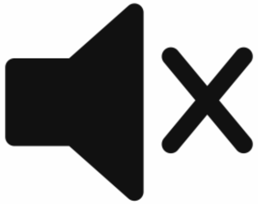
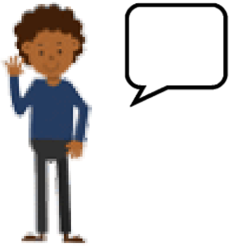
access to insurance initiative

Welcome!



For the quality of the call, you will be muted.
You can unmute yourself to ask a question.



Introduce yourself in the chat!

French translation via the **Interactio app** or at https://app.interactio.io
Event code: **A2iiIAIS**

access to insurance initiative

Artificial Intelligence (AI) and emerging regulatory expectarions - Supervisory Dialogue

7 April 2022

Dialogues

Housekeeping Rules

- This Dialogue will be RECORDED
- Please MUTE yourself whilst you are not speaking
- Please use the CHAT function to ask questions
- "RAISE HAND" when wishing to speak or ask a question
- For any technical issues, contact A2ii Secretariat via the chat function or via e-mail at (dialogues@a2ii.org)

French translation via the **Interactio app** or at https://app.interactio.io
Event code: **A2iiIAIS**

Audio Translation / Traduction Audio / Traducción de audio

1

Download the “Interactio” Application

Or stream online at https://app.interactio.io

Télécharger l’Application “Interactio”

Ou diffusé en ligne https://app.interactio.io

Descargue la aplicación “Interactio”

O stream en línea https://app.interactio.io

https://app.interactio.io

2

Enter the event code:
A2iiIAIS

Entrez le code de l'événement:
A2iiIAIS

Introduzca el código del evento:
A2iiIAIS

Manoj Pandey

Moderator

Advisor - Access to Insurance Initiative (A2ii)

Manuela Zweimueller

Head of Implementation, International Association of Insurance Supervisors (IAIS)

Jeffery Yong

Senior Advisor – FSI, BIS

Julian Arevalo

Senior Expert – Financial Innovation (EIOPA)

Financial Stability Institute | BIS

Humans keeping AI in check – emerging regulatory expectations in the financial sector

A2ii-IAIS dialogue

7 April 2022

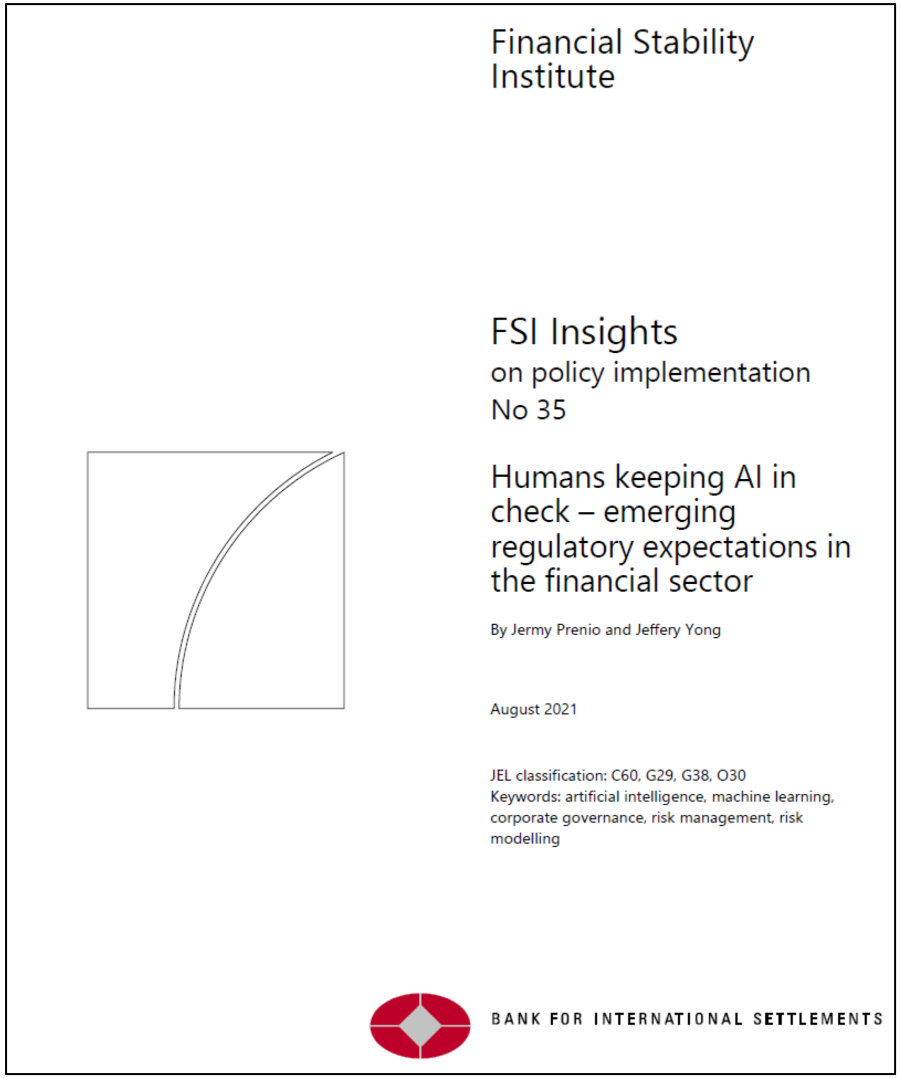
Agenda

- Introduction to FSI's policy implementation work
- Common themes in AI regulatory issuances
- Existing standards or laws
- Implementation challenges

Introduction to FSI policy implementation work

- **Objective**: to contribute to international discussions on a range of contemporary regulatory and supervisory policy issues and implementation challenges faced by financial sector authorities
- **Coverage**: analyses of different jurisdictional approaches on regulatory/supervisory topics
- **Format**: FSI Insights, FSI Briefs, Crisis Management Series etc.

Visit our webpage.


Financial Stability Institute

FSI Insights
on policy implementation
No 35

Humans keeping AI in check – emerging regulatory expectations in the financial sector

By Jermy Prenio and Jeffery Yong

August 2021

JEL classification: C60, G29, G38, O30
Keywords: artificial intelligence, machine learning, corporate governance, risk management, risk modelling

BANK FOR INTERNATIONAL SETTLEMENTS

Scope of paper

- Covers policy documents on AI governance issued by financial authorities or groupings in 9 jurisdictions
- Aims of paper
  - to provide a snapshot of existing regulatory approaches on AI governance
  - to identify emerging common regulatory themes including from relevant cross-industry, general AI guidance

Overview of AI-related issuances

	Regulation/ legislation	Guidance; guidelines	Principles	Discussion paper; others
European Union	✓(EC[1])	✓(HLEG[2])	✓(EIOPA[3])	✓(EBA[4], EIOPA[5])
France				✓(ACPR[6])
Germany			✓(BaFin[7])	✓(BaFin[8])
Hong Kong, SAR			✓(HKMA[9])	
Luxembourg				✓(CSSF[10])
Netherlands			✓(DNB[11])	
Singapore			✓(MAS[12])	
United Kingdom		✓(ICO[13])		✓(BoE/FCA[14])
United States			✓(NAIC[15])	✓(UST[16], US Agencies[17])
International			✓(OECD[18], G20[19])	

- 1 European Commission, Proposal for a regulation laying down harmonised rules on AI (April 2021).
- 2 Independent High-level Expert Group on AI (set up by the European Commission), Ethics guidelines for trustworthy AI (April 2019).
- 3 European Insurance and Occupational Pensions Authority, Artificial intelligence governance principles: towards ethical and trustworthy artificial intelligence in the European insurance sector (June 2021).
- 4 European Banking Authority, Report on big data and advanced analytics (January 2020).
- 5 European Insurance and Occupational Pensions Authority, Big data analytics in motor and health insurance: A thematic review (May 2019).
- 6 French Prudential Supervision and Resolution Authority (ACPR), Governance of AI in Finance (June 2020).
- 7 Federal Financial Supervisory Authority of Germany (BaFin), Big data and artificial intelligence: Principles for the use of algorithms in decision-making processes (June 2021).
- 8 Federal Financial Supervisory Authority of Germany (BaFin), Big data meets AI (July 2018).
- 9 Hong Kong Monetary Authority, High-level principles on AI (November 2019); Consumer protection in respect of Use of Big Data Analytics and Artificial Intelligence by Authorized Institutions (November 2019).
- 10 Financial Sector Supervisory Commission of Luxembourg (CSSF), AI: Opportunities, risks and recommendations for the financial sector (December 2018).
- 11 Netherlands Bank, General principles for the use of AI in the financial sector (July 2019).
- 12 Monetary Authority of Singapore, Principles to promote fairness, ethics, accountability and transparency (FEAT) in the use of AI and data analytics in Singapore's financial sector (November 2018).
- 13 UK's Information Commissioner's Office, draft Guidance on the AI auditing framework (February 2020) and Guidance on AI and data protection (July 2020).
- 14 Bank of England and Financial Conduct Authority, Machine Learning in UK financial services (October 2019).
- 15 National Association of Insurance Commissioners (2020), Principles on Artificial Intelligence.
- 16 US Treasury, A financial system that creates economic opportunities: nonbank financials, fintech, and innovation (July 2018).
- 17 US regulatory agencies, Request for information and comment on financial institutions' use of AI, including machine learning (March 2021).
- 18 Organisation for Economic Cooperation and Development, AI Principles (May 2019).
- 19 G20, AI Principles (June 2019).

Summary of regulatory expectations on common AI principles

Reliability / soundness	• Similar expectations as those for traditional models (eg model validation, defining metrics of accuracy, updating/retraining of models, ascertaining quality of data inputs) • For AI models, assessing reliability/soundness of model outcomes is viewed from the perspective of avoiding causing harm (eg discrimination) to consumers
Accountability	• Similar expectations as outlined in general accountability or governance requirements, but human involvement is viewed more as a necessity • For AI models, accountability includes "external accountability" to ascertain that data subjects (ie prospective or existing customers) are aware of AI-driven decisions and have channels for recourse
Transparency	• Similar expectations as those for traditional models, particularly as they relate to explainability and auditability • For AI models, external disclosure (eg data used to make AI-driven decisions and how the data affects the decision) to data subjects is also expected
Fairness	• Stronger emphasis in AI models (although covered in existing regulatory standards, fairness expectations are not typically applied explicitly to traditional models) • Expectations on fairness relate to addressing or preventing biases in AI models that could lead to discriminatory outcomes, but otherwise "fairness" is not typically defined
Ethics	• Stronger emphasis in AI models (although covered in existing regulatory standards, ethics expectations are not typically applied explicitly to traditional models) • Ethics expectations are broader than "fairness" and relate to ascertaining that customers will not be exploited or harmed, either through bias, discrimination or other causes (eg AI using illegally obtained information)

Applicability of international standards

Common principles	Applicable standards/laws
Reliability/ soundness	• Basel Core Principles (BCP) 15, Insurance Core Principles (ICP) 16, ICP 17, Basel Committee on Banking Supervision (BCBS) Principles for effective risk data aggregation and risk reporting • Minimum requirements for the use of IRB for credit risk, IMA for market risk, stress testing, technical provisions valuation
Accountability	• BCP 14, BCP 15, ICP 7, ICP 17, BCBS Corporate governance principles for banks • Minimum requirements for the use of IRB for credit risk, IMA for market risk, AMA for operational risk, stress testing, technical provisions valuation
Transparency	• ICP 17 • Minimum requirements for the use of IRB for credit risk, IMA for market risk, stress testing, technical provisions valuation
Fairness	• ICP 19, ComFrame standard 7.2a • Consumer protection laws in some countries explicitly address fairness concerns as described in AI-related issuances (ie prevent/address discriminatory outcomes)
Ethics	• BCP 29, ICP 5, ICP 7, ICP 8, BCBS Corporate governance principles for banks, BCBS Principles for the sound management of operational risk, BCBS Principles on compliance and the compliance function in banks. FSB toolkit for firms and supervisors to mitigate misconduct risk

BCP 14 Corporate governance
BCP 15 Risk management process
BCP 29 Abuse of financial services
ICP 5 Suitability of persons
ICP 7 Corporate governance
ICP 8 Risk management and internal controls
ICP 16 Enterprise risk management for solvency purposes
ICP 17 Capital adequacy
ICP 19 Conduct of business
ComFrame standard 7.2.a: The group supervisor requires the IAIG Board to ensure that the group-wide business objectives, and strategies for achieving those objectives, take into account at least the following fair treatment of customers.

Challenges in implementing the common AI themes/principles

Transparency

- If not transparent, cannot assess reliability / establish accountability
- Technical skills – both within firms and authorities to explain model
- Trade-off between 'too much' (can be mis-used by clients) and 'too little'

Reliability and soundness

- Technical issues – data quality, removing bias
- Efforts for regular and timely update – eg changes in behaviors due to Covid
- Existing regulatory requirements not fit-for-AI – what constitutes a 'change' (supervised ML learns with new data)
- Trade-off between simplicity and performance
- Cyber risk – data poisoning to alter training data set

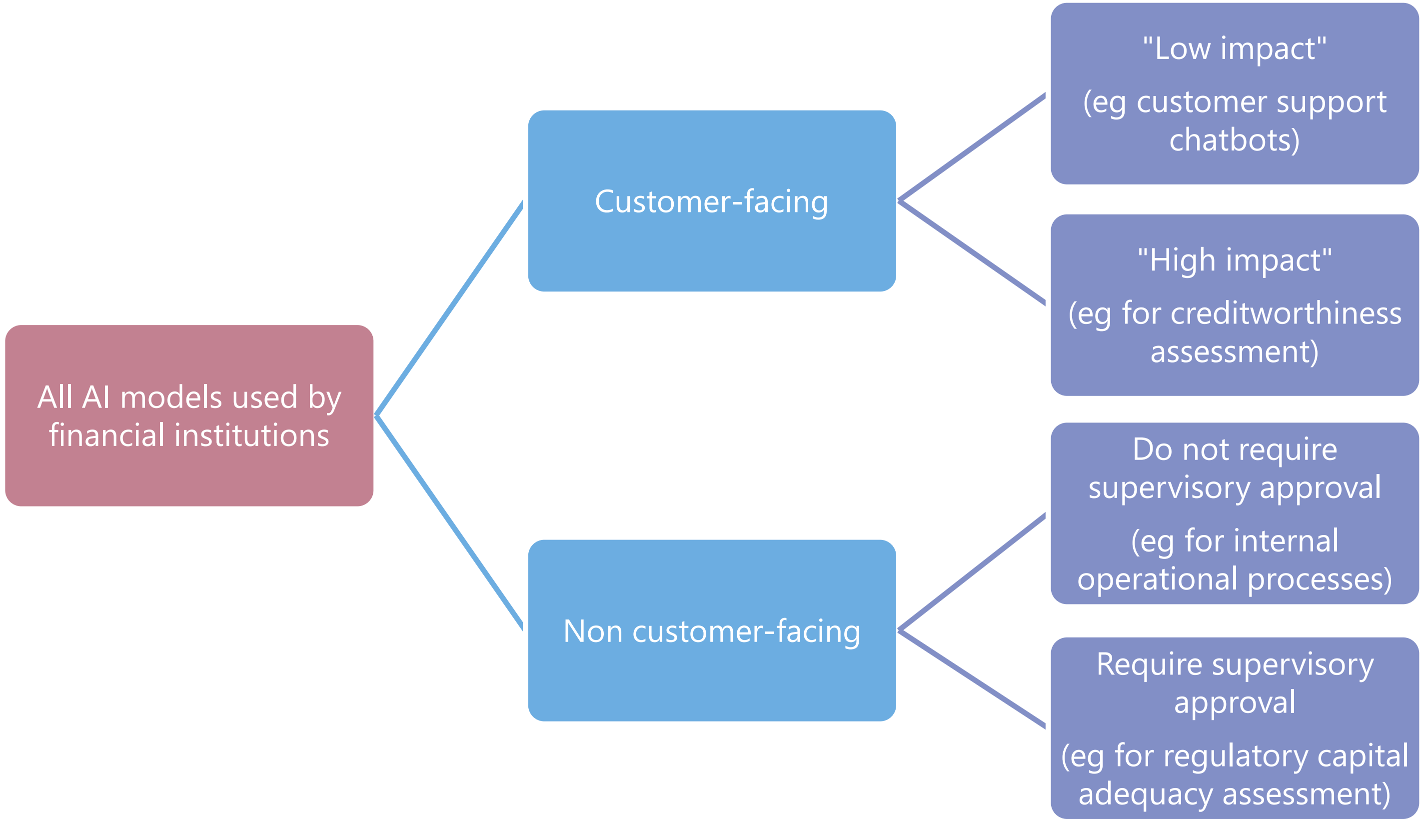
Accountability

- Unclear who is responsible at lower levels of hierarchy – eg data scientist or head of credit underwriting?
- New human risks – liable for errors if manually override model, thus increase hesitancy; easier to accept model results than to explain; human-introduced bias
- Outsourcing risks – commercial capture, accountability

Fairness and ethics

- Lack of universally accepted definitions
- Regulations that require human judgment – difficult to implement in ML as it lacks contextual understanding eg future insurance needs of a client
- Financial exclusion – eg under-represented groups not receiving good credit scores as there is no past data
- Human intervention may introduce human flaws/bias – too much human efforts negate automation benefits

Tailoring regulatory and supervisory frameworks to AI use cases



Summary of key points

- Existing requirements on governance, risk management, as well as development and operation of traditional models also apply to AI models.
- While most of the issues arising from the use of AI by financial institutions are similar to those for traditional models, the perspective might be different - scope to do more on fairness.
- The stronger emphasis on fairness in the use of AI results in calls for more human intervention to avoid unintended bias/discriminatory outcomes – humans are accountable
- The more AI model's use can potentially impact authorities' conduct and prudential objectives, the more stringent the relevant reliability/soundness, accountability, transparency, fairness and ethics requirements should be.
- Given emerging common themes on AI governance in the financial sector, there seems to be scope for financial standard-setting bodies to develop international guidance or standards in this area.

Audio Translation / Traduction Audio / Traducción de audio

1

Download the "Interactio" Application

Or stream online at https://app.interactio.io

Télécharger l'Application "Interactio"

Ou diffusé en ligne https://app.interactio.io

Descargue la aplicación "Interactio"

O stream en línea https://app.interactio.io

Google play

Available on the App Store

https://app.interactio.io

2

Enter the event code:
A2iiIAIS

Entrez le code de l'événement:
A2iiIAIS

Introduzca el código del evento:
A2iiIAIS

AI GOVERNANCE PRINCIPLES

Towards ethical and trustworthy AI in the European insurance sector

Presenter: **Julian Arevalo Carreño**
Date: **7 April 2022**

European Insurance and
Occupational Pensions Authority

USE OF AI IN THE EUROPEAN INSURANCE SECTOR

Figure 7 – Usage of BDA tools such as AI and ML

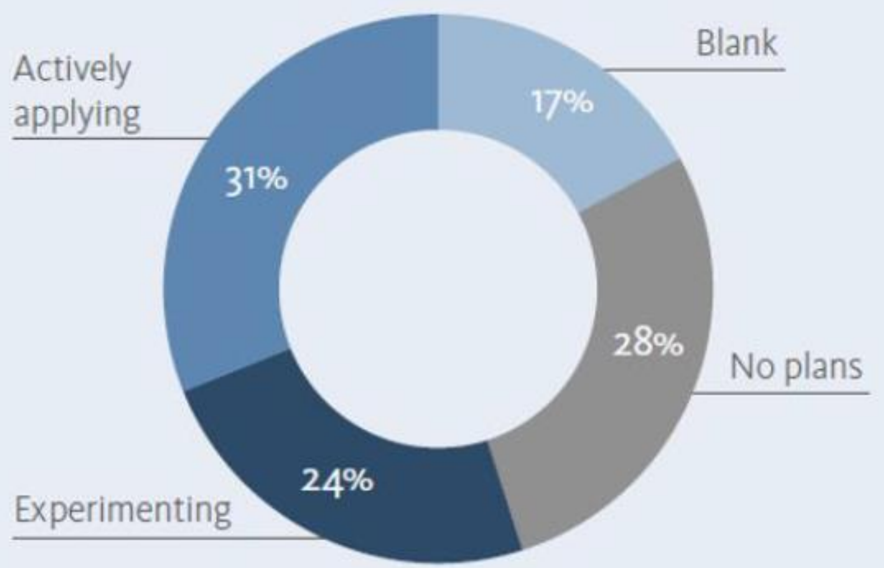
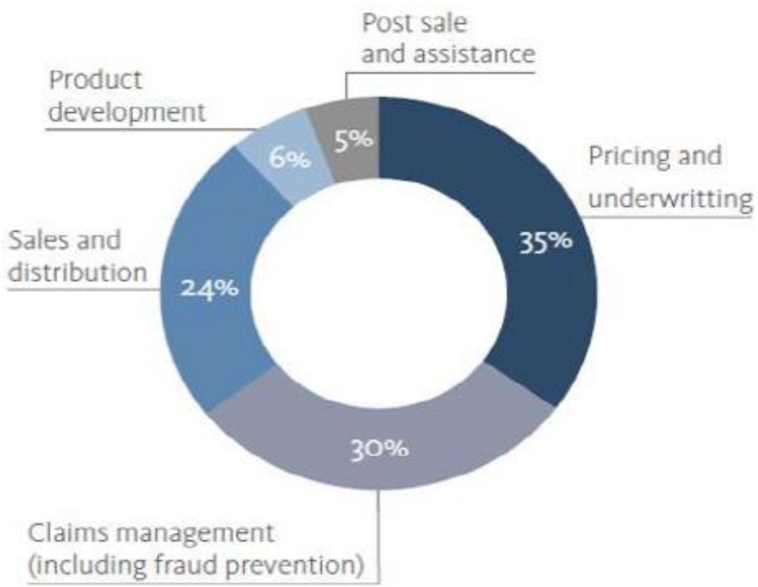


Figure 8 – Usage of BDA tools such as ML and AI across the value chain



Source: EIOPA's thematic review on the use of Big Data Analytics in the motor and health European insurance sector: https://register.eiopa.europa.eu/Publications/EIOPA_BigDataAnalytics_ThematicReview_April2019.pdf

AI USE CASES ACROSS THE INSURANCE VALUE CHAIN

Product design and development	Pricing and underwriting	Sales and distribution	Customer service	Loss Prevention	Claims management
•Historical customer and survey data analysis to inform new products •Predictive modelling of disease development patterns •Novel products, e.g. parametric and usage-based insurance	• Enhanced risk assessments combining traditional and new data sources (including IoT data) • Price optimisation: micro-segment / personalised pricing based on non-risk individual behavioural data (e.g. to estimate price elasticity, lifetime value and propensity to churn) and market competition analysis	• Digital marketing techniques based on the dynamic analysis of online search behaviour • Virtual Assistant and Chatbots that utilise Natural Language Processing (NLP) and insurance ontologies to support communication • Proactive customer communication, nudging and cross-selling of related services ("next-best action") based on consumer data from Customer Relationship Management (CRM) systems	•Call centre sentiment analysis, route cause analysis, dynamic scripting and agent allocation •Customer self-service through multiple channels using NLP, voice recognition, insurance ontology maps and chatbots •Robotic Process Automation (RPA) including Optical Character Recognition (OCR) to extract information from documents (e.g. FNOL, email with questions complaints etc.) and route them to the correct department	•Provide diagnostic advice anc coaching based on AI analytics from health and automotive big data, e.g. suggest exercise and driving behaviour changes	• Enhanced fraud analytics: claims scoring, anomaly detection, social network analytics and behavioral modelling • Loss reserving: use of AI to estimate the value losses, in particular for high-frequency claims • AI image recognition to estimate repair costs in household property insurance, business premises and automotive • Automated segmentation of claims by type and complexity and automated invoice verification and payment process

Source: AI governance principles report developed by EIOPA's Consultative Expert Group on Digital Ethics in insurance: https://www.eiopa.europa.eu/sites/default/files/publications/reports/eiopa-ai-governance-principles-june-2021.pdf

EIOPA'S CONSULTATIVE STAKEHOLDER GROUP ON DIGITAL ETHICS IN INSURANCE

Composition

- Created in October 2019
- 40 stakeholders from the insurance industry, consumers, academics and consultants
- Multidisciplinary background: actuaries, data scientists, lawyers, economists etc.

Objective

- Provide guidance and enhance trust in the use of new business model, data sources and technologies in insurance

Scope

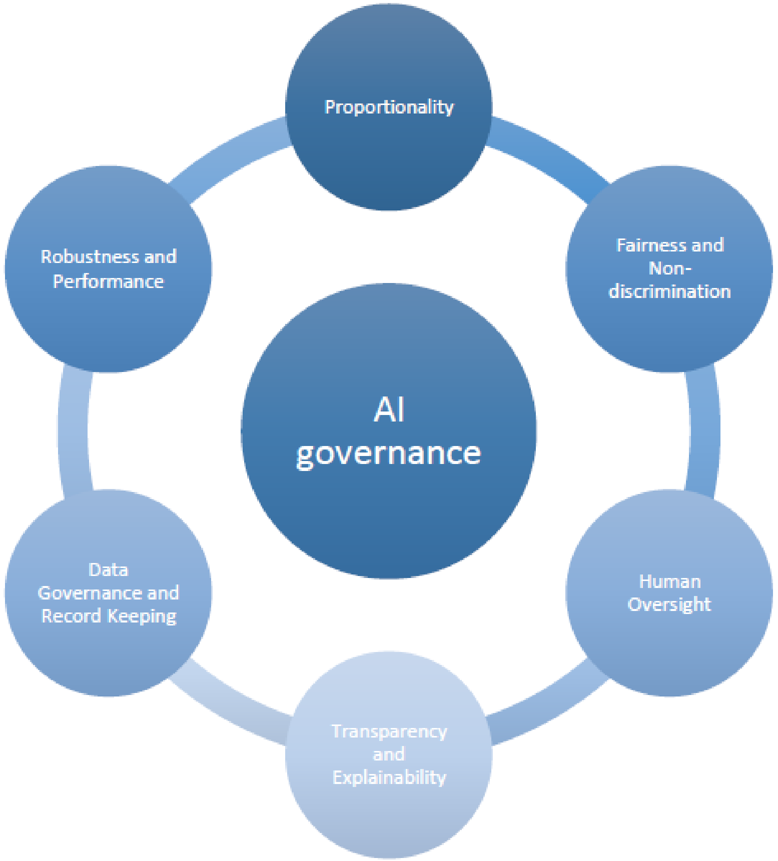
- Specific to the insurance sector
- Focus on pricing and underwriting, but also other areas of the value chain
- Retail consumers prioritised

Approach

- Principles-based approach, but include concrete examples and guidance to stakeholders

AI GOVERNANCE PRINCIPLES

- Based on Ethical and Trustworthy AI guidelines developed by the **European Commission's High Level Expert Group on AI**
- Intended to be **accommodated into existing frameworks**
- An ethical and trustworthy governance framework is achieved by a **combination of measures** and not by a single / stand-alone one

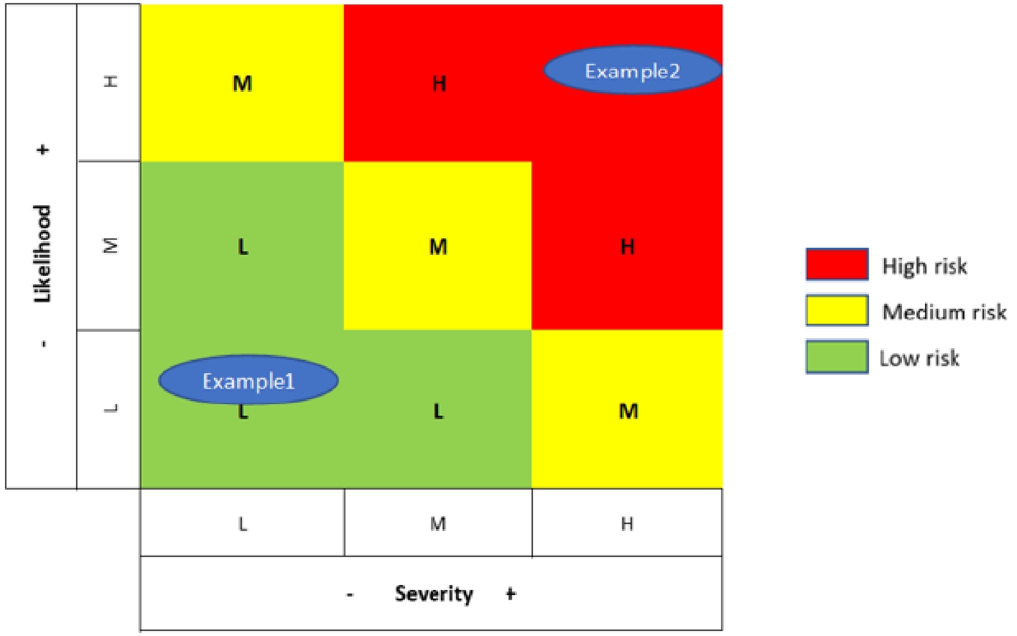


PROPORTIONALITY: AI USE CASE IMPACT ASSESSMENT

Figure 5 – AI use case 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



FAIRNESS AND NON-DISCRIMINATION

- *Take into account the **outcomes of AI systems***
- ***Balance the interests of all the stakeholders** involved (insurers, consumers, society)*
- *Insurer's **corporate social responsibility**: take into account financial inclusion issues and consider ways to avoid reinforcing existing inequalities (e.g. credit scores), especially for products that are socially beneficial.*
- *Respect the **principle of human autonomy** by developing AI systems that support consumers in their decision-making process (e.g. avoid using certain types of price optimisation practices)*
- *Dataset used should **be fit for purpose***
- ***Make reasonable efforts to monitor and mitigate biases from data and AI systems**.*
- *Insurance firms should **develop their approach to fairness and keep records***

FUNDAMENTAL RIGHTS AND INSURANCE LEGISLATION

Figure 9 – Protected classes in EU Charter of Fundamental Rights and exemptions in national legislation for insurance risk assessments

Protected characteristic in Article 21 EU Charter of Fundamental Rights[43]	Allowed for insurance risk-based pricing and underwriting, with restrictions (depends on Member State's national law)
• Sex • Race • Colour • Ethnic or social origin • Genetic features • Language • Religion or believe • Political or any other opinion • Membership of a minority group • Property • Birth • Disability • Age • Sexual orientation • Nationality	• age • disability • religion or belief[44] • sexual orientation[45]

Source: EIOPA Consultative Expert Group on Digital Ethics in insurance

- Some protected characteristics are **allowed to be use for insurance underwriting**
- Court of Justice of the European Union (Test Achat case) barred the use of **gender** as a risk factor
- **Age** is a very relevant risk factor in insurance underwriting, but can it be used for non-risk price optimisation practices?

DIRECT AND INDIRECT DISCRIMINATION

- Directive 2004/113/EC and Directive 2000/43/EC → regulate the equal treatment irrespective of gender, racial or ethnic origin, distinguishes between direct and indirect discrimination: proxis have to be "**objectively justified by a legitimate aim**" **and "appropriate and necessary"**
- European Commission guidance (2012/C 11/01): "**true risk factors on its own right**"
- AI Governance principles report:
  - **Correlation is not causation**: actuarial / risk-based pricing in insurance should be based on rating factors with a risk correlation and a causal link in compliance with anti-discrimination legislation
  - As part of their corporate social responsibility, insurance firms should assess and develop measures to mitigate the impact of rating factors such as **credit scores, location, income, occupation or level of education** on vulnerable populations and protected classes in those essential lines of business where they have a limited causal link

Figure 11 – Guidance on the necessary and appropriateness assessment of rating factors

Necessary and appropriateness assessment of rating factors and rating categories	
Necessary:Risk / claim correlation	▪ Each rating factor used for risk differentiation should have a clear correlation with claims occurrence (i.e. risk).
Appropriateness: Causal Link	▪ Each rating factor and subsequent rating categories (e.g. for the rating factor "job", the rating categories could be "blue collar" or "white collar", or more granular rating categories like "teachers, engineers, doctors, nurses etc.") should have a causal link[55] between the rating factor or rating category and claims occurrence / risk. ▪ Each rating factor and rating category should have a valid explanation or rationale for different treatment of otherwise similarly situated consumers. ▪ Each rating factor and rating category should be in line with generally accepted actuarial principles[56 57] ▪ AI systems used to predict risks based on a single or limited number of unconventional rating factors also raise significant concerns from a fairness and non-discrimination perspective.

Source: EIOPA Consultative Expert Group on Digital Ethics in insurance[58]

AVAILABLE TOOLS TO ADDRESS FAIRNESS AND NON-DISCRIMINATION

Figure 10 – Examples of fairness and non-discrimination metrics

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.[46] 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.[47] 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.[48] For example, where an insurance firms 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.[49]
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 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

- **Traditional tools (process focus):**
  - Remove bias from the training data (including proxis)
  - Use protected characteristics as "control variables" in the model to isolate each individuals predictive variable's unique contribution to explaining the outcome
- **New tools (outcome focus):**
  - Fairness metrics to measure model outcomes
  - Benchmark model outcomes (e.g. average premium in a Zip Code) with aggregated data (e.g. on diversity) at Zip code level available in the Census

TRANSPARENCY AND EXPLAINABILITY

- Explanations need to be adapted to:
    - **Concrete AI use cases**
    - **Different stakeholders**

- Explanations should be:
    - **meaningful**
    - **easy to understand** in order to help stakeholders make informed decisions

- Explainability is necessary to:
    - Ensure **accountability** of firms
    - Enable **redress mechanisms**
    - Address **bias**

Figure 14 – Transparency and explainability information to be provided to different stakeholders when using AI in pricing and underwriting (The criteria with an asterisk are further developed in Chapter IX)

		Types of stakeholders		
AI use case	Information to be provided	Consumer	Auditor and supervisor	Board
Pricing and Underwriting	Is automated decision making or AI used?	x	x	
	What datasets are used	x	x	
	Why certain criteria are chosen for underwriting and pricing i.e. causal link	x	x	
	Counterfactual explanation - most influencial rating factors	x	x	
	Reasons for using AI and consistency with corporate strategies / objectives*		x	x
	Description of how the model is integrated in the current IT system*		x	
	Staff involved in the design and implementation and core function groups*		x	x
	Data collection, preparation and post-processing methodologies*		x	
	Technical choices / arbitration and limitations / risks of the AI model chosen*		x	
	Code and data used to train and test the model*		x	
	Model performance, including KPIs*		x	x
	Model security measures*		x	
	Ethics and trustowrthy assessment*		x	x
	Documentation on compliance with regulation		x	x
	Certification by an independent body, disclosure of audit		x	x
	System logic explained to a non-expert		x	x
	Implemented third-party technologies and risks		x	x

Source: EIOPA consultative stakeholder group on digital ethics in insurance73

HUMAN OVERSIGHT

- Insurance firms should establish adequate levels of human oversight **throughout the AI system's life cycle**
- Human oversight needs to be **adapted** to concrete AI use cases
- Insurance firms should **assign and document clear roles and responsibilities** for the staff involved in AI processes

Figure 15 – Example of involvement of different staff members during the development phase of different AI application depending on their materiality

	Management / executive Board	Head of IT department	Developers of AI systems	Data protection officer (DPO)	AI / data officer	Compliance function	Risk management function	Audit function	Actuarial function
High Impact	A	A	A	C	A	A	C	I	C
Medium Impact	I	A	A	C	A	C	C	I	C
Low Impact	I	I	A	C	C	C	I	I	I

Source: EIOPA Consultative Expert Group on Digital Ethics in insurance

DATA GOVERNANCE AND RECORD KEEPING

- **GDPR and Solvency II's** data Governance (accurate, complete and appropriate requirements)
- **Reproducibility:** Could you see where things went wrong, why, and find a solution in a reasonable timeframe?
- **Each case to be assessed independently**: what is good for one thing may not be for another: record rationale for decisions, data, models, code, minutes, logs…
- **Ethical considerations** of sourcing data: Can you use it? Should you use it?
- **Addressing bias** and shortcomings in data: do you have an agreed approach?
- **Third party data** should be subject to similar requirements

Figure 17 – Record keeping requirements for high impact AI applications

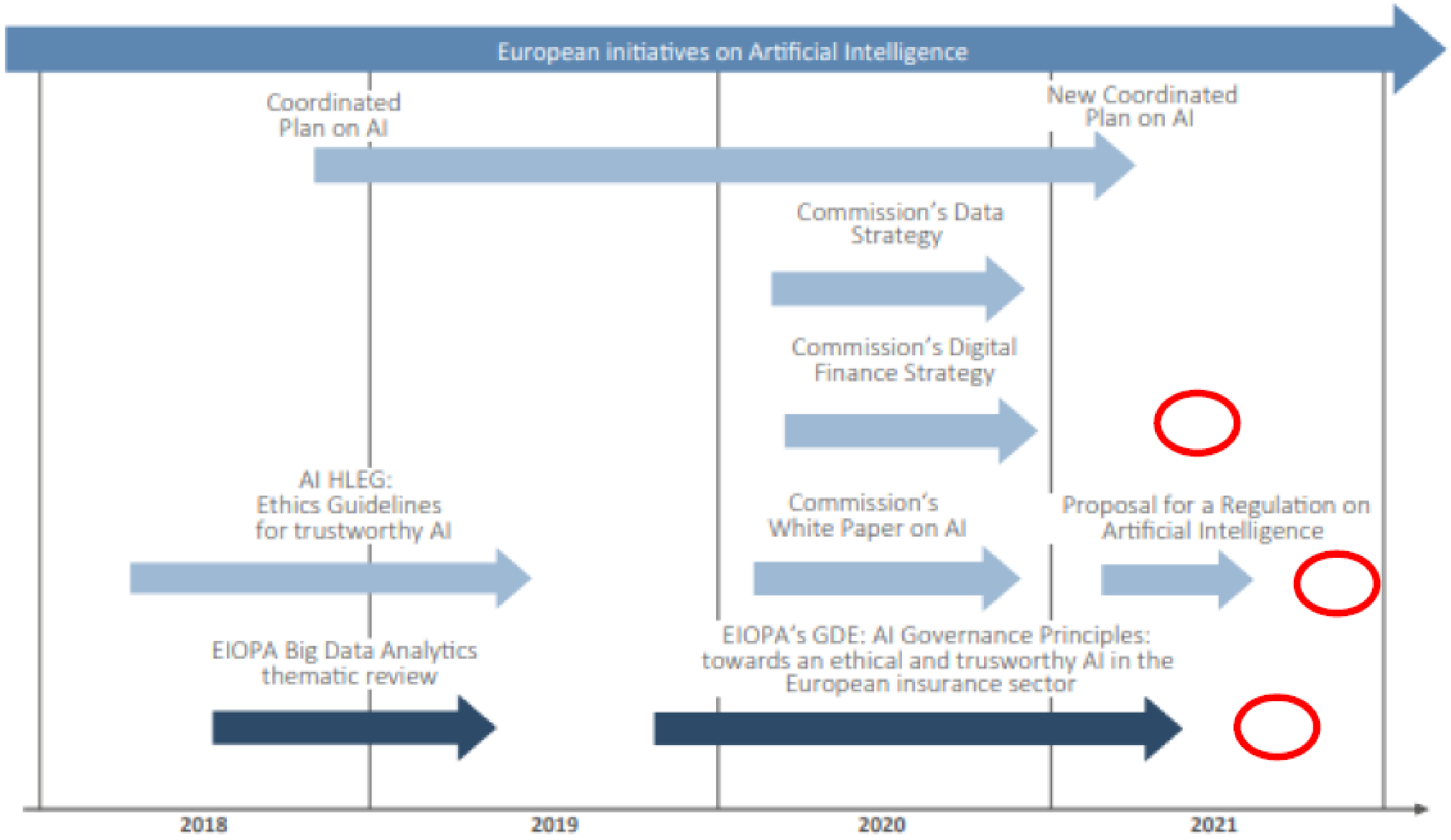
Record	Description
Reasons for using AI	Explanation of the business objective/ task pursued by using AI and its consistency with corporate strategies / objectives. Explanation how this was implemented into the AI system. This would help avoid misusage of the AI system and enable its audit and independent 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	Identify 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 allow to ensure accountability of the responsible persons.
Data collection	Document how the ground truth 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	Document 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	Record the code used to build any AI model which goes to production/exploitation. Exclusively for high impact applications, insurance firms should record the training data used to build the AI model and all the associated hyper parameters, including pseudo-random seeds. If this requirement proved to be too burdensome, insurance firms may put in place alternative measures that ensure the auditability of the AI model and the accountability of the firm using them.
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	Describe mechanisms in place (or make 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 AI use case impact assessment i.e. the potential impact on consumers and/ or insurance firms of the concrete AI use case. Explain how the governance measures put in place throughout the AI systems lifecycle address the risks included in the AI use case impact assessment and ensure ethical and trustworthy AI systems.

Source: EIOPA Consultative Expert Group on Digital Ethics in insurance

ROBUSTNESS AND PERFORMANCE

- The **calibration, validation and reproducibility** of AI systems is done on a sound manner
- Ensure that the AI systems outcomes are **stable over time and/or of a steady nature**
- **Ongoing monitoring** to ensure robustness and detect failing performance
- Performance **metrics** based on intended outcomes, including ethical ones
- **Secure and resilient**, including against **cyber attacks**
- Similar requirements for **outsourced solutions**
- Considering **fall-back plans** where appropriate

NEXT STEPS



- **European Commission's Digital Finance Strategy**
  - AI guidelines for the financial sector (postponed)
- **AI Act**
  - Debate whether insurance AI use cases should be considered as high-risk AI applications
- **EIOPA**
  - Continue looking into specific AI issues (e.g. explainability) for specific AI use cases
  - Financial inclusion: price optimisation practices and data bias

THANK YOU!

For more information visit:
https://www.eiopa.europa.eu

European Insurance and
Occupational Pensions Authority

access to insurance initiative

Thank you.

Follow us on Twitter @a2ii_org, YouTube and LinkedIn

Implementation Partner:

IAIS
INTERNATIONAL ASSOCIATION OF INSURANCE SUPERVISORS

Supported by:

Federal Ministry for Economic Cooperation and Development

Ministry of Foreign Affairs of the Netherlands

Hosted by:

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH