

Explaining decisions made with AI

Draft guidance for consultation

Part 3:

What explaining AI means
for your organisation

About this guidance

What is the purpose of this guidance?

This guidance goes into the various roles, policies, procedures and documentation that you can put in place to ensure your organisation is set up to provide meaningful explanations to affected individuals.

How should we use this guidance?

This is primarily for senior executives in your organisation. It offers a broad outline of the roles that have a part to play in providing an explanation to the decision recipient, whether directly or as a part of the decision-making process.

DPOs and compliance teams as well as technical teams may also find the documentation section useful.

What is the status of this guidance?

This guidance is issued in response to the commitment in the Government's AI Sector Deal, but it is not a statutory code of practice under the Data Protection Act 2018.

This is practical guidance that sets out good practice for explaining decisions to individuals that have been made using AI systems processing personal data.

Why is this guidance from the ICO and The Alan Turing Institute?

The ICO is responsible for overseeing data protection in the UK, and The Alan Turing Institute ("The Turing") is the UK's national institute for data science and artificial intelligence.

In October 2017, Professor Dame Wendy Hall and Jérôme Pesenti published their independent review on growing the AI industry in the UK. The second of the report's recommendations to support uptake of AI was for the ICO and The Turing to:

Explaining decisions made with AI | Part 3: What explaining AI means for your organisation

“...develop a framework for explaining processes, services and decisions delivered by AI, to improve transparency and accountability.”

In April 2018, the government published its AI Sector Deal. The deal tasked the ICO and The Turing to:

“...work together to develop guidance to assist in explaining AI decisions.”

The independent report and the Sector Deal are part of ongoing efforts made by national and international regulators and governments to address the wider implications of transparency and fairness in AI decisions impacting individuals, organisations, and wider society.

Organisational roles and functions for explaining AI

At a glance

- Anyone involved in the decision-making pipeline has a role to play in providing an explanation of a decision supported by an AI model's result.
- This includes what we have called the AI development team, as well as those responsible for how decision-making is governed in your organisation.
- We recognise that every organisation is different when it comes to the structure of their AI development and governance teams, and in smaller organisations several of the functions we outline will be covered by one person.
- Many organisations will outsource the development of their AI system. In this case, you as the data controller have the primary responsibility for ensuring that the AI system you use is capable of producing an explanation for the decision recipient.

Checklist

- We have identified everyone involved in the decision-making pipeline and where they are responsible for providing an explanation of the AI system.
- We have ensured that different actors along the decision-making pipeline, particularly those in AI development teams, those giving explanations to decision recipients, and our DPO and compliance teams are able to carry out their role in producing and delivering explanations.
- Where we are buying the AI system from a third party, we know we have the primary responsibility for ensuring that the AI system is capable of producing explanations.

In more detail

- [Who should participate in explanation extraction and delivery?](#)
- [What if I use an AI system supplied by a third party?](#)

Who should participate in explanation extraction and delivery?

People involved in every part of the decision-making pipeline, including an AI model's design and implementation processes, have a role to play in the provision of explanations to individuals who receive the decision that is supported by the result of that model's processing.

This ranges from those involved in the initial decision to use an AI system to solve a problem and the teams building the system, to those using the output of the system to inform the final decision and those who govern how decision-making is done in your organisation. Depending on the organisation, the roles outlined below might be configured in different ways, or concentrated in just one or two people.

Overview of the roles involved in providing an explanation

Product manager: defines the product requirements for the AI system and determines how it should be managed, including the explanation requirements and potential impacts of the system's use on affected individuals.

AI development team: The AI development team performs several functions, including:

- collecting, procuring and analysing the data that you will input into your AI system, which must be representative, reliable, relevant, and up-to-date;
- bringing in domain expertise to ensure the AI system is capable of delivering the types of explanations required. Domain experts could, for example, be doctors, lawyers, economists or engineers;
- building and maintaining the data architecture and infrastructure that ensure the system performs as intended and that explanations can be extracted;
- building, training and optimising the models you will deploy in your AI system, prioritising interpretable methods;

- testing the model, deploying it, and extracting explanations from it; and
- supporting implementers in deploying the AI system in practice.

Implementer: where there is a human in the loop (ie the decision is not fully automated) the implementer relies on the model developed to supplement or complete a task in their everyday work life. Implementers either directly use the model, if it is inherently interpretable and simple, or supplementary tools and methods that enable explanation of the model. The tools and methods provide implementers with a representation of the model's results, such as a score or graph, a suggestion or recommendation, and information about the rationale behind those results. Implementers take this information and consider it together with other evidence to make a decision on how to proceed.

DPO and compliance team: these ensure that the development and use of the AI system are in compliance with regulations and the organisation's own policies and governance procedures. This includes compliance with data protection law, such as the expectation within it that AI decisions are explained to individuals affected by those decisions.

Senior management: this is the team with overall responsibility for ensuring the AI system developed and used within the organisation, or procured from a third party, is appropriately explainable to the decision recipient.

Your AI system may also be subject to external audit, for example by the Information Commissioner's Office (ICO) to assess whether your organisation is complying with data protection law. Data protection includes the expectation that AI decisions are explained to individuals affected by those decisions. During an audit you will need to produce all the documentation prepared and testing undertaken to ensure that the AI system is able to provide the different types of explanation required.

As you move along the decision-making pipeline, through the roles identified above, there will be a certain amount of translation required. That is, the statistical output of the AI system will need to be translated for different audiences. Internally to your organisation this means to the implementer, DPO and compliance team, and to senior management. Externally this means translating what you have performed from a technical point of view into language and reasoning that is comprehensible to the decision recipient and the external auditor.

[AI auditing framework blog](#)

What if we use an AI system supplied by a third party?

If you are sourcing your AI system, or significant parts of it, from a third party supplier, the functions and responsibilities may look different. Whether you do this or build it yourself, you as the data controller have the primary responsibility for ensuring that the AI system you use is capable of producing an appropriate explanation for the decision recipient. If you procure the system from a third party supplier that is off the shelf and does not contain inherent explainability, you may need another model alongside it.

More information on supplementary models and techniques can be found in [Explaining AI in practice](#).

Policies and procedures

At a glance

- Whether you create new policies and procedures or update existing ones, they should cover all the 'explainability' considerations and actions that you require from your employees from concept to deployment of AI decision-support systems.
- Your policies should set what the rules are, why they are in place, and who they apply to.
- Your procedures should then provide directions on how to implement the rules set out in the policies.

Checklist

- Our policies and procedures cover all the explainability considerations and actions we require from our employees from concept to deployment of AI systems.
- Our policies make clear what the rules are around explaining AI-enabled decisions to individuals, why those rules are in place, and who they apply to.
- Our procedures set out directions on how to implement the rules set out in the policies.

In more detail

- [Why do we need policies and procedures for explaining AI?](#)
- [What should our policies and procedures cover?](#)

Why do we need policies and procedures for explaining AI?

Organisational policies and procedures are important for several reasons, they:

- help ensure consistency and standardisation;

- clearly set out rules and responsibilities; and
- support the creation / adoption of an organisational culture.

These are all highly desirable for your approach to explaining AI-assisted decisions to individuals.

You may want to create new policies and procedures, or it might make more sense to adapt and extend those that already exist, such as data protection and information management policies, or broader information governance and accountability frameworks.

How you choose to do this depends on the unique set up of your organisation. What matters is that there is a clear and explicit focus on explaining AI-assisted decisions to individuals, why this is necessary and how it is done. This will help to embed explainability as a core requirement of your use of AI.

What should our policies and procedures cover?

Both your policies and procedures should cover all the explainability considerations and actions that you require from your employees from concept to deployment of AI decision-making systems. In short, they should codify what's in the different parts of this guidance for your organisation.

Your policies should set out the what, why and who of explaining AI-assisted decisions to individuals, ie they should make clear what the rules are, why they are in place, and who they apply to. Your procedures should set out the how of explaining AI-assisted decisions to individuals, ie directions on how to implement the rules set out in the policies.

The table below summarises the key areas to cover in your policies and indicates where it is beneficial to have an accompanying procedure.

This is only a guide. Depending on what you already have in place, you may find it is more important to provide more (or less) detail in certain areas than others. There may also be additional aspects, not covered in this table that you wish to cover in your policies and procedures.

You should consult relevant staff when drafting your policies and procedures to ensure that they make sense and will work in practice.

	Policy	Procedure
Policy objective	<p>Explain what the policy seeks to achieve – the provision of appropriate explanations of AI-assisted decisions to individuals. Even if you incorporate your requirements around explaining AI-assisted decisions into an existing policy or framework, you should ensure that this particular objective is explicitly stated.</p>	N/A
Policy rationale	<p>Outline why the policy is necessary. You should cover the broad legal requirements, and any legal requirements specific to your organisation or sector. You should also cover the benefits for your organisation, and, where relevant, link the rationale to your organisation’s broader values and goals.</p>	N/A
Policy scope	<p>Set out what the policy covers. Start by clarifying what types of decision-making and AI systems are in scope. Say which departments or parts of your</p>	N/A

	<p>organisation the policy applies to. Where necessary, explain how this policy links to other relevant policies your organisation has, signposting other organisational requirements around the use of AI systems that are not within this policy's scope.</p>	
<p>Policy ownership</p>	<p>Make clear who (or which role) within your organisation has ownership of this policy, and overarching responsibility for the 'explainability' of AI decision-support systems. Explain that they will monitor and enforce the policy.</p>	<p>Detail the steps that the policy owner should take to monitor and enforce its use. Set out what checks to make, how often to make them, and how to record and sign-off this work.</p>
<p>Roles</p>	<p>Set out the specific roles in your organisation that have a stake or influence in the 'explainability' of your AI decision-systems. Describe the responsibilities of each role (in connection with explaining AI-assisted decisions to individuals) and detail the required interaction between the different roles (and departments) and the appropriate reporting lines, all the way up to the policy owner.</p>	<p>N/A</p>

Impact assessment	<p>Explain the requirement for explainability to be embedded within your organisation's impact assessment methodology. This is likely to be a legally required assessment such as a Data Protection Impact Assessment, but it could also form part of broader risk or ethical assessments.</p>	<p>Explicitly state the need to conduct the assessment (including considering explainability) before work begins on an AI decision-support system. Describe how to make the necessary explainability assessments including consideration of the most relevant explanation types, and the most suitable AI model for the context within which you will be making AI-assisted decisions about individuals.</p>
Awareness raising	<p>Explain the importance of raising awareness about your use of AI-assisted decisions with your customers or service users. Set out the information you should communicate including why you use AI for decision-making, where you do this, and simple detail on how it works.</p>	<p>Detail the specific approach your organisation takes to raising awareness. Clarify where information for your customers will be hosted, how it will be communicated or made available (eg which channels and how often), and which roles or departments are responsible for this.</p>
Data collection	<p>Underline the need to consider explanations from the earliest stages of your AI model development, including the data collection, procurement and preparation phases. Explain why this is important with reference to the benefits of</p>	<p>Set out the steps to take at the data collection stage to enhance the explainability of your AI model, including assessing data quality, structure, feature interpretability, and your approach to labelling.</p>

	interpretable and well-labelled training data.	
Model selection	Explain how considerations of explainability were factored in to the selection of your AI model in its development stage and how the algorithmic techniques you chose to use are appropriate for both the system's use case and its potential impacts.	Set out the steps taken to weigh model types against the priority of the interpretability of your AI system. Signpost how you ensured that the selected model is appropriate to fulfil that priority.
Explanation extraction	Set out the different types of explanation and outline the requirements to obtain information relevant to each one.	Signpost the various technical procedures your organisation uses to extract the 'rationale' explanation from your AI models (eg how to use a local explanation tool such as LIME). Describe how to obtain information on the other explanations, who to obtain this from, and how to record it.
Explanation delivery	Explain the need to build and deliver explanations in the way that is most meaningful for the individuals your AI-assisted decisions are about.	Detail how to prioritise the explanation types, how to translate technical terminology into plain language, the format in which to present the explanation (eg in layers), and how to assess appropriate timing of

<p>Documentation</p>	<p>Clearly state the necessity to document the justifications and choices made through the whole process of developing/acquiring and deploying an AI decision-support system. Outline the requirement to document the provision of explanations to individuals, and clarify what information to record (eg URN, time stamp, explanation URL).</p>	<p>delivery (eg before or after a decision).</p> <p>Set out the standardised method by which all stakeholders can record their justifications and choices. Explain how your organisation keeps an audit trail of explanations provided to individuals, including how this can be accessed and checked.</p>
<p>Training</p>	<p>Set requirements for general staff training on explaining AI decisions to individuals, covering why it's necessary, and how it's done. Identify roles that require more in-depth training on specific aspects of explaining AI-assisted decisions, such as preparing training data or extracting rationale explanations.</p>	<p>N/A</p>

Documentation

At a glance

- It is essential to document each stage of the decision-making process in order to provide a full explanation for how a decision was made.
- In the case of explaining AI-assisted decisions, this includes both documenting the processes behind the design and implementation of the AI system and documenting the actual explanation of its outcome.
- The suggested areas for documentation may not apply to all organisations, but are intended to give an indication of what may be helpful in providing the evidence you might need to establish how a decision was made.
- The key objective is to provide good documentation that can be understood by people with varying levels of technical knowledge and that covers the whole process from designing your AI system to the decision you make at the end.

Checklist

- We have documented what we are required to do under the GDPR.
- We have documented how each stage of our use of AI contributes to building an explanation, from concept to deployment.
- Our documentation provides an audit trail about who we give explanations to, and how we provide them.

In more detail

- [What documentation is legally required under the GDPR?](#)
- [What documentation should we provide to demonstrate the explainability of our AI system?](#)

What documentation is legally required under the GDPR?

Article 5 of the GDPR says that “The controller shall be responsible for, and able to demonstrate compliance with, paragraph 1 (‘accountability’).”

Article 12 of the GDPR requires you to provide information to the data subject in “concise, transparent, intelligible and easily accessible form, using clear and plain language...”. It also states that you can provide the information “in combination with standardised icons in order to give in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing.”

Article 13 of the GDPR requires you to provide your data protection officer’s contact details, which aligns with the responsibility explanation; the purpose for which you are processing the data subject’s personal data, as well as the legal basis for that processing, which in many cases should form part of your explanation; and the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject. You must document all of this to ensure you remain accountable.

Article 14 of the GDPR applies to cases where you have not obtained personal data from the data subject directly. You should provide data subjects with the following information in addition to that required under Article 13, within a reasonable period after obtaining the personal data, but at the latest within one month, having regard to the specific circumstances in which the personal data are processed:

- What categories of personal data you are processing.
- The source from which you obtained their personal data, and if applicable, whether it came from publicly accessible sources.

See Article 14 of the GDPR for further information on when it is not required to provide this information to the data subject. This includes when you are processing personal data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. This is subject to certain conditions and safeguards.

Article 15 of the GDPR gives data subjects an additional right of access to the personal data that you hold on them. This means you should document how you will provide them with a copy of the personal data that you process.

Article 21 of the GDPR gives data subjects the right to object at any time, on grounds relating to their particular situation, to processing of personal data concerning them, including profiling. This means you should document how you ensure data subjects are aware of this right, and how you record if they have exercised this right.

Article 22 of the GDPR gives individuals the right not to be subject to a solely automated decision producing legal or similarly significant effects, unless certain conditions apply. It obliges organisations to adopt suitable measures to safeguard individuals, including the right to obtain human intervention, to express their view, and to contest the decision. This means you need to document how you will do this.

Article 30 of the GDPR helps organisations to fulfil the accountability principle. It states that an organisation shall "...maintain a record of processing activities under its responsibility."

Article 35 of the GDPR requires organisations to carry out a Data Protection Impact Assessment (DPIA) when they are doing something with personal data, particularly when using new technologies, which is likely to have high risks for individuals. A DPIA is always required for any systematic and extensive profiling or other automated evaluation of individuals' personal aspects which are used for decisions that produce legal or similarly significant effects.

You can reconcile some of these documentation requirements through your organisation's **privacy notice**, which must contain certain elements:

- The lawful basis for the processing – one or more of the bases laid out in Article 6(1) of the GDPR.
- If applicable, the legitimate interests for the processing – these are the interests pursued by your organisation or a third party if you are relying on the lawful basis for processing under Article 6(1)(f) of the GDPR. You could also include a link to the record of your assessment of whether legitimate interests apply to the particular processing purpose.
- The rights available to individuals regarding the processing – eg access, rectification, erasure, restriction, data portability, and objection. The rights vary depending on the lawful basis for processing. Your documentation can reflect these differences.
- If applicable, the existence of automated decision-making, including profiling. In certain circumstances you will need to tell people about the logic involved and the envisaged consequences.

- If applicable, the source of the personal data. This is relevant when you didn't obtain personal data directly from an individual.

You should be aware that different documentation requirements may apply for law enforcement processing under Part 3 of the DPA 2018 and for intelligence services processing under Part 4 of the Act.

Documentation guidance in Guide to the GDPR

What documentation can help us to demonstrate the explainability of our AI system?

The list below should support you to provide an explanation to the decision recipient, and maintain an audit trail about who you give explanations to, and how you provide them.

Decision to use an AI system

- What the system is intended to be used for, so you can explain to the decision recipient the purpose for which the deployment of the AI system is envisioned.
- Who the ultimate decision recipient will be.
- What the AI system you have chosen will do from a technical perspective, in a way that the decision recipient can also understand.
- How the specifications of the system were determined, and by whom – as well as alternative specifications that were considered, and why they were not chosen.
- If your AI system has been procured from a third party or outsourced, how you can change or retool its specifications to meet changing performance and explainability needs over time.
- What trade-offs are involved for the data subject whose data will be used in the model and who will often also be the decision recipient. For example, the data subject's personal data may be used in training the AI system to produce a highly accurate model, but this use of data may not be in the data subject's interest.
- What the demographics and background of the development team are, in order to be aware of the diversity within the team responsible for designing, deploying and maintaining the system, and how that may impact on the decision for the decision recipient.

- The domain in which you will be using the AI system, and how this system has been tested and validated in that domain.
- Any other impact assessment relevant to your domain, in addition to the data protection impact assessment mentioned above under the GDPR requirements.
- The people within the organisation who have responsibility for providing explanations along the design and implementation pipeline of your AI system.

Explanation types this supports: rationale, responsibility, fairness, safety and performance, impact.

Scoping and selecting explanation types

- The processes you have set up to optimise the end-to-end accountability of your AI model.
- The setting or sector in which your AI model will be used and the bearing this has on the types of explanation you will offer.
- Justification for the explanation type(s) you have prioritised, based on your AI system's potential impact.
- Justification for how you have chosen to handle the remaining explanation types that will not be prioritised.
- How you have set up process-based and outcome-based aspects of the explanations types that you will offer.
- Justification for the depth or comprehensiveness of the explanation you will provide, given the potential impacts of the system. This includes the general risks of deploying the system, and the risks for the specific person receiving the AI-assisted decision.
- Who within your organisation is responsible for selecting the appropriate type(s) of explanation.

Explanation types this supports: rationale, responsibility, data, fairness, safety and performance, impact.

Data collection and procurement

- Documentation of where the data came from, for what purpose the data was originally collected, and for whom – this will help you explain to the decision recipient how relevant the data you have used is to the decision the AI system has made about them.
- Description of the components of the dataset together with a brief summary of why each element is being included.

- Evidence that the data is representative of the people that will be subject to the AI decisions you will make – for example through consideration by a domain expert.
- How you have made sure that that the data is reliable, accurately measured and obtained from a source of integrity.
- How you have examined your datasets for any potential inherent bias.
- Evidence that the data is recent, up-to-date and appropriately timely given the rate of change in the underlying distribution you are modelling. This will demonstrate that you have accounted for concept drift and data fluctuation from the start. The rate of change will depend on the domain you are operating in, and the specific case you are considering.
- If you use synthetic data, provide documentation on when and how it was created, and its properties – this helps you explain and justify to the decision recipient why created data has been used in the training of the model and why this is appropriate.
- The risks associated with using the data, and the risk to those whose data is included.
- How individuals can opt out of being included in the data used either to train or run the AI system.

Explanation types this supports: data, fairness, safety and performance, impact.

Data pre-processing

- How, especially in cases where social and demographic data is involved, you have ensured that the pre-processing of your data has produced a feature space which includes variables that are understandable, relevant and reasonable and that does not include variables that are opaque or difficult to understand in relation to your model's target variable.
- How you have labelled the data and why you have labelled it in that way. This should include tagging and annotating what a piece of data is, and the reasons for that tag.
- How you have mitigated any bias in your data through pre-processing techniques such as re-weighting, up-weighting, masking, or excluding features and their proxies.
- If you are using 'raw', observed, or unconventional data, documentation of what interpretively significant feature such data is supposed to indicate about the individual whose data is being processed and evidence that this has been included in the metadata.

- Who within your organisation is responsible for data collection and pre-processing.

Explanation types this supports: responsibility, data, fairness, safety and performance, impact.

Model selection

- The specific interpretability or transparency standards, conventions and requirements of the domain in which your AI system will be applied.
- How the specific type of application and the impact on individuals informs the type of model you choose.
- Explanation of how the types of data you are using, for example social or demographic data, or biophysical data, have influenced your model selection regarding its interpretability.
- Whether your use case enables you to use maximally interpretable algorithmic techniques, and if not, why not.
- When using 'black box' models, identification of the risks of using them, and provision of supporting evidence that your team has determined that your use case and your organisational capacities and resources support the responsible design and implementation of these systems.
- When using opaque algorithmic techniques such as 'black boxes', explanation of how the supplementary tools that you will use to explain the model provide a domain-appropriate level of explainability. Your documentation should demonstrate how the supplementary tool will mitigate the potential risks of using a 'black box' system, and how the use of the tool will help you to provide meaningful information about the rationale of any given outcome.
- If you use 'challenger' models alongside more interpretable models, justification of the purpose of these models and how you use them.

Explanation types this supports: rationale, responsibility, data, fairness, safety and performance, impact.

Model building, testing and monitoring

- The accuracy rate and other performance metrics chosen for the model, as well as any tuning of cost ratios to constrain error allocation, and how and why these have been selected. You should be able to explain to the decision recipient how this choice may affect the decision that you have made about them.
- If relevant, group-specific error rates and how the model has been tuned to redress any significant imbalances.

- How potential for biases in the model design have been assessed and monitored and what measures have been taken to mitigate those you have identified.
- How the model has been tested, including test results and which portions of the data have been used to train, test the model and holdout data.
- How frequently the model will be monitored, updated and re-examined after being deployed in the real world.
- How often the training data will be updated after model production and deployment. You should also document what you have put in place to establish the appropriate frequency of updates.
- Evidence that tracks each time the model has been updated, and how each version has changed, so that you can explain to the decision recipient how that particular version of the model came to the decision, and why this might differ from the output of a subsequent or prior model.

Explanation types this supports: rationale, data, fairness, safety and performance.

Tools for extracting an explanation

- When using more inherently interpretable models, the measures you have taken to ensure optimal explainability, for example the sparsity constraints placed on the feature space so that explanations can remain human understandable.
- When using supplementary interpretability tools for 'black box' models, an outline of the local and global techniques that you have used to provide explanations. This may be in the form of detailed specifications of the supplementary tools you have used.
- How you plan to combine these different explanation tools to produce meaningful information about the rationale of the system's results.
- Who is responsible for ensuring that the explanations generated by the supplementary tools are accessible to the people they are intended to inform.
- How you will translate the statistical output of your model and supplementary tools into a plain-language explanation, for example by establishing and documenting appropriate implementer training and providing users with comprehensive guidelines for responsible implementation.

Explanation types this supports: rationale, responsibility.

Explanation delivery

- Justification for which explanation types you will prioritise when you deliver the explanation to the affected individual, given the contextual factors you determine to be relevant in the particular case you are considering.
- Justification for how you prioritise the remaining explanation types.
- Documentation describing the training you have provided to implementers to enable them to use the model's results responsibly and fairly.
- Documentation describing how the implementer will be presented with the model's result, including:
 - how you present performance metrics and error rates for the model as a whole and for sub-groups if appropriate;
 - how you present uncertainty measures like error bars and confidence intervals;
 - how you use visualisation tools and present indicators of relative variable important or variable interactions; and
 - in the case of 'black box' models, how you present information from supplementary tools as well as indicators of the limitations and uncertainty levels of these tools.
- The reasonable adjustments you will make for the form in which you deliver the explanation, as required under the Equality Act 2010.
- The information you will proactively share with your customers and stakeholders, so that they are able to make informed choices in advance of engaging with the decision-making process.
- Who decision recipients can contact to query a decision.

Explanation types this supports: rationale, responsibility, data, fairness, safety and performance, impact.

The following papers can provide some ideas for how you might document these processes:

[FactSheets](#)

[Datasheets for Datasets](#)

[Model cards for model reporting](#)

[AI auditing framework blog](#)

[Understanding Artificial Intelligence Ethics and Safety](#)