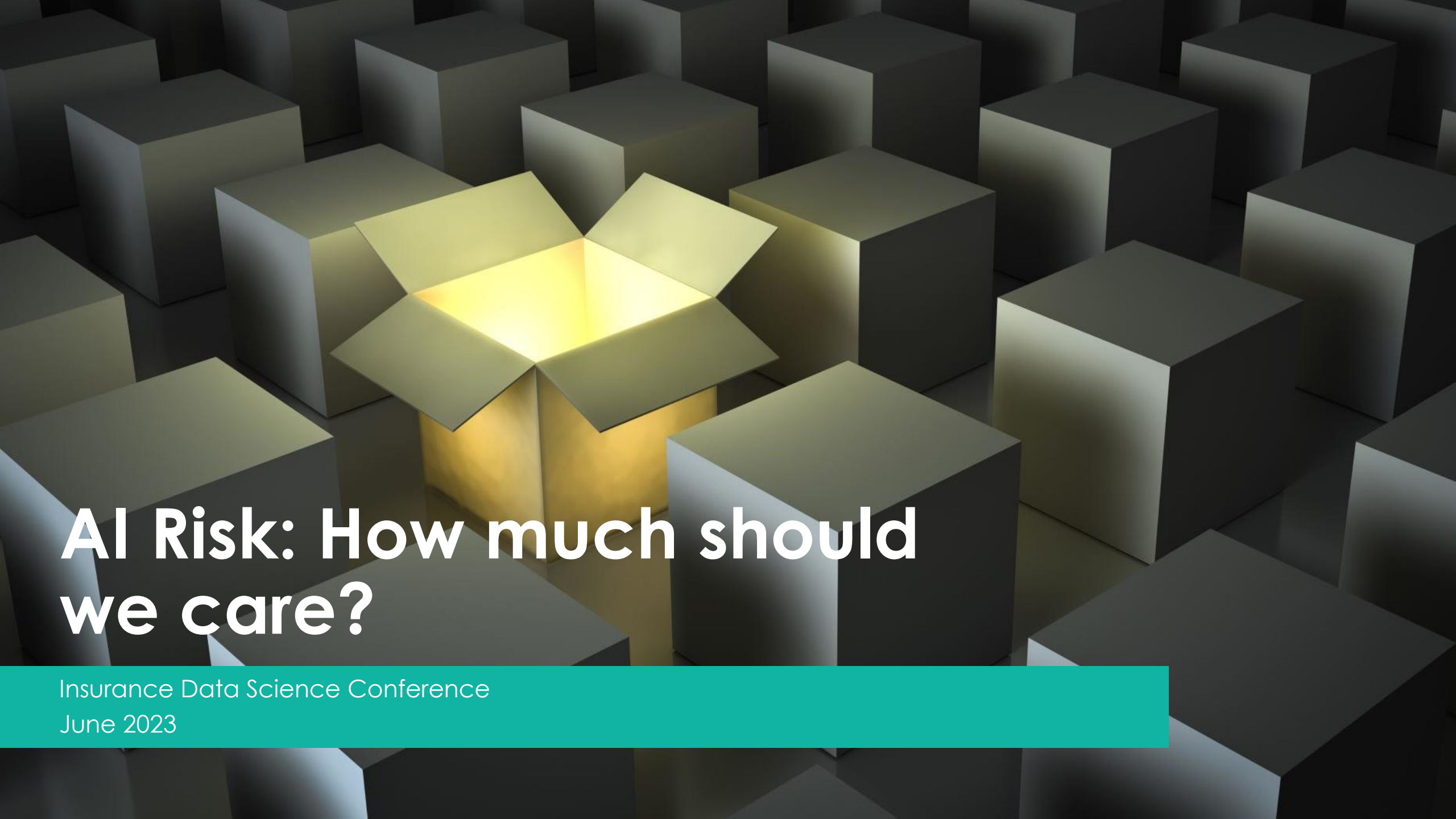


AI Risk: How much should we care?

Insurance Data Science Conference
June 2023

Presenters

- Valerie du Preez
- Paul King

We explore some of the risks related to Artificial Intelligence (AI) from an actuarial perspective based on research from **a transregional and interdisciplinary industry focus group** (see slide 9). The group considered potential issues and challenges of using AI, related specific themes covered in this deck.

The purpose of this research is to provide observations that could help inform industry and professional guidelines or discussion; or to support industry practitioners. It is not intended to replace current regulation, actuarial standards or guidelines.

The accompanying paper will be made available at www.actuartech.com/assets.

Agenda

Introduction

- Current AI landscape

Findings

- Themes in AI risk management and regulation
- Indicative example: Fairness

Conclusion

- Questions and comments

Current AI landscape

AI should be 'a global priority on par with pandemics and nuclear war': Dozens of academics - including the creator of ChatGPT - sign new letter calling for tech to be regulated

'Godfather of A.I.' leaves Google after a decade to warn society of technology he's touted

In A.I. Race, Microsoft and Google Choose Speed Over Caution

Regulation 'critical' to curb risk posed by AI, boss of ChatGPT tells Congress

Meta: Facebook owner fined €1.2bn for mishandling data

But there is development regarding regulation:

European Parliament Advances Groundbreaking AI Regulation: The European AI Act

UK Government Adopts a "Pro-Innovation" Approach to AI Regulation

G7 calls for developing global technical standards for AI

Current AI landscape

Despite the progress made in AI regulation, there are still **key areas** that require attention in the i**nsurance industry**:

- Lindholm et al. (2023):

  "**Indirect discrimination** and **fairness are major concerns** in algorithmic models. This is particularly true in insurance, where protected policyholder attributes are not allowed to be used for insurance pricing."

- Bank of England (2022):

  "There is a risk that the use of **AI** could be associated with **discriminatory decisions**. **Bias** related to protected characteristics such as race or sex, could arise inadvertently during model development."

  "There are a number of **challenges and risks** related to the use of AI, which may **amplify prudential risks** (credit, liquidity, market, operational, reputational, etc) and have implications for the safety and soundness of firms."

- Actuaries Institute and Australian Human Rights Commission (2022) - "Actuaries highlight need for ethical use of AI in insurance" with release of guidance resource on AI and discrimination in insurance in pricing and underwriting: **"While artificial intelligence (AI), promises faster and smarter decision making, safeguards are needed to prevent potential discrimination".**

Themes in AI risk management and regulation

Indicative example: Fairness

When developing and implementing AI, regulation and guidance requires of the actuary to **minimise bias** and to ensure **fairness** and that there is **no unethical discrimination**, but how does the actuary do this in a manner that also balances **accuracy?**

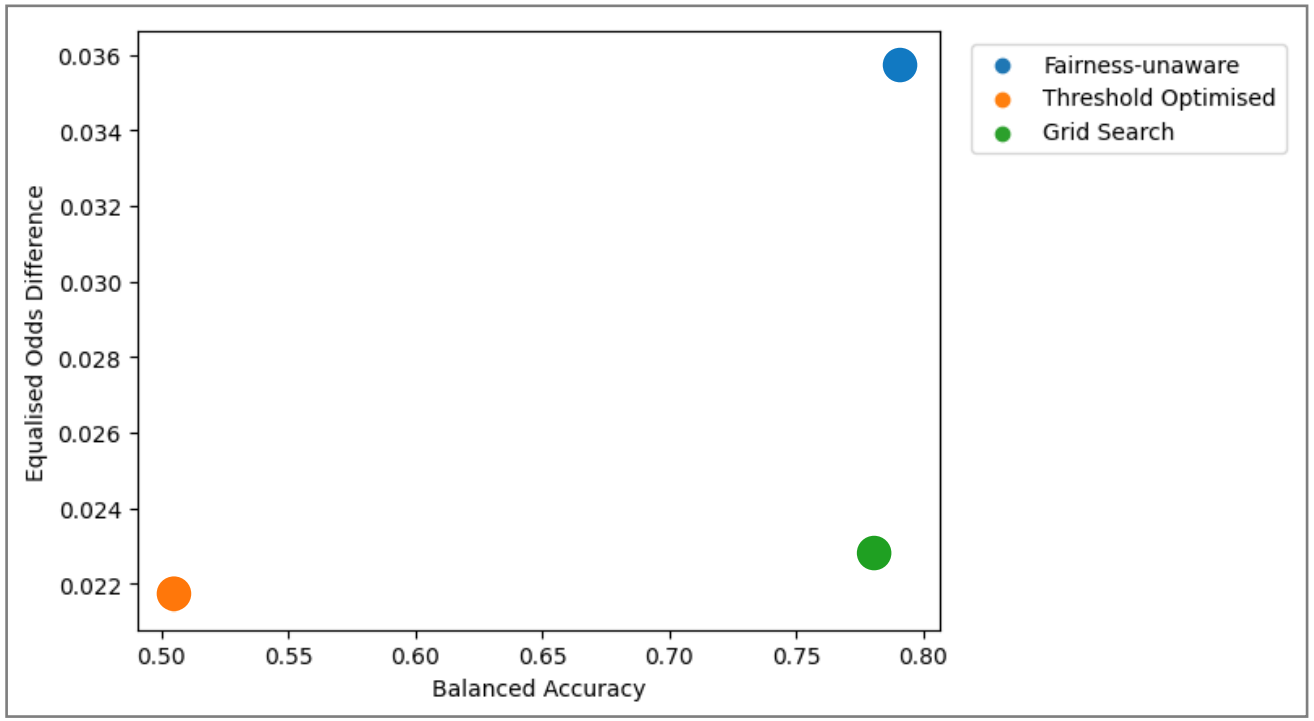


*Figure 1: Graph showing the accuracy and a fairness measure (equalised odds difference) of a GBM model with three different training approaches: **maximised accuracy** (fairness-unaware), **maximised fairness** (using the threshold optimisation method), and **balanced fairness-accuracy** (using the grid search method).*

Issue:

Unfairness can creep in in various areas in our modelling process, from sampling bias, to proxy discrimination, to unintentionally unfair outcomes for protected classes.

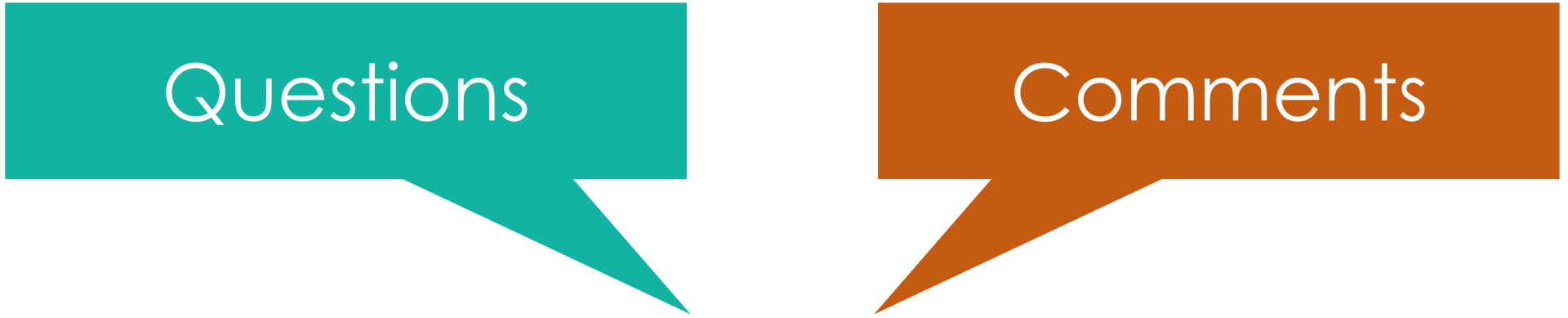
In the indicative example, we predicted whether individuals would claim using a synthetic driver telematics data. Gender is the protected feature for which we aim to manage outcomes, and was excluded as a training feature.

Best practice recommendations:

- Decide on an appropriate **definition of fairness** and suitable **fairness measures** to include, e.g. academic notions incl. demographic parity, **equalised odds**, or self-defined notions such as impact on premiums
  - Equalised odds difference (see Fig. 1) is the difference between model performance for subgroups of protected features e.g. the difference between false negatives for two subgroups (male and female) (Bird et al. 2020).
- Modify the training process to balance fairness and accuracy i.e. include **further optimisation constraints**
- Design **fair classifiers** using specific fairness notions

How do you define an acceptable level of bias?

Conclusion



Research Contributors

Research Contributors:

Shaun Bennett	Matthew Byrne	Aurelién Couloumy
Jean Dessain	Valerie du Preez	Richard Galbraith
Paul King	Victor Mutanga	Frank Schiller
Stefan Zaaiman		

Additional Contributors:

Jim Baxter – Additional Contributor	Arijit Das – Modeller
Lawrence de Jesus – Additional Contributor	Christophe Dutang - Reviewer
Ronald Richman – Advisor and Reviewer	Anton Scharnick – Additional Contributor
Fabian Transchel - Reviewer	Patrick Moehrke
Lara van Heerden	

Sources

- Actuaries Institute. 2022. Artificial intelligence and anti-discrimination: Major new publication. Available at: https://www.actuaries.asn.au/Library/MediaRelease/2022/FINALMRAHRC.pdf
- Actuaries Institute and The Australian Human Rights Commission. 2022. *Artificial intelligence and discrimination in insurance in pricing and underwriting.*
- Andrew, R. 2023. European Parliament Advances Groundbreaking AI Regulation: The European AI Act. *CityLife*. Available at: https://citylife.capetown/ai/european-parliament-advances-groundbreaking-ai-regulation-the-european-ai-act/22523/
- Bank of England. 2022. DP5/22 - Artificial Intelligence and Machine Learning. Available at: https://www.bankofengland.co.uk/prudential-regulation/publication/2022/october/artificial-intelligence
- Bird, S., Dudik, M., Horn, B., Lutz, R., Milan, V., Sameki, M., Wallach, H., and Walker, K. 2020. Fairlearn: A toolkit for assessing and improving fairness in AI. Available at: https://www.microsoft.com/en-us/research/publication/fairlearn-a-toolkit-for-assessing-and-improving-fairness-in-ai/
- Covington & Burling LLP. 2023. UK Government Adopts a "Pro-Innovation" Approach to AI Regulation. *Lexology*. Available at: https://www.lexology.com/library/detail.aspx?g=71ac875b-ec94-4881-8268-9644a02a1382
- Elias, J. 2023. 'Godfather of A.I.' leaves Google after a decade to warn society of technology he's touted. *CNBC*. Available at: https://www.cnbc.com/2023/05/01/godfather-of-ai-leaves-google-after-a-decade-to-warn-of-dangers.html
- Grant, N. and Weise, K. 2023. In A.I. Race, Microsoft and Google choose Speed over Caution. *New York Times.* Available at: https://www.nytimes.com/2023/04/07/technology/ai-chatbots-google-microsoft.html
- Jamieson, A. 2023. Regulation 'critical' to curb risk posed by AI, boss of ChatGPT tells Congress. *The Independent*. Available at: https://www.independent.co.uk/independentpremium/business/ai-regulation-sam-altman-congress-openai-chatgpt-b2340077.html
- Komiya, K. and Mukherjee, S. 2023. G7 calls for adoption of international technical standards for AI. *Reuters*. Available at: https://www.reuters.com/world/g7-calls-adoption-international-technical-standards-ai-2023-05-20/
- Liberatore, S. 2023. AI should be 'a global priority on par with pandemics and nuclear war': Dozens of academics - including the creator of ChatGPT - sign new letter calling for tech to be regulated. *Daily Mail*. Available at: https://www.dailymail.co.uk/sciencetech/article-12140103/AI-global-priority-alongside-pandemics-nuclear-war-new-letter-states.html
- Lindholm, M., Richman, R., Tsanakas, A., and Wüthrich, M.V. 2023. What is Fair? Proxy Discrimination vs. Demographic Disparities in Insurance Pricing. Available at: https://ssrn.com/abstract=4436409
- McCallum, S. 2023. Meta: Facebook owner fined €1.2bn for mishandling data. *BBC*. Available at: https://www.bbc.com/news/technology-65669839
- Prince, A.E.R. and Schwarcz, D. 2020. Proxy Discrimination in the Age of Artificial Intelligence and Big Data. *Iowa Law Review*, 105(3): 1257-1318.

Further Regulation and Industry Guidelines

- Centre for Data Ethics and Innovation. 2021. The roadmap to an effective AI assurance ecosystem. Available at: https://www.gov.uk/government/publications/the-roadmap-to-an-effective-ai-assurance-ecosystem
- Department for Science, Innovation and Technology. 2023. A Pro-innovation approach to AI Regulation. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1146542/a_pro-innovation_approach_to_AI_regulation.pdf
- Department for Science, Innovation and Technology. 2023. Initial £100 million for expert taskforce to help UK build and adopt next generation of safe AI. Available at: https://www.gov.uk/government/news/initial-100-million-for-expert-taskforce-to-help-uk-build-and-adopt-next-generation-of-safe-ai
- Digital Regulation Cooperation Forum. 2022. Auditing algorithms: the existing landscape, role of regulators and future outlook. Available at: https://www.gov.uk/government/publications/findings-from-the-drcf-algorithmic-processing-workstream-spring-2022/auditing-algorithms-the-existing-landscape-role-of-regulators-and-future-outlook
- European Insurance and Occupational Pensions Authority (EIOPA). 2021. Artificial Intelligence Governance Principles: Towards Ethical and Trustworthy Artificial Intelligence European Insurance Industry. A report from EIOPA's Consultative Expert Group on Digital Ethics in insurance. Available at: https://www.eiopa.europa.eu/system/files/2021-06/eiopa-ai-governance-principles-june-2021.pdf
- *Equality Act 2010.* Available at: https://www.legislation.gov.uk/ukpga/2010/15/contents
- Financial Conduct Authority (FCA). 2022. FG22/5 Final non-Handbook Guidance for firms on the Consumer Duty. Available at: https://www.fca.org.uk/publication/finalised-guidance/fg22-5.pdf
- Institute and Faculty of Actuaries (IFoA). 2010. Charter of the Institute and Faculty of Actuaries. Available at: https://www.actuaries.org.uk/system/files/documents/pdf/new-charter.pdf
- Institute and Faculty of Actuaries (IFoA) and Royal Statistical Society (RSS). 2019. A Guide for Ethical Data Science A collaboration between the Royal Statistical Society (RSS) and the Institute and Faculty of Actuaries (IFoA). Available at: https://www.actuaries.org.uk/system/files/field/document/An%20Ethical%20Charter%20for%20Date%20Science%20WEB%20FINAL.PDF
- *Insurance Act 2015*. Available at: https://www.legislation.gov.uk/ukpga/2015/4/contents/enacted
- Levene, M. 2023. How ISO/IEC 42001 guides Organisations toward Trustworthy AI. *AI Standards Hub*. Available at: https://aistandardshub.org/iso-iec-42001-trustworthy-ai
- Peters, S. and Van Den Brink, G. 2023. Can AI Go to Jail? – Consequences for the Corporate Governance. *EAA e-Conference on Data Science & Data Ethics*. Available at: https://actuarial-academy.com/Documents/Abstracts/E0338_Abstract_e-Conference_Recording_Peters+vandenBrink.pdf
- The European Parliament and the Council of the European Union. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L 119/1. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679
- World Economic Forum. 2020. Empowering AI Leadership: AI C-Suite Toolkit. Availability at: https://www.weforum.org/reports/empowering-ai-leadership-ai-c-suite-toolkit/
- Usher, K. on behalf of the IFoA AI Ethics Risk Working Party. 2023. Man vs Machine. *The Actuary* (May 2023). Available at: https://www.theactuary.com/2023/05/04/man-vs-machine