Department for Education

Responsible Technology Adoption Unit

Guidance for developing data analytics tools

For local authorities developing data analytics tools in children's social care

April 2024

Contents

Summary

This publication provides non-statutory guidance from the Department for Education.

It has been produced to help those who are interested in learning more about, developing or managing the development of a data analytics tool in children's social care.

It is a step-by-step guide that provides practical advice for every step of the **development lifecycle,** setting out what should be done at each stage to ensure tools are safe, reliable and trustworthy. We break down this lifecycle into the following stages:

1. **Problem identification stage**: where you identify the specific challenge you face, and how a data analytics tool might help
2. **Capability Assessment stage**: where you identify whether you have the technical and organisational capability to develop a tool
3. **Choosing a solution stage**: where you decide which type of data analytics solution best matches your needs
4. **Ethics triage stage**: where you identify ethical risks of developing your data analytics tool, and how to mitigate them
5. **Design stage**: where you design your data analytics tool to meet your users needs
6. **Development stage**: where you build and iterate your tool, building in mechanisms to ensure safe and effective use
7. **Testing stage**: where you evaluate whether the tool will meet your objectives, and measure its effectiveness, reliability and robustness
8. **Pre-deployment stage**: where you prepare your tool for live deployment
9. **Deployment and monitoring stage**: where you measure the ongoing performance of the tool as it is used
10. **Retirement stage**: where you consider how to safely take the tool out of deployment.

Who this publication is for

This guidance is for:

- data practitioners who haven't previously developed or used data analytics tools, would like additional advice and support, or are working with a third party to develop a tool.

- project or product managers working on the development of a data analytics tool.

The guidance flags where data practitioners will need to engage with other audiences (for example, information governance, procurement officers/ commissioners, children's services teams) who should be involved in advising or making decisions across the development of a tool.

How this publication should be used

Work your way through each stage of this **guidance**.

Respond to questions in the **ethics workbook** at the end of each phase.

Complete the **ethics triage self-assessment** before starting on the develop stage.

This guidance points to further advice and information in the **explainers** document.

1. Problem identification stage

This stage will help you to engage with teams across children's services who use data and insights to inform decision-making. The aim of engagement is to:

- Identify a problem that should be addressed, which improves service delivery and outcomes for children and families.
- Discuss whether better data insights could be part of the solution.

Identify and map users

Social workers, managers, senior leaders and data practitioners need to be engaged from the start of any process to ensure that their experiences are built into any data solution.

Your first step should be to map and engage with every group of users who will use or be affected by the use of a data analytics tool.

In this section, ‘users’ should be interpreted broadly. They will include those who:

- work directly or indirectly with children and families
- use data to make decisions that relate to or affected children and families.

They are likely to include:

- social care and other practitioners
- managers
- senior decision-makers
- service providers
- multi-agency partners such as the police
- data analysts or Business intelligence developers.

You should engage early and often with potential users about data analytics. It can help to inform decisions about whether a data analytics tool is appropriate, and if so, where and how.

Explore data literacy and attitudes towards data analytics

It’s important to understand the data literacy of children's services teams to understand the types of tools which may be helpful and usable. Engage with children's services teams to understand whether data analytics could be appropriate by exploring:

- their data literacy
- how they currently use and trust data and technology in their role
- how comfortable they feel about data analytics

You may find the Introduction to data analytics resource helpful to explain what data analytics is, and different tools that can be useful to support service provision.

You may ask the following questions to prompt discussion:

- Do they always have access to the information that they need?
- Could additional information help them to do their job?
- What are the main challenges in using data in their role?
- Do they have any questions or concerns about data analytics?

Identify demands and pressures in children's services

If there is interest in data analytics, you can explore how and where data analytics could be of most value.

Consider the problem you want to solve with a data analytics tool, for example:

- information overload at a decision point
- desire to move and prioritise resources
- better matching of children and families to the right services
- lack of information sharing at crucial decision points

It’s important that any data practitioner understands how the system of data recording and use currently works. Shadowing a service can help to identify how and where insights can improve service delivery.

Different types and formats of information may be required by different user groups. For example:

- a social worker may find narrative case note information most useful

- a data analyst might find structured data, e.g. drop-down menus, most helpful

Resources

Completing a user needs table can help to clarify what types of information is needed by different user groups.

Decide whether a data analytics tool is appropriate

You should draw on your user research and data maturity results to decide whether data analytics is an appropriate solution.

Problems may be addressed by different analytics solutions. You should make sure that your analytics tool best fits with your data maturity and capability.

For example:

- A children's services team with medium data maturity and less advanced technical expertise could use descriptive analytics to understand how need has changed over time.
- Children's services teams with high data maturity and technical expertise may use predictive analytics to help forecast future needs, for example, forecasting the number of care placements needed, if subject to robust scrutiny.

In making a decision to develop data analytics, it can be helpful to think about:

- Would data insights be valuable, and respond to the problems identified by users?
- What action could frontline practitioners or others in children's services take as a result of the insights?
- How could this change (and improve) how decisions are currently made? How could this change (and improve) outcomes for children and families?
- What other (non-data) solutions have you considered which could address key demands and/ or support desired outcomes?

Where you decide that data analytics is likely to improve outcomes for children, you may move ahead.

If you decide that a data analytics tool is not the right solution, it is still important to invest in improving your data maturity. It can also put you in a stronger position to develop and use data analytics in the future.

2. Capability Assessment Stage

In the capability assessment stage, you will investigate your technical and organisational capability to determine whether you have the potential to develop a data analytics tool.

You will determine your level of data maturity, and match this to the level of complexity of data analytics solutions available to you.

Complete a data maturity self-assessment for children's services

Before considering a data analytics solution, you should consider whether you have the suitable data resources to create an effective tool.

We advise that you complete a data maturity self assessment, engaging with senior leaders in children's services where appropriate.

Data maturity considers components like:

- How you use data to inform decisions and run services
- Attitudes towards data and its uses
- How data is stored, accessed, shared and quality assured
- The capacity of your authority to undertake data analysis and data science
- The governance of data, approaches to data protection and privacy.

Resources

- The Department for Levelling Up, Housing and Communities' Early Help System Guide self assessment, particularly the 'data' section. While the self-assessment is focused on early help, we suggest that it is relevant to understanding the data maturity of children's services as a whole.
- The Local Government Association's online data maturity self-assessment can also help to provide a broader understanding of your organisation's data maturity.

What your result means

A higher score on all components of the data maturity assessment

Local authorities teams should be encouraged to investigate descriptive and diagnostic analytics to improve their provision of children's social care, and you can proceed to the next Stage.

At a high data maturity, you may be able to develop a more advanced tool. This may include tools which make predictions around the likelihood of an event happening. These tools are higher risk, especially where predictions relate to individual children and families. We advise that you should demonstrate confident and effective use of descriptive and diagnostic analytics, before seeking to develop predictive analytics.

Developing a predictive analytics tool will require:

- Advanced technical expertise
- Strong engagement and support from stakeholders who are likely to be affected
- Robust governance mechanisms including an external oversight group who can effectively scrutinise and challenge throughout development
- Explicit buy-in from senior leadership in your organisation.

Any form of machine learning should only be undertaken by experienced data scientists with relevant experience. You should note that recent evaluations of machine learning have not demonstrated effectiveness in identifying individual risks to cohorts of children.

In addition, you should avoid developing a predictive tool which provides little opportunity for a practitioner or user to reach their own conclusions, for example, a tool which automatically refers children who have received a 'high risk' score for assessment, without a professional considering the evidence.

A medium score, with a lower score on some components

Local authorities may be able to develop a data analytics tool in-house. You should focus on the possibilities for descriptive and diagnostic analytics, but should avoid developing a predictive tool, as it's unlikely that you will have the technical capability and data governance processes required. A low in-house technical score doesn't mean you can't develop or use a tool. You may procure expertise from third parties including academics, private sector developers or non-profit organisations. The explainers chapter on Procuring third party expertise has more information.

A lower score across all components

Local authorities are unlikely to be able to develop a data analytics tool in-house. At a lower data maturity, you should take steps to improve maturity and capability before considering a data analytics solution.

Improve data maturity

Improving data maturity across your organisation can be a long and complex process, as it often involves changing organisational approaches to data collection and use. You may need to work with an external company or organisation to support you to do this.

- The LGA Data Maturity Self-Assessment provides several actions you can take to improve data maturity depending on the step you are at.
- You may find open source analysis templates and tools useful. For example, Data2Insight hosts data resources which have been designed to support data analysts in children's services.

Decide whether to develop a data analytics tool

After reviewing your capability, you can proceed to determine which types of data analytics tools are best suited for the particular problem you are aiming to solve.

3. Choosing a data analytics solution

Once you have carefully determined whether a data analytics solution may be suitable, this stage will help you explore potential data analytics solutions that may be able to help you to respond to the problems identified in the previous stage.

Identify the datasets needed

Data analytics tools may draw upon many different data sources. Having a range of data from multi-agency sources can help to provide a holistic view of a child, area or service.

Data practitioners should:

- identify variables that may help to better understand a problem or issue, and then consider which datasets might include these variables.
- identify and catalogue the datasets that are available and accessible to children's social care.
- note who owns the datasets, and their source.
- consider whether it would be valuable to receive additional datasets from external organisations, for example, youth justice.

Understand the quality of datasets you want to use

Data practitioners should be confident that the datasets you want to use are good enough to develop a tool. Are they accurate, complete and representative of the population under consideration?

To help you to understand the quality (and limitations) of the data, you can:

- record the size of each dataset,
- identify missing data, and
- assess whether certain data fields have been recorded in a consistent format.
- complete an audit on data to assess its accuracy, completeness and representativeness.

For some local authorities, the datasets you want to use won't be good enough to gain meaningful insights from data analytics tools.

If your data quality is poor and you don't have the resources to improve it, we would advise that you don't proceed to developing a data analytics tool. However there are

practical steps you can take to improve your data quality, the **explainers chapter on data quality** has more information.

Case study

The case study below outlines how data analysts identified how to reduce errors in data. Data practitioners mapped data cleaning to identify common errors in Looked After Children data. They found that most errors were caused by inconsistent data in the three areas:

- Data is inconsistent with previous years
- Data is missing or incorrect
- Multiple reviews are recorded for the same child

They also analysed error logs on historic returns, which helped them to understand the scale of the problem, and then identified opportunities to reduce errors.

Understand your systems architecture

Complex or legacy IT systems can inhibit your ability to implement data analytics solutions. Understanding your existing IT infrastructure and systems architecture can help you identify whether data analytics is feasible, or what changes might be needed to facilitate it.

Resources

You may find the Crown Commercial Service's Digital Transformation Guide for technology procurement in local government helpful. It provides advice on procuring software which avoids vendor lock-in.

Case study

One local authority sent a detailed questionnaire to their information analysis team to understand existing technology systems. This included:

- how data was stored
- format of storage
- how data was usually accessed.

They held interviews with information analysts to map processes for data use to help identify key challenges across aggregating, cleaning and analysing. They finally analysed

spreadsheet templates to understand data currently shared and types of analysis that could be supported.

Consider resourcing and costs required

Estimating the total costs involved in developing a tool will help demonstrate value for money, crucial when writing a strong business case. Costs involved in developing and implementing a data analytics tool can include:

- resourcing to manage and complete the design and build,
- operational costs and maintenance,
- additional costs such as public engagement, establishing an ethics committee or sourcing independent evaluation,
- resourcing to provide appropriate training for users (including social workers).
- changing your systems architecture/ infrastructure (if necessary)

Resource

The Rees Centre has developed a cost calculator tool that can be used to explore the relationship between needs, costs and outcomes of services provided to looked after children. It uses statutory return data to explore how costs might be affected when a new practice or service is introduced.

Seek feedback from children's services teams

You should check-in with your children's services team, to:

- share the user research you've completed so far
- share details about the tool you want to develop.

You may wish to ask them about the potential impacts and challenges that may arise from the development of your selected tool, and whether they have concerns that you can address during further development.

Begin a Business Case

A business case can help you to communicate with senior leaders and other key stakeholders about the importance of the project. The decision to develop or use a data analytics tool will likely require sponsorship and a business case to be signed off.

A strong business case will utilise evidence and data from the previous stages to support its assumptions, including user research that relates to:

- desirable outcomes, objectives and impacts
- undesirable outcomes
- any red lines

You should also establish how the tool will improve decision-making and positively impact a range of affected stakeholders.

Finally, identifying successes of similar projects by other local authorities' may also help you to demonstrate the value of the work in your business case. This type of knowledge sharing can take place in local government networks, such as the Local Government Association's advanced and predictive analytics network.

Resource

KPMG has produced guidance for local governments to support strong business cases.

Consider procuring third party expertise

If you have low data maturity or technical expertise, you may need to procure expertise to support the development of a data analytics tool. You should decide whether you:

- have the technical capability and resources to develop a tool in-house
- need to bring external consultants onboard to support you to improve data governance
- need to bring data science contractors/ developers onboard to build a bespoke data analytics tool
- can procure an off-the-shelf tool which meets your needs.

However, you should also consider the opportunities and risks around procuring external expertise.

Developing a tool in-house

- local authorities may have greater control over the data being used when developing a tool in-house;
- however, it requires the local authority to have a fuller understanding of data quality and infrastructure. This is necessary when monitoring the tool.

- it also requires significant investment and internal expertise, especially where a local authority is building more advanced tools in-house.

Procuring third party services or tools

- this can offer specialist data science expertise that may not be available to most local authorities.
- suppliers and consultants are likely to have relevant local government experience.
- there is limited evaluation of the effectiveness of third party tools (e.g. impact, cost and accuracy).
- lack of internal technical expertise means it can be difficult to oversee and scrutinise third parties.
- Local authorities will remain responsible for ensuring the processing and sharing of data remains compliant with GDPR. It is advised that local authorities carry out due diligence checks on the organisation which they are planning to share data.
- local authorities have a responsibility to understand how decisions are made, yet commercial sensitivities may prevent providers from sharing information about how a model is developed.

4. Ethics triage

What are the ethical risks in developing a tool and how will you address them?

At this stage, you should:

- complete the **ethics triage**
- begin relevant impact assessments
- put appropriate governance and oversight in place
- make sure that your business case has been signed off.

Complete the ethics triage

Having identified a problem and chosen a data analytics solution, we recommend that you complete the **ethics triage**. It can help you to better:

- identify specific risks associated with the tool, and inform how you respond to or mitigate risks. It advises whether you should continue, re-consider your tool, or proceed with appropriate data governance
- understand what proportionate data governance looks like, i.e. proportionate to the risk of your project
- where you can find specific advice and information to address the specific risks
- respond to relevant impact assessments (e.g. data protection, equality and children's rights).

After completing the triage, we advise you:

- record the risks identified in the ethics triage in your data protection impact assessment (DPIA) and the **ethics workbook**, outlining how you will address or mitigate them. You may decide that some risks are too high to justify developing the tool, or that a tool of lower complexity is more appropriate.
- identify appropriate measures that you can take to help address and mitigate the risks that have been identified. Some risks will reach a certain threshold that means that you shouldn't develop a data analytics tool.
- discuss the results from the ethics triage with an oversight and/ or scrutiny group for advice on how to proceed and measures that can help to mitigate the risks identified.

Begin relevant Impact Assessments

Impact assessments are a useful tool to identify risks on your project and develop ways to manage them. You do not need to complete all of the following impact assessments, but you should consider which are particularly relevant for your specific tool.

Resources

- We recommend that data analytics projects will likely require a Data Protection Impact Assessment (DPIA) because of the potential impacts they present to a child or family. The ICO has developed DPIA templates or you may use your organisation's own template.
- An Equality Impact Assessment (EIA) can help you to demonstrate that you have complied with the Equality Act 2010 and Human Rights Act 1998. EIAs are recommended where a project may present a high risk to an individual's rights.
- A Child Rights Impact Assessment can help to demonstrate how and where you have considered children's rights, their views and perspectives. The European Network of Ombudspersons for Children guidance suggests questions you may ask across the design and development of a project.

Establish appropriate governance structures

Good governance ensures that there is appropriate accountability for, and scrutiny of, decisions. You may engage with information governance teams for support and advice.

Good governance includes:

- assigning clear responsibility and accountability for the development of a data analytic tool at the outset.
- ensuring consistency and standardisation.
- developing robust mechanisms to allow for internal input and challenge, as well as external advice

Establishing scrutiny or oversight groups

Groups or networks that you may include in an oversight group include:

- internal and/ or regional local government data or technology networks who have technical expertise or experience developing data analytics tools.
- children's services practitioners, managers and senior leaders

- affected stakeholder groups
- legal experts

Establishing independent scrutiny

You may need to establish an independent scrutiny group to advise across the development and use of the tool. This includes where:

- you don’t have internal technical expertise
- the ethics triage has identified your solution as high risk, or has flagged the importance of oversight and governance

You may draw on external experts, including from local academic and civil society partnerships.

Resource

The **explainers chapter on governance and oversight** has more information on establishing effective governance structures and appropriate scrutiny.

Moving to Develop Phase

Before moving to the design stage we recommend that you have:

- completed each action in the guidance so far
- completed the first five questions in the ethics workbook under the “explore phase” heading
- received sign off on the business case from senior leaders in children's services, to allow you to progress

5. Design stage

How will your tool best meet the needs of end users, children and families?

This stage will help project teams to design a data analytics tool.

Identify end users and affected stakeholders

End users and affected stakeholders will be engaged in co-design throughout the development of the tool. Taking your solution, identify:

- end users of the tool
- those who might be affected by it. This will likely include: service users, experts by experience, and organisations or groups who could represent their perspectives.

The riskiness of your tool will dictate how broad your user engagement needs to be:

- example 1: You are developing a descriptive tool that provides evidence to help senior leaders to make strategic decisions about resource allocation. You will need to engage with senior leaders and managers.
- example 2: You are developing a tool that predicts individual children who are unlikely to be ready for school at age five. This can enable scarce resources to be targeted appropriately at the families who need it the most. You should engage heads of early intervention, your local authorities' children in care council, Special Education Needs Coordinators (SENco) and designated safeguarding leads at local primary schools.

Collect data to help evidence outcomes and effectiveness

You should have worked with key stakeholders to identify desirable outcomes, objectives and impacts for the tool in your business case. You now need to think about what evidence you may need to collect to demonstrate the effectiveness of a tool, and how you will measure it.

Establishing success criteria

Data analytics tools should demonstrably improve outcomes in children's social care. This may be that it:

- leads to better outcomes for children and families.
- improves existing decision-making processes, for example:

    - providing relevant and accurate insights to decision-makers
    - saving time by streamlining administrative tasks
    - providing value for money
- changes behaviour or attitudes

Setting accuracy thresholds

If you are developing a predictive analytics use case, you should set appropriate thresholds for accuracy including:

- precision
- recall
- accuracy against different demographic groups.

Setting an appropriate level of accuracy is contextual and can depend on the decision that a tool is supporting. For example: the risk of harm to a child may likely be higher if a tool misses a child protection case, than when it misses a child-in-need case.

Benchmarking

Data practitioners should now collect and record qualitative and quantitative data to help to evidence the status quo. This is likely to include:

- how the decision making process currently works
- current user behaviour in the decision making process
- data to evidence outcomes among children and families
    - children's outcomes
    - support they receive
    - services they have interacted with
    - demographic data, for the purposes of monitoring for biased results.

Consider acceptable data use

Data practitioners should have already identified the datasets needed to develop the tool. It's advised that you now engage with your users and affected stakeholders to consider their views on:

- acceptable data use
- how access to data should be monitored and restricted to protect children's privacy, where necessary.

As part of this process, you can share the datasets you think are necessary for developing and using the tool with your affected stakeholders.

You may find it helpful to draw on the methods below to support discussion.

Example methods to restrict data use and access

- limiting access to information that may be sensitive (e.g. only managers can access it) or surfacing information that led to a risk score (without showing the score itself) can help to avoid stigmatising a child or their family.
- pseudonymised data means children and families' names and other information that identifies them is removed. This means their identity is protected if a professional doesn't need to know it. It may be only those who cross a predefined risk threshold can be identified to the relevant child and family practitioners.
- limiting information that can be accessed by practitioners to only that which is necessary. If the tool supports decisions around assessment, more information might be needed than at the referral stage.
- using a larger dataset, which includes data on a larger number of children and families, to develop a tool may improve the representativeness of the dataset, and reduce the risk of bias. It may also help to flag children who may be vulnerable but aren't already known to children's services.

Co-design the tool's interface

Project teams may wish to conduct user research to iteratively design, develop and test prototypes with users of the tool. This will help to support user-centred design, to create a tool which supports users to make decisions without becoming an additional burden.

In addition, the design of the tool needs to align with the end users' needs. For example, if the tool is designed for social workers, it needs to mimic how social workers expect to store and retrieve information, and the types of information that they need to be readily available.

Building on the user needs table you developed in the **problem identification stage** as you design the user interface, you should make sure that you understand:

- users' data literacy

- how users normally use data
- whether users trust data and technology
- what tools and technology they have previously used, and how
- what has worked well or frustrated them in previous technology they've used
- what mechanisms or processes are needed for end users to challenge the result from a tool

You should think about how to make the tool as easy to understand and engage with as possible. Especially where practitioners or senior leaders have lower data literacy. This may necessarily mean making the tool less technically complex.

Case study

Social Finance worked with Leeds and Stockport Council to develop Family Context. It was collaboratively designed with child and family practitioners.

Social Finance found that practitioners spent a lot of their time searching for information about families they work with, which inhibited timely decision-making. They worked with practitioners to understand how information should be presented, and developed and iterated a prototype. Child and family practitioners wanted:

- Simple presentation, larger font size
- Expanded functionality to make information more digestible
- Information on when the last contact with a service occurred
- Information on how recent the information was, how far back the search went and which geographic areas had been checked

Developing multiple iterations of a tool means that users have the opportunity to make the tool as relevant and useful as possible. We will consider feedback loops to engage users and affected stakeholders in the **development stage.**

Consider wider impacts of introducing a new tool

It is important to consider how a tool might impact upon existing processes and ways of working in children's services, for example:

- How might existing data collection or recording processes need to change, to incorporate a tool?

- What might be the impact of the tool's insights in terms of resourcing and workload? If the tool helps to identify additional children and families who need support, do you have the capacity to act on those insights?

However well a tool is designed, it may be used in ways you haven't designed for or produce outputs that you may not be able to anticipate or expect.

Revisit data quality

You should have assessed the quality of the datasets you want to use in **Choosing a Use Case.**

Identifying bias

Is it likely that certain groups would be over- or under-represented in the data used to develop the tool? For example, are there historical trends or practices that mean certain demographic groups:

- Are absent from the dataset?
- Make up the majority of individuals in the dataset?

You should also consider whether there are variables in the dataset that could be used as proxies for protected characteristics.

Now, you should take steps to mitigate data quality issues, for example, through:

- data cleaning
- making the datasets more representative, possibly through additional data collection.
- recording the limitations of the data, for example, missing data.

Adopt appropriate data and development practices

Local authorities should adopt good practices for sharing and using data. For example:

Role-based access control systems

This can make sure that only authorised and named personnel are able to access higher sensitivity datasets or databases. This can minimise the risk of sensitive data about children and families being used inappropriately.

Robust logging infrastructure

This can create an audit trail of how data is accessed, used, and modified. This is invaluable in troubleshooting errors or problems caused by a data-driven tool. It can also log and flag potentially suspicious searches, e.g. family members.

An application programming interface (API)

An API defines a standard set of functions for accessing data which can be easily accessed via software code or through a browser. Making data accessible via APIs can enable sharing in a more efficient, automated, and secure way, for example by limiting sharing of a dataset to the specific records that are required.

Organisations should seek to implement authenticated APIs, rather than sharing datasets as spreadsheets via email, for example. Where data is shared externally, ensure that it is shared with an appropriate recipient, e.g. a safeguarding lead.

Good software development practices

Data practitioners should also ensure that they have good software development practices in place.

This can include:

- writing robust and comprehensive unit tests
- performing thorough code reviews
- using version control software such as Git.

Resource

You may find the following open source training materials collated by the Software Sustainability Institute can help you to build reproducible analysis.

6. Development stage

What mechanisms or processes have you developed to support safe and effective use of the tool?

This stage will guide you through the first iteration of a tool's development to create a prototype, or multiple prototypes.

Explore and build in privacy and security protections

We advise that Data practitioners explore different options for protecting data.

This will likely require you to do the following:

- Minimising data use to that which is necessary and proportionate
- Restricting access, as discussed in the **design stage** ('Consider acceptable data use')
- Appropriately de-identify sensitive datasets in line with the ICO's Anonymisation Code of Practice.
- Consider and document re-identification risks in your DPIA.
- Explore how different de-identification techniques might affect the tools' performance.
- Complete user testing to understand how de-identification techniques affect the usability and accuracy of the tool.

Resource

Follow the ICO's Anonymisation Code of Practice to help you appropriately de-identify datasets.

The **explainers chapter on data use and sharing** has more information on privacy-preserving mechanisms that can help manage and mitigate these risks.

Balancing privacy and transparency

A key challenge of providing access to sensitive data is balancing privacy and transparency. In order to make use of data, it needs to be accessible (transparency). This may necessarily compromise data subjects' privacy to some degree.

De-identification techniques include redaction, tokenization, or k-anonymity. When applied to sensitive datasets, they conceal all but the minimum amount of information

necessary from the person accessing the data. This is known as pseudonymisation or anonymisation.

Traditional de-identification techniques are limiting in that they necessarily reduce the utility of data and may be vulnerable to re-identification attacks.

A set of emerging privacy enhancing technologies (PETs) are addressing these shortcomings. They can enable data to be accessed with strong guarantees of privacy and security, whilst maintaining the data's utility.

Resource

The Responsible Technology Adoption Unit's PETs adoption guide can help data mature local authorities to explore how PETs can be applied in practice.

Incorporate explainability, transparency and scrutiny

Data analytics solutions should be built with explainability, transparency and scrutiny in mind.

- explainability is the ability to interpret, understand and explain how a data analytics tool has provided the outputs it has
- transparency is clarity over how a system works, who is responsible for its deployment, and how users are affected by it
- scrutiny is the ability for users or affected individuals to challenge outputs from a data analytics solution and, if necessary, appeal for outputs to be changed.

Together, these three concepts are essential to ensure that users and affected individuals retain their agency over the use of a data analytics tool, and are not adversely affected by unintended consequences of the system's use.

In practice this means that:

- end users have information - either via training or through a function in the tool itself - to understand how a tool has reached a conclusion, suggestion or insight.
- end users, affected stakeholders or members of the public can challenge, provide feedback/ concerns or ask a question about the tool, or its insights. This could be via an open forum or channel, or providing a technical mechanism or feature within the tool itself to provide direct challenge.
- the tool's code is open to technical scrutiny by experts, for example via Github or with data practitioner groups. In addition, technical and security experts can red-

team the tool. Red-teaming is a process of challenging a policy or intervention with the aim of reducing risks when a tool is deployed live.

- an external expert can review a tool, and predict with strong likelihood what it's going to do.

Feedback processes

Establishing different types of processes may be more or less appropriate for different stakeholders and purposes. All feedback should be considered and fed into future iterations of the tool. Any feedback process should be clear on:

- how users and other stakeholders should engage with it
- how users and other stakeholders can raise or flag concerns
- whether the process is anonymous
- how users and other stakeholders will be included in any resolution

7. Testing stage

Do the results from testing meet the threshold you set for success? Is the tool sufficiently effective, reliable and robust to be used in children's social care?

This stage will help you undertake appropriate 'systems testing' of the tool. It will guide your decision on whether the tool is ready to be used in a live setting. In any testing and review stages, you should test both:

- the outputs from the tool
- how the tool itself is being used (testing with end users)

The testing stage should ideally be separated into two testing parts.

- alpha testing is "lab-based" tests against historic data, to ensure that the tool is technically functional
- beta testing is operational tests in a real-world environment

If you have procured an off the shelf tool, you should make sure that it is tested and evaluated in your context.

What effectiveness looks like

Testing should provide clear evidence that the tool is:

- likely to improve outcomes for children and families (as set out in your Business Case)
- likely to work better than existing approaches/ delivers a new capability
- accurate
- robust and reliable
- effective for its intended purpose

If testing does not provide clear evidence that the tool is effective, the project may need to return to the **development stage** for redesigning or it may not be able to proceed at all.

Complete alpha testing

Alpha testing can help you to understand the accuracy and security of a tool. When writing tests, map out different outcomes that might occur from using the tool, and write test cases that consider these ‘edge’ cases or exceptions.

If you are alpha testing, you could use historic data to test the prototype. You should test:

- accuracy, which includes precision and recall rates.
- accuracy on different demographic groups. This can help identify where there may be different and unfair outcomes for different groups.
- security of a tool and its underlying infrastructure. Cybersecurity standards like the ISO27001 or through external penetration (pen) testing are good ways to test security. Pen testing simulates a cyberattack to evaluate the security of a system, and identify potential weak spots.

Complete beta testing

You should be mindful of the limitations of only testing with historic data (e.g. alpha testing), as results can be misleading. You are likely to get more meaningful performance metrics by testing the tool on live data. This is why beta testing is important. For example, a tool that uses a predictive model trained on historic data will not account for recent policy changes designed to impact outcomes.

Beta or ‘scenario testing’ means piloting the tool in a live context to evaluate its accuracy ‘in the field’. Here, the tool will be presented with new and unfamiliar live data. A limited number of users will use the tool, however testing will take place in a way that does not yet impact on members of the public.

To keep testing separate, you may wish to:

- use live data which runs in parallel to existing processes
- establish a ring-fenced area or sandbox
- anonymise the names of individual children or families.

The exact approach to beta testing will depend on the technology being developed. It may include:

- running new systems in parallel to existing ones on a common population (if possible) to assess differences in outputs and outcomes.
- testing a data-driven intervention in a limited geographic area.

- taking a phased approach to live deployment, for example, trialling the tool on a real subset of live data for a set period of time. You could:
    - pass data-driven insights to users responsible for decision-making over a limited time period
    - then assess the extent to which use of the data-driven insights is improving the quality of decision-making.

Provide training for end users

Children's services teams need to understand the reasons behind any result or recommendation from a tool, in order to combine the tool's insights with their own expertise.

How training should be provided

- those who have developed the tool should provide appropriate training
- if you are working with a third party developer, the developer should provide training to users and data practitioners for when the developer is no longer involved.
- information governance teams may provide support.

What information users should understand

- what the tool is intended to do, the expected benefits, outcomes and impacts
- information about its capabilities and limitations, and mechanisms that have been put in place to mitigate the risks
- initial results from systems testing (described above)
- how they should appropriately use and disclose sensitive data.

What users should be shown

- how the tool works
- how and where it should be incorporated into the decision making processes
- how users should consider the information provided by the tool - considering its capabilities and limitations.
- how users can directly challenge a result

Testing with end users

You should make sure that a broad range of user skill levels and abilities are included in user testing.

After any test or review period, you should also seek feedback from users. You could explore:

- what users like and dislike about the tool
- how they think it could be improved
- whether users feel that the tool is intuitive and easy to use
- whether they think it has a positive or negative effect in supporting their decision-making
- how easily they think the tool fits into the existing decision making process
- whether users trust the results
- whether they feel comfortable to flag a concern or suspected error in the tool

Incorporating feedback

The tool may need to be adapted if users:

- don't understand how it works or it doesn't fit in well with their existing decision-making process,
- don't trust its results,
- don't feel confident to challenge its results.

Moving to pre-deployment

You should share the results from alpha and beta testing with end users, affected stakeholders and technical experts for feedback and review.

We advise that you only proceed to the **pre-deployment stage** if:

- results from technical and user testing indicate that the tool is effective, accurate and reliable
- the tool has been validated by experts outside of your project team

8. Pre-deployment stage

Do you have appropriate governance in place for live deployment?

This stage will help you to ensure that the right governance is in place ahead of deployment.

Revisit resourcing, governance and oversight

This will help to make sure that you have adequate capacity and resource assigned to the tool, post-deployment. You should:

- make sure that there is sufficient expertise, capacity and resource assigned to named individuals for the ongoing monitoring and review of a tool once it is live.
- revisit any risk log or assessment to make sure that it is up to date, and effective in identifying, monitoring and raising risks to appropriate levels of seniority.
- for additional accountability, your oversight or scrutiny group should remain in place to monitor and scrutinise the results of the tool once it has been deployed

Handover between a third party and local authority

Where you have procured third party expertise, you should ensure that there is a detailed handover from the third party to you.

In addition, the third party should provide thorough training and upskilling for data practitioners who will be responsible for reviewing, updating and maintaining the tool. Training and upskilling may include:

- sharing technical details about the tool, including what it's intended to do and its function
- the tools' limitations and the potential risks in using the tool. Third parties should outline ways that data practitioners can mitigate and address risks.
- how practitioners should review, test and update the tool, and schedules for reviewing it.

Prepare for evaluation

You should consider how the tool will be evaluated in a live environment to:

- help demonstrate effectiveness

- identify where it needs to be improved or adapted

You should bring an independent evaluator onboard to evaluate the effectiveness of the tool, where you are developing a high risk or novel use case. For example, the tool you have developed identifies individual children and families, or makes predictions. You could engage with local universities, academic institutions or an evaluation consultancy to do this.

Process evaluation

Process evaluation focuses on:

- identifying the strengths and weaknesses of the processes surrounding the use of a tool
- considering users' perceptions and reactions to the tool
- making recommendations for adjusting the structure and/ or implementation of the tool.

Impact or outcome evaluation

Impact/ outcome evaluation aims to provide an overall assessment of the outcomes of a tool in terms of its benefits, merits and worth. Outcome evaluation may be a longitudinal process taking place over several months or even years.

You should:

- Compare outcomes with the benchmarking data you collected at the beginning of the project
- Measure outcomes against the success criteria that you set at the **design stage.**

Controlled trials

Well-designed controlled trials can provide significant confidence in the efficacy of a data analytics tool, ultimately leading to better outcomes for children.

Local authorities should communicate transparently about any controlled trial, to minimise the perception that a trial may be "experimenting" on children. Having an independent expert evaluator conduct the trial can build greater trust in the process.

Resource

Refer to the **explainers chapter on testing and evaluation** for more information about process and impact evaluations.

Prepare to publicly communicate about the project

Those involved in the development of a tool are encouraged to complete and publish information about your tool.

As part of any public communication and engagement, you should:

- establish and resource a forum or process for public stakeholders. This is so that they can provide general feedback about questions or concerns they have about a tool.
- review and update any impact assessments to ensure that they are accurate. These should be published as part of the transparency standard.

You should make sure that you have relevant sign off, including from the corporate communications team, before you publish this information.

Resource

The Responsible Technology Adoption Unit and the CDDO developed an algorithmic transparency standard template. Completing the template will help you to provide clear information about any data analytics tool you use, and why you use them.

You can find guidance on how to complete the transparency template.

Checks ahead of deployment stage

Before moving to the implement phase:

- you should have completed the first eleven questions in the **ethics workbook**
- senior leaders in children's service should agree that the tool can and should be deployed live

9. Deployment and monitoring stage

How is the tool effective in supporting more informed decision-making?

This stage will help you to:

- complete continual review and monitoring of the tool,
- complete appropriate evaluation of the tool in a live setting,
- make iterations and adaptations as necessary, and with the right governance in place.

Establish set intervals to review and update the tool

It is necessary to establish regular points for data practitioners or other technical experts to review the tool. This can help to make sure that it performs as expected and identify how it can be improved.

Regular monitoring can also help to identify any unintended consequences that weren't foreseen. It is likely that you will need to monitor the tool more regularly immediately after it has been deployed, for example, monthly.

Reviewing may include:

- testing the tool's accuracy, including on different demographic groups
- testing the tool's security
- reviewing data quality of input data
- re-training the tool
- end user testing and feedback

Special consideration should be given to a tool which is dependent on up-to-date data, as it is likely to need to be updated at semi-regular intervals to remain effective. For example, models developed using machine learning algorithms will need to be retrained on more up-to-date training data.

In addition, end users and affected stakeholders should provide feedback on the tool. With these stakeholders, you may consider:

- whether the tool is still relevant and necessary
- whether the outcomes and objectives for the tool remain important

- whether they understand how to - and feel comfortable to - flag concerns about the tool.
- whether there are appropriate forums and communication channels in place to support open feedback about the tool
- whether there is evidence that the tool is being used.

Begin process and / or impact evaluations

Relevant experts should begin an evaluation of the tool in a live context. You should understand:

- whether and how behaviour has changed since the tool has been used
- users' views on the functionality of the tool
- whether the tool's insights are changing decision-making
- whether insights informing decision-making are improving outcomes in children's social care

User observation and interview

Users of the tool can be interviewed and observed as they use the tool. This can enable an understanding of:

- its impact on their work
- their perception of whether it is providing benefits, for example by improving their decision-making, or helping them use their time more efficiently.

Publish information about the tool

You should publish details of your tool (via the algorithmic transparency standard or otherwise), as well as any impact assessments you completed.

Public information should be reviewed and updated as necessary, to ensure that it is accurate. In addition, publishing the results of any evaluations can help to demonstrate transparency, and encourage trust.

Make sure that you also communicate the results from evaluation and testing in a way which is appropriate for different stakeholders.

Resource

The Responsible Technology Adoption Unit and the CDDO developed an algorithmic transparency standard template. Completing the template will help you to provide clear information about any data analytics tool you use, and why you use them.

You can find guidance on how to complete the transparency template.

Update and iterate the tool

It's important to make sure that you are continuing to review, test and iterate the tool once it has been deployed. You may decide to adapt the tool in the future, for example:

- expanding the geographical area that the tool is deployed in
- including further datasets
- impacting on a larger number of children and families
- using more advanced analytics
- sharing with additional external partners

Scope changes

If the scope changes, it is very important to make sure you are amending and updating data governance processes to give sufficient consideration to adaptations. This includes, but is not limited to:

- engaging with end users and affected stakeholders to seek their views on the proposed amendments
- revisiting the ethics triage tool, and identifying if governance processes need to be strengthened
- engaging with information governance teams to make sure that it meets legal requirements
- reviewing and updating any impact assessments, including a DPIA
- reviewing and updating the information you provide to the public

10. Retirement

Where there is evidence that the tool is no longer relevant or is not meeting its intended objectives, it may need to be wound down and then retired.

You should consider how the tool is integrated into your workflows and other IT systems. How would these be impacted by taking the tool out of service?

A phased approach, where the tool is first deprecated before being decommissioned may be beneficial.

Resource

The government's service manual provides advice on retiring a product or service.

Department for Education

Responsible Technology Adoption Unit

About this publication:

enquiries https://www.gov.uk/contact-dfe
download www.gov.uk/government/publications

Follow us on Twitter: @educationgovuk
Connect with us on Facebook: facebook.com/educationgovuk