



**RESPONSIBLE
AI INNOVATION IN
LAW ENFORCEMENT**
AI Toolkit

Responsible AI Innovation in Action Workbook



Funded by
the European Union

REVISED FEBRUARY 2024



unicri
United Nations
Interregional Crime and Justice
Research Institute



Funded by
the European Union

DISCLAIMER

The contents of this document are for information purposes only. INTERPOL and UNICRI assume no liability or responsibility for any inaccurate or incomplete information, nor for any actions taken in reliance thereon. The published material is distributed without warranty of any kind, either express or implied, and the responsibility for the interpretation and use of the material lies with the reader. In no event shall INTERPOL or UNICRI be liable for damages arising from its use.

INTERPOL and UNICRI take no responsibility for the content of any external website referenced in this publication or for any defamatory, offensive or misleading information which might be contained on these third-party websites. Any links to external websites do not constitute an endorsement by INTERPOL or UNICRI, and are only provided as a convenience. It is the responsibility of the reader to evaluate the content and usefulness of information obtained from other sites.

The views, thoughts and opinions expressed in the content of this publication belong solely to the authors and do not necessarily reflect the views or policies of, nor do they imply any endorsement by, INTERPOL or the United Nations, their member countries or member states, their governing bodies, or contributory organizations, nor does it imply any endorsement. Therefore, INTERPOL and UNICRI carry no responsibility for the opinions expressed in this publication.

INTERPOL and UNICRI do not endorse or recommend any product, process, or service. Therefore, mention of any products, processes, or services in this document cannot be construed as an endorsement or recommendation by INTERPOL or UNICRI.

The designation employed and presentation of the material in this document do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations, UNICRI or INTERPOL, concerning the legal status of any country, territory, city or area of its authorities, or concerning the delimitation of its frontiers or boundaries.

The contents of this document may be quoted or reproduced, provided that the source of information is acknowledged. INTERPOL and UNICRI would like to receive a copy of the document in which this publication is used or quoted.

OVERVIEW

WHAT

This workbook is designed to support law enforcement agencies in planning, developing or procuring, using, and monitoring their AI systems responsibly. It contains structured exercises to help agencies along the path towards responsible AI innovation. It may also serve as a basis for documenting and supporting the implementation and application of the best practices explored in the Principles for Responsible AI Innovation.

WHEN

This workbook is designed to help guide brainstorming sessions at every stage of the AI life cycle.

WHO

This workbook is designed for use by the team within the agency that is responsible for the AI project. It is recommended that such a team should include staff who understand the technical, ethical and legal issues related to the project.

Table of Contents

DISCLAIMER	1
OVERVIEW	2
Where to Start?	5
Factsheet: The AI Life Cycle	6
The AI Project	8
Planning Stage	9
EXERCISE 1: PRINCIPLES IN ACTION DURING PLANNING	10
EXERCISE 2: STAKEHOLDER ENGAGEMENT	18
EXERCISE 3: GAPS & NEEDS ANALYSIS	20
EXERCISE 4: VALUE MAPPING	21
EXERCISE 5: DECIDING HOW TO OBTAIN THE AI SYSTEM	22
EXERCISE 6: IDENTIFYING POSSIBLE MISUSES OR UNAUTHORIZED USES	26
EXERCISE 7: RISK RESPONSE	27
Design & Development Stage	30
EXERCISE 1: PRINCIPLES IN ACTION DURING DESIGN AND DEVELOPMENT	31
EXERCISE 2: STAKEHOLDER ENGAGEMENT	40
EXERCISE 3: DATA CHECK	41
EXERCISE 4: KEY ASPECTS OF DESIGN & DEVELOPMENT	43
EXERCISE 5: RISK MONITORING	43

Procurement Stage **46**

EXERCISE 1: PRINCIPLES IN ACTION DURING PROCUREMENT	46
EXERCISE 2: STAKEHOLDER ENGAGEMENT	55
EXERCISE 3: PROCUREMENT PROCESS CHECKLIST	56
EXERCISE 4: RISK MONITORING	57

Use & Monitoring Stage **60**

EXERCISE 1: PRINCIPLES IN ACTION DURING USE & MONITORING	61
EXERCISE 2: STAKEHOLDER ENGAGEMENT	71
EXERCISE 3: DEPLOYMENT PROCESS CHECKLIST	72
EXERCISE 4: RISK MONITORING	73

Where to Start?

Implementing and maintaining AI innovation practices that align with the principles for responsible AI innovation involves understanding and applying the principles, identifying and engaging with the relevant stakeholders, checking the results, and restarting if needed – as many times as necessary throughout the AI life cycle. While this may seem daunting, agencies should keep in mind that the path to responsible AI innovation will be travelled one step at a time. ▶ *Learn more about this in the “Putting the Principles into Practice” section of the **Principles for Responsible AI Innovation** and “The Right Mindset” in the **Organizational Roadmap**.*

This workbook offers a framework to support the teams within a law enforcement agency who are in charge of an AI project as they travel this path. It can serve as a basis for reflection, decision-making and record-keeping at each stage of the AI life cycle.

The suggested approach to using the workbook is to:

1. Identify or designate the team or staff members in charge of the AI project within the law enforcement agency.
2. Fill out the key details in the AI project cover sheet for record-keeping purposes.
3. Identify which stage of the AI life cycle you are at.
4. Complete the exercises provided for the relevant stage, together with (internal or external) experts in the relevant areas (for example, law, ethics, AI). **Ensure that each exercise is either filled out or signed off by the most appropriate unit in the law enforcement agency.**
5. Implement any necessary changes in the AI project resulting from the completed activities.
6. Repeat the exercise if necessary to continue aligning the AI project with the principles for responsible AI innovation.
7. Check off the completed exercises using the table in the AI project cover sheet for record-keeping purposes.
8. Give the workbook to the person or team responsible for verifying and approving the decision.

This suggested pathway can apply to *any stage of the AI life cycle* which the project team/staff members wish to focus on. In other words, as the AI project evolves, the workbook can be passed between the relevant teams/staff members within the agency, and the process suggested above may be repeated at each stage.

However, it is important to note that law enforcement agencies are encouraged to engage with the workbook in a way that best suits the needs and the processes they have in place. **Some agencies or units may already have processes in place that overlap with the exercises included here, or that fulfil a similar function, and the workbook is not designed to replace or subsume these processes. It also does not in any way intend to substitute or override any legal obligations.** This workbook should be read as part of the AI Toolkit as a whole. It does not provide an exhaustive description of all the activities that agencies may need to engage in to observe the principles for responsible AI innovation, and neither does it specify any obligations which may be contained in the laws, regulations and procedures that apply to these activities.

Factsheet: The AI Life Cycle

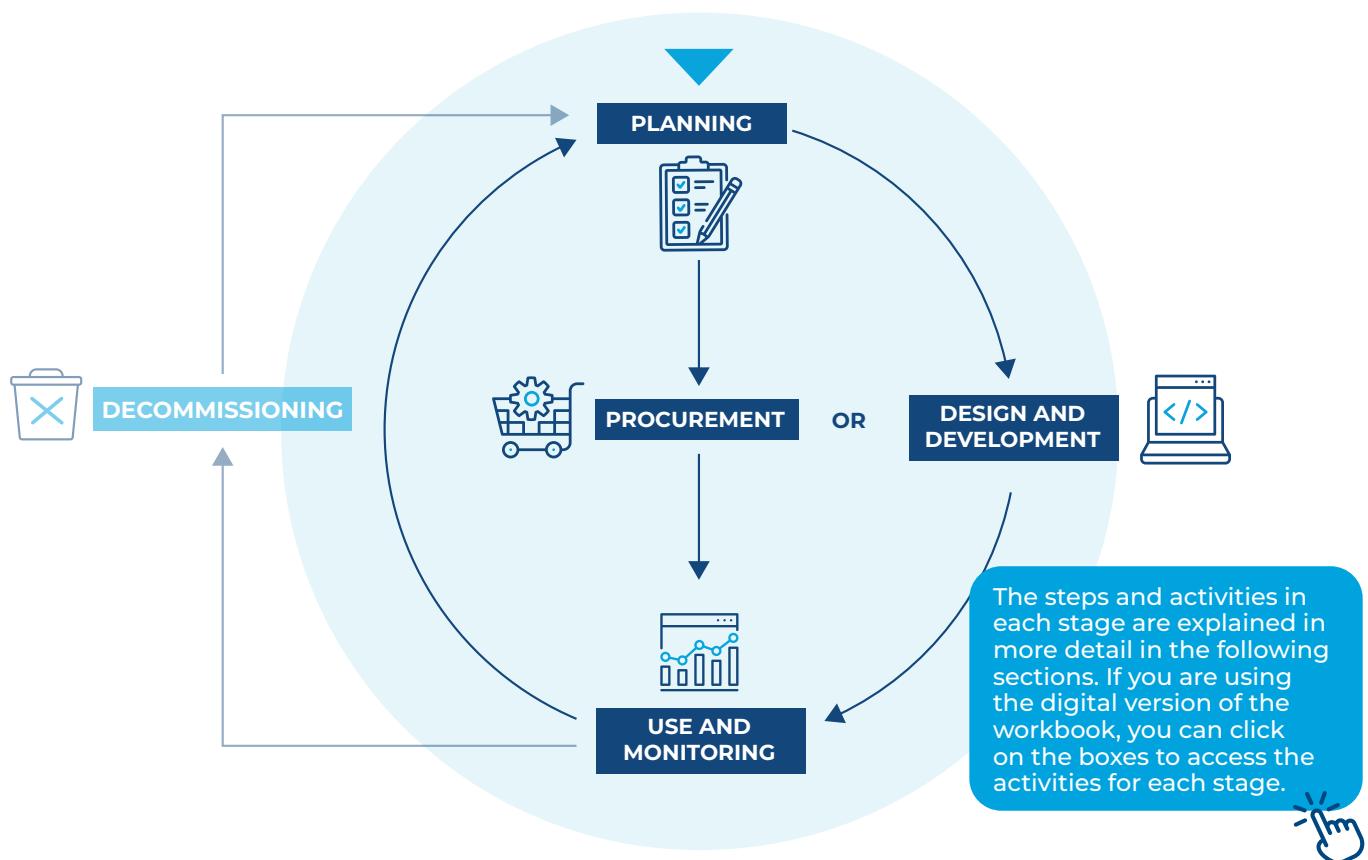


Figure 1 - The life cycle of an AI system

The AI life cycle involves all stages of an AI project, from the conceptualization of an AI system to its use and beyond. The term 'cycle' is used because it is a continuous process in which the stages are repeated iteratively until the AI system is no longer in use. While the specific configuration of the AI life cycle may differ between AI systems, it generally includes three main stages: (1) planning; (2) development / procurement; (3) use & monitoring.

1. During **PLANNING**, law enforcement agencies identify and understand the problem they wish to address and establish whether implementing an AI system is an adequate response to that problem.
2. If the agency decides to produce the AI system in-house, the planning stage is followed by the **DESIGN & DEVELOPMENT** stage. During this stage, the team in charge carries out the suggested activities for designing and developing the AI system.

If the agency decides to outsource the development of an AI system or buy it 'off-the-shelf', the planning stage will be directly followed by the **PROCUREMENT** stage. During this stage, the team in charge will follow the necessary processes to acquire the AI system from a third party.

3. The **USE & MONITORING** stage consists of implementing the AI system in the intended context and observing and supervising its use.

If the monitoring process shows that the AI system is malfunctioning or does not adhere to the principles for responsible AI innovation, the team(s) in charge can either decide to improve it, which leads them back to the planning stage, or to stop using it – i.e. **DECOMMISSION** it, following the appropriate processes. ► *Learn more about this in "The Right Mindset" section of the **Organizational Roadmap**.*

After decommissioning, the agency may decide to initiate a new AI life cycle.

The AI Project

KEY DETAILS	
PROJECT NAME	
START DATE	
PROJECT MANAGER	
SIGNATURE	
TYPE OF DEVELOPMENT	<input type="checkbox"/> Internal <input type="checkbox"/> Hybrid <input type="checkbox"/> External <input type="checkbox"/> Off-the-shelf
PROJECT DESCRIPTION	

Have all exercises for this stage been completed?	COMPLETED BY (Signature and Date)	CHECKED BY (Name of person responsible & Date)
1. PLANNING		
2. DESIGN & DEVELOPMENT (when applicable)		
3. PROCUREMENT (when applicable)		
4. USE & MONITORING		

Planning Stage

The planning stage is the first stage of the AI life cycle. During this stage, agencies discuss their needs, the scope of the AI system, and the available options for obtaining such a system, including whether to develop it internally or procure it from a third party.

The planning stage is an extremely valuable opportunity for law enforcement agencies to align their AI project with ethical standards and human rights. It can be helpful to take the **Principles for Responsible AI Innovation** into account at this stage in order to ensure that they are considered throughout the following steps. Exercise 1 is an ideal starting point for putting these principles into practice.

During this stage, law enforcement agencies are also encouraged to conduct a requirement gathering exercise. The requirement gathering phase set out in the following activities includes a gap/needs analysis which should be completed before an agency decides to develop an AI system internally, procure it, or engage with a technology partner to develop the system jointly. This is because these requirements will determine the design, development and procurement processes. Requirement gathering sessions should answer the most relevant questions about an AI system before design or procurement begins.

The following exercises are designed to guide you through the planning stage in order to maximize the usefulness of the project and optimize your AI strategy, in line with the principles for responsible AI innovation:

Exercise 1

PRINCIPLES IN ACTION DURING PLANNING

Exercise 2

STAKEHOLDER ENGAGEMENT

Exercise 3

GAPS & NEEDS ANALYSIS

Exercise 4

VALUE MAPPING

Exercise 5

DECIDING HOW TO OBTAIN THE AI SYSTEM

Exercise 6

USE CASE ANALYSIS

Exercise 7

IDENTIFYING POSSIBLE MISUSES OR UNAUTHORIZED USES

Exercise 8

RISK RESPONSE

If you are using the digital version of the workbook, you can click on the boxes to access the exercises directly.



EXERCISE 1: PRINCIPLES IN ACTION DURING PLANNING

Use the questions in this exercise to assess activities during the planning stage according to each of the principles for responsible AI innovation. This will help with identifying strengths and weaknesses and next steps, and record keeping.

GENERAL		LAWFULNESS *
		<ol style="list-style-type: none">1. What laws and regulations apply in the context of the AI project? What requirements and limitations do they impose that should be considered at this stage?2. What measures are in place to ensure and verify compliance with the applicable laws and regulations?
		<ol style="list-style-type: none">3. Do any activities in the planning stage interfere with human rights? If so, specify which activities interfere with human rights, which rights are at stake and who is affected.
<i>If the answer to the previous question is yes:</i>		<ol style="list-style-type: none">3.1 Does the law authorize interference with human rights? If so, identify the legal basis.3.2 What is the purpose of the interference? Is it legitimate, according to the law?

LEGITIMACY, PROPORTIONALITY & NECESSITY

- 3.3 Is the interference necessary to achieve the legitimate purpose?
- 3.4 Is the interference the least intrusive way to achieve the legitimate purpose?
- 3.5 Is there a fair balance between the interference and the legitimate purpose?

* These questions do not, in any way, substitute the need for a full and independent legal assessment. Answering these questions requires consulting internal or external legal experts that can conduct a careful assessment of legal and regulatory compliance and perform a human rights impact assessment.

MINIMIZATION OF HARM**ROBUSTNESS & SAFETY**

1. In what ways may the AI system be unreliable or unsecure? What risks and harms could result from this?
2. What risks may the AI system pose to the safety of people and the environment, including those resulting from dual use, misuse or malicious use?
3. What measures can be put in place to eliminate, reduce, or mitigate the risks and harms identified?

ACCURACY

4. What level of accuracy should the AI system have in order to be accepted?

5. What risks and harms could result from an inaccurate AI system? What measures can be put in place to eliminate, reduce, or mitigate them?

HUMAN AND ENVIRONMENTAL WELL-BEING

6. How might the development or use of the AI system impact human well-being (for example, the physical or mental health of individuals, levels of employment, trust in systems, socio-political stability, criminal justice systems)?

7. How might the development or use of the AI system impact the environment?

8. What measures can be put in place to eliminate, reduce, or mitigate the negative impacts?

EFFICIENCY

9. What is the expected added value of this AI system? Are there other, less complex, less risky, and/or less costly ways of achieving the same goals?

10. What is the expected cost of this AI system (consider all of the resources required to develop, use, and maintain the system)?

HUMAN AUTONOMY	
HUMAN CONTROL & OVERSIGHT	<ol style="list-style-type: none">1. Will the AI system make decisions autonomously? If so, do the outputs of the system have the potential to affect the lives of individuals? Is there a guarantee that humans can intercept or evaluate the output?2. What are the foreseen measures and gaps in terms of human control and oversight?
HUMAN AGENCY	<ol style="list-style-type: none">3. Will the AI system support or augment people's ability to make decisions?4. Can the AI system be used to manipulate and/or control individual behaviour?5. Will the users of the AI system receive all the necessary information and training to remain independent in their decisions?
PRIVACY	<ol style="list-style-type: none">6. What are the foreseen effects on individuals' privacy at all stages of the AI life cycle?7. Will the development of the AI system entail access to personal data? If so, will data subjects be able to control their personal data as per the applicable law?8. Will the AI system access personal data during use? If so, what safeguards will be put in place to ensure that privacy and data protection rights are respected?

TRANSPARENCY & EXPLAINABILITY	<p>9. What mechanisms will be put in place to ensure that the AI system's purpose, capabilities, and limitations will be clear to users and those who are subjected to the system?</p> <p>10. Will the system be developed in such a way as to be explainable? If so, how?</p> <p>11. What are the risks of a lack of transparency and explainability in the system? What mitigation or corrective strategies are in place to address these risks?</p>
-------------------------------	--

EQUALITY & NON-DISCRIMINATION	FAIRNESS
	<p>1 How are "fairness" and "unfair bias" defined in relation to the AI system? What are the relevant categories of unfair bias?</p>
PROTECTION OF VULNERABLE GROUPS	<p>2. Where might unfair bias be produced in the AI system?</p> <p>3. How are people protected against the adverse effects of potential unfair bias?</p> <p>4. How is 'vulnerability' defined in the context of the AI project? Are there any vulnerable groups potentially affected by the AI system? If so, who are the potentially vulnerable groups?</p> <p>5. What are the positive and negative effects of introducing the AI system for vulnerable groups?</p>

DIVERSITY & ACCESSIBILITY	<ol style="list-style-type: none">6. How is 'diversity' defined in the context of the AI project? Which diversities are taken into account?7. What accessibility concerns are related to the introduction of the AI system?
CONTESTABILITY & REDRESS	<ol style="list-style-type: none">8. What challenges might users and those affected by the AI system encounter if they wish to challenge its decisions or obtain compensation for related harms?9. What technological measures could be put in place to allow users and those affected by the AI system to challenge a decision?10. What organizational measures could be put in place to allow users and those affected by the AI system to challenge a decision?

GOOD GOVERNANCE	
TRACEABILITY & AUDITABILITY	<ol style="list-style-type: none">1. How will the traceability of the AI system's processes and user interaction be ensured? Are there mechanisms in place to document the ethical and technical decisions made throughout the system's development and use?2. What are the available mechanisms that could be used to ensure the auditability of the AI system?3. What internal and external audit cycles are in place and what do they entail?
ACCOUNTABILITY	<ol style="list-style-type: none">4. What mechanisms are available to ensure that accountability is integrated into the AI system's design?5. Is the scope of the responsibility of relevant officers or teams clearly defined?

CONCLUSIONS AND ASSESSMENT

If possible, the following section should be assigned to personnel who were not involved in answering the previous questions.

Use this section to reflect on the answers to the previous questions and conclude whether each principle is sufficiently fulfilled. Based on your conclusions, determine the steps recommended to ensure the AI project aligns with the principles for responsible AI innovation. This section also serves the purpose of record-keeping and communication between the different teams and personnel involved in the AI project.

CONCLUSIONS & ASSESSMENT

1 – WORKING TEAM

Is the principle of **LAWFULNESS** sufficiently fulfilled? Provide a justification.

Suggested next steps:

Is the principle of **MINIMIZATION OF HARM** sufficiently fulfilled? Provide a justification.

Is the principle of **HUMAN AUTONOMY** sufficiently fulfilled? Provide a justification.

Is the principle of **FAIRNESS** sufficiently fulfilled? Provide a justification.

Is the principle of **GOOD GOVERNANCE** sufficiently fulfilled? Provide a justification.

2 – MANAGEMENT DECISION

- Proceed
- Do not proceed
- Take steps to improve the following principle(s):

Signature:

Date:

EXERCISE 2: STAKEHOLDER ENGAGEMENT

The use of an AI system has two sets of stakeholders – those responsible for the use of the system and those that may be directly or indirectly affected by its use. A clear overview of all stakeholders will help to better understand exactly who is affected and if and how this impact is even or uneven. ▶ Learn more about this in the “Putting the Principles into Practice” section of the **Principles for Responsible AI Innovation**.

STEP 1: IDENTIFYING THE STAKEHOLDERS

To help you identify the stakeholders, the exercise below has been divided into two broad categories: direct and indirect stakeholders. **Direct stakeholders** actively use, engage, or interact with an AI system, while **indirect stakeholders** are those affected by or associated with its use. Identifying the stakeholders is an important way of evaluating the impact of the system.

Reflect upon and answer the questions below:

DIRECT STAKEHOLDERS	
END USERS	<ul style="list-style-type: none">• What individual or team will be directly and actively using this AI system?• If the system's decision requires a level of interpretation or analysis, what individual or team will be responsible for this process?
DEVELOPERS	<ul style="list-style-type: none">• If an AI system is being developed internally, who is involved in the design and development of the system?• If it is developed in conjunction with a partner or procured off the shelf, who are the other parties involved?
PROGRAM MANAGER & SYSTEM ADMIN	<ul style="list-style-type: none">• Who is the individual in charge of this project and who authorizes the use of the system?• Who will be in charge of ensuring that this system is integrated with other existing technology infrastructure?
MONITORING TEAM	<ul style="list-style-type: none">• What team is responsible for managing, adjusting, or monitoring the use of this AI system through the development, deployment and use?• What team is responsible for decommissioning the AI system?
INDEPENDENT AUDITORS	<ul style="list-style-type: none">• Who are the potential independent auditors who may check the reliability of this system?• Who are the potential independent auditors who may check for compliance with the principles for responsible AI innovation?

INDIRECT STAKEHOLDERS

REGULATORS

- Are there international, national, and regional frameworks regulating the development and use of the AI system?
- Is there an obligation for this system to be licensed or inspected by the authorities?

INTEREST GROUPS

- Who are the interest groups, professional bodies, legal advisors, human rights defenders, privacy advocates or civil society groups that may be interested in and/or likely to challenge the use of this AI system or call for its regulation?

INDIVIDUALS AND COMMUNITIES

- Who is likely to be affected by the use of this AI system in the short or long term?
- Are there any specific subgroups that may be disproportionately affected by the use of this AI system?
- Are there any individuals whose personal data will be used to train and test the AI system?

STEP 2: INVOLVING THE STAKEHOLDERS

Once you have identified the stakeholders, set out a plan to engage with them throughout the AI project.

EXERCISE 3: GAPS & NEEDS ANALYSIS

A gaps and needs analysis involves comparing actual outcomes with desired outcomes in relation to the law enforcement agency's needs. In other words, for agencies to understand their needs and whether they can be fulfilled with the help of an AI system, they must first assess their current performance without AI and the potential or desired result of the integration of AI.

This process involves four steps:

1. Identifying the **current/actual performance**.
2. Identifying the **desired result**.
3. Identifying the **gaps/needs to be filled**.
4. Identifying **workable solutions**.

Fill out the following diagram to conduct your own gaps and needs analysis.



EXERCISE 4: VALUE MAPPING

Value mapping is an exercise that enables law enforcement agencies to assess the worth/value of an AI technology for their specific needs. It involves identifying and aligning the potential benefits and costs of implementing the technology, considering factors such as increased efficiency, cost savings, competitive advantage, and overall agency objectives, to determine its true value proposition within the organization.

Complete the table below to conduct your value mapping exercise.

KEEP IN MIND:

- How will this AI system benefit your agency?
- Is it necessary?
- Why is it the preferred solution?
- Can the problem be addressed in a different way without the use of the AI system?

	USER PROFILE Some AI systems will require special training or authorization for users. Outline the qualifications required for the users of this system.	RELIEF What needs or requirements does this AI system address? For example, speed, cost, convenience, ease of use, or reducing stress rates. 
	GAINS Describe how this AI system will be of benefit to your agency.	ADVANTAGES List all the departments or units that would benefit from the use of this AI system. 

EXERCISE 5: DECIDING HOW TO OBTAIN THE AI SYSTEM

There are four possible approaches for a law enforcement agency to obtain an AI system:

- **internal development:** building the AI system internally
 - **external development:** outsourcing the development of the AI system entirely to a technology provider
 - **hybrid or joint development:** building the AI system jointly with a technology provider
 - **“off-the-shelf” procurement:** procuring an externally developed AI system that is ready to use.

Having established your needs and outlined the use case for the AI system in the previous exercises, you should now consider the best option for obtaining such a system. Many factors need to be weighed up in order to make this decision. Below are some questions to serve as a basis for discussion:

1. What are the available options?

2. What is the available budget?

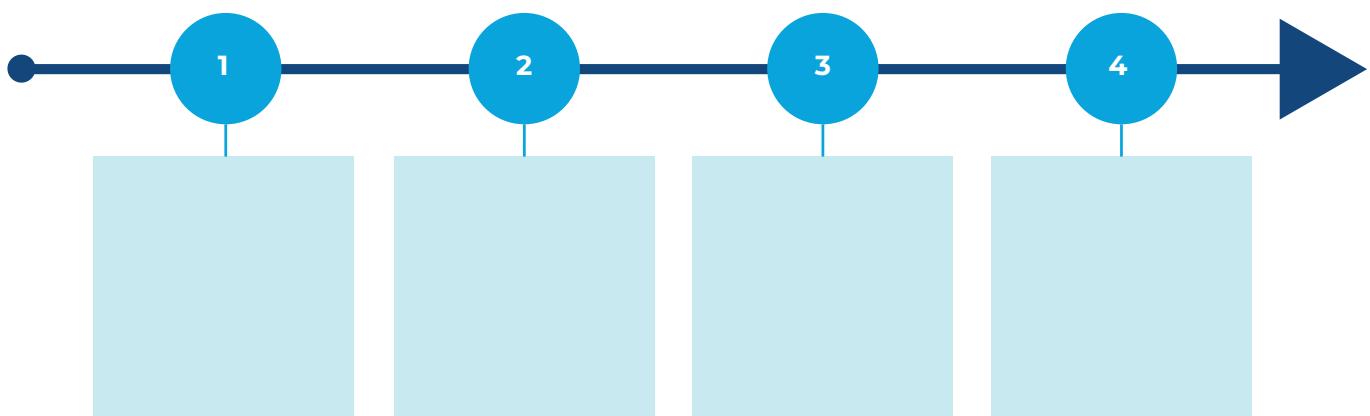
It is important to have a conversation about the budget for any AI system. You should estimate what the project will cost and whether it is within your team's available budget. It is possible that during the development and implementation of an AI system, the system may turn out not to be useful or may fail to work as anticipated. At this point, responsible AI innovation requires you to terminate the project despite the costs already incurred.

KEEP IN MIND:

- What is the estimated cost of building this AI system internally / collaborating with a technology partner / purchasing an externally developed system?
- What is the estimated cost of deployment and use of the AI system, including training, integration into the existing system, maintenance and other key necessary features?
- In each case, this should include a reflection on other options that do not include the use of an AI system (including their costs and benefits).

3. What is the timeline for the chosen option?

The template below can be adjusted to lay out initial ideas for the steps required for your projected development plan. Carefully consider the various steps required for each type of development and create a realistic schedule for your desired objectives. This should be as detailed as possible.



A projected timeline for bringing an AI system into operation is helpful for putting things into perspective and determining which type of development would best suit the anticipated time-frame and your agency's needs.

KEEP IN MIND:

- Is this timeline realistic?
- Are there other options more suitable for your current needs regarding timing?
- Is there sufficient time throughout the process to ensure that appropriate safeguards are implemented to introduce the AI system responsibly?
- Are the resources currently available sufficient for this project's ambitions within the given timeframe? (i.e., expertise, computational, financial)
- What additional resources are required to fulfill these needs?

4. What trade-offs should be considered?

Deciding between the different options is not simply a matter of whether or not the necessary funding, technical expertise and time is available to pursue a particular approach. Each of the four approaches also has a specific set of pros and cons that the agency should carefully weigh up. For example, an externally or jointly developed system can raise issues around data protection or access to relevant training data. It may also affect the ability of the law enforcement agency to ensure transparency, auditing, and explainability of the AI system used (e.g., for reasons of proprietary algorithms or confidentiality).

EXERCISE 6: IDENTIFYING POSSIBLE MISUSES OR UNAUTHORIZED USES

Misuse of the AI system consists of an improper or wrong use of the system. Unauthorized use, on the other hand, refers to use of the system without approval or official permission. Consider your preferred approach to AI development and identify any possible or potential misuse or unauthorized use of this type of AI system.

In choosing which type of AI development approach to employ, the possibility of misuse and security breaches should be a primary consideration. This is because the risks involved with each type of approach are different.

Agencies should conduct a thorough assessment of the possible threats, taking into consideration factors such as the organization's security standards, the expertise of internal staff and external vendors, and the level of customization required. Ultimately, the choice of a particular development approach should be guided by the organization's specific needs and the level of security threat which each development option may pose.

Below are some questions to guide you in completing this section. Answering these questions requires engaging subject matter aspects and carrying out an information security risk assessment, cybersecurity risk assessment or similar risk assessments.

Who are the possible unauthorized users?

How could they gain access to the AI system? At what stage and under which circumstances?

What could happen if this AI system is accessed by an authorized user? Could the confidentiality of data be affected?

Can this AI system be misused?

What could happen if this AI system is used by a malicious end user? Could the integrity of data be affected?

What mitigation strategies can be put in place to ensure that this AI system is not accessed by unauthorized persons?

What mitigation mechanisms can be put in place to avoid misuse of this AI system?

Is another possible approach to obtain the AI system more suitable in terms of maximizing the security of the system?

EXERCISE 7: RISK RESPONSE

A crucial step in the planning stage is identifying and assessing any risks related to the AI system at a project level. These are independent processes that should be conducted as per your agency's risk management policy. ▶ *Learn more about this in the "Processes" section of the **Organizational Roadmap**.*

Accordingly, the AI Toolkit's **Risk Assessment Questionnaire** should be completed at this stage (as well as any other similar assessments in your agency's Risk Management Policy). Use the results of the **Risk Assessment Questionnaire** to define a comprehensive risk response that can adequately prevent and mitigate any risks.

KEEP IN MIND:

The various possible courses of action that can be taken to respond to risks are classified as follows: risk acceptance; risk mitigation; risk avoidance; risk monitoring; risk auditing; or risk reporting. Generally, a risk response should include a combination of several actions from different categories. ▶ Learn more about these actions in the **Risk Assessment Questionnaire**.

Details of the Risk Assessment Questionnaire	
Completed on:	Responsible person/team:
RISK RESPONSE	
1- Working Team	
What EXTREME RISKS were identified in the Risk Assessment Questionnaire?	What is the suggested risk response?
What HIGH RISKS were identified in the Risk Assessment Questionnaire?	What is the suggested risk response?
What MEDIUM RISKS were identified in the Risk Assessment Questionnaire?	What is the suggested risk response?
What LOW RISKS were identified in the Risk Assessment Questionnaire?	What is the suggested risk response?

2 - Management decision

Signature:

Date:

Design & Development Stage

If the decision is made to develop the AI system in-house, the team in charge should perform the activities and processes necessary to fulfil the principles for responsible AI innovation.

The design and development stage starts with the identification/collection and preparation of the training data in parallel with selecting the most appropriate algorithm architecture. A correct exploration of the training data and preparation for the development stage helps the data science team identify the constraints and possible errors or unfair biases in the data. In some cases, the data preparation may also involve some data wrangling or engineering, where data engineers produce a working data set and convert it into a format that can be used for modeling purposes, depending on the selected algorithm.

Once the data is ready and the algorithm selected, the next steps include modelling and evaluating the model. The developers start by carrying out an iterative training process, defining and adjusting the model settings to optimize it to best represent that data. To know if the model is good, the developers need to conduct a model evaluation, monitoring relevant metrics such as accuracy. When the model meets the performance evaluation metrics, the subsequent stage consists of assessing that the system meets the requirements of accountability, security, transparency etc. This process is essential to identifying gaps, fixing bugs, and iterating on the model to ensure it aligns with all the principles for responsible AI innovation.

|▶ Learn more about data requirements and model performance evaluation in the **Technical Reference Book**.

The following exercises will guide you through the design & development stage by encouraging you to reflect on the various details and complexities you will encounter:

Exercise 1PRINCIPLES IN ACTION
DURING DESIGN &
DEVELOPMENT**Exercise 2**STAKEHOLDER
ENGAGEMENT**Exercise 3**

DATA CHECK

Exercise 4KEY ASPECTS OF
DESIGN &
DEVELOPMENT**Exercise 5**RISK
MONITORING

If you are using the digital version of the workbook, you can click on the boxes to access the exercises directly.



EXERCISE 1: PRINCIPLES IN ACTION DURING DESIGN AND DEVELOPMENT

Use the questions in this exercise to assess activities during the design and development stage according to each of the principles for responsible AI innovation. This will help with identifying strengths and weaknesses and next steps, and record keeping.

KEEP IN MIND:

- You do not need to start from scratch if you have completed Exercise 1 during the planning stage. Use the questions in this exercise to reassess your previous answers with the new information available.

LAWFULNESS *	
GENERAL	<ol style="list-style-type: none">1. What laws and regulations apply in the context of the design and development of the AI system? What requirements and limitations do they impose that should be considered at this stage?2. What measures are in place to ensure and verify compliance with the applicable laws and regulations?
LEGITIMACY, PROPORTIONALITY & NECESSITY	<ol style="list-style-type: none">3. Do any activities in the design and development stage interfere with human rights? If so, specify which activities interfere with human rights, which rights are involved and who is affected. <p><i>If the answer to the previous question is yes:</i></p> <ol style="list-style-type: none">3.1 Does the law authorize interference with human rights? If so, identify the legal basis.3.2. What is the purpose of the interference? Is it legitimate according to the law?3.3. Is the interference necessary to achieve the legitimate purpose?3.4. Is the interference the least intrusive way to achieve the legitimate purpose?3.5. Is there a fair balance between the interference and the legitimate purpose?

* These questions do not, in any way, substitute the need for a full and independent legal assessment. Answering these questions requires consulting internal or external legal experts that can conduct a careful assessment of legal and regulatory compliance and perform a human rights impact assessment.

MINIMIZATION OF HARM	
ROBUSTNESS & SAFETY	<ol style="list-style-type: none">What measures are in place to ensure the AI system's security in the case of attacks or unexpected situations and environments?What measures are in place to ensure the AI system is reliable when using different inputs and in different contexts?What mechanisms are in place to eliminate, reduce, or mitigate the risks the AI system poses to the safety of people and the environment, including those resulting from dual use, misuse or malicious use?
ACCURACY	<ol style="list-style-type: none">How is the accuracy of the AI system measured? Is its consistency across demographics verified? Is this element submitted to independent third-party testing?Is the training data accurate, representative of the population, up-to-date and adequate for the intended purpose?Is the accuracy rate sufficient for the intended use of the AI system? What measures are in place to eliminate, reduce, or mitigate the negative consequences of any inaccurate results?

HUMAN AND ENVIRONMENTAL WELL-BEING	<p>7. How might the AI system under development impact human well-being, both positively and negatively? What safeguards are in place to reduce any negative impacts on human well-being?</p>
EFFICIENCY	<p>8. How might the AI system under development affect the environment, both positively and negatively? What safeguards are in place to reduce any negative effects on the environment?</p> <p>9. How does the AI system's efficiency compare with other AI systems, or solutions that do not involve the use of AI?</p> <p>10. Is the AI system achieving the defined goal efficiently in terms of financial, time, and human and environmental resources? Will it maintain such efficiency when used?</p>

HUMAN AUTONOMY	
HUMAN CONTROL & OVERSIGHT	<ol style="list-style-type: none">1. What is the appropriate level of human control for the AI system?2. What measures are in place to ensure meaningful and informed human control and oversight of the system?3. Is there a stop / abort command that can be safely employed in a timely manner by a human operator if needed?
HUMAN AGENCY	<ol style="list-style-type: none">4. What measures are in place to ensure that the AI system provides relevant and sufficient information for independent human decision-making?5. What measures are in place to ensure that the AI system does not manipulate or deceive individuals in their interaction with the system or in their decision-making?

PRIVACY

6. What measures are taken to ensure data minimization, security and other personal data protection principles regarding the training data?
7. Who can access the training data? How is it safeguarded from unauthorized access?
8. What measures are taken to ensure data minimization, security and other personal data protection principles regarding the data to be processed by the AI system when in use?
9. Who will be able to access the data that will be processed by the AI system? How is it safeguarded from unauthorized access? Is there a mechanism in place for flagging up privacy breaches?
10. What privacy enhancing features have been incorporated into the AI system (encryption, anonymization, etc.)?

TRANSPARENCY &
EXPLAINABILITY

11. Is the AI system being developed in such a way as to make it explainable? How? Are there mechanisms in place to inform users of the reasons and criteria behind the AI system's results?
12. Is the AI system transparent for users in terms of the way it will affect their decision-making? Are its purposes, capabilities and limitations clear to users and those affected by the system's use?

FAIRNESS	
EQUALITY & NON-DISCRIMINATION	1.What are the possible sources of unfair bias in the data processing, design or development of the AI system? Is there a risk of such unfair bias being reflected in the system's outputs? 2 What is the potential negative impact of such biased outputs and who would be at risk as a result? What mechanisms are in place to protect those individuals or groups? 3.What mechanisms are in place at this stage to avoid the use of the AI system creating or reinforcing unfair bias?
PROTECTION OF VULNERABLE GROUPS	4. What are the positive and negative effects of the AI system's introduction on vulnerable groups?
DIVERSITY & ACCESSIBILITY	5 Is the AI system designed to accommodate a wide range of individual preferences and abilities? How does the AI system affect individuals with varying types and levels of ability?
CONTESTABILITY & REDRESS	6. Is the AI system designed to allow accessible, efficient, and effective contestability and redress?

GOOD GOVERNANCE	
TRACEABILITY & AUDITABILITY	1. How is the AI system's development documented?
ACCOUNTABILITY	2. Does the AI system's design include features that ensure traceability of its use?
	3. Have mechanisms been established to facilitate auditing by independent third parties?
	4. What steps have been taken to ensure that the AI system is designed with accountability in mind?
	5. What is the mechanism used to document ethical, legal and technical decisions throughout the development process for the AI system?
	6. Is the scope of relevant officers' or teams' responsibility clearly defined?

CONCLUSIONS AND ASSESSMENT

If possible, the following section should be assigned to personnel who were not involved in answering the previous questions.

Use this section to reflect on the answers to the previous questions and conclude whether each principle is sufficiently fulfilled. Based on your conclusions, determine the steps recommended to ensure the AI project aligns with the principles for responsible AI innovation. This section also serves the purpose of record-keeping and communication between the different teams and personnel involved in the AI project.

CONCLUSIONS & ASSESSMENT

1 – WORKING TEAM

Is the principle of **LAWFULNESS** sufficiently fulfilled? Provide a justification.

Suggested next steps:

Is the principle of **MINIMIZATION OF HARM** sufficiently fulfilled? Provide a justification.

Is the principle of **HUMAN AUTONOMY** sufficiently fulfilled? Provide a justification.

Is the principle of **FAIRNESS** sufficiently fulfilled? Provide a justification.

Is the principle of **GOOD GOVERNANCE** sufficiently fulfilled? Provide a justification.

2 – MANAGEMENT DECISION

- Proceed
- Do not proceed
- Take steps to improve the following principle(s):

Signature:

Date:

EXERCISE 2: STAKEHOLDER ENGAGEMENT

STEP 1: IDENTIFYING STAKEHOLDERS

Did you identify the direct or indirect stakeholders during the planning stage?

*If yes, you should now reassess your answers with the new information available at this stage.
If no, complete step 1) of the stakeholder engagement exercise included in the planning stage.*

STEP 2: INVOLVING THE STAKEHOLDERS

Did you involve the direct or indirect stakeholders during the planning stage?

*If yes, you should now reassess if and how the direct and indirect stakeholders should be involved.
If no, plan how you will engage them throughout the AI project.*

[Click here to access Exercise 2 of the planning stage](#)



EXERCISE 3: DATA CHECK

Reflect on and answer the questions below. These are some of the crucial aspects to consider when building a model to ensure that it is aligned with the principles for responsible AI innovation. These questions can help you address upcoming challenges and guide you in determining the next steps on your responsible AI journey.

DATA SELECTION	DATA PRIVACY	DATA EXPLORATION
<ul style="list-style-type: none">• Who collected the data and how was it collected?• Is this data relevant to the purpose of the system? Can this data reply to the question proposed to be replied to by this system?• Was the data collected in a lawful manner?• Was the data collected in an ethical manner?• Is the data transparent and well-documented?• Can the quality of the data be properly assessed? If so, can a quality benchmark be determined?• How reliable is this data? In terms of completeness, and potential sources of error or uncertainty.• Is this data set representative of diverse perspectives?• Does the data require wrangling or cleaning?	<ul style="list-style-type: none">• Are there any data privacy concerns?• Does this data contain sensitive information?• Was personal data permitted to be used for developing the model in question?• Was personal data collected under a valid ground for processing under the applicable law?• Is personal data anonymized? If not, are there any requirements to provide access to the data upon request of the data subject?• Are there security measures in place to effectively manage and comply with data protection regulations? If so, outline them.• Are there any concerns relating to data transfers?• How long and under what safeguards will the data be stored?	<ul style="list-style-type: none">• What are the key features and their relevance to the problem?• Are there missing values, outliers, or inconsistencies?• Which variables are most correlated to the target variable?• What are the limitations of this data?• Has bias been detected in the data set and how will this be addressed?• How might external factors influence the data?• How might the data evolve or change over time?• Is the analysis and data processing well-documented?• Are there clear records of data transformations and pre-processing steps?• Can insights derived from the data be explained?

KEEP IN MIND:

- Machine learning algorithms are usually trained using historical data. Given the nature of law enforcement, the type of data required to build AI systems with machine learning may be considered sensitive. Sometimes, the data may include non-public records on crime or personal data which would be used in building a repository. Any use of these types of data could potentially expose the law enforcement agency to legal and information security risks, even if the training and test data sets have been anonymized. One example where the data is extremely sensitive is when building AI systems to detect child sexual abuse or exploitation. Machine learning models would traditionally need to be trained using sensitive images, which can pose legal issues.
 - This would become much more complicated if the AI system were developed by external technology providers, as they would need to access these extremely sensitive data sets, which would be considered a crime in multiple jurisdictions. There are techniques that can be implemented to solve these issues, such as using synthetic data to train the algorithms, but the issues must be fully understood if adequate solutions are to be adopted. ► *Learn more about algorithms and data in the **Technical Reference Book**.*

EXERCISE 4: KEY ASPECTS OF DESIGN & DEVELOPMENT

Note down relevant aspects of the design and training phase, as well as the testing performed either internally in the lab or under real conditions after integration into the law enforcement infrastructure. This will help you to gather key information to share with stakeholders for transparency purposes. Below are some prompts to guide you in completing this activity.

MODELLING/ TRAINING	LAB TESTING	FIELD TESTING
<ul style="list-style-type: none"> Identify the challenges during model training and explain how you addressed them. What type of data was used for training? What types of models or ensemble of models were used? 	<ul style="list-style-type: none"> What was the result of your internal testing? Do the lab conditions simulate the real conditions? Does the model achieve its intended purpose? Which performance metrics were used? Does the model's performance require improvement? 	<ul style="list-style-type: none"> What were the results of the field testing? Is the model similarly accurate and precise when tested using real data? What resources are required to run the model? Does the model's performance require improvement?

EXERCISE 5: RISK MONITORING

Before you start: Did you carry out the **Risk Assessment Questionnaire** and complete Exercise 7 (Risk Response) in the planning stage?

If yes, you can proceed with this risk monitoring exercise. If no, you should first perform the **Risk Assessment Questionnaire** and complete Exercise 7 in the planning stage.

[Click here to access Exercise 7 of the planning stage](#)



STEP 1:

REEVALUATING THE RISK ASSESSMENT QUESTIONNAIRE AND THE RISK RESPONSE

Risk monitoring should be an ongoing process throughout the AI life cycle. Accordingly, you should now reevaluate the results of the **Risk Assessment Questionnaire** and Exercise 7 (Risk Response) of the planning stage and update them with the new information available at this stage.

KEEP IN MIND:

- Have any risk-affecting changes been identified in the AI system's use case or in the environment in which it is intended to operate? If so, what are these changes and which new risk responses do they require?
 - Look at the **Risk Assessment Questionnaire** completed during the planning stage. Are the purposes, scope and constraints of the **Risk Assessment Questionnaire** the same, or do they need to be updated?
 - Are the actions outlined in the Risk Response exercise still being implemented properly? What has their impact been? Do they need to be reconsidered?
 - Repeat the **Risk Assessment Questionnaire** as necessary. What main changes in the risk levels have been identified? What new risk responses do they require?

STEP 2: ADJUSTING THE COURSE OF ACTION

Once you have revisited the **Risk Assessment Questionnaire** and the Risk Response exercises carried out during the planning stage and updated your answers, you may need to reconsider the actions that are necessary for preventing or mitigating risks in the future.

KEEP IN MIND:

- Outline any new risk response actions which may be needed. If necessary, complete the tables in the Risk Response exercise again and pass on the information to the management.
- Identify any future circumstances in which the **Risk Assessment Questionnaire** will need to be updated.
- Communicate any relevant updates to the responsible personnel in your agency.
- Set out a plan to decommission the AI system should this become necessary.

Procurement Stage

During this stage, the team or staff members responsible for procurement should follow the applicable processes in the agency or unit. At this stage, those in charge should request relevant information regarding the AI system's functions and limitations from the technology providers, as well as the results of various assessments such as accuracy, robustness, and bias tests.

The following exercises are suggested for putting responsible AI innovation into practice during the procurement stage:

Exercise 1

PRINCIPLES IN ACTION DURING PROCUREMENT

Exercise 2

STAKEHOLDER ENGAGEMENT

Exercise 3

PROCUREMENT PROCESS CHECKLIST

Exercise 4

RISK MONITORING

If you are using the digital version of the workbook, you can click on the boxes to access the exercises directly.



EXERCISE 1: PRINCIPLES IN ACTION DURING PROCUREMENT

Use the questions in this exercise to assess activities in the procurement stage according to each of the principles for responsible AI innovation. This will help with the identification of strengths and weaknesses, next steps, record keeping, etc.

KEEP IN MIND:

- You do not need to start from scratch if you have completed Exercise 1 during the planning stage. Use the questions in this exercise to reassess your previous answers with the new information available.

GENERAL		LAWFULNESS *
1. What laws and regulations apply in the context of the procurement of the AI system? What requirements and limitations do they impose that should be considered at this stage?		
2. What measures are in place to ensure and verify compliance with the applicable laws and regulations?		
3. Do any activities in the procurement stage interfere with human rights? If so, specify which activities interfere with human rights, which rights are involved and who is affected.		
<i>If the answer to the previous question is yes:</i>		
3.1. Does the law authorize interference with human rights? If so, identify the legal basis.		
3.2. What is the purpose of the interference? Is it legitimate according to the law?		
3.3. Is the interference necessary to achieve the legitimate purpose?		
3.4. Is the interference the least intrusive way to achieve the legitimate purpose?		
3.5. Is there a fair balance between the interference and the legitimate purpose?		

* These questions do not, in any way, substitute the need for a full and independent legal assessment. Answering these questions requires consulting internal or external legal experts that can conduct a careful assessment of legal and regulatory compliance and perform a human rights impact assessment.

MINIMIZATION OF HARM

ROBUSTNESS & SAFETY

1. Can the vendors produce test results for the AI system's reliability, security and/or safety? Was this testing carried out internally or by independent third parties?

2. What mechanisms did the vendors put in place to identify, eliminate, reduce, or mitigate the risks that the AI system may pose to the safety of each stakeholder?

3. Are there additional mechanisms that need to be put in place by the agency to ensure the reliability, security and safety of the AI system during use? Which mechanisms? Are they sufficient to minimize harm?

4. Can the vendors demonstrate the accuracy of the AI system and its consistency across demographics? Has the AI system's accuracy been submitted to independent third-party testing?

5. Have the vendors provided information on the data used to train the AI system? Does the training data correctly represent the population where the AI system will be used?

6. Have the vendors provided information on possible variations in accuracy depending on the conditions in which the AI system is used? What mechanisms will the agency put in place to ensure proper use of the AI system?

7. Is the accuracy rate sufficiently high considering the intended use of the AI system?

ACCURACY

HUMAN AND ENVIRONMENTAL WELL-BEING	<p>8. How could the AI system being procured affect human well-being, both positively and negatively? What safeguards are in place to reduce any negative effects on human well-being?</p> <p>9. How might the AI system being procured affect the environment, both positively and negatively? What safeguards are in place to reduce any negative effects on the environment?</p>
EFFICIENCY	<p>10. How does the efficiency of procuring and using the AI system, in terms of financial, time, and human and environmental resources, compare with other AI systems or solutions that do not involve AI?</p>

HUMAN CONTROL & OVERSIGHT	HUMAN AUTONOMY
	<p>1. What is the adequate level of human control for the AI system being procured?</p> <p>2. What measures are in place to ensure meaningful and informed human control and oversight of the system? Can the vendors demonstrate these measures?</p> <p>3. Is there a stop / abort command on the AI system being procured that can be used by a human operator safely and in a timely manner if needed?</p>

HUMAN AGENCY

4. What measures are in place to ensure that the AI system provides relevant and sufficient information for independent human decision-making?

5. What measures are in place to ensure that the AI system's outputs do not negatively influence individuals' interaction with the system or their decision-making?

6. What measures have vendors taken to ensure data minimization, security and other personal data protection principles regarding the training data?

7. What measures have been/will be taken to ensure data minimization, security and other personal data protection requirements regarding the data that will be processed by the AI system?

PRIVACY

8. Who will be able to access the data that will be processed by the AI system? How is it safeguarded from unauthorized access? Is there a mechanism for flagging up privacy breaches?

9. Which privacy enhancing features have been incorporated into the AI system (encryption, anonymization, etc.)?

TRANSPARENCY & EXPLAINABILITY

10. Has the AI system been developed in such a way as to make it explainable? Can vendors demonstrate this?

11. What are the risks of a lack of transparency and explainability in the system? What mitigation or corrective strategies are in place to address these risks?

12. Is the AI system transparent for users in terms of the way it will affect their decision-making? Are its purposes, capabilities and limitations clear to users and those affected by the system?

FAIRNESS**EQUALITY &
NON-DISCRIMINATION**

1. What potential sources of unfair bias might have arisen in the course of the development of the AI system being procured, and how might they affect the AI system's outputs? Have the vendors informed the users of any potential issues?

2. What could be the negative impact of such biased outputs and who would be at risk as a result? What mechanisms are in place to protect these individuals or groups?

3. What mechanisms are in place at this stage to avoid the use of the AI system creating or reinforcing unfair bias?

PROTECTION OF VULNERABLE GROUPS	<p>4. What vulnerable groups might be affected by the AI system, and how? Is there a mechanism in place to ensure that vulnerable groups are not harmed by the system, directly or indirectly?</p>
DIVERSITY & ACCESSIBILITY	<p>5. Has the AI system been designed to accommodate a wide range of individual preferences and abilities?</p> <p>6. How might the AI system affect individuals with varying types and levels of abilities? What measures are in place to prevent or minimize any negative impact?</p>
CONTESTABILITY & REDRESS	<p>7. Has the AI system been designed to allow accessible, efficient, and effective contestability and redress? If not, can measures be put in place to that effect?</p>

GOOD GOVERNANCE	
TRACEABILITY & AUDITABILITY	<ol style="list-style-type: none">1. Was the AI system's development documented? Have the vendors made that documentation available?2. Does the AI system's design include features that allow for traceability of its use?3. Have mechanisms been established to facilitate auditing by independent third parties?
ACCOUNTABILITY	<ol style="list-style-type: none">4. Has the AI system been designed to ensure accountability? How? Can the vendors demonstrate this?5. Is the scope of relevant law enforcement officers' or teams' responsibility for adopting the AI system clearly defined?

CONCLUSIONS AND ASSESSMENT

If possible, the following section should be assigned to personnel who were not involved in answering the previous questions.

Use this section to reflect on the answers to the previous questions and conclude whether each principle is sufficiently fulfilled. Based on your conclusions, determine the steps recommended to ensure the AI project aligns with the principles for responsible AI innovation. This section also serves the purpose of record-keeping and communication between the different teams and personnel involved in the AI project.

CONCLUSIONS & ASSESSMENT

1 – WORKING TEAM

Is the principle of **LAWFULNESS** sufficiently fulfilled? Provide a justification.

Suggested next steps:

Is the principle of **MINIMIZATION OF HARM** sufficiently fulfilled? Provide a justification.

Is the principle of **HUMAN AUTONOMY** sufficiently fulfilled? Provide a justification.

Is the principle of **FAIRNESS** sufficiently fulfilled? Provide a justification.

Is the principle of **GOOD GOVERNANCE** sufficiently fulfilled? Provide a justification.

2 – MANAGEMENT DECISION

- Proceed
- Do not proceed
- Take steps to improve the following principle(s):

Signature:

Date:

EXERCISE 2: STAKEHOLDER ENGAGEMENT

STEP 1) IDENTIFYING STAKEHOLDERS

Did you identify the direct and indirect stakeholders during the planning stage?

*If yes, you should now reassess your answers with the new information available at this stage.
If no, complete step 1) of the stakeholder engagement exercise included in the planning stage.*

[Click here to access Exercise 2 of the planning stage](#)



STEP 2) INVOLVING THE STAKEHOLDERS

Did you involve the direct or indirect stakeholders during the planning stage?

*If yes, you should now reassess if and how the direct and indirect stakeholders should be involved.
If no, plan how you will engage them throughout the AI project.*

EXERCISE 3: PROCUREMENT PROCESS CHECKLIST

Reflect on and answer the questions below. These are some of the key aspects to consider when aligning the procurement process with the principles for responsible AI innovation. These questions can help you decide on a specific vendor and determine the next steps.

SELECTION OF TECHNOLOGY PARTNERS

1. Is the vendor selection in line with appropriate national licensing and accreditation requirements?
2. Are there any contractual or other restrictions (such as intellectual property or confidentiality) in the arrangement with a potential vendor which may affect the ability to audit algorithms, test data, or other code upon demand and at a granular level?
3. What criteria have been set for the vendor to assess bias and other relevant risks?
4. Are technology partners/vendors required to present the results of any evaluations conducted during testing to help assess the AI system's performance?
5. Will the vendor be providing training for users? Is this training covered by the vendor's contractual obligations?
6. Has a standardized procurement process been created/followed?
7. Have the following aspects been completed?
 - Review of purchase requisition
 - Implementation of the solicitation process
 - Evaluation of vendors
 - Awarding the contract
 - Order management
 - Invoice approvals and record keeping
 - Other internal impact assessments (i.e., of Human Rights or Data Protection)

PROCUREMENT PROCESS

ALGORITHMIC AUDIT & SUPPORT

8. Are there procedures in place to ensure that the algorithms used in AI systems procured off the shelf can be independently audited? If so, explain these procedures.
9. Will testing be carried out using appropriate test standards (lab tests and, if possible, field tests)?
10. Are there procedures in place to ensure that regular software upgrades are included in the vendor's/service provider's contractual obligations?

EXERCISE 4: RISK MONITORING

Before you start: Did you carry out the **Risk Assessment Questionnaire** and complete Exercise 7 (Risk Response) in the planning stage?

*If yes, you can proceed with this risk monitoring exercise. If no, you should first perform the **Risk Assessment Questionnaire** and complete Exercise 7 in the planning stage.*

[Click here to access Exercise 7 of the planning stage](#)



STEP 1:

REEVALUATING THE RISK ASSESSMENT QUESTIONNAIRE AND THE RISK RESPONSE

Risk monitoring should be an ongoing process throughout the AI life cycle. Accordingly, you should now reevaluate the results of the **Risk Assessment Questionnaire** and Exercise 7 (Risk Response) of the planning stage and update them with the new information available at this stage.

KEEP IN MIND:

- Have any risk-affecting changes been identified in the AI system's use case or in the environment in which it is intended to operate? If so, what are these changes and which new risk responses do they require?
- Look at the **Risk Assessment Questionnaire** completed during the planning stage. Are the purposes, scope and constraints of the **Risk Assessment Questionnaire** the same, or do they need to be updated?
- Are the actions outlined in the Risk Response exercise still being implemented properly? What has their impact been? Do they need to be reconsidered?
- Repeat the **Risk Assessment Questionnaire** as necessary. What main changes in the risk levels have been identified? What new risk responses do they require?

STEP 2: ADJUSTING THE COURSE OF ACTION

Once you have revisited the **Risk Assessment Questionnaire** and Risk Response exercises carried out during the planning stage and updated your answers, you may need to reconsider the actions that are necessary for preventing or mitigating risks in the future.

KEEP IN MIND:

- Outline any new risk response actions which may be needed. If necessary, complete the tables in the Risk Response exercise again and pass on the information to the management.
 - Identify any future circumstances in which the **Risk Assessment Questionnaire** will need to be updated.
 - Communicate any relevant updates to the responsible personnel in your agency.
 - Set out a plan to decommission the AI system should this become necessary.

Use & Monitoring Stage

The use and monitoring stage begins with the deployment of the AI system. The deployment of the AI system includes all activities needed to make it available for use, including software installation and setup.

This stage covers the entire period during which the AI system is available, regardless of whether it is being used or not.

As well as using the AI system for its intended purpose, this stage involves overseeing and controlling the AI system and its use. Monitoring allows end users, developers, auditors and system administrators to keep an eye out for errors or bugs and correct them to improve the AI system's predictions and precision (i.e., improving databases to reduce bias).

The following exercises are suggested for putting responsible AI innovation into practice during the use & monitoring stage:

Exercise 1

PRINCIPLES IN ACTION DURING USE & MONITORING

Exercise 2

STAKEHOLDER ENGAGEMENT

Exercise 3

DEPLOYMENT PROCESS CHECKLIST

Exercise 4

RISK MONITORING

If you are using the digital version of the workbook, you can click on the boxes to access the exercises online.



EXERCISE 1: PRINCIPLES IN ACTION DURING USE & MONITORING

Use the questions in this exercise to assess activities in the procurement stage according to each of the principles for responsible AI innovation. This will help with the identification of strengths and weaknesses, next steps, record keeping, etc.

KEEP IN MIND:

- You do not need to start from scratch if you have completed Exercise 1 during the previous stage(s). Use the questions in this exercise to reassess your previous answers with the new information available.

GENERAL		LAWFULNESS *
		<ol style="list-style-type: none">1. What laws and regulations apply in the context of the use and monitoring of the AI system? What requirements and limitations do they impose that should be considered at this stage?2. What measures are in place to ensure and verify compliance with the applicable laws and regulations?

LEGITIMACY, PROPORTIONALITY & NECESSITY

3. Do any activities in the use and monitoring stage interfere with human rights? If so, specify which activities interfere with human rights, which rights are involved and who is affected.

If the answer to the previous question is yes:

3.1 Does the law authorize interference with human rights? If so, identify the legal basis.

3.2. What is the purpose of the interference? Is it legitimate according to the law?

3.3. Is the interference necessary to achieve the legitimate purpose?

3.4. Is the interference the least intrusive way to achieve the legitimate purpose?

3.5. Is there a fair balance between the interference and the legitimate purpose?

* These questions do not, in any way, substitute the need for a full and independent legal assessment. Answering these questions requires consulting internal or external legal experts that can conduct a careful assessment of legal and regulatory compliance and perform a human rights impact assessment.

MINIMIZATION OF HARM

ROBUSTNESS & SAFETY

1. What processes are in place for monitoring and testing the robustness and safety of the AI system, including flagging up misuse or malicious use?

2. What is the strategy for managing any failure of the AI system? Is there a backup plan?

3. Are users trained to monitor the AI system's reliability?

4. Could the AI system pose a risk for users and other stakeholders? How are these safety risks communicated and what is the process for reporting them?

5. What measures are in place for testing and monitoring accuracy during use?

6. How is the AI system's accuracy communicated to users and other stakeholders?

7. What is the process for flagging up accuracy issues? Are users trained to identify and flag up inaccuracy issues?

ACCURACY

HUMAN AND ENVIRONMENTAL WELL-BEING**EFFICIENCY**

8. What strategies are in place for assessing and safeguarding individual and environmental well-being during use of the AI system?

9. In terms of well-being, what are the risks that users and other stakeholders should be aware of? How are these risks communicated and how can they be reported?

10. What mechanisms are in place for continuously measuring the AI system's efficiency in terms of financial, time, and human and environmental resources?

11. Are users trained to measure the AI system's efficiency and to use it in the most efficient way possible?

HUMAN CONTROL & OVERSIGHT**HUMAN AUTONOMY**

1. How and when are users or other stakeholders able to interfere with the AI system's processes?

2. Who is responsible for controlling and overseeing the use of the AI system? What mechanisms are in place for this? How and how often is their effectiveness verified?

3. How is the AI system assessed for any adverse effects? Is this assessment made by independent parties?

HUMAN AGENCY

4. What are individuals' roles in deploying and using the AI system? Do they have a proper understanding of the AI system they are interacting with?
5. Are there mechanisms in place to enable users or other stakeholders to challenge or reasonably disagree with the AI system's outputs?
6. Are there mechanisms in place to enable stakeholders to request information about the AI system?
7. What mechanisms are in place to assess to what extent users rely on the AI system and how it affects their decision making? Are there measures in place to prevent overreliance on the system?

PRIVACY

8. How does the AI system impact the stakeholders' privacy (including users)? How is this communicated to stakeholders?
9. What mechanisms are in place for monitoring the impact on privacy and assessing any privacy violations? Is there a mechanism for reporting privacy issues?
10. What mechanisms are in place to ensure the security of the data which is fed into and processed by the AI system?

TRANSPARENCY & EXPLAINABILITY

11. Are the individuals affected by the outcomes of the AI system aware that an AI system is being used?

12. Are the users or the individuals affected by the outcomes of the AI system able to understand the criteria behind these results?

13. How are the limitations and strengths of the AI system communicated to stakeholders? Are they able to understand them?

14. What mechanisms are in place to assess the system's transparency and explainability? Is there a process in place to provide feedback in this regard?

FAIRNESS	
EQUALITY & NON-DISCRIMINATION	<ol style="list-style-type: none">1. Are users aware of any potential biases in the AI system? How are these potential biases communicated to other stakeholders?2. What mechanisms are in place for testing and monitoring potential bias in the AI system?3. What mechanisms are in place for preventing any biases in the AI system having a negative impact on individuals and groups?4. What mechanisms are in place for flagging up any issues of unfair bias in the AI system?
PROTECTION OF VULNERABLE GROUPS	<ol style="list-style-type: none">5. Who is most at risk of harm from the use of this AI system? What are these risks and how are they communicated to stakeholders?6. What are the positive and negative effects of introducing the AI system for vulnerable groups?7. What mechanisms are in place to monitor the positive and negative effects of the AI system on vulnerable groups?8. What safeguards and mitigation measures are in place to protect vulnerable groups from any negative effects? What are the processes that can be used to flag up any issues?

DIVERSITY & ACCESSIBILITY	<p>9. What mechanisms are in place for assessing the usability of the AI system by individuals with varying types and levels of ability?</p> <p>10. What mechanisms are in place for monitoring the accessibility of the AI system? Is there a mechanism in place for flagging up any accessibility issues?</p>
CONTESTABILITY & REDRESS	<p>11. What mechanisms are in place to enable stakeholders to contest the AI system's outputs?</p> <p>12. Are the mechanisms for contestability and redress monitored? If so, how and how often?</p>

GOOD GOVERNANCE	
TRACEABILITY & AUDITABILITY	<p>1. What measures are in place to monitor traceability mechanisms?</p> <p>2. Is there a mechanism in place for logging user behaviour?</p> <p>3. Has the AI system's audit been completed? Is there a routine audit mechanism in place?</p> <p>4. How often are audits conducted?</p> <p>5. What process has been set out for implementing the results and feedback from audits?</p>

ACCOUNTABILITY

6. Is the scope of relevant law enforcement officers' or teams' responsibility for using and monitoring the AI system clearly defined?

7. What mechanisms are in place for assessing any incidents related to the use of the AI system and how will responsibility be assumed in such cases?

CONCLUSIONS AND ASSESSMENT

If possible, the following section should be assigned to personnel who were not involved in answering the previous questions.

Use this section to reflect on the answers to the previous questions and conclude whether each principle is sufficiently fulfilled. Based on your conclusions, determine the steps recommended to ensure the AI project aligns with the principles for responsible AI innovation. This section also serves the purpose of record-keeping and communication between the different teams and personnel involved in the AI project.

CONCLUSIONS & ASSESSMENT

1 – WORKING TEAM

Is the principle of **LAWFULNESS** sufficiently fulfilled? Provide a justification.

Suggested next steps:

Is the principle of **MINIMIZATION OF HARM** sufficiently fulfilled? Provide a justification.

Is the principle of **HUMAN AUTONOMY** sufficiently fulfilled? Provide a justification.

Is the principle of **FAIRNESS** sufficiently fulfilled? Provide a justification.

Is the principle of **GOOD GOVERNANCE** sufficiently fulfilled? Provide a justification.

2 – MANAGEMENT DECISION

- Proceed
- Do not proceed
- Take steps to improve the following principle(s):

Signature:

Date:

EXERCISE 2: STAKEHOLDER ENGAGEMENT

STEP 1: IDENTIFYING STAKEHOLDERS

Did you identify the direct and indirect stakeholders during the previous stage(s)?

*If yes, you should now reassess your answers with the new information available at this stage.
If no, complete step 1) of the stakeholder engagement exercise included in the planning stage.*

STEP 2: INVOLVING THE STAKEHOLDERS

Did you involve the direct or indirect stakeholders during the previous stage(s)?

*If yes, you should now reassess if and how the direct and indirect stakeholders should be involved.
If no, plan how you will engage them throughout the AI project.*

EXERCISE 3: DEPLOYMENT PROCESS CHECKLIST

Reflect on and answer the questions below. These are some of the key aspects to consider when aligning the deployment process with the *principles for responsible AI innovation*. These questions can help you prepare to make the AI system available for use.

SYSTEM

1. How would the deployment of this AI system be conducted; would it be a phased approach?

2. Identify the step-by-step deployment plan for this AI system.

3. Is there any unique technological infrastructure or AI system which must be used alongside this AI system? If so, explain what.

4. Identify the key team responsible for ensuring the correct deployment of this AI system.

5. Identify the team responsible for the monitoring and control of this AI system.

6. Identify all possible independent auditors of this AI system.

KEY ACTORS

UPSKILLING

7. Is training for end-users and other teams essential?

8. If yes, outline the kind of training, upskilling or certification required. This may include cloud certifications, data analytics, data visualization, etc.

EXERCISE 4: RISK MONITORING

Before you start: Did you carry out the **Risk Assessment Questionnaire** and complete Exercise 7 (Risk Response) in the previous stage(s)?

If yes, you can proceed with this risk monitoring exercise. *If no*, you should first perform the Risk Assessment Questionnaire and complete Exercise 7 in the planning stage.

STEP 1:**REEVALUATING THE RISK ASSESSMENT QUESTIONNAIRE AND THE RISK RESPONSE**

Risk monitoring should be an ongoing process throughout the AI life cycle. Accordingly, you should now reevaluate the results of the **Risk Assessment Questionnaire** and Exercise 7 (Risk Response) of the previous stage(s) and update them with the new information available at this stage.

KEEP IN MIND:

- Have any risk-affecting changes been identified in the AI system's use case or in the environment in which it is intended to operate? If so, what are these changes and which new risk responses do they require?
- Look at the Risk Assessment Questionnaire completed during the planning stage. Are the purposes, scope and constraints of the **Risk Assessment Questionnaire** the same, or do they need to be updated?
- Are the actions outlined in the Risk Response exercise still being implemented properly? What has their impact been? Do they need to be reconsidered?
- Repeat the **Risk Assessment Questionnaire** as necessary. What main changes in the risk levels have been identified? What new risk responses do they require?

STEP 2: ADJUSTING THE COURSE OF ACTION

Once you have revisited the **Risk Assessment Questionnaire** and Risk Response exercises carried out during the planning stage and updated your answers, you may need to reconsider the actions that are necessary for preventing or mitigating risks in the future.

KEEP IN MIND:

- Outline any new risk response actions which may be needed. If necessary, complete the tables in the Risk Response exercise again and pass on the information to the management.
- Identify any future circumstances in which the **Risk Assessment Questionnaire** will need to be updated.
- Communicate any relevant updates to the responsible personnel in your agency.
- Set out a plan to decommission the AI system should this become necessary.



How to cite this publication: UNICRI and INTERPOL. (Revised February 2024). Toolkit for Responsible AI Innovation in Law Enforcement: **Responsible AI Innovation in Action Workbook**.

© United Nations Interregional Crime and Justice Research Institute (UNICRI), 2024

© International Criminal Police Organization (INTERPOL), 2024

2025 Update: This concerns a correction to minor errors contained in the previous version.



www.interpol.int
www.unicri.it



INTERPOL_HQ



@INTERPOL_HQ
@UNICRI



INTERPOL HQ
UNICRI



INTERPOL
UNICRI



@INTERPOL
@UNICRIHQ

www.ai-lawenforcement.org