Department for
Science, Innovation
& Technology

AI Management Essentials

Public Consultation

Issue date – 6th November 2024

Closing date – 29th January 2025

Introduction

This consultation introduces the Department for Science, Innovation and Technology's (DSIT) AI Management Essentials tool (AIME). AIME is a resource that is designed to provide clarity to organisations around practical steps for establishing a baseline of good practice for managing artificial intelligence (AI) systems that they develop and/or use.

The effective management of AI systems is important to ensuring that organisations can unlock the benefits of innovative technologies while mitigating risks and potential harms. Alongside the increasing uptake of AI across sectors, recent years have seen a proliferation of frameworks and tools designed to support organisations to manage and mitigate risks associated with AI systems. However, navigating this landscape of resources can be complex and resource-intensive, especially for smaller organisations that may lack knowledge of AI management practices, or have limited time and resources to meaningfully engage with these frameworks.

To address this, AIME distils key principles from existing AI regulations, standards and frameworks to provide an accessible starting point for organisations to assess and improve their AI management systems. The tool contains a self-assessment questionnaire designed to highlight the strengths and weaknesses of an organisation's AI management system. The final version of AIME—which will be developed after this consultation—will be accompanied by a scoring system and recommended actions for mitigating issues highlighted by the tool.

This consultation is seeking feedback on the design, content and utility of the AI Management Essentials tool, to ensure that it is fit for purpose, and will help businesses to implement effective AI management processes across their organisation.

Contents

Introduction 1

General information 3

Why are we consulting? 3

Consultation details 3

Guidance for using AI Management Essentials 4

Introduction to AI Management Essentials (AIME) 4

Who AIME is designed for 4

Why DSIT has developed AIME 4

Why businesses should use AIME 5

What AIME looks like 6

Who the AIME self-assessment should be completed by 6

How to complete the AIME self-assessment 6

Self-assessment questions 8

1. AI system record 8

2. AI policy 9

3. Fairness 10

4. Impact assessment 12

5. Risk assessment 13

6. Data management 15

7. Bias mitigation 17

8. Data protection 18

9. Issue reporting 19

10. Third party communication 21

Annex A: Glossary 23

General information

Why are we consulting?

This consultation introduces and invites feedback on the AI Management Essentials tool. Building on previous engagement with industry and regulators through workshops and pilots, it looks to better understand how DSIT can enable businesses of different sizes and sectors to implement robust AI management systems.

DSIT will analyse the responses from this consultation, using feedback to further refine the AI Management Essentials tool to ensure that it is fit for purpose and supports organisations in assessing and implementing responsible AI management practices.

Consultation details

Issued: 6th November 2024

Respond by: 23:55, 29th January 2025

Enquiries to: ai-assurance@dsit.gov.uk

Consultation reference: AI Management Essentials

Audiences:

The government invites feedback from any interested party, but in particular from representatives of start-ups and Small-to-Medium Enterprises (SMEs) who develop and/or use AI systems.

Territorial extent:

All of the UK.

Guidance for using AI Management Essentials

Introduction to AI Management Essentials (AIME)

AI Management Essentials is a self-assessment tool designed to help businesses establish robust management practices for the development and use of AI systems. The tool is not designed to evaluate AI products or services themselves, but rather to evaluate the organisational processes that are in place to enable the responsible development and use of these products.

ⓘ **AI systems:** products, tools, applications or devices that utilise AI models to help solve problems. AI systems are the operational interfaces to AI models - they incorporate technical structures and processes that allow models to be used by non-technologists. More information on how AI systems relate to AI models and data can be found in DSIT's Introduction to AI Assurance.

Who AIME is designed for

AIME can be used by any organisation that develops, provides or uses services that utilise AI systems as part of its standard business operations. AIME is sector agnostic and may be used by organisations of different sizes. However, it is primarily intended for SMEs and start-ups that encounter barriers when navigating the evolving landscape of AI management standards and frameworks. For larger organisations, AIME can be used to assess AI management systems for individual business divisions, operational departments or subsidiaries.

ⓘ **AI management system**: the set of governance elements and activities within an organisation that support decision making and the delivery of outcomes relating to the development and use of AI systems. This includes organisational policies, objectives, and processes among other things. More information on assuring AI governance practices can be found in DSIT's Introduction to AI Assurance.

Why DSIT has developed AIME

Over the last few years, there have been a proliferation of standards and frameworks designed to help organisations effectively manage AI systems. While these resources offer important guidance, our engagement with industry suggests that many organisations find it challenging to navigate this landscape and engage with these resources. To address this, DSIT has developed AIME to deliver practical support and greater clarity for businesses. AIME distils key principles from existing AI regulations, standards and frameworks to provide an accessible resource for organisations to assess and improve their AI management systems and practices.

We conducted a literature review of key frameworks and standards and have based the tool on three prominent frameworks:

- ISO/IEC 42001,
- the NIST Risk Management framework,
- the EU AI Act.

We prioritised these international frameworks, in part, to ensure the interoperability of the tool. It worth noting that AIME does not seek to replace these frameworks, nor does completing the AIME self-assessment represent compliance, but it provides a starting point for implementing commonly regarded best practices in AI management.

AIME will also complement and support other existing international efforts to identify and mitigate risks posed by AI systems, such as the OECD Reporting Framework for the G7 Hiroshima Process Code of Conduct for organisations developing advanced AI systems, which is currently under development. In particular, the OECD's G7 Reporting Framework will primarily seek to facilitate effective action and greater transparency among companies developing the most advanced AI systems, complementary to the UK's AIME, which will provide an accessible starting point for organisations of any size to assess and improve their AI management systems.

After a thematic analysis to identify common themes and principles across these documents, we distilled key information into a series of questions for organisations to self-assess and identify actions to improve their management systems.

Over the past year, we have iterated and tested a prototype of AIME in three targeted pilots with industry organisations. These pilots were followed by three workshops, where we presented and iterated the tool with regulators and policy makers; government departments; and SMEs via techUK. This feedback has informed the ongoing development of the tool, and this consultation seeks to gather information to refine it further.

Why businesses should use AIME

The tool will not be mandatory to use but will help organisations to embed baseline good practice within their AI management systems. It is designed to provide clarity on what is needed to demonstrate responsible AI management systems and will help organisations to identify the strengths and weaknesses of their internal processes. The tool also provides practical actions for improving management systems.

AIME does not provide formal certification. However, working through the tool will help organisations to assess and improve their AI management processes, and become better positioned for a foundational level of compliance with the standards and frameworks that inform it.

In the future, there may be opportunities to explore embedding AIME into public sector procurement frameworks for AI products and services.

What AIME looks like

We expect the final version of AIME to include three components:

1. A self-assessment questionnaire;
2. A rating for each section of the self-assessment, to provide users with a concise view of their AI management system health, calculated on self-assessment answers;
3. A set of action points and recommendations for improvement, generated by self-assessment answers.

Only the self-assessment questionnaire is included in this consultation. The ratings and recommendations will be developed by DSIT following this consultation. The outputs will be made available alongside the final version of the AIME tool.

AIME is organised into three thematic areas:

- **Internal processes**: these questions assess the overarching structures and principles underlying your AI management system.
- **Managing risks**: these questions assess the processes through which you prevent, manage, and mitigate risks.
- **Communication**: these questions assess your communication with employees, external users and interested parties.

Each section begins with a motivating statement to represent good practice that the following questions are designed to interrogate.

Who the AIME self-assessment should be completed by

The assessment should be completed by an individual or individuals who have wide-reaching knowledge of an organisation's governance practices. For example, a CTO or software engineer may have relevant expertise for answering more technical questions, whilst involving your AI Ethics Officer or HR Business Manager may also be helpful, if you have colleagues with these roles or similar in your organisation.

How to complete the AIME self-assessment

Please note, you do not need to complete the self-assessment in order to respond to this consultation. We welcome general feedback on the design, content and usability of this tool.

For those who do wish to complete the self-assessment, please work your way through the questions in order, starting from Section 1. Respond to questions by checking **one** of the multiple-choice boxes provided below them. You can either do this by hand on a printed copy or digitally using a PDF mark-up tool. The time taken to complete this section will vary across organisations, depending on your expertise and your organisation's existing governance structures.

Depending on your answer to a given question, you may not be required to respond to all subsequent questions in that section. Where this is the case, this will be clearly stated beside the relevant answer box. If no option to skip is provided, please proceed to the next question as usual. Questions that are conditional on a previous answer are indented.

For questions containing technical or specialised terminology, a short explanation of these terms is provided in-line in a grey text box marked with an ⓘ. A glossary of key terms used throughout the self-assessment and this document is also available in Annex A for reference.

Self-assessment questions

Internal Processes

1. AI system record

We maintain a complete and up-to-date record of all the AI systems our organisation develops and uses.

1.1 Do you maintain a record of the AI systems your organisation develops and uses?

a. ☐ Yes *If a, then continue to 1.2*

b. ☐ No *If b, then skip to next section*

ⓘ **AI system record**: an inventory of documentation, assets and resources related to your AI systems. This may encompass, but is not limited to, content referenced throughout this self-assessment, including: technical documentation; impact and risk assessments; AI model analyses; and data records. In practice, an AI system record may take the form of a collection of files on your organisation's shared drive, information distributed across an enterprise management system, or resources curated on an AI governance platform.

1.2 What proportion of the AI systems that you develop and use are documented in your AI system record?

a. ☐ All

b. ☐ The majority

c. ☐ Some

1.3 Do you have an established process for adding new systems to your AI system record?

a. ☐ Yes

b. ☐ No

1.4 If you procure or access AI systems from third party providers, do you request and receive documentation, assets and resources for your AI system record from them?

a. □ Yes, always
b. □ Yes, sometimes
c. □ No

1.5 How frequently do you review and update your AI system record?

a. □ Twice a year or more
b. □ Once a year
c. □ Less than once a year

Internal Processes

2. AI policy

We have a clear, accessible and suitable AI policy for our organisation.

2.1 Do you have an AI policy for your organisation?

a. □ Yes *If a, then continue to 2.2*
b. □ No *If b, then skip to next section*

ⓘ **AI policy:** information that provides governance direction and support for AI systems according to your business requirements. Your AI policy may include but is not limited to: principles and rules that guide AI-related activity within your organisation; frameworks for setting AI-related objectives; and assignments of roles and responsibilities for AI management.

2.2 Is your AI policy available and accessible to all employees?

a. □ Yes
b. □ No

2.3 Does your AI policy help users evaluate whether the use of an AI is appropriate for a given function or task?

a. □ Yes
b. □ No

2.4 Does your AI policy identify clear roles and responsibilities for AI management processes in your organisation?

a. ☐ Yes
b. ☐ No

2.5 How frequently do you review and update your AI policy?

a. ☐ Twice a year or more
b. ☐ Once a year
c. ☐ Less than once a year

Internal Processes

3. Fairness

We seek to ensure that the AI systems we develop and use which directly impact individuals are fair.

3.1 Do you develop or use AI systems that directly impact individuals?

a. ☐ Yes *If a, then continue to 3.2*
b. ☐ No *If b, then skip to next section.*

ⓘ **Direct impact**: we encourage you to judge 'directness' of impact in the context of your own organisational activities. As a starting point, we suggest that the following categories of AI systems should be considered to have direct impact on individuals:

1. AI systems that are used to make decisions about people (e.g. profiling algorithms);
2. AI systems that process data with personal or protected characteristic attributes (e.g. forecasting or record linking algorithms that utilise demographic data or personal identifiers);
3. AI systems where individuals affected by the system outputs are also the end-users (e.g. chatbots, image generators).

3.2 Do you have clear definitions of fairness with respect to these AI systems?

a. ☐ Yes, for all *If a, then continue to 3.3*
b. ☐ Yes, for some *If b, then continue to 3.3*
c. ☐ Not for any *If c, then skip to next section.*

ⓘ **Fairness:** a broad principle embedded across many areas of law and regulation, including equality and human rights, data protection, consumer and competition law, public and common law, and rules protecting vulnerable people.

Section 7 focuses on bias mitigation. We differentiate unfairness from bias, where bias is statistical phenomenon that is characteristic of a process such as decision-making, and unfairness is an outcome of a biased process being implemented in the real world.

3.3 Do you have mechanisms for detecting or identifying unfair outcomes or processes with respect to these AI systems and your definitions of fairness?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

3.4 Do you have processes for monitoring fairness of AI systems over time and mitigating against unfairness?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

3.5 How frequently do you review your process(es) for detecting and mitigating unfairness?

a. ☐ Twice a year or more
b. ☐ Once a year
c. ☐ Less than once a year

Managing risks

4. Impact assessment

We have identified and documented the possible impacts of the AI systems our organisation develops and uses.

4.1 Where appropriate, do you have an impact assessment process for identifying how your AI systems might impact…

4.1.1 The legal position or life opportunities of individuals?

a. □ Yes
b. □ No

4.1.2 The physical or psychological wellbeing of individuals?

a. □ Yes
b. □ No

4.1.3 Societies and the environment?

a. □ Yes
b. □ No

If a to any of 4.1, continue to 4.2
If b to all of 4.1, skip to next section

ⓘ **AI impact assessment:** a framework used to consider and identify the potential consequences of an AI system's deployment, intended use and foreseeable misuse.

4.2 Do you document potential impacts of your AI systems?

a. □ Yes, for all
b. □ Yes, for some
c. □ Not for any

4.3 Do you communicate the potential impacts to the users or customers of your AI systems?

d. □ Yes, for all
e. □ Yes, for some
f. □ Not for any

Managing risks

5. Risk assessment

We effectively manage any risks caused by our AI systems.

5.1 Do you conduct risk assessments of the AI systems you develop and use?

a. □ Yes, for all *If a, then continue to 5.1*
b. □ Yes, for some *If b, then continue to 5.1*
c. □ Not for any *If c, then skip to 5.3.1*

ⓘ **AI risk assessment**: a framework used to consider and identify a range of potential risks that might arise from the development and/or use of an AI system. These include bias, data protection and privacy risks, risks arising from the use of a technology (e.g. the use of a technology for misinformation or other malicious purposes) and reputational risk to the organisation.

5.2.1 Are your risk assessments designed to produce consistent, valid and comparable results?

a. □ Yes
b. □ No

5.2.2 Do you compare the results of your risk assessments to your organisation's overall risk thresholds?

a. □ Yes
b. □ No

5.2.3 Do you use the results of your risk assessment to prioritise risk treatment?

d. □ Yes
e. □ No

5.3.1 Do you monitor all your AI systems for general errors and failures?

a. ☐ Yes
b. ☐ No

5.3.2 Do you monitor all your AI systems to check that they are performing as expected?

a. ☐ Yes
b. ☐ No

5.4 Do you have processes for responding to or repairing system failures?

a. ☐ Yes, for all *If a, then continue to 5.5*
b. ☐ Yes, for some *If b, then continue to 5.5*
c. ☐ Not for any *If c, then skip to 5.6*

5.5 Have you defined risk thresholds or critical conditions under which it would become necessary to cease the development or use of your AI systems?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

5.6 Do you have a plan to introduce necessary updates to your risk assessment process as your AI systems evolve or critical issues are identified?

a. ☐ Yes
b. ☐ No

Managing risks

6. Data management

We responsibly manage the data used to train, fine-tune and otherwise develop our AI systems.

6.1 Do you train, fine-tune or otherwise develop your own AI systems using data?

a. ☐ Yes *If a, then continue to 6.2*
b. ☐ No *If b, then skip to next 6.6*

6.2 Do you document details about the provenance and collection processes of data used to develop your AI systems?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

ⓘ **Data provenance**: information about the creation, updates and transfer of control of data.

6.3 Do you ensure that the data used to develop your AI systems meet any data quality requirements defined by your organisation?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

ⓘ **Data quality**: broadly, the suitability of data for a specific task, or the extent to which the characteristics of data satisfy needs for use under specific conditions. Further information can be found on the Government Data Quality Hub.

6.4 Do you ensure that datasets used to develop your AI systems are adequately complete and representative?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

ⓘ **Data completeness**: the extent to which a dataset captures all the necessary elements for use under specific conditions. In practice, ensuring data completeness may involve replacing missing data with substituted values or removing data points that may compromise the accuracy or consistency of the AI system it is used to develop.

ⓘ **Data representativeness**: the extent to which a data sample distribution corresponds to a target population. In practice, ensuring data representativeness may involve undertaking and responding to statistical data analysis that quantifies how closely your sample data reflects the characteristics of a larger group of subjects, or analysis of data sampling and collection techniques.

6.5 Do you document details about the data preparation activities undertaken to develop your AI systems?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

ⓘ **Data preparation**: includes any processing or transformation performed on a dataset prior to training or development of an AI system. In practice, this may include, but is not limited to: any process used to ensure data quality, completeness and representativeness, converting or encoding dataset features, feature scaling or normalisation, or labelling target variables.

6.6 Do you sign and retain written contracts with third parties that process personal data on your behalf?

a. ☐ Yes, always
b. ☐ Yes, sometimes
c. ☐ No

Managing risks

7. Bias mitigation

We mitigate against foreseeable, harmful and unfair algorithmic and data biases in our AI systems.

7.1 Do you take action to mitigate against foreseeable harmful or unfair bias related to the training data of your AI systems?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

ⓘ **Bias**: the disproportionate weighting towards a particular subset of data subjects. Whilst bias is not always negative, it can cause a systematic skew in decision-making that results in unfair outcomes, perpetuating and amplifying negative impacts on certain groups.

7.2 If you procure AI as a Service (AIaaS) or pretrained AI systems from third party providers to use or develop upon, do you have records of the full extent of the data that has been used to train these systems?

a. ☐ Yes, for all *If a, then continue to 7.3*
b. ☐ Yes, for some *If b, then continue to 7.3*
c. ☐ Not for any *If c, then skip to 7.4*

ⓘ **AI as a Service**: a service that outsources a degree of your AI system functionality to a third party. AIaaS are often delivered as ‘off-the-shelf’ solutions with supporting infrastructure such as online platforms and APIs that allow for easy integration into existing business operations. Cloud-based AI software and applications provided by large tech companies are archetypal examples of AIaaS.

ⓘ **Pre-trained**: refers to machine learning systems that have been initialised by training on a large, general dataset, and can be fine-tuned to accomplish specific downstream tasks.

7.3 **If you procure AIaaS or pretrained AI systems from third party providers, do you conduct appropriate due diligence on the data used to train or develop these systems to mitigate against foreseeable harmful or unfair bias?**

a. □ Yes, for all
b. □ Yes, for some
c. □ Not for any

ⓘ **Due diligence:** this may include requesting and reviewing the results of bias audits conducted by the developer of the 'off-the-shelf' AI system to determine if there is unfair bias in the input data, and/or the outcome of decisions or classifications made by the system.

7.4 **Do you have processes to ensure compliance with relevant bias mitigation measures stipulated by international or domestic regulation?**

a. □ Yes
b. □ No

Managing risks

8. Data protection

We have a "data protection by design and default" approach throughout the development and use of our AI systems.

8.1 **Do you implement appropriate security measures to protect the data used and/or generated by your AI systems?**

a. □ Yes
b. □ No

ⓘ **Data protection security measures**: see ICO guidance on data security under UK GDPR for a further information.

8.2 **Do you record all your personal data breaches?**

a. □ Yes
b. □ No

If a, then continue to 8.3
If b, then skip to 8.4

8.3 Do you report personal data breaches to affected data subjects when necessary?

a. □ Yes
b. □ No

8.4 Do you routinely complete Data Protection Impact Assessments (DPIAs) for uses of personal data that are likely to result in high risk to individuals' interests?

a. □ Yes
b. □ No

ⓘ **Data Protection Impact Assessment**: see ICO guidance on DPIAs for further information.

8.5 Have you ensured that all your AI systems and the data they use or generate is protected from interference by third parties?

a. □ Yes
b. □ No

Communication

9. Issue reporting

We have reporting mechanisms for employees, users and external third parties to report any failures or negative impacts of our AI systems.

9.1 Do you have reporting mechanisms for all employees, users and external third parties to report concerns or system failures?

a. □ Yes *If a, then continue to 9.2*
b. □ No *If b, then skip to 9.4*

9.2 Do you provide reporters with options for either anonymity or confidentiality or both?

a. □ Yes
b. □ No

ⓘ **Anonymity:** in practice, providing anonymity requires excluding any personal data collection from the reporting procedure.

ⓘ **Confidentiality:** in practice, providing confidentiality requires preventing anyone other than the intended recipient from connecting individual reports to a reporter.

9.3 Have you identified who in your organisation will be responsible for addressing concerns when they are escalated?

a. □ Yes
b. □ No

9.4 Are your reporting procedures meaningfully transparent for all employees, users and external third parties?

a. □ Yes
b. □ No

ⓘ **Transparency**: refers to the communication of appropriate information about an AI system to relevant people, in a way that they understand. In practice, making reporting procedures transparent requires clearly informing reporters about: how they can expect their report to be processed; how their report is processed; when their report has finished being processed; and any outcomes to which the report can be directly attributed.

9.5 Do you respond to concerns in a timely manner?

a. □ Yes
b. □ No

ⓘ **Timely**: timeliness is subjective and will depend on the nature of your organisation and concerns. As a rule of thumb, you could consider “timely” to mean no more than 72 hours. This is the amount of time in which you are required to report a data breach after becoming aware of it under GDPR.

9.6 Do you document all reported concerns and results of any subsequent investigations?

a. □ Yes
b. □ No

Communication

10. Third party communication

We tell every interested party how to use our AI systems safely and what the systems' requirements are.

10.1 Have you determined what AI system technical documentation is required by interested parties across your relevant stakeholder categories (e.g. developers, AI system end-users, academic researchers, etc)?

a. □ Yes, for all *If a, then continue to 10.2*

b. □ Yes, for some *If b, then continue to 10.2*

c. □ Not for any *If c, then skip to 10.3*

ⓘ **Technical documentation**: a written description of or guide to an AI system's functionality. For instance, technical documentation content may include: usage instructions; technical assumptions about its use and operation; system architecture; and technical limitations. Manuals, code repositories and model cards are examples of technical documentation.

10.2 Do you provide technical documentation to interested parties in an appropriate format?

a. □ Yes, for all

b. □ Yes, for some

c. □ Not for any

ⓘ **Appropriate format**: broadly, this means that documentation is tailored to your interested parties' needs and expected level of understanding.

10.3 Have you determined what AI system non-technical documentation is required by interested parties across your relevant stakeholder categories?

a. ☐ Yes, for all *If a, then continue to 10.4*
b. ☐ Yes, for some *If b, then continue to 10.4*
c. ☐ Not for any *If c, then stop*

ⓘ **Non-technical documentation**: a written description or analysis of the benefits or issues associated with the use of an AI system outside of its operational processes. Impact assessments and risk assessments are examples of non-technical documentation.

10.4 Do you provide non-technical information to your users and other relevant parties?

a. ☐ Yes, for all
b. ☐ Yes, for some
c. ☐ Not for any

Annex A: Glossary

AI as a Service (AIaaS): a service that outsources a degree of your AI system functionality to a third party. AIaaS are often delivered as 'off-the-shelf' solutions with supporting infrastructure such as online platforms and APIs that allow for easy integration into existing business operations. Cloud-based AI software and applications provided by large tech companies are archetypal examples of AIaaS.

AI impact assessment: a framework used to consider and identify the potential consequences of an AI system's deployment, intended use and foreseeable misuse.

AI management system: the set of governance elements and activities within an organisation that support decision making and the delivery of outcomes relating to the development and use of AI systems. This includes organisational policies, objectives, and processes among other things. More information on assuring AI governance practices can be found in DSIT's Introduction to AI Assurance.

AI policy: information that provides governance direction and support for AI systems according to your business requirements. Your AI policy may include but is not limited to: principles and rules that guide AI-related activity within your organisation; frameworks for setting AI-related objectives; and assignments of roles and responsibilities for AI management.

AI risk assessment: a framework used to consider and identify a range of potential risks that might arise from the development and/or use of an AI system. These include bias, data protection and privacy risks, risks arising from the use of a technology (e.g. the use of a technology for misinformation or other malicious purposes) and reputational risk to the organisation.

AI systems: products, tools, applications or devices that utilise AI models to help solve problems. AI systems are the operational interfaces to AI models – they incorporate technical structures and processes that allow models to be used by non-technologists. More information on how AI systems relate to AI models and data can be found in DSIT's Introduction to AI Assurance.

AI system record: an inventory of documentation, assets and resources related to your AI systems. This may encompass, but is not limited to, content referenced throughout this self-assessment, including: technical documentation; impact and risk assessments; model analyses; and data records. In practice, an AI system record may take the form of a collection of files on your organisation's shared drive, information distributed across an enterprise management system, or resources curated on an AI governance platform.

Anonymity: in practice, providing anonymity requires excluding any personal data collection from the reporting procedure.

Confidentiality: in practice, providing confidentiality requires preventing anyone other than the intended recipient from connecting individual reports to a reporter.

Bias: the disproportionate weighting towards a particular subset of data subjects. Whilst bias is not always negative, it can cause a systematic skew in decision-making that results in unfair outcomes, perpetuating and amplifying negative impacts on certain groups.

Data completeness: the extent to which a dataset captures all the necessary elements for use under specific conditions. In practice, ensuring data completeness may involve replacing missing data with substituted values or removing data points that may compromise the accuracy or consistency of the AI system it is used to develop.

Data representativeness: the extent to which a data sample distribution corresponds to a target population. In practice, ensuring data representativeness may involve undertaking and responding to statistical data analysis that quantifies how closely your sample data reflects the characteristics of a larger group of subjects, or analysis of data sampling and collection techniques.

Data preparation: includes any processing or transformation performed on a dataset prior to training or development of an AI system. In practice, this may include, but is not limited to: any process used to ensure data quality, completeness and representativeness, converting or encoding dataset features, feature scaling or normalisation, or labelling target variables.

Data Protection Impact Assessment: see ICO guidance on DPIAs for further information.

Data protection security measures: see ICO guidance on data security under UK GDPR for a further information.

Data provenance: information about the creation, updates and transfer of control of data.

Data quality: broadly, the suitability of data for a specific task, or the extent to which the characteristics of data satisfy needs for use under specific conditions. Further information can be found on the Government Data Quality Hub.

Direct impact: we encourage you to judge 'directness' of impact in the context of your own organisational activities. As a starting point, we suggest that the following categories of AI systems should be considered to have direct impact on individuals:

1. AI systems that are used to make decisions about people (e.g. profiling algorithms);
2. AI systems that process data with personal or protected characteristic attributes (e.g. forecasting or entity resolution algorithms that utilise demographic data or personal identifiers);
3. AI systems where individuals impacted by the system output are also the end-users (e.g. chatbots, image generators).

Fairness: a broad principle embedded across many areas of law and regulation, including equality and human rights, data protection, consumer and competition law, public and common law, and rules protecting vulnerable people. We differentiate unfairness from bias, where bias is statistical phenomenon that is characteristic of a process such as decision-making, and unfairness is an outcome of a biased process being implemented in the real world.

Non-technical documentation: a written description or analysis of the benefits or issues associated with the use of an AI system outside of its operational processes. Impact and risk assessments are examples of non-technical documentation.

Pre-trained: refers to machine learning systems that have been initialised by training on a large, general dataset, and can be fine-tuned to accomplish specific downstream tasks.

Technical documentation: a written description of or guide to an AI system's functionality. For instance, technical documentation content may include: usage instructions; technical assumptions about its use and operation; system architecture; and technical limitations. Manuals, code repositories and model cards are examples of technical documentation.

Transparency: refers to the communication of appropriate information about an AI system to relevant people, in a way that they understand. In practice, making reporting procedures transparent requires clearly informing reporters about: how they can expect their report to be processed; how their report is processed; when their report has finished being processed; and any outcomes to which the report can be directly attributed.