Welcome!

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

Introduce yourself in the chat!

Spanish translation via the Interactio app or at https://app.interactio.io
Event code: A2iiAIS
Artificial Intelligence (AI) and emerging regulatory expectations - Supervisory Dialogue

7 April 2022
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)

Spanish translation via the **Interactio app** or at [https://app.interactio.io](https://app.interactio.io)

Event code: **A2iiIAIS**
Audio Translation / Traduction Audio / Traducción de audio

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
</tr>
</thead>
</table>
| **Download the “Interactio” Application**
Or stream online at [https://app.interactio.io](https://app.interactio.io)  
Télécharger l’Application “Interactio”
Ou diffusé en ligne [https://app.interactio.io](https://app.interactio.io)  
Descargue la aplicación “Interactio”
O stream en línea [https://app.interactio.io](https://app.interactio.io) | **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)

Jeffery Yong  
Senior Advisor - F (FSI, BIS)

Awelani Rahulani  
Head of Department: Fintech at the Financial Sector Conduct Authority (FSCA) in South Africa

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

Julian Arevalo  
Senior Expert – Financial Innovation (EIOPA)

Spanish translation via the Interactio app or at https://app.interactio.io  
Event code: A2iiIAIS
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](#).
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

<table>
<thead>
<tr>
<th>European Union</th>
<th>Regulation/legislation</th>
<th>Guidance; guidelines</th>
<th>Principles</th>
<th>Discussion paper; others</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ (EC¹)</td>
<td>✓ (HLEG²)</td>
<td>✓ (EIOPA³)</td>
<td>✓ (EBA⁴, EIOPA⁵)</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>✓ (ACPR⁶)</td>
<td>✓ (BaFin⁷)</td>
<td>✓ (BaFin⁸)</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>✓ (BaFin⁷)</td>
<td>✓ (HKMA⁹)</td>
<td>✓ (CSSF¹⁰)</td>
<td></td>
</tr>
<tr>
<td>Hong Kong, SAR</td>
<td>✓ (HKMA⁹)</td>
<td>✓ (CSSF¹⁰)</td>
<td>✓ (BoE/FCA¹⁴)</td>
<td></td>
</tr>
</tbody>
</table>

- 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).
- 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).
- 19 G20, AI Principles (June 2019).
# Summary of regulatory expectations on common AI principles

## Reliability / soundness
- Similar expectations as those for traditional models (e.g., 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 (e.g., 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 (i.e., 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 (e.g., 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 (e.g., AI using illegally obtained information)
## Applicability of international standards

<table>
<thead>
<tr>
<th>Common principles</th>
<th>Applicable standards/laws</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliability/</td>
<td>• Basel Core Principles (BCP) 15, Insurance Core Principles (ICP) 16, ICP 17, Basel Committee on Banking Supervision (BCBS) Principles for effective risk</td>
</tr>
<tr>
<td>soundness</td>
<td>data aggregation and risk reporting</td>
</tr>
<tr>
<td></td>
<td>• Minimum requirements for the use of IRB for credit risk, IMA for market risk, stress testing, technical provisions valuation</td>
</tr>
<tr>
<td>Accountability</td>
<td>• BCP 14, BCP 15, ICP 7, ICP 17, BCBS Corporate governance principles for banks</td>
</tr>
<tr>
<td></td>
<td>• Minimum requirements for the use of IRB for credit risk, IMA for market risk,AMA for operational risk, stress testing, technical provisions valuation</td>
</tr>
<tr>
<td>Transparency</td>
<td>• ICP 17</td>
</tr>
<tr>
<td></td>
<td>• Minimum requirements for the use of IRB for credit risk, IMA for market risk, stress testing, technical provisions valuation</td>
</tr>
<tr>
<td>Fairness</td>
<td>• ICP 19, ComFrame standard 7.2a</td>
</tr>
<tr>
<td></td>
<td>• Consumer protection laws in some countries explicitly address fairness concerns as described in AI-related issuances (ie prevent/address discriminatory outcomes)</td>
</tr>
<tr>
<td>Ethics</td>
<td>• 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</td>
</tr>
</tbody>
</table>

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

All AI models used by financial institutions

Customer-facing

- “Low impact”
  (eg customer support chatbots)

- “High impact”
  (eg for creditworthiness assessment)

Non customer-facing

- Do not require supervisory approval
  (eg for internal operational processes)

- Require supervisory approval
  (eg for regulatory capital adequacy assessment)
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

2

Enter the event code: A2iiIAIS

Entrez le code de l'événement: A2iiIAIS

Introduzca el código del evento: A2iiIAIS

https://app.interactio.io
AI GOVERNANCE PRINCIPLES

Towards ethical and trustworthy AI in the European insurance sector

Presenter: Julian Arevalo Carreño
Date: 7 April 2022
USE OF AI IN THE EUROPEAN INSURANCE SECTOR

## AI USE CASES ACROSS THE INSURANCE VALUE CHAIN

<table>
<thead>
<tr>
<th>Product design and development</th>
<th>Pricing and underwriting</th>
<th>Sales and distribution</th>
<th>Customer service</th>
<th>Loss Prevention</th>
<th>Claims management</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Historical customer and survey data analysis to inform new products</td>
<td>- Enhanced risk assessments combining traditional and new data sources (including IoT data)</td>
<td>- Digital marketing techniques based on the dynamic analysis of online search behaviour</td>
<td>- Call centre sentiment analysis, route cause analysts, dynamic scripting and agent allocation</td>
<td>- Provide diagnostic advice and coaching based on AI analytics from health and automotive big data, e.g. suggest exercise and driving behaviour changes</td>
<td>- Enhanced fraud analytics: claims scoring, anomaly detection, social network analytics and behavioral modelling</td>
</tr>
<tr>
<td>- Predictive modelling of disease development patterns</td>
<td>- 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</td>
<td>- Virtual Assistant and Chatbots that utilise Natural Language Processing (NLP) and Insurance ontologies to support communication</td>
<td>- Customer self-service through multiple channels using NLP, voice recognition, insurance ontology maps and chatbots</td>
<td>- Loss reserving: use of AI to estimate the value losses, in particular for high-frequency claims</td>
<td>- AI image recognition to estimate repair costs in household property insurance, business premises and automotive</td>
</tr>
<tr>
<td>- Novel products, e.g. parametric and usage-based insurance</td>
<td></td>
<td>- Proactive customer communication, nudging and cross-selling of related services (“front-best action”) based on consumer data from Customer Relationship Management (CRM) systems</td>
<td>- 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</td>
<td></td>
<td>- Automated segmentation of claims by type and complexity and automated invoice verification and payment process</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Severity</th>
<th>Impact on consumers</th>
<th>Impact on insurance firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of consumers affected</td>
<td>Business continuity</td>
<td></td>
</tr>
<tr>
<td>Consumer interaction and interests</td>
<td>Financial impact</td>
<td></td>
</tr>
<tr>
<td>Types of consumers (e.g. vulnerable consumers)</td>
<td>Legal impact</td>
<td></td>
</tr>
<tr>
<td>Human autonomy</td>
<td>Reputational impact</td>
<td></td>
</tr>
<tr>
<td>Anti-discrimination and diversity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance line of business relevance</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Likelihood</th>
<th>Evaluation or scoring, including profiling and predicting</th>
<th>Automated decision making with legal or similar significant effect</th>
<th>Systematic monitoring</th>
<th>Model complexity/combining datasets</th>
<th>Innovative use or applying new technological or organisational solution</th>
<th>Type and amount of data used</th>
<th>Outsourcing datasets and AI applications</th>
</tr>
</thead>
</table>

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

<table>
<thead>
<tr>
<th>Protected characteristic in Article 21 EU Charter of Fundamental Rights⁴³</th>
<th>Allowed for insurance risk-based pricing and underwriting, with restrictions (depends on Member State’s national law)</th>
</tr>
</thead>
</table>
| • 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 belief⁴⁴  
• sexual orientation⁴⁵ |

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 appropriate-ness assessment of rating factors and rating categories**

<table>
<thead>
<tr>
<th>Necessary and appropriateness assessment of rating factors and rating categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Necessary: Risk / claim correlation</td>
</tr>
<tr>
<td>- Each rating factor used for risk differentiation should have a clear correlation with claim occurrence (e.g. risk)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Appropriateness: Causal Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 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 with the rating factor or rating category and claim occurrence / risk</td>
</tr>
<tr>
<td>- Each rating factor and rating category should have a valid explanation or rationale for different treatment of other risk-similarly situated consumers</td>
</tr>
<tr>
<td>- Each rating factor and rating category should be in line with generally accepted actuarial principles</td>
</tr>
<tr>
<td>- AI systems used to predict risks based on a single or limited number of environmental rating factors also raise significant concerns from a fairness and non-discrimination perspective</td>
</tr>
</tbody>
</table>

Source: EIOPA Classification Report Group on Digital White on Insurance®
AVAILABLE TOOLS TO ADDRESS FAIRNESS AND NON-DISCRIMINATION

- **Traditional tools (process focus):**
  - Remove bias from the training data (including proxis)
  - Use protected characteristics as “control variables” in the model to isolate each individual's 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

---

**Figure 1 - Examples of fairness and non-discrimination metrics:**

<table>
<thead>
<tr>
<th>Fairness Metric</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic Parity</td>
<td>The goal of &quot;Demographic Parity&quot; 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. For example, in the context of a recruitment scenario, &quot;Demographic Parity&quot; could mean that male and female candidates are invited to job interviews at equal rates, proportionately to the number of applications.</td>
</tr>
<tr>
<td>Calibration</td>
<td>Another approach aims at equal positive and negative predictive values for all subgroups. Such calibration guarantees that the predicted 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 ethnic backgrounds because the predictive values are comparable across all subgroups.</td>
</tr>
<tr>
<td>Equalized Odds</td>
<td>This fairness definition requires equal true positive and true negative rates for all subgroups. For example, where an insurance firm uses AI systems to scan through CVs and job applications in recruitment processes, &quot;Equalized Odds&quot; would ensure that the chances for men and women to be invited to the job interview are equal.</td>
</tr>
<tr>
<td>Equalized Opportunities</td>
<td>This relaxed version of &quot;Equalized Odds&quot; is often used in practice because it reduces the computational complexity when working with large real-world datasets. &quot;Equalized Opportunities&quot; 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, where the objective is to inform men and women at equal rates about an insurance offer, &quot;Equalized Opportunities&quot; 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 relevant may differ, however.</td>
</tr>
<tr>
<td>Individual Fairness</td>
<td>Any definition mentioned above builds on a group level, based on one or several protected attributes. A completely different approach is &quot;Individual Fairness&quot; which abandons the idea of group memberships and suggests instead that unfamiliar 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.</td>
</tr>
</tbody>
</table>

Source: EIOPA, Consultative Expert Group on Digital Ethics in Insurance
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)

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

Source: EIOPA consultative stakeholder group on digital ethics in insurance
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 applications depending on their materiality**

<table>
<thead>
<tr>
<th></th>
<th>Management/executive Board</th>
<th>Head of IT department</th>
<th>Developers of AI systems</th>
<th>Data protection officer (DPO)</th>
<th>AI/data officer</th>
<th>Compliance function</th>
<th>Risk management function</th>
<th>Audit function</th>
<th>Actuarial function</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Impact</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>I</td>
<td>C</td>
</tr>
<tr>
<td>Medium Impact</td>
<td>I</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>I</td>
<td>C</td>
</tr>
<tr>
<td>Low Impact</td>
<td>I</td>
<td>I</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>I</td>
<td>I</td>
</tr>
</tbody>
</table>

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
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
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

2

Enter the event code: A2iiIAIS

Entrez le code de l'événement: A2iiIAIS

Introduzca el código del evento: A2iiIAIS
A2ii-IAIS supervisory dialogue on Artificial Intelligence and Regulatory Expectations

7 April 2022
Agenda

FSCA Overview

BrandsEye description and case study
About the Financial Sector Conduct Authority

Who we are

- The FSCA was established to be the dedicated market conduct regulator in South Africa's Twin Peaks regulatory model implemented via the Financial Sector Regulation Act.

- Our mandate includes all financial institutions that provide a financial product and/or a financial service as defined in the Financial Sector Regulation Act.

- The FSCA's mandate is expressed through the following strategic objectives:
  - Ensure the stability of financial markets;
  - Drive transformation of the financial sector to improve access;
  - Promote fair treatment of financial customers through a robust regulatory framework;
  - Provide financial education and literacy in order to have informed customers; and
  - Assist in maintaining the efficiency and integrity of financial markets through innovation.

Our Vision

- To foster a fair, efficient, and resilient financial system that supports inclusive and sustainable economic growth in South Africa.
What is BrandsEye (DataEQ) tool?

**Definition**

- BrandsEye is a reputation management and social media monitoring system that searches the social web for mentions and uses crowdsourced analysis to judge sentiment. Using a unique combination of Artificial intelligence BrandEye is able to priorities the conversation that requires attention and action.

- The tool gained global recognition in 2016 for analyzing public tweets to predict both Brexit and the US elections providing meaningful, predictive insights gained from analyzing social media at scale.

**BrandsEye (DataEQ) Features and Capabilities**

- Media monitoring capability
- Competitor benchmarking
- Unlimited Dashboards
- Opinion Based insights
- Unlimited users
- Detailed Metrics Reporting
BrandEye (DataEQ) within the FSCA

Overview

• The Financial Sector Conduct Authority (FSCA) commissioned BrandsEye to create a report on the Authority’s social media posts over the March - August 2021 period.
• The report aims to measure the performance of these posts and their impact on the FSCA’s public perception.
• The report also aims to provide into mechanisms to improve engagement and sentiment towards future posts.

Work done

• The FSCA has tracked public social media conversation of about 214 FSPs, between May 2020 and April 2021.
• Banking, long-term Insurance and insurance contribute the largest volumes of conversation about FSPs.
• The FSCA monitored consumer mentions containing themes related to the Treating Customers Fairly framework
• Measuring the completeness and transparency of financial advertising.

Key Outcomes

Proactive monitoring:
- Establishing industry benchmarks
- Deep dive analysis into market conduct trends
- Identifying new and unregulated entities

Active communication:
- Market conduct thought-leadership
- Interventions in social media discourse
- Keeping the public informed about ongoing matters
- Driving public awareness on issues like unclaimed benefits

Active listening:
- Elevation of consumer voice resulting in FSCA protecting the public against scams
Questions
Thank you.

Follow us on Twitter @a2ii_org, YouTube and LinkedIn