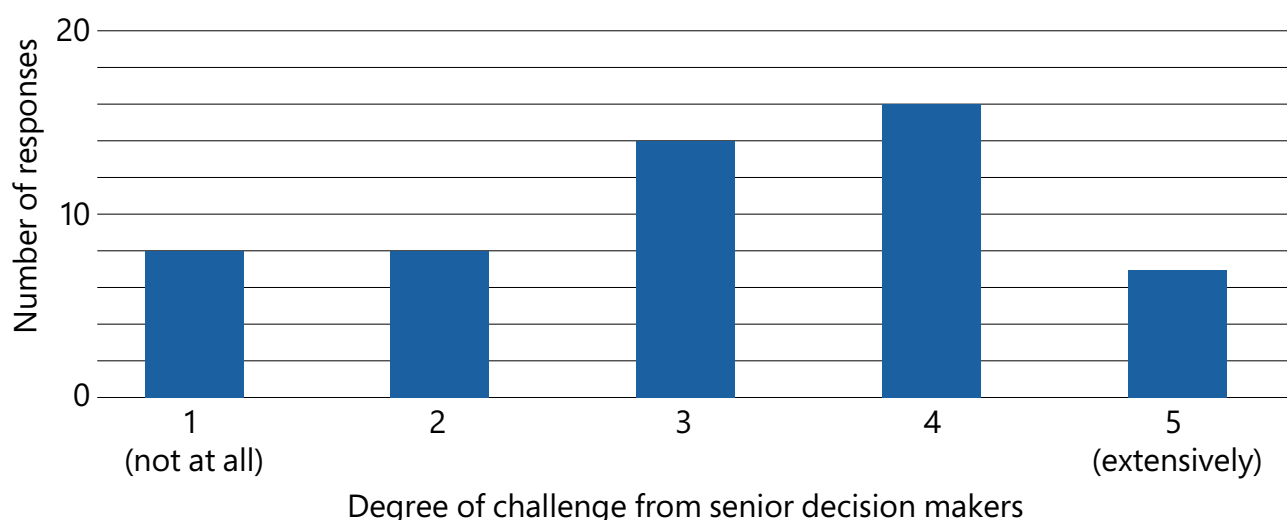
Participants from three interviews raised the risk that rapid advances in the AI/ML techniques being applied may be outpacing changes in actuary skillsets, making actuarial oversight increasingly difficult. Another interviewee expressed reservations about the

⁸ 2 GI; 2 Life; 1 Pensions, 1 Finance & Investment and 1 'Other'; about 16% of people who gave a substantive answer

limited technical knowledge of regulators, which could lead them to rely on more intuition or rule-of-thumb -based strategies for judging the appropriateness of analytical practices.

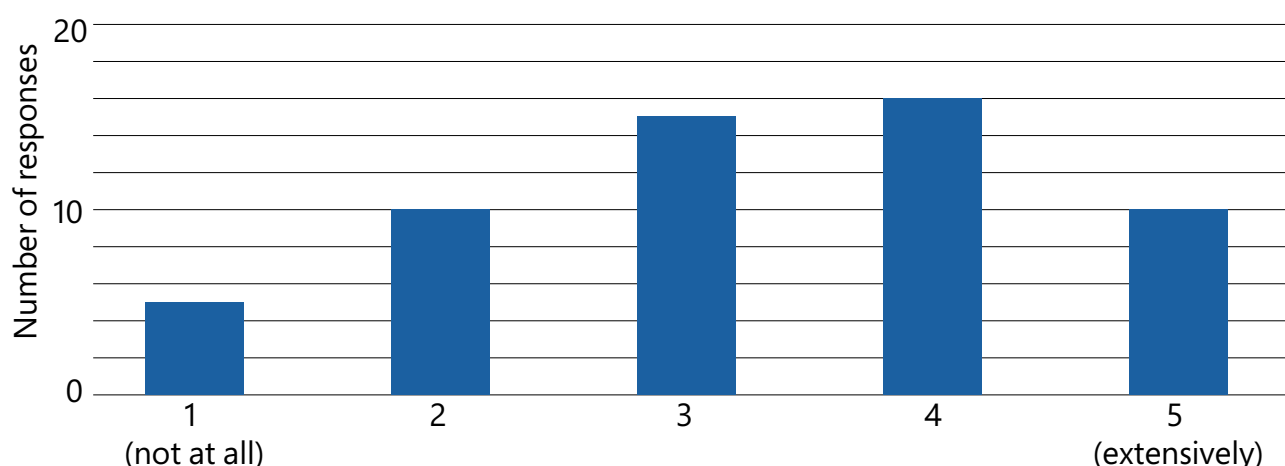
Survey respondents indicated there were reasonably effective mechanisms to ensure a process of challenge and applying 'healthy scepticism' to models was in place. When asked about the extent to which senior decision makers provided challenge in relation to AI/ML models, the mean rating was 3.1 on a scale from 1 to 5, with little variation between actuarial fields (see Figure 11). There was a similar answer to a survey question asking about the extent to which caveats and risks are communicated to decision makers with a rating of 3.3 on a scale from 1 to 5 where 1 was 'not at all' and 5 was 'extensively' (see Figure 12).

Figure 11 - Degree of challenge from senior decision makers when receiving AI/ML analysis



The data in this plot is based on responses to the question: "To what extent do decision makers in your organisation provide challenge and critical input when presented with Artificial Intelligence and Machine Learning-based actuarial analysis?", for which the response options ranged from 1 (not at all) to 5 (extensively). 51 of 104 survey respondents did not answer this question.

Figure 12 - Effectiveness of communication of caveats and risk to decision makers



The data in this plot is based on responses to the question: "In your organisation, how effectively would you say that caveats and risks relating to Artificial Intelligence and Machine Learning-based are communicated to decision makers?" for which the response options ranged from 1 (not at all) to 5 (extensively). 48 of 104 survey respondents did not answer this question.

Two interviewees highlighted the risk that a failure to explain the workings of an AI/ML model could make it more difficult for actuaries to understand the kinds of checks and tests that need to be performed to validate its outputs, thereby heightening the possibility that they allow it to pass their review without having been adequately tested.

4.4 Reproducibility

Only five interviews mentioned the idea that AI/ML analytical outputs need to be reproducible. Two of these mentioned the use of version-controlled workbooks, recording which versions of different software packages have been used in a given model so it does not behave unexpectedly if these packages are updated by default, and 'setting seeds' so that any pseudo-randomness generated as part of the modelling process will follow exactly the same pattern when the model is re-run, and therefore make it reproducible. One of the two interviewees said that the issue with ensuring pseudo-random numbers are reproducible occurs in many aspects of actuarial work, but is more commonly encountered with AI/ML models than other types of models they had worked with. One interviewee emphasised the importance of version control tools like Git to record all assumptions, data and parameters for any given model run.

There were different interpretations of the term reproducibility amongst participants. For some it referred to two independent users generating exactly the same results when running the same model under a pre-specified, and easily re-created, set of conditions. For others, reproducibility was achieved when model outputs were approximately replicated using alternative modelling techniques.

The findings from the survey were consistent with the relatively low level of importance placed on reproducibility in the interviews. In a question asking about the importance of different aspects of the AI/ML quality assurance process, reproducibility was ascribed the second least importance (see Figure 10).

4.5 Consideration of regulation in governance and quality assurance process

Three interviewees spoke in detail about how regulatory requirements influenced governance in their organisation, or would be likely to do so in the future. One interviewee discussed how their organisation was striving to incorporate into its governance process the principles promoted in a recent white paper published by the UK Department for Science Innovation & Technology⁹ describing the UK government's pro-innovation approach to AI regulation.

⁹ <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>
see paragraph 10.

5. Explainability

5.1 Importance of explainability when using AI/ML techniques

In all but one interview, the point was made that explainability is an important factor to consider in the context of AI/ML analysis to a greater extent than is usually the case for traditional actuarial analysis. The one exception to this stated that their team was trying to understand the reasons for their models' behaviour but indicated this was not always possible and that they were placing higher priority on ensuring predictive accuracy. This interviewee then said that their organisation is based in a jurisdiction where regulation of their industry is more relaxed than the UK.

The need for explainability was also reported to vary depending on the importance of the model in question. For example, one interviewee emphasised that their organisation tends to place more focus on explainability for AI/ML models that are used for more business-critical activities, which are therefore likely to involve simpler techniques. More complicated techniques tend to be used for less business-critical activities and are therefore subject to less pressure to be explainable.

In five interviews where the AI/ML analysis was used to inform insurance or reinsurance pricing, interviewees stated that they often favour models that are easier to explain over slightly more accurate alternatives.

Consistent with the extensive mentions of explainability among interviewees, the most frequently mentioned risks of AI/ML by survey respondents were issues relating to explainability and the idea of avoiding 'black box' AI/ML models. 20 respondents¹⁰ mentioned explainability in this question. Of these, four said they mitigate the risk by opting for simpler and more explainable AI/ML models. Other approaches included developing "explainers" in model documentation, using more visualisations, being more careful when checking model inputs, and avoiding using AI/ML tools that they do not understand.

5.2 Ease of explainability for AI/ML models

Some interviewees indicated that that they did not find it difficult to ensure that their ML models were explainable, usually in cases when the modelling technique being used was more amenable to explanation.

For example, some interviewees were aware of an AI/ML model that has resonated with the actuarial community because of its focus on explainability. These interviewees explained there was an organisation which had developed an AI/ML model that allows reserving actuaries to inspect a large number of diagnostic indicators, with the most important variables highlighted to enable the actuary to focus their attention on more useful information. This model had reportedly been designed with explainability as a core consideration – it allows users to clearly determine the reasons for different decisions taken by the model at different stages of processing. Users could also test the impact of changing these decisions or the underlying assumptions.

¹⁰ About 44% of substantive responses to this question, comprising 12 GI, 6 Life and 2 'Other'

Another interview mentioned that a research team within their organisation was in the process of developing a method for layering a generative AI mechanism on top of an existing decision tree model, such that the generative AI tool could present and explain how the decision tree model operated. While there were seen to be potential benefits from layering AI mechanisms on top of existing models, it also introduced the risks inherent in a generative AI tool.

In one interview, where the organisation in question used ML for assessing credit risk, the model technique used (decision trees) was relatively easy to explain because of its simplicity. These interviewees said that simpler forms of ML were less likely to necessitate an increased focus on explainability because, in many cases, they were easy to explain by their nature.

Some participants characterised some types of AI/ML, particularly those underpinned by neural networks, as inherently unexplainable, although they did not suggest that this lack of explainability meant that AI/ML techniques of this nature should never be used.

5.3 Different approaches to explainability

Some participants suggested that they faced increased risks when using models with poor explainability because it was harder to ensure they comply with regulatory requirements, and likewise to demonstrate to actuarial reviewers that model outputs were reasonable. One interviewee noted that when working with a model whose mechanics were not fully understood, it was difficult to anticipate situations where its outputs might be highly erroneous. Even though such instances were infrequent, the inability to predict their occurrence, or the nature of the errors that may ensue, meant they could have severely negative impacts.

Other interviewees viewed explainability as a compliance requirement and therefore saw it as less of a concern for less heavily regulated aspects of their business.

There were also discrepancies in how different interviewees appeared to interpret the term “explainability”, particularly in the context of LLMs. Some interviewees characterised LLMs as a ‘black box’ whose complex inner workings were impossible to understand. One interviewee stated that this difficulty in understanding when LLMs were likely to make mistakes – even if the mistakes may be rare – led to their organisation abandoning a plan to use ChatGPT for extracting information from text reports in a context where the consequences of mistaken analysis could be catastrophic.

Some research participants considered explainability to mean understanding the internal workings of a model. Others considered a model explainable if they could demonstrate and understand how its outputs change relative to different inputs. AI/ML models were also more explainable where the recipients of the explanation had better understanding of the techniques involved; one interviewee expressed a view that explainability can be understood as being relative to the technical competence of a reviewer, customer or regulator. Overall, the research for this project suggested there was no widely accepted benchmark for what constitutes an “explainable” model.

6. Data privacy

The concept of data privacy and security was mentioned in all but two interviews. The participants of two further interviews simply confirmed that their organisations did not use personal data so data privacy was not a major concern. Another interview discussed the importance of “ring-fencing” personal data used in ML models so it could only be accessed by people needing to use it. Three interviews mentioned the use of anonymisation as a means of mitigating against data privacy risks, in one case specifically in relation to data shared with ChatGPT. Two interviewees discussed how their governance process was designed to safeguard data privacy and security. The data privacy risks associated with sharing information with online LLMs were mentioned in only two interviews.

Interviewees generally took the view that data privacy concerns were of equal relevance for both AI/ML and traditional analysis, as AI/ML was seen as relating primarily to how data was processed, not the type of data that was processed. However, some interviewees described AI/ML techniques that greatly expanded the content and diversity of the datasets fed into subsequent modelling stages, meaning they indirectly led to more data privacy concerns. For example, natural language processing could be used to process unstructured data that might have otherwise been inaccessible.

One interviewee said they were aware of a project to create a repository of highly sensitive personal health data for access, in principle, by any organisation that can demonstrate their suitability to handle the data correctly (i.e. subject to constraints designed to protect data privacy and to the consent of data subjects). The project envisaged organisations and individuals submitting coded analysis requests that would be executed only if the data they return is sufficiently aggregated and without any personally identifiable information. The interviewee suggested this could attract the interest of health insurance companies but cautioned that care would be needed to ensure robust data protection mechanisms were in place. For example, even where data is extracted in aggregated form, it can still be possible to infer individual details from aggregated datasets (making it unlikely to be allowable under GDPR).

In contrast to these extensive discussions of data privacy in the interviews, only a single survey respondent mentioned GDPR or data privacy. Another mentioned they had once refrained from engaging in an AI/ML project because of a failure to receive consent to use third-party data.

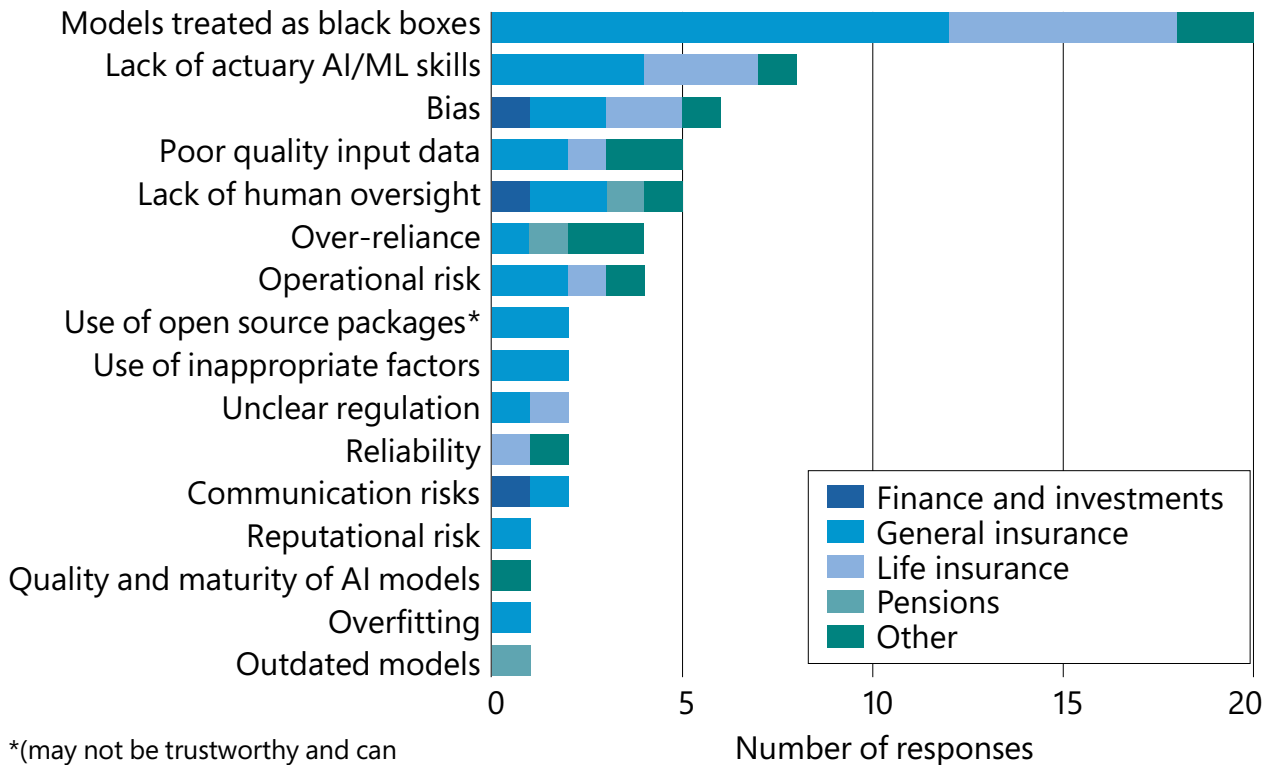
7. Increased risks from use of AI/ML techniques

7.1 Overview of risks of use of AI/ML techniques

Figure 13 sets out participants' views of the risks introduced or increased by the use of AI and ML in actuarial work, ranked by the frequency referenced.

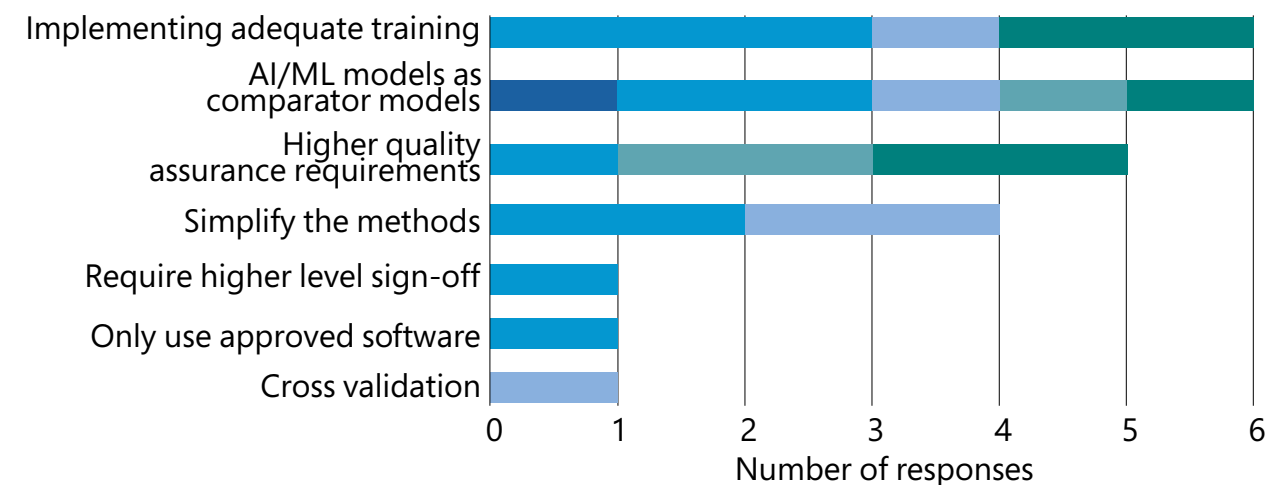
Figure 13

Risks



*(may not be trustworthy and can impair version control)

Mitigations



Data from this plot are based on manual categorisation of open-text responses to the question: "What risks have been increased or introduced by the use of Artificial Intelligence and Machine learning in actuarial work? What measures do you take to address these risks?"

Some of these risks were already discussed in previous sections; others are discussed further in the remainder of this section.

7.2 Lack of actuaries with AI/ML skills

Interviewees discussed how AI/ML model reviewers, including actuarial reviewers, need to have a higher level of data science expertise than reviewers of non AI/ML analysis.

Some interviewees specifically highlighted the need for actuaries to move away from Microsoft Excel and towards more AI/ML friendly software, such as R or Python. Interviewees reported that the model checks required, and the types of model risks that they need to look out for, were typically more numerous, complex, and technically demanding than was the case for traditional analysis. It was apparent that this had implications on the skills sought when hiring actuaries, and on the training and education provided to them.

Likewise, some of the interviewees and survey respondents linked the growing need for AI/ML skills among actuaries to labour market dynamics. When asked about the main AI/ML-related challenges facing actuaries, 17 respondents¹¹ discussed the increasing pressure on actuaries to develop AI/ML capabilities and, in some cases, to justify their continuing involvement in analysis given the expanding role played by non-actuary data scientists, who were potentially cheaper to employ. Some respondents stated that actuarially relevant AI/ML analysis is increasingly being performed by teams in which actuaries tend to be outnumbered by data scientists. Seven survey respondents¹² stated that actuaries often lacked the technical skills to handle advanced AI/ML techniques and/or that the relevant skills were often so scarce that 'key person risks' arose because of the small number of people capable of understanding and operating a model.

7.3 Bias and discrimination

Interviewees frequently raised the need to avoid bias and discrimination against certain groups, specifically with regard to gender-based price discrimination.

One interviewee commented they did not consider members of executive boards were provided with enough technical information about companies' use of AI/ML to confidently judge whether the risk of discriminating on the grounds of gender when pricing insurance products had been avoided. They added that this risk was higher than with traditional techniques due to lower explainability.

The interviewees from one organisation stated that they had a mathematically watertight and legally defensible method for proving their models did not discriminate on the basis of gender. Some other interviewees considered the risk of indirect gender discrimination in pricing, such as using health conditions that heavily correlated with a certain gender as predictor variables in a model. Another interviewee opined that the lower transparency typically found with ML models made it more likely that they could inadvertently discriminate on the basis of protected characteristics like gender, without the model builders being aware.

11 8 GI, 3 Life, 4 Pensions, 2 Finance & Investment; about 31.5% of the people who provided a substantive answer to the question

12 Around 16.2% of those who supplied a substantive answer to the question; comprising 3 GI, 3 Life and 1 'Other'

One interviewee whose work mainly covered population health management mentioned that when performing AI/ML analysis in this area, they were required to assess the impacts of different health interventions on health inequalities between different groups, for example by assessing whether a given intervention would exacerbate or reduce existing health disparities between races or ethnic groups.

One interviewee commented that there was a lot of inherent bias in the data against people with certain protected characteristics in big health datasets¹³ and in the application of public health models. The interviewee said that ML could help to deal with this by shedding more light on differences in health interventions that worked best for specific groups.

One interviewee discussed the risk that non-traditional data sources could inadvertently restrict access to financial services for small subsets of the population where there are exceptions to general statistical patterns. For example, a wheelchair user may not have been able to access health or life insurance products if insurance companies relied heavily on data from step-counting apps or similar metrics to price their products.

7.4 Lack of human oversight and automation

One interviewee gave the view that there were huge risks in implementing a fully automated process for actuarial or analytical purposes that is totally dependent on an LLM, with no human oversight. The interviewee stated that this could lead to catastrophic results given the unpredictability of the mistakes that LLMs could make, even if these mistakes were rare. The interviewee stressed that human involvement was fundamental for the associated governance framework for ensuring accuracy, regulatory compliance and ethical practice. This view was echoed by other interviewees.

One interviewee also stated that it was wrong to assume that human involvement is a perfect solution for minimising risks because humans can make mistakes just like models. This interviewee said there was a tendency for AI/ML models to be held to an unfair standard in this regard, wherein decision makers may demand a much higher level of accuracy from an AI/ML model than they could reasonably expect from a human performing the same task.

¹³ Data may have over-representation of certain groups, who may therefore be better represented in research studies whilst those with protected characteristics are underrepresented.

7.5 Poor data quality

Risks associated with data labelling were mentioned, such as the burden of the process and the potential for inaccuracies. One interviewee cited an example of an ML model designed to identify skin cancer based on photos of patients' skin. The model appeared to perform well when tested, but its developers eventually realised that its strong performance was almost entirely due to its ability to detect ink marks around cancerous skin that had been drawn by doctors who had already identified the location of a tumour.

Interviewees also expressed reservations about potential biases or systematic inaccuracies in the data used to train ML models, leading to flawed patterns of judgement and low generalisability of models to different contexts. Some interviewees noted that this could be a bigger issue for AI/ML models than for more traditional actuarial analysis because of the larger, potentially less structured and less heavily audited, datasets that could be processed by AI/ML models compared to traditional models.

The participants of one interview raised the issue of fake news and deep fakes becoming increasingly sophisticated, potentially to the point of being able to trick AI/ML-based tools designed to monitor and prevent fraud or to respond to certain news developments. This risk was described as being higher in light of the fact that the AI/ML models in question sometimes needed to respond to new events in real time, reducing the scope for humans to override their decisions. This point aligns with the fact that some survey respondents reported using natural language processing for the purposes of detecting fraud or scraping news articles to understand market-wide trends.

7.6 Over-reliance on AI/ML analysis

In six interviews, participants highlighted that AI/ML is sometimes seen as a "magical" solution to analytical problems to the extent that there is a risk of some actuaries being overly optimistic about the kinds of problems it could solve. Participants suggested that actuaries may rely on these techniques excessively without understanding the caveats and applying appropriate validation. Two of these emphasised that AI could not serve as a replacement for human judgement and decision making, and it was important for actuaries to avoid the pitfall of assuming that an AI/ML model could do their thinking for them.

One interviewee highlighted the risk that advances in AI/ML techniques might outpace practitioners' understanding, making it more difficult for them to understand the limitations of what AI/ML could achieve or suppressing their sense of healthy scepticism by showcasing particularly impressive AI applications that would lead to exaggerated impressions of what AI is capable of. In one case, the issue of over-reliance was mentioned as the interviewee's primary concern about ChatGPT.

In the survey responses, there were four mentions of over-reliance on AI/ML models in the question about risks¹⁴. Two of these suggested corresponding mitigations. One was to ensure that actuaries have a sufficient level of skill to understand models; the other was to run AI/ML models in parallel with other methods so as to compare results and assess reasonableness. As noted, while respondents' mean rating was 3.3 (on a scale from 1

¹⁴ About 9% of the substantive responses to the question, comprising 2 'Other'; 1 General Insurance and 1 Pensions

to 5) for the effectiveness of how caveats about AI/ML were communicated to decision makers, a small number of respondents said risks and caveats were not communicated well.

Reliance on LLMs

Interviewees discussed the level of reliance that can be placed on LLMs and the risks and benefits they could produce. Some interviewees emphasised the importance of remembering that LLMs are not sentient, that they do not possess “General Intelligence”, that their function is simply to find patterns in long sequences of text, and that this is not equivalent to “actual thinking”. However, these interviewees did not further explain the risks this could give rise to.

Many interviewees mentioned the potential for over-reliance on LLM’s outputs as set out above - a reviewer might be so enamoured by the capabilities showcased by an LLM that they fail to appreciate its fallibility and the need to subject its outputs to rigorous checking and quality assurance.

When determining the appropriate scale, extent and form of these checks, the kinds of factors participants cited as needing to be considered included information about how accurately the LLM had been found to perform similar tasks in the past or known quirks in the LLM’s behaviour that might affect its performance on specific types of inputs. Participants expressed the view that it would not be necessary to consider whether an LLM has “General Intelligence” or is “actually thinking” because the answer to these questions would reveal nothing about the LLM’s reliability beyond what can be derived from systematic testing or other mathematical techniques for probing its behaviour, meaning that they are irrelevant to quality assurance and risk management.

7.7 Insufficient holdout data

One respondent highlighted the risk of not having enough data in the holdout set when validating a model after parameter tuning. In general, when supervised ML models are developed, model builders typically divide their data into:

- Training set
- Validation set
- Holdout set.

The training set is used to train the model. The validation set is employed to tune the model's parameters and optimise its performance; this can be repeated many times to fine-tune the model. The holdout set is used to test the model's accuracy. Because the model is not exposed to the holdout set during the training or validation stages, it can be used to derive an unbiased estimate of the model's accuracy, free from any bias that could have been introduced during the parameter tuning phase.

This process presents a trade-off: using a larger portion of the data for the training and validation sets may allow the model to be a better fit to the data presented, as it has more data from which to learn. However, this leaves a smaller holdout set, which could lead to a less reliable estimate of the model's true performance at the end of the process. The interviewee in question implied that the temptation to reduce the holdout set in this way at the expense of reliably assessing the model's accuracy could lead to holdout sets being inappropriately small in some cases. This risk would be particularly high when suitable data is scarce.

7.8 Lack of supporting infrastructure

One interviewee described a failure to build up the required data infrastructure meaning a company did not have the ability to handle vast datasets and to provide the extremely high amounts of computing power required for some particularly advanced ML models. Likewise, three survey respondents mentioned the importance of a supportive and secure IT infrastructure when deploying AI/ML models.

7.9 Other risks identified in survey responses

Whilst the survey responses and interviews discussed a number of risks in relation to AI/ML, survey respondents had not typically avoided using AI/ML due to these risks. The mean response to a question asking whether they had ever refrained from a piece of AI/ML analysis due to legal, regulatory, or ethical issues was 1.8 on a scale from 1 to 5.

Respondents whose answer to the above question was anything other than 1 (indicating that they have never shelved a piece of AI/ML analysis due to legal or ethical concerns) were then asked to describe the issues that had led their organisations to avoid AI/ML analysis. The responses covered issues which have already been addressed elsewhere in this report, notably communication, bias, and governance challenges.

Appendix A – Technical terms

A glossary of definitions is presented here for some of the technical terms used in the report. This glossary provides definitions used for the purpose of the research. We recognise, however, that there may not be commonly accepted definitions for some terms defined here.

Artificial Intelligence and Machine Learning (AI/ML). Techniques that allow computers to learn from data without being explicitly programmed or reprogrammed. It involves algorithms that can adapt and improve over time, learning from experiences (in the form of data inputs) to predict outcomes or make decisions. AI/ML algorithms identify patterns and relationships within data, which can then be used to predict future data or inform decision-making processes.

Classification. An ML task where the output is a category. In this task, the model is trained with a set of examples (input-output pairs) and it learns to predict the output category for a new, unseen example.

Clustering. An ML technique that groups similar ‘instances’ on the basis of features, such that instances belonging to the same group (or cluster) are more similar to each other than those in other groups. For example, clustering could be used to define groups of policyholders who have high similarity on key characteristics.

Collinearity. A statistical phenomenon in which two or more predictor variables in a regression model are highly correlated. This can make it difficult to determine the effect of each predictor variable on the outcome variable and may lead to unstable and unreliable estimates of the model parameters.

Data Scientist. A professional who analyses and interprets large datasets of both structured and unstructured data using scientific methods, employing mathematical, statistical, and computational techniques to gain insights in order to inform data-driven decisions.

Deep Triangle. In actuarial science, this refers to a method that applies deep learning to ‘triangle’ data, such as loss development triangles in insurance. It involves using multi-layer neural networks to make predictions based on the complex, hierarchical structure of the data.

Decision Trees. Decision trees are a type of model used to make predictions based on a series of binary choices or splits. The method is comparable to the well-known game “20 questions”, where each question is designed to narrow down the possibilities until a final answer is reached. A decision tree follows a similar approach, with each split in the tree representing a question or decision about the data, which guides the algorithm to the final prediction.

Explainability. This describes how intrinsically understandable an AI tool is both with and without the application of any technique to explain the tool's workings. Simple AI tools which make use of more simple statistical models and techniques generally have a higher degree of explainability than complex models using more advanced techniques. Black-box AI tools for example can still be explainable but are more likely to require detailed post-hoc analysis in order to explain how a tool has produced a given result.

Extreme Gradient Boosting (XGBoost). A scalable ML system for tree boosting that is known for its efficiency and effectiveness. This is a supervised learning method that enhances weak prediction models (such as decision trees) by combining them into a strong predictive model.

Feature Importance. A technique used in ML that assigns a score to input variables ('features') based on how useful they are in predicting a target variable. It helps to understand which features are most influential in making predictions and can inform the selection or elimination of features for model development.

Fine-Tuning. Fine-tuning in ML refers to the process where a pre-trained model, initially trained on a large, diverse dataset, is subsequently trained on a more specific, often smaller, dataset. This process adapts the broad knowledge of the base model to be more applicable and precise in a particular context or task. Essentially, it's about refining a general-purpose model to perform well on a specific problem or dataset.

Forecasting. This refers to the process of predicting future events or trends based on historical data. In the context of AI/ML models this includes using ML algorithms that identify patterns in past data and extrapolate into the future.

Generalised Linear Models (GLMs). A flexible generalisation of ordinary linear regression that allows for response variables that do not follow a normal distribution. This includes various types of regression models for different types of prediction problems. Generalised linear models can technically be viewed as a type of ML model, but they are not typically regarded as such.

Generative AI. This refers to a category of AI techniques and models that are designed to generate new and original content - such as images, text, audio - that is similar to the data it has been trained on. These models are built using advanced ML techniques, particularly deep learning, and they are capable of producing content that exhibits certain patterns, styles, or characteristics learned from the training data.

Large Language Models (LLMs). These are a type of natural language processing model that are trained on extensive amounts of text data. They can generate human-like text ranging from poems and short stories to programming code and instruction manuals. GPT – the model underpinning ChatGPT developed by OpenAI – is likely the most well-known LLM at the time of writing. Other highly sophisticated LLMs currently in use include LaMa (developed by Meta) and PaLM (developed by Google).

Linear Regression. A statistical method that models the relationship between a dependent variable and one or more independent variables. In simple terms, it tries to draw a line that assumes there is a linear relationship between predictor variables and a dependant variable, and estimates what that relationship is (e.g. paying a worker 10% more increases productivity by 20%). This produces linear equations that can then be used to predict future values outside of the training dataset. Linear regression is not generally viewed as an ML technique.

Neural Network. A neural network is an ML model designed to recognise patterns in data. It consists of layers of interconnected nodes, often referred to as “neurons”, that process and transmit information. Each neuron converts inputs to outputs according to its own specific rules. Through iterative training on datasets, these rules are optimised to improve the predictive accuracy of the overall network.

Natural Language Processing (NLP). An area of AI that deals with the interaction between computers and humans through language. NLP enables machines to interpret, generate, and respond to human languages in a way that is both meaningful and useful.

Regularised Regression. This is a type of regression analysis that introduces a ‘penalty term’ that increases when the model includes more predictor variables or places greater reliance on a given predictor variable, thereby leading the model to behave more conservatively when deciding how much a given predictor variable should be used to inform predictions. This penalty term encourages simpler models that are more likely to generalise well to new data.

Reinforcement Learning. A type of ML where a model learns to make decisions by performing certain actions in a simulated environment to maximise a ‘reward’ signal that is defined by the model builder as a function of the consequences of the decisions taken by the model. The learning process involves trial and error, with successful actions being reinforced to improve the future performance of the model.

SHAP (SHapley Additive exPlanations) Analysis. This refers to a unified measure of feature importance, developed based on cooperative game theory, that attributes the change in prediction outcome to each contributing feature. It interprets the impact of having a certain value for a given feature in comparison to the prediction we would make if that feature took some baseline value. SHAP values are fair, consistent, and locally accurate attributions that provide interpretable analysis of model predictions. Not only does SHAP analysis highlight the importance of a feature, but it also indicates the direction of association (positive or negative) for individual instances.

Supervised learning. This is a ML technique where the model is trained using labelled data, which means that each data point contains features and an associated output. This allows the model to make decisions or predictions based on examples with known outcomes.

Tree boosting. This refers to a ML technique that combines multiple decision trees to make a more accurate and robust predictive model.

Unsupervised learning. This is a ML approach where the model is trained using unlabelled data, where unlabelled data does not have explicit or predefined output labels or target variables. This means that the model doesn't have information about the correct answers and discovers patterns or relationships without predefined or known outcomes.

Appendix B – Use of AI/ML techniques in different areas in actuarial fields

Use of different AI/ML techniques across fields

The plots below show the number of people who said they used a given technique in each area within a field. Note that one individual could report using a given technique in multiple fields. Therefore, the row totals in the plots will not reconcile with the 'number of respondents using technique' plot in the main text because the former avoids double counting these multi-areas uses.

These plots are based on responses to survey questions 9, 12, 15 and 18 (see the full survey here), which asked respondents to provide binary responses indicating whether they used a given technique in a given area within their fields. Note that while the basic form of the questions was the same for all respondents, the specific areas that respondents were asked about varied depending on their field. For 'Other' actuaries, the questions asked about the use of different techniques without specifying any breakdown across different areas.

Figure 14 - Use of different techniques among General Insurance actuaries

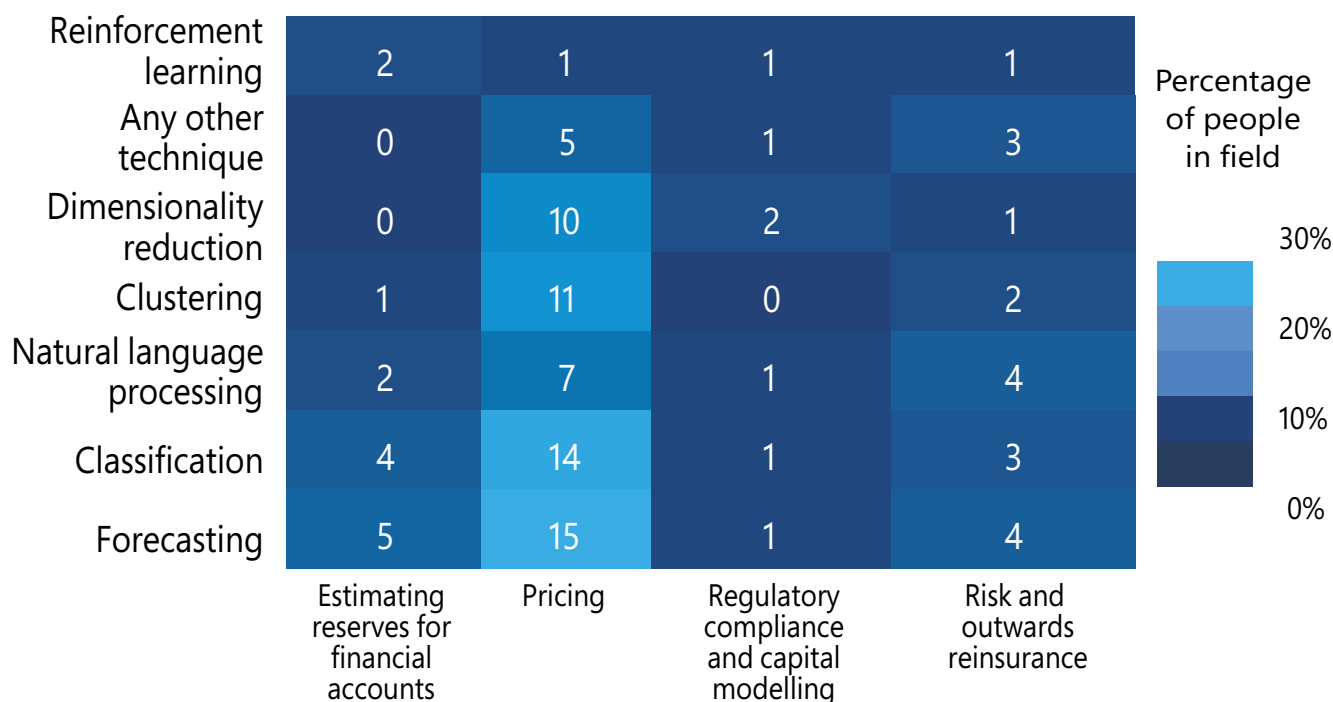


Figure 15 - Use of different techniques among Life actuaries

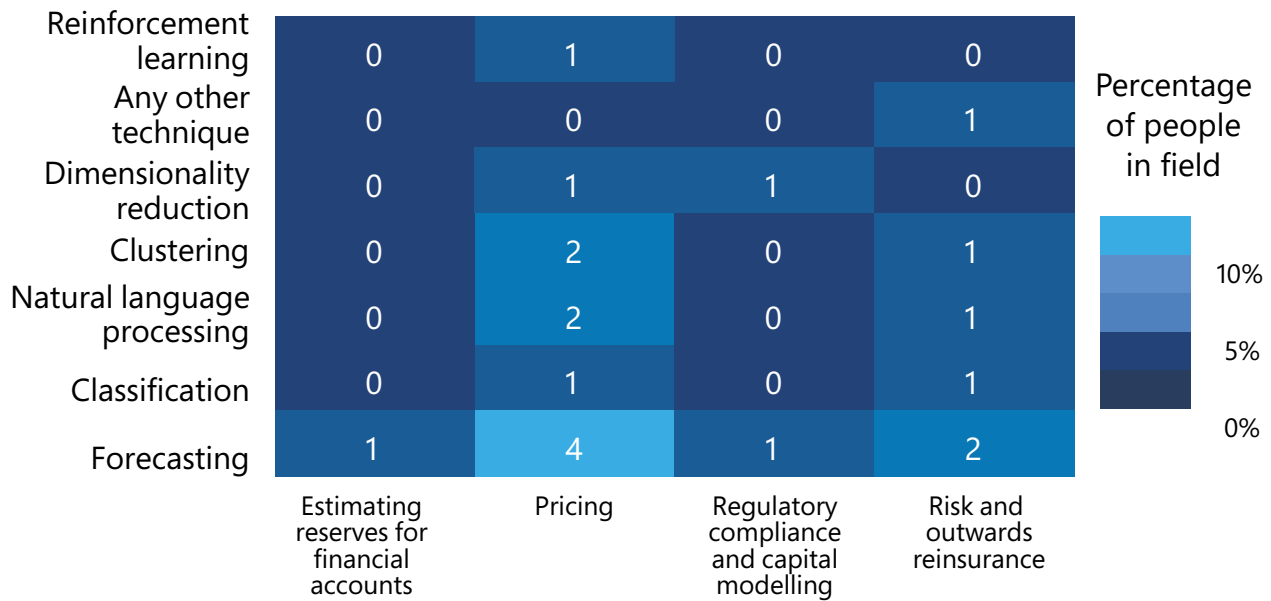


Figure 16 - Use of different techniques among Finance & Investment actuaries

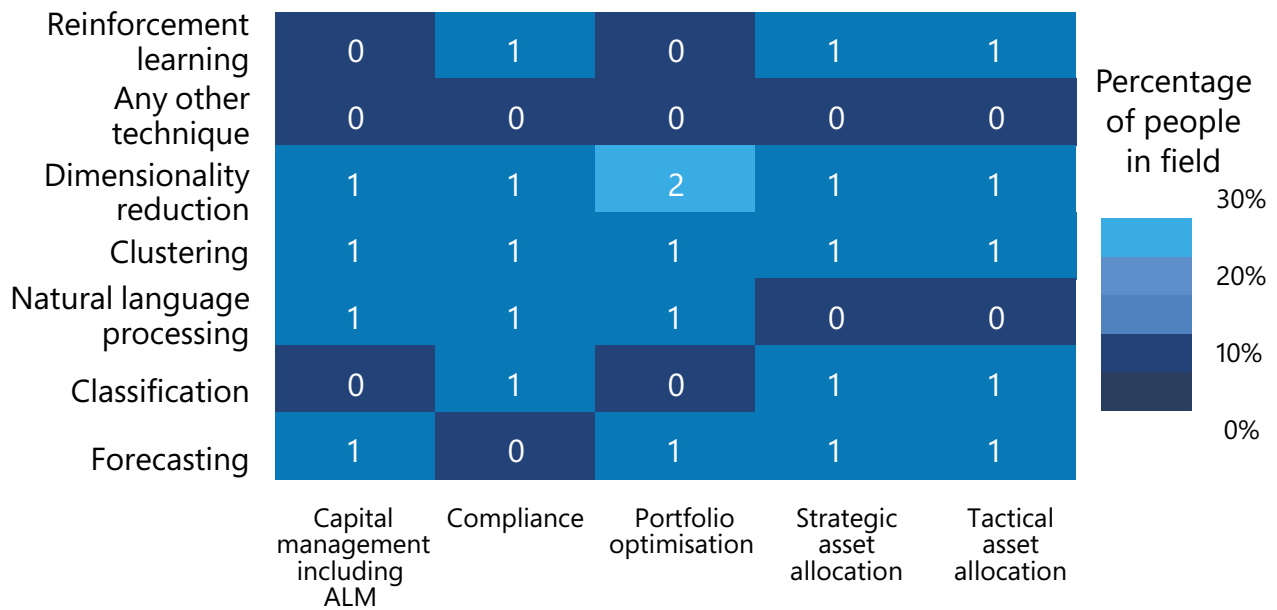


Figure 17 - Use of different techniques among Pensions actuaries

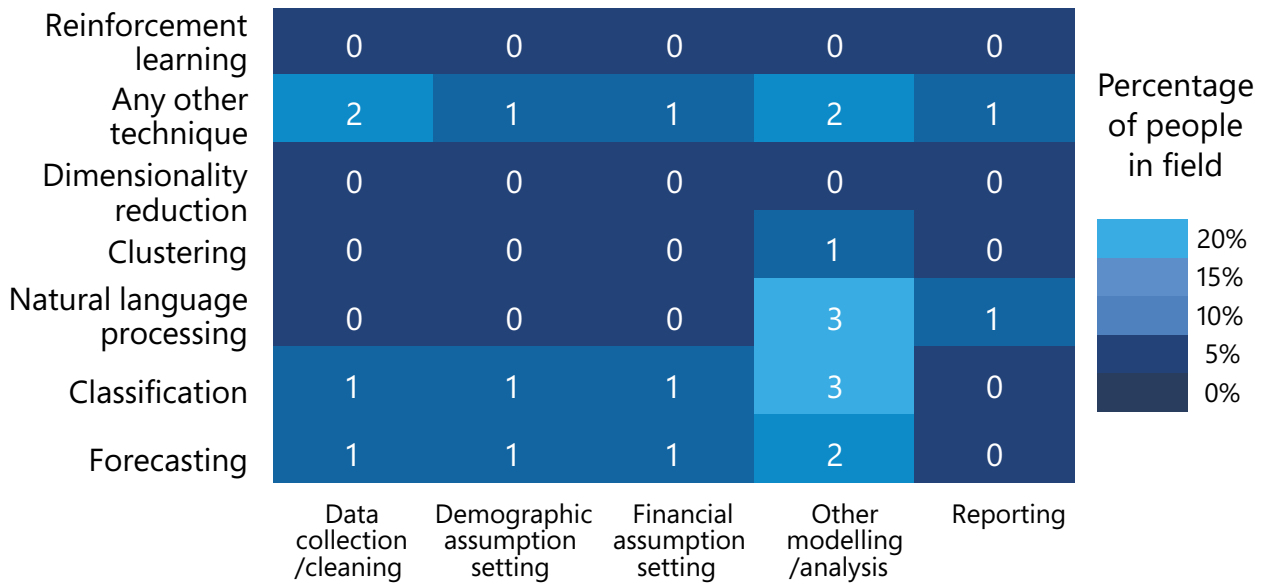
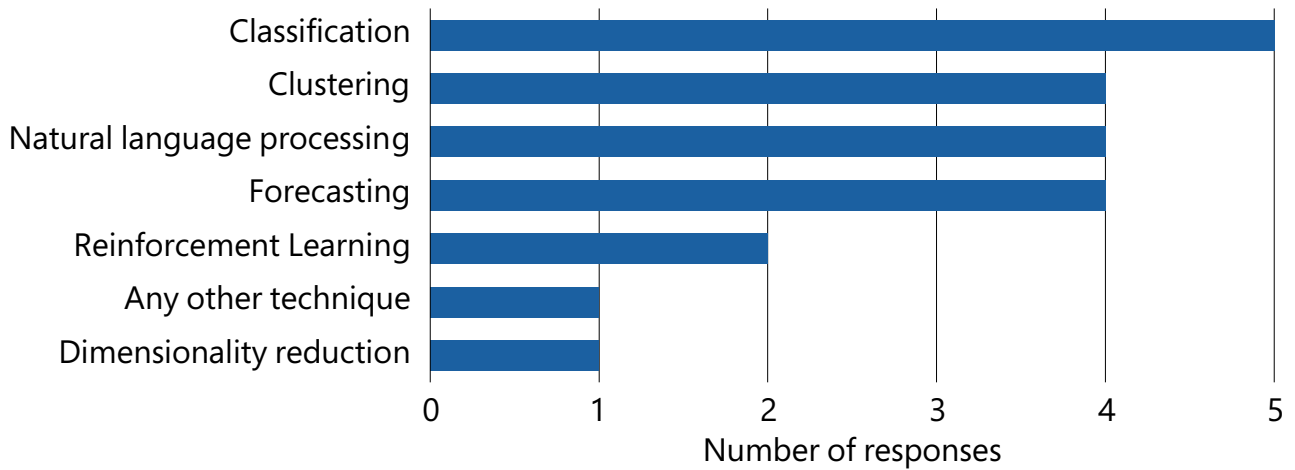


Figure 18 - Use of different techniques among "Other" actuaries







Financial Reporting Council

**Financial
Reporting Council**

8th Floor
125 London Wall
London, EC2Y 5AS

+44 (0)20 7492 2300
www.frc.org.uk

Follow us on
 Twitter @FRCnews
or  LinkedIn