



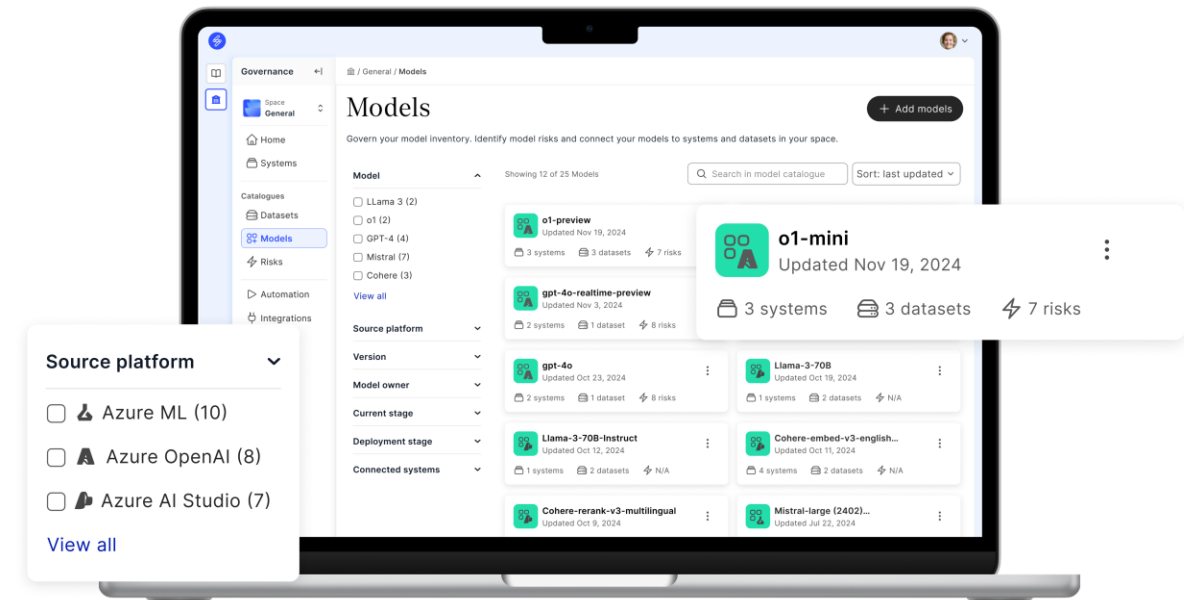
AI Risk Awareness Session

Saidot x Scottish Government
31 January, 2025

Saidot in brief

- Forerunners of responsible AI, founded in 2018
- Based in Helsinki, operating globally
- Diverse and multinational team of 20+ experts
- Mission to enable responsible AI innovations and governance efficiency
- 190+ organisations use Saidot Library as their AI Governance knowledge base
- Clients from different industries, including banking, healthcare, governmental, telecom, media and others
- Saidot's experts have been contributing to the development of different standards and regulations in EU, UK and US
- Recently announced collaboration with Microsoft

Unlocking AI's promise responsibly.





Introductions

We're a team of 20+ mission-driven AI experts with backgrounds in AI, ML engineering, policy, law, product, design, and business.

Your trainers today:



Iiris Lahti

Head of Services & Customer Success, experienced data & AI transformation leader



Edla Aittokallio

AI Governance Specialist, AI risk and governance expert, and manager of our AI Risk Library.

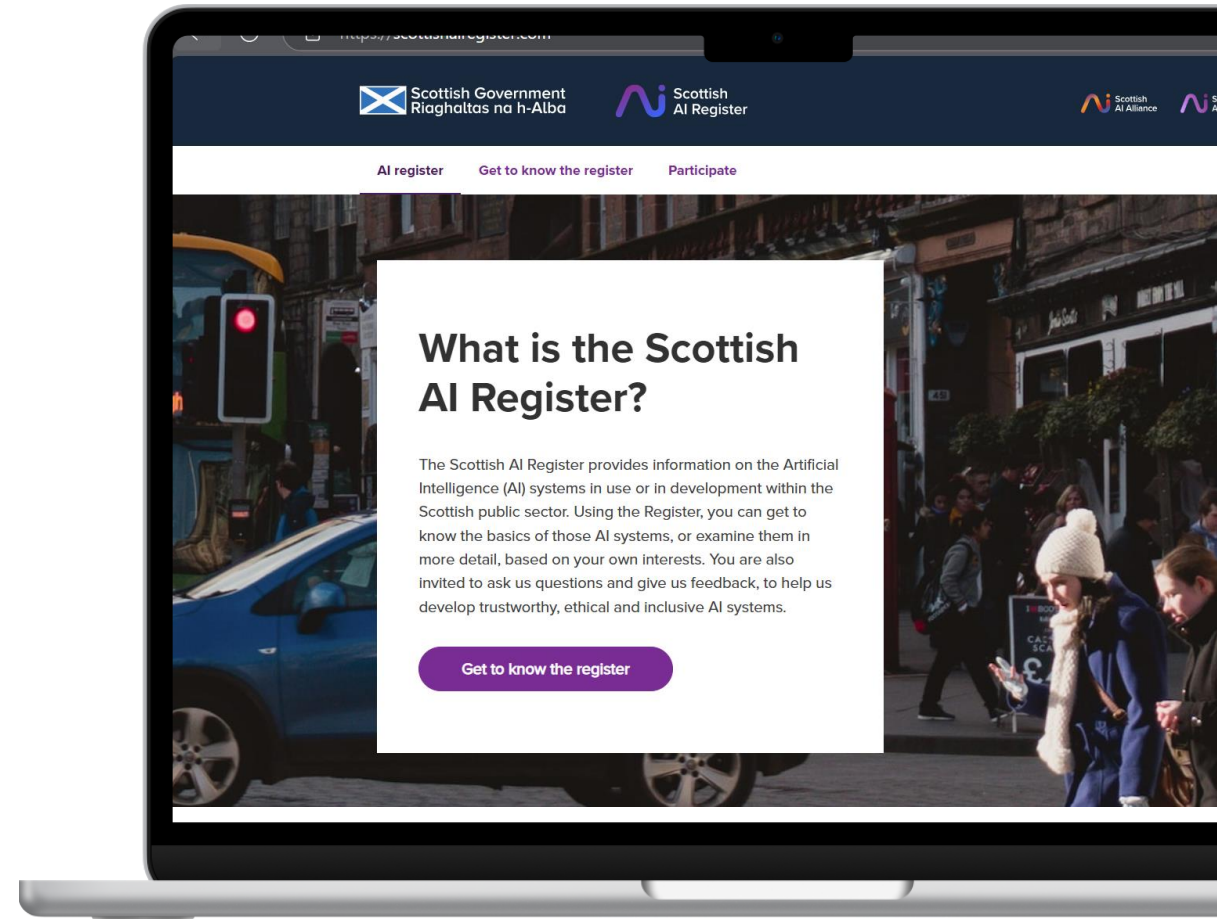
Our journey together

To enable responsible use of AI in Scotland

- Scottish AI register
- Saidot Library as a knowledge base for AI policies, risks, models and evaluations

💡 What's next

- Internal AI inventory for better possibility for risk and compliance management





Part I: AI Risk Landscape

30 January, 2025

Aims of the session

- The training is aimed at anyone interested in knowing more about the AI risk landscape.
- It is designed to provide a general overview of the AI risk landscape and its broader implications.
- It provides a comprehensive insight into the risks associated with AI, emphasises the importance of addressing these risks in today's world, and introduces the key steps and roles in effective AI risk management.

A light purple rounded square card with a white border, containing the text 'AI risk landscape' and '30 January, 2025'.

AI risk landscape

30 January, 2025

A light blue rounded square card with a white border, containing the text 'AI risk management' and '30 January, 2025'.

AI risk management

30 January, 2025



Challenge

How can we drive AI innovation and use in public services while managing the risks and ensuring compliance?

The recent evolution of AI

The AI value chain is becoming more complex

- Both the AI value chain and the world are getting more complex.
- The evolution of AI has been rapid during the past years from LLM's to AI components, products and agents.

In-house
algorithms
and Machine
Learning

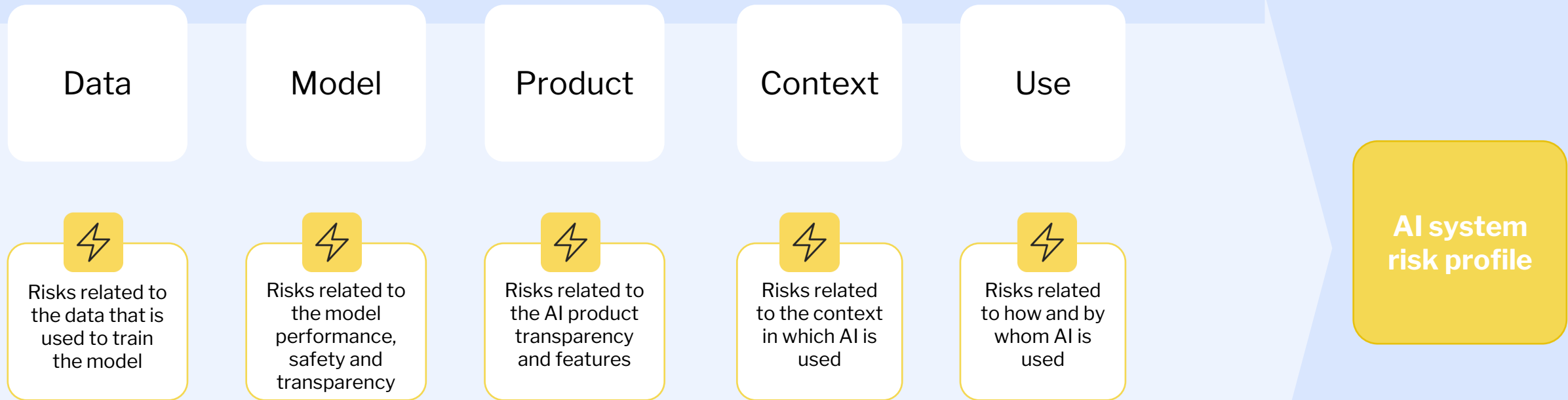
The burst of
LLM's

AI platforms
and tailored
AI UI

AI products
and
components

AI agents

AI value chain and the complexity of risk management



Why everyone should be aware of AI risks?

Risk management ensures safe and responsible AI adoption in a constantly evolving risk landscape.

Universal impact

- **AI is transforming work** across industries.
- AI risk management is **relevant to everyone**, regardless of industry or role, because its influence spans society.
- Effective AI governance is **still evolving**.

Consequences

- AI affects **critical life decisions**, including hiring, healthcare, and credit approval.
- **Malicious actors** increasingly use AI, amplifying security threats.
- **Serious risks** like bias and system failures continue to emerge.

Accountability

- **Fragmented AI value chains** spread accountability between AI system **developers, deployers, and users**.
- **Multiple organisational roles** share accountability within an organisation.



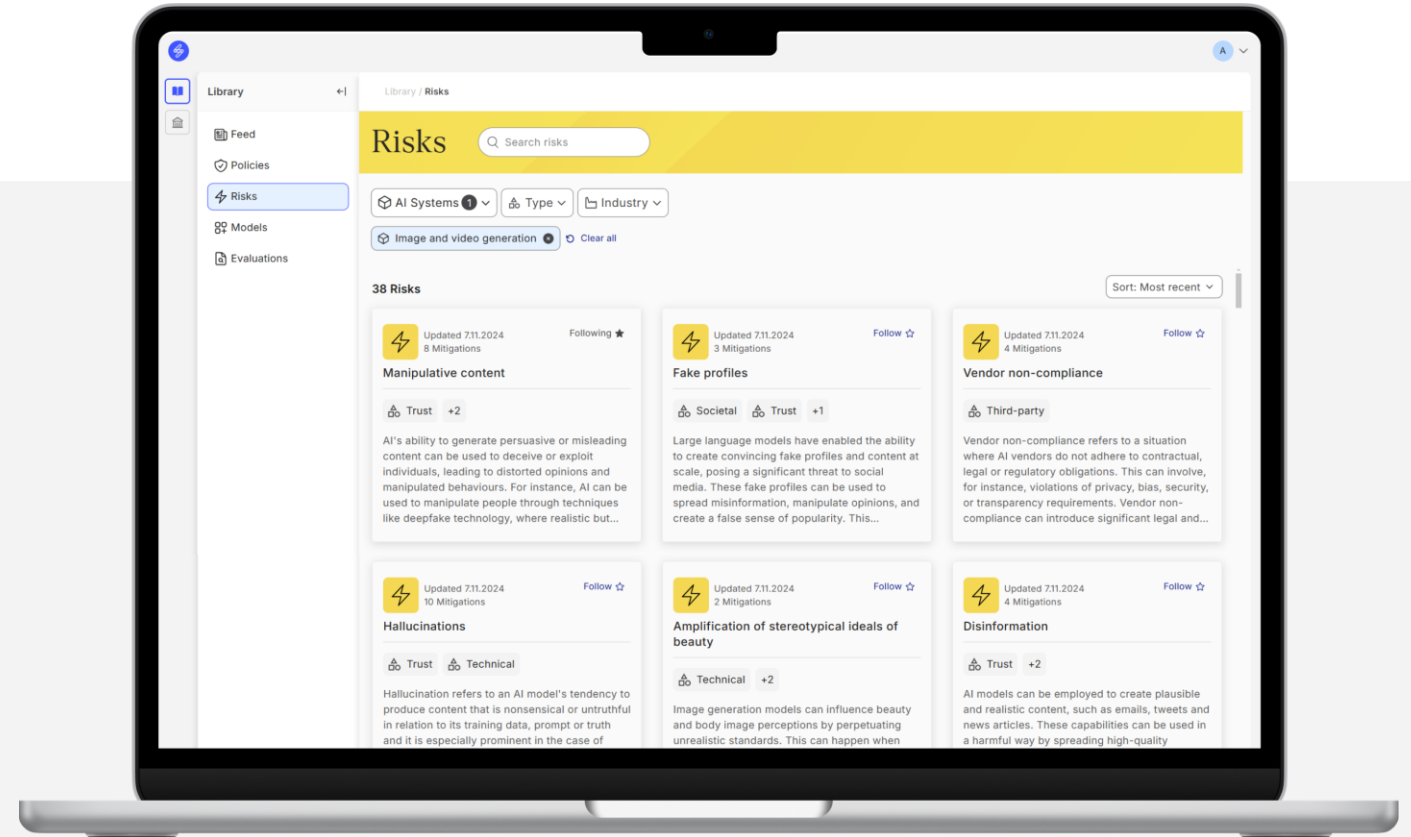
Situation

Over the past years, new AI risks have surfaced and there is wider awareness of the potential harms and uncertainties.

Understanding AI risks

Common AI risks

- ⚡ Hallucinations
- ⚡ Harmful, biased and stereotypical content
- ⚡ Safety and performance issues
- ⚡ Third-party vendor reliability
- ⚡ Reputational harm
- ⚡ Overreliance
- ⚡ Cybersecurity threats
- ⚡ Privacy infringements
- ⚡ Contractual confusion
- ⚡ Copyright issues



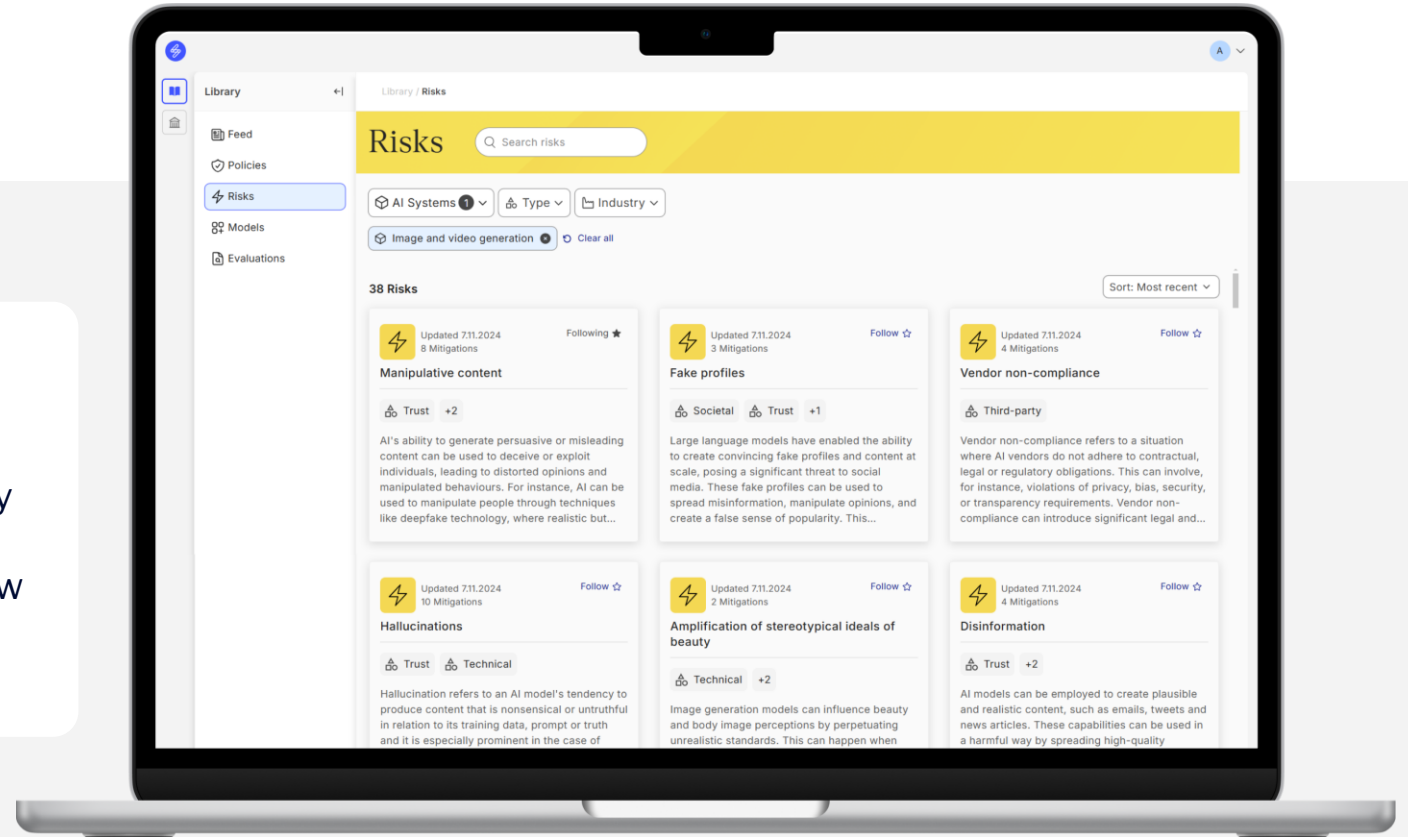
Source: Saidot AI Risk Library

Understanding AI risks

Saidot AI Risk Library



The Saidot Risk Library helps organisations understand and manage 150+ AI-related risks. It provides detailed descriptions of each risk, how they manifest, and their impact, along with practical mitigation strategies. Users can filter, sort, and follow specific risks for updates.



Source: Saidot AI Risk Library

Risk types

Classifying risks by type helps organisations understand the nature of the potential harm and facilitates targeted mitigation measures.

 **Risk type:** to the category of potential negative outcomes or threats associated with the use of the AI system.

Risk type

Legal

Cybersecurity

Environment

Technical

Trust

Fundamental rights

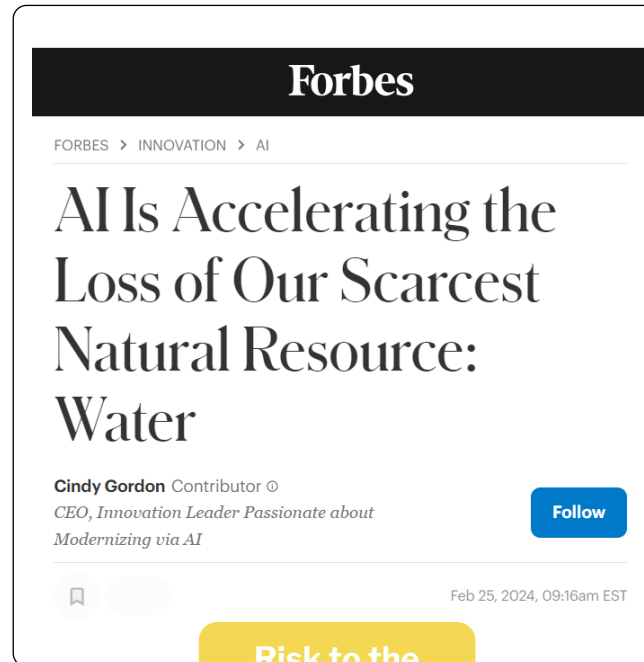
Privacy and data protection

Societal

Third-party

Business

Health and safety



Risk to the environment

Risk sources

AI risk source helps organisations understand the factors contributing to and creating a risk

 **Risk source:** an element which, alone or in combination, has the potential to give rise to risk.

Risk source

Data

Model

Product

Use

Context

Regulation

Other



Risk from use



Note

AI risks are real, but with the right strategies, they can be managed and mitigated.

Misuse of deepfakes

 Risk type Trust

Deepfake is AI-generated or manipulated image, audio, or video content that resembles existing persons, objects, places, or other entities or events and would falsely appear to a person to be authentic or truthful. Deepfake technology can be harmful if it is misused.

Risk mitigation examples

- AI literacy training
- Deepfake detection
- Deepfake policies and guidelines
- Labels (E.g. disclaimers, signatures and watermarks)
- Transparency
- Safety and content filters

Source: Saidot AI Risk Library



The image shows a screenshot of a BBC News article. At the top, the BBC logo is visible. Below it, a navigation bar includes links for Home, News, Sport, Business, Innovation, Culture, Arts, Travel, Earth, Video, and Live. The main headline reads "AI Brad Pitt dupes French woman out of €830,000". Below the headline, it says "22 hours ago" and "Share Save". The author is listed as "Laura Gozzi, BBC News". The main image is a portrait of Brad Pitt smiling, with a "Getty Images" watermark in the bottom right corner. Below the image, there is a quote: "It's awful that scammers take advantage of fans' strong connection with celebrities," a spokesperson for Brad Pitt said. At the bottom of the screenshot, the URL "https://www.bbc.com/news/articles/ckgnz8rw1xgo" is displayed.

Harmful or toxic content

Risk type

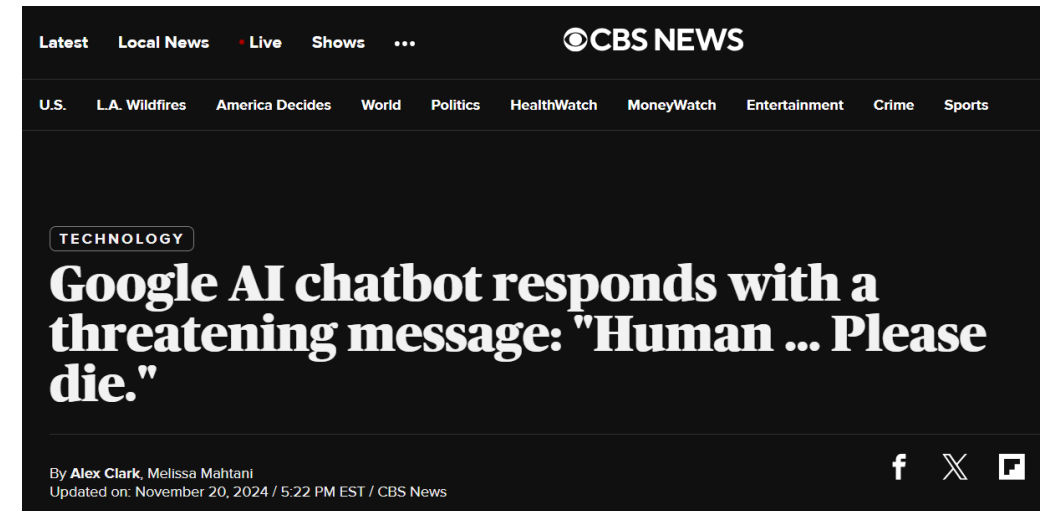
Trust

Fundamental rights

AI systems may generate or can be used to generate **harmful, offensive, inappropriate, "explicit", or spurious content**. This can include, for instance, derogatory comments targeting protected attributes, rude or offensive language, violent scenarios, and content promoting danger or gore.

Risk mitigation examples

- Content policy
- Hardcoded responses
- Harmful content detection
- Feedback collection of AI-generated outputs
- Safety and content filters
- Input or prompt filtering



<https://www.cbsnews.com/news/google-ai-chatbot-threatening-message-human-please-die/>

Bias amplification

Risk type

Technical


Fundamental rights

Bias in the context of AI refers to unfair or discriminatory outcomes in model outputs, stemming from a systematic error in decision-making processes. It occurs when certain groups are disproportionately favoured or disadvantaged due to inaccuracies or imbalances in how the AI system processes information.

Risk mitigation examples

- Bias mitigation and detection plan
- Model fairness assessment
- High-quality and representative training data
- Development team diversity
- Data gathering and labelling
- Impact assessment



AMNESTY INTERNATIONAL  ENGLISH

WHO WE ARE WHAT WE DO

< RESEARCH  

November 12, 2024
Index Number: EUR 18/8709/2024

Denmark: Coded Injustice: Surveillance and Discrimination in Denmark's automated welfare state

<https://www.amnesty.org/en/documents/eur18/8709/2024/en/>

Hallucinations

 Risk type

Technical

Trust

Hallucination refers to an AI model's tendency to produce content that is nonsensical or untruthful in relation to its training data, prompt or truth and it is especially prominent in the case of generative AI models. Hallucinations can emerge in different ways, such as as false information and irrelevant responses.

Risk mitigation examples

- Testing and validation
- Retrieval augmented information
- High-quality and representative training data
- Feedback collection of AI-generated outputs
- Use case restriction
- Input or prompt filtering

WIRED

SECURITY POLITICS GEAR THE BIG STORY BUSINESS SCIENCE CULTURE IDEAS MERCH

BENJ EDWARDS, ADS TECHNICA BUSINESS OCT 30, 2024 9:00 AM

OpenAI's Transcription Tool Hallucinates. Hospitals Are Using It Anyway

In health care settings, it's important to be precise. That's why the widespread use of OpenAI's Whisper transcription tool among medical workers has experts alarmed.



PHOTOGRAPH: MELJAN ZIVKOVIC/GETTY IMAGES

<https://www.wired.com/story/hospitals-ai-transcription-tools-hallucination/>

Concentration of power

 Risk type **Societal**

The development of large-scale AI models requires significant computational resources, which may result in a **concentration of power** in the AI sector, potentially leading to monopolies or oligopolies among larger technology companies. This concentration of power can give these companies the economic power to impact political decision-making on how AI is developed and used in society.

Risk mitigation examples

- International collaboration
- AI literacy training
- Use of open-source models
- Transparency

Mark Zuckerberg's Geopolitical Free Speech Gambit

The Meta CEO's U.S. political changes will have complicated global consequences.

By [Rishi Iyengar](#), a reporter at *Foreign Policy*.



<https://foreignpolicy.com/2025/01/10/mark-zuckerberg-meta-fact-check-hate-speech-trump/>

Contractual and confidentiality infringements

 Risk type

Legal

The use of AI and improper data handling by an organisation may result in [violations of contractual agreements](#), compromise [data privacy](#), and breach [confidentiality obligations](#). This can happen, for instance, by sharing internal code or confidential business information with an AI tool.

Risk mitigation examples

- Training programs
- Access limitation
- Data handling and privacy policies
- Acceptable use policy
- Data Protection Impact Assessment
- Data use instructions

Forbes

BREAKING

Samsung Bans ChatGPT Among Employees After Sensitive Code Leak

Siladitya Ray Forbes Staff

Siladitya Ray is a New Delhi-based Forbes news team reporter.

Follow



May 2, 2023, 07:17am EDT

Updated May 2, 2023, 07:31am EDT



TOPLINE Samsung Electronics has banned the use of ChatGPT and other AI-powered chatbots by its employees, Bloomberg [reported](#), becoming the latest company to crack down on the workplace use of AI services amid concerns about sensitive internal information being leaked on such platforms.



<https://www.forbes.com/sites/siladityaray/2023/05/02/samsung-bans-chatgpt-and-other-chatbots-for-employees-after-sensitive-code-leak/>

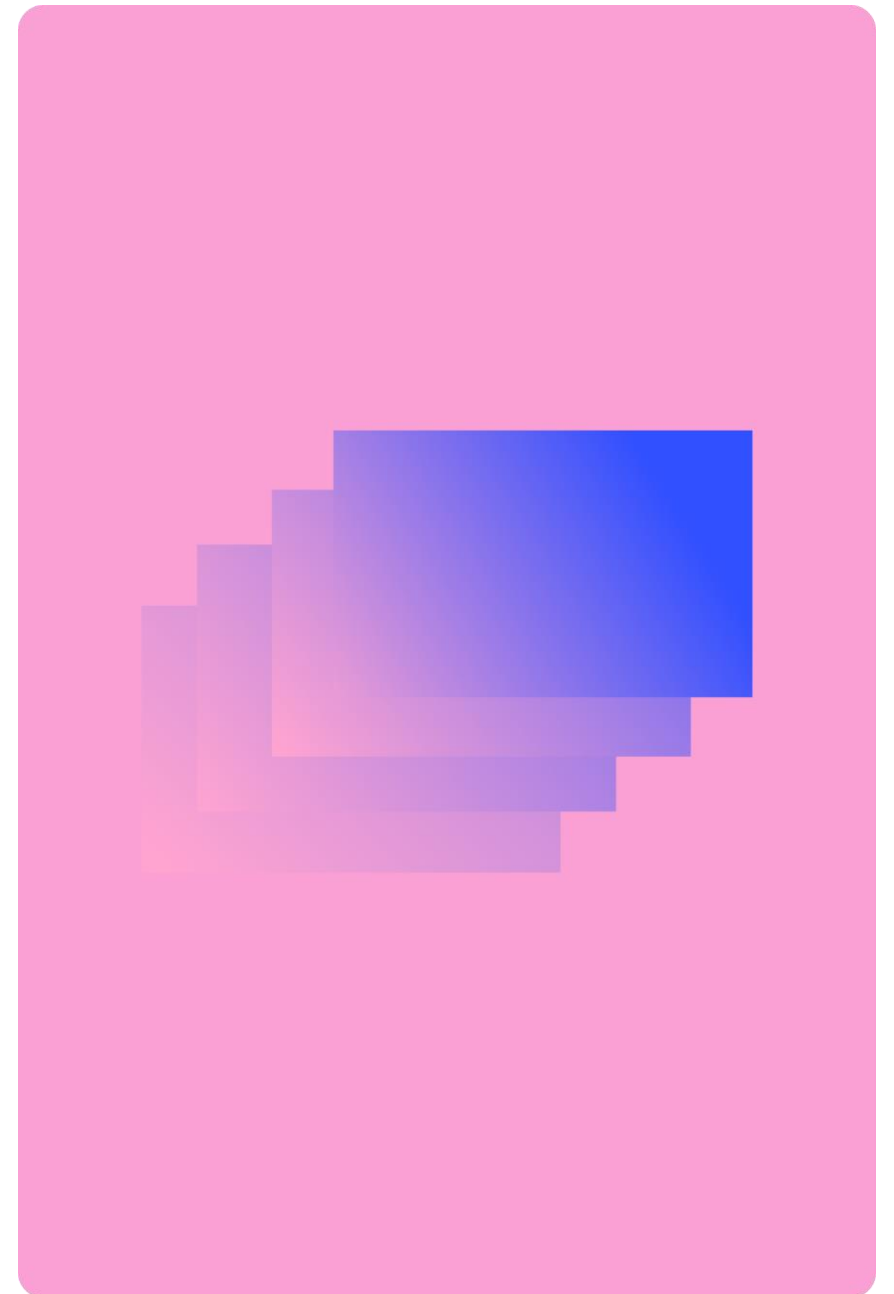
Examples of risk mitigations organisations can take

General:

- AI literacy trainings and employee upskilling
- AI policy and clear working models and guidelines for responsible use of AI
- Aligning AI use with strategic targets, industry standards and risk appetite

AI system specific:

- Human oversight
- Transparency and explainability
- Contractual use case restriction
- Data Protection Impact Assessment
- Performance and safety evaluations
- Bias and harmful content detection
- Model evaluation and choice
- Cybersecurity audits and threat detection



How to use AI responsibly and consider the risks as an individual

- Understand your **responsibility** as AI user and consumer
- Understand the **AI policy** and guidelines Scottish Government has committed to
- **Be aware** of the models and tools you use: Purpose, impact, benefits, risks and what actions you can take
- Improve your **AI literacy** to understand the risks and to be able to question the outcomes of the AI model and tools
- Require your service providers to **explain** their AI policy and guidelines and how the models have been trained
- **Be consistent** in the way you give permission to access your data or the data you give to AI tools when prompting





Question

Which other AI risks have you encountered when using or developing AI solutions?



How to effectively manage AI risks?

The importance of risk management

Risk management is vital for responsible AI development and regulatory compliance

Responsible AI

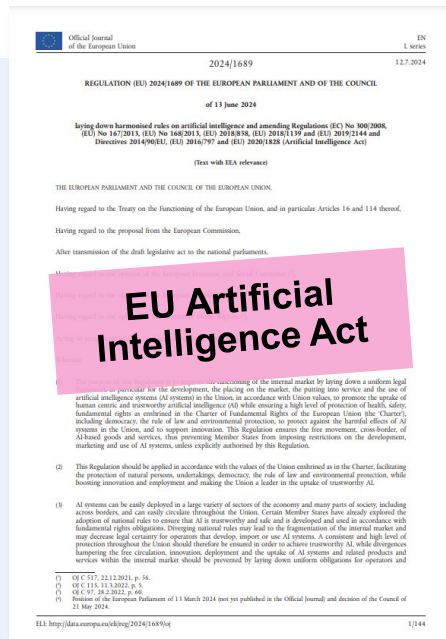
- **Best practice in responsible AI development** and an essential component of AI governance.
- Crucial to **prevent and reduce possible negative impacts** of AI on individuals, organisations, and society as a whole.
- Ensures that AI technologies are used **responsibly and safely**.
- **Builds trust** among users and stakeholders by showing a commitment to addressing potential issues related to AI.

Regulatory compliance

- Often **mandatory for legal and policy compliance**.
- **Positive developments** in the EU and UK in AI regulation and policy.
- **Risk management considerations** also often integrated.
- Ongoing advancements in AI regulation highlight the **growing recognition of risk management** as essential for responsible AI development and deployment.

AI risk management in legal and policy instruments

EU and Europe



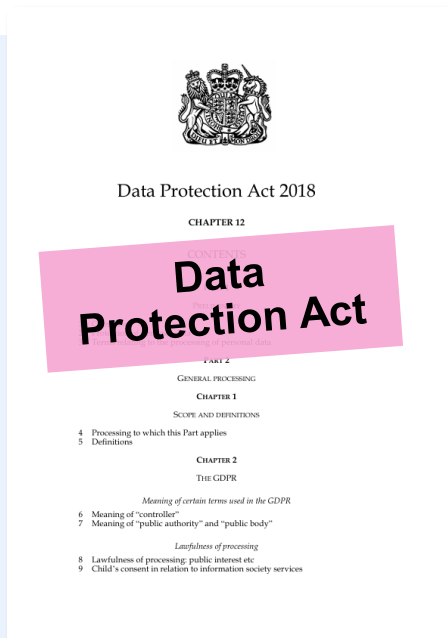
- Unified regulatory and legal framework for all sectors and types of AI.
- Risk-based approach.
- Unacceptable systems, high-risk systems, AI systems with transparency risk and general-purpose AI models.
- Risk management requirement for high-risk AI systems



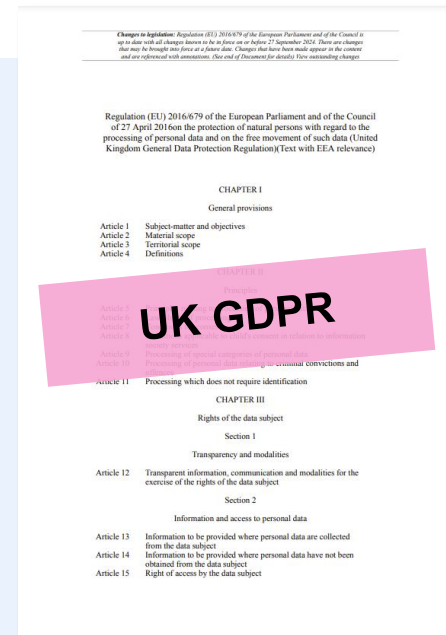
- Convention to ensure AI systems respect human dignity, rights, democracy, and the rule of law throughout their lifecycle
- Obligations related to ensuring effective remedies and the assessment and mitigation of risks and adverse impacts.

AI risk management in legal and policy instruments

UK



- Rules on how personal information is used by organisations, businesses or the government
- UK's implementation of the General Data Protection Regulation (GDPR).



- UK GDPR sits alongside an amended version of the DPA 2018.
- Same key principles, rights and obligations
- However, implications for the rules on transfers of personal data between the UK and the EEA.

AI risk management in legal and policy instruments

GOV.UK



Policy paper
**AI Opportunities Action Plan:
government response**

- **Policy paper** addressing the government's response to the AI Opportunities Action Plan
- Emphasis on cutting-edge, secure, and sustainable AI infrastructure
- The right regulatory regime that addresses AI risks and actively supports innovation

GOV.UK



Guidance
**Generative AI Framework for
HMG (HTML)**

- **Guidance** on the responsible, ethical, and effective use of generative AI within UK government organisations.
- Ten principles guide public sector use of generative AI to boost productivity while managing risks and ensuring legal and ethical compliance.

Guidance

Algorithmic Transparency Recording Standard

A standardised way of recording and sharing information about how the public sector uses algorithmic tools.

- **Guidance** on helping public sector organisations provide information on the algorithms they use and their purposes.
- Documentation on risks, mitigations, and impact assessments.

Model risk management principles for banks

Supervisory statement | SS1/23

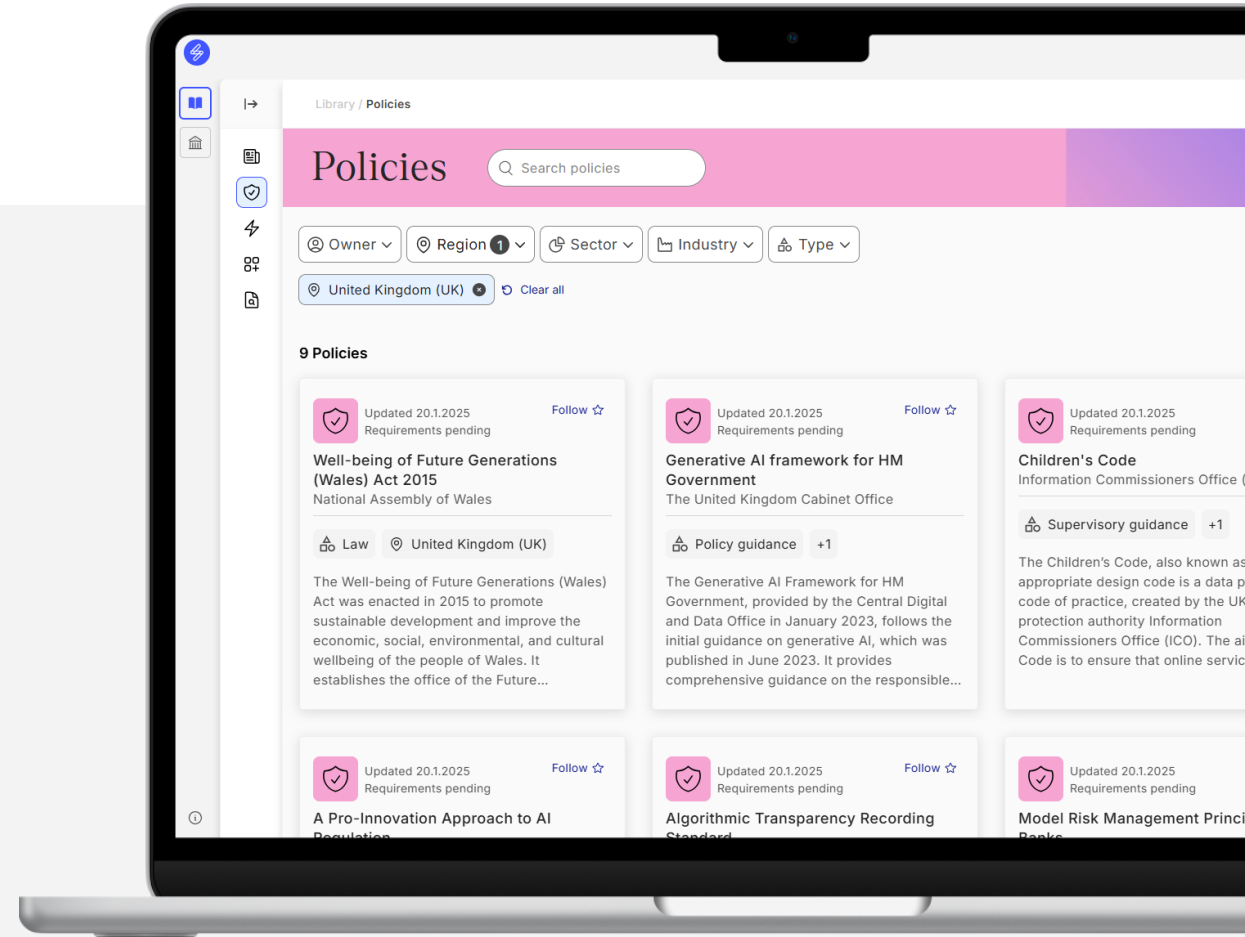
- **Supervisory statement** outlining the Prudential Regulation Authority's expectations for banks to manage model risk as a distinct discipline.
- Five key principles for implementing a robust model risk management framework across all model and risk types.

Understanding AI policies

Saidot AI Policy Library



Saidot Policy Library helps organisations and individuals to understand the variety of different AI, data and digital policies contained in the Library. The Policy Library contains an ever-growing collection of policies, policies ranging from laws to standards, frameworks, guidelines, and best practices.





Part II: AI Risk Management

30 January, 2025

Aims of the session

- This training is tailored for AI system owners, data scientists, legal professionals, compliance officers, and anyone closely involved with AI systems and their risk management processes. It is ideal for those who want to deepen their expertise and apply risk management practices effectively.
- The workshop explores the AI risk management process and methodology, providing a step-by-step guide on how to perform effective risk management using the Saidot platform.
- Participants will gain practical knowledge in identifying, evaluating, and treating AI risks in their specific contexts.

AI risk
landscape

30 January, 2025

**AI risk
management
workshop**

30 January, 2025

Saidot AI Governance Framework

AI policies

Direct and support the development and use of AI in line with business objectives and legal obligations.

AI lifecycle management process

Follow best-practice lifecycle model stage-specific requirements.

Third-party management

Manage AI partners and the use of third-party AI components effectively.

Data management

Align and connect with data management practices to ensure AI system data quality.

Risk management

Apply risk-based governance and robust AI risk management processes and practices.

Compliance management

Analyse and implement legal and policy requirements to ensure AI compliance.

Roles and responsibilities

Identify stakeholders involved in AI governance and assign responsibilities.

Tools and resources

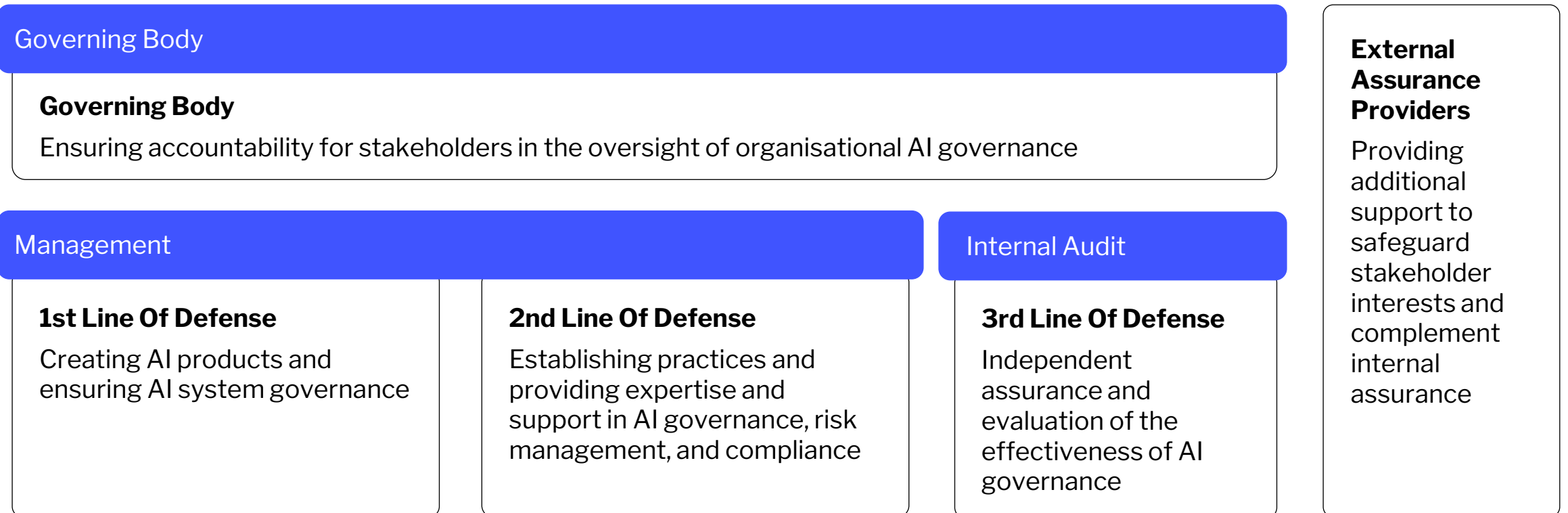
Leverage shared expertise, data, and tools for governance quality and efficiency.

Transparency and documentation

Ensure transparent communication with AI users and other AI stakeholders through documentation standard and AI inventory.

Assigning accountabilities for AI governance

Three Lines of Defense of AI Governance



Source: Saidot based on 'Three Lines of Defence' model by the Institute of Internal Auditors (IIA)

In practice: AI governance lifecycle stages

AI Governance tasks in different Lifecycle Stages

Initiation

- Register system
- Determine context
- Determine intended use
- Classify system according to risk and business impact level
- Determine roles and responsibilities

Design & Development

- Determine system components
- Assign and manage suppliers
- Perform data governance
- Determine applicable policies
- Perform risk management
- Evaluate safety and performance
- Implement policy-specific requirements and controls

Validation

- Create review
- Conduct review and report findings
- Create and assign finding remediation tasks
- Implement remediation tasks
- Approve the system

Deployment

- Configure and share transparency
- Setup human oversight and monitoring

Operation & Monitoring

- Monitor system risks and performance
- Process post-market feedback
- Implement remediation tasks

Re-evaluation

- Re-evaluate the system and report findings
- Create and assign finding remediation tasks
- Implement remediation tasks
- Re-approve the system

Retirement

- Archive system

Risk management in the AI lifecycle stages

Various tasks in AI lifecycle management centre around risk management.



Before risk management

Proper system documentation and risk level analysis sets you up for effective AI risk management.

What steps to take before risk management

1. Document system in your AI inventory

- Document your AI system and its relevant components (system type, model, function, capability...)

2. Analyse the Risk level of your AI system

- Risk classification:** the perceived overall risk level of an AI system, based on its specific use context, intended purpose, and the associated regulatory and business risks.
- Right-size governance:** AI governance actions and measures that are required, considering the AI system's risk level and risk profile, applicable policies and compliance requirements, and industry and company-specific documentation standards.



The screenshot shows the Saidot Governance interface for a system named 'Image Generation for Marketing'. The interface is divided into several sections:

- System Context:** Includes 'Organisation's Role' (with options for Provider, Deployer, and Other), 'Region' (with tags for Europe (EU) and United Kingdom (UK)), 'Industry' (with tag for Consumer products), 'Function' (with tag for Sales and marketing), and 'System Type' (with tag for Image and video generation).
- Capabilities:** Includes tags for Text to image and Image generation.
- Tags:** A field to add tags relevant to the system.
- Lifecycle:** A vertical timeline showing key events from May 26, 2024 (Initiation) to Nov 12, 2024 (Retirement). Key events include 'Approved for design and development', 'Design and development', 'Verification and validation', 'Approved for deployment', 'Deployment', 'Operation and monitoring', 'Re-evaluation', and 'Approved for retirement'. The 'Retirement' event is highlighted in blue.
- Team:** Shows 'All members' and '6 team members', with 'Lumi Kajaste' listed as the 'System owner'.

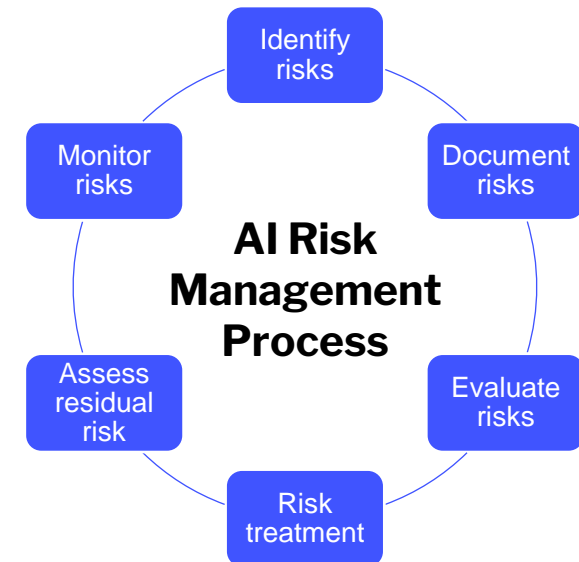
Risk management for AI systems

Identify, evaluate, and treat specific risks throughout the lifecycle of the AI technology

🧠 **AI risk management:** the identification, evaluation, and mitigation of specific risks of an AI system.

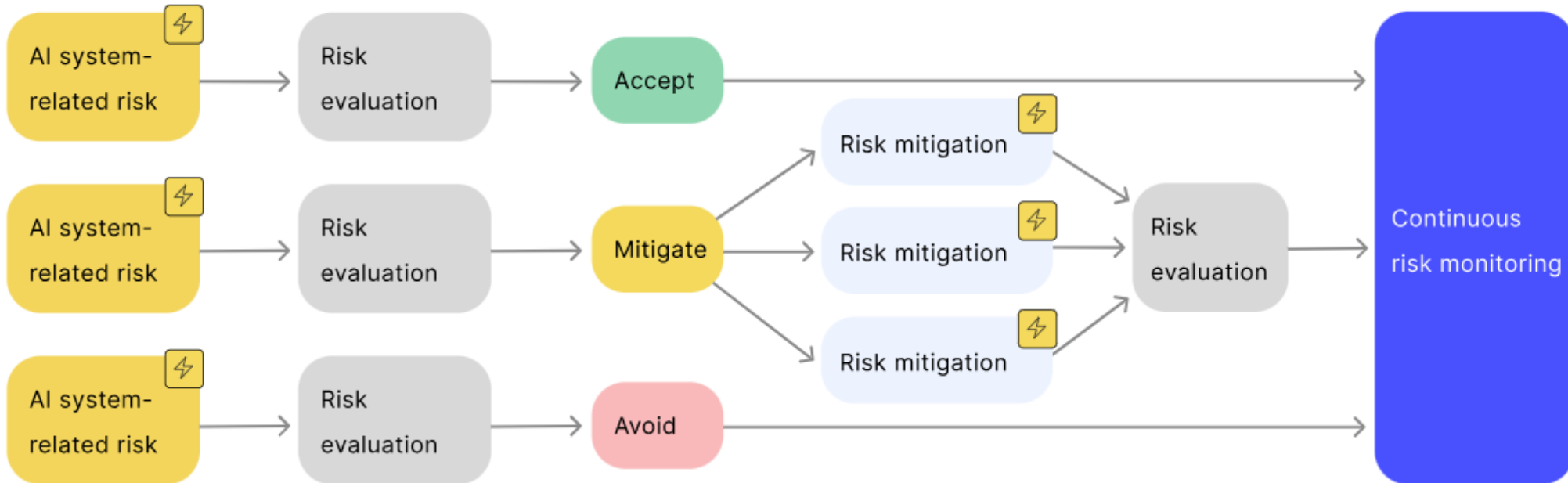
Why is risk management important?


- AI risk management is essential to **prevent and reduce possible negative impacts of AI** on individuals, organisations, and society as a whole.
- Risk management ensures that AI technologies are used **responsibly and safely**, in compliance with relevant regulatory instruments.
- Effective AI risk management **builds trust among users and stakeholders** by showing a commitment to addressing potential issues related to AI.



Risk management framework

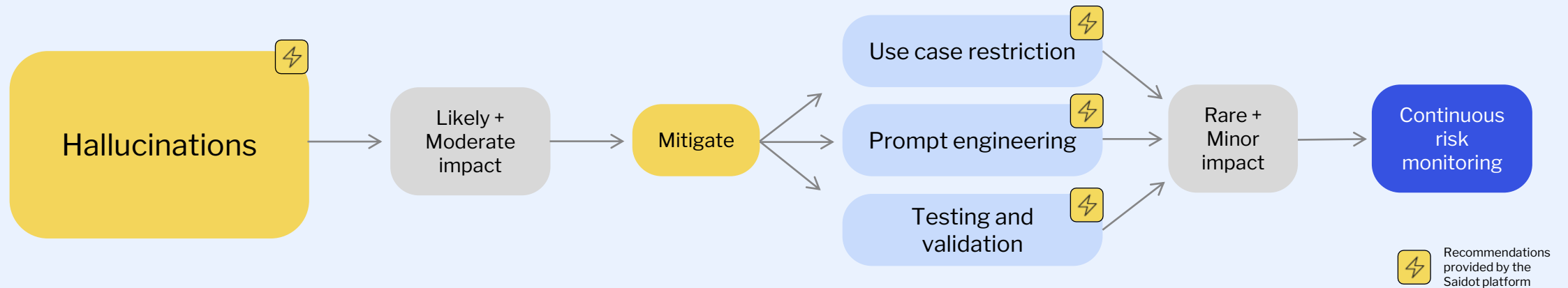
- 1 Identify risks
- 2 Document risks
- 3 Evaluate risks
- 4 Treat risks
- 5 Assess residual risk
- 6 Monitor risks



 Recommendations provided by the Saidot platform

Example: Risk management

Managing the risk of hallucinations using mitigations recommended by the Saidot platform



1. Identify risks

Saidot platform helps identify relevant risks for your registered AI systems.



How to identify AI risks?

Your organisation can identify relevant risks for the registered AI system on the **Saidot platform**.

- The Saidot platform **recommends risks** from the comprehensive Saidot Library based on the system's context and components.
- Some **context-specific risks** might not have been captured by the platform or risk recommendations. Therefore, it is important to take accountability in identifying these, by e.g. scrolling through the Risk Library yourself.

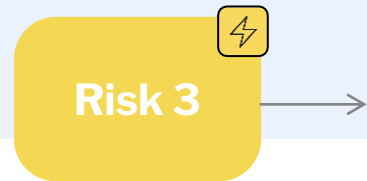
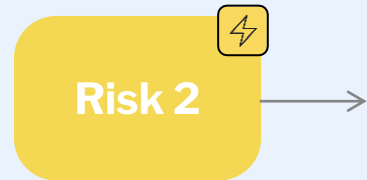
Risk recommendations are based on:

- Risks associated with the same model your AI system uses
- Risks relevant to the industry your AI system operates in
- Risks related to the function of your AI system
- Risks associated with the type of system your AI employs

 [Identify risks](#)

2. Document risks

Your organisation can document the identified risks on the Saidot platform




How to document AI risks?

- Documenting risks involves recording the risk's **name**, **description**, **owner**, **type**, **source**, and **consequences**.
- Context-specific risks can be documented on the Saidot platform as **custom risks**.
 - These may include more contextual or specific risks you have identified as relevant within your organisation.
- Documenting risks facilitates the organised and accessible tracking and evaluation of risks.



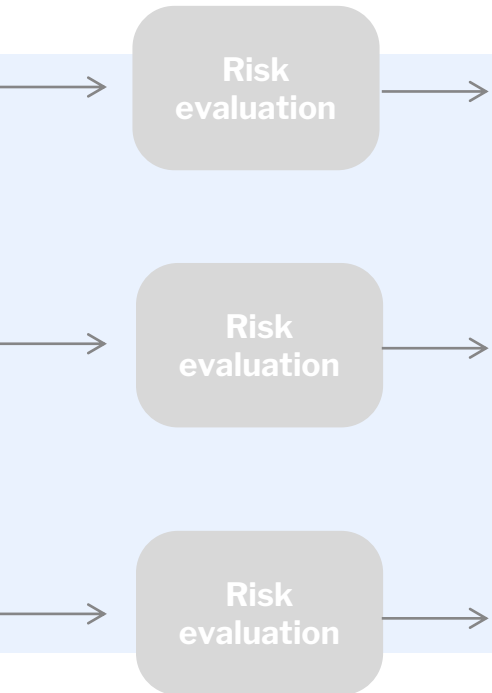
Freya, Risk owner

 **Risk owner:** the person managing the risk, often the individual or entity within the organisation who is accountable for overseeing the risk and its treatment.

 [Document risks](#)

3. Evaluate risks

Evaluating the AI risks helps you see the severity and probability of facing the risk's consequences.



How to evaluate AI risks?

- Risk evaluation includes assessing the **inherent likelihood** and **impact** of each risk, providing a baseline understanding of the severity and probability of possible adverse events.
- Evaluating the marginal risk enables your organisations to **differentiate AI specific risks** from those already existing in the organisation.
 - 🧠 **Marginal risk:** how the introduction of AI technology changes the risk.
- This comprehensive assessment helps your organisation **effectively prioritise specific risks** as well as available **risk management resources**.

Probability × Impact

Likelihood	Almost certain – 5	5	10	15	20	25
	Likely – 4	4	8	12	16	20
	Possible – 3	3	6	9	12	15
	Unlikely – 2	2	4	6	8	10
	Rare – 1	1	2	3	4	5
		1 Negligible	2 Minor	3 Moderate	4 Major	5 Severe
		Impact				

💡 Saidot's risk evaluation methodology builds on well-recognised international standards.

Inherent and marginal risk

Evaluate risks by assessing their inherent and marginal risk


Inherent risk

 **Inherent risk:** The inherent likelihood times the impact of each risk.

Probability × Impact

Likelihood	Almost certain – 5	5	10	15	20	25
	Likely – 4	4	8	12	16	20
	Possible – 3	3	6	9	12	15
	Unlikely – 2	2	4	6	8	10
	Rare – 1	1	2	3	4	5
		1	2	3	4	5
	Impact					
	Negligible	Minor	Moderate	Major	Severe	

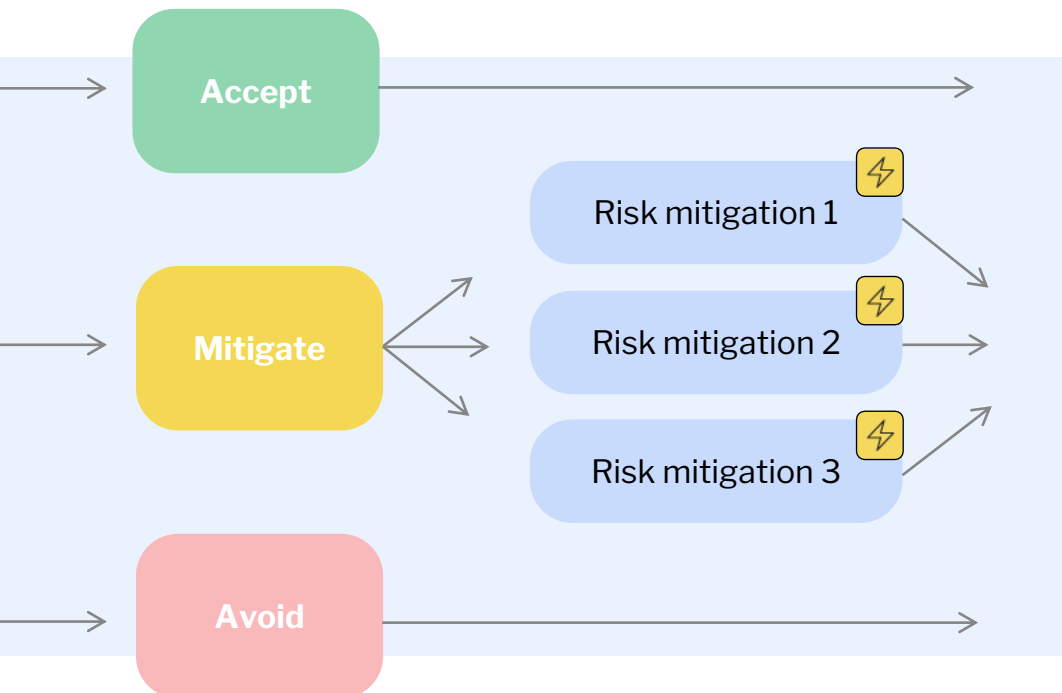
Marginal risk

 **Marginal risk:** How much higher or smaller the risk is when using AI, compared to situation of not using AI.

Significantly smaller	Significantly reduced risk level due to the introduction of AI technology.
Somewhat smaller	Somewhat reduced risk level due to the introduction of AI technology.
Same	Similar risk level due to the introduction of AI technology.
Somewhat higher	Somewhat higher risk level due to the introduction of AI technology.
Significantly higher	Significantly higher risk level due to the introduction of AI technology.

4. Treat risks

Treat risks based on risk evaluation outcomes, by accepting, mitigating, or avoiding them




How to treat AI risks?

- You can decide to **accept, mitigate**, or **avoid** an AI risk.
- The input of the risk treatment strategy is based on the **risk evaluation outcomes** in the form of a **prioritised set of risks** to be treated based on inherent risk scores.
- Effective risk management enables right-size governance.
 - 🧠 **Right-size governance:** Deploying a treatment strategy that is proportional to an AI system's risks, based on the risk evaluation.

Risk treatment strategies

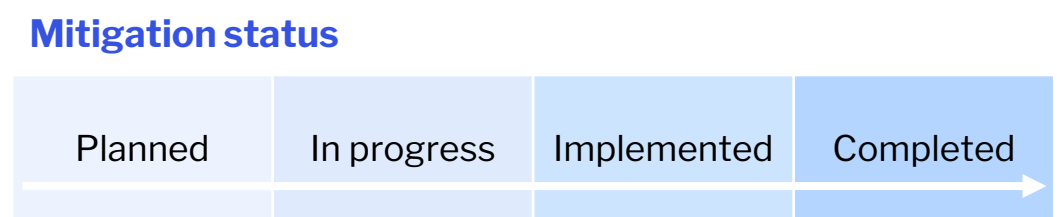
Risks can be accepted, mitigated, or avoided

Risk treatments

 **Risk treatment:** Selecting and implementing options for addressing specific risks.

Evaluation scale	Evaluation	Treatment strategy options
Low	Acceptable as is	<u>Accept</u>
Medium	Tolerable under control	<u>Mitigate</u> / Transfer / Avoid
High	Unacceptable	<u>Mitigate</u> / Transfer / Avoid

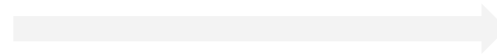
Treatment strategy	Description
Avoid	Eliminate the risk by either abstaining from the activities that introduce it or by removing its source.
Mitigate	Reduce the risk effects by altering its probability and/or impact. This means you reduce the risk effects by guided risk mitigation.
Accept	Decide to accept the risk by informed decision. This means you accept the risk without mitigations. Document your risk acceptance justification.



Sources: Adapted based on ISO/IEC 27005

Example: Risk treatment

Privacy infringement	20	25
4	8	12
3	Inherent risk: 2	
2	Severe impact x possible	
1	2	3
	4	5
		15



Marginal risk: **Significantly higher**

Risk treatments

Evaluation scale	Evaluation	Treatment strategy options
Low	Acceptable as is	<u>Accept</u>
Medium	Tolerable under control	<u>Mitigate</u> / Transfer / Avoid
High	Unacceptable	<u>Mitigate</u> / Transfer / Avoid

Example mitigations:

- ⚡ Data handling and privacy policies
- ⚡ Privacy-by-design
- ⚡ Anonymisation
- ⚡ Encryption

Residual risk: **9**

Public Benefits and Welfare Assistance Chatbot

⚡ Risk: Privacy infringement

Details:

- Chatbot that helps **citizens** apply for **public assistance programs**.
- The bot may collect **sensitive data** on personal finances, family status, and employment history.


Write a message

5. Assess residual risk

Residual risk indicates the effectiveness of the treatment strategy, and its impact on the inherent risk level and informs decision-making on the next steps



How to assess residual risk?

-  **Residual risk:** the remaining risk level after the adequate treatment strategy is carried out.
- The scales to evaluate residual risk level (**impact** and **likelihood**) are the same as those used to evaluate the inherent risk level.

		Probability × Impact				
Likelihood	Almost certain – 5	5	10	15	20	25
	Likely – 4	4	8	12	16	20
	Possible – 3	3	6	9	12	15
	Unlikely – 2	2	4	6	8	10
	Rare – 1	1	2	3	4	5
			1	2	3	4
		Negligible	Minor	Moderate	Major	Severe
		Impact				

 [Assess residual risks](#)

6. Monitor risks

Continuously monitoring all aspects of the risk managing process ensures it stays up-to-date



Continuous
risk
monitoring

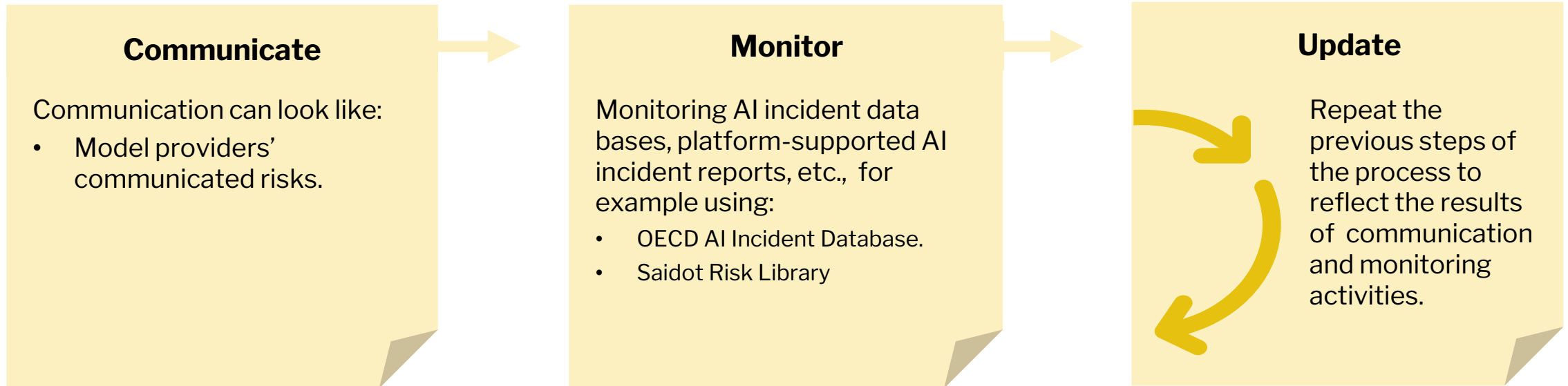
How to monitor AI risks?

- Regular **monitoring and review of risks**, their **treatments** and the **risk management process** is a vital part of your organisation's risk management strategy.
- On the Saidot platform, risk monitoring is facilitated by risk and mitigation **recommendations**.

 [Monitor risks](#)

In practice: Risk monitoring

In practice, risk monitoring relies on a variety of tasks and sources



Inherent risk

The inherent likelihood and impact of each risk.

Probability × Impact

Likelihood	Almost certain – 5	5	10	15	20	25
	Likely – 4	4	8	12	16	20
	Possible – 3	3	6	9	12	15
	Unlikely – 2	2	4	6	8	10
	Rare – 1	1	2	3	4	5
		1	2	3	4	5
	Impact					
	Negligible	Minor	Moderate	Major	Severe	

Marginal risk

How much higher or smaller the risk is when using AI, compared to situation of not using AI.

Risk treatments

Actions to treat the risk and reduce the probability or impact.

Evaluation scale	Evaluation	Treatment strategy options
Low	Acceptable as is	<u>Accept</u>
Medium	Tolerable under control	<u>Mitigate</u> / Transfer / Avoid
High	Unacceptable	<u>Mitigate</u> / Transfer / Avoid

Residual risk

Risk level after treatments.

Monitor risk

Continuously monitoring all aspects of the risk managing process

Saidot AI Governance Framework

AI policies

Direct and support the development and use of AI in line with business objectives and legal obligations.

AI lifecycle management process

Follow best-practice lifecycle model stage-specific requirements.

Third-party management

Manage AI partners and the use of third-party AI components effectively.

Data management

Align and connect with data management practices to ensure AI system data quality.

Risk management

Apply risk-based governance and robust AI risk management processes and practices.

Compliance management

Analyse and implement legal and policy requirements to ensure AI compliance.

Roles and responsibilities

Identify stakeholders involved in AI governance and assign responsibilities.

Tools and resources

Leverage shared expertise, data, and tools for governance quality and efficiency.

Transparency and documentation

Ensure transparent communication with AI users and other AI stakeholders through documentation standard and AI inventory.

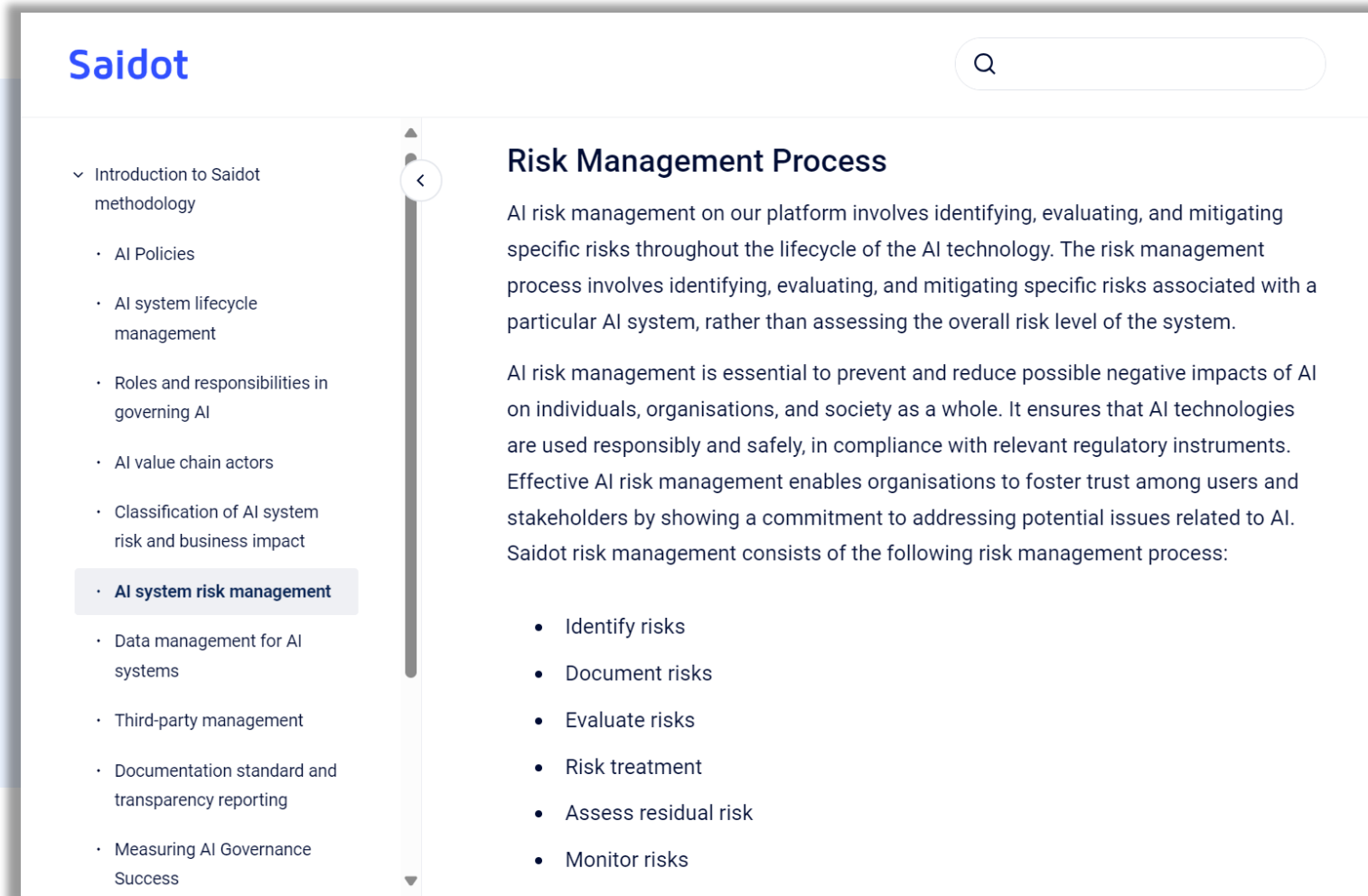
Knowledge base

Need more information? Saidot knowledge base is a comprehensive guide to all things AI governance

Your guide to using Saidot

- Detailed overview of the **methodology behind the AI governance framework**, split into individual chapters for each element.
- **Step-by-step guide** to using the Saidot platform.

 [Knowledge base](#)



The screenshot displays the Saidot knowledge base interface. On the left is a navigation menu with the following items: 'Introduction to Saidot methodology', 'AI Policies', 'AI system lifecycle management', 'Roles and responsibilities in governing AI', 'AI value chain actors', 'Classification of AI system risk and business impact', 'AI system risk management' (highlighted), 'Data management for AI systems', 'Third-party management', 'Documentation standard and transparency reporting', and 'Measuring AI Governance Success'. The main content area is titled 'Risk Management Process' and contains the following text: 'AI risk management on our platform involves identifying, evaluating, and mitigating specific risks throughout the lifecycle of the AI technology. The risk management process involves identifying, evaluating, and mitigating specific risks associated with a particular AI system, rather than assessing the overall risk level of the system. AI risk management is essential to prevent and reduce possible negative impacts of AI on individuals, organisations, and society as a whole. It ensures that AI technologies are used responsibly and safely, in compliance with relevant regulatory instruments. Effective AI risk management enables organisations to foster trust among users and stakeholders by showing a commitment to addressing potential issues related to AI. Saidot risk management consists of the following risk management process:'. Below this text is a bulleted list: 'Identify risks', 'Document risks', 'Evaluate risks', 'Risk treatment', 'Assess residual risk', and 'Monitor risks'. A search bar is visible in the top right corner of the interface.



Thank you.
Kiitos.

Contact us

Iiris Lahti
iiris@saidot.ai