AUTHORED BY
ASTRID LANGEVELD-VOS DR.
CHIARA ZILIANI
FRANK HELDENS
LUCIA HROVATIN
PETER GALJAARD DR.
ROBERT ZERGENYI DR.
SALVINO MARIGO

POSITION PAPER

Auditing a Digital Insurance World

Artificial Intelligence and Machine Learning audits within Insurance Firms

JUNE 2023

European Confederation of Institutes of Internal Auditing

Index

Executive summary

The market context in which insurance companies operate is fundamentally changing. The use of data and Artificial Intelligence (AI) algorithms is growing significantly and is expected to be a key currency of future success. With the huge quantities of data created across the insurance value chain, AI provides tremendous opportunities for further automation of processes, development of new, more customer-centric products and the assessment of insurance risks. With these new possibilities, processes are becoming more complex and risks need to be handled. AI algorithms may have a direct impact on people and therefore ethical and privacy questions arise, which in turn brings regulators and industry bodies to the discussion to avoid adverse effects, without stifling the innovation and potential of AI.

Insurance companies must achieve the right balance between improving their operations with the new solutions which AI will make possible and managing the corresponding risks. This requires rigorous risk assessment and management of the development, implementation and use of AI. The importance is reflected by various legislation currently under development across the world, including the European Union's AI Act, which includes penalties of up to 6% of total worldwide annual turnover. With these regulatory requirements and the potential reputational implications, AI risk management cannot be completely diversified or assessed proportionally. No matter the size of the insurance company, it can be catastrophic for reputation and business if customers are harmed by AI. That's why Internal Audit should play a role in providing assurance and advice on mitigating risks arising from implementing AI.

The Internal Audit function can, according to its mandate, help organizations with the balancing act between risk mitigation and business innovation. This could include developing strategies for assurance to govern AI, data privacy and security, reviewing processes for potential bias and ensuring compliance with relevant laws and regulations. In addition, internal auditors can provide insights and advice for companies in understanding and mitigating the risks associated with AI adoption and use.

Internal Audit should be involved from the start of new AI implementations to provide advice on how to implement AI securely, according to policies and regulation. Following a top down approach is wise, starting with auditing the AI strategy, governance and test individual instances, algorithms and models, starting with high risk AI. This will ensure that the development is being conducted in an efficient and effective manner and that controls are in place tailored to the risks related to the specific AI implementation.

Internal Audit should not only provide assurance over the process of developing AI, but also perform risk-based deep dives to ensure AI implementation is compliant and working effectively. Auditing AI includes technical aspects, data governance and quality, ethical themes and business application. Therefore, a multidisciplinary audit team should be formed. The team should have representatives from IT audit, data science, business audit and specific technical expertise such as actuaries, as well as ethics, to ensure each aspect is thoroughly assessed. Hence, Internal Audit departments should upskill their staff where needed, to stay ahead of key new developments, and be able to independently assess the risks, plan and execute audits as required. Our research has shown that most Internal Audit departments are at an early state of establishing the required skills and processes, and often not keeping up with the rapid development in use of AI in the Insurance industry. For these reasons, this paper contains a proposal of an AI Audit Program, where the most important AI related risks, possible root causes and testing strategies are identified.

1. Introduction

In this position paper, we elaborate on the progress and relevance of AI within the European insurance industry, the upcoming legislation and risk response. This is supported by a survey which provides the perspective from the Three Lines of Defence (First Line: Business, Second Line: Risk Management and Third Line: Internal Audit) and their current state of readiness to manage the risks related to AI. We then provide suggestions to Internal Audit for a solid audit response on AI, to help the insurance industry prepare for 'trustworthy AI' and future legislation.

1.1. Consensus on a specific definition of AI is work in progress

In general, AI is the branch of computer science that focuses on creating machines or software for tasks that normally require human interpretation or more. It combines the principles of computer science, mathematics, linguistics, psychology and neuroscience.

Consensus on a more specific definition is still work in progress. This is illustrated by the definition that is included in the European AI Act: *"software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human defined objectives, generate outputs such as content, predictions, recommendations or decisions influencing the environments they interact with*". In the appendix of the proposed Act, it distinguishes three types of AI techniques and approaches: machine learning (including supervised, unsupervised, reinforcement and deep learning), logic and knowledge-based including expert systems, and statistical approaches (European Commission, 2021).

Many of the amendments to the AI Act, also mentioned in section 2.4, are related to narrowing the scope. Others argue that the AI Act should use a broader definition or concept of "automated and algorithmic decision-making" to truly show the socioeconomic impact of AI systems on individuals. Nevertheless, the consensus seems to be that the dynamic nature of the AI algorithms landscape needs to be reflected in the definition.

For this paper, we use the OECD definition of an AI-system as *"a machine-based system that is capable of influencing the environment by producing an output (predictions, recommendations or decisions) for a given set of objectives. It uses machine and/or human-based data and inputs to (i) perceive real and/or virtual environments; (ii) abstract these perceptions into models through analysis in an automated manner (e.g., with machine learning), or manually; and (iii) use model inference to formulate options for outcomes. AI systems are designed to operate with varying levels of autonomy."*(OECD, AI Principles, 2019).

Regardless of the exact definition, AI systems still remain computer programs. The software itself might be very smartly coded and use advanced technologies, it's actual intelligence can be questioned. A human mind grows to understand the meanings. Computer software, also when called AI, calculates probabilities.

1.2. Technology drives AI systems

Insurance was always based on statistical methods and one of the industries which relied most on actuaries and mathematicians. Developments took place gradually; some actuarial models relying on mathematical and statistical methods developed many years ago. Further, computing power gradually increased with a doubling every two years from 1960 to 2010, following Moore's law on computational power.

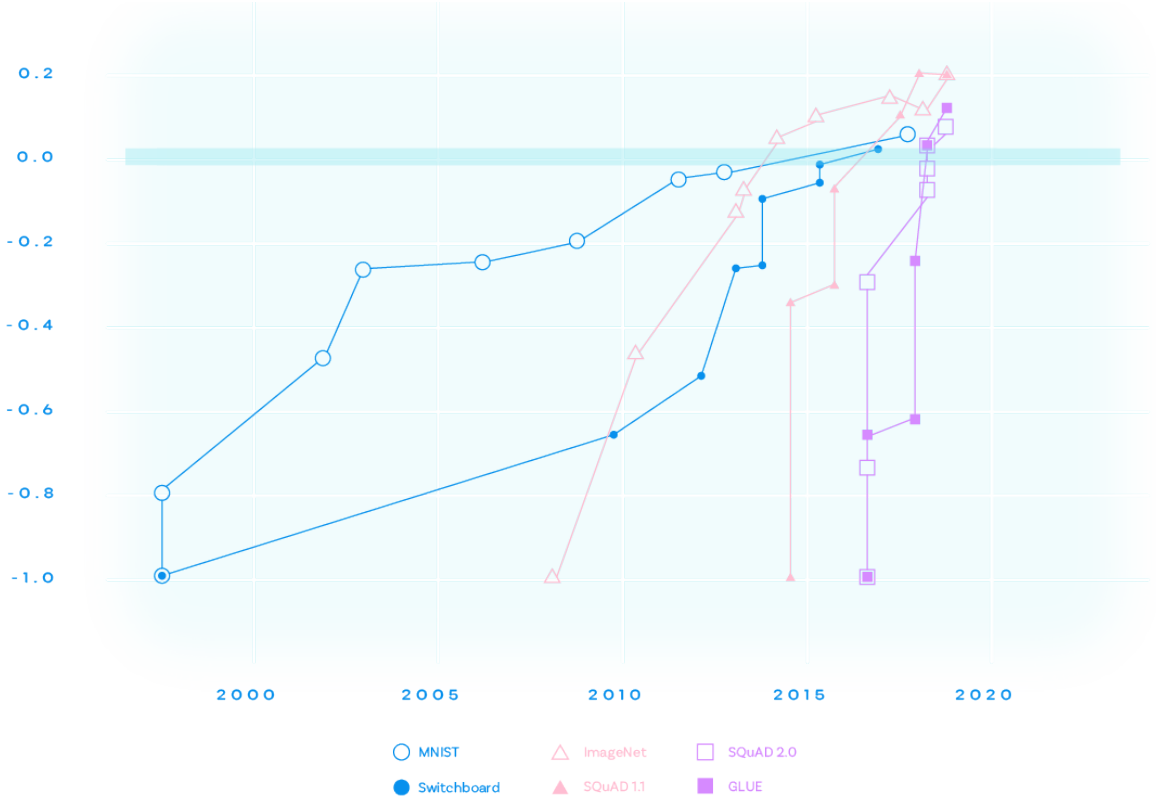
Since 2010, cloud computing has become more pervasive, enabling flexible scaling of computing power and storage. This significant increase in processing capabilities and the

development of new AI algorithms to process more and more data, is resulting in an exponential growth in accuracy. Several performance benchmarks are available to measure accuracy of AI Algorithms. Comparison of these performance benchmarks shows that AI algorithms already surpassed human performance benchmarks in handwriting, speech and image recognition, reading comprehension and language understanding between 2015 and 2019. The graph illustrates benchmark saturation over time for popular benchmarks, normalized with initial performance at minus one and human performance at zero (Douwe Kiela, 2021).

Today, the use of Natural Language Processing (NLP) and generative AI models such as Google's BERT, OpenAI's GPT-4, Facebook's RoBERTa, and Microsoft's MT-DNN has enabled even more advanced capabilities in generating new content (text, speech, picture or video).

These powerful large language models (LLMs) have immense upside potential, but also bring new risks to bear, such as copyright infringement, insensitive content creation or over-confidence in the reliability of output created.

Figure 1. Benchmark saturation over time for popular benchmarks (Douwe Kiela, 2021).



1.3. Social Impact of AI

It is difficult to predict and quantify the future social impact of AI. But considering the speed of development to be able to replace human tasks, simple and even advanced, it will likely have significant impact on the labour market, innovation, products, human behaviour, economic value chains and the way people interact with machines. It is however not yet clear what the impact will be precisely. Research on one side focusses on AI for a social good and a positive impact in areas such as transportation, healthcare, communication, translation,

access to wealth and inclusion, and improvement to the quality of everyday life (Nenad Tomašev, 2020).

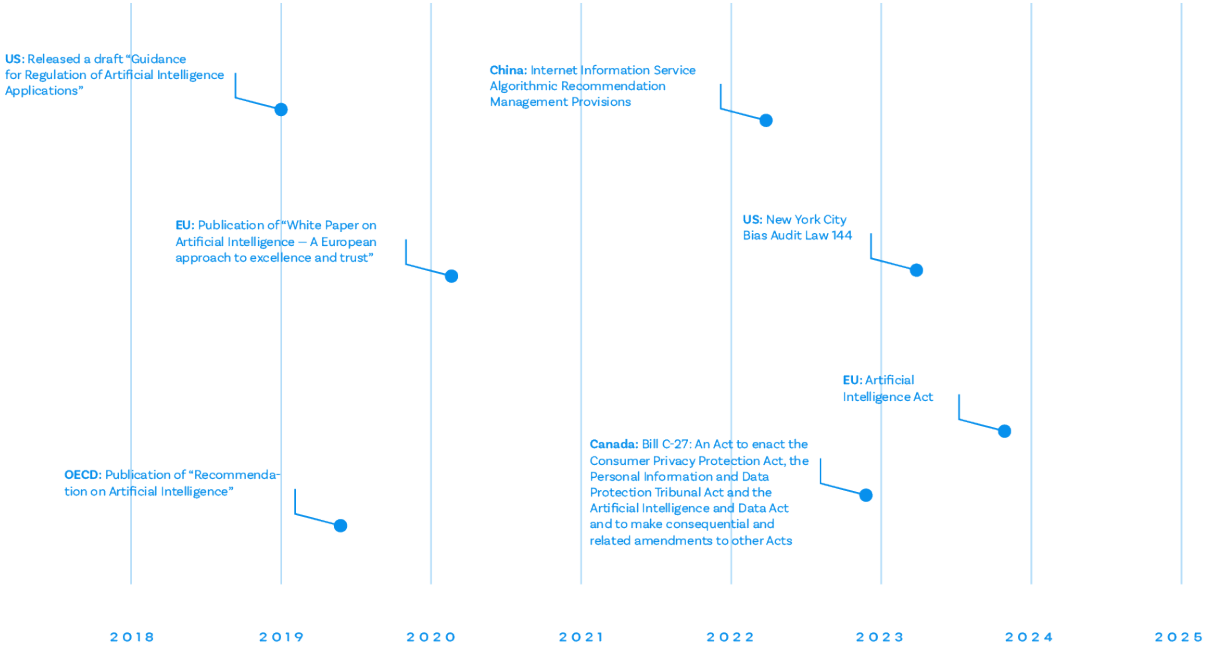
Other research shows potential risks and pitfalls of AI. AI can perpetuate and amplify existing prejudices and social inequalities, especially when using historical training data. Short term risks such as bias, privacy violations, unethical use, deep fakes, hallucination have made it on the radar of industry bodies, public and private organisations and policy-makers. Italy for example became the first European country to ban the use of ChatGPT in March 2023. At the same time, an open letter was signed by major industry leaders and experts to pause the training of AI systems more powerful than GPT-4 for at least 6 months. Current AI research also lacks a systematic discussion of how to manage long-tail risks from AI systems, including speculative long-term risks (Dan Hendrycks, 2022).

1.4. Legislation is rapidly developing

In view of all these developments, the question arises how AI can be regulated to prevent harm to people, while still unlocking the potential power of AI. It should be pointed out that the use of AI is not unregulated, as existing data protection and privacy laws already set some boundaries; however new more prescriptive AI laws are in the works. Many (voluntary) principles are already developed and countries, regulators and industry bodies around the world are rapidly moving forward with legislation in the field of AI.

Below are just a few examples of principles and regulations, of which the European AI act is likely to have the most impact for companies operating in the European market. The common denominator in all regulatory requirements is that they address concerns about transparency of machine decision-making and ethics. Considering the fragmented regulatory landscape it is key for companies to stay ahead of all applicable regulations.

Figure 2. AI Principles and Legislation

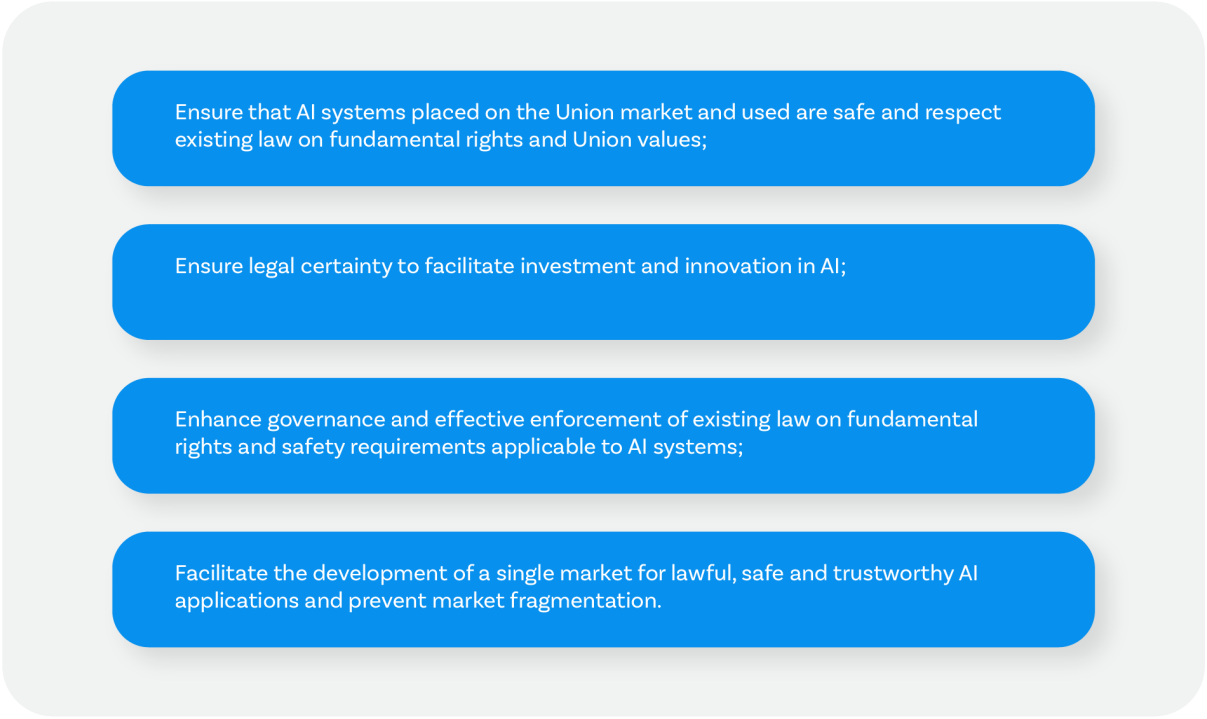


The OECD's "Recommendation on Artificial Intelligence" published in 2019 is seen as the backbone of the development of further regulation (OECD, Recommendation on Artificial

Intelligence, 2019). In May 2019, several countries, including the United States, adopted these recommendations as the first set of intergovernmental principles for trustworthy AI. The principles promote inclusive growth, human-centred values, transparency, safety and security and accountability. The Recommendation also encourages national policies and international cooperation to invest in research and development and support the broader digital ecosystem for AI.

In terms of regulation, China took the lead in March 2022 as the first country with an AI regulation (Stanford University, 2022). China's approach includes some global best practices around algorithmic system regulation, such as provisions that promote transparency and user privacy controls. The regulation is however under criticism because the Chinese approach also requires compliance with an ethical business code. This might be a way for the Chinese government to influence or control online information flows. Nevertheless, China has taken the lead in designing AI regulation.

The European Union is rapidly developing new regulations on digitalization like the Digital Markets Act, Digital Services Act and the Artificial Intelligence Act. The latter in particular is likely to impact all companies operating in Europe as the EU chose the so-called 'Regulation' as the legal instrument for this Act. Hence the AI Act will be immediately applicable law in all countries of the EU as soon as the European Parliament and the Council of Europe agree with the final text. The AI Act is built on the following objectives:



The AI Act recognizes three classes of risk:

- Unacceptable risk AI applications are prohibited. Examples of this risk type are practices using techniques beyond a person's consciousness, social scoring techniques likely to cause physical or psychological harm, activities exploiting vulnerabilities of specific groups, or the use of real-time remote biometric systems in publicly accessible spaces.

- High risk AI applications essentially consist of two lists of industries and activities that to date are recognized as high risk. The first list describes AI systems used as products or safety components of products covered by the sectorial Union law (for example, machinery, personal protective equipment, radio equipment, medical devices, transportation). The second list consists of "other" AI applications where risks have already materialized or are likely to materialize (for example, biometric identification of natural persons, supply of water, recruitment, access to public benefits and services, access to or assessment in educational and vocational training, creditworthiness, asylum and border control, administration of justice). For high risk applications, a conformity assessment must be performed on requirements around the risk management system, data and data governance, technical documentation, record keeping, transparency to users, human oversight and the system's accuracy, robustness and cybersecurity.
- Low or minimal risk applications are AI systems that are not prohibited or have a high risk. Before placement on the market, low risk systems have the possibility, but not the obligation, to follow a code of conduct on a voluntary basis.

At this moment the AI Act is still in the process of finalization. On May 11, 2023 the AI act was brought one step closer as the Internal Market Committee (IMCO) and the Civil Liberties Committee (LIBE) adopted a draft negotiating mandate (European Parliament, 2023). This position will be voted for in a plenary session in Week 24. More than 3,000 amendments are however still being tabled in the political arena of the EU, which must be considered before the AI Act is formally accepted. Formal acceptance is expected to be completed in early 2024, after which a two-year implementation period will start. The speed of the current developments in AI may however overtake the speed of the development of this regulation.

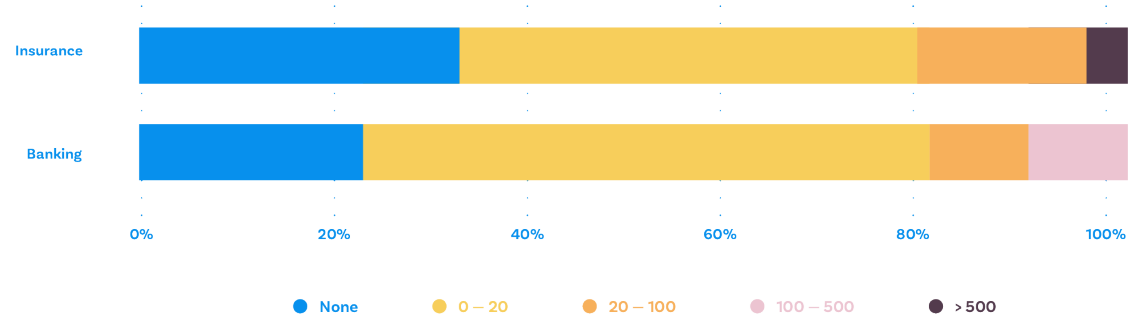
2. AI within the insurance industry

2.1. AI is becoming increasingly prominent in insurance firms

Technological innovation is changing organisations and business processes, and the insurance industry is included. To assess the ongoing changes within the insurance industry, an online survey was created. Respondents, data professionals working in insurance companies, affinity industries (i.e., banking sector) or others (e.g., public administration, pharmaceutical, advisory), covered over 80 entities. Out of the total, 28 belong to the Insurance industry and are the reference throughout the paper. Moreover, to facilitate comparisons and avoid scale issues, the companies were clustered into small, medium and large. Further information on the survey questions and definitions can be found in Annex 1.

As mentioned, technological changes affect companies in different ways and at varying pace, and deploying and exploiting AI systems within companies is following this course. This is displayed in the following graph, in which the distribution for the Insurance sector is compared with Banking. The comparison shows a common pattern, in which most entities have implemented up to 20 AI systems and only a few - medium and large – are aware of the deployment of integration of more than 20.

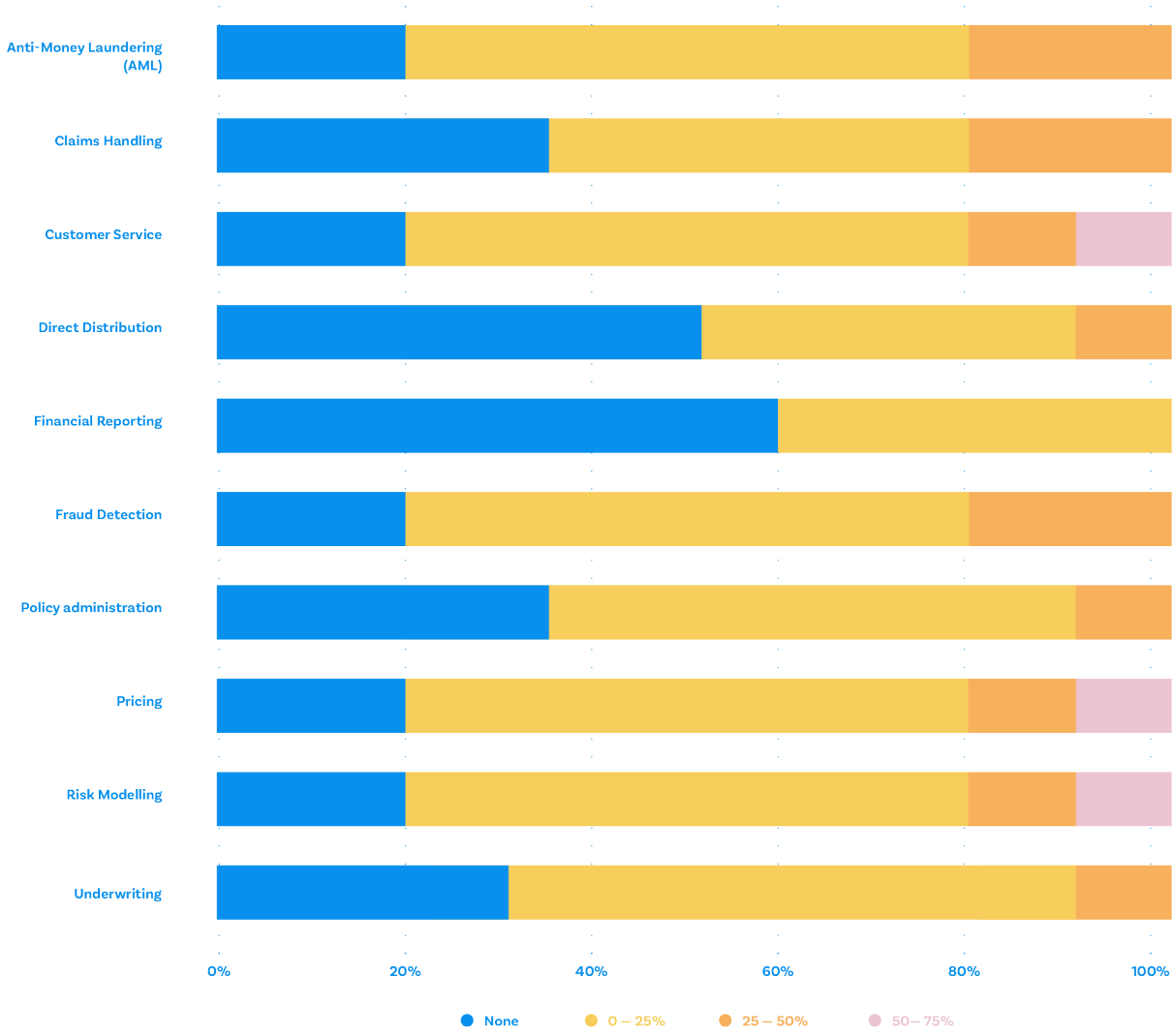
Figure 3. How many AI systems are you aware of in your company?



Excluding small companies without AI systems, the other companies use AI in various core business areas. The areas currently most backed by AI are customer-facing processes, such as direct customer service, where the aim is to improve customer satisfaction and reduce costs. Also, pricing and compliance-related areas show a more substantial use of AI systems. Whereas the former can be strengthened with AI due to the actuarial and mathematical nature, the latter might exploit AI in anomaly detection and similar tasks.

There is less implementation of AI systems in core processes such as financial reporting, policy administration, direct distribution and underwriting. In these areas, companies may be hesitant to rely on AI and consider that human intervention is required and needed. However, this pattern follows any implementation cycle of new technologies, and, in the coming years, core insurance processes will likely gradually migrate towards AI systems.

Figure 4. How much is each core area supported by AI?
— Large and medium companies



2.2. AI systems have different use cases, depending on how the algorithm is trained

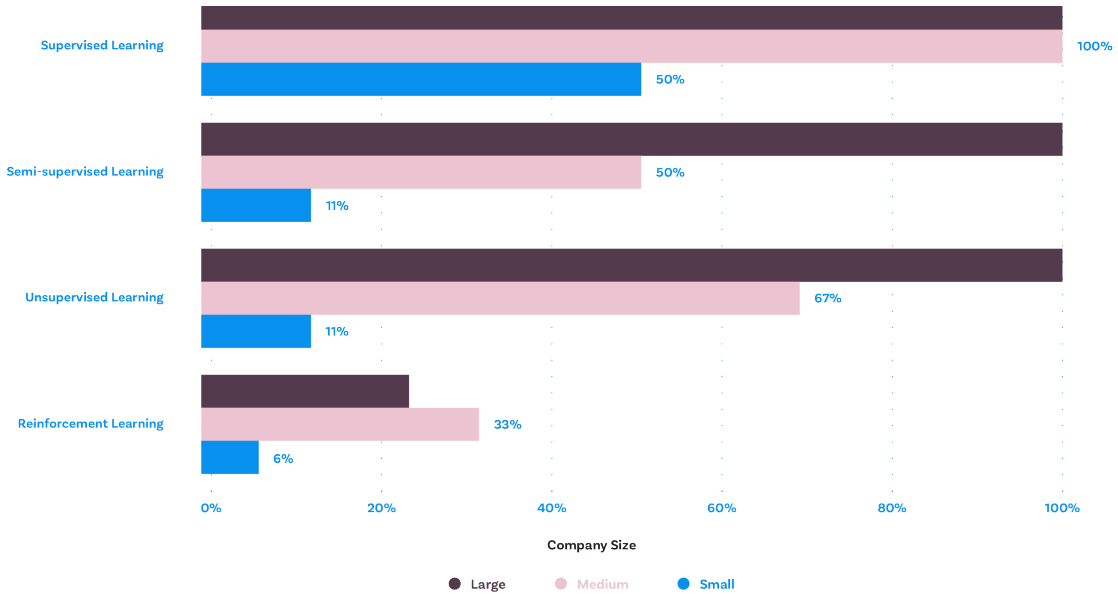
Before we further investigate some use cases of AI systems, it is worthwhile noting that they are typically categorized into four main groups: supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. As per the OECD definition of AI and many regulatory regimes, even simple mathematical algorithms that make predictions, recommendations or decisions need to be governed. The table below briefly explains the key characteristics and lists the major algorithms used in the Insurance industry as of today:

AI Category	Supervised learning	Unsupervised learning	Semi-supervised learning	Reinforcement learning
Definition	Supervised machine learning relies on labelled input and output training data. The algorithm is trained to make predictions on new (labelled) data. The quality of the predictions depends on training data quality, as well as the model's complexity.	Unsupervised machine learning algorithms analyse and cluster unlabelled datasets. These algorithms discover hidden patterns or group data without the need for human intervention.	Combination of supervised and unsupervised learning. The algorithms use a small amount of labelled data and a large amount of unlabelled data. The mixed data provides the benefits of both unsupervised and supervised learning, avoiding the challenge of finding a large amount of labelled data.	Reinforcement learning algorithms involve an agent that should learn to perform a specific task. The learning process follows a series of trial-and-error interactions aiming at the highest total reward.
Examples of algorithms used in the Insurance industry	**Linear regression/ logistic regression** • Risk assessments • Customer segregation • Customer surveys and analytics • Forecasting **Decision trees** • Risk Assessment • Root Cause analytics • Pricing **Support Vector Machines** • Identify fraudulent claims • Predict customer behaviour • Generate customer insights • Provide personalized recommendations **Machine learning for Natural Language Processing (NLP)** • Process automation, Optical Character Recognition (OCR) • Automatic Speech Recognition (ASR) • Chatbots • Customer self service • Authentication with voice recognition	**Clustering: K-means clustering, K-nearest neighbours (K-NN)** • Claims Fraud Detection • Performance monitoring • Underwriting • Customer segmentation **Association Rule Learning (Apriori, Eclat, etc.)** • Identify patterns in customer claims data **Principal Component Analysis (PCA)** • Predict customer churn **Self-Organizing Maps (SOM)** • Fraud detection **Autoencoders** • Claims fraud **Deep Belief Networks (DBN)** • Fraud detection	**Naïve Bayes value Estimation** • Fraud detection • Value estimation • Chatbots **Naïve Bayes Classifier** • Natural language processing tasks (social media, customer analytics) • Sentiment Analysis (call center, claims, complaints) **Neural networks** • Claims processing • Claims rooting, triage • Chatbots • Image recognition for claims, repair costs **Transformer neural network** • Automated claims assessment • Improve underwriting effectiveness • Support agents in the sales process **Generative Adversarial Networks (GANs) and Generative Pre-trained Transformers (GPTs)** • Improved chatbots and other digital interactions (e.g., ChatGPT)	**Model Based value estimation** • Assess the risk of a particular policyholder and to accurately calculate premiums • Predict future losses to better manage risk • Detect errors in policyholder data • Error and fraud identification **Q-learning** • Predict customer churn • Automate underwriting processes, such as determining risk factors, setting premiums, and detecting fraud

Insurance companies deploy multiple types of algorithm categories for different use cases. Supervised learning systems are the most used and their broad use may be justified by the interpretability of results. Among the supervised learning algorithms, regression models and decision trees stand out, independent of company size.

Deep learning techniques like 'deep belief networks' or 'transformer neural networks' are seen less frequently due to their computational costs, the need for suitable expertise and large (labelled) datasets. Large and medium-sized companies display interest in also utilising semi-supervised and unsupervised systems to address tasks such as natural language processing (NLP), clustering and anomaly detection. Lastly, the low implementation rate of reinforcement learning algorithms suggests a misalignment between business needs and the technological solution.

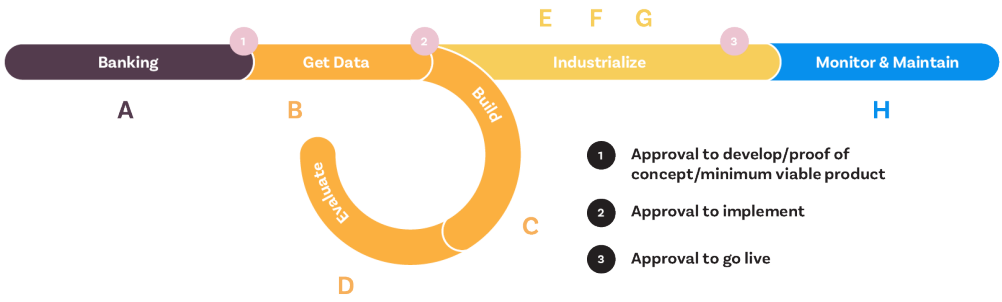
Figure 5. Which of these AI systems are deployed in your company?



2.3. Internal auditors must understand the AI development cycle

The development of AI is not, as is traditional software development, mainly a coding job. It has parallels with actuarial model management, but typically follows a more agile delivery process. McKinsey uses the model below to show the broad scope of AI development:

Figure 6. The AI development cycle (McKinsey, 2020)



A. Designing the solution, **B.** Obtaining reliable data required to build and train model, **C.** Building a model that achieves good performance in solving the problem specified during ideation, **D.** Evaluating performance of model and engaging business regularly to ensure business fit, **E.** Moving model to production environment, **F.** Deploying model where it starts with being used by the businesses, **G.** Inventory management of all models, **H.** Live monitoring in production

For internal auditors, it is important to understand the concept of data engineering tasks and concepts for AI models in development and production, as controls are needed during the whole lifecycle of processes based on AI implementation and operations. Internal auditors should not only focus on the more traditional software development controls like design reviews and approvals or user acceptance testing. Controls around data-usage, data pre-processing, model training, outcome fairness and live monitoring in production are equally important. Internal auditors should also understand that building and managing AI systems is a team effort. Business owners, data owners, data engineers, data scientists, risk management and ethics specialists must all work together to make sure that AI systems meet requirements.

3. Recommendations for AI risk management

Internal auditors can help organisations to find and keep balance. With new technologies, new risks arise. With all the potential that AI technology has to revolutionize the Insurance industry, a balancing act starts between risk mitigation and business innovation. Internal auditors should take the company's AI roadmap and maturity level into account to make maximum impact.

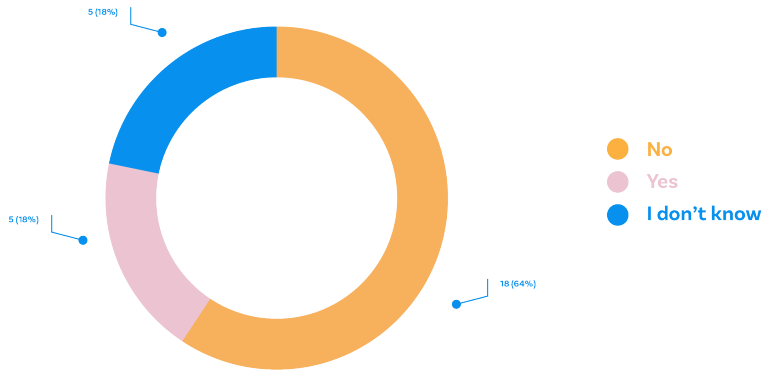
3.1. Include AI in governance, policies and procedures

As always, risk management should start with an AI vision and policy which is made concrete in an AI strategy that should be aligned to the business strategy and risk appetite. This provides transparency within the company on how AI should be used, and how it should not be used. Clear guidance helps the organization to find opportunities to harvest the benefits of AI, whilst staying true to the company strategy and risk appetite. A policy must provide guidance on the company's definition of AI, the kind of AI applications that are (not) allowed, the extent of decision-making or influence that is allowed for AI systems, and processes and controls to follow when implementing AI systems.

Depending on the company's AI strategy, AI clauses can be specified in a separate AI policy to stress the importance of AI for the business or can be incorporated into existing policies, such as model governance policies. Moreover, privacy policies address consequences and regulatory demands for privacy and security of personal identifiable information (i.e. GDPR) and should be updated to the application of AI.

Specifically, consent for customer data processing is key. Consent is only meaningful and valid when the customer knows exactly what data they are consenting to share and how the data is used in AI systems. There are several risks discussed with consent on AI and big data practices, which can erode the role of informed consent as it pertains to the use of personal information (Adam J. Andreotta, 2022).

Figure 7. Does your company have a formal AI policy?

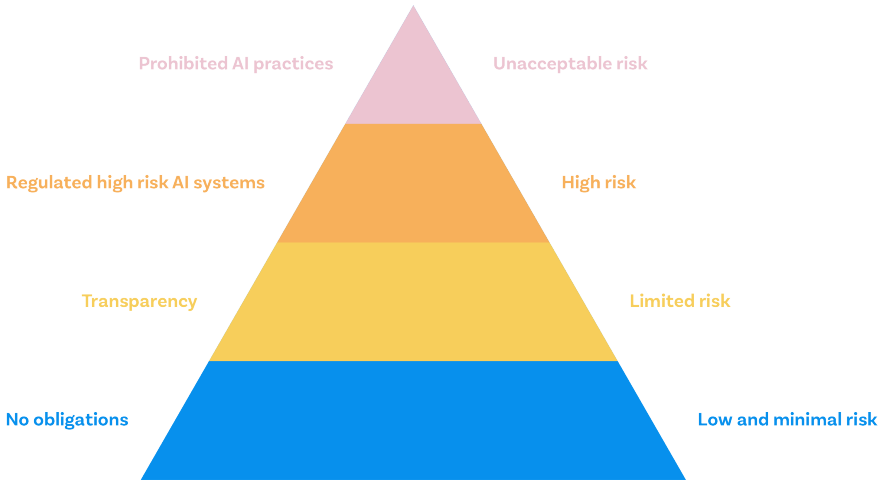


Internal auditors should align their audit approach to the company's AI maturity and risk appetite. The approach can be designed in multiple ways: from addressing the importance of an AI policy when companies are just getting started with AI, to testing if the AI policy is in line with the company's values and if the policy is followed when implementing AI systems for mature companies. Whereas the survey indicates a good coverage in terms of Data Management policies, the formal policies around AI are recent or still absent as illustrated by the graph. In light of this context, internal auditors can have a significant impact by addressing the need to introduce an AI policy and by refining existing ones.

3.2. Focus on high risk AI systems in governance, policies and procedures

The aforementioned balancing act between harvesting AI opportunities and managing the related risks requires a risk-based approach, in which additional controls are only implemented when needed to reduce the risks to an acceptable level. Such a risk-based approach is also mandated by the European AI Act:

Figure 8. Risk-based approach in AI (subject to changes) (EPRS, 2022)



The AI Act requires companies to set up and maintain an inventory and risk assessment of AI systems in use. AI systems classified as 'High risk system' are to be published in a European database before being put into operation. In addition the AI Act defines specific requirements for High risk AI systems regarding:

- Risk management system: A risk management system must be established and maintained for high-risk AI systems throughout the AI system's lifecycle. Identified risk management measures must be sufficient to reduce residual risks to an acceptable level and residual risks must be communicated to users.
- Data and data governance: Training, validation and testing data sets are subjected to appropriate data management and data governance practices. Requirements are for example set regarding the collection and preparation of these data sets, examination of possible biases, and the identification of possible data gaps or shortcomings and how they can be addressed.
- Technical documentation: The AI contains an appendix with the information that, at minimum, should be included in the technical documentation. The technical documentation must be drawn up before the system is placed on the market or put into service and must be kept up-to date.
- Record-keeping: High risk AI systems must be designed with automatic recording of events. These logging capabilities must ensure traceability of the AI systems functioning throughout its lifecycle.
- Transparency and provision of information to users: The AI system's operation must be sufficiently transparent for users to interpret its results. Also AI systems must be accompanied by instructions for use with information, available and comprehensible to users, regarding the characteristics, capabilities and limitations of performance of the high-risk AI system.
- Human oversight: The AI system must be designed with measures to facilitate effective oversight by humans. Individuals responsible for the human oversight must fully understand the capabilities and limitations of the AI system. They must also be able to correctly interpret the results of the AI system and remain aware of the possible tendency of over-relying. And they must be able to decide not to use, or even stope, the AI system.
- Accuracy, robustness and cybersecurity: High risk AI systems must be designed and developed in such a way that they achieve an appropriate level of accuracy,

robustness and cybersecurity, and perform consistently in those respects throughout their lifecycle. High-risk AI systems that continue to learn must be developed in such a way to ensure that possibly biased feedback loops are addressed with appropriate mitigation measures.

The first step for companies to be compliant with the AI Act, is to establish a formal risk management cycle for AI. This formal risk management cycle must be included in the AI development process. During development, risks should be identified, assessed and mitigated when necessary. However, the survey's outcomes, as illustrated below, highlight a marked immaturity in this regard.

Figure 9. How many of the deployed AI systems have a risk assessment?

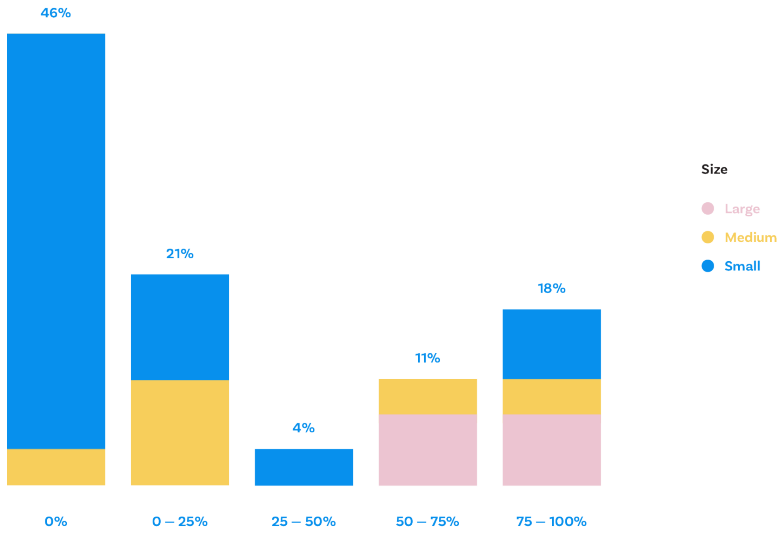
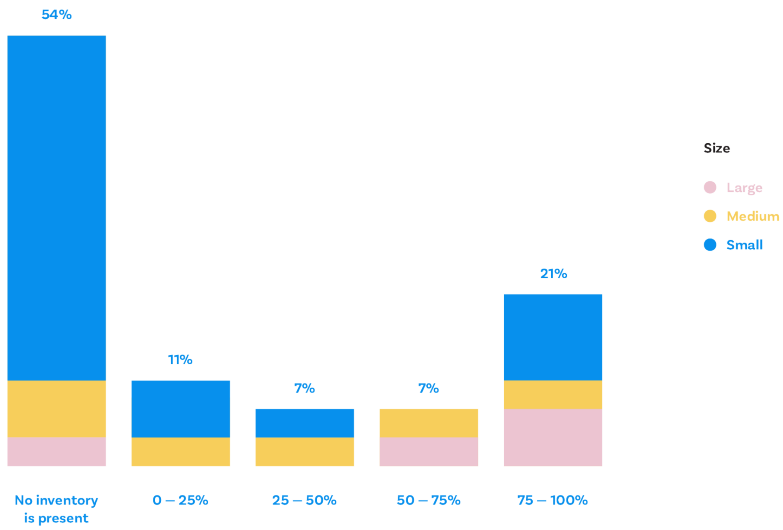


Figure 10. How many of the deployed AI systems are defined and documented in an inventory?



Whereas large companies have good coverage, medium and small entities still need to catch-up in this area. Therefore, internal auditors can play a fundamental role in getting AI risks managed and ensuring a swift implementation of external regulation. Depending on the company's AI maturity, internal auditors should stress the importance of incorporating the risk assessment into existing processes or confirm that all relevant AI risks have been appropriately addressed during the risk assessment.

Looking at the risks related to AI systems, many are already known software development or model management related risks. There are however new aspects, like bias or model drift, that should be addressed as well. The overview below shows typical AI risks and resilience factors, clustered around a framework presented by the IIA in their Global Perspective and Insights series on AI:

Figure 11. AI risks and resilience factors framework (IIA, 2018).

4. Recommendations on auditing AI systems

Following a top-down approach, it seems wise for internal auditors first to assess the governance for AI at organization level and if needed to advise about policy definition and implementation, inventory practices and risk assessments for AI application in the organization. In addition, it can also be valuable to deep dive into AI systems and their development cycle. In this section we focus on how internal auditors can develop an approach for audits on AI systems. A concrete proposal of an AI Audit Program is described in the next chapter.

4.1. Invest in AI knowledge and experience

To accomplish its role and provide assurance on AI, Internal Audit should recognize that new skillsets are required. Collectively, internal audit departments must have a sufficient understanding of AI, of how the organization is using it, and of the risks that AI represents to the organization. The survey results show an interesting pattern. Although internal expertise is still very limited among large and medium companies, internal audit departments are not too keen on getting external support to audit AI systems.

Figure 12. How familiar are your auditors in testing AI systems?

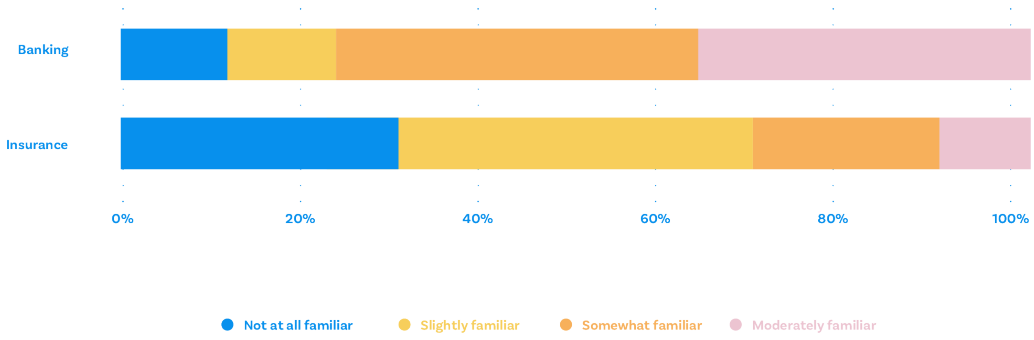
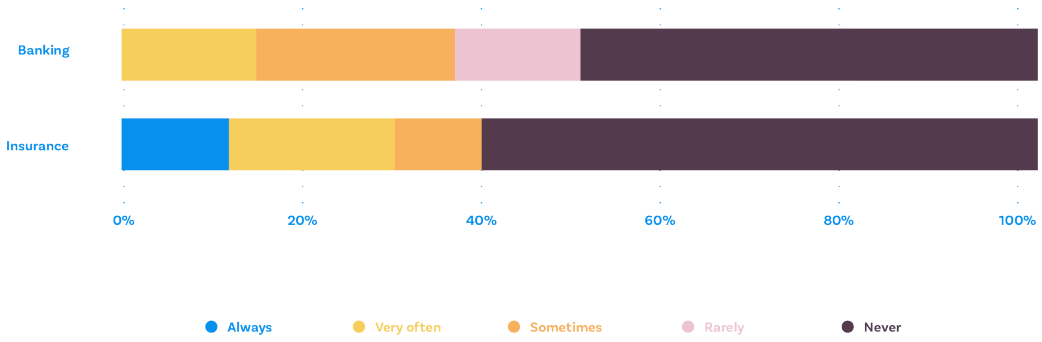


Figure 13. How often do you use external expertise when auditing AI systems?



Hence Internal Audit departments should invest in AI knowledge and expertise. Depending on the size of the internal audit department, they can consider embedding this within the team by training existing staff or hiring new employees with a machine learning or data science background. Alternatively, organizations might bring in external expertise when performing audits on AI systems or develop a guest auditor program. This can ensure the necessary knowledge and expertise and at the same time help to develop the AI skillset of the internal audit team itself.

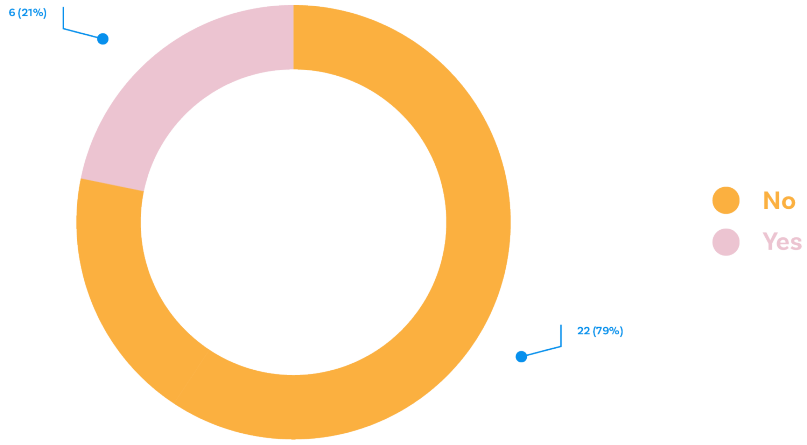
Auditing AI however does not only require knowledge about technical AI aspects. Knowledge and experience regarding data governance, data quality, ethics and business processes are

equally important. Therefore, a multidisciplinary audit team should be formed to cover all necessary areas. A well-equipped audit team should have representatives from IT audit, data science, business audit, and specific technical expertise such as actuaries, as well as ethics and external regulations to ensure each aspect is thoroughly investigated.

4.2. Adopt and adapt an AI audit framework

Survey results illustrated below suggest that assessing a company's AI framework is still not a common practice. While large companies have assessed both the framework and the AI systems in the last 24 months, this was not the case for medium and small companies. However, this tendency could be explained by the consistently different levels of maturity in the introduction of AI systems.

Figure 14. Has the Internal Audit department assessed your company's AI framework or strategy in the last 24 months?



When performing an audit, internal auditors should first establish their audit framework including control objectives and controls to be tested. International organizations have already been working on standards and audit frameworks for AI and algorithms. Internal auditors can adopt and adapt these frameworks to establish an audit framework with the control objectives and controls matching the organizational requirements. Examples of recently published guidance on audit approaches towards AI are listed below.

The International Organization for Standardization (ISO) has been developing around 30 standards on big data and AI:

- Two published standards on AI and ethics
- Three under development on AI ethics topics
- Other 15 standards under development on AI

The Dutch Association of chartered IT-auditors (NOREA) developed the "Guiding Principles Trustworthy AI investigations" (NOREA, 2021) to support IT-auditors in performing ex-ante or ex-post investigations of algorithmic systems. The Guiding Principles are structured according to the Cross Industry Standard Process for Data Mining (CRISP-DM). For each phase of the CRISP-DM process, 5 risk categories are established: Governance, Ethics, Privacy, Performance and Security, and 119 related key considerations have been developed.

In 2018 **ISACA** (Information Systems Audit and Control Association) suggested the use of the Cross Industry Standard Process for Data Mining (CRISP-DM) framework as a viable solution for an Audit Framework (ISACA, 2018). Based on the CRISP-DM model, the steps follow the needs of the machine learning process: gain an understanding of the business, gain an understanding of the data, prepare the data, complete modeling, evaluate, deploy. By following the CRISP-DM approach, a level of audit assurance can be obtained by a high-level review, with more assurance provided if subject matter experts examine each step in more depth.

DNB (Dutch Central Bank – De Nederlandsche Bank) published in 2019 a guidance document containing general principles for the use of artificial intelligence (AI) in the financial sector (DNB, 2019). Financial undertakings using AI should adhere to principles of sound and controlled business operations. In the Guidance Document, DNB has formulated several general principles regarding the use of AI in the financial sector divided over six key aspects of responsible use of AI: soundness, accountability, fairness, ethics, skills and transparency.

One of the most prominent examples of guidance for Internal Audit functions providing assurance on Artificial Intelligence comes from IIA. In 2017, the Institute of Internal Auditors issued a special edition of the series "Global Perspectives and Insight", entitled "Artificial Intelligence – Considerations for the Profession of Internal Auditing" (IIA, 2017). The paper, first of a series of three (followed by The IIA's Artificial Intelligence Auditing Framework Practical Applications, Part II (IIA, 2017) and Artificial Intelligence Part III (IIA, 2018)), explores the concept of artificial intelligence, and presents a high-level overview of considerations for the internal auditing profession about AI Strategy, Governance (including data architecture and infrastructure, data quality, and performance measurement) and the Human Factor.

Following IIA's guidance, "*Internal audit can help an organization evaluate, understand, and communicate the degree to which artificial intelligence will have an effect (negative or positive) on the organization's ability to create value in the short, medium, or long term. Internal audit can engage through at least five critical and distinct activities related to artificial intelligence:*

- *For all organizations, internal audit should include AI in its risk assessment and consider whether to include AI in its risk-based audit plan.*
- *For organizations exploring AI, internal audit should be actively involved in AI projects from their beginnings, providing advice and insight contributing to successful implementation.*
- *For organizations that have implemented some aspects of AI, internal audit should provide assurance over the management of risks related to the reliability of underlying algorithms and data on which the algorithms are based.*
- *Internal audit should ensure the moral and ethical issues that may surround the organization's use of AI are being addressed.*
- *Like the use of any other major system, proper governance structures need to be established and internal audit can provide assurance in this space*".

Internal auditing should approach AI as it approaches everything — with systematic, disciplined methods to evaluate and improve the effectiveness of risk management, control, and governance processes related to AI".

The Audit Framework proposed by IIA is comprised of three components: AI Strategy, Governance, and the Human Factor as briefly explained below.

AI Strategy

Each organization needs a well-defined AI strategy to capitalize on opportunities provided by AI. Internal audit should look at the organization's AI strategy, including its investment in AI research and development and plans to address AI threats and opportunities. AI can provide a competitive advantage, and internal audit should help management and the board formulate a deliberate AI strategy consistent with the organization's objectives.

Governance

AI governance refers to the set of structures, processes, and procedures that an organization implements to direct, manage, and monitor its AI activities to achieve its objectives. In particular, accountability and oversight must be established with reference to:

- data architecture and infrastructure (the way that data is accessible, information privacy and security throughout the data lifecycle, roles and responsibilities for data ownership and use);
- data quality (completeness, accuracy and reliability of the data);
- performance measurement (definition of metrics to tie AI activities to business objectives and illustrate whether they are effectively supporting the achievement of those objectives);

The Human Factor

The human factor component is concerned with identifying and managing the risk of unintended human biases in AI design, testing AI to ensure that results reflect the original objective, ensuring transparency in AI technologies despite their complexity, and ensuring that AI output is being used legally, ethically, and responsibly. An area of concern for Internal Audit is linked with the increasing use of the most advanced AI technologies, powered by algorithms that are less and less transparent and more difficult to understand. This lack of transparency presents a challenge for organizations which should develop strategies and techniques to ensure clarity and accountability in the use of AI.

A proposal of a concrete audit program to support IA functions in assessing how organizations are leveraging on AI is reported in the next chapter. It elaborates and deepens the three macro categories depicted above and identifies specific risks to be addressed in an audit engagement, the related possible root causes, and some examples of possible testing strategies. Such program is structured in 7 key risk areas: 1 - Strategy & Governance, 2 - Legal & Compliance, 3 - Development of AI systems, 4 - Operations Management for AI systems, 5 - Security and Data Protection, 6 - Human Capital, 7 - Sustainability. Please refer to chapter 6 for details.

Based on a sample of real audit engagements conducted in major insurance groups, AI-related risks depicted in the IIA framework in fact materialized. The following are typical audit findings raised during the engagements:

- *Unclear or blurred accountabilities and controls for models;*
- *1st, 2nd and 3rd line governance for models/ AI instances not clear;*
- *Gaps in AI model validation (no validation policy, no validation performed, or validation performed with weaknesses);*

- *No inventory or inventory not complete, not updated;*
- *Risk assessment of AI models not complete, not updated;*
- *Operational gaps on model operations;*
- *Issues around Third-Party management;*
- *Data quality issues with training data;*
- *Missing controls on model drift;*
- *Real time monitoring of models not in place;*
- *Models exposed to external attacks;*
- *Missing monitoring controls for models;*
- *If models are used for robot automation, access issues;*
- *Segregation of duties conflicts.*

4.3. Determine the audit scope and test approach

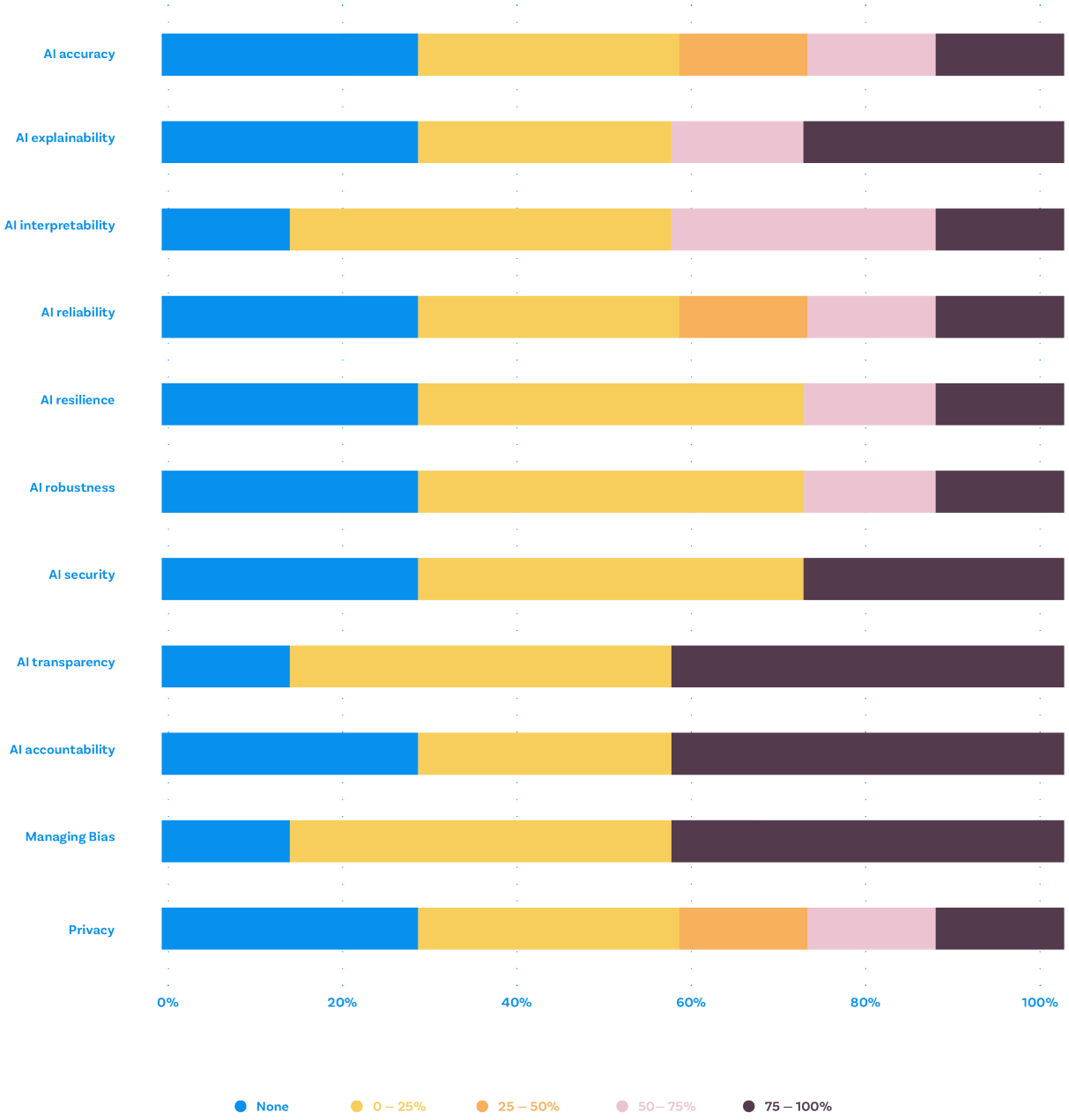
In addition to selecting appropriate audit standards, scoping of the audit is key in helping the organization manage the AI risks. Auditors normally have two options when performing an audit on AI. They can choose to focus on a process, being the AI development cycle, or they can focus on the outcome of a process: the AI system that was developed. Considering the maturity of the AI development and risk management cycles, internal auditors are likely to start with deep dive audits on the AI systems that were developed. Starting with the AI systems that pose the highest risk for the organization, internal auditors can provide valuable insights for those key risk systems and at the same time help mature the development cycle.

In addition, internal auditors must choose from which perspective they want to audit the AI system. They can choose to include all AI risk areas mentioned in section 4.2 in their audit scope. It is also possible to focus on a subset of risk areas. The objective of the audit, the nature of the AI system and other audits in the audit plan are key elements to consider when deciding on the AI risk areas that should be audited.

Our survey shows that internal auditors from the participating insurance firms focus on a selection of AI risks areas for their audits, instead of addressing all areas in every audit. The highest focus is on risks such as accountability, transparency, and managing bias. Least focus is on more technical topics such as security, robustness, and resilience (see Figure 15).

Internal auditors should also decide on the approach they apply for testing the AI system. Internal auditors can choose to perform a desktop review. This is the more traditional test approach in which the internal auditor follows the audit trail from the development process to test if all requirements are met. They can also choose to perform a full review approach. This approach is not focused on the development process, but on independently testing the actual behavior of the model. The internal auditor can for example use the available test and train data and train the model to establish if this results in similar model performance. The internal auditor may even independently build a reference model and compare the model performance with the actual model being audited. In a hybrid approach, the internal auditor combines elements of both the desktop and the full review approach.

Figure 15. In how many of the AI-related audits did you Internal Audit department consider any of the following characteristics?



Internal auditors should carefully consider their approach. Choosing for a full review will significantly increase the level of assurance provided but it is more difficult and costly to perform. Thus, it requires the availability of the original data and a strong knowledge in AI development from internal auditors. It is therefore not surprising that the majority of respondents – from both Insurance and Banking sectors – mainly applied the Desktop review or the Hybrid approach instead of the Full review.

Figure 16. Different review-approaches and their levels of feasibility.

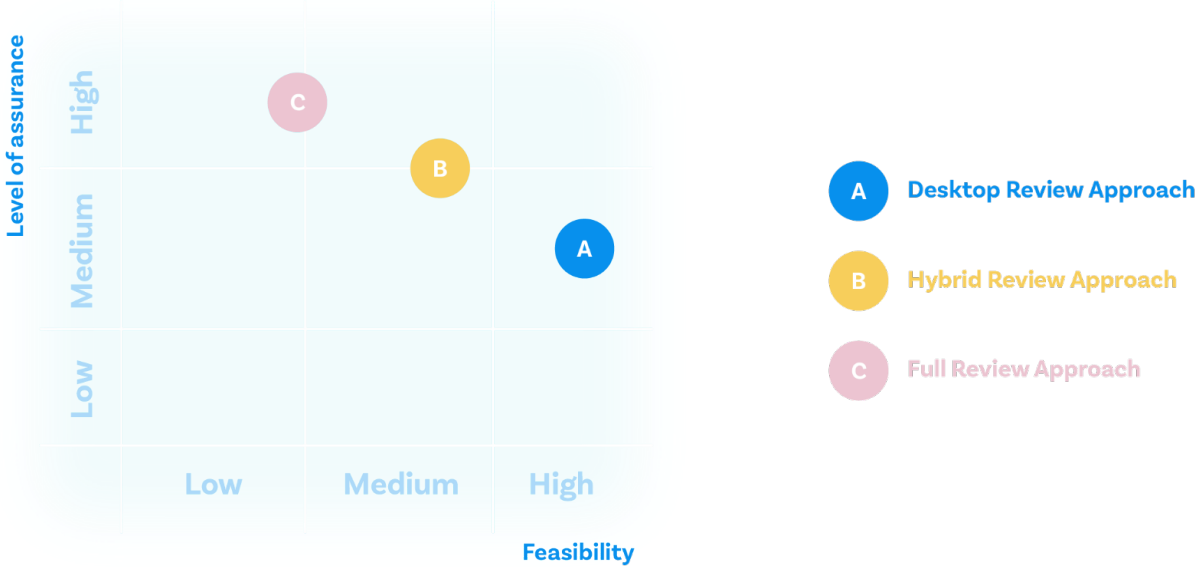
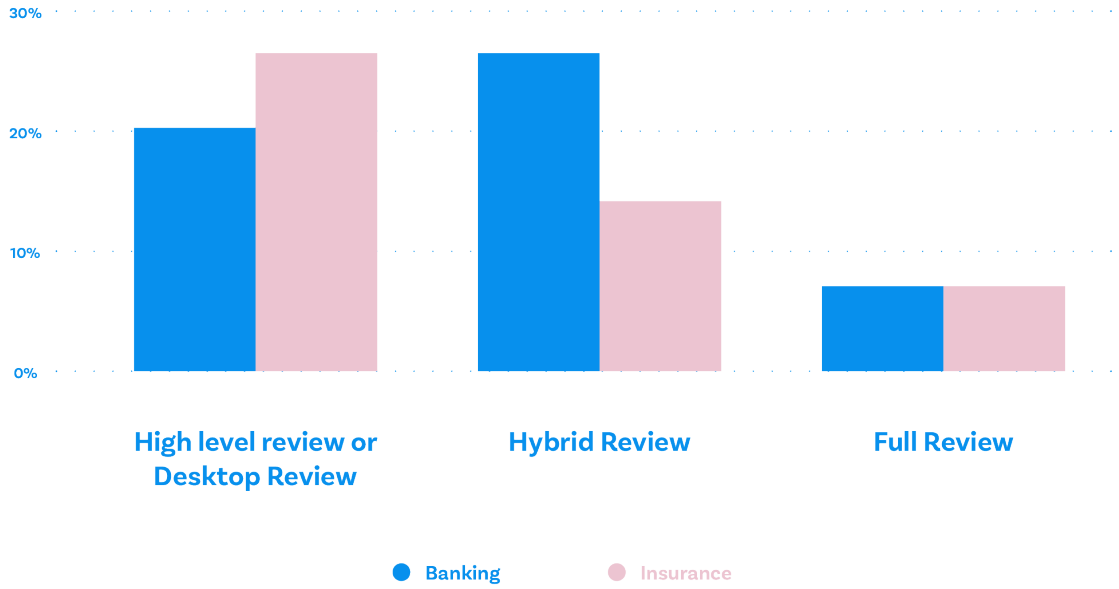


Figure 17. What was your preferred testing approach in the AI-related audit(s)?



5. A concrete proposal of an AI Audit Program

A concrete audit program[1] is presented below in order to support IA functions in assessing how organizations are leveraging on AI. It elaborates and deepens the three macro categories depicted above (Strategy, Governance and Human factor) and identifies specific risks to be addressed in an audit engagement, the related possible root causes, and some examples of possible testing strategies. Such program is structured in 7 key risk areas: 1 - Strategy & Governance, 2 - Legal & Compliance, 3 – Development of AI systems, 4 - Operations Management for AI systems, 5 - Security and Data Protection, 6 - Human Capital, 7 - Sustainability.

Risk Area	1 - Strategy & Governance
Risk	Weaknesses in the strategy definition and cascading with reference to the vision and implementation of AI, and bad governance of AI initiatives due to a lack of commitment and/or consideration of AI, and/or governance frameworks (e.g., policies, procedures, guidelines) and/or due to unclear accountability, leading to a lack of support in the company, failure of implementation projects and/or underperformance compared to competitors in the market.
Possible root causes	Lack of a clear vision (document) about the risks and opportunities of AI for the company Lack of an overall defined, approved and communicated strategy that fulfils the vision Insufficient setup and/or poor organizational commitment for AI projects Lack of appropriate performance metrics and/or monitoring for AI application Lack of defined and approved policies, procedures, operating models, roles, and responsibilities
Possible testing strategy	Has an AI strategy been defined and documented? How are AI initiatives communicated within the organization? How is the accomplishment of the objectives and the AI strategy monitored? Are AI-specific standards and procedures incorporated into internal requirements? Have a governance framework and operating models been defined, with a RACI matrix? Were industry leading practices (e.g., COSO/ COBIT/ ISO etc.) used as an input for the review / update process? Do policies and procedures sufficiently address AI risks and opportunities and require periodic what-if analysis and/or scenario planning? Did Senior Management approve a code of conduct to define how to use and develop AI solutions? Is the organization trained to it?

[1] Please note that the audit program presented is a comprehensive one and not specifically focusing on regulatory compliance with the upcoming EU act on AI

Risk Area	2 - Legal & Compliance
Risk	Compliance breaches due to insufficient governance and regulation analysis or acknowledgement of legal requirements with reference to AI initiatives leading to adverse publicity in media and/or non-compliance with external regulations and resulting potential regulatory fines.
Possible root causes	Missing awareness or neglect of ethical aspects in the AI application Missing analysis or acknowledgement of legal and supervisory requirements Inappropriate use of, or access to, data Inadequate record keeping for services and transactions undertaken by the organization
Possible testing strategy	Did you carry out a fundamental rights impact assessment where there could be a negative impact on fundamental rights? Could the AI system affect human autonomy by interfering with the user's decision-making process in an unintended way? In case of a chat bot or other conversational system, are the human end users made aware that they are interacting with a non-human agent? Have all relevant external regulations as well as all relevant international standards for the use of AI been identified and is the organization aware of it? Are AI initiatives prepared for compliance with new technology regulations, such as the EU's General Data Protection regulation (GDPR) or the Proposal for AI Regulation? Are international standards (e.g., ISO, IEEE) adopted? Is the necessity to use personal data in AI use cases assessed prior to doing so? Is consideration given to how personal data will be used to train an AI system, as well as the purposes for processing for every stage of the lifecycle where personal data will be processed? Is there a process in place to determine whether the data used by the AI solution need to be anonymized / tokenized or whether synthetic data should be used?

Risk Area	3 – Development of AI systems
Risk	Bad implementation and failure of AI initiatives (including poor performance and robustness over time, possible biases, lack of interpretability and explainability) due to technical malpractice, leading to operational inefficiencies, financial losses, reputational damage.
Possible root causes	Unstructured design phase Inadequate collection and preparation procedures for the development of AI systems Missing key steps in the development process for AI systems Change management processes not adequate and/or change activities carried out not in a coordinated way
	Has a risk assessment to determine appropriateness of features (e.g. consider issues of ethics, fairness / unwanted-bias, identification if in the

Possible testing strategy	data set or created features there are sensitive features, etc) been performed in the design phase? Has a process for the development of AI systems been defined? Does the process include: • Assessment process to confirm relevance of features; • Analysis to detect features that may cause correlations that appear to be, but are not actually, valid; • Monitoring and periodical testing of the outputs to (re)-validate the appropriateness of selected features, including proxies, to prevent and/or detect unethical, unfair or unwanted-biased results Are minimum data specified for the use case, proportional to the complexity of the problem being solved? Does data specification include: • Data definitions (including name, source, type, description, format, how the data will be used), • Minimum data volume and partitioning requirements, • Data ownership (e.g., owner, producer, consumer etc), • Expected ranges (e.g., range of permitted values, time-series), • Expected statistical attributes (e.g., expected distributions); and/or • Assessment of potential issues. Are standard data quality controls in place to ensure that data used by the AI system are complete, accurate, reliable, and timely? Is there a plan in place to train, calibrate and optimize the performance of the AI systems, including success criteria? Does the plan describe the strategy for the model training including: • Benchmarking, challenger models, and comparisons to pre-deployment models using standard performance and fairness measures; • Back-testing • Sensitivity analysis to test the robustness under different conditions; and/or • Testing to confirm the validity of (or identify new) assumptions, limitations and weaknesses identified. Do the tests consider all the fairness measures and interpretability requirements? Is any explainability technique implemented, such as SHAP values[2] or LIME[3]? Is the output of reports/performance monitored against expectations and maximum deviation metrics? Has a Change Management process for AI systems been defined, as well as Release Management process and related request modalities and templates? Are requests for change to the AI solutions categorized, appropriately approved, planned, and scheduled? Are changes: • tested by authorized independent personnel who did not develop them?

[2] SHAP (SHapley Additive exPlanations) values are used whenever you have a complex model (e.g., a neural network, or anything that takes some features as input and produces some predictions as output) and you want to understand what decisions the model is making. This algorithm was first published in 2017 by Lundberg and Lee and it is a brilliant way to reverse-engineer the output of any predictive algorithm.

[3] LIME (Local Interpretable Model-agnostic Explanations) is a method for explaining predictions of Machine Learning models, developed by Marco Ribeiro in 2016. It works for any kind of Machine Learning model, Model-agnostic, and aims at explaining only a small part of the Machine Learning function, Local.

- authorized for the migration into the production environment by independent personnel who did not develop them?
- migrated into production environment by an authorized independent IT personnel who did not develop it?

Risk Area	4 – Operations Management for AI systems
Risk	Bad implementation and failure of AI initiatives due missing compliance with governance frameworks (including lack of human oversight and accountability), leading to operational inefficiencies, financial losses, reputational damage.
Possible root causes	Missing clarity on the explanation of the processes, services and decisions delivered or assisted by AI to the individuals affected by them Improper or inhomogeneous approaches for the use of AI initiatives Lack of operating models, roles and responsibilities Lack of staff motivation to work in 1st line departments where AI is heavily involved resulting in poor oversight and increased difficulty to attract and hire people.
Possible testing strategy	Are business and functional specifications for AI business processes and interfaces with other applications/systems collected, documented, and approved? Has an adequate level of human supervision been put in place by the process owner? The design should take into consideration the level of criticality and the risk assessment of the AI initiative, as well as the level of reliability of the implemented AI solution. Based on this analysis the correct level of automation should be identified and validated by the relevant stakeholders (e.g., human-in-the-loop vs human-out-of-the-loop). Are there documented audit trails/history logs for AI systems, providing sufficient information to understand what AI decisions and amendments were made and why, and to allow replicability of the results? Does the AI Center of expertise of the Company or any other governance mechanism in place maintain an accurate firm-wide inventory of AI initiatives/systems within the application inventory tool to ensure the enterprise application inventory is up to date? Are responsibilities related to the monitoring of the business performances of AI initiatives appointed to specific resources/teams of resources? This pretains to both economic and process performances, as well as adequacy of achievement of the desired business objectives. Were business KPIs defined over the relevant steps of the process, referring to both economic and process performances, as well as adequacy of achievement of the desired business objectives? Have requirement criteria for recalibration / retraining been identified? Is there a job rotation or training program in place with specific reference to AI oversight to avoid the loss of staff motivation working with AI?

Risk Area	5 - Security and Data Protection
Risk	Failure of implementation of AI systems due to a lack in AI system resilience (missing identification of alternative solutions to ensure the business continuity) or inadequate data protection, change management and/or

	monitoring processes which may lead to business interruption and/or malfunctioning.
Possible root causes	Lack/inadequate access security management processes Lack of implemented framework and of the maintenance plan's periodical review Lack in AI system resilience Inadequate risk assessment and mitigation procedures in the development of the AI systems Insufficient (or absent) measures to monitor and suspend the system when necessary
Possible testing strategy	Are access rights determined in line with the firms Information Security policy requirements, including, but not limited to rhe principles of Need to Know, Need to Do, Least Privilege? Has a plan for the secure disposal and/or decommissioning of the AI system been defined and documented? Disposal and decommissioning of AI systems must also be performed: • In agreement with the impacted business areas, • In compliance with the firm's data retention schedule; and • Other relevant Information Security and Physical Security requirements (e.g., permanent destruction or sanitization of data). • The plan must be reviewed and approved by an independent individual with sufficient authority (e.g., CIO, CISO or their delegated representatives). Are the Business Continuity Plans updated to ensure the AI system is included in the recovery strategy? Were assurance tasks (e.g., Vulnerability assessment, Penetration Test, Web Application Scanning) executed to support the accreditation of new or modified systems during the implementation of AI solutions carried out? Was an analysis of data sources, especially those outside/online performed to assess the risk for adversarial attacks or data poisoning? In case, verify that the design of data requirements includes to filter and analyse data received from outside/online sources used as training data for training/improving AI models (e.g., When receiving input data from online input for the training of a continuous learning algorithm) to avoid data poisoning attacks on training data that can manipulate results of a predictive model.

Risk Area	6 - Human Capital
Risk	Insufficient technical capabilities and know-how for the development and use of AI systems on one side, and vanishing business knowledge on the other side, due to lack of human capabilities, knowledge/collaboration tools and/or inadequate training on innovation leading to innovation backlog, security deficiencies and/or business disruptions
Possible root causes	Lack of an adequate skillset to manage AI technologies Failure to educate all employees in mandatory information security procedures on AI solutions in use Lack of an adequate training on project management for the implementation of AI systems Loss of business knowledge and data literacy over time in 1st line by absence of human practice (replaced by AI)

Possible testing strategy	Are currently available technical skills and competencies of internal and external resources assessed with reference to the requirements of AI systems? This includes capability to manage business continuity in case of AI disruption. Are there minimum skill requirements defined? Do people in charge of AI development have their stated qualifications? Are adequate training programs in place related to AI initiatives? Is an adequate level of business practice ensured to 1st line despite the use of AI?

Risk Area	7 - Sustainability
Risk	Insufficient technical capabilities and know-how for the development and use of AI systems on one side, and vanishing business knowledge on the other side, due to lack of human capabilities, knowledge/collaboration tools and/or inadequate training on innovation leading to innovation backlog, security deficiencies and/or business disruptions
Possible roots causes	Energy intensive processes or use of unsustainable energy
	Use of rare or unsustainable materials
Possible testing strategy	Does the Ai system involve or require the potential use of environmentally unsustainable materials (e.g. rare materials, or materials with high impacts on the environment or on the communities)?
	Has the use of environmentally unsustainable materials in relation to AI systems been considered in Company's ESG Strategy?

Annexes

Annex 1: Questionnaire for the ECIIA Paper on Auditing Artificial Intelligence and Machine Learning

A.1.1. AI definition

The questionnaire followed the **OECD** definition of an **AI-system** as "*a machine-based system that can influence the Environment by making recommendations, predictions or decisions for a given set of objectives. It does so by utilising machine and/or human-based inputs/data to:*

i. perceive real and/or virtual environments;
ii. abstract such perceptions into models manually or automatically;
iii. *use Model Interpretations to formulate options for outcomes*" (OECD, 2019)

Specifically, approaches and techniques listed in Annex I of the **EU AI Act** (2022) were considered in scope.

A.1.2. Company Information

1. Please, provide the name of Your company
2. How many auditors are there in Your Internal Audit department? (Headcount) Excluding Operations and Leadership
3. How many IT (Information Technology) auditors are there in Your Internal Audit department?
4. How many Data experts are there in Your Internal Audit department?
5. How many AI systems are you aware of in Your company? [None, 0-20, 20-100, 100-500, >500]

Which of these AI solutions are deployed in Your company? [Categorization based on Swiss Re, SIGMA: Machine Intelligence in Insurance]

- Supervised Learning = a model trained using data which are labelled (i.e., tagged with the correct answer). Relationships are inferred from the sample and used to map new examples
- Unsupervised Learning = a model that discovers the hidden structures in the unlabelled data on its own. Used for clustering and association.
- Semi-Supervised Learning = a model trained using a combination of a small amount of labelled data and a large amount of unlabelled data
- Reinforcement Learning = goal-oriented model (agent) that answers the question how can this be optimised? And learns from the environment
- I don't know

6. Which of these supervised AI-systems are deployed in your company? Select all that apply

Supervised Learning = a model trained using labelled data (i.e., tagged with the correct answer). Relationships are inferred from the sample and used to map new examples

- Regression models
- Generalized Linear Models
- Machine Learning Models with Boosting Gradient
- Deep Learning models
- Decision trees
- Other

7. Which of these unsupervised AI-systems are deployed in your company? Select all that apply
    - Clustering
    - Principal Component Analysis
    - NLP (e.g., BERT)
    - Deep Learning models
    - Other

8. To which industry does Your company belong?
    - Bank
    - Asset management company
    - Insurance
    - Other

9. Which of these insurance products does Your company offer? (Select all that apply)

[Categorization based on DIRECTIVE 2009/138/EC, Annex I and II]

- Individual LIFE
- Individual NON-LIFE – Property–Casualty (P&C)
- Individual health
- Commercial – Group Life/Pension
- Commercial – P&C Commercial
- Reinsurance

10. How much is each area supported by AI systems?

Example: If all customer interactions are supported by AI systems, then that equals to 100%
[0%, 0-25%, 25-50%, 50-75%, 75-100%]

- Customer service
- Claims handling
- Fraud detection
- Direct distribution
- Anti-Money Laundering (AML)
- Policy Administration
- Underwriting
- Financial Reporting
- Pricing
- Risk modelling

11. Is there any missing area in question 12? If yes, please specify which and the relative AI coverage

A.1.3. AI system risk assessment

1. Does Your company have a formal Data Management policy? [Yes/No/I don't know]
2. Based on the Data management policy, how many core business processes have a well-implemented data quality management guideline? [0%, 0-25%, 25-50%, 50-75%, 75-100%]Does your company have an AI policy? [Yes/No/I don't know]
3. 3.1 (conditional) When did Your company implement the formal AI policy? (year)?

4. How many of the deployed AI systems are defined and documented in an inventory?
5. [0-20%/ 20-40% / 40-60% / 60-80% / 80-100%/ No inventory is present]
6. How many of the deployed AI systems have a risk assessment? [0-20%/ 20-40% / 40-60% / 60-80% / 80-100%]

A.1.4. Auditing AI system

1. How many audits did Your Internal Audit department perform in the last 24 months in total?
2. Has the Internal Audit department assessed Your company's AI framework or strategy in the last 24 months? [Yes/No/I don't know]
3. Has the Internal Audit department assessed AI systems implementations in the last 24 months? [Yes/No/I don't know]
    a. How many AI systems did the Internal Audit department audit?
4. In how many of Your AI-related audits did Your Internal Audit department consider any of the following characteristics? [0-20%/ 20-40% / 40-60% / 60-80% / 80-100%]

AI Accuracy	Whether a model is correctly capturing a relation that exists in training data
AI reliability	Whether a model consistently generates the same results, with acceptable statistical error
AI robustness	Whether a model has a minimum sensitivity to variations in uncontrollable factors
AI resilience	Whether a morel can withstand unexpected changes in its environment or use
AI explainability	Whether a description of how model predictions are generated can be done
AI interpretability	Whether a description of the model's output can be done
Privacy	Norms and practices that help to safeguard values
Managing bias	A model can manage systemic, computational and human bias
AI transparency	Information is available to a user when interacting with an AI-system
AI accountability	Expectations for the responsible party if a risky outcome is realized
AI security	Counter measures to defeat machine learning attacks

5. What was your preferred testing approach?
    - High level review or Desktop Review = testing the controls around the model
    - Hybrid Review = testing the design, input, and output of the model
    - Full review = testing the behaviour of the model (rebuilding, stress-testing, quantitative validation)
    - Which level(s) did the Internal Audit department include in the AI-related audit(s)? Select all that apply [Source: Derisking AI: Risk management in AI development | McKinsey]

- Design level: testing the model requirements, intended uses, and specifications
- Input level: testing data quality, data treatment, and features Input level: testing data quality, model's assumptions and features
- Development level: model's assumptions and hyperparameters
- Output level: testing accuracy, bias and other quantitative measures
- Implementation level: testing system documentation, processing code, production environment
- Monitoring level: model drift, calibration and benchmarking
- Other

6. How familiar are the auditors in Your Internal Audit department in testing AI systems? [from not at all familiar to Extremely familiar]

7. How often do you use external expertise when auditing AI systems? [from Never to Always]

Annex 2: Bibliography

Adam J. Andreotta, N. K. (2022). AI, big data, and the future of consent. AI & SOCIETY volume 37.

AFME. (2018). Artificial Intelligence: Adoption in Capital Markets.

AFME. (2018). Considerations on the Ethical Use of Artificial Intelligence in Capital Markets.

Dan Hendrycks, M. M. (2022). X-Risk Analysis for AI Research. Cornell University, Computer Science > Computers and Society.

DNB. (2019). General principles for the use of Artificial Intelligence in the financial sector.

Douwe Kiela, M. B. (2021). Dynabench: Rethinking Benchmarking in NLP. Computer Science.

Dutch Association of insurers. (2021). Ethical Framework for data-driven applications by insurers.

EBA. (2022). Risk assessment of the European banking system.

EBF. (2019). AI in the banking industry: EBF position paper.

EBF. (2020). AI inception impact assessment consultation: EBF response.

EIPOA. (2016). Big data analytics in motor and health insurance: a thematic review.

EPRS. (2022). EU Legislation in Progress briefing. Retrieved from https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI%282021%29698792_EN.pdf

European Commission. (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence. Retrieved from https://ec.europa.eu/newsroom/dae/redirection/document/75788

European Parliament. (2023). AI Act: a step closer to the first rules on Artificial Intelligence. Retrieved from https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence

Goodfellow, I., Bengio , Y., & Courville, A. (2016). Deep learning.

IIA. (2017). Global Perspective and Insights: Artificial Intelligence – Considerations for the Profession of Internal Auditing.

IIA. (2017). Global Perspective and Insights: The IIA's Artificial Intelligence Auditing Framework Practical Applications (Part II).

IIA. (2018). Global Perspective and Insights: Artificial Intelligence, Internal Audit's Role, and Introducing a New Framework (part III).

ISACA. (2018). The Machine Learning Audit—CRISP-DM Framework.

IVASS. (2021). The Governance of Artificial Intelligence in the insurance sector between ethical principles, board responsibility and corporate culture.

Juliana Sandu, M. W. (2022). Time to audit your AI algorithms. Maandblad voor Accountancy en Bedrijfseconomie, issue 96.

Kop, M. (2021). The right to process data for machine learning purpose in EU. Harvard Journal of Law & Technology, Volume 34.

Martinez Plumed, F. B.-O. (2021). Research Community Dynamics behind Popular AI Benchmarks. NATURE MACHINE INTELLIGENCE, ISSN 2522-5839, https://www.nature.com/articles/s42256-02, 581-589.

McKinsey. (2020). Derisking AI by design: How to build risk management into AI development. Retrieved from https://www.mckinsey.com/capabilities/quantumblack/our-insights/derisking-ai-by-design-how-to-build-risk-management-into-ai-development

Nenad Tomašev, J. C. (2020). AI for social good: unlocking the opportunity for positive impact. Nature Communications.

NOREA. (2021). Guiding Principles Trustworthy AI Investigations.

OECD. (2019). AI Principles. Retrieved from https://oecd.ai/en/ai-principles

OECD. (2019). Recommendation on Artificial Intelligence. Retrieved from https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

OECD. (2020). The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector.

Schaller, R. R. (1997). Moore's law: past, present and future. IEEE Spectrum, vol. 34, no. 6,, 52-59.

Stanford University. (2022). Translation: Internet Information Service Algorithmic Recommendation Management Provisions. Retrieved from https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/

Swiss Re. (2020). Machine intelligence in insurance: insights for end-to-end enterprise transformation. Sigma.

The Geneva Association. (2020). Promoting Responsible Artificial Intelligence in Insurance.

About ECIIA

The European Confederation of Institutes of Internal Auditing (ECIIA) is the professional representative body of 34 national institutes of internal audit in the wider geographic area of Europe and the Mediterranean basin.

The mission of ECIIA is to be the consolidated voice for the profession of internal auditing in Europe by dealing with the European Union, its Parliament and Commission and any other appropriate institutions of influence. The primary objective is to further the development of corporate governance and internal audit through knowledge sharing, key relationships and regulatory environment oversight.

About ECIIA Insurance Committee

ECIIA set up an Insurance Committee[1] in 2012 with Chief Audit Executives of the largest European Insurance companies. The mission of the ECIIA Insurance Committee is: "*To be the consolidated voice on behalf of the profession of Internal Audit in the Insurance sector in Europe by representing and developing the Internal Audit profession as part of good corporate governance, achieving thought leadership through publications on relevant topics and by interacting with the Regulators, as required, and any other appropriate institutions of influence at European level.*"

ECIIA represents around 55.000 internal auditors and around 13.000 are active in the insurance sector. The paper describes the results of a survey and discussions amongst the Committee members.

We want to thank the Committee members and the workgroup members that wrote this paper:

Peter Galjaard, Senior IT Auditor at Achmea; Frank Heldens, Senior IT Auditor at Achmea; Lucia Hrovatin, Behavioural Data Scientist at Group Audit Zurich Insurance Company Ltd; Astrid Langeveld-Vos, CAE at Achmea; Salvino Marigo Head of IT & GOSP Audit at Assicurazioni Generali S.p.A.; Robert Zergenyi, Global Head Internal Audit - Operations and Technology and Zurich Global Ventures at Zurich Insurance Company Ltd; Chiara Ziliani, Head of Group Audit Analytics at Assicurazioni Generali S.p.A.

European Confederation of
Institutes of
Internal Auditing

Avenue des Arts 41
1040, Brussels—Belgium
TR: 84917001473652

LINKEDIN WEBSITE EMAIL