Guide to Risk Management and Responsible Use of Artificial Intelligence Tools in the Public Sector

Public Consultation Draft**

*Illustration created using Microsoft Designer's Image Creator tool.

משרד החדשנות, המדע והטכנולוגיה
Ministry of Innovation, Science & Technology

ייעוץ וחקיקה
OFFICE OF LEGAL COUNSEL AND LEGISLATIVE AFFAIRS
المشورة والتشريع

משרד המשפטים
MINISTRY OF JUSTICE | وزارة العدل

מערך הדיגיטל הלאומי
نظام الديجيتال الوطني
Israel National Digital Agency

June 2025

**This is an informal translation of the Hebrew document published by the Israel National Digital Agency on June 3, 2025. Minor modifications have been made for improved clarity and readability.

Guide to Risk management and Responsible Use of Artificial Intelligence Tools in the Public Sector

Public Consultation Draft

Table of Contents

1 Introduction

1.1 Background

In recent years, the use of artificial intelligence (A) technologies has expanded across both the public and private sectors, in Israel and around the world. The public release of OpenAI's ChatGPT in late 2022 enabled, for the first time, simple and user-friendly use of generative artificial intelligence applications. Since then, AI has continued to evolve rapidly, including the recent development of AI agents capable of performing complex tasks based on detailed instructions.

AI technologies offer vast potential to drive economic growth, enhance productivity, raise living standards, and improve the efficiency of public administration. In addition, they can significantly advance AI-based innovation in the public sector and in industry, helping to modernize infrastructure, streamline government and municipal services and support inclusive growth, sustainable development and social well-being. These advances also position Israel to strengthen its global leadership in innovation and technology.

However, these opportunities come with various challenges and risks – operational, social and more – and raise various legal issues, particularly within the public sector. Realizing their significant potential requires addressing these challenges and managing associated risks thoughtfully and systematically.

This Guide is the first government publication in Israel that sets forth best practices for public sector entities seeking to incorporate AI systems in their operations. It addresses organizational, technological and business considerations, and outlines a structured process for assessing, managing, and minimizing risks, to reduce uncertainty in AI implementation. Following these practices may also assist organizations in meeting international certification standards.

This Guide was prepared in collaboration with the Israel National Digital Agency, the Legal Counsel and Legislative Affairs Department of the Ministry of Justice and the AI Policy and Regulation Center at the Ministry of Innovation, Science and Technology. The development process included consultation with public sector stakeholders, including the National Cyber Directorate, the Privacy Protection Authority, the Accountant General, the Government Procurement Administration, various Chief Data Officers (CDOs) and experts and counterparts from leading countries in the field.

1.2 Target Audience

The Guide is designed for public sector bodies looking to incorporate AI in their operations and in the processes they manage. It identifies the **roles and responsibilities** of key actors in the process, and provides each of them with detailed guidelines on the risk management procedures relevant to their roles. This Guide also includes guidelines for **end users**, i.e., public sector employees who use AI tools.

1.3 Core Principles

The development of this Guide was based on a proactive and responsible approach to AI use, grounded in a clear methodology, with appropriate management of both benefits and risks. The Guide is based on several principles:

- **Risk-based approach–** responsible use of AI systems does not mean completely eliminating all risk. The Guide promotes a tiered risk management approach: the higher the risks, the more robust risk mitigation measures and control processes are required.
- **General –** this Guide serves as a general recommendation, which may be adapted to specific organizational contexts and needs, supporting customized risk management practices.

- **Process-oriented**–the guide outlines a risk-management process, but does not mandate specific fixed thresholds for acceptable risk.

- **Dynamic –** the Guide will be periodically updated, based on the needs of public bodies, technological developments, regulatory developments in Israel and around the world, and evolving best practices.

- **Alignment with government policy principles –** the Guide is based on the Digital Agency's guidelines on ICT risk management.[1] It is consistent with the "Artificial Intelligence Regulation and Ethics Policy Principles" ("**AI Regulatory Principles Document**") prepared by the Ministry of Innovation, Science and Technology and the Legal Counsel and Legislative Affairs Department at the Ministry of Justice, which reviews the risks and challenges arising from the development and use of artificial intelligence systems in the private sector and contains ethics and regulatory policy recommendations.[2]

- **Consistency with regulatory frameworks and leading standards** – the Guide draws on AI risk management methods set forth in relevant international standards, as well as internationally accepted frameworks, including: (1) the OECD Principles on responsible use of AI; (2) the Council of Europe's Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law[3] (CAI Convention); and (3) the convention's accompanying methodology - HUDERIA; (4) similar frameworks from the United States, the United Kingdom and Canada.

- **Prioritization of the government cloud environment –** the government of Israel provides a range of AI services through its "Nimbus" cloud infrastructure, which includes both direct services from the government cloud providers (Layer One), and third-party vendor applications in the cloud marketplace (Layer Five)[4] including some of the world's most advanced generative AI applications. Marketplace services are hosted securely in Israel. They are pursuant to Israeli law and contractual terms set forth in the Nimbus tender and with government ministries' requirements. The use of AI applications within these environments is therefore strongly encouraged.

1.4 What does the Guide include?

The Guide has two main parts: the first part describes the foundational principles for the responsible use of AI in the public sector. The second part presents an organizational model for responsible use of AI, including a description of the key AI actors and their respective roles and responsibilities. Two central roles are the "**AI Governance Officer**", tasked with establishing and implementing an organization-wide policy for responsible use of AI, and the "**Business Process Manager"** who spearheads the efforts to implement an AI system for a specific business need.

The Guide includes several annexes: details on the roles and responsibilities of the AI Governance Officer and of the Business Process Manager, guidelines for end users, and a comprehensive risk management methodology, references to international standards, and a glossary of key terms. The Guide does not cover

[1] Includes additional guidelines, in particular guidelines from the Government Cyber Defense Unit (Yahav) regarding secure use of AI-based chat.

[2] Among other things, the document recommends a sectoral regulation approach (as opposed to broad, horizontal regulation). For an example of this approach, see the Interim Report of the Inter-Agency Taskforce for Examining the Use of Artificial Intelligence and Machine Learning in the Financial Sector.

[3] The Convention is intended to address challenges that arise throughout the lifecycle of AI systems. It applies primarily to uses of AI in the public sector, but also requires member states to address the risks posed to human rights, democracy, and the rule of law by private sector uses of AI. The State of Israel was a part of the negotiations that led to the elaboration of the Convention and signed it on September 5, 2024. Israel has not yet ratified the Convention, such that it is not formally binding.

[4] The catalog of products approved for purchase in Level 5 is available here.

public authorities' legal obligations in relation to the use of AI systems. Legal guidance is being prepared by the Legal Counsel and Legislative Affairs Department at the Ministry of Justice to assist public sector legal departments.

The Israel National Digital Agency provides assistance and support services to implement responsible use of AI in the public sector, such as training via its digital skills academy ("HaDigitalit"), a risk management process implementation support, and an expert advisory center. Furthermore, the Digital Agency will work with various bodies to improve the guide as the needs evolve.

This version is open for public comment. Feedback, questions or suggestions may be submitted via e-mail to: ResponsibleAI@digital.gov.il.

2 Responsible Use in the Organization

"Responsible AI" is an approach that encourages organizations to implement AI in an informed manner. It requires a holistic perspective spanning the lifecycle of an AI tool, from initial design and development to its use by the end user. It addresses, inter alia, the tool's expected impacts, both positive and negative.

This approach is described in the OECD recommendations on **responsible stewardship of trustworthy AI**, published in 2019 and updated in 2024. The recommendations include several non-binding principles of responsible use of AI. Below is a summary of the principles:[5]

1. **Inclusive growth, sustainable development, and well-being** – measures to promote AI in pursuit of beneficial outcomes for human beings and the environment, such as augmenting human skills and capabilities, addressing the exclusion of underrepresented populations, reducing gaps and social inequalities, protecting the environment, and streamlining planning and construction processes.
2. **Upholding the rule of law, human rights and democratic values, including fairness and privacy** – the development, implementation and use of AI systems should conform to democratic values and the rule of law in a manner that respects human rights (including dignity, equality, privacy and autonomy). It is also important to address phenomena amplified by AI such as disinformation. To uphold these principles, the relevant actors must establish appropriate mechanisms, such as human involvement and oversight and risk management, in accordance with the law, the context and available technological capabilities.
3. **Transparency and explainability** – This refers to transparency towards the public regarding AI systems used by the organization. Transparency may include providing relevant information, including for example on how the system operates (e.g., its capabilities and limitations), disclosing when interactions with an AI system occur, providing information about the data sources used and the system's underlying reasoning, and enabling affected parties to challenge a negative outcome.
4. **System security and safety** – when developing and using AI systems, it is crucial to ensure they are reliable and safe throughout their entire lifecycle. Systems should operate properly and not create unreasonable security or safety risks, under both foreseeable and unforeseeable conditions, in cases of misuse or other adverse conditions. Furthermore, mechanisms must be put in place to address risks caused by the system.
5. **Accountability** – organizations are expected to be accountable for the proper functioning of an AI system, and to respect the above principles, based on their roles and subject to technological limitations. To this end, they must develop risk management mechanisms and internal rules of conduct for dealing with risks (including social risks, such as biases, privacy, worker rights and violation of intellectual property, safety etc.).

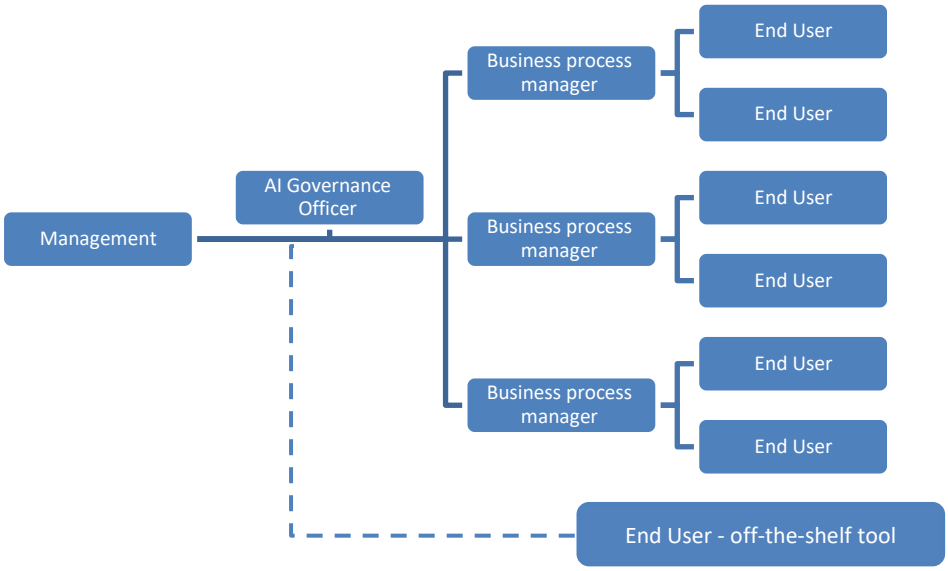
The AI Regulatory Principles Document is based on the same principles, with minor adjustments. These are also referenced in the TELEM Forum AI Strategy (2025 Report).

These principles convey that responsible use of AI is not merely a risk management methodology or a check-the-box exercise. Rather, it requires, fundamentally, an organizational culture that internalizes and embraces these values. It calls for acute awareness of the implications of AI, literacy and understanding of the technology, and stewardship of development and implementation processes of AI systems, rooted throughout organization. Responsible use of AI depends on creating organizational workflows and allocating roles and responsibilities that foster this goal. Future versions of the Guide will expand on this.

[5] This is not a literal translation of the Principles.

3 AI Governance, Key Roles and Responsibilities

The proposed structure is comprised of four layers of main actors, as specified below:



3.1 Senior Management

An organization's senior management is responsible for leadership and steering of responsible use of AI processes. To this end, it should allocate sufficient operational and budgetary resources. Among other things, it should **appoint the organization's AI Governance Officer** and empower them to carry out the actions described herein, establish public engagement processes, and report periodically on the organization's activities in the field.

3.2 AI Governance Officer

The AI Governance Officer leads the development and implementation of the organization's procedures for responsible use of AI. This Guide does not prescribe who should fill this role. While the organization's CDO (Chief Data Officer) is the default recommendation, other suitable officers could fulfill this role (e.g., CIO, or VP for Strategy and Policy). An organization could also establish an internal or external governance committee to exercise these functions. What matters is that the AI Governance Officer have relevant expertise on responsible use of AI and that they be afforded the internal organizational support and means necessary to fulfill their duties. For a detailed description of the AI Governance Officer's responsibilities, see Annex A.

3.3 Business Process Manager

The Business Process Manager is an officer within the organization who seeks to implement an AI solution in their field. For example, they can be a department head responsible for regulating a specific field, the chair of a benefits review committee, or an official responsible for planning, budgeting or internal control processes. For a detailed description of the Business Process Manager's responsibilities, see Annex B.

3.4 End Users

End users are public sector employees that use an AI system to perform their duties. The AI system in question could be one that operates in the organization's internal computing environment (such as the organizational cloud environment) with approved AI tools, or it could be an off-the-shelf product available to the general public. Regardless of the system type, end users are expected to use AI tools responsibly and thoughtfully. For guidelines for end users within an organization, see Annex C.

Annex A – AI Governance Officer: Roles and Responsibilities

Main role: establishing and implementing the organization's AI policy, including governance, management processes and reporting.

The following table details the roles of the AI Governance Officer, namely policy development, policy implementation and day-to-day management.

Proposed Actions	Details
1 – Policy Development	
1.1 Responsible use	Developing an overarching policy for the responsible use of AI within the organization. The policy should integrate the responsible AI principles (Chapter 2 of this Guide) (with necessary adaptations), and the risk management methodology (Annex D). It should include internal processes and mechanisms enabling consultation, coordination and implementation of the policy.
1.2 AI data governance	Developing data governance policies, specifically, for data used by AI systems and models, to address specific risks, including: (1) the rights and interests of data subjects and the right to privacy in particular, (2) training the models on diverse, representative, and up-to-date data to reduce the risk of algorithmic biases, false or harmful outputs. Ensuring functional continuity and safeguarding organizational data when off-the-shelf AI tools are used for various processes. The organization should favor the use of AI tools it has approved rather than private accounts. It should create organizational solutions to preserve organizational information and knowledge management and processes, for situations in which the person managing the process leaves the organization or changes roles.
1.3 Policy on off-the-shelf AI tools and AI systems in the organization	Establishing authorization and use policies regarding products and applications, in coordination with relevant parties within the organization, such as the CISO, DPO, accounting, etc. Promoting the acquisition of AI services in government cloud environments, (as opposed to applications available to the general public), for greater security. The use of off-the-shelf tools available within Nimbus, should be in accordance with relevant Finance and Economy Directives ("TAKAM") rules;[6] the use of off-the-shelf tool not on Nimbus is subject to TAKAM rule 16.12.1.2 - Project Nimbus - General Guidelines for the Procurement of Public Cloud Services. In coordination with the organization's CISO, consider whether to prohibit or limit the use of off-the-shelf AI-based tools. The policy should be made

[6] For AI services offered by AWS and Google, see TAKAM Rule 16.12.2 "Supply of AWS and Google Public Cloud Services to Government Ministries"; for AI services offered by Salesforce under a central tender, see TAKAM Rule 16.2.4 "Supply of Customer Relationship Management (CRM) Services on the Cloud"; for AI services offered under Nimbus by third party suppliers, see TAKAM Rule 16.2.4 "Supply of Customer Relationship Management (CRM) Services on the Cloud". TAKAM Rule 7.10.7 "Contract for the purchase of AI services".

	known within the organization. Additionally, consider, in consultation with relevant actors in the organization, whether their use can and should be blocked, for information security and cyber protection reasons, or if they were developed or operate in non-democratic countries.
1.4 Permission management policy	Formulating a policy for managing end users' permissions and use of external and internal AI tools.
2 – Policy Implementation	
2.1 Implementation of rules for management of data used by AI	Implementing data management rules (developed in accordance with Section 1.2 above), specifically, for organizational data used by AI systems and models.
2.2 Fostering AI literacy	Deepening knowledge within the organization as to the benefits and risks of AI tools, by establishing an organizational AI literacy program that includes internal and external training sessions and advanced studies (including trainings provided by the National Digital Agency's "HaDigitalit" school) for Business Process Managers and end users.
3 - Policy Implementation -	
3.1 Risk-based oversight processes	**Systems with no identified risk** Conducting an annual sample test with the systems' Business Process Manager to ensure that no risks have arisen.
	Low risk system Receiving an annual report from the Business Process Manager.
	Medium risk systems **1)** Authorizing the use of the system in accordance with the risk management plan submitted by the Business Process Manager; **2)** Receiving reports on updated risk evaluation, at least twice a year; **3)** Receiving immediate reports in case of indications of risk materialization.
	High risk systems **1)** Authorizing the use of the system and the risk mitigation plan, and follow-up once every quarter; **2)** Receiving reports on risk reevaluation at least twice a year; **3)** Receiving immediate reports in case of indications of risk materialization.
3.2 Creating an updated situational report on risk mapping and rating	Maintaining an up-to-date situational report of the various risks for organizational systems that use AI; reviewing the reports from Business Process Manager and other AI actors within the organization, such as identifying information leaks stemming from inputs of protected data into external AI tools. When creating the situational report, interactions between AI systems within the organization that create or compound risks should be identified.

3.3 Handling AI incidents	Determining ways of handling AI incidents.
3.4 Transparency	Publish information on the organization's website on AI systems currently used, in accordance with the responsible use and risk management policy (see more on the Risk Management Methodology).[7]
3.5 Reporting of AI Incidents	Reporting to the senior management and relevant parties within the organization (e.g., Data Protection Officer) when an AI risk materializes. In particular, if an AI-related information security risk materializes or if any information leak or security problem is suspected, it must be reported to the relevant bodies (e.g., National Cyber Directorate, Privacy Protection Authority, Cyber Protection Unit in the National Digital Agency).
3.6 Post-incident inquiries	Conducting proactive and thorough inquiries into AI incidents, including tracing the various decisions related to the system's specification and operation and the manner in which risks were managed. Insights – according to the findings (1): considering the changes to the way the system is operated and risks are managed; (2) in case of serious incidents, the considering the temporary or permanent discontinuation of the system; (3) in cases where significant or serious damage is caused, consulting with the organization's senior management and informing the public and those are impacted.
3.7 Internal Reporting	Annual reporting and documentation to the Director General or senior management on AI-integrated systems used by the organization, their benefits and risks, the risk management and oversight processes that were implemented, and any AI incidents that have occurred. Submitting recommendations for improving responsible use of AI processes.

[7] AI systems used in the government will be published on the AI Watch website created by the National Digital Agency. It is intended to serve as a registry of AI systems and will provide up-to-date data on the use of AI in various organizations, for government and residents alike.

Annex B – Business Process Manager: Roles and Responsibilities

Main role: business design and implementation of an AI-integrated system, and conducting systematic risk management.

The following table details the roles of the Business Process Manager, namely, defining the business need and required specifications, risk management and providing instructions to end users.

Proposed Actions	Details
1 – Defining the business need and required specifications	
1.1 Collection of information	Defining the business need for an AI system (efficiency, accuracy in decision-making, improved data analysis, enhanced delivery of services to citizens, etc.). Assessing the current situation and exploring alternatives (both AI-based and non-AI options). Initial assessment of existing AI tools that may meet the business need, with the assistance of the CDO and relevant parties from the IT division. Specification and documentation of the particular business need and of the initial assessment performed (alternatives, relevant tools). **Note**: In some cases, non-AI tools that meet the needs efficiently and effectively may be available. It is important, at the onset of the process, to articulate with precision the rationale for turning to an integrated AI system
1.2 Consultation	Engaging with key AI actors within the organization such as the DPO, CDO, IT division manager, project/system/product manager, Business Process Manager, cloud manager, CISO, and the customers. Generally, depending on the nature of the project the legal department should be involved, preferably as early as possible, to assess the legal situation.
1.3 Business design	Setting forth with accuracy the requirements of the AI system, if the decision is made to incorporate an AI system following the above process.
2 – Risk management	Implementing the risk management process specified in Annex D. Documenting the findings (mapping the benefits and risk of a specific system in a comprehensive, detailed and clear manner, formulating a risk mitigation plan).
3 – Instructions to end users	For a government cloud-based AI tool, relevant instructions for end users must be obtained from the AI Governance Officer or the supplier, as applicable. If such instructions cannot be provided, end users must be provided with the information included in the End User Guide (Annex C), as a default. These should be complemented by additional instructions, as may be deemed necessary in light of the product specifications and the risk management process (Annex D).

Annex C – End User Guide

This chapter is intended for end users of AI-based systems, i.e., public sector employees who use AI applications, including generative AI. It includes general recommendations for responsible use of AI systems within the public sector. The recommendations are relevant to the use of both off-the-shelf tools available to the general public and tools made available specifically to the organization. Some of the recommendations may also be relevant to the use of customized applications developed specifically for the organization.[8]

In general, when an AI system is made available to employees by the organization, the user can expect to receive instructions that address the following topics (as well as additional guidelines) from the AI Governance Officer or the relevant Business Process Manager. What follows is intended to provide a general, complementary framework, as necessary.

This Guide does not include an analysis of any legal issues or of applicable legislation. It is intended to emphasize key points regarding the use of AI system that employees are encouraged to follow.

Background: Characteristics and Limitations of AI Systems

Users should be aware of the limitations of AI systems. The following is a description of the main limitations.

Some limitations are particularly common in generative AI systems, i.e., systems that can create new content – such as text, images, audio or video – based on patterns learned from existing data (e.g.: ChatGPT, Gemini or Midjourney). Some applications have dedicated capabilities, for example "personal assistant" applications with generative capabilities, conversational search engines and meeting transcription and summary.

AI models differ from one another, due to different model designs or training on different datasets. Each model has its own characteristics - including strengths, weaknesses, and biases. Below are a few common limitations that should be noted:

- **"Hallucinations" – credibility and accuracy limitations**: AI systems sometimes provide uncredible and inaccurate information, occasionally without disclaimer. For example, the results of text models sometimes contain citations of inaccurate and even fabricated sources.
- **Limited objectivity:** When AI models are trained on databases containing information about human beings from certain cultural backgrounds, they may lack cultural inclusiveness, particularly regarding minority groups. On controversial issues, AI systems may lack representation of a variety of perspectives and reinforce dominant narratives.
- **Harmful content and discrimination:** A central challenge for AI-based systems is the risks of discrimination and biases. This could be caused by various factors, such as databases that are insufficiently inclusive or that reflect existing biases, when algorithms use data that can be used to discriminate based on one's affiliation (nationality, gender, etc.) or other information that correlates with such data. As the use of AI systems increases, so does the risk of these phenomena occurring on a large scale, compared to an individual human decision.
- **Risks With Respect to Protected Information:**

[8] As specified below, sensitive or protected information must not be submitted to a tool operating outside the organizational environmental including the organizational cloud environment.

- **Use of copyrighted information:** AI systems sometimes collect or provide information protected by intellectual property rights such as copyright or commercial confidentiality. This could result in violation of rights and unauthorized use of materials, exposing the State to legal action.
- **Use of personal information and production of aggregated data:** Depending on the AI system's settings and on the frequency of use, when users input large quantities of data over time, such data can be aggregated in the system. This could include information about the public authority and its operations, personal information of data subjects, confidential business data, and inferences) resulting from the cross-referencing of different types of data that were inputted into and processed by the AI system.

  For example, in many cases, the companies that provide free AI-based applications aggregate by default the data fed by end users via prompts and reserve the ability to use this data to train their models as well as for commercial uses. Some applications allow this data to be removed even under the free membership, but this does not guarantee that the information will not be used.

- **Anthropomorphizing**: Some models, and language models in particular, communicate with the user using natural, human language. This feature can lead users to mistakenly attribute authority or expertise to the AI system, especially with regard to sensitive subjects, and mistakenly assume that the system "understands" the meaning of things or has moral intentions, when it is in fact merely processing probabilities.

General guidelines for end users

1. Ensure that the relevant AI tools have not been prohibited for a general or particular use by the organization's cybersecurity officer, AI Governance Officer or relevant Business Process Manager.
2. When using generative AI, determine whether this is an off-the-shelf AI-based tool **external** to the organization (public or belonging to another organization), or a tool that operates in an environment **dedicated** to the organization (for example, in the organization's cloud or its information systems).

   2.1. When an off-the-shelf tool **operates in an external environment to the organization,** rather than an **environment dedicated** to the organization, such as using an online chatbot used with a private license:

   2.1.1. **Do not** use generative AI applications developed and operating in non-democratic countries. The use of applications developed in countries that promote principles of responsible use of AI is recommended.

   2.1.2. **Do not** input, in external systems, protected information such as personally identifiable information, sensitive information, information that must not be disclosed under the Freedom of Information Act,[9] or information protected by commercial confidentially or intellectual property, legal privilege or classified information.

   2.1.3. **Do not** input prompts that indicate an intention to take an action concerning a specific individual or concerning a specific sensitive governmental action.

   2.1.4. AI technologies are relatively new and evolving dynamically, and there is significant increase in the variety and number of applications available to the general public. Their respective data usage policies vary. It is therefore recommended to review the AI

[9] For example, see information that may not be disclosed under s.9 of the Freedom of Information Law, 5758-1998.

system's policy and instructions for use, if available, in order to understand the tool's limitations and the way it uses and data.

2.1.5. **Do not** independently create an automatic interface between an external AI tool and the organization's work environment, without consulting and receiving approval from relevant parties within the organization such as the CISO, the DPO (Data Protection Officer) or the person in charge of AI risk management.

2.2. When the tool is **operated in the organizational IT environment** (including the organization's cloud), it must be used in accordance with the product policy and terms of use that will be reflected to end users by the Business Process Manager.

3. When using AI in decision-making processes, **the outcomes should be used to support human decisions, and should not serve as the sole source of information or final decision**. A limited exception can be made in the case of a dedicated product that has been integrated into the organization's operations pursuant to a risk management plan, with a customized policy and after receiving all required approvals, including legal. In such a case, all applicable guidelines must be followed.

4. When the output of an external or integrated AI system appears to reveal personal, sensitive or classified information) - **this must be accurately documented and reported** to the Business Process Manager entrusted with using the system or to organization's AI Governance Officer.

5. **Do not assume the content is correct without checking** – when a factual output is received, for example by off-the-shelf products based on generative AI, it is important to cross-reference it with credible sources of information, such as publications in scientific literature or well-known websites, and to check its credibility before using it as such.

6. **Ask questions in your field of expertise** – it is recommended, as much as possible, to write prompts on issues that are within or directly related to the user's field of expertise or occupation, to enable direct review of the outputs. For example, a person who deals with property appraisals should not rely on generative AI answers to questions on taxation, because they do not necessarily have the tools to judge the answer's credibility and accuracy. In such cases, it is recommended to consult relevant experts. If applicable, consult with colleagues also on the credibility of the system's outputs.

7. **Transparency and documentation** – if generative AI outputs are used materially, it is recommended, as much as possible, to disclose this in the final product. Do not present content generated by AI as if it is the product of original human creation. Furthermore, when the system is used to assist in decision-making (e.g., in the case of dedicated systems implemented in the organization's operations), use of the system must be documented.

8. **Check if information protected by intellectual property is used** – generative AI outputs may contain copyrighted information.[10] Use discretion in using such information. If there is any risk, it is recommended to identify the tool's sources of information, if available, and check whether the content requires a citation and reference to the sources. For non-textual products, a citation and reference may be insufficient - it is recommended to consult the organization's legal department.

9. **AI literacy –** learn how to use generative AI tools and understand their advantages, limitations and risks. It is recommended to frequently attend courses and read articles on the subject.

[10] For example, where the user notices that the output contains quotes from written works and intends to use the product, e.g., in the organization's official documents, the organization's legal counsel should be consulted to examine whether this constitutes "fair use".

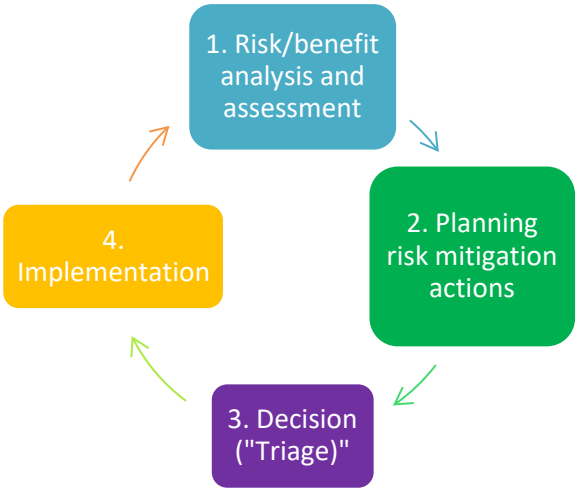
Annex D – AI Risk Management

This annex consists of five sections:

1- Risk Management Methodology

This section provides a uniform foundation for developing a risk management methodology. The organization's AI Governance Officer may, at their discretion, expand the model or change the methodology (provided it includes the core elements described herein). The main phases of an AI risk management process outlined below are: benefit analysis and assessment, risk identification, risk ranking, decision (triage) and risk management.

Risk identification and management is an iterative process. It begins with the business design and articulation of required specifications, and is performed routinely throughout the AI system's lifecycle, twice a year or as necessary given the level of risk, according to the following phases:



The following table describes the actions included in each phase.

Phase description	Action
Phase 1 - Risk/benefit analysis and assessment	Assessing, in a holistic manner, the benefits and risks of the requested AI system. Phase 1 consists of four sub-phases: **identification** of benefits and risks, systematic **analysis** of each risk and benefit, **consultation** and testing, and overall **assessment**.
Identification	• Identify the relevant risk/benefit categories (see parts 2 and 3 of this annex) as well as other emerging risk and benefits; • Map potential risks and benefits arising from the AI system; • Obtain relevant information from the system's supplier • Document the outcomes.

<table>
<tr><th>Phase description</th><th colspan="2">Action</th></tr>
<tr><td rowspan="2">Analysis</td><td>Analyze the benefits including:
Scope of impact
Duration
Magnitude
Likelihood of realization
See Part 2 of this annex.
Document the results of the analysis.</td><td>Analyze the risks including:
Scope of impact
Duration
Magnitude
Likelihood of realization
See Part 3 of this annex.
Document the results of the analysis.</td></tr>
<tr><td colspan="2">The analysis of both benefits and risks should consider potential impact on:
Service recipients in the organization
Suppliers
Partners in the system's planning and implementation
Units within the organization
The general public
It should also take into account the protection of rights and the public interest</td></tr>
<tr><td>Consultation</td><td colspan="2">Consult with all relevant parties, depending on the circumstances:
Legal aspects: Legal Department;
Privacy aspects: DPO (in charge of privacy protection in the organization)
Risks arising from cloud use: Cloud manager;
Implementation of technical risk mitigation components: IT Division manager;
Risks arising from challenges in data quality and accessibility: CDO (Chief Data Officer); on the systems' impact on other systems that are being planned or developed: PMO (Project Manager or the person in charge of planning).
Other relevant actors:
- The organization's partners, including suppliers;
- Service recipients, including subordinates
- Business parties expected to use the system;
- The general public
Document the outcome</td></tr>
<tr><td>Assessment</td><td colspan="2">Assess the overall benefit: very low/low/medium/high/very high benefit
Assess the overall risk: very low/low/medium/high/very high. See examples here.
Note: in general, when the system poses one "very high" risk, this may indicate that the system as a whole poses a very high risk. In other cases, the overall risk and benefit assessment is subject to the Business Process Manager's discretion, pursuant to the organization's policy on the matter.
Document the outcome</td></tr>
</table>

Phase description	Action
Phase 2 – Risk Mitigation Planning	Once the benefits and risks have been mapped, it is necessary to plan risk mitigation measures, among other things to determine whether to recommend implementing the system (see Phase 3 – Triage). Phase 2 consists of four sub-phases: **mapping** of measures, **consultation**, **cost assessment** and **documentation**.
Mapping	• Map the measures at your disposal – see detailed examples of reduction measures here.
Consultation	• Consult with the parties specified in Phase 1 (as needed in the circumstances). • Generally, the higher the risks, the more comprehensive the consultation should be.
Cost assessment	• Examine the costs of the risk mitigation plan including: - examining alternatives; - estimating the cost of risk mitigation measures, in relation to their expected benefits.
Documentation	Document the outcomes, including the risk mitigation plan.
Phase 3 – Decision ("Triage")	Integrating all information obtained thus far to decide whether to recommend implementing the system, in accordance with the policy established by the AI Governance Officer. The triage process includes the following sub-phases: **consultation, recommendation/decision** and **documentation.**
Consultation	Consult the parties specified in Phase 1 (as necessary in the circumstances). **Note:** when the risks associated with a given AI system are high, it is critical to verify that the organization has the requisite capacity to manage them.
Recommendation/Decision	Recommend a course of action and obtain the requisite approvals pursuant to the organizational procedures as determined by the AI Governance Officer. If it is decided to move forward with the AI system, a detailed risk mitigation plan must be drafted, to enable ongoing oversight and control for the safe and responsible use of the system.
Documentation	Documentation of the final decision, including the details of the risk mitigation plan.
Phase 4 – Implementation	This stage begins when the organization deploys the system. It includes **implementing** the risk mitigation plan, **oversight, reporting, making adjustments** as needed and proactive actions of **transparency** toward the public.
Implementation	Implement the risk measures set forth in the risk mitigation plan.
Real-time oversight	Collect data from the system; Address comments on the operation of the system from impacted stakeholders, received through various channels (public comments, the media, dialogue with civil society organizations etc.) Oversight of system functioning: by taking samples or conducting ad hoc tests, periodically or as needed, based on to developments, needs and risk level.

Phase description	Action
Reporting to AI Governance Officer	Report regularly to the AI Governance Officer, as set forth in the risk mitigation plan. Reporting to the AI Governance Officer if an AI incident occurs. **Note:** reports should make recommendations as to any necessary adjustments.
Adjustments	Make adjustments to the AI system in accordance with the AI Governance Officer's instructions. The adjustments may include, for example, changes to product specifications or how data is collected or labeled.
Public transparency	Public transparency includes several main components: • Continuous transparency regarding the use of AI systems by the organization, the types of information used by a system, and the system's general mode of operation, including its capabilities and limitations. This information should be accessible to the public. This component can be adjusted as appropriate based on the impacted stakeholders, the level of risk arising from the system, and available technologies. • Transparency towards the general public (or impacted stakeholders) if an AI incident occurs – in accordance with the AI Governance Officer's instructions. • Transparency and increased awareness of cases of direct or significant interactions with an AI system. Further details on this subject will be included in the legal guide to be attached to this document at a later stage.

2- Classification of benefits

Defining the business needs for an AI system starts with assessing its potential benefits. These can be divided into different categories. Below is a non-exhaustive list of key categories that should be identified. Additional benefits relevant to the organization's operations or expected uses can be considered.

Productivity via automation of various processes (data analysis, services to citizens, enforcement) and more efficient public policy. Examples: shortened waiting times/processing times, improving the quality of decisions, continuous work outputs, reduced burden on employees which enables them to focus on core tasks, lower risk of human errors. These and other advantages of AI systems are expected to improve the productivity of public servants as well as residents and businesses that receive public services, improve the quality of public services and enhance public trust in State institutions.

Proactivity in the design and supply of public policy and services. Example: actively reaching out to residents and businesses to realize benefits and exercise rights by activating a team of AI agents, and drafting proactive policies in preparation for emergencies by activating AI models for smart and early forecasts.

Personalization, i.e., adjusting public services to residents and businesses based on their unique needs.

Added benefits: when AI systems are used responsibly in order to promote the organization's goals, the residents receive better public service, thereby enhancing its reputing and increasing public confidence.

Use of AI can lead to valuable outcomes. For example, a call for proposals by the National Digital Agency and the Ministry of Innovation, Science and Technology, supports several highly promising public sector AI uses. See also the OECD database for use cases from other countries.

The following actions are recommended as part of the risk management methodology:

- Anticipating and defining the various expected benefits, with as much specificity as possible, including estimating the potential return on investment;
- Assessing the benefits and their relative value (from "very low" to "very high");
- Evaluating their likelihood of realization;
- Comparing risks and benefits holistically.

3- Risk Classification

In all processes, some basic risks can arise, such as human errors, slowness, poor prioritization, etc. Integrating an AI system into the process could reduce or compound these risks, or create new ones. Therefore, the added risk must be examined in relation to the basic situation and in the aggregate. For example, a reduction in system errors from 15% by humans to 1% by AI is computed as a 1% risk, but constitutes a significant improvement over the existing situation, which must be taken into account. A fundamental aspect of risk identification and analysis is determining whether the AI system creates direct or indirect risks, i.e., whether the risk originates in the AI system or if it is caused by other reasons such as an organizational or "business" process, another IT system, etc.

The following is a non-exhaustive list of the main risks that should be identified and managed. Other risks relevant to the organization's operations or expected uses of the tool should be examined.

- **Operational risk**: AI systems are based on complex internal processes, advanced technology, and human interactions prone to disruptions and failures. For example, they are based on the collection and analysis of large amounts of data, such that inaccurate, outdated or biased data may harm the quality of the system and cause errors or distorted outcomes. The system's "reasoning" could also be difficult to trace.[11]
- **Economic risk**: mistakes caused by the use of AI systems could result in economic damages, including loss of income, collection shortfall, discontinuation or failure to start a business activity (e.g., license revocation) and excess expenses. These damages can impact both the public body and the service recipients, including residents, businesses, and other organizations.
- **Risk of health and safety damage:** for example, mistakes in a system that assists in the process of issuing regulatory approvals for medications, or directing traffic, may cause actual harm to people and property.
- **Risk of environmental damage:** for example, mistakes in a system that assists in granting permits for air emissions and environmental conditions in business licenses, or perform autonomous controls for oversight and enforcement purposes. Such mistakes may lead to excess pollution and lack of enforcement.
- **Data security risks**: the use of AI systems exposes the organization to unique data security risks, which may compromise data privacy and integrity and lead to the disclosure of sensitive information. These damages may result from attacks on the tools themselves – hostile actors attempting to exploit vulnerabilities in the AI system and disrupt outcomes, or to extract information from the data management systems. Unauthorized access to information could occur in the system or in supporting infrastructure, such as databases and cloud services, resulting in the disclosure of sensitive data or the misuse of models in a way that exposes sensitive data. In addition, AI models could be vulnerable to

[11] Due to the "black box" (opacity) of AI systems.

model deception attacks, which can disrupt model predictions and impact decisions that are critical to the organization and information security.

- **Damage to public trust and the organization's reputation:** mistakes, biases or disinformation, whether externalized to end users, such as office chatbots for public use, or used for internal needs, could damage the organization's reputation and the public's trust in the organization.

Legal aspects: AI systems occasionally raise legal questions, for example, with respect to administrative law (questions regarding the scope of authority, the duty to give reasons), preventing privacy violations, preventing biases and discrimination, and protection of intellectual property. Based on the AI system and the legal questions stemming from its planned or actual use, the organization's legal department must be consulted in order to identify the system's legal requirements. The legal review must be conducted autonomously and independently. At the same time, it is important to inform the Business Process Manager of the legal requirements and questions that need to be addressed, so that these can be integrated in the design of the mitigation plan (for example, to implement means for handling bias and privacy risks) and at the decision-making stage (triage).

Exceptional risks and types of uses that should be avoided – in cases where AI tools pose a particularly high risk of harm or danger, or that could substantially and significantly harm basic rights, **a separate examination is required to decide whether to advance the project.** By way of example, in the EU AI Act enacted in 2024, certain uses of AI systems are explicitly prohibited, such has AI systems that use subliminal, manipulative, or deceptive techniques; take advantage of the weaknesses of a person or group of people, due to their age, disability or economic status; rank people based on their social behavior or personal characteristics in a way that could lead to harmful or unfair treatment in contexts specified in the law; or systems that draw conclusions and analyze emotions in the workplace or in educational institutions (other than for medical or safety purposes).

If a system being considered poses an exceptional risk or may substantially and significantly harm rights as described above, and in particular with respect to systems such as the ones described above, the organization's **legal department should be consulted** in order to determine whether a legal impediment or a significant legal obstacle renders the AI system inappropriate for use. To the extent that the legal aspects can be addressed, **relevant high-ranking policy officials must then be consulted.**

4- Examples of risk management methodology implementation

The following example relates to an AI system expected to shorten waiting times for a certain service, by a certain percentage. The percentages and scoring below are **illustrative**. The Business Process Manager should exercise discretion, based on the organization's risk management methodology and instructions from the AI Governance Officer, and determine the appropriate scoring on a case-by-case basis.

Benefit Assessment – Shortening Waiting Times

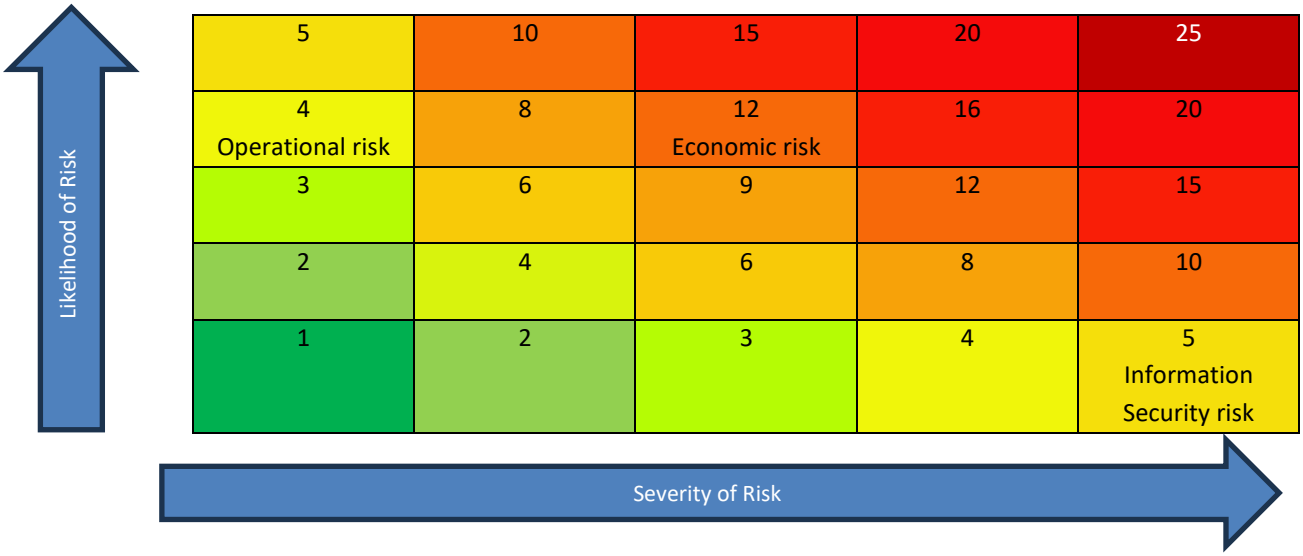
Assessment	Potential Impact	Duration	Magnitude	Likelihood
Very low	Few individuals	One-time	Less than 2%	Less than 50%
Low	Tens to hundreds	Rare	2-5%	51-60%
Medium	Thousands of individuals from vulnerable groups	Occasionally	6-7%	60-74%
High	Tens of thousands	Frequently	8-9%	75% and above
Very high	Hundreds of thousands	Regularly	10% and above	Certain or near-certain

Risk Assessment

Assessment	Potential impact	Risk Duration	Magnitude	Likelihood
Very low	Few individuals	Hours	Less than 2%	Less than 50%
Low	Tens to 49	Days	2-5%	51-60%
Medium	50-99	A week	6-7%	60-74%
High	100-999	Two weeks	8-9%	ומעלה 75%
Very high	Over 1,000 or hundreds from vulnerable groups	Three weeks or more	10% and above	Certain or near-certain

After assessing the above parameters, it is recommended to assign a numerical score **to each risk separately**, relating to the severity of the risk and the likelihood of its realization (with 1 being the lowest and 5 the highest). The weighted score for each risk is calculated by multiplying severity score (1-5) by the likelihood scope (1-5).

Following the risk identification and assessment process, it is recommended to create a "heat map" to help prioritize the attention that should be given to addressing each risk. For example, assume three risks were identified in the system: data security, economic and operational. The severity of the operational risk is very low (Level 1) but the likelihood of its realization is relatively high (Level 4). The data security risk is very high, but the likelihood of its realization is very low. The economic risk is medium (Level 3) but the likelihood of its realization is relatively high (Level 4). Visualizing the different risk levels helps the Business Process Manager deal with various risks adaptively. Furthermore, the AI Governance Officer can use tables they receive from all Business Process Managers within the organization, to create an organizational heat map and derive practical insights.



For a detailed bank of ICT risks with recommended controls, see the National Digital Agency's risk bank here.

5- Recommendations and Risk Reduction Actions

Risk management theory generally establishes four strategies for dealing with risks:

1. Avoiding the risk for very high risks that do not justify the expected benefit.

2. Ignoring the risk for very low risks that are not material to the organization or to the people or organizations associated with it;
3. Transferring the risk, when possible, to a third party such as via disclosure or insurance;
4. Reducing the risk by taking reduction measures to mitigate material risks as appropriate in the circumstances, when such actions are possible, together with continuous oversight and control of their implementation to operate the system safely and responsibly.

A variety of actions can be taken to reduce risks during the business design, development, and deployment phases of AI systems. The measures described below are general and can address multiple types of risks posed by the use of AI systems. As mentioned, implementing these measures can reduce the severity of different risks.

They can be divided into main groups: technical and architectural measures; adding components to the AI product; and business/organizational measures. **Below are examples of measures from each category:**

a. **Technical and architectural measures**

In order to reduce the risk, it is important to incorporate technical and architectural solutions at the planning stage. Here are some possible technical and architectural measures for risk reduction:

1. **Incorporating RAG or CAG techniques in the system's architecture** – a combination of information retrieval from credible sources with a generative model is expected to reduce risks of "hallucinations" and mistakes by the system, and increase the organization's ability to explain the output produced by the system. Furthermore, these techniques allow the system to operate based on credible, up-to-date information sources, provided by the organization or by external sources, and reduce the risk of unreliable output or false information.
2. **Incorporating an AI Agent dedicated to a specific task** – a combination of specialized agents in an AI system for dedicated tasks, with a clearly defined mission (such as internet search, performing calculations, etc.), can reduce risks posed by mistakes and deceptions in the AI model, and increase the system's accuracy. The operation of AI agents can also help pinpoint when human involvement is required and call upon an individual qualified to make a decision.
3. **AI guardrails –** defining technical rules that prevent the system from dealing with prohibited subjects or producing problematic answers. Guardrails block unwanted content in advance (e.g.: sensitive topics, harmful language, unauthorized decisions).
4. **Reasoning –**models that combine self-reasoning techniques provide a degree of explainability and the ability to trace to some extent the system's process for generating the output, thus reducing risks caused by entering inputs that are inaccurate or not adjusted to the user's business need.
5. **Using privacy-enhancing technologies (PET) to reduce privacy risks –** Privacy-enhancing technologies enable the processing, analysis, and sharing of data in a manner that protects the privacy of data subjects.[12] Using these technologies, coupled with privacy by design principles, can minimize the risk of breaches to privacy.
6. **Temperature control –** Temperature is a parameter in generative models that defines the output's level of randomness and creativity. Low temperature (0.1-0.3) tends to result in consistent, safe and careful outputs, whereas high temperature (0.7-1.0) tends to result in more diverse, albeit less predictable ones, which can sometimes be wrong or otherwise problematic. In some language

[12] For more information, see the guide to privacy-enhancing technologies (PETs), authored by the Privacy Protection Authority.

models, temperature control can be performed by end users using natural language within the prompt, while some models need this to be defined technically as part of the API call to the model.

b. **Product Components**

In the design of an AI-based product, it may possible to incorporate certain components that reinforce the transparency of use as well as the capability to oversee and learn:

1. **Disclosure –** a component whereby the system informs users that it is based on AI and reflect its limitations. This action can increase human control and shape how the system is used. Note that in some cases, the AI component in an ICT system is relatively minor; also, the risks posed by its operation could be low, carrying a negligible impact on the work of end users. Accordingly, discretion should be used in deciding when adding a disclosure component is appropriate.
2. **User Feedback/Reporting -** Adding a component that enables users to report problems, misunderstandings, or harm caused by using the system is expected to strengthen organizational oversight. Where appropriate, it can even allow users to appeal the resulting output within the organization.

c. **Business/Organizational Measures**

Another way to reduce risks is through mechanisms for enabling human judgment, professional oversight, promoting a culture of responsible use, and continuous learning throughout the lifecycle of the system. For example:

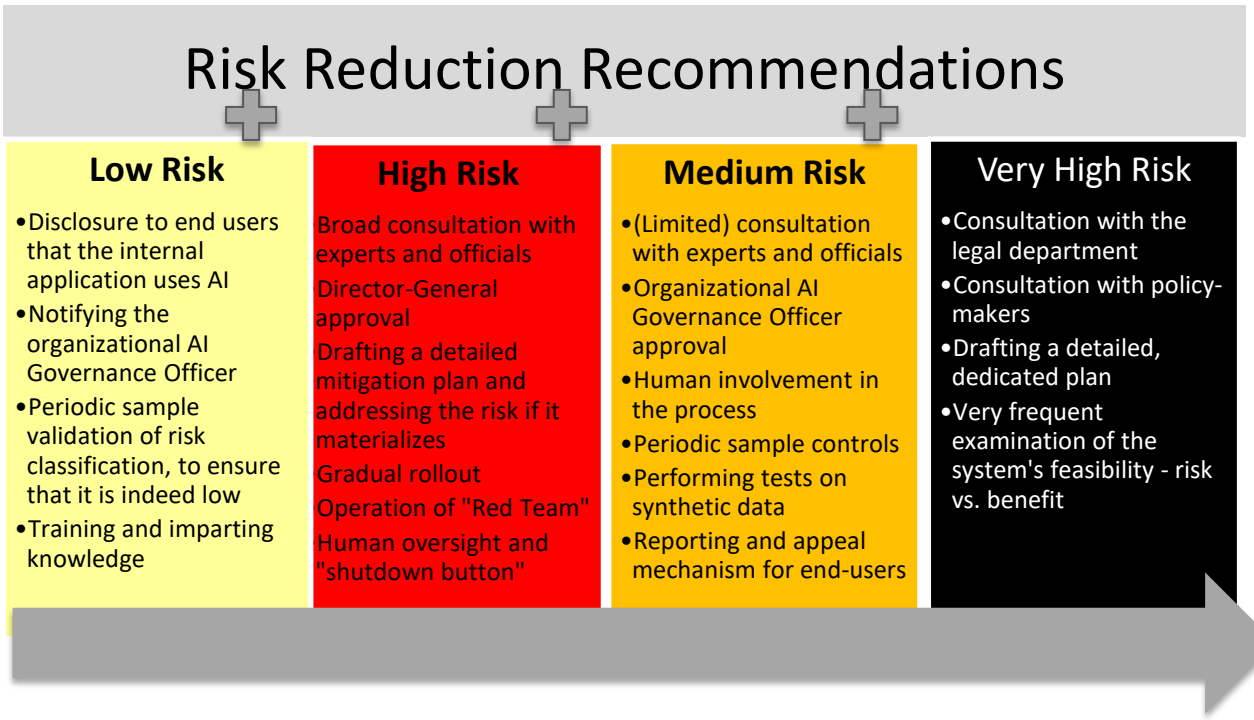
1. **Human involvement** – a combination of human elements within the system's operations and the AI-based business process. Ensuring proper human oversight throughout the AI system's lifecycle, including control of the development, deployment, use and decisions made by the system, can serve as an ongoing operational control measure. Involvement may be carried out in advance (ex ante), during the operation (real-time), or retrospectively (ex post), depending on the level of risk and the context of use.
2. **Consultations with expert –** incorporating experts from relevant fields such as ethics, law, privacy (DPO), equality and inclusion, and cyber protection, at the business design and development stages, allows to identify risks in advance and build mechanisms to reduce them.
3. **Sample controls** – Performing sample quality checks of the system's outputs, periodically or in sensitive cases, is expected to help identify risks throughout the system's life.
4. **Using a "red team" and conducting controlled attacks** – operating a red team to test the safety, reliability, and robustness of the system through controlled attacks on the system, can assist in identifying vulnerabilities, resilience failures, weaknesses, biases, and possible risks.
5. **Data Governance** – An AI system's trustworthiness depends directly on the data on which the models are based. In order to ensure quality governance of the data that is used by an organizational AI system, the following measures may be taken:
   - Using the right data for a given model – the data should be up-to-date, complete, representative, legal and reliable.
   - Access management, licensing, implementation of mechanisms to handle unbalanced data and ensure data quality and integrity.
   - Taking organizational steps to identify and reduce data biases.

   In implementing the above measures, the rights and interests of data subjects must be taken into account.

6. **Increased awareness –** an organization can perform several actions such as training, workshops and publication of policy documents and internal case studies to users and decision-makers. Furthermore, the organization may take measures to encourage reporting when AI risks materialize.

Recommended actions by severity of risk

The following diagram presents recommendations for action based on the system's risk level. These include control, notification, and action if there is a concern or if the risk has materialized. Each risk level includes the actions recommended for the risk levels below it. For optimal management of identified risks, risk management tools should be defined to allow for continuous oversight control and operation of the system in a safe and responsible manner.



Annex E – International Standards

Leading countries and standards bodies around the world are working to develop standards and guidelines for responsible use of AI. The following common principles can be found in many of these:

a. **Risk management-based approach towards AI use,** to identify, assess, prioritize and manage potential risks throughout the system's lifecycle. This approach focuses on adjusting the mitigation activities to the actual level of risk, while balancing between risk management and technological innovation and the expected benefits.

b. **Transparency** in the use of AI tools.

c. **Continuous improvement of standards and guidelines** - the texts are updated from time to time.

d. Protection of **human rights** as a basic condition for operation and use.

e. **Stakeholder participation** is an essential component in creating an open, transparent and inclusive process, which allows for optimal management of AI risks and ensures broad benefits.

f. **Oversight** and control mechanisms are tailored to the level of risk and to obstacles.

Below are some relevant standards:

1) **ISO standards** – ISO 23894 guidance on risk management concerning AI system; ISO 8000-51 for data governance within the organization; ISO 42001 on AI system management; ISO 31000 on general risk management.

2) **IEEE standards** – including P7000 on ethical design of products, P7003 on algorithmic biases, P7009 on fail-safe design of autonomous systems.

3) **CEN/CENELEC standards** - A series of standards that European standards bodies have been asked to prepare, in accordance with European Union legislation in the field (the AI Act), to assist in the implementation of said legislation.

4) **NIST.AI 100-1 -** A framework document for managing AI risks, prepared by the US National Institute of Standards and Technology (NIST).

Annex F - Glossary

The following is a glossary of common terms from the world of AI that appear in the guide.[13]

Term	Definition
Artificial Intelligence (AI) system	An artificial intelligence system is a machine-based system that, for explicit or implicit purposes, infers from the input it receives how to produce outputs such as predictions, content, recommendations, or decisions that can affect physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptability after deployment. [14]
Generative AI	AI systems that are capable of generating new content such as text, images, video, audio, and more, based on examples or data they have been trained on. These models are capable of producing creative outputs without requiring specific training for each task. There are different types of applications that **generate new content**, including: • Visual: generating images, videos or graphics. • Audio: generating artificial speech, soundtracks, or narration. • Textual: writing creative, professional or technical texts. The tools rely on large databases and machine learning to generate content that simulates human output. Examples: DALL-E (images), Resemble AI (voice), Jasper or ChatGPT (text).
AI Agent	An AI system that operates autonomously or semi-autonomously to perform tasks or achieve defined goals while interacting with its environment. Generally, the term refers to the combination of a language model along with an implementation that allows it to perform actions – in an agent system, there is also a routing agent that refers tasks to different agents according to the tools they have.
RAG or CAG technique (Retrieval-Augmented Generation/Cache-Augmented Generation)	Approaches in which users can provide an AI model with a specific data set that should be taken into account, while processing the question that was asked. This allows the user to receive more accurate, current, and detailed answers to complex questions.
Off-the-shelf AI-based tools	AI-based tools, software, or platforms that are ready to use "out-of-the-box." These are general solutions suitable for broad and common uses, and in some situations, they allow a degree of customization.

[13] Also see ICT Glossary.
[14] Principles for Policy, Regulation and Ethics in the field of AI

Term	Definition
Conversational search engines	Tools that combine traditional search engines with natural language processing (NLP) technologies, allowing users to have a natural and intuitive conversation with the system to search for information. Instead of settling for rule-based results, the engine responds conversationally, explains the information, and offers advanced context based on understanding complex queries. Examples: ChatGPT, Microsoft Bing Chat.
"Personal assistant" applications with generative AI capabilities	Programs or apps designed to help users manage their daily lives or business tasks. Using generative AI capabilities, they can plan meetings, suggest ideas, provide tailored insights, and even create content on demand. Examples: Siri, Google Assistant, Microsoft Copilot.
Meeting transcription and summary	Tools that use Automatic Speech Recognition (ASR) and natural language processing technologies to accurately transcribe conversations or meetings and to produce concise, clear summaries. They include advanced features such as speaker identification, speaker separation, noise filtering, transcription, key insights, and generating actionable tasks from the conversation. Examples: Otter.ai, Microsoft Teams transcription, Notion AI.
Prompt	An instruction given to a generative AI system. The prompt is written in text and is used to carry out a dialogue with the system.
AI Incident	An event in which an (expected or unexpected) AI risk has materialized.