Data for Children Collaborative WITH UNICEF

Ethical Assessment

Developed in partnership with
Ethical Intelligence Associates Ltd

THE UNIVERSITY of EDINBURGH
Data-Driven Innovation
Part of the Edinburgh & South East Scotland City Region Deal

THE UNIVERSITY of EDINBURGH

unicef for every child

Scottish Government
Riaghaltas na h-Alba
gov.scot

Data for Children Collaborative WITH UNICEF

Contents

Introduction

Welcome to the Data for Children Collaborative's ethics and safeguarding training pack. This includes everything you need to successfully complete the ethical assessment and safeguarding training for your project. We hope that you find this tool useful in helping you to reflect on any ethical challenges you may face as your project progresses, and to remind you of the importance of protecting children and their rights at all times.

The Data for Children Collaborative with UNICEF is a joint partnership between UNICEF, The Scottish Government and the University of Edinburgh's Data Driven Innovation Programme, which seeks to improve outcomes for children locally, nationally and globally. We draw on the strengths of all partners to bring insight and solve problems using data and data science techniques.

Using data for children presents a unique set of ethical challenges. Not only should we ensure that we are collecting, using and storing data in an appropriate way, but we also need to also be mindful of children's vulnerabilities. Successfully delivering an ethical Data for Children project requires more than just legal compliance. We want to ensure that every project is being delivered in the right way *and* for the right reasons.

Within this guide, you will find our ethics assessment and safeguarding training. After completing both the ethics and safeguarding assessments we will sign off your 'Driving Licence' and allow you to begin with your project, confident that you have acknowledged and accepted the ethical and practical risks of working with children and data.

Ethics is not static, nor is it just a tick box exercise. We hope that this process can help to highlight and address potential issues at various points throughout the project lifecycle. We encourage you to use this pack whenever you may need it.

Compass

Our compass represents our guiding principles that help us to ethically navigate our projects. These are steadfast and unchanging, and remain at the core of everything we do. Keep these principles in mind when completing your ethical assessment. You should be able to highlight how each stage of your project has adhered to each compass point.

Remember that, above all, your actions should be for every child.

For Every Child

We are open in our intended outcomes for children and clearly communicate the goals of all our projects.

With Transparency

We show integrity in our projects by holding our methods to the highest standards.

We strive to positively impact children by regarding them as our priority stakeholder and treating them, and their data, with respect.

With Trust

We use data from reliable sources and protect all personal data in our projects.

With Safe Data

Developed in partnership with Ethical Intelligence Associates Ltd.

Data for Children Collaborative WITH UNICEF

Road Map

Please refer to the Compass and the Highway Code as the two supporting documents for this Road Map. The Compass provides overall guidance of our core principles. The Highway Code provides a detailed set of prompts and example answers to be considered while answering the roadmap questions. Click on any of the questions to lead you to the corresponding Highway Code information.

The Road Map is split into three sections:

- START aims to support you in defining any ethical issues before your project begins
- ON THE ROAD helps you to revisit existing and highlight any new ethical issues throughout the duration of the project
- DESTINATION supports the final stages of the project, thinking about the consequences of communicating your findings and how this could influence your stakeholders

As you work through these questions, please remember that there **is no such thing as too much detail.** Be as **specific** and **accurate** as you can and include as much relevant information as possible. You can refer to the 7 Responsible Data for Children Principles in your answers:

Purpose-Driven

People-Centric

Participatory

Protective of Children's Rights

Proportional

Professionally Accountable

Prevention of Harms across the Data Life Cycle

START - To be completed at the start of any project.

PROJECT:

1.1 State your project goal.

1.2 Define the scope of the project.

1.3 How does your project intend to benefit children?

1.4 To what purposes will findings be used to inform processes, and if, (and only if) already negotiated, for which stakeholders?
Please note: Do not identify changing policy or programmes as a direct outcome if this hasn't been agreed with Government or Programme Organisations.

1.5 List all likely stakeholders for this project, and your level of engagement with them.

Stakeholder	Level of Engagement

1.6 What are the potential effects (both positive and negative) of the outcomes listed in question 1.4 on each of the stakeholders as stated in question 1.5? Be as specific as possible with your answers.

Stakeholder	Potential Positive Effects	Potential Negative Effects

1.7 Are any conflicts likely to arise when ensuring positive outcomes/impacts for one class of stakeholder versus another? If yes, what are they?

1.7.1 How have you reconciled these conflicts? Detail specific strategies.

1.7.2 Are there any stakeholders noted in 1.5 for which an additional duty of care may be required, for example children, or who may be considered a 'vulnerable population?' If yes, detail who, why, and how this care will be provided.

DATA:

1.8 Is it necessary for your project to be able to identify individuals? Please justify your answer.

1.9 What are the sources of your data?

1.10 For each source, if collecting primary data, explain the negative impacts this may have on stakeholders. Include data controller / processor roles.

1.10.1 For each source, if the data controller/processor is not the Collaborative, ensure that the original method of collection is fit for purpose and aligns to the Collaborative's values.

Source	Controller / Processor	Method of Collection

1.10.2 For each source, does the intended use of the data within this project align to the purpose the subjects consented to when it was collected?

Source	Subjects	Consent Required

1.10.3 For each source, are each of the data variables you are requesting justified by the scope of the project?

Source	Variables	Justification

1.11 For each source, how will any data be stored throughout the project lifecycle and upon project completion?

Source	Retention

1.12 For each source, has the data been anonymised to the greatest degree possible? Detail how each data set has been anonymised.

Source	Anonymisation Method

1.12.1 If using multiple data sets, could the merging of these datasets re-identify individuals? Have any measures been taken to prevent this? Detail below.

1.13 Is there any reason why your project methodology should not be made public or shared for usage throughout the Collaborative?

RESULTS:

1.14 Who are you required to share your findings with?

1.15 What are other applications, both positive and negative, of any methodologies developed within this project?

1.16 What are the implications of your project using data science to make predictions or assumptions about children? How can these be resolved?

1.17 How may communicating your findings bring about unintended harm to stakeholders? Please identify stakeholder groups, respective potential harms and how you plan to mitigate any risks.

Stakeholder Group	Potential Harm	Risk Mitigation

1.18 Are there possible adjustments that can and should be made to communications that may mitigate against; misunderstandings, misrepresentation and misapprehension of data and limitations?

1.19 Will the project still be beneficial if you cannot communicate findings to those outside of the required group listed in 1.14? Explain your answer.

ON THE ROAD – to be completed at periodic stages throughout the project lifecycle.

This stage of the ethical assessment is designed to help you reflect on your progress so far. It provides you with an opportunity to revisit your original answers and track any changes that have occurred since your project began. Remember that children are always your key stakeholder, and the project should always be protective of children's rights.

2.1 Has the proposed benefit to children, as detailed in 1.3, changed in any way? If yes, explain how and why these changes have been made.

2.2 Have any new stakeholders been added to the project? If yes, explain who and why they have been included.

2.3 Has any of your data science methodology changed since the last assessment? If yes, detail how you came to this decision, and how it aligns to the original project goal and purpose as detailed in 1.1 and 1.2.

2.3.1 Have any new data sources been used? If yes, please complete questions 1.8 – 1.12.1 for your new data source(s).

2.3.2 Have any new parties been given access to the data? If yes, detail who and why this decision was made.

2.3.3 Are there any parties that still have access to the data that no longer need it? If yes, detail the steps you will take to ensure that they no longer have access.

2.4 Have you encountered any unexpected results since beginning your project? Explain your answer.

2.5 Are you aware of any new ethical issues that may arise as a result of these changes? If so, how do you plan to mitigate against these?

Reflecting back to the 7 RD4C principles, can you confidently answer that your project is still…

Purpose-Driven

People-Centric

Participatory

Protective of Children's Rights

Proportional

Professionally Accountable

Prevention of Harms across the Data Life Cycle

DESTINATION – to be completed before dissemination of findings.

At this point in the project lifecycle, it is crucial to think about the impact of your work. These questions are tailored to encourage you to reflect on your project outputs and their possible effects when shared outside of the project team.

3.1 Has your project output achieved or diverted from your intended project goal? Please detail your answer.

3.2 Given your project output, can you now identify any additional stakeholders? If yes, explain who and why they have been included.

3.3 Have you stored or disposed of the data, and any copies of the data, as originally stated? If no, explain why.

3.4 Given the final outcome of your project, who will your results be communicated to?

3.5 Will the communication of your outcomes negatively impact any of the project stakeholders? If yes, how will you resolve this issue?

3.6 What impacts, both positive and negative, will communicating your findings have on children?

Reflecting back to the 7 RD4C principles, can you confidently answer that communicating your findings is still …

Purpose-Driven

People-Centric

Participatory

Protective of Children's Rights

Proportional

Professionally Accountable

Prevention of Harms across the Data Life Cycle

Highway Code

When leveraging data to generate insights that can improve children's lives, the Road Map will guide you through important questions at each of the critical project stages. These are not meant to be easy questions, and at times it may be beneficial to consult some additional resources before coming to an answer. This is the purpose of the Highway Code: a useful collection of additional questions, key definitions and external resources to make sure you feel confident in your decision-making, and to encourage positive ethical engagement throughout the project lifecycle. Throughout every step in the Roadmap you can refer to the Highway Code for more information. Please remember that there **is no such thing as too much detail.** Be as **specific** and **accurate** as you can and include as much relevant information as possible.

Data *for* Children

At the Data for Children Collaborative with UNICEF, we are dedicated to improving outcomes for children locally, nationally and globally. For us, every project should be for children. That's why it is important that everyone involved in our projects is committed to promoting and protecting children's rights.

Data can be leveraged in so many ways to better the lives of children across the globe. With fast-paced technological advances and innovative data science techniques, we are well positioned to make positive impacts. However, it is important to acknowledge that combining the worlds of data and children creates its own unique opportunities and risks.

Many data systems have been designed with consenting adults in mind, and therefore do not focus on the unique needs and vulnerabilities of children. There is a higher duty of care required for the use and handling of children's data – they may lack understanding of how their data is generated and recorded, the associated risks and benefits, and how vulnerable their data can leave them. It is our responsibility to ensure that we always put children's rights, needs and interests first.

Often, children's rights can get lost in the data rush. Here is a useful resource to introduce or refresh you on the Convention on the Rights of the Child. Having a good understanding of the key principles will help keep you focused when answering these questions; ensuring that your project aligns to the Collaborative vision and keeps children at the core of all that you're doing.

The Ethical Research Involving Children (ERIC) Compendium is another fantastic user guide that walks you through common ethical issues that can arise when carrying out research with children. There are useful case studies that can help you engage with ethical decision-making in real project contexts.

Responsible Data for Children Principles

The questions in our Highway Code are informed by the Responsible Data for Children (RD4C) principles, resulting from rigorous research and broad consultations conducted by UNICEF and The GovLab at New York University to highlight and support best practice in data responsibility. Taken together, these principles outline a commitment to steward the data collected, stored, prepared, shared, analysed and used to save children's lives, defend their rights and help them fulfil their potential from early childhood through adolescence. We briefly list the principles below so that you can use them as prompts for reflection and focus throughout your answers.

Purpose-Driven

A responsible data practice starts by being purpose-driven. When seeking to handle data actors should identify and specify why the data is needed and how the intended or potential benefits relate to improving children's lives. If there is no clearly articulated benefit for children, actors should not collect data, store, share or analyse it.

People-Centric

Much of the data used for drawing insights to improve children's lives involves or is generated by people. The insights from it have the potential to impact the lives of children in many ways, both positive and negative. Actors must thus ensure the needs, interests and expectations of people—including children and their caregivers in particular—are prioritized by those handling data about them. Actors should take a people-centric approach to the consideration of opportunities and risks of data initiatives—prioritizing the consideration of data practices' effects on people over potential efficiency gains or other process-oriented objectives. This entails some combination of the following criteria: children and/or their caregivers have consented to the data use, children and/or their caregivers have a clear understanding of how this work will be conducted, the work is demonstrably serving children's interests, and/or the work is required by law or institutional mandate. In addition, actors need to be context sensitive, paying attention to and acting according to the legal, cultural and community contexts in which any given project exists.

Participatory

Responsible data is participatory. It seeks and builds with inputs from those who use and are affected by data, namely children, their caregivers, and the communities in which they live. Accordingly, actors should inform and engage with individuals and groups. In seeking input, actors should pay attention to marginalized and vulnerable population segments as well as to the inputs of partners, donors and other key stakeholders.

Protective of Children's Rights

When it comes to children, responsible data practices begin by recognizing their distinct needs and requirements. Children's rights must be realized in order for them to develop to their full potential. Realizing these rights can be complex given children's inherent vulnerabilities, the likelihood that others are making impactful decisions on their behalf, and the future prospects they can achieve if supported effectively by those working in their interest.

Proportional

In the data space, less can sometimes be more. When developing and implementing data initiatives, actors should always consider necessity and whether there is proportionality in the breadth of data collection and duration of data retention in order to achieve the intended purpose. The collection and retention of data should be relevant, limited and adequate to what is necessary for achieving intended purposes. The importance of targeting and minimizing collection is true of all data, but especially true of data related to children, given potential and actual vulnerabilities.

Professionally Accountable

Data responsibility rests upon broader foundations of professional accountability. To ensure that the practices and principles described above are put in action, and the unique considerations of responsible data for children are operationalized within institutional processes, organizations and partners should collect, process, and use data within a more general culture of data responsibility. Such a culture has many elements, but one of the most important is to establish and clearly define the role of organization-wide data stewards. Data stewards are an emerging role; they are individuals or groups whose duties cut across departments and functions, and whose broad remit is to oversee responsibility and accountability in the way data is handled.

Prevention of Harms across the Data Life Cycle

Data is not static but exists on a cycle. As part of a commitment to data responsibility, actors should assess and seek to prevent risks across the full data life cycle, including the collecting, storing and preparing, sharing, analysing and using stages. This concept is called end-to-end data responsibility. It is essential for preventing harm to children and ensuring trust.

Data Lifecycle

These 7 principles, and all your ethical decision-making, span the entire data lifecycle. You should consider potential risks and benefits for children at each of these key points. Note that for your project, these stages may not be sequential or discrete, so continual ethical reflection will help you ensure that your project remains child focused.

1. Planning: how a data system will be developed, what type(s) of data will be collected, and for what intended uses
2. Collecting: the process by which data is generated or extracted
3. Storing and Preparing: holding and cleaning processes to enable data sharing, analysis, and use
4. Sharing: the transfer of data between different systems and/or stakeholders
5. Analysing: the interpretation of data, whether through algorithmic or human analysis, to inform some type of decision
6. Using: the ultimate action taken (if any), as well as any eventual archiving or destruction of the data

START:

PROJECT

1.1 "State your project goal"

Clearly stating your project goal will provide you with a reference point throughout your decision making, keeping you focused on what it is your project looks to achieve. Note that all projects should be purpose-driven.

Some key points to consider:

- How does your project goal provide positive outcomes for children?
- What are you hoping to achieve by carrying out this project?
- Why does this project need to be carried out in the first place?
- What problem will it address, and how will it solve it?

1.2 "Define the scope of the project"

The aim of this question is to provide a 'bigger picture' analysis of your project by looking at its applications and implications before the project has started. This is an important exercise, allowing you to think about what can and cannot be done within your project constraints. It may be useful to highlight issues that will lie 'out-of-scope' too. Mapping this out will help to assess whether your project is proportional.

Some questions to think about:

- How do children fit into the scope of your project, both directly and indirectly?
- What background knowledge and skills are needed (if any)?
- What types of data are you likely to use/need? (e.g. personal, group, administrative)
- What theories/methods will be used? What are their strengths/weaknesses?
- What are the limitations of your project?
- Are there any questions that you are not able, or willing, to answer?

1.3 "How does your project intend to benefit children?"

Projects within the Data for Children Collaborative with UNICEF are committed to using data to improve outcomes for children locally, nationally and globally. Remember that one of the RD4C principles is to be people-centric. Children and their needs and interests should be at the centre of any of our projects.

This question looks to assess the practical 'what' question – *what* your project will provide that will improve outcomes for children. What is the end output that can be used to achieve your project goal? There may not be just one answer, so list any outcomes that might be appropriate.

For example: this project will provide positive outcomes for **children** by...

- Improving health care
- Communicating knowledge to the wider public
- Informing professionals

Reference, where you can, how this project will help to realise or safeguard children's rights.

1.4 "To what purposes will findings be used to inform processes, and if, (and only if) already negotiated, for which stakeholders?"

Please note: Do not identify changing policy or programmes as a direct outcome if this hasn't been agreed with Government or Programme Organisations. In these instances, note that the findings will be used to advocate for change. Only in the case where it is agreed or known with certainty that findings will be used should this be noted in this section as 'informing policy or programmes'.
Keep in mind the scope of the project and be realistic with what you think can be achieved given your timeframe and resources. Remember to highlight, at every stage, how this links back your project being purpose-driven to improve **children's** lives.

1.5 "List all likely stakeholders for this project."

A stakeholder is a person or organization with an interest in a project activity and who is affected by the project outcomes. At the Data for Children Collaborative with UNICEF, **children** should always be the priority stakeholder. We also encourage our projects to be participatory, paying attention to marginalized and vulnerable population segments as well as to the inputs of partners, donors and other key stakeholders.

Please clearly number your stakeholder groups in your response so that you can reference these in later responses.

Be as specific as possible with your answers. For example:

Children > in Scotland > in low-income households > between the ages of 5-10

Stakeholder groups can include:

- **Children** as data subjects
  - This group refers specifically to children whose data was used in the project
- **Children** as stakeholders potentially impacted by the project
  - This group includes all children that could be impacted by the results of the project
- Parents, Guardians and Families
  - Parents, legal guardians and caregivers may fall under their own stakeholder group. However, the wider family should also be considered. Many families today are mixed to include stepparents, grandparents and other members

- Professionals
    - This includes those professionally involved with the research, for example: doctors, teachers, researchers
- Communities
    - Communities are not only formed geographically – they can also include communities within schools, churches and other groups
- Politicians
    - This includes anything from small city councils to larger governmental bodies, including policy makers and individual politicians
- Financial Stakeholders
    - Anyone with a financial stake in the project - including funders and investors
- General Public
    - Everyone not mentioned previously, who might still be affected by publication and implementation of the project findings

You can find more information here.

Please also indicate your level of engagement with your stakeholders. This can be listed as high/medium/low depending on how often or actively you will engage with each group.

Stakeholder	Level of Engagement

1.6 "What are the potential effects (both positive and negative) of the outputs listed in question 1.4 on each of the stakeholders as stated in question 1.5? Be as specific as possible with your answers."

In question 1.4, you have listed how your findings will be used once the project is completed. For each potential output, it is important to consider how these could bring about positive and negative outcomes. This can often be carried out as part of a 'harm' and 'benefit' analysis – although in some cases it may not be this straightforward. This is an opportunity for you to assess and seek to prevent risks across the full data life cycle.

Remember that children are always the key stakeholder, and all outcomes should be protective of children's rights.

When detailing the outcomes, it is imperative to include as much detail as possible. Specifics should include:

- Exact nature of the risk/benefit
    - E.g. not just identifying 'stigma' as a general risk but noting how it may arise (in dissemination of findings, in distribution of aid...)
- To whom each risk/benefit refers
    - E.g. including, where known and relevant, age, location, socio-economic status of the stakeholder group

Stakeholder	Potential Positive Effects	Potential Negative Effects
Children in low-income households in Scotland between the ages of 5-10	Agreed government strategy to lower cost of healthy food – meaning children in low income households have better access to nutritious foods Reduction of risk of child obesity and potential adolescent/adult obesity as a result	Dissemination of findings may inaccurately target individuals within predicted 'at-risk of obesity' groups Labelling groups 'at-risk of obesity' could lead to negative mental health and social issues e.g. bullying/discrimination/low-self esteem

1.7 "Are any conflicts likely to arise when ensuring positive outcomes/impacts for one class of stakeholder versus another? If yes, what are they?"

In question 1.6, potential impacts were highlighted alongside which stakeholders these are likely to affect. It is important to assess how these impacts may work in tandem or in conflict – and how to mitigate these issues.

It may be useful to revisit the Convention on the Rights of the Child here, to ensure you have a clear understanding of how to protect children's rights. Think about how you will prioritise the interests of children at each stage.

Some things to consider:

- If one group must benefit over another, children should always be the priority stakeholder
- Are there possible research outcomes that benefit one group but also present risks to that group or others?
- What measures can be put in place to mitigate this conflict of interest?
- How will you justify your reasoning?

1.7.1 "How have you reconciled these conflicts? Detail specific strategies."

Please detail, as specifically as possible, how you plan to reconcile any conflicts that you have identified in question 1.7. Remember that, above all, we have a responsibility to protect children's rights. Try and make sure that your strategies are people-centric rather than data focused, and highlight exactly how they will prevent harm.

1.7.2 "Are there any stakeholders noted in 1.5 for which an additional duty of care may be required, for example children, or who may be considered a 'vulnerable population.' If yes, detail who, why, and how this care will be provided."

A vulnerable group is a cohort that may be at greater risk of harm due to personal or circumstantial factors and/or may have a lesser capacity to safeguard their own interests. They are groups for which a greater duty of care is required. Children are generally considered as requiring a greater duty of care. Taking a people-centric approach will ensure that you have accounted for the specific needs of all the groups involved in your project.

Aside from children, other vulnerable groups can include:

- Immigrants/migrants/asylum seekers
- Persons in institutional settings
- Persons living in poverty
- Religious, linguistic, and ethnic minorities
- Persons with disabilities
- The elderly
- Members of the LGBTQ+ community
- Single parents

Project Compass-Guided Questions:

- Are children the top priority? If no, why not?
    - Remember, children and their wellbeing should be the primary focus of the project.
- Who else might your project be prioritizing?
- How can you better structure your project in order to ensure that children are the key focus throughout the project lifecycle?
- How have you worked to ensure that your project is protective of children's rights?

DATA

1.8 "Is it necessary for your project to be able to identify individuals? Please justify your answer."

Unless necessary for the project, subject data should be kept anonymous. By keeping data anonymous, you will be protecting the rights of the children involved.

When thinking about whether your data should be anonymous, it may be helpful to consider:

- Would the ability to identify individuals bring about positive outcomes for children?
- Can you successfully justify a need to identify individuals?
- Will anonymous data be sufficient in allowing you to achieve your project goals?
- What steps will you take to ensure that individuals cannot be identified in your dataset?
- What would the implications be if children could be identified in your data set?
- Would merging databases potentially re-identify individuals?

If you are still convinced that identification will be necessary to complete your project, you will need to make a strong case to justify this. This will include being able to evidence that you have a lawful basis for collecting this data, a comprehensive data storage plan, and strict compliance with data protection principles. You will need to detail who will be professionally accountable for this data, and that it is proportionate given your project goal.

It is important to remember that children may be less aware of the risks involved with having their personal data collected and processed, so you should be prepared to explain your reasoning in a way that is accessible to children. This could involve explaining to children, in language they can understand:

- Their rights over their own personal data
- Why we require their personal data and what we intend to do with it
- Any risks that may occur in the processing of their data and how we are safeguarding against these

1.9 "What are the sources of your data?"

When thinking about using data ethically, it is key to detail exactly where each part of your dataset is coming from. Transparency about data sourcing will ensure that the data being used is ethical, trustworthy, reliable and safe. It will also help you to identify any risks across the data life cycle.

List your sources in as much detail as possible.

1.10 "For each source, if collecting primary data, explain the negative impacts this may have on stakeholders. Include data controller/processor roles."

If you are collecting primary data, it is imperative that the rights of the child are a top priority. Data collection, particularly regarding **children**, is subject to rigorous ethical assessment. If you have identified a need to carry out primary data collection, please consult with the delivery director for further guidance on next steps to take.

To determine whether you are the **data controller**, you need to assess which organisation decides:

- To collect the personal data
- The legal basis for collecting this data
- Which items of personal data should be collected
- The purpose(s) of the data
- Which individuals the data is collected about
- Whether this data is disclosed (and if so, who to)
- Who is allowed access to the data (and why this is the case)?
- How long the data is retained for
- Whether non-routine amendments to the data can be made

If you are the **data processor**, you may decide:

- Which IT systems and methods are used to collect personal data
- How this personal data is stored
- Security details for the personal data
- How personal data can be transferred from one organisation to another
- How to retrieve personal data about specific individuals, if needed
- Methods to ensure that a data retention schedule is followed
- How to delete/dispose of the data

It is important to ensure that you are clear on who is to be held professionally accountable for these decisions.

1.10.1 "For each source, if the data controller/processor is not the Collaborative, ensure that the original method of collection is fit for purpose and aligns to the Collaborative's values."

Source	Controller / Processor	Method of Collection

If you are not collecting the data yourself, it is still important to do your due diligence. Even if a data set is available this does not guarantee that the data was collected in line with our values as outlined in our compass and seven principles. Take some time to research the methods of data collection and the original purpose for collection.

1.10.2 "For each source, does the intended use of the data within this project align to the purpose the subjects consented to when it was collected? If no - Stop. Review GDPR and fix."

Source	Subjects	Consent Required

Required by GDPR Article 7. Data subjects must be informed of the purpose for which their data is being collected and their consent should be; freely given, given via a clear, affirmative act and easy to withdraw. Data usage should remain aligned to the purpose given at the point of collection.

1.10.3 "For each source, are each of the data variables you are requesting justified by the scope of the project? If no - minimize the number of variables obtained."

Source	Variables	Justification

Keeping the data variables requested to a minimum reduces the risk of subjects being re-identified and ensures that your data is proportional to the project goal. You should be able, if asked, to justify the need for each of the data variables requested. NB: 'Potential control variable' is an allowed justification.

1.11 "For each source, how will any data be stored throughout the project lifecycle and upon project completion?"

Source	Retention

It is important to consider the entire life cycle of the data, from project initiation to project completion.

Think through, and answer:

- Where live data will be held
- How files will be organised/structured
- What backup/recovery provisions are in place
- Any security risks
- How the data will be accessed
- How the data will be transferred
- When the data be destroyed
- How to justify the duration of storage

1.12 "For each source, has the data been anonymised to the greatest degree possible? Detail how each data set has been anonymised."

Source	Anonymisation Method

As previously discussed, data should be anonymous unless absolutely essential for the project. Ensuring the safety of personal data is critical in protecting children's rights.

For each data source, note the method of data anonymisation, giving justification if necessary. There are varying methods of data anonymisation, such as:

- Removal
- Substitution
- Distortion
- Generalisation
- Aggregation
- Rounding and suppression
- Restriction

1.12.1 "If using multiple data sets, could the merging of these datasets re-identify individuals? Have any measures been taken to prevent this?"

Merging different data sets, even if they are anonymised, can make it easier to discern the identity of specific individuals. Whilst this can be a useful exercise in providing better insights about certain behaviours and activities, it can unintentionally put people's private data at risk. This could cause harm at any point during the data life cycle.

This is not only an issue for personal data, but also for certain types of aggregated group data. This could include:

- Gender disaggregated data
- Age disaggregated data
- Child-headed households
- Locations of children
- Children experiencing violence

Whilst these types of aggregated data do not identify specific children, they can expose groups of children to harm if the data is handled irresponsibly.

Detail how this could occur in your project, and steps you will take to prevent this from happening.

1.13 "Is there any reason why your project methodology should not be made public or shared for usage throughout the Collaborative?"

This question encourages you to stop and think about your project so far. This is a reflective exercise, allowing you to assess your project methodology and how it aligns to the Collaborative vision.

Think about:

- Is your project people-centric and designed to maximise benefits for **children?**
- If everyone were to use the same methods as you, would this be a good or a bad thing?
- Have you cut corners or used methods/data that you wouldn't be proud to share?
- What would you think if you found out that another team/company was working in this way?
    - Would you have cause for concern?

Data Compass Guided Questions:

- How have you decided that these types of data will be the best to help you benefit **children?**
- Is the way you are using your data **safe?**
- Is your data sourcing **trustworthy?**
- Are you being **transparent** with your data sourcing and methodology?

RESULTS

1.14 "Who are you required to share your findings with?"

These are the people that you will communicate your findings to regardless of the outcome. It is important to identify this group, as this will ensure that the project is transparent from start to finish. It may be useful to highlight key people within stakeholder groups that you can communicate with, and who can be held professionally accountable for discussing how to share the results of the project.

Some people that you may be required to share findings with could include:

- The Data for Children Collaborative with UNICEF
- Your internal project manager
- Representatives of the client
- Funding bodies

1.15 "What are other applications, both positive and negative, of any methodologies developed within this project?"

Any new methodology is at risk of being extended in both positive and negative ways. Consider how your methods could be transferred to other locations, data subjects and sectors to improve outcomes for children in a similar way to our project. Consider the impact of your methodology in the hands of others who may not have the same intent. It is important to think through potential harms even at this point in the data life cycle.

Are the appropriate data sharing or intellectual property sharing agreements in place to ensure that your methodology will be used by the right people for the right reasons?

1.16 "What are the implications of your project using data science methodology to make predictions about children? How can these be resolved?"

Your project might be using innovative data science techniques to make predictions about children that could have both positive and negative effects. This is a good opportunity to reflect on how people-centric your methodology is.

For example, using methods that can predict when a child might be at risk of being overweight or obese can have positive effects in that steps can be taken to improve that child's health. However, making that prediction and publishing results on what these indicators are for child overweight and obesity could lead to negative targeting of children who fall into this category.

Predictions are not always correct, either. Think about the implications for children if they are wrongly categorised – will this put any of their rights at risk?

1.17 "How may communicating your findings bring about unintended harm to vulnerable groups and stakeholders? Please identify stakeholder groups, respective potential harms and how you plan to mitigate any risks."

Every bit of information that is unearthed can seem important and beneficial, but sometimes it's necessary to step back and think about the unintended negative impact that the results of your project could bring about. Are your findings still proportional to the original project goal?

Some considerations:

- Do the results favour short-term harm instead of long-term good?
- Is your project child-centric and designed to maximise benefits for children?
- Sometimes we think we're using the information gained to bring about good, when in reality it may have negative effects. How will you measure the impacts of your work once the project is over? How will you resolve any issues that arise?
- How are you going to overcome any of these issues? Answer in as much detail as possible.

Stakeholder Group	Potential Harm	Risk Mitigation

1.18 "Are there possible adjustments that can and should be made to communications that may mitigate against; misunderstandings, misrepresentation and misapprehension of data and limitations?"

Is there a way of communicating your findings that would limit any assumptions or interpretations that may be made from the results? Is there a smaller audience to communicate your findings with that would limit any negative impact? Do you have the appropriate data retention and deletion procedures in place?

Your explanations and justifications should be communicated in plain, understandable language that those from a non-technical background can easily understand. This will ensure that your project is transparent and preventing harms that could arise from your results.

1.19 "Will the project still be beneficial if you cannot communicate findings to those outside of the required group listed in 1.14?"

Perhaps you and your team have concluded that your results are best left unsaid. But what do you do now? This is a good time to reflect on the project as a whole, and what your next steps may be.

- Was the project still beneficial?
- Could the results from this project be used as the foundation of another project?
- Are there any other ongoing projects in the Data for Children Collaborative with UNICEF that would benefit from the results of your project?

Results Compass Guided Questions:

- Have you thought about the impact of the results for the target cohort of children and children more broadly?
- How will you ensure transparency if results are not as intended?
- Under what conditions would you be willing to publish?
- If there is a circumstance where you wouldn't feel comfortable publishing your results and is there any way to mitigate this?
- Could this be altered by re-assessing the scope and aims of your project?
- Will publishing your results be a safe thing to do?

You have now completed the "Start" phase of your ethical roadmap, but this is not the end of the journey. We encourage you to look back and reflect on your answers throughout your project life cycle. This can help to keep you focused on your initial project aims and goals, as well as ensuring that you are aligning to the Responsible Data for Children principles as your project progresses.

ON THE ROAD:

2.1 "Has the proposed benefit to children, as detailed in 1.3, changed in any way? If yes, explain how and why these changes have been made."

Your project should use your proposed benefit to **children** as its key motivation throughout the project lifecycle. However, sometimes new findings, methods or input from stakeholders may mean that your initial project 'goalposts' shift slightly. If the impact to **children** is no longer beneficial, you should not move forward with these changes. If the intended benefit has altered slightly, you should be able to justify your reasoning and highlight how it aligns to the Collaborative vision. It may be helpful to include information about the people involved in the decision making process.

It is important that you ensure your project remains people-centric and protective of children's rights at all times.

2.2 "Have any new stakeholders been added to the project? If yes, explain who and why they have been included."

As your project progresses, it may be necessary to include new individuals or groups to the team. As you did in questions 1.5 and 1.6, you should think about your level of engagement with the new party and any positive or negative outcomes they could face as a result of your project. Thinking through these questions will help to prevent harms across the data lifecycle.

You can use the tables below to help structure your answer.

Stakeholder	Level of Engagement

Stakeholder	Potential Positive Effects	Potential Negative Effects

2.3 "Has any of your data science methodology changed since the last assessment? If yes, detail how you came to this decision, and how it aligns to the original project goal and purpose as detailed in 1.1 and 1.2."

Your project should be purpose-driven. As a result, any changes that you make along the way should always align to your project goals and purpose.

Some considerations:

- Why is this method better suited than the one you had previously chosen?
- What new skillsets or training is needed to successfully carry out this method?
- Who was involved in making this decision? Why?
- Can you provide evidence as to why this is the best option?
- Can you explain how this assists you in achieving your project goal?

2.3.1 "Have any new data sources been used? If yes, please complete questions 1.8 – 1.12.1 for your new data source(s)."

Introducing new data sources or new data sets may be necessary to fulfil your project aims. If you are planning on using any new data sources, even at this stage in the project, it is important to give detailed answers about how this data will be used and managed. This will highlight if the new data sets are proportional to your project goals, as well as determining professional accountability.

As before, answer questions 1.8-1.13 in as much detail as you can. You should be able to highlight why this data is best suited to address a gap or an issue in your project so far. Justify your reasoning where possible.

2.3.2 "Have any new parties been given access to the data? If yes, detail who and why this decision was made."

Just as if any new stakeholders have been added to the project, if any new parties have been given access to the data it is important to document this. You should make sure that there access is granted in a proportionate way, and is purpose-driven.

Think about the type of data they have access to, and any potential harms that could result from this. Make sure your new team are aware of the sensitivities in handling children's data.

2.3.3 "Are there any parties that still have access to the data that no longer need it? If yes, detail the steps you will take to ensure that they no longer have access."

It may be the case that there are people who have completed their phase of work, or are no longer involved in the project. To be professionally accountable, you should be monitoring who has access to data at various points throughout the project. Use this question as a prompt to review data access permissions for the data sets you are using.

If an individual has had access to data when they no longer need it, please explain why this is the case.

2.4 "Have you encountered any unexpected results since beginning your project? Explain your answer."

Sometimes we are faced with challenges or issues that we could not have accounted for at the start of a project. This question is an opportunity for you to reflect on any unexpected results that have been unearthed in your project thus far.

If these unexpected results could present harms across the data lifecycle, or threaten children's rights in any way, you should share these concerns. Remember that protecting and helping children is at the core of everything we do.

2.5 "Are you aware of any new ethical issues that may arise as a result of these changes? If so, how do you plan to resolve these?"

During the project lifecycle, some changes may be necessary in order to make sure that you are positioned to achieve your project goals. It can be impossible to try and predict for everything that could happen before your project has even started!

However, these changes could surface new ethical issues, and it is essential to consider the potential impacts that these could have, particularly on children and their rights.

Think through your initial answers, and then consider:

- The types of data you are now using
- The people now involved
- The ultimate positive impact on **children**
- Whether your changes are proportionate to the original project scope
- How you will justify your reasoning for any changes you have made
- What measures you will put in place to mitigate any new risks

<u>On the Road Compass Guided Questions:</u>

- Are you confident that your project is still **child focused** and providing positive impacts **for children**?
- How can you ensure that your decisions, changes and new ways of working are **transparent** and accessible?
- Are you still using and handling your data in a **safe** and **trustworthy** way?

DESTINATION:

3.1 "Has your project output achieved or diverted from your intended project goal? Please detail your answer."

Now you are at the end of your project and can assess its outputs, you should consider how closely they align to the intended purpose and project goal. Perhaps they have exceeded your expectations, or perhaps you didn't quite meet your objectives. Similarly, some things may have changed along the way that have resulted in a diversion from the original project aims. Remember that all projects should be people-centric and project driven – so your outputs should always be useful and beneficial for children.

In any of these cases, it is important for you to look back at what you originally set out to do and evaluate how successful you have been. It may be helpful to think about any implications that could arise if your outputs are not as intended.

3.2 "Given your project output, can you now identify any additional stakeholders? If yes, explain who and why they have been included."

Now that your project outputs are completed, you may be able to identify additional stakeholders who can benefit from your work. You should list any potential new parties and why you think they should be included. You should reference, where possible, how their participation is proportional to the project.

3.3 "Have you stored or disposed of the data, and any copies of the data, as originally stated? If no, explain why."

As before, you should have a clear understanding of who is professionally accountable for data management, including storage and disposal at the end of the project. Refer back to your original answers and ensure that you have dealt with any remaining data in the appropriate way. If you plan to adjust these measures in light of your project outcomes, please detail your new data management strategy and why these decisions were made.

Remember that children's data is particularly sensitive and requires stringent protection.

3.4 "Given the final outcome of your project, who will your results be communicated to?"

Your project outcome could influence who you communicate your results to. For example, you may have unearthed results that could cause harm to children if disseminated publicly. Similarly, you may have discovered insights that could be beneficial to more people than you had originally thought. You should list any additional parties that you wish to communicate to, and justify your answers where necessary. Anyone added should be proportionate to the project goal, and in children's best interest.

3.5 "Will the communication of your outcomes negatively impact any of the project stakeholders? If yes, how will you resolve this issue?"

Just as it is important to assess the positive impacts of your findings, it is also important to reflect on whether your results could have negative effects on those involved. Refer back to your stakeholder list and think about any harms that could surface when you publish your results. This is a useful time to really focus on ensuring that your project is people-centric. If you can identify any negative impacts, you should provide a detailed risk mitigation plan, with clear professional accountability.

For example:

- Could any negative assumptions or predictions be made about your stakeholders as a result of communicating your findings?
- Does communicating your findings favour one stakeholder group at the expense of another? (Not including children, as they are always the priority stakeholder)
- How can you present your results to minimise any risks?

3.6 "What impacts will communicating your findings have on children?"

The overarching goal of all of our projects is to have a positive impact on **children**. Take this opportunity to reflect on how you are communicating your findings and why this is in **children's** best interests. Detail, where possible, how this is in line with your original project purpose. This is a chance to highlight how your project results align with the Collaborative's vision to use data science to improve outcomes for **children**.

Think through what your new insights can achieve. Are you:

- Informing policy?
- Providing a research paper?
- Creating a new tool?

Most importantly: how will **children** benefit from these outputs? Remember that your answers should be people-centric.

It is important to prevent harms across the data life cycle. Your findings may have negative impacts on **children** if communicated to the wrong audiences in the wrong way, or even communicated at all. If there is any reason that you think your findings could have negative effects on children, either directly or indirectly, you should not proceed further. Above all, we have a responsibility to protect children and their rights, even if this means not communicating the project results.

Destination Compass Guided Questions:

- How are you planning on communicating your results to ensure that your insights are **transparent** and accessible?
- Have you stored or disposed of any remaining data in a **safe** and **trustworthy** way?
- Are the decisions you have made to communicate or not communicate your findings in **children's best interest** and documented in a **transparent** way?
- How can you ensure that your findings are **trustworthy**?
- Have you taken necessary measures to ensure that any data included in your results is **safe** and does not expose children to any harm?

Data for Children Collaborative WITH UNICEF

© Copyright

This document was developed in partnership with Ethical Intelligence Associates, Ltd.

Data for Children Collaborative WITH UNICEF

@dataforchildren

www.linkedin.com/company/dataforchildrencollaborative/

THE UNIVERSITY of EDINBURGH
Data-Driven Innovation
Part of the Edinburgh & South East Scotland City Region Deal

THE UNIVERSITY of EDINBURGH

unicef for every child

Scottish Government
Riaghaltas na h-Alba
gov.scot