Copenhagen Fintech

PA

Responsible AI

PA Consulting & Copenhagen Fintech

Bringing Ingenuity to Life.
paconsulting.com

1

Introduction

Together, PA Consulting and Copenhagen Fintech interviewed several Nordic financial services leaders to explore how they can set the right guardrails for generative AI success.

From our conversations, we found that while even experienced hands can struggle to keep pace with the new wave of governance and ethical complexities brought about by generative AI, there are four broad steps that leaders can tale to ensure their usage is responsible, risk-managed, and more likely to deliver a return.

This guide enables financial institutions to future-proof AI deployments by embedding ethical principles into strategy, governance, and operations, while anticipating regulatory, reputational, and operational risks.

Preparing for the wave

The introduction of ChatGPT in November 2022 created a storm of technological transformation, fundamentally changing the way that the financial sector had to think about their operations.

While most technologies follow the well-known hype cycle phases of innovation trigger, generative AI (under different names) has continuously remained at the top of inflated expectations since this introduction.

In the Nordic financial services sector, machine learning and AI had been on the agenda long before November 2022; particularly in areas like fraud detection, credit scoring, and process automation.

However, even the most mature players are struggling to keep pace with the unprecedented speed at which new generative AI capabilities are being released.

Key challenges include:

Rapid technology turnover – the release cycles of foundation models are now measured in months, not years

Talent and skills gaps – even large Nordic banks report shortages of AI engineering talent, prompting accelerated hiring and upskilling programs

Governance at scale – the jump from structured data AI to generative AI handling unstructured text, code, and imagery introduces new governance complexities, ranging from bias management to regulatory compliance, requiring new controls and cross-functional oversight

This combination of rapid technological evolution, early-stage adoption, and growing strategic investment underpins the sentiment voiced by one Nordic banking leader:

> **“We are already preparing ourselves for the next wave, because we have only just seen the beginning of what AI can do.”**

Thomas Eatmon, Head of Responsible AI at Danske Bank

From strategy to measurement

Based on our interviews, there is a broad array of strategies at play. These range from experimenting and getting experience with AI while the organisation is still in early stages, through to seeking positive business cases and outcomes.

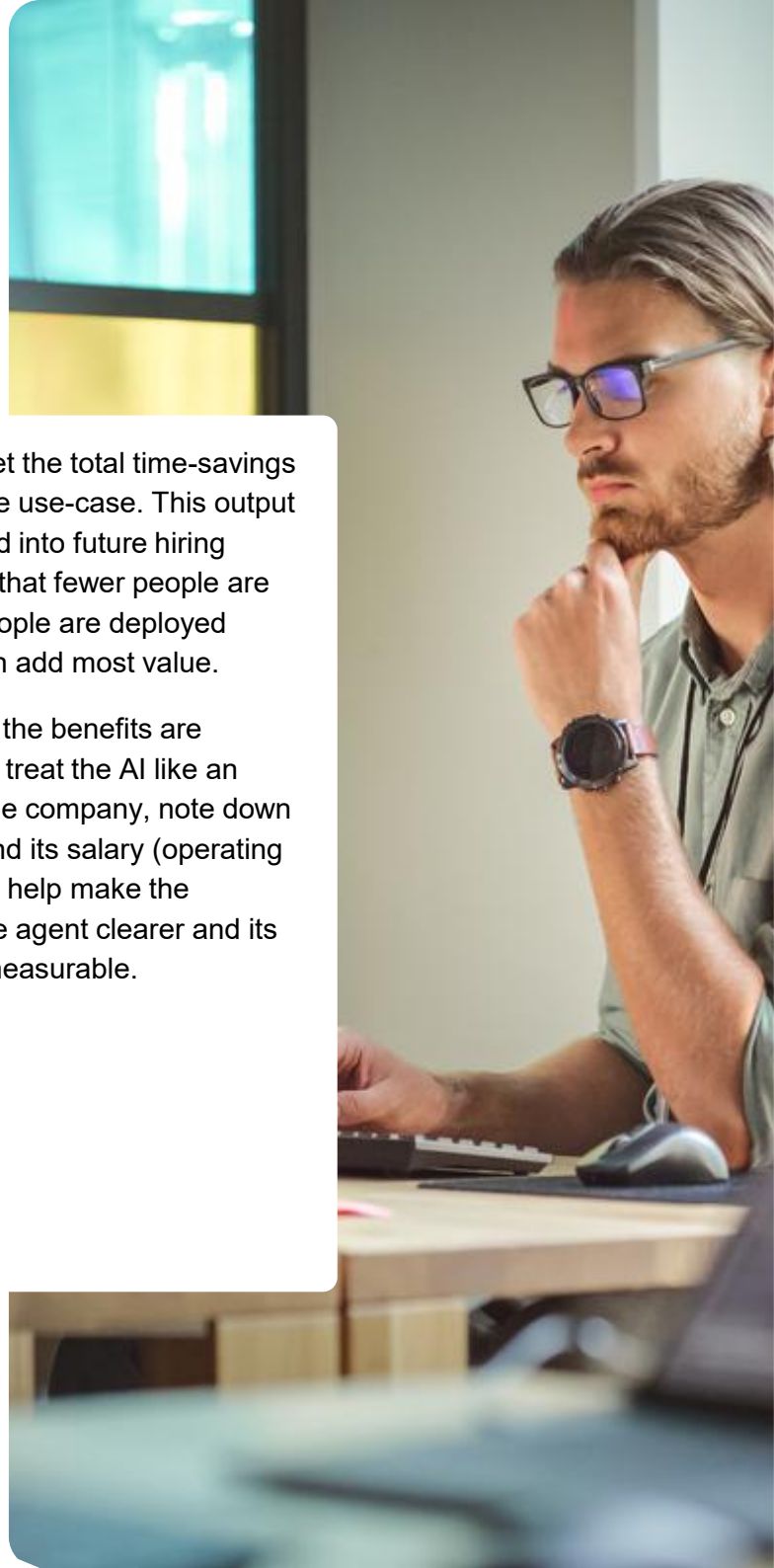
One thing that is becoming clear is that it is important to consider all aspects of the use-case when making the business case. It is easy to get caught up in efficiency gains alone – but usually there are more benefits to be gained.

For example, consider an AI-chatbot for customer-facing employees that can support them in case-handling and with customer interactions. It is likely that every employee will become more efficient to some degree. It is also likely that there are other potential benefits to be gained here. For example, the employees could be happier and feel more supported; case handling could become more consistent between call centres; and employees may make fewer mistakes.

And while there are many other benefits to AI, it is clear that efficiency is a central but intangible part of AI. Yet many organisations struggle with measuring the personal efficiency gain for employees.

The most straightforward way to approach this is to gather data on AI usage for that use-case, and gain an estimate on how much time users are saving with AI. This can then be multiplied to get the total time-savings estimate for the use-case. This output can be factored into future hiring processes, so that fewer people are hired – and people are deployed where they can add most value.

The other way the benefits are estimated is to treat the AI like an employee at the company, note down what it does and its salary (operating cost). This can help make the workload of the agent clearer and its impact more measurable.



Balancing innovation and risk

The first step of the journey towards responsible AI is to have the discussions at a strategic level of what your organisational ambitions are, and what the risk appetite is.

Being able to articulate, as an organisation, if you want to be a first mover, a close follower, or a late adopter is an important first step – and will also be a guide in determining your risk appetite.

Defining a risk appetite can be an uncomfortable topic, but as financial institutions, whether credit or insurance, your businesses are built on evaluating and taking on acceptable risks. Having those discussions about AI at a broad level will prevent you needing to start the discussion of risk from scratch every time you pick up a new use-case.

A risk appetite cannot be non-existent, so discussing what types of mistakes you can tolerate and which ones you cannot, while maturing within the AI space, enables you to go into the development and tackle innovation with open eyes and clear guidelines.

When we see organisations fail at getting past the first couple of basic generative AI use-cases,
it is rarely down to talent gaps. It is often because there is not enough clarity on the risks; nor is there governance strong enough to make the organisation feel they fully understand the risks they are taking on.

Our belief is that effective and efficient governance is an enabler for innovation, not a force that works against it.

> “Good governance is a powerful enabler for innovation.”
>
> Susanne Gildberg

2

AI governance and risk management

This chapter details how to structure responsible oversight and manage organisational risk.

a.

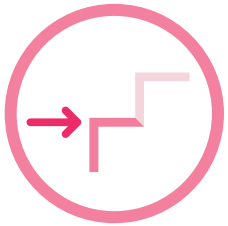
Clarify the playing field

- Establish core principles (e.g., fairness, transparency, accountability). How do we ensure that it is followed (controls)? Joint effort – led from the top.
- Set boundaries: Define clearly what is in scope and what is out of bounds?

b.

Address cross-cutting risks

- Workforce strategy: build diverse, multidisciplinary teams with clear roles.
- Joint taskforces: early-stage collaboration between tech, legal, compliance, and business functions.
- Modular governance: break risks into reusable components to future-proof governance frameworks.



c.

Early-stage interventions

- Introduce triage checkpoints before full deployment—AI should be reviewed like other critical investments.

a.

Clarify the playing field

As organisations adopt more sophisticated AI in more parts of the business, the potential risks demand careful oversight. It is, again, important to note that to allow innovation to happen, it is important to have the governance in place to evaluate and control risks, and that governance is an enabler of innovation when done right.

The first step in responsible AI governance is to establish core principles such as fairness, transparency, and accountability; and also to describe what these mean to you as an organisation, so it becomes clear and actionable.

We will also need to determine what controls will be in place to ensure the organisation is acting in accordance with these principles. Statements are easy to make, but hard to relate to day-to-day implementation and use of AI. Across the organisation, senior leadership needs to embed these values into culture, processes, and decision-making.

As part of this process it is also important to determine if there are any uses of AI or types of AI that are out of bounds. It could, for example, be that using AI to generate targeted marketing for individuals is out of bounds for the organisation. This could be because it is viewed to be risky, or it could be a temporary hiatus until the AI maturity level has risen enough that we trust it is being done correctly.

Defining certain uses as out of bounds helps to focus the innovation on use-cases that will be greenlit once completed. As limiting as it may seem when descoping AI use-cases, nothing is worse for innovation than spending time, resources, and energy on developing AI only to have all the work and motivation be discarded.

> **“It is essential that risk appetite and risks are defined and discussed at strategic level/C-level, so we do not have to start from scratch with each use-case. By clearly outlining the boundaries for acceptable risk, we enable informed decision-making and avoid reinventing the risk assessment for every single solution.”**

Ida Bach, Head of GenAI at AL Bank

b.

Address cross-cutting risks

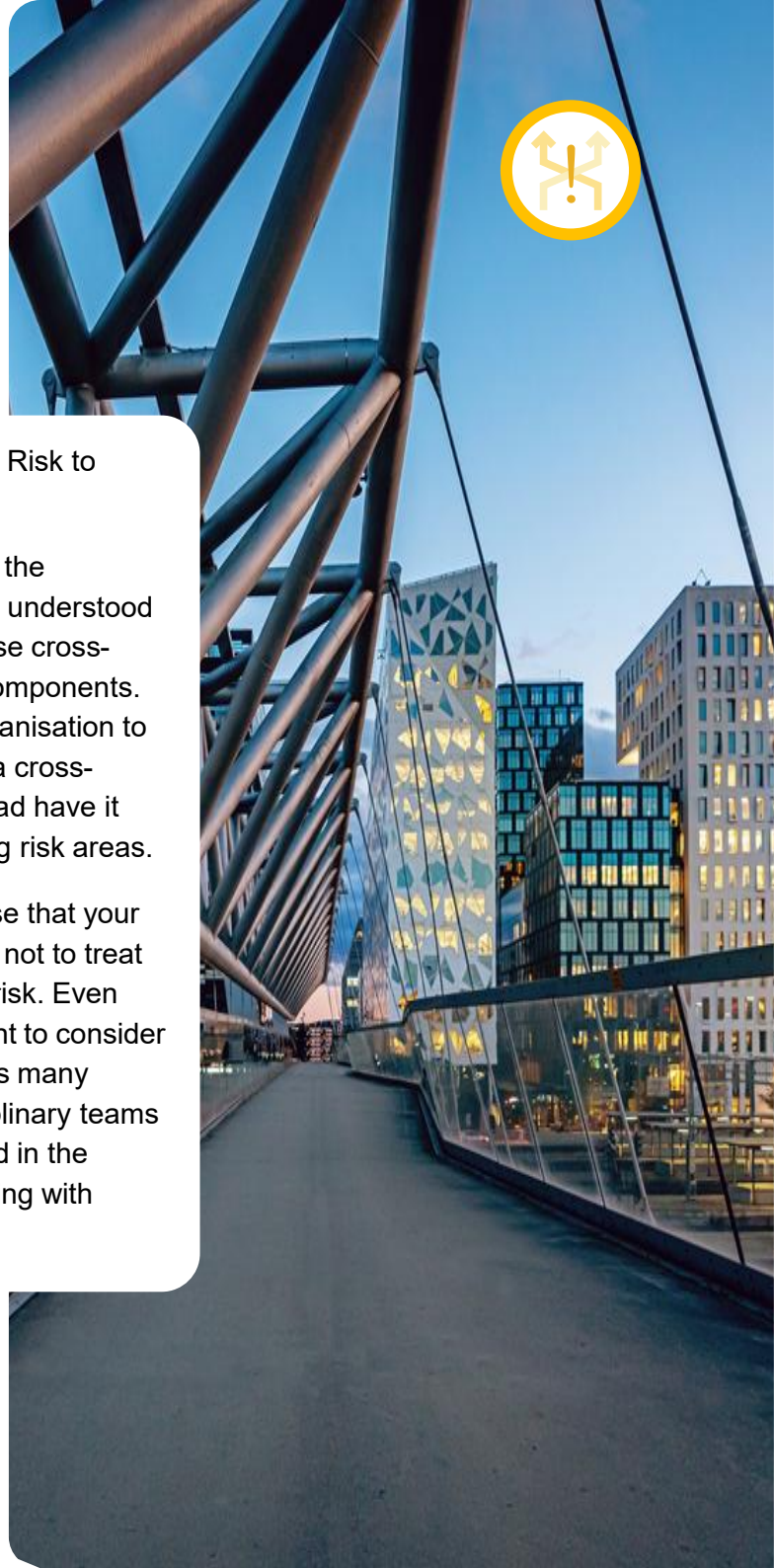
With the advent of AI, the risk taxonomy must be revisited. The diagram shows how AI introduces several new risks across multiple areas. Some of these are entirely new risks; some are new twists on existing risks.

One way to approach this, from the perspective of the risk taxonomy, is to introduce AI as a cross-cutting risk. We have previously seen organisations take this approach with other types of risks that cut across multiple areas, such as ESG risk. The benefits of this approach is that a multidisciplinary team or committee can be created to address new risk types, which demands expertise from many different areas to consider. These people can then learn from each other, and ensure that nothing in-between is missed.

A concrete example of this is that for AI there is a risk of bias within these highly complex models. Here, Model Risk should step in and investigate the model using MLOps and metrics that can unmask some of the inherent biases in the model. This, however, only works if the Model Risk team knows what to look for. Here, the interplay with different capabilities comes into play. Compliance is trained in thinking in this way, and has the ability to pose specific questions relating to bias, such as “Can the transcription model understand all regional accents?”, which can then be tested by model risk. Here we need Compliance to be wary of differential treatment in all its aspects, and not just the usual bias-checks within credit models for loans, premium models for insurance, and Model Risk to perform the check.

In time, it is likely that the organisation will have understood and broken down these cross-cutting risks into its components. This will allow the organisation to no longer treat AI as a cross-cutting risk, and instead have it integrated into existing risk areas.

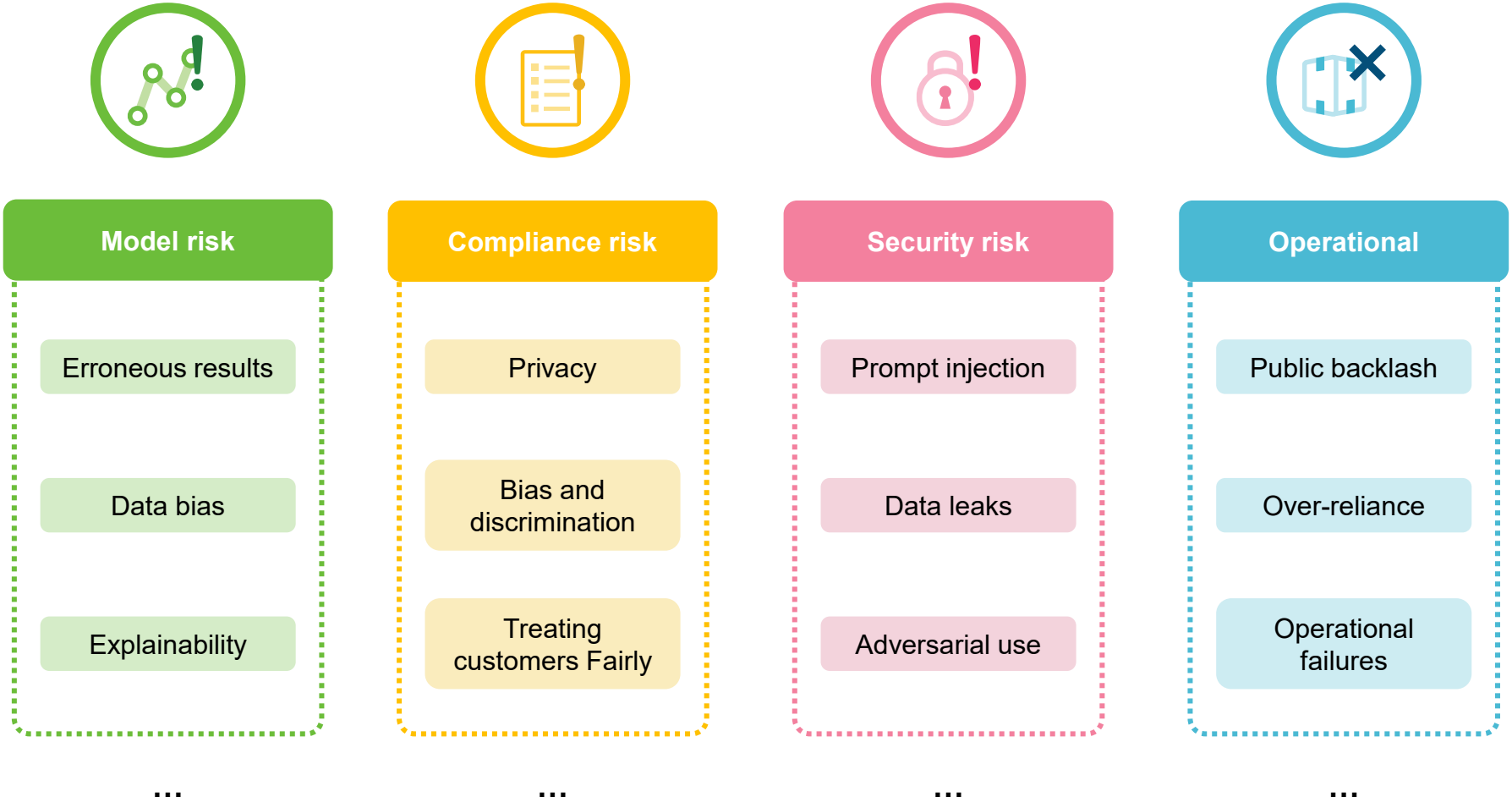
It may also be the case that your organisation chooses not to treat AI as a cross-cutting risk. Even then, it is still important to consider that AI reaches across many areas, and multidisciplinary teams are highly encouraged in the beginning when working with AI risks.



“We usually say that we are not afraid of technology, so we do not go out with a message to companies telling them not to use artificial intelligence because it is dangerous. We simply say that you need to be aware of the risks involved when you choose to use it.”

Camilla Neuenschwander, DFSA

Generative AI is a cross-cutting risk that impacts multiple areas



C.

Early-stage interventions

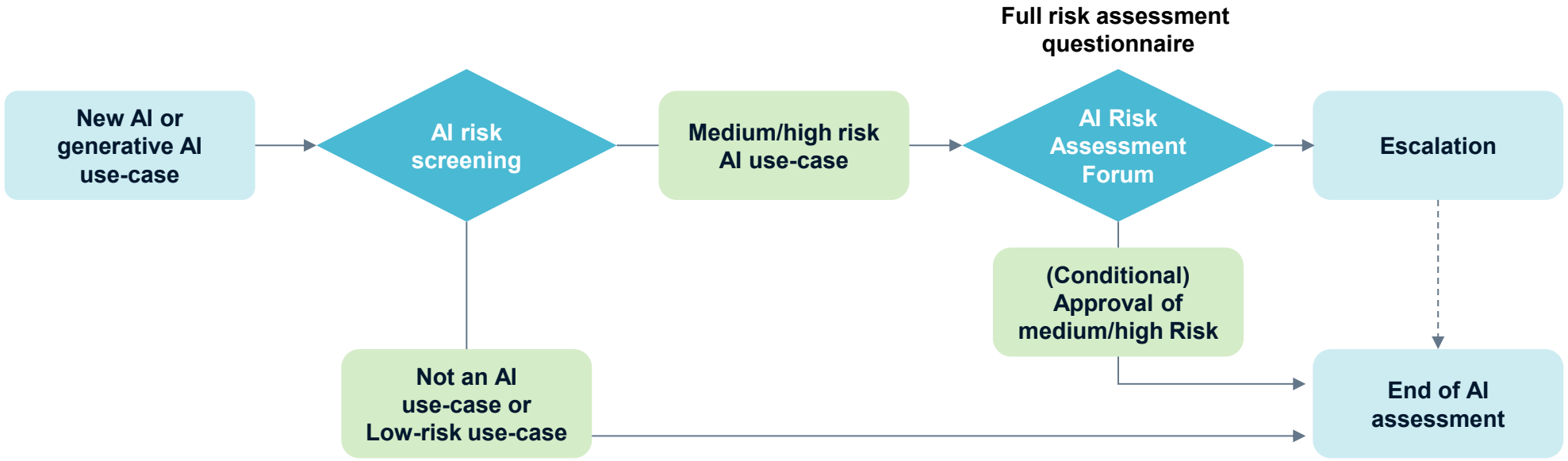
A change recommended by almost all experts we have interviewed, is that it is crucial to conduct early screening with a multidisciplinary set of experts from Risk, Compliance and other defense functions. This is because AI is such a rapidly evolving field, both within the technology, the regulations, the possibilities, and likely the organizations own policies that it is hard to keep up with the space.

As such, it is beneficial to gather a cross-functional team with each of their respective areas covered and consider briefly each new (gen)AI use-case before it is built, to ensure that the business owners and tech people side are fully understanding what concerns and risks the solution may have, and for everyone in this AI forum to stay up to date on the latest developments in the organisation.

The business, now understanding the potential risks, can then work on building AI in such a way that the risks are mitigated, and if that is not possible stop the development before time is wasted.

Example.

AI governance that can align expectations and AI risk understanding



Screening questionnaire:

- Type of AI
- Potential consequence to people
- Use of data
- Level of oversight

3

Technology – How to build AI responsibly

This chapter dives into the infrastructure, data, and technical guardrails needed to implement responsible AI.

a.

Creating trust in AI systems

- Accept the black box (to a degree): Transparency is important, but full interpretability may not always be possible.
- Human in the loop vs human-at-the-top (A2A).

b.

Rethinking data governance

- GenAI changes the game: Unstructured text is now a first-class asset.
- Access control and dataset classification: Who can use what data – and for what purpose?
- Versioning and drift: How to manage changes in data over time and resolve conflicting truths across documents.

c.

Control and success metrics

- Evolving metrics: Precision and recall are not enough – AI can now help evaluate itself. (Champion AI to evaluate other use cases)
- User acceptance: Ultimately, users behave like rational agents – choosing the fastest, easiest tool. Governance must account for actual usage patterns, not ideal ones.

a.

Creating trust in AI systems

When embarking on your AI journey, it is critical to understand that this is no governance exercise alone. The same factors that make AI so powerful and flexible also make it difficult to interpret and trust.

To fully utilize AI, it is important that we learn to accept the black box (within reason): Transparency is important, but full interpretability will not be possible.

This requires a paradigm shift in how we interact with and put our trust in AI systems. Fortunately, there is a technological toolbox available to support us to trust our AI systems, even though it is quite different from the way we are used to treating models and systems.

It is also important to measure AI systems against what they replace – be it individual humans or less advanced systems. When the accuracy or performance of the AI models becomes demonstrably better, then switching to a less interpretable model is often the right thing to do. But it is much harder to switch to AI purely for efficiency gains, if we cannot argue that the performance is at least not worse. This, again, is why it is so important to consider benefits beyond efficiency. This new paradigm can often feel like giving up some control, especially if AI is not well understood in the organisation. One of the ways that currently can generate trust is by placing a human-in-the-loop.

The next wave of AI we are seeing emerge requires yet another paradigm shift: agents collaborating with each other. To allow ourselves to gain the benefits of this wave, we often must place a human-on-top, instead of in-the-loop. This is where a human must be equipped with tools giving insight into the performance of the models at a broader level, to intervene and override when necessary.

The human-on-top must be sufficiently in control, and it must also be demonstrated to the rest of the organisation that they are indeed in control. This can only be achieved through robust governance, which has adapted to the AI future, and state-of -the-art techniques that allow AI to oversee AI and report to humans.

> **“If you truly want an agent factory, then governance and risk management need to run in a super streamlined way beneath the development layer, and you must have human oversight at the top, constantly monitoring what’s happening.”**

Michael Munck, CEO at 2021.AI

b.

Rethinking data governance

Data is what gives AI information about how the world works. However, the way we think about data could benefit from being updated, so it can best suit AI purposes.

First and foremost, AI needs not only what is classically called data, which is the stuff in our SQL databases and excel sheets. Generative AI requires that we have a good grasp of what have previously been documents written by humans for other humans, such as SOPs, policies, instructions, guides, and so on. Generative AI needs these to understand how the job should be performed correctly, and to help humans navigate the complex process landscape.

It is likely that during this process of categorising and cleaning up this category of data, you will discover that it has errors, does not follow best practice, and that different documents will give conflicting instructions. In this process:

You will likely realise that it is easy to use AI to identify and correct conflicting instructions and clean up the document

You must make sure that only the latest versions are included in the AI training data, as the AI will get confused about which practice is the valid way to do things and which are outdated

Accept that this is a new opportunity to correct the mistakes and misinterpretations of the past, and fix this for both humans and AI.

C.

Control and succes metrics

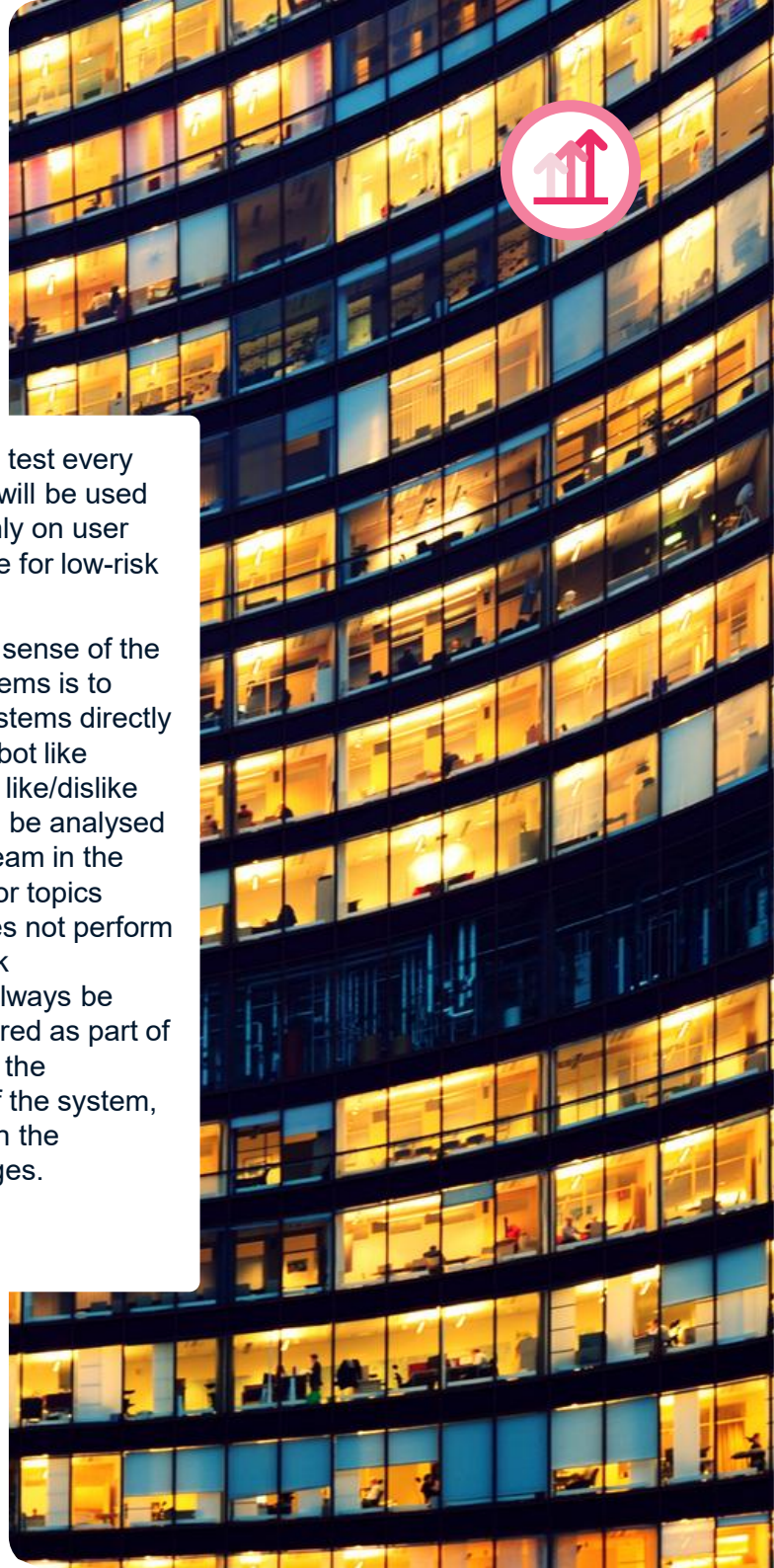
Traditional ML/AI systems often rely on large annotated datasets, to train or improve the performance of an algorithm through a dedicated loss function.

Most GenAI applications skip this training, as the foundational model knowledge typically suffices, and additionally avoid overwriting valuable parts of the neural network. Instead, they inject knowledge into the model through methods such as retrieval-augmented generation, where context relevant to the user prompt is injected into the query to the LLM. This leads to large savings during development of these applications but also comes at a cost. Developers often lose track of their accuracy metrics as they move into this new territory.

One way to approach this issue is to set up a large set of tests, such as question-answer pairs, for how the model should respond to certain questions. While time-consuming, this method is feasible and impactful. One challenge is that multiple phrasings of the same response may have the same semantic meaning. To address this, LLMs can be used to cut through the different phrasings to evaluate if the semantic meaning is the same. As a result, we now have a system of LLMs evaluating LLMs. During such evaluations it is prudent to use state-of-the-art or challenger models for evaluating the test scenarios.

Another approach towards evaluation of generative AI tools is to look towards user testing and get more qualitative feedback on metrics such as user trust, transparency, and explainability. Such testing must always be carried out with the intended user groups and SMEs to get an accurate picture. While this method is generally considered less resource-intensive, it is also considered less accurate as it is typically impossible to test every scenario that the tool will be used in. As such, relying only on user testing should be done for low-risk use cases.

A third option to get a sense of the accuracy of such systems is to integrate feedback systems directly into the tools. In chat-bot like systems this can be a like/dislike button which can then be analysed by the development team in the logs to identify areas or topics where the system does not perform well in. Such feedback mechanisms should always be integrated and monitored as part of the MLOps, to ensure the continued accuracy of the system, and to avoid drift when the underlying data changes.



Bringing Ingenuity to Life.

About PA.

We believe in the power of ingenuity to build a positive human future.

As strategies, technologies, and innovation collide, we create opportunity from complexity.

Our diverse teams of experts combine innovative thinking and breakthrough technologies to progress further, faster. Our clients adapt and transform, and together we achieve enduring results.

We are about 4,000 strategists, innovators, designers, consultants, digital experts, scientists, engineers, and technologists. And we have deep expertise in consumer and manufacturing, defence and security, energy and utilities, financial services, government and public services, health and life sciences, and transport.

Our teams operate globally from offices across the UK, Ireland, US, Nordics, and Netherlands.

PA. Bringing Ingenuity to Life.

Discover more at paconsulting.com and connect with PA on LinkedIn and X

Copenhagen Office

Portland Towers
Göteborg Plads 1
DK-2150
Nordhavn
+45 39 25 5000