



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

Responsible AI Innovation in Action Workbook



Funded by
the European Union



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 anykind, 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 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



Where to Start?	6
Factsheet: the AI Life Cycle	8
The AI Project	10
Planning Stage	11
Design & Development Stage	39
Procurement Stage	58
Use & Monitoring Stage	76

Annex: Risk Response Measures	94
--------------------------------------	-----------

A decorative graphic at the bottom of the page features a network of glowing blue nodes connected by thin lines, set against a dark blue background with scattered light blue particles.

1.

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 IN 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 the moment.
4. Complete the exercises provided for the relevant stage, together with (internal or external) experts in the relevant areas (for example, law, ethics, AI).
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.

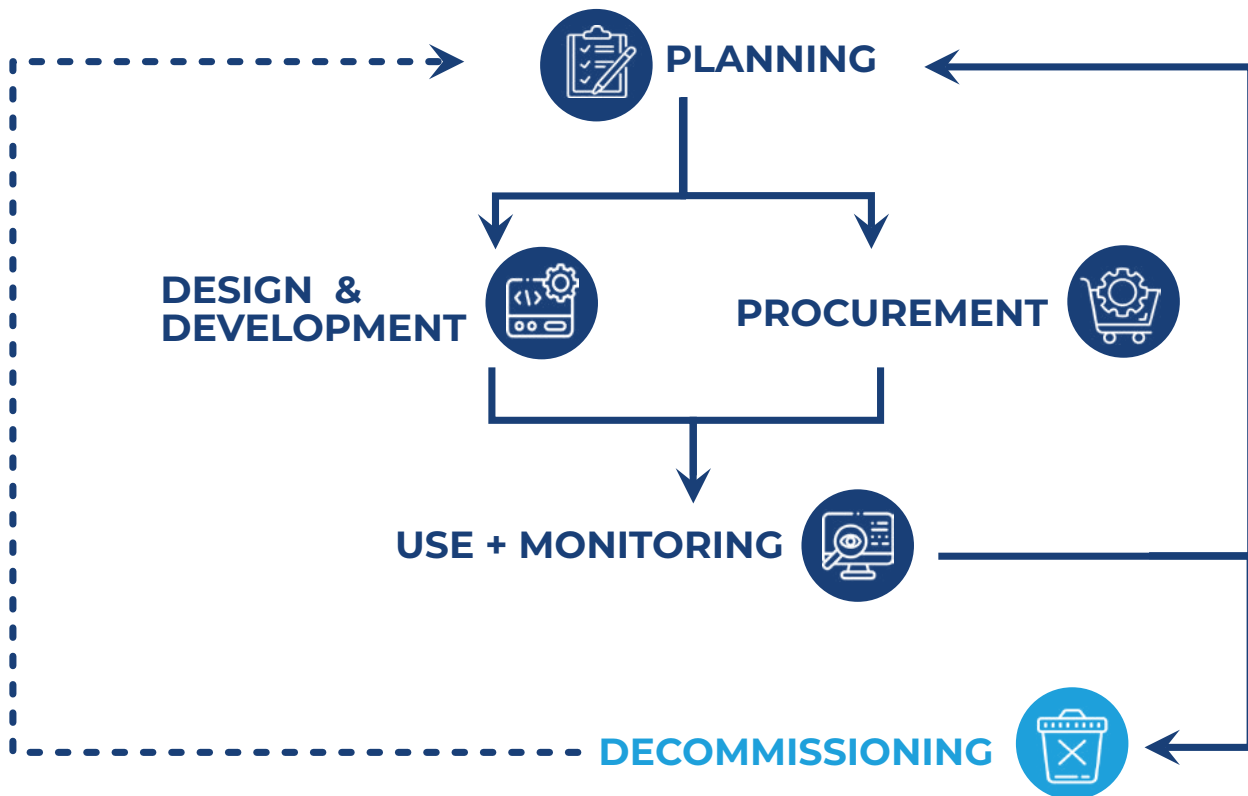
This suggested pathway can apply to *any stage of the AI life cycle* which the project team/staff members wish to focus on. In other words, as the AI project evolves, the workbook can be passed between the relevant teams/staff members within the agency, and the process suggested above may be repeated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

2.

Factsheet: the AI Life Cycle

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.



If you are using the digital version of the workbook, you can click on the boxes to access the activities for each stage

-
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 steps and activities in each stage are explained in more detail in the following sections. .

3. The AI Project

KEY DETAILS

PROJECT NAME

START DATE

PROJECT MANAGER

SIGNATURE

TYPE OF DEVELOPMENT	Internal	External
	Hybrid	Off-the-shelf

PROJECT DESCRIPTION

Have all exercises for this stage been completed?

COMPLETED BY
(SIGNATURE / DATE)

CHECKED BY (NAME OF PERSON RESPONSIBLE & DATE)

1 PLANNING

2 DESIGN & DEVELOPMENT
(when applicable)

3 PROCUREMENT
(when applicable)

4 USE & MONITORING

4. 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 and use case 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*:



If you are using the digital version of the workbook, you may 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.

LAWFULNESS *	
GENERAL	<p>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?</p> <p>2. What measures are in place to ensure and verify compliance with the applicable laws and regulations?</p>
LEGITIMACY, PROPORTIONALITY & NECESSITY	<p>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.</p> <p><i>If the answer to the previous question is yes:</i></p> <p>3.1 Does the law authorize interference with human rights? If so, identify the legal basis.</p> <p>3.2 What is the purpose of the interference? Is it legitimate, according to the law?</p>



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?



PLANNING



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	<p>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?</p> <p>2. What are the foreseen measures and gaps in terms of human control and oversight?</p>
HUMAN AGENCY	<p>3. Will the AI system support or augment people's ability to make decisions, or will it replace them in those decisions?</p> <p>4. Will the users of the AI system receive all the necessary information to make decisions independently?</p>
PRIVACY	<p>5. What are the foreseen effects on individuals' privacy at all stages of the AI life cycle?</p> <p>6. Will the development of the AI system entail access to personal data? If so, will data subjects be able to control their personal data?</p>



PRIVACY

7. 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

8. 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?

9. Will the system be developed in such a way as to be explainable? If so, how?

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

FAIRNESS	
EQUALITY & NON-DISCRIMINATION	<p>1. How are “fairness” and “unfair bias” defined in relation to the AI system? What are the relevant categories of unfair bias and where might this bias be produced?</p> <p>2. How are people protected against the adverse effects of potential unfair bias?</p>
PROTECTION OF VULNERABLE GROUPS	<p>3. How is ‘vulnerability’ defined in the context of the AI project?</p> <p>4. What are the positive and negative effects of introducing the AI system for vulnerable groups?</p>
DIVERSITY & ACCESSIBILITY	<p>5. How is ‘diversity’ defined in the context of the AI project? Which diversities are taken into account?</p> <p>6. What accessibility concerns are related to the introduction of the AI system?</p>



7. 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?

8. What technological measures could be put in place to allow users and those affected by the AI system to challenge a decision?

9. 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

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

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?



PLANNING



CONCLUSIONS & ASSESSMENT

1 – WORKING TEAM

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

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.

Suggested next steps:

2 – MANAGEMENT DECISION

Signature:

Date:



PLANNING



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.

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	DEVELOPERS
<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? 	<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?



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?

Step 2) Involving the stakeholders

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

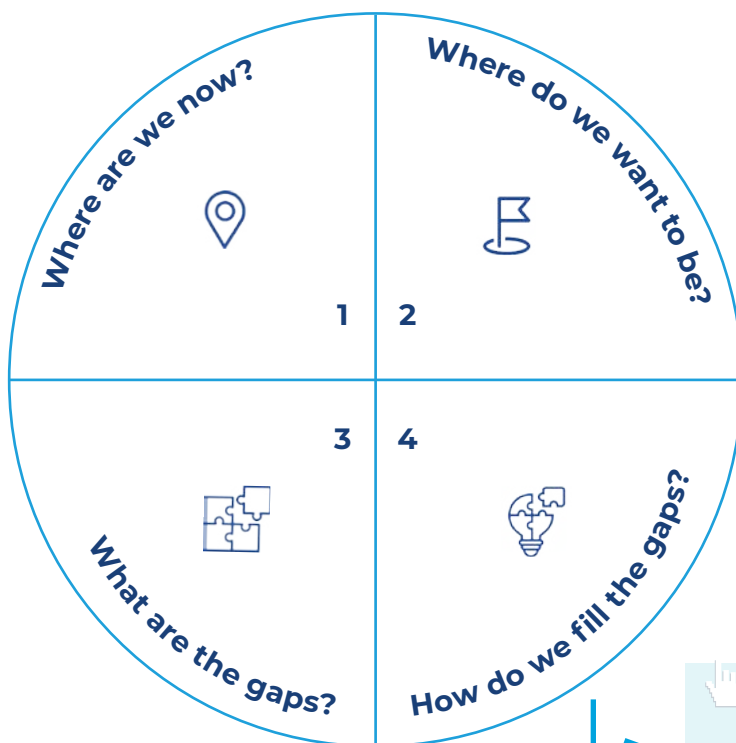
EXERCISE 3: Gaps & Needs Analysis

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



Keep in mind:

- Answering this question requires identifying and defining your use case. [Click here to access the "Use Case Analysis" exercise.](#)



PLANNING



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.



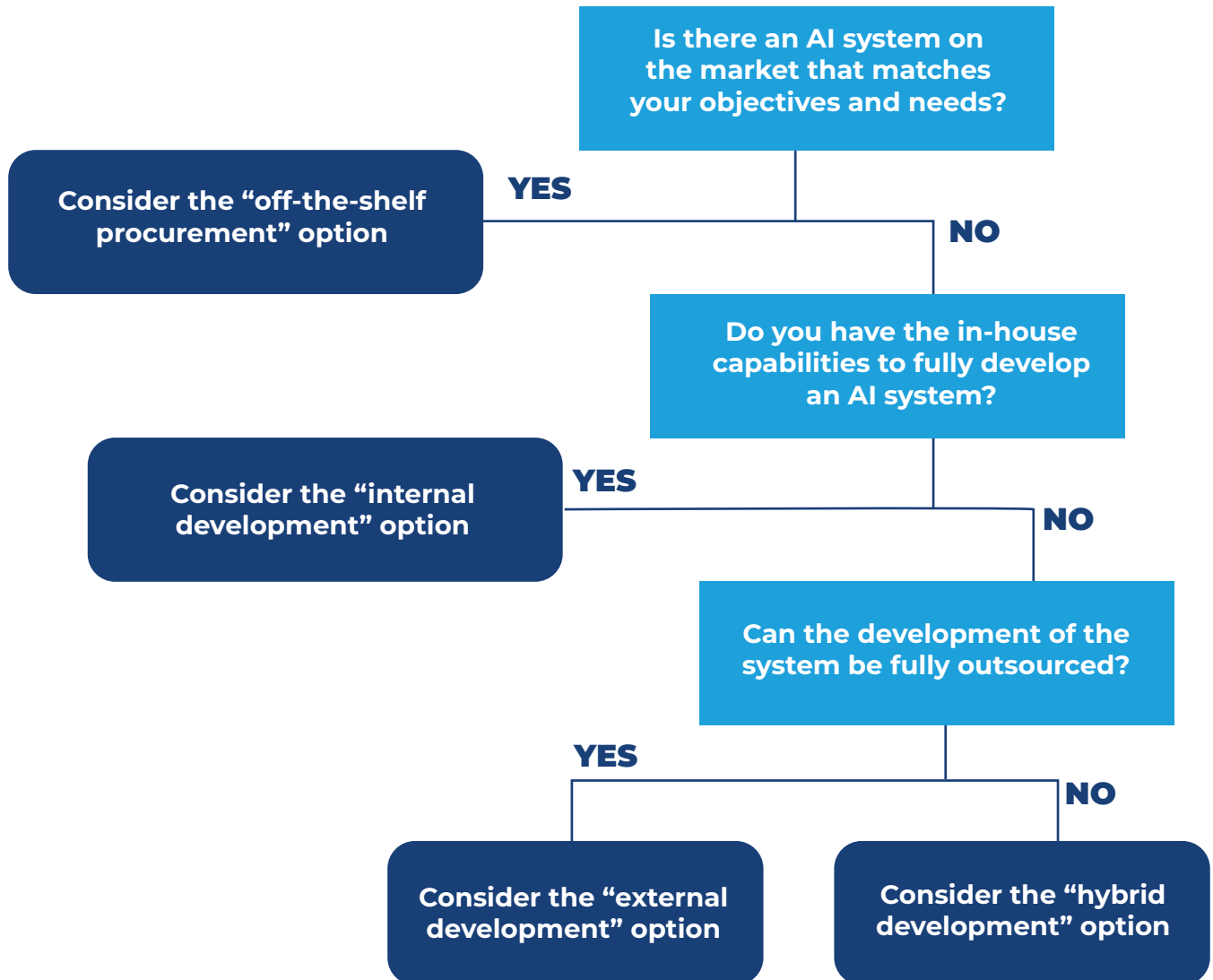
ADVANTAGES

List all the departments or units that would benefit from the use of this AI system.





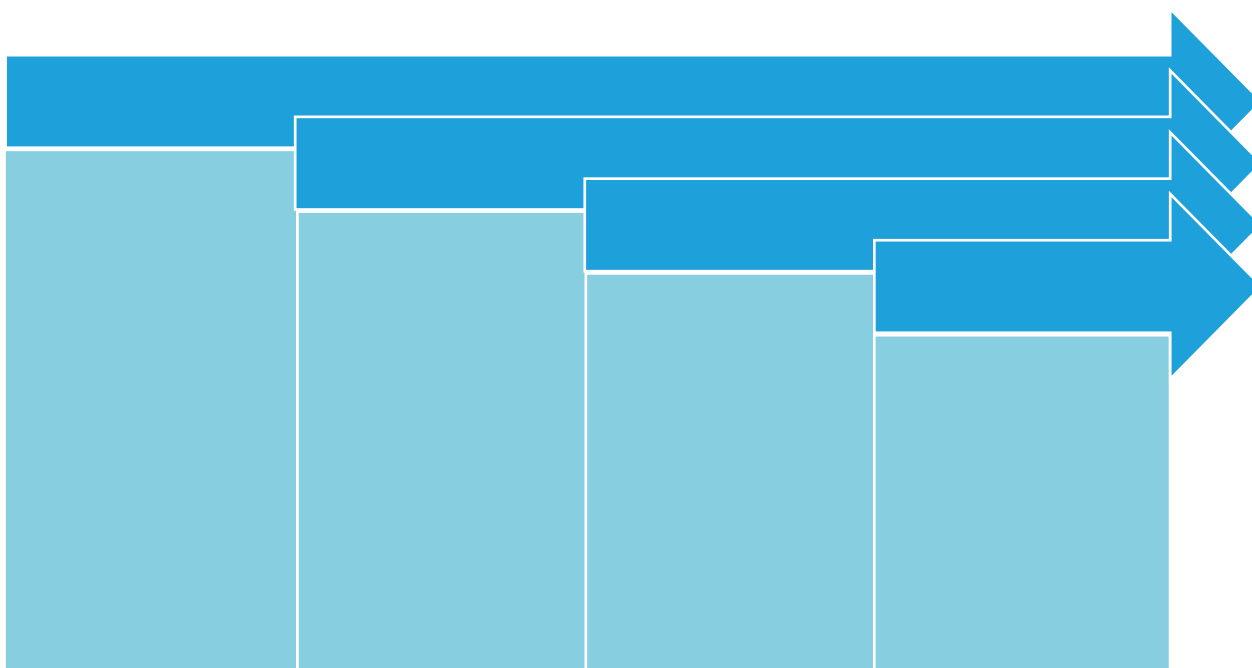
The following decision tree may help you to answer this question:





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 which would be more suitable for your current needs?
- 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?



EXERCISE 6: Use Case Analysis

After completing the gaps and needs analysis and deciding how to obtain the AI system, you should now consider a use case analysis. A use case analysis is the practice of identifying all the requirements and information needed to develop a process, system or software tool. It also includes brainstorming, defining the system’s uses and users, the conditions of its use and any constraints (ethical or legal) it may be subject to.

Consider the template below as an idea pad to set out what you intend the AI system to do. Feel free to be idealistic in your brainstorming session but be specific and precise as to the system requirements and functionality. Outline the proposed use of the system. Remember that the proposed use will guide the success metrics and test parameters to be set at the validation and testing stage. Below are some prompts to guide you in completing this activity.

WHO	WHERE & WHEN
<ul style="list-style-type: none"> Who is/are the potential users of the AI system? Who is responsible for monitoring? 	<ul style="list-style-type: none"> Where and when would this AI system be used? When is the use of this system likely to pose a risk in terms of ethics and human rights (state the risks)?
<ul style="list-style-type: none"> What are the advantages of using an AI system and what will it help to achieve? What are the current IT systems in use that could be replaced or complemented by the AI system? 	<ul style="list-style-type: none"> How would this system be obtained (internal development, external development, hybrid development or off-the-shelf procurement)? How would it be funded? How would this system improve performance in comparison with the methods previously employed by end users in order to carry out their tasks?
WHAT	HOW

EXERCISE 7: Identifying Possible Misuses or Unauthorized Uses

Consider your preferred approach to AI development and identify any possible or potential misuses 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.*

<ul style="list-style-type: none"> • 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? 	<ul style="list-style-type: none"> • 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?
<ul style="list-style-type: none"> • What mitigation strategies can be put in place to ensure that this AI system is not accessed by unauthorized persons? 	<ul style="list-style-type: none"> • What mitigation mechanisms can be put in place to avoid misuse of this AI system?

EXERCISE 8: 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 [ORGANIZATIONAL ROADMAP](#).*

Accordingly, the AI Toolkit's Risk Assessment should be completed at this stage (as well as any other similar assessments in your agency's Risk Management Policy). After carefully completing the Risk Assessment, you will be prepared to take informed action to respond to such risks. This is the focus of the current exercise.

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. [These actions are explained further in the ANNEX.](#)

Details of the Risk Assessment

Completed on:	
Responsible person/team:	



Risk Response

1- Working Team

<p>What EXTREME RISKS were identified in the Risk Assessment?</p>	<p>What is the suggested risk response?</p>
<p>What HIGH RISKS were identified in the Risk Assessment?</p>	<p>What is the suggested risk response?</p>
<p>What MEDIUM RISKS were identified in the Risk Assessment?</p>	<p>What is the suggested risk response?</p>

Risk Response (cont.)

<p>What LOW RISKS were identified in the Risk Assessment?</p>	<p>What is the suggested risk response?</p>
--	--

2 – Management decision

	<p>Signature:</p> <p>Date:</p>
--	--------------------------------



5.

Design & Development Stage

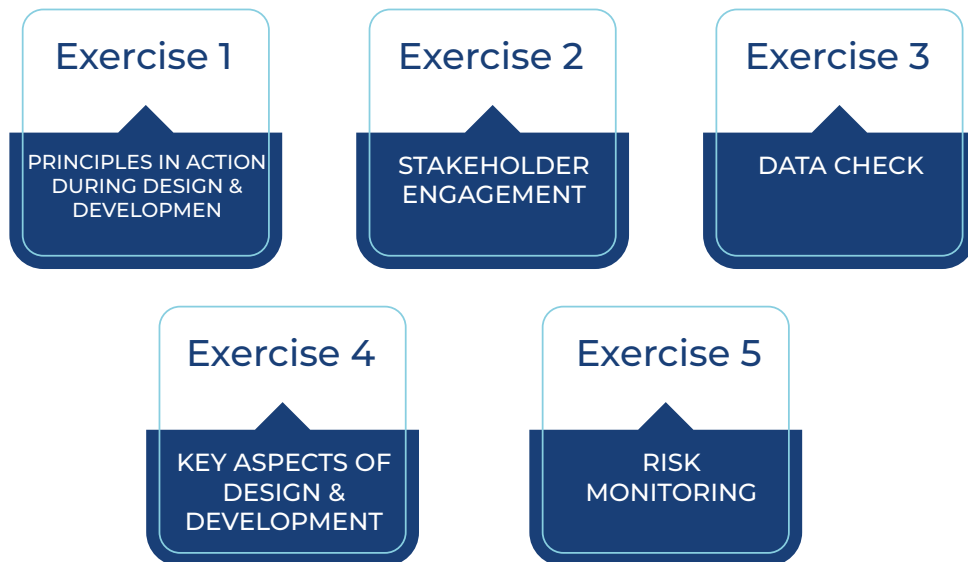
If the decision is made to develop the AI system in-house, the team in charge should perform the activities and processes necessary to fulfil the *Principles for Responsible AI Innovation*.

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:



If you are using the digital version of the workbook, you may 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.

LAWFULNESS*

GENERAL

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?

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.

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 measures are in place to ensure the AI system's security in the case of attacks or unexpected situations and environments?
2. What measures are in place to ensure the AI system is reliable when using different inputs and in different contexts?
3. 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

4. How is the accuracy of the AI system measured? Is its consistency across demographics verified? Is this element submitted to independent third-party testing?
5. Is the training data accurate, representative of the population, up-to-date and adequate for the intended purpose?
6. 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 ENV. 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> <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>
EFFICIENCY	<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

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

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.)?



HUMAN AUTONOMY (cont.)

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?

<p>PROTECT OF VULNERABLE GROUPS</p>	<p>4. What are the positive and negative effects of the AI system's introduction on vulnerable groups ?</p>
<p>DIVERSITY & ACCESS.</p>	<p>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?</p>
<p>CONTEST. & REDRESS</p>	<p>6. Is the AI system designed to allow accessible, efficient, and effective contestability and redress?</p>



GOOD GOVERNANCE

TRACEABILITY & AUDITABILITY

1. How is the AI system's development documented?
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?

ACCOUNTABILITY

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 & ASSESSMENT

1 – WORKING TEAM

<p>Is the principle of LAWFULNESS sufficiently fulfilled? Provide a justification.</p> <p>Is the principle of MINIMIZATION OF HARM sufficiently fulfilled? Provide a justification.</p> <p>Is the principle of HUMAN AUTONOMY sufficiently fulfilled? Provide a justification.</p> <p>Is the principle of FAIRNESS sufficiently fulfilled? Provide a justification.</p> <p>Is the principle of GOOD GOVERNANCE sufficiently fulfilled? Provide a justification.</p>	<p>Suggested next steps:</p>
--	-------------------------------------

2 – MANAGEMENT DECISION

	<p>Signature</p> <p>Date:</p>
--	-------------------------------



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?• Was the data collected in a lawful manner? Was personal data collected under a valid ground for processing under the applicable law?• Has bias been detected in the data set and how will this be addressed?• Can the quality of the data be properly assessed? If so, can a quality benchmark be determined?• Does the data require wrangling or cleaning?	<ul style="list-style-type: none">• Are there any data privacy concerns? Was personal data permitted to be used for developing the model in question?• 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 or the origin of the data in question?• How long and under what safeguards will the data be stored?	<ul style="list-style-type: none">• Is there a data dictionary for the data set to be used for modelling?• What types of features do we need and how many?• Which variables are most correlated to the target variable?



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 and complete Exercise 7 in the planning stage.

🔗 [Click here to access the exercise in the planning stage](#)

Step 1: Reevaluating the Risk Assessment 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** and the *risk response exercise* and update them with the new information available at this stage.

Keep in mind:

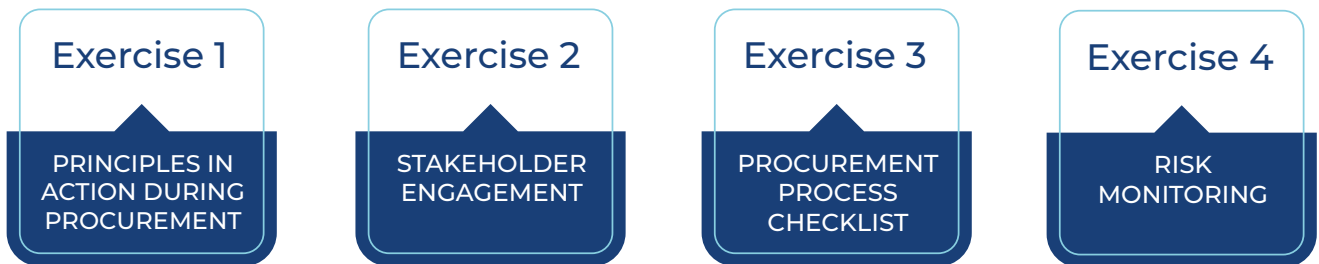
- Have any risk-affecting changes been identified in the AI system's use case or in the environment in which it is intended to operate? If so, what are these changes and which new risk responses do they require?
- Look at the Risk Assessment completed during the planning stage. Are the purposes, scope and constraints of the risk assessment 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 as necessary. What main changes in the risk levels have been identified? What new risk responses do they require?

6.

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:



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.

LAWFULNESS

GENERAL

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?

LEGITIMACY, PROPORTIONALITY & NECESSITY

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.

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. Can the developers 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 developers put in place to identify, eliminate, reduce, or mitigate the risks that the AI system may pose to the safety of each stakeholder?
3. What mechanisms will be put in place by the agency to ensure the reliability, security and safety of the AI system during use? Are they sufficient to minimize harm?

ACCURACY

4. Can the developers 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 developers 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?

<p>ACCURACY</p>	<p>6. Have the developers 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?</p> <p>7. Is the accuracy rate sufficiently high considering the intended use of the AI system?</p>
<p>HUMAN AND ENV. WELL-BEING</p>	<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>
<p>EFFICIENCY</p>	<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 AUTONOMY

HUMAN CONTROL & OVERSIGHT

1. What is the adequate 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? Can the developers demonstrate these measures?
3. Is there a stop / abort command that can be used by a human operator safely and in a timely manner if needed?

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?

PRIVACY

6. What measures have developers 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?

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.)?



10. Has the AI system been developed in such a way as to make it explainable? Can developers 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 are the potential sources of unfair bias arising from the development of the AI system and how might they affect the AI system's outputs? Have the developers 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?

PROTECT. OF
VULNERABLE GROUPS

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?



DIVERSITY & ACCESS.	<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>
CONT. & 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	<p>1. Was the AI system’s development documented? Have the developers made that documentation available?</p> <p>2. Does the AI system’s design include features that allow for traceability of its use?</p> <p>3. Have mechanisms been established to facilitate auditing by independent third parties?</p>
ACCOUNTABILITY	<p>4. Has the AI system been designed to ensure accountability? How? Can the developers demonstrate this?</p> <p>5. Is the scope of relevant officers’ or teams’ responsibility clearly defined?</p>



CONCLUSIONS & ASSESSMENT

1 – WORKING TEAM

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

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.

Suggested next steps:

2 – MANAGEMENT DECISION

Signature

Date:

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	<p>1. Is the vendor selection in line with appropriate national licensing and accreditation requirements?</p> <p>3. What criteria have been set for the vendor to assess bias and other relevant risks?</p> <p>4. Are technology partners/vendors required to present the results of any evaluations conducted during testing to help assess the AI system's performance?</p> <p>5. Will the vendor be providing training for users? Is this training covered by the vendor's contractual obligations?</p> <p>6. Has a standardized procurement process been created/followed?</p>
PROCUREMENT PROCESS	<p>7. Have the following aspects been completed?</p> <ul style="list-style-type: none"> 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)

EXERCISE 5: Risk Monitoring

Before you start: Did you carry out the Risk Assessment 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 and complete Exercise 7 in the planning stage.

🔗 [Click here to access the exercise in the planning stage](#)

Step 1: Reevaluating the Risk Assessment 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** and the *risk response exercise* 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 completed during the planning stage. Are the purposes, scope and constraints of the risk assessment 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 as necessary. What main changes in the risk levels have been identified? What new risk responses do they require?

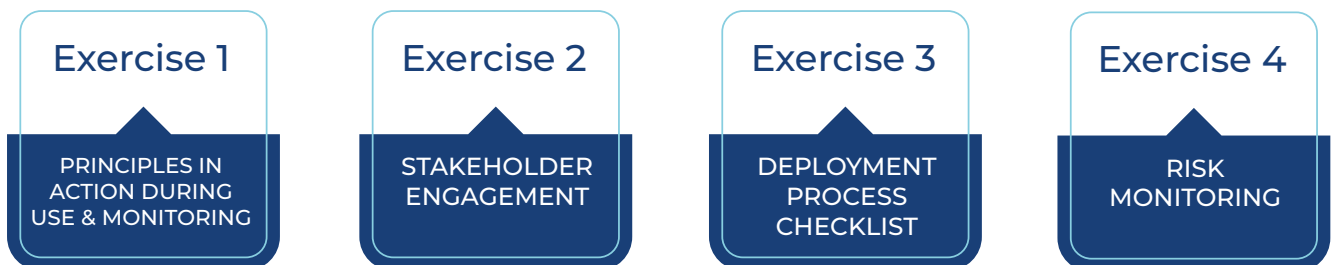
7.

Use & Monitoring Stage

The use and monitoring stage begins when the AI system is *deployed*, i.e. made available to primary stakeholders – end users, monitoring and control teams, developers and ICT teams, system admin, and independent auditors. It 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 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:



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 Use & Monitoring

Use the questions in this exercise to assess activities during the use and monitoring 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.

LAWFULNESS*

GENERAL

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 at stake 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.

LEGITIMACY, PROPORTIONALITY & NECESSITY

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? How often are audits conducted?
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?

ACCURACY

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?

HUMAN AND ENV. WELL-BEING	<p>8. What strategies are in place for assessing and safeguarding individual and environmental well-being during use of the AI system?</p> <p>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?</p>
EFFICIENCY	<p>10. What mechanisms are in place for continuously measuring the AI system's efficiency in terms of financial, time, and human and environmental resources?</p> <p>11. Are users trained to measure the AI system's efficiency and to use it in the most efficient way possible?</p>



HUMAN AUTONOMY

HUMAN CONTROL & OVERSIGHT

1. How and when are individuals 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?

HUMAN AGENCY	<p>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?</p>
PRIVACY	<p>8. How does the AI system impact the stakeholders' privacy (including users)? How is this communicated to stakeholders?</p> <p>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?</p> <p>10. What mechanisms are in place to ensure the security of the data which is fed into and processed by the AI system?</p>



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

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

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

9. What mechanisms are in place for assessing the usability of the AI system by individuals with varying types and levels of ability?

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?

GOOD GOVERNANCE

TRACEABILITY & AUDITABILITY

1. What measures are in place to monitor traceability mechanisms?
2. Is there a mechanism in place for logging user behaviour (including officers)?
3. Has the AI system's audit been completed? Is there a routine audit mechanism in place?
4. What process has been set out for implementing the results and feedback from audits?

ACCOUNTABILITY

5. Is the scope of relevant officers' or teams' responsibility clearly defined?
6. What mechanisms are in place for assessing any incidents and how will responsibility be assumed in such cases?



CONCLUSIONS & ASSESSMENT

1 – WORKING TEAM

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

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.

Suggested next steps:

2 – MANAGEMENT DECISION

Signature

Date:

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 to primary stakeholders – end users, monitoring and control teams, developers and ICT teams, system admin, and independent auditors.

SYSTEM LAUNCH & ROLLOUT	1. How would the rollout of this AI system be conducted; would it be a phased approach?
	2. Identify the step-by-step roll-out plan for this AI system.
	3. Which are the key teams responsible for ensuring the adoption of this AI system?
	4. Is there any unique technological infrastructure or AI system which must be used alongside this AI system? If so, explain what.

EXERCISE 5: Risk Monitoring

Before you start: Did you carry out the Risk Assessment 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 and complete Exercise 7 in the planning stage.

🔗 [Click here to access the exercise in the planning stage](#)

Step 1: Reevaluating the Risk Assessment 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** and the *risk response exercise* 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 completed during the planning stage. Are the purposes, scope and constraints of the risk assessment 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 as necessary. What main changes in the risk levels have been identified? What new risk responses do they require?

Annex:

Risk Response Measures

Once a risk has been assessed and identified, the organization and/or the risk owner is prompted to choose a course of action to deal with the risk. When the level of risk is in the Medium, High or Extreme category, it is particularly important that adequate measures are taken to minimize the severity of the risk's impact and/or its likelihood. These measures should be sufficient to lower the risk level to Low and safeguard individuals and communities from continuous and disproportionate harm.

The risk response may include one or more of the following types of measures:

- **Risk acceptance:** Making an informed and conscious decision to take on a specific risk with the aim of leveraging the opportunities offered by the AI system. It is crucial that these accepted risks are in the Low category and are continuously monitored and reviewed to ensure the good governance principles of traceability, auditability and accountability are adhered to.
- **Risk mitigation:** Implementing measures designed to decrease the level of risk. These strategies can include various processes, policies, practices, or devices intended to decrease the probability and/or severity of a particular event or situation, or eliminate the source of the risk.
- **Risk avoidance:** Making a conscious and informed decision to refrain from participating in or to terminate the creation, deployment, or use of an AI system in order to prevent a specific risk. Risk avoidance may be based on the outcomes of a risk assessment and/or legal and regulatory responsibilities.
- **Risk monitoring:** This involves the constant observation, supervision, and evaluation of the risk status to identify any instances where the use of the AI system is creating harm and other negative effects to individuals and communities. It may also involve documenting all processes associated with monitoring the AI system in order to guarantee traceability and auditability. This stage is also valuable as it enables any risk

mitigation measures previously implemented to be assessed for efficacy, and allows you to determine whether new measures need to be introduced.

- **Risk auditing:** This is an objective, systematic, and thoroughly documented procedure which is carried out by a third party in order to independently assess the adequacy and effectiveness of the AI risk management framework in terms of protecting individuals and communities from harm and other negative effects.
- **Risk reporting:** This is a communication method whereby certain internal or external stakeholders are informed of the current status of the risks associated with the AI system and how they are being managed.
- **Risk financing:** This involves planning the allocation of funds to address or alter any potential financial consequences of the risks identified.

How to cite this publication: UNICRI and INTERPOL. (June 2023). Toolkit for Responsible AI Innovation in Action Workbook

© United Nations Interregional Crime and Justice Research Institute (UNICRI), 2023

© International Criminal Police Organization (INTERPOL), 2023



www.interpol.int
www.unicri.it



INTERPOL_HQ



@INTERPOL_HQ
@UNICRI



INTERPOL HQ
UNICRI



INTERPOL
UNICRI



@INTERPOL
@UNICRIHQ