

# DGX Working Group on Artificial Intelligence

**MVP - good practice for new AI products and services for governments and central government departments**

July 2022

<b>Introduction</b>	3
Methodology	3
Minimal Viable Product (MVP)	4
Using the MVP	4
Steps to help guide project teams in buying or building new AI products or services	4
<b>Use Cases</b>	4
<b>1. Mitigate risks with project design, governance and strong teams</b>	6
<b>2. Project Design</b>	6
Understand and use your AI values	7
Start with a discovery phase	8
Build your AI team, set out roles and responsibilities	8
Funding	10
Develop a communications narrative and managing your stakeholders	10
Get your data ready for AI and plan your AI modelling phase	11
<b>3. AI Programme Governance</b>	12
Frameworks to support governance	12
Steps that can help guide teams on governance	14
<b>4. Manage and Assure your Data</b>	15
<b>5. Laws and regulations</b>	16
<b>6. Ensure inclusivity and preventing bias</b>	17
<b>Next steps</b>	18
<b>Resources</b>	19

## Introduction

The pandemic was a stress test on governments worldwide and their ability to deliver digital services securely and at scale. As the case for digital government is more clear than ever before, governments are looking to improve digital services at national and sub-national levels to build resilience against future shocks and embed trust between citizens and the state. Artificial Intelligence (AI) is now viewed as part of the toolbox of digital tools that support better service delivery. However, leveraging artificial intelligence (AI) technologies also comes with ethical and operational considerations for teams in government and central government agencies.

In response, the Digital Government Exchange (DGX) Working Group on AI was established in December 2021 to share experiences, build collective approaches and exchange knowledge of Artificial Intelligence (AI) within the DGX network.

The working group is currently chaired by the UK's Government Digital Service (GDS). It comprises members from Australia, China, Germany, Japan, Singapore, Sweden, and the World Bank and World Economic Forum.

The objectives of the DGX Working Group on Artificial Intelligence (AI) were to:

- develop draft good practice to support teams in government and central government agencies
- leverage the expertise of its members to create real-world use cases for the DGX community

## Methodology

The group decided to approach this project using agile practices, co-designing a survey and through several co-working sessions to understand the unique experience of DGX member country contexts and to follow agile best practice through continuous improvement and iteration of the MVP based on feedback.

The group and others in the DGX network filled out the co-designed survey. The information gathered from these surveys validated the general approach and provided the necessary information to form the basis of the best practice. Finally, desk research was carried out, looking at academic and multilateral papers on AI frameworks and country case studies on AI.

## Minimal Viable Product (MVP)

We have created a [Minimum Viable Product](#) (MVP) - a piece of good practice to help teams consider what to do when they build or buy AI projects. The areas in this MVP are based on the common thematic areas derived from the co-designed survey as well as secondary research, looking outward at good practices from the international GovTech and AI communities.

This MVP is for:

- Senior Civil Servants who want to understand more about how their teams can adopt AI responsibly and appropriately and to help them make decisions that they are responsible for
- programme managers who are looking at building AI into their product or service
- procurement officials who may be involved in buying AI based technology to help them understand what's expected when working with private sector organisations
- officials developing programmes that will include an AI workstream

## Using the MVP

The MVP can be used in 3 ways :

1. To analyse the common challenges and considerations for governments are working with, as they implement their own AI initiatives.
2. To understand where there are unique or country-specific challenges based on context, structure or other factors.
3. To 'self-serve' through a series of thematically based criteria that a public sector team can assess itself against to understand what they need to consider when building and/or buying AI products.

It is essential to reflect that this work is an MVP, we expect to iterate this work based on feedback from across the DGX network.

## Steps to help guide project teams in buying or building new AI products or services

The steps outlined under each section are designed to help guide you through the AI delivery process, mitigate the introduction of any cognitive or other types of biases and manage expectations. They are by no means definitive but rather steps you need to consider as you progress throughout your journey.

## Use Cases

AI DGX members have also shared use cases on various AI projects. These are attached to this MVP.

## What is Artificial Intelligence (AI)?

In broadest of definitions, artificial intelligence (AI) refers to systems that imitate or mimic human intelligence to perform actions. Some AI systems can constantly improve themselves through their interactions with users and through the information they collect.

AI systems work by initially absorbing labelled training data, looking for data patterns, and using these patterns to “generate outputs, such as content, predictions, recommendations, or decisions influencing the environments they interact with.”<sup>1</sup>

Some examples might include;

- using chat bots to provide real-time guidance for common government services, reducing wait times
- tracking spread of a disease, like COVID, and ensuring that appropriate resources are available to treat patients areas that might need it most, or treating patients with heart conditions<sup>2</sup>
- reducing the work-load in agencies by using machine-based risk assessment in e.g. business registration and income tax
- receiving automated benefits for specific life events like childbirth, retirement and bereavement
- AI supported language processing to support writing documents

## Opportunities and challenges of AI

Many governments are working out how to capitalise on the opportunities AI offers in areas like service delivery whilst addressing the challenges it raises around ethical considerations, such as minimising bias, inaccuracies, and unintended consequences of using AI.

Artificial Intelligence (AI) can boost productivity by freeing up people to work elsewhere on value added products and services, drive efficiencies by easing the operational burdens of routine or labour -intensive activities, and accelerate innovation by creating time and space for collaboration and connection. AI can also be used to support user needs through its ability to collect more and better qualitative and quantitative data.

In the private sector AI-based automation is slowly being included as one of the technology options to help it manage its service. Its adoption is evident in parts of the private sector like banking, where users may no longer expect to proactively apply for certain essential products such as credit card loans, they can now just apply online. However, anecdotal

---

<sup>1</sup> [REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL: Laying Down Harmonised Rules on Artificial Intelligence \(Artificial Intelligence Act\) and Amending Certain Union Legislative Acts](#)

<sup>2</sup> [3D heart scans on the NHS to speed up disease diagnosis](#)

evidence suggests in the public sector, users do not hold the same level of comfort with AI and the risks are more profound. Governments and public sector organisations need to take a more values based approach to ensure that any bias is mitigated by design.

For digital leaders, the challenge is threefold:

- the lack of quality data, in machine readable formats that conforms to metadata standards, underpinned by a robust, secure infrastructure to support AI models,
- the technical expertise required to support the delivery of AI-based projects, and
- the potential ethical challenges and operational difficulties AI products and services pose.

## 1. Mitigate risks with project design, governance and strong teams

To address some of the issues raised, digital leaders should focus on creating a set of core values and standards to be agreed and adhered to across government for all AI projects, in order to:

- set strong project foundations with a common understanding of what the ethical deployment of AI technology looks like
- develop appropriate levels of governance to ensure security measures are in place and build in accountability
- build in robust levels of public participation and consultation when required to mitigate risks around AI bias where it involves different types of data (for example, using different levels of personal data)
- consider differing risk tolerances for different applications (for example the difference in risk tolerance for using AI for non-critical note keeping via automatic transcription of meetings, versus using AI to inform decisions that may influence citizen payments)

## 2. Project Design

The foundation of all AI projects is to start with a strong understanding of the values that need to shape and underpin each aspect of an AI project delivery during its design, delivery and long-term development.

Most countries agree that there must be a set of agreed values or principles to guide AI projects. For many, these guidelines are still in the process of being developed. Where there are gaps or no existing frameworks, countries have based their values on legal or constitutional foundations as the basis for all AI projects, setting out the values that drive outcomes and outputs. An implied or specific value common to most is accountability. Governments or government agencies' project teams are encouraged to inform the public that AI is being deployed to automate the service to support legal and ethical requirements.

Some of the values that drive AI delivery could be:

1. **Accountability:** where there are effective governance and oversight mechanisms for any project.<sup>3</sup>
2. **Transparency:** the actions, processes and data are made open by publishing information about the project in a complete, open, understandable, easily-accessible and in a free format for users, and how decisions are made throughout the project
3. **Sustainability:** ensuring sustainability of the model and how it mitigates bias over the longer term
4. **Fairness:** mitigate biases which may influence your model's outcome and ensure that the project and its outcomes respect the dignity of individuals<sup>4</sup> and communities that may be impacted by decisions
5. **Privacy and Security:** (in the area of data protection)
6. **Diversity and Inclusion:** ensuring diversity and inclusion not only in regards to the AI solution itself but also the project team working on its deployment. Diversity safety mechanisms ultimately minimise forms of bias and protect marginalised groups which may be underrepresented or misrepresented.<sup>5</sup>

There may be other values that are context or country-specific.

## Understand and use your AI values

Governments should:

- understand what values and ethical frameworks your government has to support AI projects and clarify terms and values that may not be clear with the owners of your country's framework or guidance or other appropriate experts and identify AI ethics teams within your organisation/government
- where there are no values specifically relating to AI projects, understand what legislation or constitutional elements may inform the values that your product or service may need to follow
  - engage with relevant thematic experts (for example, data or AI ethicists, academia, other public sector experts or the private sector) to get a rounded view of what elements should inform AI values
  - look at best practice guidance on implementing a set of AI ethics (for example, the [Alan Turing Institute](https://www.alanturinginstitute.org/)) to cross reference against country specific considerations such as legislation or within the regulatory environment
- establish agreement from a project perspective on how AI ethics issues, such as evaluating datasets for bias, are applied and how AI ethics issues are resolved

---

<sup>3</sup> <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020>

<sup>4</sup> <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020>

<sup>5</sup> 'Successful system design starts with ensuring key questions about "why, for whom, and by who" are appropriately answered and with continuous revisiting throughout the development process.'  
<https://www.weforum.org/whitepapers/a-blueprint-for-equity-and-inclusion-in-artificial-intelligence>

## Start with a discovery phase

Projects involving AI should:

- start with a discovery phase, which includes following and applying Agile at the project's inception point.
- understand what problem you are trying to solve - your projects should be hypothesis-led, rather than technology lead
- assess if AI is the right type of solution to meet your needs and solve the problem you have identified
- apply user research in every aspect of your project to establish user needs ('user needs' are the needs that a user has of a service, and which that service must satisfy for the user to get the right outcome for them) and uncover potential harms for both users and stakeholders
- make an initial assessment of potential savings through streamline business processes or capability-enhancements to innovate areas like service delivery
- understand the broader strategic, operational impacts and the wider digital ecosystem -consider where your AI program sits in the organisation, any other services it may impact, and what services it may connect to more broadly across government as well as possible challenges in terms of supply of data or legal considerations (see relevant sections below for more details)
- outline how diverse and inclusive participation will be built into the project lifecycle (this includes involving the whole multidisciplinary team as well as looking outwards to participation and input from relevant stakeholders and users) to ensure there is collaborative deliberation on project choices - AI might not necessarily be the best solution to the defined problem given risks/cost-benefit/impact analysis
- do an early assessment of risks (with possible mitigations) to understand what issues might emerge

Projects should also outline public benefits. These questions may be helpful to reflect on before starting a project (example from the UK):

- map out the direct benefits for individuals in this project (for example, saving time when applying for a government service)
- how does the project deliver positive social outcomes for the wider public
- how can you measure and communicate the benefits of this project to the public
- understand what are the groups that would be disadvantaged by the project or would not benefit from the project - set out what mitigations you might apply

## Build your AI team, set out roles and responsibilities

AI is a relatively new field for many governments and central government agencies. A growing practice within central government agencies is focused on developing multi-



disciplinary teams to support governance practices, particularly related to interventions relevant to data collection, model development, deployment, and stakeholder interaction.

A newer role has emerged in recent years - an AI engineer - which requires individuals to gain familiarity in aspects such as data pipeline setup, Machine Learning model deployment and optimisation, and traditional software engineering skills. Such engineers work alongside data scientists and product managers to develop, scale and fine-tune AI products.

Projects involving AI should:

- set out who is the owner of the model - and who will maintain it post-delivery
- develop a robust plan for knowledge transfer and training to ensure the model can be sustainably managed including
- consider from a model design perspective who is responsible for every part of the project for the AI outputs (see capability) and how implementers and designers of AI systems will be held accountable.
- as part of the design of the roles and responsibilities, set out who can change and modify the code
- at a project level, establish who you may need to contact on AI issues at both organisational and strategic matters on legal and data issues
- ensure diversity of the team in skills, subject expertise, and lived experiences

Multidisciplinary teams are the best way to deliver digital projects, especially those that involve AI. It allows everyone to understand what is required, builds skills across the team on AI projects and supports better stakeholder engagement. Different voices in a team can also help reduce bias.

Teams need to have the following skills or roles:

- **data architects** to develop and shape the use of data to meet user and business needs
- **data scientists** to understand and leverage data value to meet business and user needs. They will also be familiar with methods for cleaning and maintaining, and assuring data, and will help with good data standards for ongoing project management
- **data engineers** to help develop cybersecurity, software and infrastructure development skills to support scaling, integration and maintaining AI systems (for example, cloud engineering, distributed systems)
- **user researchers** to understand what users needs are throughout the delivery cycle
- **product management skills** to help teams understand user requirements, develop strategies for products and platforms, and shape the delivery team's efforts towards a sustainable and effective direction. This role plays a significant role influencing and promoting the product across all levels of the organisation.

- **business focused Subject Matter Experts (SMEs) and policy experts** to ensure that the business, policy and IT are working hand-in-hand to enable a business driven solution (rather than tech driven)
- **subject experts in the domain that you are building the AI project or service** for, to help teams understand what the business challenges are (for example, if you are making a social care service, a social worker will help you understand business needs)
- **a role dedicated explicitly to stakeholder management /communications** to support transparency and showcase accountability on the use of data (some of these skills in these roles already exist in agile, multi-disciplinary teams). This role should also represent the project at different levels within the organisation to promote, support and influence decision making processes.
- **legal experts to understand the rules, relevant national or international laws and regulations**
- **an ethicist** to assess and make decisions on any inputs into the AI model. An ethics role should be situated appropriately in the organisational hierarchy so that they support leaders to make the right decisions at various points throughout delivery.

Simply having these roles will not guarantee success. The impact of AI projects extends beyond project teams, and technical model owners and decision makers need to take into account the wider impact on policy/business areas.

## Funding

At the outside of any project, teams or organisations should:

- consider how much it will cost to build or buy an AI product or service and train and educate staff to support and maintain AI projects and services over the longer term
- understand and build in costs for long-term maintenance of both the product and the team itself
- ensure there is appropriate funding to cover regular user research (for example, the UK's approach is at least 2 hours every 6 weeks) of the product to inform future iterations and improvements
- include a requirement to demonstrate the highest standard of AI ethics when bidding for funding for AI projects

## Develop a communications narrative and managing your stakeholders

Project teams should consider:

- how open and transparent reporting and engagement is crucial to success and identify which groups, communities, civil societies, NGOs, academia and the public who have an interest in algorithms

- set a clear plan for engaging and communicating with these stakeholders from the outset of your work, setting out how you will engage with them in the design process
- design and set out a clear explanation of the model that everyone in the team can understand and articulate it to users (it might be helpful to do user research to test that messaging resonates with end-users)
- ensure that your communications are in plain language to enhance accessibility, secure participation and help gather feedback from a wider section of society

Teams should also develop a narrative for users to explain the use of AI throughout the development cycle. It should cover the following points:

- set out why and how you are using AI in this project and your business objectives for your organisation and subject area and how it supports user needs (providing convenience, greater efficiency, and more innovative solutions)
- outline what measures you have taken to support transparency, and accountability (meets legal requirements and legislation) and sets out how users expect their data to be used
- create a road map that takes into account stakeholders' needs and sets out progress against the goals you've described
- set out your long-term plans for the on-going development of the project
- establish and communicate upfront how you will address ethical and safety concerns and how they are weaved into the management and design of the algorithm (and your model)
- outline clearly the risk criteria developed against the project, with well signposted information how your governance structures will be leveraged to mitigate each risk
- create channels of "contestability" along the AI implementation and usage

## Get your data ready for AI and plan your AI modelling phase

The success of any AI project begins with quality data. Teams should have a data governance system in place:

- assess the quality of data to ensure that it is valid to the project, relevant to the problem you are trying to solve, representative of your users, and consistent (format and from the data collected)
  - what are your data sources - how accurate and how complete is the data
  - what data labels are going to be used and how will they be tested to prevent bias
- ensure that any data collection considers user needs - is it representative of the users you want to reach
- understand what data structure you need, how often the data is updated, where it's located, and how quickly you can access it

- Identify what type of databases you need to store your data on (your choice of databases will depend on the different data sources required to inform your model and the complexity of the project)
- ensure that plans are in place to support models trained and tested on relevant, accurate, and generalisable datasets
- ensure that you have enough data for the model to learn from - and that you have permission to use it
- understand if there are any privacy and legal considerations or constraints - does it include any sensitive data, for example
- consider what modelling approaches could be suitable based on the available data

**Responsible AI testing microservices - Singapore's Government Technology Agency (GovTech)** The "black-box" nature of machine learning models, especially neural networks, can be problematic as AI-driven decisions cannot be explained and may be unknowingly biased. GovTech's Data Science and AI Division (DSAID) is collaborating with Meta to develop comparable metrics and testing tools for explainability and fairness, applicable to deep neural networks used in computer vision.

### 3. AI Programme Governance

Governance means several different things to different parts of government. This section focuses on the systems and methods by which an AI project and project strategy is defined, authorised, and monitored and that the model is achieving its business objectives.

#### Frameworks to support governance

Some governments take a risk-based governance framework specific to AI projects and projects. Risk-based governance frameworks help set out a mechanism to help identify and handle risks, particularly those involving different user groups and stakeholders. It supports a more inclusive approach to set out, frame, gauge, manage or mitigate and communicate essential risk issues, often marked by complexity. For those engaged with the private sector to either buy or build AI projects, a risk-based approach is weaved into the contractual agreements, and the level of accountability within the project or project is tiered according to the risk of the system.

A strong element of both approaches - either using existing frameworks or AI-specific risk-based governance frameworks - is the focus on user experience and user research.

There are various approaches that have been taken. In a programme-based example, one government has appointed an autonomous AI ethics group to create suggestions for ethical

implementation of the project. The group is led by the central government agency and includes stakeholders from the public sector, academia and NGOs.

Another way to govern AI projects is through a process-based governance framework<sup>6</sup> (PBG), as a basis for AI model development. A PBG looks to integrate the values and principles that will underpin a team's AI design process alongside the operational and technical steps of AI development to provide a holistic view of the governance procedures and reporting lines against the control flow as it is executed and evaluated. A PBG could include:

- team members/roles involved in a governance/oversight capacity under each action
- which stages within the workflow require intervention and targeted consideration
- specific timeframes for follow-up actions, assessments and monitoring and evaluation
- well defined principles for recording and logging activity

Other approaches focus on using existing common digital, data and technology principles or structures designed for Digital and ICT projects which can also take into risk-based governance structures and accountability governance.

As part of project governance, a growing practice suggests that people responsible for legal, policy and risk management, should work closely alongside project and data-science teams and subject matter experts who understand the wider business context to mitigate risks, particularly around AI model design and data flows. These roles are sometimes included as part of a multi-disciplinary service team. As a general practice, a multidisciplinary service team is responsible for ensuring that the system hosting the algorithm meets existing government standards, including policies on cybersecurity and cost-effectiveness. The service team must also assess the algorithm's potential impact and abuses in the context of its application within the public sector and work with approving authorities and domain experts within the government to mitigate potential risks.

**Working with data:** Nigeria, Zambia, Mozambique and Democratic Republic of the Congo are working on a project on learning algorithms to better predict population growth to help conduct censuses in fragile and conflict zones with little data availability. A machine learning algorithm analysed satellite image data to identify features such as settlement boundaries, buildings, transport networks, waterways and industrial areas to predict the population density of an area. A hybrid census model was developed and officials were able to better plan vaccination campaigns and the roll-out of other services.<sup>7</sup>

---

<sup>6</sup> [Alan Turing Institute AI Ethics Guide](#)

<sup>7</sup> [Supporting Artificial Intelligence with geospatial data in Nigeria: GRID3 joins Data Science Nigeria's AI Bootcamp](#) and <https://www.nature.com/articles/s41467-022-29094-x> (see also attached case studies)

## Steps that can help guide teams on governance

The following steps should be taken to ensure good governance of an AI project.

### *Adopt risk management approaches:*

- use Agile programme practices to ensure the user experience is woven into delivery
- consider a risk prioritisation plan with specific project controls throughout the delivery and post-delivery cycle, such as evaluating datasets for bias and clearly explaining your model's performance measure
- decide how and when you evaluate the model during its design, development and post-deployment and what approaches can be applied to mitigate risks like bias, interpretability
- establish a data reporting mechanism that captures information on how data flows are managed and operated across the different stages of the project
- engage with Civil Societies and NGOs, academia to ensure their views are incorporated into the governance of the project to support transparency and mitigate in risks in bias
- ensure mechanisms are in place for users to report issues and concerns across different stages of the project
- set out how any 'human review is included as part of the governance of the project

### *Adopt oversight and assurance processes:*

- as you progress through building and evaluating your machine learning model and deploying and maintaining your model, plan how your project will be assured by AI experts in ethics and technology.
- if your organisation is taking a centralised approach to governance, consider what mechanisms are in place to support oversight by other stakeholders, like civil society, or NGOs
- make sure that your project is focused on user needs to help mitigate user concerns, check in with your users to make sure they are being included at every stage of the design and delivery cycle, and that they understand what is being developed
- consider if there is a need for a specific role to support governance processes - does this need to be an independent process depending on the data sets involved within your AI product or service

### *Post deployment actions:*

- teams should consider what indicators need to be tracked to help them review their model's performance over time
- ensure that there's a clear explanation of any changes in your model's performance and what you have done to mitigate any impacts

- set out a description of the IT architecture that helps holistic integration between organisations (e.g. utilising AI platforms or strategic partnerships). The life cycles of AI solutions are managed.
- consider on-going validation of ethical values and principles through the development of your project or project, not only during the development of your project but also post-launch to mitigate bias
- understand and set out what team structures need to be in place to support long-term user needs
- evaluate the qualitative and quantitative results after deployment, assess what impact they may have and how they might be mitigated

## 4. Manage and Assure your Data

Biases in the datasets can create mistrust, reduce people's ability to participate in the economy and society, and create inequalities even if sensitive variables such as gender, race, or sexual orientation are removed.

Efforts to mitigate biases range from examining the effects of the bias on a model's performance, assessing the performance of algorithms on different users groups and giving feedback from the discovery to model development teams, or conducting primary research on these datasets. Some countries use internal ethics frameworks for automated decision-making to work out specific methods for assessing and mitigating data biases. If the data collected warrants it and to mitigate the potential for bias, teams may consider using neutral parties like Non-Government Organisations (NGOs), Civil Societies, or Academia to act as neutral bodies to review the research, assess how the data is managed and collected and provide on-going external governance. Research can help gain multi-disciplinary perspectives, including from ethicists, social scientists, and other humanities thinkers.

**Managing Data: Australia's Emergency Situation Awareness (ESA) software:** The Commonwealth Scientific and Industrial Research Organisation has designed a software that detects unusual behaviour in the Twitter stream. It can be used to assess how a disaster event is being broadcast and help disaster managers evidence of pre-incident activity. Once the event happens it can provide, near real-time notice of an incident occurring, first-hand reports and understand community response to an emergency warning.

The following steps will help you continue to manage and assure your data:

- keep testing for any impacts on the broader digital ecosystem as well as any possible impacts on the service/domain you work in
- assess your data collection pipelines to ensure clean data input to support reliable modelling and performance

- set out your testing plan and approach to ensure that there are mechanisms that cover who and how will monitor the model's performance and how often it's assured over time
- plan and clearly set out how your data is going to continue to be validated throughout delivery
- build-in regular reviews of your data over time to ensure consistency, and ongoing validation post-delivery (so plan for the longer term)
- secure the long-term supply of data-streams through agreements or licences

## 5. Laws and regulations

The legal implications of AI are still emerging as the application of AI in both the public and private sectors is still emerging. Countries have different legal approaches to setting these regulations and identifying essential needs, which vary from population-wide consultations to regulatory guidelines developed by specialist agencies.

Some countries favour taking their national constitutions, others supranational (e.g. European Union (EU) laws like the General Data Protection Regulation (Regulation (EU) 2016/679) ('GDPR')), and/or guidelines developed by international expert agencies.

There is also the recently launched [EU AI Act](#) which sets out three risk categories;

- 1) the use of applications and systems that create unacceptable risk,
- 2) high-risk applications, like loan applications which may introduce bias, and
- 3) as concerns unregulated applications or systems.<sup>8</sup>

The following steps that may help you meet your legal requirements:

- understand what laws, rights and responsibilities you need to meet throughout delivery of your AI model and post implementation
- ensure what you are planning to deliver meets relevant national legal accessibility standards and [W3C](#) standards
- if human intervention is needed due to an unexpected decision made by the system, ensure that any help desk or call centre details are clearly signposted, along with an easy step-by-step guide and detailed information for users about the complaint process
- consult your legal teams to ensure that any automated decision-making at any stage of the process does not breach any rights, legal rules or regulations or if there are any wider legislative implications or impacts

---

8



[Creating positive social outcomes -The UK's National Health Service \(NHS\)](#) Heart Disease is one of the UK's biggest killers. A standard angiogram procedure which involves injecting dye into your heart can be costly and time and resource intensive. Teams worked in a multidisciplinary way, with radiographers and other service domain experts to create [HeartFlow](#), to help doctors understand how a patient's heart is functioning. It's a system that analyses scans of those who may have coronary heart disease and then creates a personalised 3D model of the heart. This helps doctors spot any disruption by blockages and provide more personalised treatment.

## 6. Ensure inclusivity and preventing bias

Many countries are still developing guidance for bias testing to ensure that AI projects contribute toward advancing human capabilities and the inclusion of underrepresented populations. While there are various methods for testing biases in datasets, this practice requires that data models have equal predictive value across different groups or that models have equal rates of false positive and false negative results across groups.

Measuring potential environmental implications of AI projects and planning for their mitigation remains challenging, and consideration needs to be given to the impact of performance errors on different stakeholders. Some governments take into account any direct and indirect impacts such as physical, financial and emotional impacts as well as loss of competitiveness or compromises to public infrastructure. Other governments have taken this a step further to consider human, social and environmental impacts.

The following steps can you help mitigate biases:

- develop a technical strategy and a portfolio of tools that can help you identify potential sources of bias and reveal the traits in the data that affect the accuracy of the model
- do user research early and often - not only during delivery, but also post-deployment to ensure each iteration meets user needs
- assure against bias by expanding sample data sets or by mixing teams from different socioeconomic backgrounds
- engage in fact-based conversations around potential human biases. This could take the form of running algorithms alongside human decision-makers, comparing results, and using “explainability techniques” that help pinpoint what led the model to reach a decision to understand why there may be differences.
- stay up-to-date on this fast-moving field of research (for example, the [Partnership on AI](#) and the [Alan Turing Institute's Fairness, Transparency, Privacy group](#))

## Next steps

The next steps for the work is for the AI Digital Exchange Working group to review and provide feedback on the MVP framework.

## Resources

1. [Understanding artificial intelligence ethics and safety](#) (UK)
2. [Den gemensamma värdegrunden för de statsanställda](#) (Stockholm)
3. [Algorithm Charter for Aotearoa New Zealand](#) (NZ)
4. [Understanding artificial intelligence ethics and safety](#) (Alan Turing Institute)
5. [Ethics, Transparency and Accountability Framework for Automated Decision-Making](#) (UK Government)
6. [A Blueprint for Equity and Inclusion in Artificial Intelligence](#) (World Economic Forum)
7. [AuroraAI national artificial intelligence programme](#) (Finland)
8. [Public Sector Playbook](#) (Singapore)
9. [AI Playbook for the US Federal Government](#) (United States of America)
10. [Four Principles of Explainable Artificial Intelligence](#) [DRAFT] (United States of America)
11. [Guidance For Creating Trustworthy AI](#) (Germany)