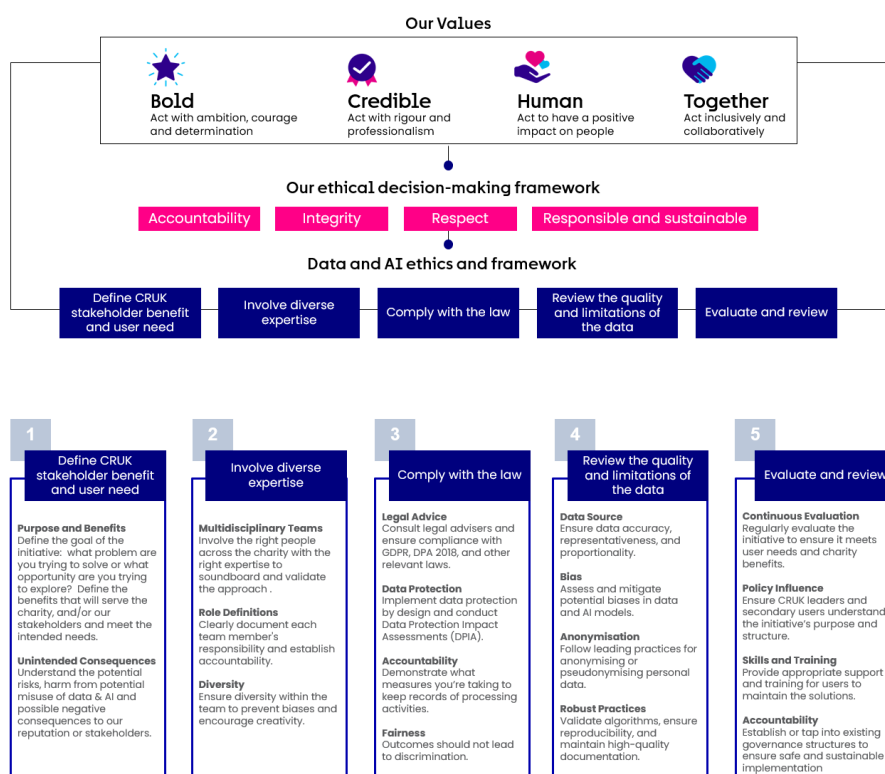# Our data and artificial intelligence ethics: principles and framework

Introduction

Our data and artificial intelligence (AI) ethics principles and framework is underpinned by the three pillars of our Policy Framework – our resources, our people and places, and our reputation – and interlocks specifically with our Code of Ethics. The existing CRUK Requirement, Artificial Intelligence Use, will be updated with the finalised framework.

We've considered our organisational values – bold, human, credible and together – in relation to data and AI ethics. And we've developed a five-step framework to support any future data and AI initiatives.

Our Values

Bold	Credible	Human	Together
Act with ambition, courage and determination	Act with rigour and professionalism	Act to have a positive impact on people	Act inclusively and collaboratively

Our ethical decision-making framework

Accountability | Integrity | Respect | Responsible and sustainable

Data and AI ethics and framework

Define CRUK stakeholder benefit and user need | Involve diverse expertise | Comply with the law | Review the quality and limitations of the data | Evaluate and review

---

1 — Define CRUK stakeholder benefit and user need

Purpose and Benefits
Define the goal of the initiative: what problem are you trying to solve or what opportunity are you trying to explore? Define the benefits that will serve the charity, and/or our stakeholders and meet the intended needs.

Unintended Consequences
Understand the potential risks, harm from potential misuse of data & AI and possible negative consequences to our reputation or stakeholders.

---

2 — Involve diverse expertise

Multidisciplinary Teams
Involve the right people across the charity with the right expertise to soundboard and validate the approach .

Role Definitions
Clearly document each team member's responsibility and establish accountability.

Diversity
Ensure diversity within the team to prevent biases and encourage creativity.

---

3 — Comply with the law

Legal Advice
Consult legal advisers and ensure compliance with GDPR, DPA 2018, and other relevant laws.

Data Protection
Implement data protection by design and conduct Data Protection Impact Assessments (DPIA).

Accountability
Demonstrate what measures you're taking to keep records of processing activities.

Fairness
Outcomes should not lead to discrimination.

---

4 — Review the quality and limitations of the data

Data Source
Ensure data accuracy, representativeness, and proportionality.

Bias
Assess and mitigate potential biases in data and AI models.

Anonymisation
Follow leading practices for anonymising or pseudonymising personal data.

Robust Practices
Validate algorithms, ensure reproducibility, and maintain high-quality documentation.

---

5 — Evaluate and review

Continuous Evaluation
Regularly evaluate the initiative to ensure it meets user needs and charity benefits.

Policy Influence
Ensure CRUK leaders and secondary users understand the initiative's purpose and structure.

Skills and Training
Provide appropriate support and training for users to maintain the solutions.

Accountability
Establish or tap into existing governance structures to ensure safe and sustainable implementation

Contents

Who is the framework for?

Using an ethics framework for data and AI initiatives helps us make sure our activities align with our values and mission, and the broader societal good. It also helps us to act responsibly, transparently and with respect for all stakeholders involved.

This framework is for all employees, project teams and stakeholders involved in internally led data and AI initiatives, including technical experts, non-technical staff, senior leadership and external partners. It supports decision-making and makes sure that everyone has a clear understanding of the guidelines and principles involved in the use of data and AI.

What is the framework for?

This framework provides guidance on ethical data practices, transparency and compliance for designing, implementing and running data and AI initiatives. It helps teams identify potential risks, protect stakeholder rights and foster trust in technology projects. It also helps make sure that data and AI are used in ways that maximise Cancer Research UK stakeholder benefit and minimise potential harm or unintended consequences.

How should the framework be used?

This framework should be applied at all stages during data and AI initiatives, from initial discovery to planning, implementation and evaluation. It provides practical steps and considerations for teams to follow, such as defining CRUK stakeholder benefit, assessing risks, ensuring transparency and monitoring outcomes. The framework is also integrated into decision-making processes to provide actionable insights and support compliance with legal and regulatory standards.

When should the framework be used?

This framework should be used for all data and AI initiatives that involve significant impact or potential risk to Cancer Research UK. This includes projects that handle personal data, involve automated decision-making or affect public or stakeholder perceptions. For smaller scale projects or routine data activities, simpler governance and ethical guidelines may suffice, but the overarching principles of this framework should still be considered.

Read some use cases and scenarios

SUGGESTED INCLUSION: SELF ASSESSMENT CONFIDENCE RATING

Why is it Important to follow an ethics framework?

- **Trust and transparency:** Using an ethics framework fosters trust among stakeholders by ensuring that data and AI practices are transparent, fair and accountable.
- **Risk management:** An ethics framework helps to identify and mitigate against ethical risks associated with data and AI initiatives to make sure the potential negative impacts are minimised.
- **Legal and regulatory compliance:** This framework ensures that data and AI initiatives comply with legal and regulatory requirements to avoid potential legal issues and reputational damage.

1. Define Cancer Research UK stakeholder benefit and user need

When starting a data initiative, clearly define its purpose and benefits. Identify how it will serve Cancer Research UK and meet user needs, while still aligning with our values.

1.1 Define the benefit to Cancer Research UK

This can relate to research, policy, philanthropy, marketing, fundraising or another area of our work.

- Does this project benefit Cancer Research UK, our stakeholders and society at large?
- Have you documented the expected benefits and how they will be measured and communicated transparently?

If no clear benefit can be identified, **stop the project**.

Continually revisit the expected benefits throughout the initiative lifecycle.

1.2 Consider unintended consequences

- What are the potential risks if this project is not implemented?
- What harm might come from misusing the data or model?
- Could individuals face negative consequences because of this initiative?
- What are the environmental implications of the initiative? How could they be mitigated?

If multiple risks cannot be mitigated, **pause the project** until they're addressed.

1.3 Align with our values

- How does the design and implementation of the initiative or algorithm adhere to our organisational values?

1.4 Justify the use of donations to benefit our supporters

- Can you demonstrate the value for money of your initiative?
- Is there effective governance and decision-making oversight to ensure the success of the initiative?
- Are the benefits to Cancer Research UK and the wider public transparent?
- Have the benefits been communicated to all stakeholders?
- Do you have evidence to demonstrate all of the above?

Commented [SG2]: I'm working through this doc (testing it out on a potential use case to see where there are sticking points). Do we have any guidance on how to assess environmental impact? I know there are some environmental concerns on the use of genAI but I've no idea how I would quantify this, nor find strategies for mitigation. If this guidance exists can we link it?

1.5 Understand and articulate the benefits

It's helpful to use the **user story format** by completing the following sentence:

"As a [user type], I need to [user need or task] so that [desired outcome]."

Example: "As a supporter, I need to know how my donations are being used so that I can see the impact of my contributions."

2. Involve diverse expertise

Working in diverse, multidisciplinary teams with wide-ranging skillsets helps to prevent bias and encourages more creativity and diversity of thought in any initiative. Involve others from your team or wider network to make sure you have the right skills and expertise in place.

2.1 Get the right expertise

Make sure that your initiative has the appropriate expertise and diverse perspectives to identify potential risks and deliver successful outcomes. If the necessary expertise isn't available, **pause the project** until the gaps are filled.

- Do you have the expertise you need?
    - **Technical experts**: Data scientists, engineers and technical specialists who understand the technical aspects and implications of the project.
    - **Domain experts**: People with deep knowledge of specific areas of the organisation, such as marketing, fundraising, healthcare or research.
    - **Policy and legal advisers:** People who can help ensure compliance with data protection laws and ethical standards.
    - **User representatives and advocates**: Patient advocates or representatives from the community to help make sure decisions reflect real-world impact.
- Have you used a **Responsibility Assignment Matrix (RACI)** to determine who is responsible, accountable, consulted and informed at each stage of the project?

2.2 Ensure diversity within your team

- Do you have a truly diverse team (including special categories and different roles within the organisation)?

2.3 Involve external stakeholders

- Do you need to consult external experts or partners?
  For example, external expertise may be needed for:
    - advanced AI model validation
    - legal consultations on complex data-usage scenarios
    - input from patient or stakeholder advocacy groups

2.4 Consider your governance structure

- What senior or external oversight is there for your initiative?
  For example, the Data Governance Forum, patient advocate groups and user groups.

2.5 Ensure transparency

- Where it's appropriate, have you published information on expert consultations and other relevant information.
- AI specific: Have you maintained comprehensive documentation of the AI models? This should include the purpose, design, data sources and decision-making processes. And have you regularly published reports on the performance and impact of the AI models, highlighting key metrics, successes and areas for improvement?

3. Comply with the law

Embarking on a data and AI initiative that's ethical is not the same as it being legal. Legality is your duty, obligation and the minimum standard that an initiative should meet. The current data and AI areas that legislation covers are:

- **Personal data:** Defined in Section 3(2) DPA 2018 (a wider explanation is detailed in Article 4 of the GDPR).
- **Data protection by design and by default:** See Article 25 of the GDPR. Also use Article 35 of the GDPR to complete a Data Protection Impact Assessment.
- **Accountability principle:** This principle demonstrates what measures you're taking to ensure everything is documented, as seen in Article 5(2) of the GDPR and Article 30 of the GDPR on keeping records of processing activities.
- **Fairness principle:** Data analysis or automated decision-making must not result in outcomes that lead to discrimination as defined in the Equality Act 2010.

- **Additional regulation:** Consider additional [relevant legislation and codes of practice](#).

You must consult the [Data Protection section](#) of our Legal Services Hub and our Data Privacy team for advice and compliance as they can assist with a Data Protection Impact Assessment and other queries regarding the above. If you haven't done this, **pause the project** until you have.

4. Review the quality and limitations of the data

Insights from new technology are only as good as the data and practices used to create them.

- Is the data for the initiative accurate, representative, proportionally used and of good quality?
- Are you able to explain its limitations?

If the data isn't good enough, **pause the project**.

4.1 Data source

- What data source(s) is being used? Consider:
  - Do individuals or organisations providing the data know how they will be used?
  - If you're repurposing data for analysis without individual consent, is the new purpose compatible with the original reason for collection?
  - Can all metadata and field names be clearly understood?
  - Do you understand how data for the initiative are generated?
- What processes do you have in place to ensure and maintain data integrity?
  - Are there any caveats? And how will these be considered for any future policy or service which uses this work as an evidence base?

If the data sources and processes are not good enough, **pause the project**.

4.2 Determining proportionality

Use the minimum data necessary and consider pseudonymised and/or synthetic data.

- How can you meet the initiative's aim using the minimum data possible?
  - If you're using personal data, what measures are in place to control access?

- o If you're repurposing data, would the people who provided that data approve of its secondary use?
- o Can you justify the use of this data to the general public?
- o AI specific: How does the sensitivity and necessity of the input data align with the initiative's aim?

4.3 Bias (fairness)

- Has the data used to train a model been assessed for potential bias?
    - o Have you considered historical bias, selection bias and proxy variables?
    - o Have you documented the measures you've taken to mitigate bias?
    - o AI specific: Which tools and software are being used to detect and measure bias? And how do you plan to conduct regular fairness audits and assessments of AI systems?
    - o AI specific: Have you used fairness-aware algorithms to guide the design of the AI system as well as traditional accuracy metrics?
    - o AI specific: What scenario analysis will you conduct to understand how different decisions within the AI system could potentially amplify existing biases or introduce new ones?

4.4 Anonymisation

If you anonymise or pseudonymise personal data before linking or analysis, follow the ICO's Anonymisation: managing data protection risk code of practice and R&I's Anonymisation Assessment to document your methods. You can find more technical advice on the UK Anonymisation Network's website.

- If data is anonymised:
    - o Can you demonstrate that the data has been de-identified to the greatest degree possible?
    - o Can the data be matched to any other datasets that will make individuals easily identifiable? If so, what measures have you taken to mitigate against this?

4.5 Robust practices

- Does the data meet the FAIR principles: findable, accessible, interoperable, reusable?
- If necessary, how can you (or external scrutiny) validate that the algorithm is achieving the correct output decision?

- Have you ensured there is adequate human oversight during all the lifecycle stages of the initiative?
- Have you followed the three requirements for reproducibility? I.e. can you apply the same logic (code, model or process) to the same data in the same environment so that they can be easily reproduced, tested and audited? Can you supply evidence of this?
- How will your documentation be kept?
- AI specific: What continuous monitoring and feedback systems will you use to track the performance of AI models and allow users to report inconsistencies or errors in the AI's outputs over time?
- AI specific: What rate of error are you proposing to accept and then quality check?

4.6 Data sharing

If your data is non-sensitive and non-personal, and if data sharing agreements with the supplier allow it, make the data open and assign it a digital object identifier. This gives others access to the data and the code, so the analysis can be reproduced. You can also publish data on Find open data and the UK Data Archive.

4.7 Model sharing

When sharing models, make sure it doesn't endanger either the privacy of those whose data was used to train them or the integrity of the task being undertaken.

- Can you openly publish your methodology and metadata about your model, or the model itself? For example, on Confluence or SharePoint?

4.8 Transparency

- Is it appropriate to present your model to the Cancer Research stakeholders? If so, how are you planning to do it?
- Have the necessary classification labels been adopted?

4.9 Explainability

- Can the workings in a machine learning algorithm be explained in a way that everyone can understand?
- Can you explain what your initiative does and how it was designed to a non-expert audience?

- Can you describe the process and the aim of your algorithm, including the variables used and the outcomes the achieve, without using technical terms?

5. Evaluate and review

Throughout the initiative, continuously evaluate how findings and data models are being used and monitored. Also consider how your initiative influences the objectives of the Marketing, Fundraising and Engagement, Philanthropy, Chief Operating Office, Policy, Information and Communication and Research & Innovation directorates.

5.1 Evaluate the initiative

- At the beginning of the initiative, ask yourself:
  o Are we doing the right thing?
  o Is this benefiting Cancer Research UK?
- During the initiative, ask yourself:
  o Have we designed it well?
  o Has the user need changed? If so, how?
  o Has there been a change of circumstances that has affected the initial understanding of the benefit to Cancer Research UK? If so, how can we adjust it?
  o Are we regularly gathering user experience feedback on their perception of, and interaction with, the initiative?
- After the initiative, ask yourself:
  o Is it still doing the thing we designed it for?
  o Has the user need changed? If so, how?
  o Does the initiative continue to benefit users?
  o Is the current human oversight sufficient?
  o Have you mitigated against any unanticipated harms that have emerged during the initiative?
  o What has been done to make sure there are clear protocols for escalating issues, biases or concerns from the initiative to staff?
  o Have you scheduled regular development / reviews of the initiative to make sure it's accurate and relevant?

5.2 Make sure the initiative is sustainable

- Do the users have the support and tested training they need to maintain the initiative?

5.3 Ensure accountability

Governance structures are essential to guarantee the safe and sustainable implementation of an initiative.

AI specific: Encourage employees to question AI outcomes and provide feedback as this can improve the accuracy and reliability of the system.

- How often will you update the governing authorities?
- AI specific: Are the areas, roles and responsibilities for employees clearly defined? This includes the need for human judgment, empathy and ethical considerations that cannot be fully automated.
- AI specific: Make sure AI isn't used as a sole author for any work.

5.4 Allow for Cancer Research stakeholders scrutiny

- What channels have you established for Cancer Research stakeholder engagement and scrutiny throughout the initiative?
- Have you documented and shared the progress and case studies from your initiative with stakeholders?

Glossary of terms

Accountability:
The obligation to take responsibility for the outcomes of data and AI initiatives. This includes ensuring that processes are fair, ethical and legally compliant.

Algorithm:
A set of rules or instructions given to a computer to help it solve a problem or perform a task. In AI, algorithms are often used to process data and make predictions.

Anonymisation:
The process of removing personally identifiable information from datasets, so people can't be identified from the data.

Artificial intelligence (AI):
A branch of computer science focused on creating systems that can perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making and language translation.

Bias:
Systematic error in a dataset or model that can lead to unfair outcomes, particularly when these models reinforce existing inequalities or stereotypes.

Data governance:
The management of the availability, usability, integrity and security of data. It includes policies, processes, roles and standards that ensure data are handled responsibly throughout its lifecycle.

Data privacy:
The protection of personal information shared by individuals. It involves ensuring that data is handled responsibly, securely and in compliance with legal regulations.

Data Protection Impact Assessment (DPIA):
A process designed to help organisations identify and minimise the privacy risks of data processing activities.

Data source:
The origin of the data being used in a project. This could be internal databases, third-party providers or publicly available data.

Ethical review:
A process for evaluating whether a project meets an organisation's ethical values, including considering its fairness, transparency and public benefit.

Ethical risk:
The potential for a project or initiative to cause harm, discrimination or other negative outcomes that are not aligned with an organisation's ethical values.

Fairness:
In AI and data science, fairness refers to ensuring that systems don't favour one group of people over another, avoiding discrimination based on race, gender or other personal characteristics.

GDPR (General Data Protection Regulation):
A legal framework that sets guidelines for the collection and processing of personal information from people who live in the European Union.

Historical bias:
A bias within data because of how it was selected, which means it is not truly representative

Initiative:
A project or programme undertaken by an organisation to achieve a specific goal. In the context of this framework, an initiative refers to any data or AI project aimed at improving operations, providing services or achieving a CRUK stakeholder benefit that aligns with our mission

Machine learning:
A subset of AI that involves training computers to learn from data and make decisions without being explicitly programmed to perform those tasks.

Predictive analytics:
The use of data, statistical algorithms and machine learning techniques to predict future outcomes based on historical data.

Public benefit:
The positive impact or value that a project or initiative provides to society, stakeholders or a particular community.

Selection bias:
A bias within data because of how it was selected, which means it is not truly representative

Stakeholders:

Individuals or groups who have an interest in the outcome of a project. This could include employees, supporters, donors and the people or causes they serve.

Synthetic data:

Data that are artificially generated rather than obtained by direct measurement or collection. They can be used to test systems without exposing real data.

Transparency:

The practice of being open and clear about how decisions are made, especially in data processing and AI, so that stakeholders and the public can understand and trust the processes.

User need:

The requirements or problems that users, such as supporters, cancer patients or stakeholders, expect to be addressed by a data or AI initiative.

Proxy variable

Proxy variable is an indirect measurement that is used when the variable is unobservable or cannot be measured directly.