

Getting started

How to get started with artificial intelligence (AI) and the steps to take for a successful test and roll-out

First published 17 June 2025

4 mins read

Delivery team

The delivery team should be structured to work fast, so make sure it is led by someone suitably senior to open doors and move things along in your force. That person will usually also serve as information asset owner for the duration of the project and should link in with the information security team regarding any training available to support them in this role.

If it is a high-profile or higher-risk project, consider a dedicated team rather than one composed of officers or staff juggling other duties.

The team should be multidisciplinary, spanning operational, IT, data science, data engineering, testing and evaluation skills. Some forces have joint senior reporting officers for data-driven innovation projects: one from operational policing and another from IT. It is essential that your force's IT specialists are well-represented and respected in these sorts of projects. You may also find it beneficial to designate a team member as the main point of contact with the supplier.

If you are unsure of what skills you need, consult the National Police Chiefs' Council (NPCC) [AI Playbook](#), the Office of the Police Chief Scientific Adviser (OPCSA) or Police Digital Service (PDS). In addition to your IT department, you may be able to find some of the technical skills you need in your service desk and (for testing and evaluation) your force analysts.

If your force lacks the capabilities, you could consider:

- partnering with other forces (if you are not doing that already)
- seconding expertise from other forces
- recruiting university graduates – not necessarily employees but through graduate internship opportunities or through [Knowledge Transfer Partnerships](#) (a programme that promotes collaboration on innovation with academics and graduates)

Training, fine-tuning and preparing to test

Training large generative artificial intelligence (AI) models is something only large companies like OpenAI can afford. The slightly older generation of machine learning models, which can run on central processing units or smaller graphics processing units, may be trainable on relevant data (ideally from multiple forces). This is impossible for models that require large amounts of training data, such as large language models (LLMs), which need that pre-training to learn the meaning of words in context.

A pre-trained model can, however, be fine-tuned to perform a specialised task and return information in the desired format. Fine-tuning is where a pre-trained model undergoes additional training on a smaller, more focused data set.

The accuracy of pre-trained and fine-tuned systems can be further enhanced through retrieval augmented generation (RAG), where you supplement the LLM with specialised and up-to-date data sources that may not have been available during the training and fine-tuning process. RAG can significantly limit the risk of an AI hallucinating, although note that RAG will not be applicable for every type of LLM.

All AI-based tools and systems, regardless of how they have been trained, should be tested on your force's data.

Preparing to test

Delivery team members with data science or engineering expertise will need to identify from your force's systems the data required for each data feature. The [Machine learning guide for policing](#) sets out the typical stages of preparing organisational data for machine-learning in an accessible way. These include:

- importing
- labelling
- cleaning data
- splitting data for training (if applicable)
- validation
- testing

Whatever type of AI you are bringing in, you will need to pay attention to the following.

Data quality

Consider whether this is an AI application where:

- the data needs to be pristine to deliver reliable outputs (for example, automated case-progression updates for victims of crime)
- this is less of a consideration (LLMs that have been trained on the entire internet)

There may be workarounds where a particular subset of data is not satisfactory that will enable the tool or system to still go ahead, such as multi-factor identification where names are spelled inconsistently across systems.

Data sufficiency

There may also be workarounds for not having enough data, such as cross-validation techniques or using algorithms to generate a synthetic data set. Your data scientists can explore these.

Data access

In addition to guarding against data leaking out into the wider world, ensure that the right internal access controls are in place to protect your force's most sensitive data from being shared too widely within your organisation ('overprivileged access').

Ahead of testing, make sure that vetting and other data access agreements for any external parties involved are completed before training and testing begins.

Tags

Artificial intelligence