OPINION ON AI GOVERNANCE AND RISK MANAGEMENT

EIOPA-BoS-25-360
06 August 2025

European Insurance and
Occupational Pensions Authority

1. LEGAL BASIS

1.1. The European Insurance and Occupational Pensions Authority (EIOPA) provides this Opinion on the basis of Article 29(1)(a) of Regulation (EU) No 1094/2010[1]. This Article mandates EIOPA to play an active role in building a common Union supervisory culture and consistent supervisory practices, as well as in ensuring uniform procedures and consistent approaches throughout the Union.

1.2. EIOPA delivers this Opinion on the basis of Articles 17, 20 and 25 of the Directive (EU) 2016/972 (Insurance Distribution Directive),[2] Articles 41, 46 and 82 of the Directive 2009/138/EC (Solvency II Directive),[3] Articles 4, 5, 6, 11 and 12 of the Regulation (EU) 2022/2554 (Digital Operational Resilience Act),[4] Articles 258 and 260 of the Commission Delegated Regulation 2015/35,[5] and Articles 6, 7, 8 and 9 of the Commission Delegated Regulation 2017/2358.[6]

1.3. This Opinion is addressed to the competent authorities, as defined in Article 4(2) of the Regulation (EU) No 1094/2010, and covers the activities of both insurance undertakings and intermediaries (hereafter jointly referred as 'undertakings'), insofar as they may use AI systems within their respective areas of competence in the insurance value chain.

2. CONTEXT, OBJECTIVE AND SCOPE

2.1. Artificial Intelligence (AI) is expected to play a pivotal role in the ongoing digital transformation in all industries, including the insurance sector, where there is a trend towards the increasing use of AI systems throughout the insurance value chain, including pricing, underwriting, claims management and fraud detection.

[1] Regulation (EU) No 1094/2010 of the European Parliament and of the Council of 24 November 2010 establishing a European Supervisory Authority (European Insurance and Occupational Pensions Authority), amending Decision No 716/2009/EC and repealing Commission Decision 2009/79/EC (OJ L 331, 15.12.2010, p. 48).

[2] Directive (EU) 2016/97 of the European Parliament and of the Council of 20 January 2016 on insurance distribution (OJ L 26, 2.2.2016, p. 19).

[3] Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance (OJ L 335, 17.12.2009, p. 1).

[4] Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011 (OJ L 333, 27.12.2022, p. 1–79).

[5] Commission Delegated Regulation (EU) 2015/35 of 10 October 2014 supplementing Directive 2009/138/EC of the European Parliament and of the Council on the taking-up and pursuit of the business of Insurance and Reinsurance (Solvency II) (OJ L 12, 17.1.2015, p. 1).

[6] Commission Delegated Regulation (EU) 2017/2358 of 21 September 2017 supplementing Directive (EU) 2016/97 of the European Parliament and of the Council with regard to product oversight and governance requirements for insurance undertakings and insurance distributors (OJ L 341, 20.12.2017, p. 1).

2.2. AI offers significant opportunities for the insurance sector such as faster and automated claims handling processes, the development of more accurate and granular risk assessments, or combatting customer fraud more efficiently. However, AI can also bring new risks or increase existing ones, in particular due to the limited explainability of some AI systems, which among other things can increase the risk of bias and discriminatory outputs.

2.3. In July 2024 the Regulation (EU) 2024/1689 (the AI Act)[7] was published in the Official Journal of the European Union. The AI Act applies to all sectors of the economy and aims at ensuring a high level of protection of fundamental rights, health, and safety. The AI Act follows a risk-based approach, classifying AI systems according to different risk levels.

2.4. Among other high-risk AI systems which may be used by undertakings, the AI Act identifies as high-risk the use of AI systems for risk assessment and pricing in relation to natural persons in the case of life and health insurance. Providers and deployers of high-risk AI systems will need to comply with a comprehensive set of governance and risk management requirements foreseen in the AI Act. Limited derogations are introduced to address overlaps with existing sectoral insurance legislation.

2.5. The remaining AI systems in insurance that are not prohibited AI practices and that are not considered to be high-risk, without prejudice to Articles 6(3), 6(4) and 7 of the AI Act, continue to operate subject to existing sectoral legislation without new requirements, with the exception of certain transparency requirements (e.g. need to inform the customer that she/he is interacting with an AI system), the need to promote staff AI literacy, and the development of voluntary codes of conduct.

2.6. The objective of this Opinion is to provide further clarity on the main principles and requirements foreseen in the insurance sectoral legislation that should be considered in relation to those insurance AI systems that are not considered as prohibited AI practices or high-risk under the AI Act. Although insurance legislation such as the Insurance Distribution Directive and the Solvency II Directive, on which this Opinion is based, applies to all AI systems used in insurance, to avoid regulatory complexities and overlaps this Opinion does not cover prohibited AI practices or high-risk AI systems under the AI Act.

2.7. The Opinion follows a principle-based approach and is in line with the underlying principles and requirements of the AI Act and other international initiatives in this area.[8]

[7] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

[8] In addition to the AI Act, the expectations set out in this Opinion are aligned with the work of other international standard setting bodies such as the Organisation for Economic Co-operation and Development (OECD) (link), the G20 (link), or the International Association of

2.8. This Opinion does not set out new requirements and does not seek to alter the scope of the AI Act by extending the requirements of the AI Act for high-risk AI systems to all AI systems used in insurance. It also does not alter the scope of application of the existing insurance sectoral legislation mentioned in the Opinion.[9]

2.9. This Opinion provides guidance on how to interpret various provisions of the existing insurance sectoral legislation in the context of AI systems which were either non-existent or not widely used when that legislation was approved. It sets out high-level supervisory expectations relating to governance and risk-management systems that undertakings should develop, following a risk-based and proportionate approach, to ensure a responsible use of AI systems. It follows a holistic approach by highlighting the key principles that need to be observed, and which can be embedded into existing risk management frameworks and adapted to the specificities of different AI systems used in the insurance value chain.

2.10. To ensure consistency at European level, this Opinion is based on the definition of AI system adopted in the AI Act,[10] including the European Commission's AI Office Guidelines on the AI system definition.[11] Further clarifications on the AI system definition may be provided by the AI Office at a later stage. EIOPA is engaging with the AI Office and other relevant stakeholders to provide a sectoral perspective. Nevertheless, it is important to highlight that existing insurance sectoral legislation requires adequate and proportionate governance and risk management measures when using mathematical models, regardless of whether they are considered AI systems or not.

Insurance Supervisors (IAIS) (link). The Opinion also leverages on the AI governance principles report developed by EIOPA's stakeholder group on digital ethics in insurance in 2021 (link).

[9] This Opinion focuses on the main provisions in insurance sectoral legislation within EIOPA's remit that are relevant to the use of AI systems, but it should be noted that other legislations such as Regulation (EU) 2016/679 (General Data Protection Regulation; GDPR) may also include provisions relevant to the use of AI systems.

[10] Article 3(1) of the AI Act defines AI system as "a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments".

[11] Paragraph 42 of the Commission Guidelines on the definition of an artificial intelligence system established by Regulation (EU) 2024/1689 (AI Act) (link) states that "systems used to improve mathematical optimisation or to accelerate and approximate traditional, well established optimisation methods, such as linear or logistic regression methods, fall outside the scope of the AI system definition. This is because, while those models have the capacity to infer, they do not transcend 'basic data processing'. An indication that a system does not transcend basic data processing could be that it has been used in consolidated manner for many years. This includes, for example, machine learning-based models that approximate functions or parameters in optimization problems while maintaining performance. The systems aim to improve the efficiency of optimisation algorithms used in computational problems. For example, they help to speed up optimisation tasks by providing learned approximations, heuristics, or search strategies."

3. AI GOVERNANCE AND RISK MANAGEMENT FRAMEWORK

RISK-BASED APPROACH AND PROPORTIONALITY

3.1. According to Article 41 of the Solvency II Directive, insurance undertakings need to have in place an effective system of governance which provides for a sound and prudent management of the business, which shall be proportionate to the nature, scale, and complexity of the operations of the insurance or reinsurance undertaking. In a similar line, Article 25 of the Insurance Distribution Directive (IDD) requires undertakings to maintain, operate and review a process for the approval insurance products which shall be proportionate and appropriate to the nature of the insurance product. Furthermore, Articles 5 and 6 of the Digital Operational Resilience Act (DORA) require financial entities to have in place an internal ICT governance and risk management frameworks, in accordance with the principle of proportionality as set out in Article 4 of the DORA.

3.2. As a first step, for those AI systems that are within the scope of this Opinion, undertakings should assess the risk of the different AI systems used; it is acknowledged that there are varying levels of risks amongst those AI systems that are not prohibited or considered as high-risk under the AI Act. Therefore, undertakings should assess their risks and develop governance and risk management measures adequate and proportionate to the characteristics and risks of the specific use of AI systems at hand. The assessment should be conducted in a manner that is proportionate to the potential impact of the specific AI system on customers and on the undertakings themselves i.e. those AI systems where a minimal impact can be anticipated may be subject to a relatively simple and streamlined assessment, whereas those likely to have a greater impact should undergo a more comprehensive evaluation.

3.3. The impact assessment should take into account criteria such as the processing of data on a large scale, the sensitivity of the data, the number of customers (including vulnerable customers) affected, the extent to which the AI system can act autonomously, the extent to which the AI system is used in consumer-facing applications or for purely internal purposes that do not involve decision-making with a direct customer impact, or the potential adverse impact that an AI system could have on the individual (e.g. right to non-discrimination). Insurance-specific criteria should also be taken into consideration, including where certain categories of personal data need to be used (e.g. the age of the customer) to underwrite risks in insurance, or the extent to which an AI system is used in a line of business that is important for the financial inclusion of customers or which is compulsory by law.

3.4. Undertakings should also assess prudential considerations such as the extent to which an AI system is used in critical activities that can impact the business continuity of an insurance undertaking. The extent to which an AI system can have an impact on the financial position of an undertaking (e.g. substantial number of claims, contracts, Gross Written Premiums, solvency ratios etc.), or on the legal obligations of an undertaking is also relevant. Reputational risks that could potentially

arise from the use of AI systems should also be considered. Annex I provides additional examples of indicators that may be used to assess the impact of the use of AI systems.

3.5. As a second step, taking into account the impact assessment mentioned in the previous paragraphs, undertakings should develop a set of proportionate measures that aim to ensure the responsible use of the AI system. This implies that governance and risk management measures may be tailored to the specific use of AI systems at hand to achieve the desired outcome i.e. the proportionality principle is applicable to all the governance and risk management measures described in this Opinion.

3.6. For example, in application of the principle of proportionality, the supervisory expectations in terms of governance and risk management (e.g. data governance, human oversight or explainability etc.) for AI systems that have a low or very limited impact on customers or undertakings themselves would be very limited, as opposed to those AI systems that pose higher risks and that would be subject to more stringent expectations. More specifically, and taking into account that there are often trade-offs between the accuracy and explainability of AI systems, for certain AI systems such as those used to process images, videos, or text, for which it is not possible to comprehensively explain how a certain output was obtained and for which there are no suitable alternatives, complementary risk management measures such as data governance or human oversight may be developed to compensate for a lack of explainability.

RISK MANAGEMENT SYSTEM

3.7. In line with Article 41 of the Solvency II Directive, Article 25 of the IDD and Articles 4, 5, and 6 of the DORA, in order to ensure a responsible use of AI systems that maximise the benefits and minimises the risks of AI systems, undertakings should develop risk-based and proportionate governance and risk management systems, considering the following areas:

- Fairness and ethics
- Data governance
- Documentation and record keeping
- Transparency and explainability
- Human oversight
- Accuracy, robustness and cybersecurity

3.8. The responsible use of AI systems is not achieved by a standalone measure, but by a combination of different risk management measures. This holistic approach implies that the AI governance and risks management measures mentioned above should be tailored to specific AI systems used, that they are complementary to one another, and that cross-references and dependencies between them will be common, as reflected in this Opinion.

3.9. Undertakings need to define and document the approach to the use of AI systems across the organisation, including the governance and risk management measures that should be applied

throughout the entire lifecycle of an AI system. Undertakings may leverage in this regard on existing or updated Enterprise Risk Management (ERM), model risk management, Product Oversight and Governance (POG) frameworks, or other policy or strategy approaches (e.g. through specific IT, Data, or AI frameworks etc.), insofar as these reflect the key principles outlined in this Opinion. The undertaking's approach to AI systems should be regularly reviewed, in particular if the number, type and materiality of AI systems used within the organisation changes.

3.10. The approach to AI systems should also include frameworks where the roles and responsibilities of different staff and the interplay between them are clearly defined (see also Human oversight section below).

3.11. Undertakings are ultimately responsible for the AI systems that they use, regardless of whether the AI systems are developed in-house or in collaboration with third party service providers. However, third-party service providers also have a role to play; undertakings should obtain adequate information and assurances from third-party service providers about the characteristics, capabilities, data used to train and test the AI systems, and the limitations of the AI systems used. Where it is challenging to implement certain AI governance and risk management measures (e.g. data governance or explainability) due to the intellectual property rights of third-party service providers, undertakings should mitigate consequent risks by implementing complementary governance measures and by adopting other measures such as including appropriate clauses in contracts and service level agreements, conducting external audits, or performing due diligence testing and monitoring.

FAIRNESS AND ETHICS

3.12. Article 17 of the IDD stipulates that insurance distributors shall always act honestly, fairly and professionally in accordance with the best interests of their customers. Moreover, EIOPA's 2023 Supervisory Statement on Differential Pricing Practices[12] outlines certain pricing practices that are not considered compliant with the requirement to treat customers fairly and also provides guidance on the governance and risk management measures that insurers need to develop to mitigate risks.

3.13. Following a risk-based and proportionate approach, undertakings should adopt a customer-centric approach to the use of AI systems throughout their entire lifecycle and across the value chain so that customers are treated fairly and according to their best interest. This includes developing a corporate culture, documented in policies and procedures, that includes ethics and fairness guidance and training for relevant staff (see the human oversight section below).

[12] https://www.eiopa.europa.eu/system/files/2023-03/EIOPA-BoS-23-076-Supervisory-Statement-on-differential-pricing-practices_0.pdf

3.14. Undertakings should also adopt sound data governance policies (see the data governance section below), including by making reasonable efforts to remove biases in the data, including potential unlawful proxy discriminatory variables[13]. The outputs of AI systems should also be meaningfully explainable to identify and mitigate potential bias (see the explainability section below).

3.15. The outcomes of AI systems should also be regularly monitored and, where appropriate, audited, including with the use of fairness and non-discrimination metrics (see examples of metrics for higher risk uses of AI systems in Annex I).

3.16. Adequate redress mechanisms (e.g. procedures to submit complaints)[14]should also be in place to enable customers to access and seek redress when they have been harmed by an AI system.

DATA GOVERNANCE

3.17. According to Article 260(1)(a)(ii) of the Commission Delegated Regulation 2015/35, the risk management system should include policies regarding the sufficiency and quality of relevant data for underwriting and reserving processes. Article 82 of the Solvency II Directive stipulates that data shall be complete, accurate and appropriate for calculating the technical provisions. Furthermore, Article 6(1) of the Commission Delegated Regulation 2017/2358 required that manufacturers test their insurance products appropriately, including scenario analyses where relevant.

3.18. Following a risk-based and proportionate approach, undertakings should implement a data governance policy for AI systems in compliance with applicable insurance and data protection legislation.

3.19. The data used to train and test the AI system should be complete (e.g. sufficient historical information), accurate (e.g. no material errors) and appropriate (e.g. consistent with the purposes for which it is to be used). Any limitations of data in this regard should be duly documented and addressed. In particular, undertakings should make reasonable efforts to remove biases in the data in line with the undertaking's policy.

3.20. Sound data governance should be applied throughout the AI system life cycle for data collection, data processing and post processing.

---

[13] Proxy discrimination arises when sensitive customer characteristics (e.g. ethnicity), whose use is not permitted, are indirectly inferred from other customer characteristics (e.g. location) that are considered legitimate. As noted by the Commission Guidelines on the application of Council Directive 2004/113/EC to insurance in the light of the judgment of the Court of Justice of the European Union in Case C-236/09 (Test-Achats) (link), proxies should be removed unless their use is objectively justified by a legitimate aim, and it is appropriate and necessary. The Commission explains this situation with the following examples: price differentiation based on the size of a car engine in the field of motor insurance should remain possible, even if statistically men drive cars with more powerful engines. On the contrary, it is not possible to price differentiation based on the size or weight of a person in relation to motor insurance (men are commonly taller and heavier than women).

[14] Redress mechanisms do not need to be specific to AI systems; complaints mechanisms that undertakings already have in place would normally suffice to enable customers to seek redress when they have been harmed by the use of an AI system in insurance.

3.21. If the undertaking makes use of external data acquired from a third party, the same data quality standards should apply (see also paragraph 3.11 above regarding third parties).

DOCUMENTATION AND RECORD KEEPING

3.22. Article 258(1)(i) of the Commission Delegated Regulation 2015/35 requires that insurance undertakings maintain adequate and orderly records of the insurance undertaking's business and internal organisation. Furthermore, Article 9 of the Commission Delegated Regulation 2017/2358 requires that relevant actions taken by undertakings in relation to their product approval process are duly documented, kept for audit purposes and made available to the competent authorities upon request.

3.23. Following a risk-based and proportionate approach, undertakings should keep appropriate records of the training and testing data and the modelling methodologies to enable their reproducibility and traceability.

3.24. An example of the types of records and documentation that could be kept and reviewed on a regular basis for higher risk uses of AI systems is provided in Annex I.

TRANSPARENCY AND EXPLAINABILITY

3.25. Pursuant to Article 20(1) of the IDD, undertakings shall provide the customer with objective information about the insurance product in a comprehensible form to allow the customer to make an informed decision. Any contract proposed shall be consistent with the customer's insurance demands and needs. Furthermore, Article 258(h) of the Commission Delegated Regulation 2015/35 stipulates that insurance undertakings shall establish information systems which produce complete, reliable, clear, consistent, timely and relevant information concerning the business activities, the commitments assumed and the risks to which the insurance undertaking is exposed. Moreover, according to Article 8 of the Commission Delegated Regulation 2017/2358 manufacturers shall carefully select distribution channels that are appropriate for the target market, thereby taking into account the particular characteristics of the relevant insurance products.

3.26. Following a risk-based and proportionate approach, undertakings should ensure that the outcomes of AI systems can be meaningfully explained. Different approaches can be used to this extent, such as using explainable AI algorithms instead of more opaque ("black box") ones, or using complex AI systems only for the purpose of challenging and fine-tuning traditional mathematical models. Local and global model-agnostic explanatory tools[15] may also be used to

[15] For example, LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (Shapley Additive Explanations) are two techniques that could be used for explainability in AI. Both focus on providing local explanations, meaning they aim to explain how specific data points or regions within the input data impact the output of an AI system.

explain the inner functioning of complex AI systems, but the assumptions and limitations of these tools should be duly documented and addressed. Statistical or stochastic explanations may also be used instead of deterministic ones when duly justified and documented.

3.27. Undertakings should adapt the explanations to specific uses of AI systems. For certain uses where there are no suitable alternatives, if the complexity of the AI system hinders the full transparency and explainability, the undertaking should put in place, where necessary, complementary risk management measures such as stronger guardrails and increased human oversight. Undertakings should comprehensively secure and test - before release as well as on an ongoing basis - those uses of AI systems that could have a high impact on customers or the solvency of the undertaking.

3.28. The explanations should also be adapted to the needs of different recipient stakeholders. For example, undertakings should be able to provide to competent authorities and auditors a global and comprehensive explanation about the functioning of the AI system. For customers, in addition to being informed that they are interacting with an AI system, upon the customer's request, the influence of the AI system on the decision that has a material impact on them should be clarified using simple, clear and non-technical language to allow them to make informed decisions. Where relevant, insurance intermediaries should also be informed by insurance undertakings when a decision is made on the basis of an AI system so that they can comply with their legal obligations towards customers.

HUMAN OVERSIGHT

3.29. According to Article 46 of the Solvency II Directive, insurance undertakings shall have in place effective internal control systems at all levels of the insurance undertaking. Furthermore, Article 258(2) of the Commission Delegated Regulation 2015/35 requires the insurance undertakings to develop policies on internal control, internal audit and where relevant outsourcing. Also, Article 7 of the Commission Delegated Regulation 2017/2358 requires that manufacturers continuously monitor and review insurance products.

3.30. Following a risk-based and proportionate approach, undertakings should put in place effective internal control systems during the entire lifecycle of the AI system. Roles and responsibilities should be defined in policy documents, including escalation procedures, involving relevant staff in the necessary steps of the AI system lifecycle, in particular:

- Administrative, management or supervisory body (AMSB) members are responsible for the overall use of AI systems within the organisation, and need to have sufficient knowledge of how AI systems are used in their organisation and the potential risks. They are responsible for defining and internally communicating the vision and policy towards the development and use of AI systems within the organisation.
- The compliance and audit functions verify that the use of AI systems within the organisation is compliant with all applicable laws and regulations.

- The Data Protection Officer verifies that personal data processed by AI systems is processed in compliance with the applicable data protection regulations.
- The actuarial function is responsible for the controls on AI systems that fall under its responsibilities (e.g. for coordination of technical provisions calculation, opinion on the overall underwriting policy).

3.31. Undertakings may decide to create other organisational arrangements that fit their business model, such as outsourcing certain oversight functions (not the responsibility) while ensuring independence between the relevant entities. Undertakings could also decide to appoint an AI officer who provides oversight and advice to all functions, or create an AI or data committee which comprises members with the necessary expertise and ensures coordination, or establish a dialogue with social partners about the implications of the use of AI systems in the undertaking.

3.32. Sufficient training should be provided to relevant staff adapted to their respective roles and responsibilities to ensure that the human oversight of AI systems is effective.

3.33. Human oversight by the relevant staff should support the identification and mitigation of potential biases, in line with the policy of the undertaking. Appropriate guardrails should be established to ensure that the AI system functions as intended, respects customers' rights and maintains a high standard of safety.

ACCURACY, ROBUSTNESS AND CYBERSECURITY

3.34. According to Article 46 of the Solvency II Directive, the insurance undertaking should put in place an effective internal control system. Article 258(1)(j) and (3) of the Commission Delegated Regulation 2015/35 stipulates that the security, integrity and confidentiality of the information shall be safeguarded depending on the nature of the information. Furthermore, Articles 6 to 10 of the DORA require financial entities to have in place sound, comprehensive and well-documented ICT risk management frameworks, Article 11(4) and (6) lays down uniform requirements concerning the security of information and communication technologies (ICT) for the financial sector, including the requirement to establish, implement, maintain and test business continuity plans and Article 12 sets out requirements on business continuity and fall back plans.

3.35. Following a risk-based and proportionate approach, undertakings should define the levels of accuracy, robustness and cybersecurity of AI systems. The AI system should perform consistently in those respects throughout their lifecycle., regardless of whether they have been developed in-house or purchased from third-party service providers (see also paragraph 3.11 above regarding third parties).

3.36. The undertaking should use metrics, including, where appropriate, fairness metrics, to measure the performance (e.g. accuracy, recall etc.) adapted to the AI system in question. These metrics

should support ongoing monitoring and enable the timely identification and remediation of issues such as model drift or data degradation.

3.37. When testing an AI system, undertakings should assess potential connections to other AI systems via Application Program Interface (APIs), since this could impact the overall security and performance of the AI system.

3.38. AI systems should be resilient against attempts by unauthorised third parties to alter their use, outputs or performance by exploiting system vulnerabilities (e.g. data poisoning or adversarial attacks). To this extent, undertakings should have adequate and up-to-date IT infrastructure as well as fall-back plans to ensure ICT business continuity.

4. MONITORING BY EIOPA

4.1. Two years following the publication of this Opinion, EIOPA will look into the supervisory practices of competent authorities with a view to evaluate supervisory convergence.

4.2. EIOPA will continue collaborating with competent authorities to facilitate smooth implementation of regulation applicable to the use of AI in the insurance sector and support competent authorities in their supervisory work.

4.3. Based on the proposed AI governance framework in this Opinion, EIOPA envisages to subsequently develop more detailed analysis on specific AI systems or issues arising from the use of AI systems in insurance and provide further guidance, as appropriate.

4.4. EIOPA will continue monitoring market developments via different tools in close collaboration with stakeholders.

4.5. This Opinion will be published on EIOPA's website.

Done at Frankfurt am Main, on DayMonthYear.

[signed]

For the Board of Supervisors

Petra Hielkema

Chairperson

ANNEX I – EXAMPLES OF IMPACT ASSESSMENT INDICATORS, RECORD KEEPING AND FAIRNESS METRICS

This Annex includes practical examples of impact assessment indicators, record keeping and fairness metrics which have been extracted from the AI governance principles report developed by EIOPA's stakeholder group on digital ethics in 2021.

These examples have been included here for illustrative purposes only and are not intended to be considered a prescriptive guidance.

Indeed, while these examples represent the views of EIOPA's stakeholder group on digital ethics and the undertaking concerned should ultimately develop the governance and risks management framework that best adapts to the nature, scale and complexity of their business model, they illustrate possible practical ways on how to implement some of the high-level principles included in this Opinion.

1. EXAMPLES OF IMPACT ASSESSMENT INDICATORS

	AI Use case Impact Assessment	
	Impact on consumers	Impact on insurance firms
Severity	Number of consumers affected	Business continuity
	Consumer interaction and interests	Financial Impact
	Types of consumers (e.g. vulnerable consumers)	Legal impact
	Human autonomy	Reputational impact
	Anti-discrimination and diversity	
	Insurance line of business relevance	
Likelihood	Evaluation or scoring, including profiling and predicting	
	Automated-decision making with legal or similar significant effect	
	Systematic monitoring	
	Model complexity/combining datasets	
	Innovative use or applying new technological or organisational solution	
	Type and amount of data used	
	Outsourcing datasets and AI applications	

Source: EIOPA Consultative Expert Group on Digital Ethics in insurance

2. EXAMPLE OF RECORD KEEPING FOR HIGHER RISK USES OF AI SYSTEMS

Record	Description
Reasons for using AI	Explanation of the business objective / task pursued by using AI and its consistency with corporate strategies / objectives. Explanation of how this was implemented into the AI system. This would help avoid misuse of the AI system and enable its audit and independent review.

Integration into IT infrastructure	Description of how the model is integrated in the current IT system of the organisation and document any significant changes that could eventually take place
Staff involved in the design and implementation of the AI model	Identification of all the roles and responsibilities of the staff involved in the design and implementation of the AI model as well as their training needs. This would ensure accountability of the responsible persons.
Data collection	Documenting how the ground truth[16] was built including how consideration was given to identifying and removing potential bias in the data. This would include explaining how input data was selected, collected and labelled.
Data preparation	Records of the data used for training the AI model, i.e. the variables with their respective domain range. This would include defining the construction of training, test and prediction dataset. For built (engineered) features, records should exist on how the feature was build and the associated intention.
Data post processing	Description of processes in place to operationalize the use of data and to achieve continuous improvement (including addressing potential bias). Records should specify the timing and frequency of data improvement actions.
Technical choices / arbitration	Documenting why a specific type of AI algorithm was chosen and not others, as well as the associated libraries with exact references. The limitation / constraints of the AI model should be documented and how they are being optimised alongside their supporting rationale. Ethical, transparency and explainability trade-offs that may apply together with their rationale should also be recorded.
Code and data	Recording the code used to build any AI model which goes to production/exploitation. Exclusively for high impact applications, insurance undertakings should record the training data used to build the AI model and all the associated hyper parameters, including pseudo-random seeds.[17] If this requirement proved to be too burdensome, insurance undertakings may put in place alternative measures that ensure the auditability of the AI model and the accountability of the undertaking using them.

[16] Real world data used to train and test the AI system.

[17] Pseudorandom number generator is a deterministic computational process that has one or more inputs called "seeds", and it outputs a sequence of values that appears to be random according to specified statistical tests.

Model performance	Explanations should include, inter alia, how performance is measured (KPIs) and what level of performance is deemed satisfactory, including scenario analysis and timing and frequency of reviews and / or retraining of the model. Ethical, transparency and explainability trade-offs that may apply together with their rationale should also be recorded.
Model security	Description of mechanisms in place (or marketing reference to) to ensure the model is protected from outside attacks and more subtle attempts to manipulate data or algorithms themselves: how robust is the model to manipulation attacks (especially important in auto ML models)
Ethics and trustworthy assessment	Description of the impact assessment of the AI system used i.e. the potential impact on customers and/or insurance undertakings of the concrete AI systems used. Explain how the governance measures put in place throughout the AI systems lifecycle address the risks included in the impact assessment and ensure ethical and trustworthy AI systems.

Source: EIOPA Consultative Expert Group on Digital Ethics in insurance

3. EXAMPLES OF FAIRNESS METRICS

The table below does not represent an exhaustive list; the fairness metrics may vary from one use of AI system to another (e.g. underwriting, fraud detection, chatbots etc.), and this is a field of ongoing scientific research.

Some of the group fairness metrics mentioned in the table below could contradict the concept / metric of actuarial fairness in insurance underwriting, where customers bearing the same risk are charged the same price.

The fairness metrics should therefore be treated with caution and undertakings should ultimately adopt the fairness metrics that best suit their business model and uses of AI systems, taking into account risk based and proportionality considerations.

Fairness metric	**Description**
Demographic Parity	The goal of "Demographic Parity" is to assign the positive outcome at proportionally equal rates to each subgroup of a protected class where the

	positive outcome refers to the favourable decision.[18] For example, in the context of a recruitment scenario "Demographic Parity" could mean that male and female candidates are invited to job interviews at equal rates, proportionately to the number of applications.
Calibration	Another approach aims at equal positive and negative predictive values for all subgroups.[19] Such calibration guarantees that the predictive values across subgroups correspond to the scores which represent the probability of predicting the positive or the negative outcome. For example, in a medical diagnosis scenario, a calibrated model could ensure equal levels of confidence in the predictions for patients of different gender or ethical backgrounds because the predictive values are comparable across all subgroups.
Equalized Odds	This fairness definition requires equal true positive and true negative rates for all subgroups.[20] For example, where an insurance undertaking uses AI systems to scan through CVs and job applications in recruitment processes, "Equalized Odds" would ensure that the chances for men and women to be invited to the job interview are equal.[21]
Equalized Opportunities	This relaxed version of "Equalized Odds" is often used in practice because it reduces the computational complexity when working with large real-world datasets. "Equalized Opportunities" only requires the error rates for the favourable outcome to be the same but allows deviations for the unfavourable outcome. For example, in online marketing when the objective is to inform men and women at equal rates about an insurance offer, "Equalized Opportunities" could ensure that relevant segments of both groups are shown the information at equal rates. The rate of exposure to people for whom the offer is actually irrelevant may differ, however.
Individual fairness	All definitions mentioned above bind on a group level, based on one or several protected attributes. A completely different approach is "Individual Fairness" which abandons the idea of group memberships and suggests instead that any similar individuals should be treated similarly. For example, all the individuals

[18] Dwork, C., Hardt, M., Pitassi, T., Reingold, O., Zemel, R. (2011),

[19] Crowson, C., Atkinson. E., Therneau, T., Lawson, A., Lee, D. and MacNab, Y. (2016),

[20] Hardt, M., Price, E. and Srebro, N. (2016).

[21] This fairness metric is already used by some companies such as Linkedin: https://engineering.linkedin.com/blog/2021/using-the-linkedin-fairness-toolkit-large-scale-ai.

	with the same risk profile should pay the same premium for the same insurance product.

Source: EIOPA Consultative Expert Group on Digital Ethics in insurance