GMCA GREATER MANCHESTER COMBINED AUTHORITY

MANCHESTER CITY COUNCIL

Wigan Council

Ministry of Housing, Communities & Local Government

ISLE of WIGHT COUNCIL

East Sussex County Council

STOCKPORT METROPOLITAN BOROUGH COUNCIL

Department for Education

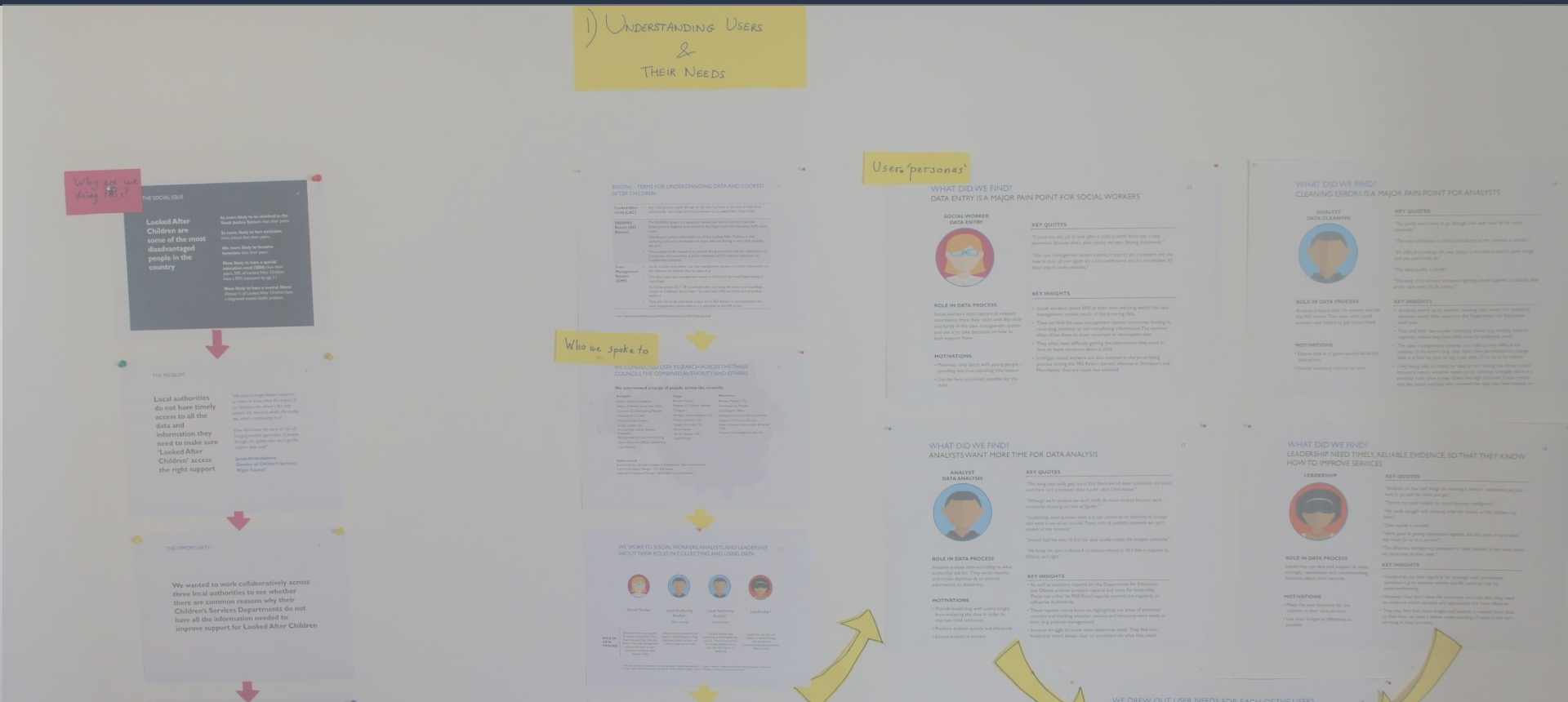
West Berkshire COUNCIL

milton keynes council

SOCIAL FINANCE

QUALITY DATA ON CHILDREN IN CARE - ALPHA

FINAL REPORT



EXECUTIVE SUMMARY

- This project was a collaboration between 9 councils, led by GMCA, with MHCLG, DfE and Social Finance, aiming to explore shared solutions to improve the quality of data on children in care, to enable councils to trust their data in using it to improve services

- We explored solutions both for avoiding errors in data and improving the cleaning process and concluded that the strongest common need with the clearest route to a shared solution was to improve the identification of errors and the cleaning of data

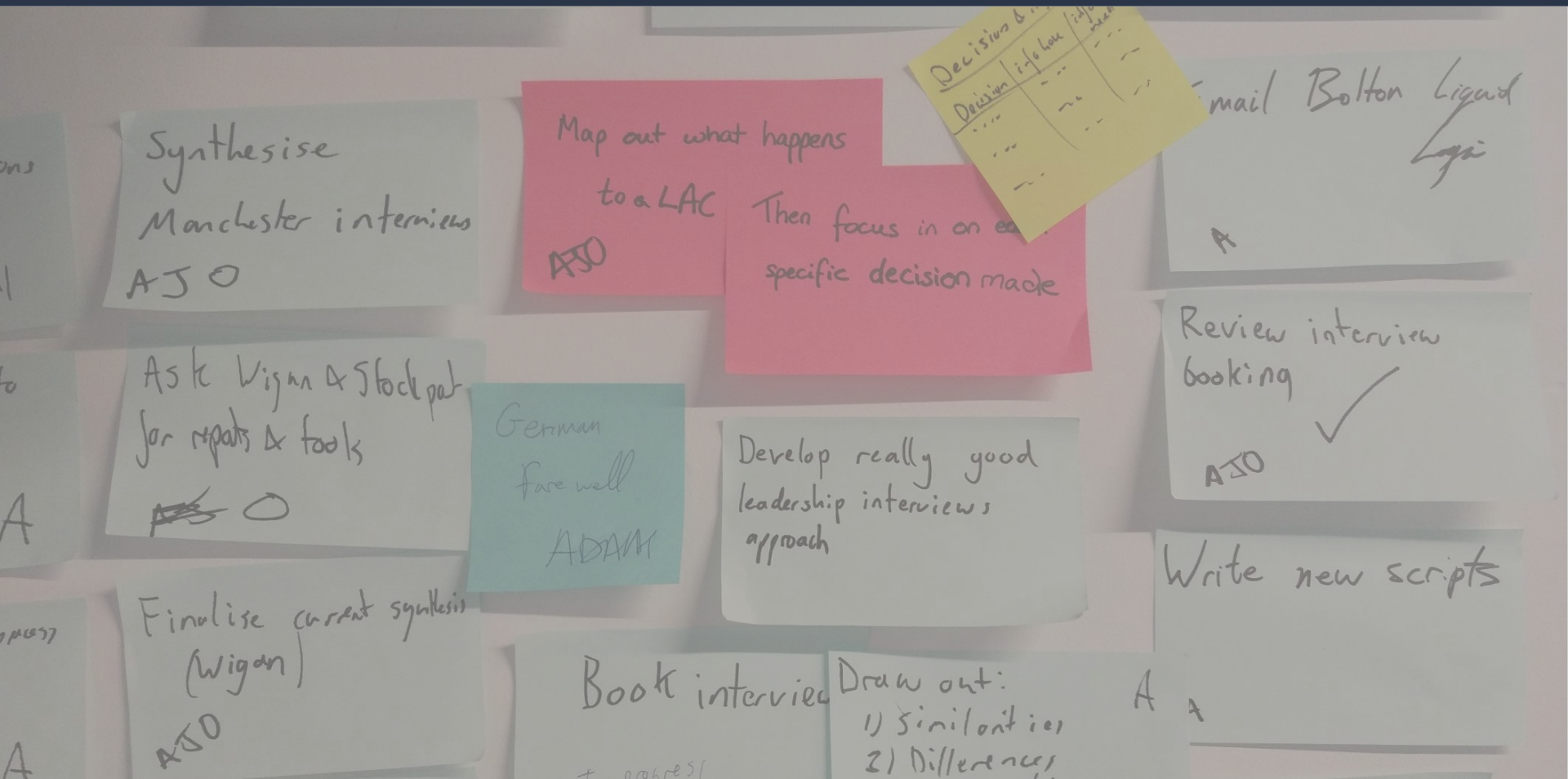
- We prototyped and iteratively tested a solution that would help analysts identify errors in children in care data using the DfE's SSDA903 validation rules and automatically identify the correct information for the 37% of errors which are placement related, using Ofsted data

- To ensure a prototype can be shared easily it needs to a) not require data sharing and b) not require installation of a software. We successfully prototyped an innovative solution to overcoming these barriers using a browser-based tool which runs on local data

- We conclude that we have a design for a tool that every council in the country could use to improve the quality of data in children in care and are proposing to progress to beta to build this and share with all. This would save significant amounts of the months each year analysts spend cleaning data, but even more importantly would help ensure leadership trust their data and can use evidence to improve vital services for these children

CONTENTS

1. PROJECT BACKGROUND



BACKGROUND AND CONTEXT

THE PARTNERSHIP

This alpha project was a cross-council partnership of seven councils led by GMCA and funded by the Local Digital Fund to improve the quality of data on children in care, to enable better evidence and improve support

We collaborated with MHCLG and the DfE, and delivered it with the support of Social Finance, following the principles of the Local Digital Declaration

GMCA GREATER MANCHESTER COMBINED AUTHORITY

MANCHESTER CITY COUNCIL

Wigan Council

STOCKPORT METROPOLITAN BOROUGH COUNCIL

SOCIAL FINANCE

Department for Education

Ministry of Housing, Communities & Local Government

ISLE of WIGHT COUNCIL

West Berkshire COUNCIL

East Sussex County Council

milton keynes council

Looked After Children are some of the most disadvantaged people in the country

4x more likely to be involved in the Youth Justice System than their peers

5x more likely to face exclusion from school than their peers

40x more likely to become homeless than their peers

More likely to have a special education need (SEN) than their peers. 59% of Looked After Children have a SEN statement by age 11

More likely to have a mental illness Almost ½ of Looked After Children have a diagnosed mental health problem

Local authorities do not have timely access to all the data and information they need to make sure Looked After Children access the right support

"We need to target limited resources so I need to know what the impact of our decisions are, where's the cost, where's the demand, what's the quality like, what's contributing to it?

If we don't have this we're at risk of bringing another generation of people through the system who don't get the support they need"

James Winterbottom
Director of Children's Services
Wigan Council

WHY SOLVE THE PROBLEM?

Situation

Support for children in care needs to improve

The support for two-thirds of children in care "requires improvement" or is "inadequate" as assessed by Ofsted. The impact of this is poor outcomes across health, mental health, employment, housing and crime, costing local authorities and central government departments billions each year

This must be done whilst needs rise and budgets fall

Over the past decade the number of children in care has risen 20%, whilst budgets supporting them have reduced in real teams by 25%

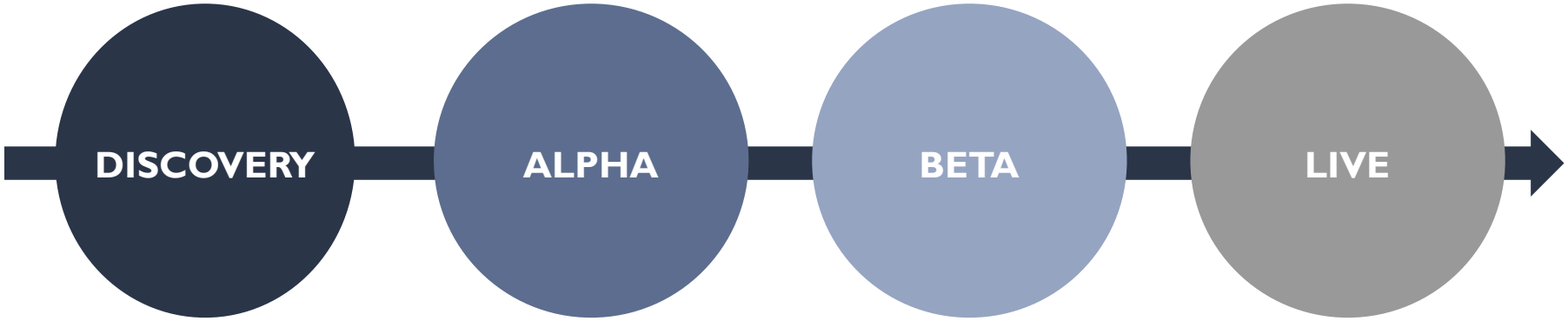
Implication

Major changes are needed

For changes to be effective, they need to be based on evidence. This evidence requires good quality data

We need to improve data quality to enable this change

PROJECT ROADMAP



Jan 2019 to Mar 2019

Dec 2019 to March 2020

Explore the problem:

- Research with users of data on children in care to understand what needs to improve
- Identified data quality as a key problem

Explore solutions to meet user needs

- Decide how to build the service/product
- Test prototypes
- Find out whether a solution is technically possible

Developing a working version of the product/ service e.g. minimum viable product (MVP)

Continuously improve and iterate the product / service

LEARNINGS FROM DISCOVERY

DISCOVERY RECAP – GMCA, GM COUNCILS, SOCIAL FINANCE AND MHCLG COLLABORATED TOGETHER

Project goal

Understand why councils don't have timely access to all the data and information they need to make sure Looked After Children access the right support and how we could improve this in a common way

Local Authority Partners

GMCA GREATER MANCHESTER COMBINED AUTHORITY

Wigan Council

STOCKPORT METROPOLITAN BOROUGH COUNCIL

MANCHESTER CITY COUNCIL

DISCOVERY RECAP – CORE USERS

There are four key users of data on children in care

We did 29 interviews with these users

LEADERSHIP

SOCIAL WORKERS

ANALYSTS (data cleaning)

ANALYSTS (data analysis)

DISCOVERY RECAP – WE LOOKED IN PARTICULAR AT THE "SSDA903" RETURN

The SSDA903 return is the key dataset on children in care. Every council must submit this set of data to the Department for Education on a yearly basis

Background

Frequency: Once a year, all Children's Services Departments in England have to complete the SSDA903 statutory return (903 return)

About: This is a set of data reports on any child who has been Looked After in their authority at any point during the financial year

Purpose: this return aims to provide the government with the necessary information to: (1) evaluate the outcome of policy initiatives and (2) monitor objectives on Looked After Children

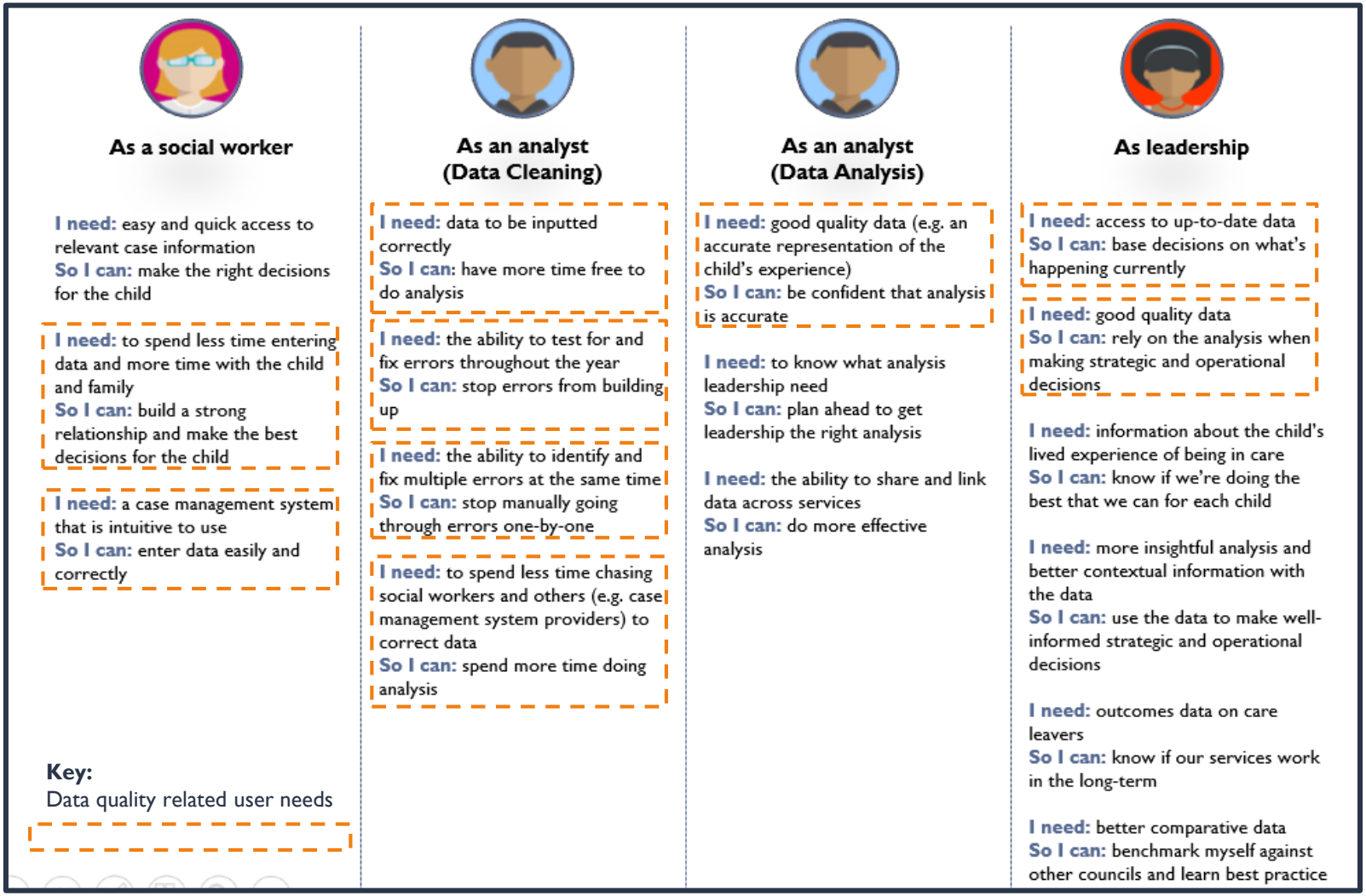
Reporting process and format: All data has to be uploaded to the Department for Education's online portal in either xml format or in multiple csv files

Data cleaning: Analysts, business support, quality assurance, social workers and others work to fix errors in the data

National report: Once the DfE have finalised all the data, they produce a pdf report, *Children looked after in England (including adoption)*, along with data tables, in November

DISCOVERY RECAP – USER NEEDS

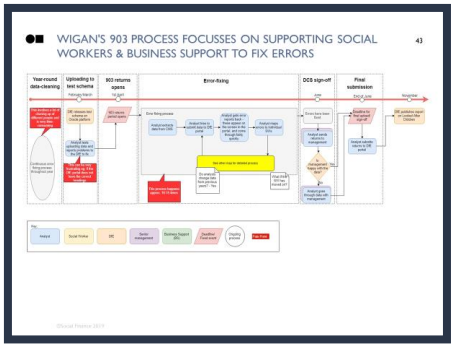
Our 29 interviews gave us a longlist of user needs. From these we identified data quality as a key common barrier across all: stopping leadership trusting analysis and wasting time for analysts and social workers. We therefore focused our work in on data quality

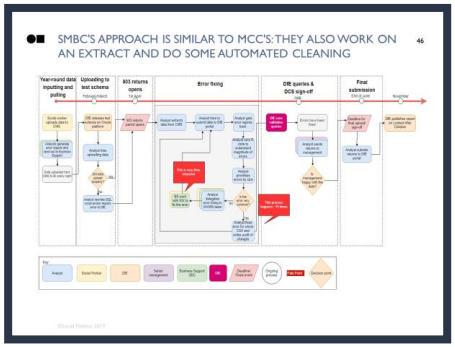


DISCOVERY RECAP – PROCESS MAPPING & ANALYSIS

Having identified that data quality was a core problem, stopping councils getting timely access to all the data and information they need on children in care, we explored this further through:

1. Mapped the cleaning and submission in each council

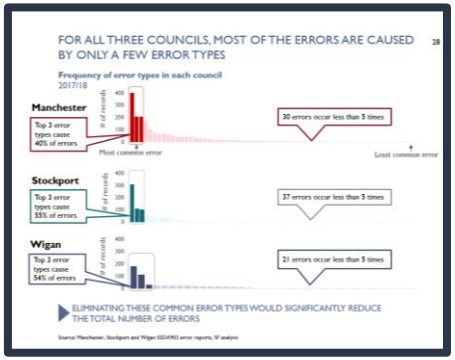


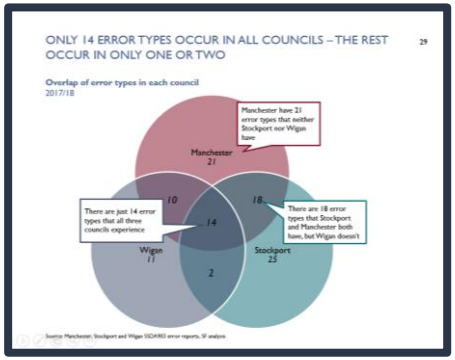


Showing:

- The cleaning process is complex and intensive
- There is significant variation between councils, but no clear best practice

2. Analysed the errors that each council sees in their data





Showing:

- Just 3-5 error types cause >50% of errors in each council
- However these vary between councils, suggesting a common solution for all councils can't just target specific error types

DISCOVERY RECAP – FINDINGS

Findings summary:

1. Data quality is one of the major barriers stopping councils improving services with evidence. It is also a major pain point for analysts and social workers
2. This pain point was common across the three councils
3. Leadership have low trust in data, particularly when uncleaned
4. However analysts can only identify errors to clean data during a limited window when the SSDA903 submission is live

Implication: This means that, throughout most of the year, data quality is poor and leadership don't trust data, hindering their use of evidence to improve services and outcomes. Over the course of Discovery, we pivoted to focus on data quality as key barrier of the use of data to improve services

Conclusion: Progress to alpha to explore how to ensure good data quality all year round. We identified that helping analysts identify errors year-round would be valuable for them and leadership, and could be one approach

ALPHA OVERVIEW

ALPHA GOAL:

Test approaches to help councils ensure good quality data on children in care all year round…

…so they can improve support using evidence and analysis teams can save time on data cleaning

THE OBJECTIVES OF THE ALPHA PHASE WERE TO:

1. Test how common the user needs identified in Discovery are across other councils
2. Understand the impact that meeting these needs would have
3. Explore solutions to meet user needs
4. Prototype and test solutions to meet user needs

PARTNERS

IN ALPHA, WE EXPANDED THE NUMBER OF COUNCILS FROM 3 TO 9 TO ENSURE WE COULD BUILD A COMMON SOLUTION FOR ALL COUNCILS

Key:

- **3** Local councils involved in Discovery and Alpha
- **3** Local councils involved in Alpha
- **3** Additional local councils consulted in Alpha

** Consulted local councils took part in the project at a later stage, to confirm user needs and test the prototype.*

OUR COLLABORATION WAS ENABLED BY CLOSE COMMUNICATION AND WORKING IN THE OPEN

Show & Tells

Fortnightly

- To test findings of user research with project partners and refine them
- To test ideas of solutions
- To take decisions about going forward or not with a solution

Drumbeats

Weekly

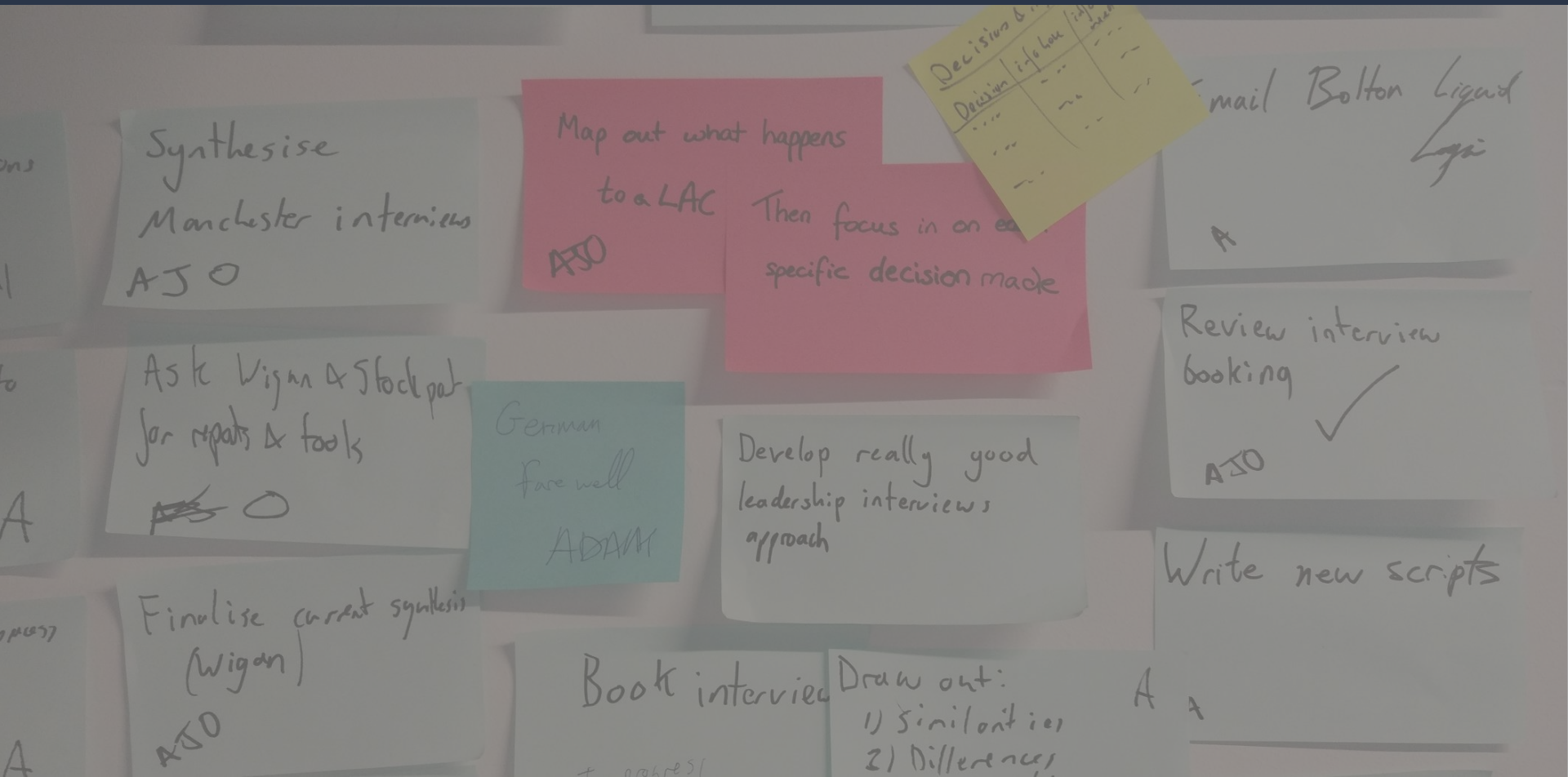
- To update project partners and others interested with weekly progress, learnings and challenges
- To celebrate achievements and milestones

1-2-1 calls

Ad-hoc

- To keep partners across councils, MHCLG and DfE up-to-date
- To share learnings and help steer our work
- To make key decisions on priorities and direction

2. USER RESEARCH REPORT



USER RESEARCH OVERVIEW

WE RESEARCHED WITH 9 DIVERSE COUNCILS, TO ENSURE THAT OUR FINDINGS ARE REPRESENTATIVE

	COUNCIL					CHILDREN SERVICES		
	Type	Popula-tion (est.)	Urban vs. Rural	Political affiliation	IMD rank[1]	# LAC[2]	OFSTED rating	CMS used
Manchester City Council	Met	575K	Urban	Labour	2	1290	Requires Improvement	Liquidlogic
Stockport Metropolitan Borough Council	Met	290K	Urban	NOC[3]	154	361	Good	Liquidlogic
Wigan Council	Met	318K	Urban	Labour	97	448	Good	Liquidlogic
West Berkshire Council	Unitary	158K	Rural	Conservative	289	172	Good	careworks an Advanced company
Isle of Wight Council	Unitary	140K	Rural	Conservative	80	243	Good	Integrated Children's System
East Sussex County Council	County	555K	Rural	Conservative	93	600	Outstanding	Liquidlogic
milton keynes council	Unitary	270K	Urban	NOC[3]	172	381	Requires Improvement	Liquidlogic
Suffolk County Council	County	758K	Rural	Conservative	99	866	Outstanding	Liquidlogic
Buckinghamshire Council	County	542K	Rural	Conservative	145	515	Inadequate	Liquidlogic

1. *The 2019 Index of Multiple Deprivation is a UK government qualitative study of deprived areas in English local councils. The figure above is the average rank per Local Authority District, out of **317**. Source: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019*
2. *Estimates given by analysts in each council.*
3. *No overall control*

USERS AND USER RESEARCH AIMS

	 ANALYSTS	 SOCIAL WORKERS	 LEADERSHIP
Role in the 903 data process	• Extract and clean 903 data, and submit returns to the DfE • Analyse 903 data and create reports	• Record 903 data in case management system	• Use 903 data to inform decision-making
Demo-graphics	• Gender: 7 Males / 5 Females • Age Range: 30-50 • Digital literacy: High	• Gender: 2 Males / 6 Females • Age Range : 25-50 • Digital literacy: Variable	• Gender: 4 Males / 6 Females • Age: 35-50 • Digital literacy: Variable
Alpha aim	1 Test how common the user needs identified in Discovery are across other councils 2 Understand the impact that meeting these needs would have 3 Explore solutions to meet user need 4 Prototype and test solutions to meet user needs	3 Explore solutions to meet user needs	2 Understand the impact that meeting these needs would have

APPROACH: USER RESEARCH METHODOLOGY WITH EACH USER GROUP

	ANALYSTS	SOCIAL WORKERS	LEADERSHIP
Approach	• 1-2-1 semi structured interviews • Written questions through email • Moderated usability testing • Re-analysed interviews from Discovery	• 1-2-1 semi structured interviews, mainly remote	• Re-analysed interview notes from Discovery
Rationale behind this approach	• Interviews relevant to understand broad context, processes and challenges • Remote interviews allowed us to do more interviews in various geographical areas • Emails relevant for follow-up and technical questions	• Interviews relevant to understand broad context and workflow • Remote interviews more appropriate due to SW's limited availability	• Discovery had captured rich insights which addressed our alpha research questions, that we needed to build on to save leadership and team's time
Number	• 15 interviews with 10 people • 52 emails • 9 interview transcripts from Discovery	• 7 interviews with 7 ppl. • 2 interview transcript from Discovery	• 10 interviews transcripts from Discovery

Our user research highlighted the roles of additional user groups including commissioning and IT teams. We therefore carried out a second phase of research to capture their insights, engaging 3 participants through a combination of interviews and follow-up questions by email.

USER RESEARCH DEEP DIVE

1. Test how common the user needs identified in Discovery are across other councils
2. Understand the impact that meeting these needs would have
3. Explore solutions to meet user needs
4. Prototype and test solutions to meet user needs

1 USER RESEARCH GOALS AND APPROACH

We wanted to understand whether the findings in Discovery resonate across local councils and whether poor data quality is indeed a shared problem that should be solved collaboratively. Our specific questions were:

1a | What type of errors do each council have and how do they compare at an aggregate level?

1b | To what extent do analysts in other councils face challenges in identifying and resolving data errors year round?

1c | What other evidence can we draw from our existing network of 30+ councils regarding this issue?

To answer these, we:

 | Interviewed analysts in 3 additional councils beyond the discovery partners to understand their experience of identifying and resolving data errors year round

 | Carried out quantitative analysis to examine the magnitude and type of errors faced by 6 local councils

 | Analysed relevant reports and surveys and gathered knowledge from domain experts across the partners

1a ERROR ANALYSIS FOR EACH COUNCIL

WE FOUND A SIMILAR PATTERN TO DISCOVERY WHERE ~50% OF ERRORS ARE CAUSED BY A FEW ERROR TYPES, BUT WITH A LONG TAIL OF LESS COMMON ERRORS

Source: Manchester, Stockport, Wigan, West Berkshire, Isle of Wight, East Sussex 2017/18 SSDA903 error reports, SF analysis

While quick gains can be derived in each local council from focusing on the few error types, eliminating all data errors is a much more complicated task

1a ERROR ANALYSIS BETWEEN COUNCILS

WE FOUND THAT THE MAJOR ERROR TYPES DIFFER SIGNIFICANTLY ACROSS 6 LOCAL COUNCILS

No.	Error message (truncated)	Manchester	Stockport	Wigan	West Berkshire	East Sussex	IOW	Total
1	One or more data items in first episode do not match open episode at end of last year.	413	111	112	13	0	8	657
2	Distance is not valid. Please check a valid postcode has been entered	206	103	0	56	126	32	523
3	Ofsted URN is required	67	40	180	16	81	64	448
4	More than one review has been held on the same day	0	307	0	0	0	0	307
5	Placement provider code is not a valid code	208	9	0	0	2	86	305
6	A Strengths and Difficulties Questionnaire score completed	3	2	22	0	230	27	284
7	Local Authority's (LA) of placement is not valid or is missing.	175	36	0	37	2	24	274
8	One or more data items relating to CLA for 12 months have been left blank	6	6	0	111	7	38	168
9	Episodes need to be loaded for this child before further validation is possible	1	0	13	81	16	31	142
10	There is a missing continuous episode in the previous year	41	51	15	4	23	4	138

Source: Manchester, Stockport, Wigan, West Berkshire, Isle of Wight, East Sussex 2017/18 SSDA903 error reports, SF analysis

This heatmap of the 10 most common errors, where more darker green indicates more occurrences, shows that errors differ significantly between councils

Given the disparity in error types across local councils, any scalable solution will have to focus on errors in general, rather than specific error types

1b OUR INTERVIEWS WITH NEW COUNCILS[1] ECHO THE FINDINGS MADE DURING DISCOVERY

Learning	Quotes
Analysts can only identify errors effectively during a limited window when the SSDA903 portal is open	*"We don't clean data year round because it's really difficult to identify errors in the first place"* *"There's nothing easy to replicate the error messages in the 903 portal that we only have access to during certain times of the year"*
Analysts find the process of completing the SSDA903 return time consuming and intensive	*"People were doing stupid hours – 10 to 12 hours plus weekends and there is no work life balance at all.* *"Last year's [SSDA 903 return] was an absolute nightmare, with the platform change and all that.."*
Analysts have limited technical ability and resources to identify errors year round	*"We tried to write some validation rules for simple errors using Business Objects, but we are still a long way from covering those written by DfE."* *"Creating data quality report is an area that we are lacking. It's a strain on our time and resources."*

All 6 local councils struggle to keep their data on looked-after children in good quality year round due to limited resources and technical ability

1. The new councils to the partnership we tested with are Isle of Wight, West Berkshire and East Sussex

1c WIDER EVIDENCE SHOWS THE LEADERSHIP NEED FOR BETTER DATA QUALITY IS COMMON…

- From our extensive networks across councils (e.g. 35+ councils across the North West and South East Regional Information Groups, 30+ councils Social Finance have partnered with) we see that the quality of data on children in care is a significant need for all

- 38% of children's services leaders identified data quality and usefulness as a top barrier for improving services with data and digital in a NLGN and Social Finance survey of 61 senior children's services leaders across the country

- This has also been widely highlighted in national research:

  "However, the key message from most councils was a bigger challenge around the importance of good quality data." - Local government social care data standards and interoperability discovery report, Local Government Association

1c …AND THAT THE DATA QUALITY NEEDS OF ANALYSTS ARE COMMON ACROSS COUNCILS

Common analyst needs

- Networks of analysts, such as the Children's Social Care Data google group have hundreds of members, and thousands of conversation threads. These largely focus on statutory returns and on handling data quality issues
- Analysts' pain points have been widely highlighted in national research:

  "The process of preparation of data for submission to the DfE has been reported to be onerous and time consuming for most local authorities, with the process of preparing data taking up to three months" - Use of children's social care data at the local and regional area level, Dr Lisa Holmes, Nuffield Family Justice Observatory

Common process – Due to statutory requirements, every council submits exactly the same data in the same way. Statutory requirements drive much of children's services and have been the driver for all previous common tools e.g. the CHAT, the only children's services data tool to successfully scale across councils (used by 151 of 152), helps councils respond to the statutory requirements around Ofsted

Common systems – 90% of local authorities use one of three case management systems (CMS), and all CMS are based around the same "ICS" core, meaning similar issues are experienced

1 CONCLUSION

Our findings from user research suggests that poor quality of children in care data is a significant problem across local councils

1a

What type of errors do each local council have and how do they compare at an aggregate level?

→ There are many types of errors that cover each local council's children in care data and significant differences between them, meaning that to create a common solution, we need can't just focus on specific error types

1b

Do analysts in other local councils face challenges in identifying and resolving data errors year round?

→ Yes, our interviews with analysts in other local councils reveal significant challenges in identifying errors year round due to resource and capability constraints

1c

What other evidence can we draw from our existing network of 30+ local councils regarding this issue?

→ Findings from other surveys, reports and forum corroborate the view that quality children in care data is a high priority need for all

1. Test how common the user needs identified in Discovery are across other councils
2. Understand the impact that meeting these needs would have
3. Explore solutions to meet user needs
4. Prototype and test solutions to meet user needs

2 USER RESEARCH GOALS AND APPROACH

To get a sense of the impact that meeting these user needs would have, we needed to assess the impact data quality has on analysis and decision-making. Therefore, we needed to understand:

2a | How important the 903 return data is for analysis and decision-making – both internally and externally to the council?

2b | How current errors impact analysis?

To answer these, we:

 | Interviewed analysts to understand what they think about the impact that each specific error has on data analysis and decision-making

 | Analysed user interview transcripts with analysts (9) and leadership (10), to understand the relationship between data quality, errors and decision-making

2a HOW IMPORTANT IS THE 903 DATA FOR ANALYSIS/DECISION-MAKING?

THE 903 RETURN COVERS THE MAJORITY OF DATA THAT LEADERSHIP HAVE AVAILABLE FOR ANALYSIS ON CHILDREN IN CARE

LEADERSHIP

The data contained in the 903 return forms the basis of most internal reporting used by leadership, as it is the main dataset available

- **Reporting:** Internal quarterly or annual performance reports about children in care mostly build on data from the 903 return
- **Ad-hoc analysis**: Analysts also use this data to answer leadership's ad-hoc requests
- **Dashboards**: Where councils have self-service dashboards on children in care (e.g. Stockport's Tableau dashboards) these are also largely based on the data that feeds the 903 return

Leadership find the 903 data useful, despite several limitations

- **Useful:** *"All that data is needed" / "I actually like the 903 return data you know, as raw data I like it"*
- **Despite limitations:** The 903 data is too process-oriented and doesn't tell much about the quality of practice - for which leadership rely on qualitative audits and feedback. They also need additional strategic data e.g. evolution in needs, referrals' quality or outcomes.

2a HOW IMPORTANT IS THE 903 DATA FOR ANALYSIS/DECISION-MAKING?

FROM OUR WIDER RESEARCH, WE KNOW IT IS ALSO USED BY STAKEHOLDERS FOR VARIOUS PURPOSES, AS THE ONLY NATIONAL DATASET ON CHILDREN IN CARE

Benchmarking and rating

- Many Councils and Regional Benchmarking Groups use it for benchmarking: *"I can't think of an example recently where [leadership] wanted to measure something different to a national indicator" [based on the 903 return dataset]* (Analyst)
- OFSTED rely on this data during their inspections, to assess and rate children services departments

Forecasting

- *Example*: Suffolk and Cambridgeshire councils are using it in a collaborative project to forecast the future number of children in care, so that they can effectively plan placement requirements

Informing *policy makers*

- 903 return data contributes to the DfE LAIT tool and the DfE's National Pupil Database which are used by DfE, academics, charities and local authorities to inform policy

Evidencing *effectiveness & value for money*

- *Example*: Essex council and the Greater London Authority used the 903 returns to make the investment case for Multi-Systemic Therapy, enabling large scale investment in these services which resulted in 350+ children staying out of care

2b HOW DO ERRORS IN THE 903 DATA IMPACT ANALYSIS?

WHILST NOT ALL 903 DATA ERRORS MATERIALLY AFFECT ANALYSIS, ACCORDING TO ANALYSTS, MORE THAN HALF OF TOP ERRORS DO HAVE MEDIUM TO HIGH IMPACT ON DATA ANALYSIS

	HIGH IMPACT		MEDIUM IMPACT	LOWER IMPACT
	31% of errors		**23% of errors**	**46% of errors**
Error examples	**Distance between home and placement is not valid.**	**Child started to be looked after […] but no review was recorded within that time**	**Episode commenced before the start of the current collection year but there is a missing continuous episode in the previous year**	**Date of birth of child's mother is not a valid date**
Why is it important for leadership?	Councils generally aim to place children as close as possible to home, to help maintain relationships. Long distance placements can be a major concern	Councils are statutorily required to perform reviews on time, so leadership need to ensure they are doing this	It helps to get a longitudinal view of a child's journey. It is medium in impact, as if it is due to late recording, it would have little impact on analysis.	The important information for leadership is the % mothers under 18 – which can be already captured through motherhood status
Quotes from analysts	*"Distance between home and placement is an important performance indicator"*	*"This is very important because the review time scale is one of the big performance information for all authorities"*	*"It depends if it is caused by inconsistency or late data entry. If it is the latter, then it bears no impact on analysis as the data in our CMS is accurate even though it doesn't match DfE's system"*	*"I don't think knowing the date of birth specifically is useful for analysis"*

2b HOW DO ERRORS IN THE 903 DATA IMPACT ANALYSIS?

THESE ERRORS CONTRIBUTE TO LEADERSHIP LACKING TRUST IN DATA

We found a widespread distrust in data accuracy across leadership, disincentivising them from using data analysis to inform the decisions they take on services for children in care. Service decisions are therefore more likely to be based on anecdotal evidence.

Quotes from Leadership

"Do you think the data is accurate? It is variable (laughs)"

"It is written here that there are 3 children in residential care, I can guarantee all 3 are incorrect"

"I suppose any data on LAC is really important (…) but then we need to look at the quality of that data"

"What info do you need (…)? Well, I suppose there's the basic stuff of being able to trust the information"

Note: These comments are extracts from the Discovery interview transcripts, which did not focus specifically on data quality. Most of them were unprompted, which reinforce their value and suggests that data quality is an important concern for leadership.

2 CONCLUSION

Our findings from user research suggests that improving the 903 return data quality will lead to leadership further using data to inform their decisions

2a

How important is the 903 return data for analysis and decision-making?

→ It is the main dataset used by leadership for performance management

→ It is also used at national and regional level to improve practices on children in care through benchmarking, rating, forecasting, etc.

2b

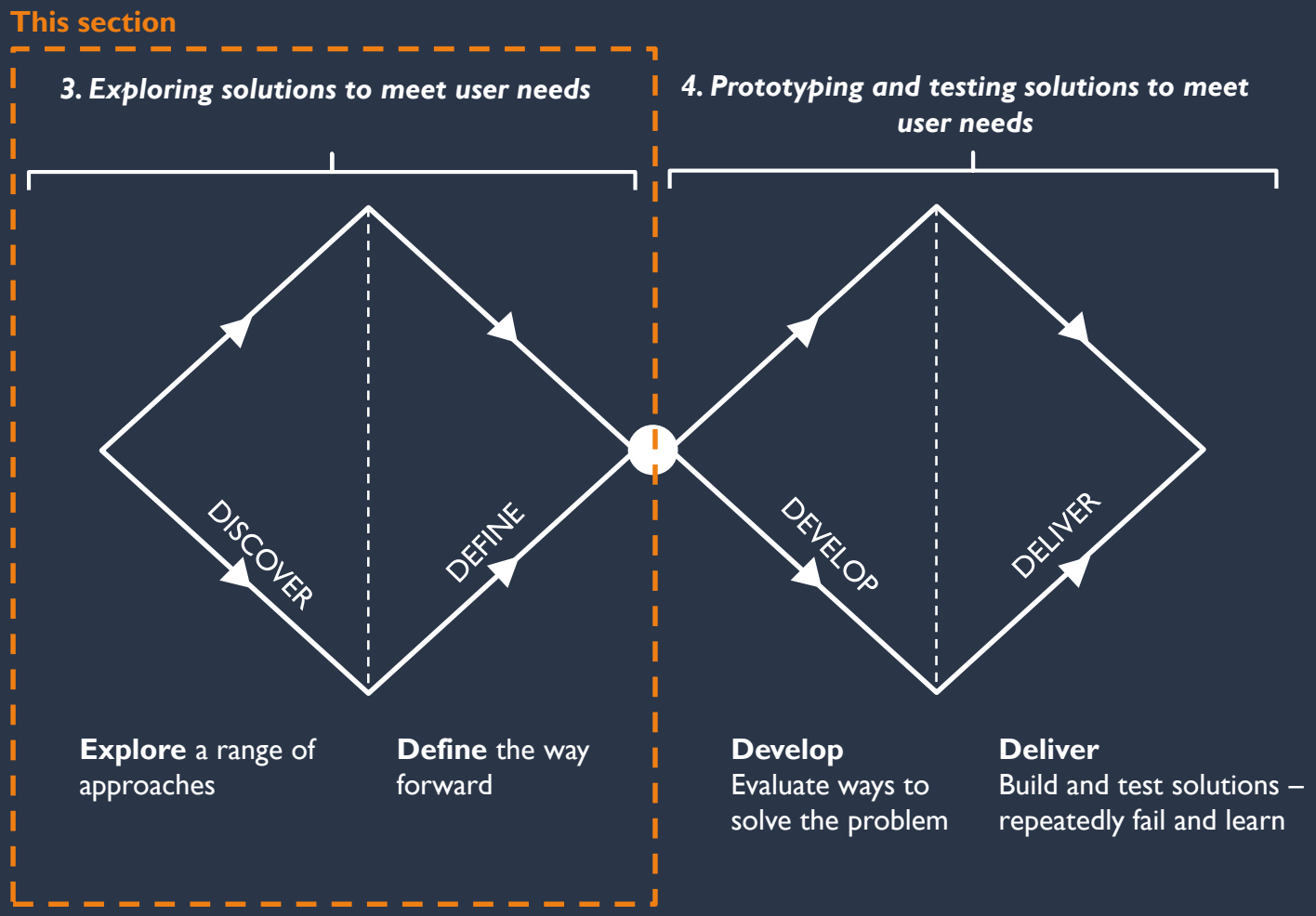
How do errors impact analysis?

→ Whilst not all errors have a material impact on analysis, more than half of them do have medium or high impact

→ Errors in the 903 return reduce leadership's trust in data and their confidence in using data to inform decisions

1. Test how common the user needs identified in Discovery are across other councils
2. Understand the impact that meeting these needs would have
3. Explore solutions to meet user needs
4. Prototype and test solutions to meet user needs

WE USED THE DOUBLE DIAMOND APPROACH TO EXPLORE POTENTIAL SOLUTIONS



3 OUR APPROACH

To improve data quality, solutions must either help avoid data errors, or improve error cleaning. We explored both of these options:

To answer these, we did the following user research:

3a | Can we help avoid data errors?

- Interviews with 3 analysts, 7 social workers, 2 commissioning team members & 1 system's team member
- Written questions to analysts (5 emails)

3b | Can we help improve error cleaning?

- Interviews with 8 analysts
- Written questions to analysts (20 emails)

3a OUR APPROACH TO: "CAN WE HELP AVOID DATA ERRORS?"

Exploring whether we can help avoid errors was a twofold process:

i. Understand the data recording problems

- **What** errors occur?
- **Why** do they occur?

ii. Brainstorm and assess how to tackle these problems

- **What options** could help avoid data errors?
- **What is the potential** for each of these options?

3a WHAT ERRORS OCCUR?

THERE ARE THREE MAIN GROUPS OF DATA ERRORS

	PLACEMENT ERRORS	YEAR-TO-YEAR INCONSISTENCIES	LOCAL AUTHORITY-SPECIFIC ERRORS
% total errors	37%	20%	43%
Definition	These are errors about the placement of a child looked-after (mainly about the provider and location)	When some information in this year's 903 return does not match with information in last year's return	These are errors that happen mainly in just one or two councils, due to local processes or challenges
Error examples	• Distance between home postcode is missing • Placement postcode is not valid • OFSTED URN is required • Placement provider code is not a valid code • LA of placement is not valid or missing	• One or more data item in the first episode do not match open episode at end of last year • There is a missing continuous episode in the previous year	• More than one review has been held on the same day • Children [meeting a set of criteria] should have a SDQ score completed
Impact on analysis	30% of those have high to medium impact	45% of those have high to medium impact	57% of those have high to medium impact

3a WHY DO THESE ERRORS OCCUR?

THERE ARE SPECIFIC DRIVERS FOR EACH OF THESE MAIN ERROR TYPES

PLACEMENT ERRORS

There are three main causes:

- **Diffuse responsibility:** complex and sequential workflow, involving up to 5 teams in some councils, depending on the placement type
- **Duplication:** In some councils, the same information has to be recorded by 2 people
- **Access to data**: It can be difficult to get information like placement address in some cases e.g. for children's homes.

YEAR-TO-YEAR INCONSISTENCIES

There are two main causes:

- **Error cascades** – missing data elsewhere can cause year-to-year inconsistencies
- **Late-recording** – if events occurring at the end of last year's return window aren't recorded in time then they aren't included in last years return, but will be in this year's

LOCAL AUTHORITY-SPECIFIC ERRORS

The reasons why these errors occur vary significantly depending on each error. Below are two examples:

- **Lack of validation**: The error "*More than one review has been held on the same day*" used to occur in one council, due to a lack of validation, a problem that is resolved now
- **Mistake in the upload process**: The error "*Strength and Difficulties Questionnaire score completed is missing*" was frequent in one council due to a temporary mistake in the upload process

3a WHY DO THESE ERRORS OCCUR?

INTERVIEWS WITH SOCIAL WORKERS SHOWED THAT, BEYOND IMMEDIATE CAUSES, OTHER FACTORS CONTRIBUTE TO LOWER OVERALL DATA QUALITY

Heavy workload and reduced support

- ***Heavy workload**: "We're always busy on the go, it is not always a priority"*
- ***Reduced support**: "We don't have a lot of admin support anymore"*

Non-optimal tech resources, making it complicated to record "on the go"

- ***Laptops' efficiency**: The answer would be giving us decent laptops. They gave us ours 5-6 years ago […] They're quite big and heavy. It takes 20 minutes to turn them on."*
- ***Access to the system:** "We don't have remote access to LiquidLogic but we're often out of the office with families"*

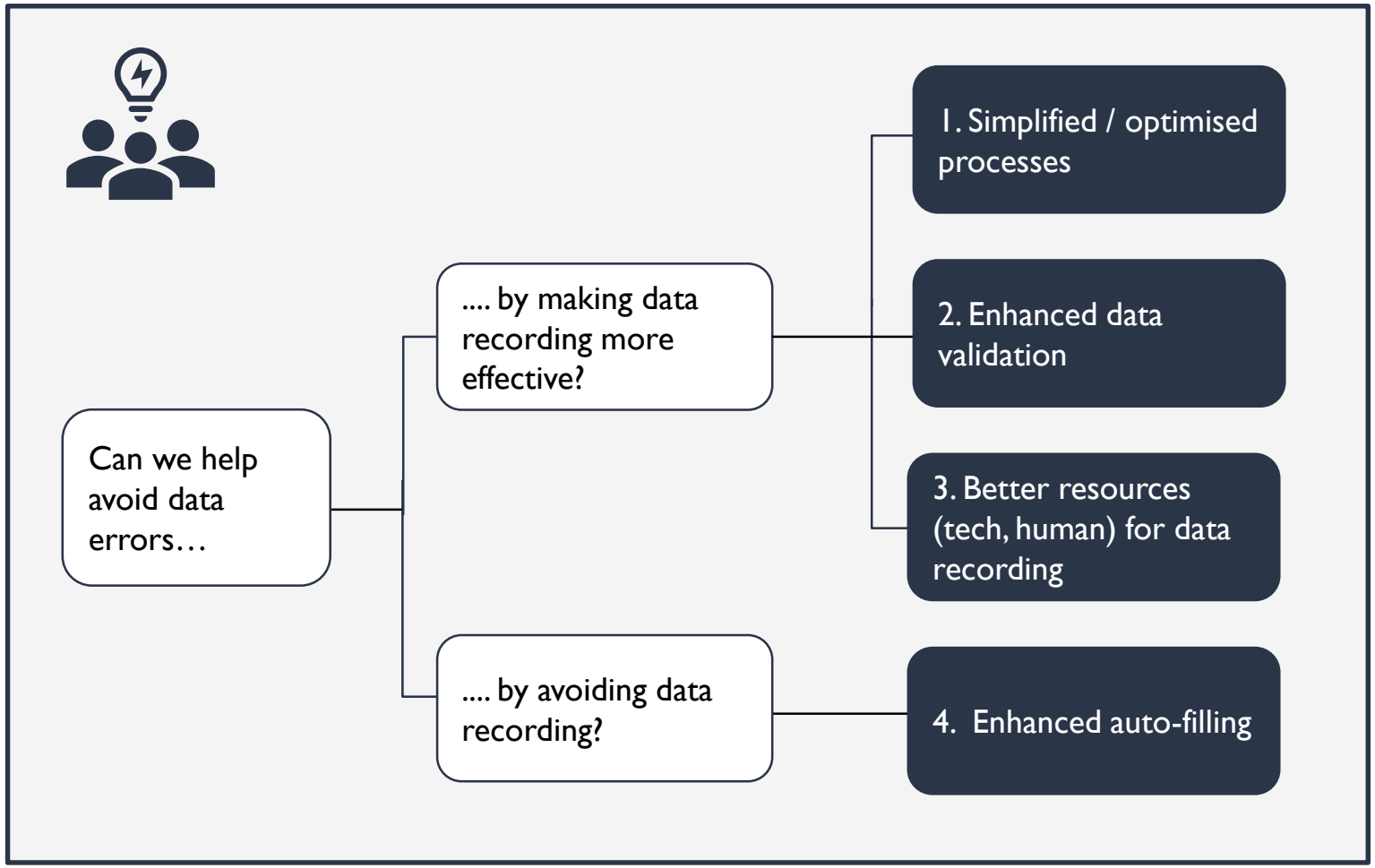
Frustration associated with data recording

- ***Frustrating conditions:** "Sometimes the whole process, it can be a real frustration" [due to the amount of information to collect and the lack of resources to do so]*
- ***Non-fulfilling activity:** "Recording reduces SWs' sense of autonomy and fulfilment. Just churning things through a machine"*

Limited benefits of 903 good data quality for SWs

- ***Limited benefits**: "Having drop downs is useful to a certain extent, but the most important is the narrative."*
  "Conversations are quicker and easier. Let's stop thinking we should digitise everything"

3a BUILDING ON THESE FINDINGS, WE BRAINSTORMED HOW THESE ERRORS COULD BE AVOIDED



3a HOW COULD THESE 4 OPTIONS HELP AVOID DATA ERRORS?

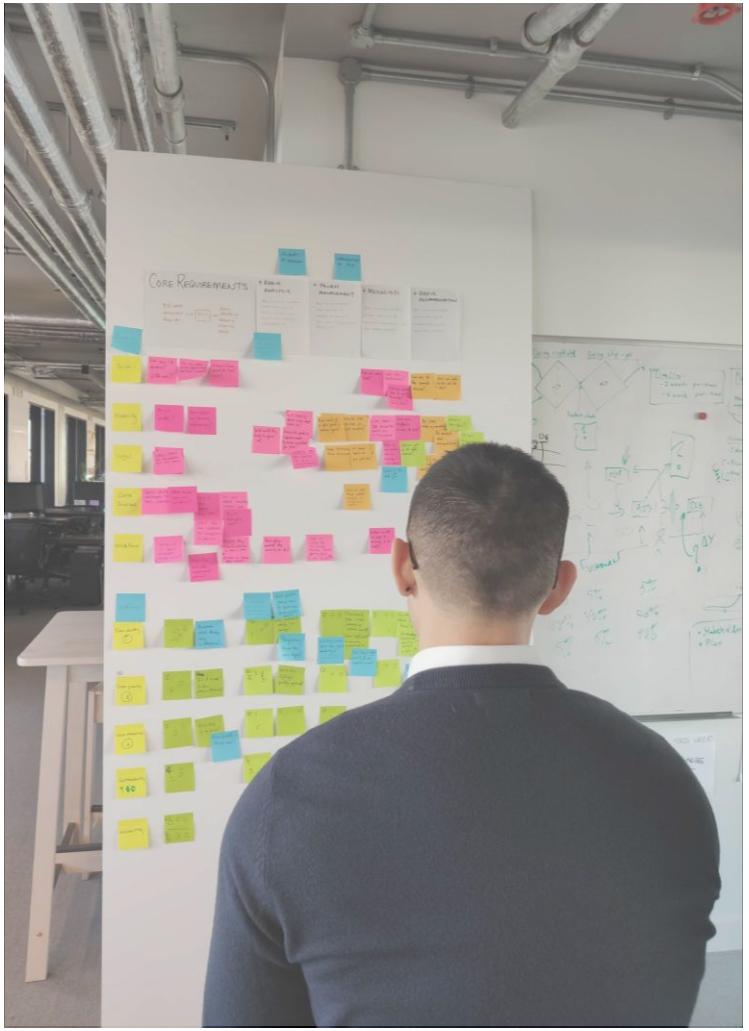
1. Simplified / optimised processes	• **Learning:** User research on placement errors suggests that councils having less errors have easier processes. We assume that these findings could apply to other error types. • **Opportunity:** We assume that other councils could adopt similar processes or learn from it to reduce the number of errors.
2. Enhanced data validation	• **Learning:** Most errors can be technically validated, as they are machine-readable (they are either: missing information, invalid format or incoherent logic) • **Learning**: Currently, the use of validation varies across councils. • **Opportunity:** There is potential to do more. The use of validation for "missing information" is debatable though, as it may not be beneficial to impede staff recording data from progressing if they don't have part of the information.
3. Better resources (tech, human) for data recording	• **Learning:** The conditions for data recording increase the risk of delays and errors. These findings are relatively consistent across councils. • **Opportunity**: Resources have been decreasing over the past decades, there is definitely scope for improvement here. With more time and better conditions for people to record information, we would expect mainly an impact on missing data.
4. Enhanced auto-filling	• **Learning:** Currently, the use of auto-filling varies across councils. • **Opportunity**: There is scope to do more. For example, in one council, SWs can update a placement address by searching through a list, but they have to manually re-enter 4 datafields. Having auto-filled information would be an efficient way of avoiding duplication and manual entry.

3a WE ASSESSED THE POTENTIAL OF THESE 4 OPTIONS ON IMPACT, FEASIBILITY, COST AND COMMONALITY

	Impact	*Feasibility*	*Cost*	*Commonality*
1. Simplified / optimised processes	✓ Less data interdependency, increased ownership, leading to less incorrect + missing data	✓ Changing data recording process on a few error types in one council seems easily feasible	✓ Medium costs on human resources mainly around change management	✗ Data recording processes are highly council-specific
2. Enhanced data validation	✓ Less incorrect data	✓ Applied to one council on a few error types, it is quite straightforward	✓ Low costs on human resources (IT team mainly)	✗ Systems, workflows, and current use of data validation are highly council-specific
3. Better resources for data recording	✓ Massive impact going much beyond data quality: SW feel valued and less frustrated, gain time. Less missing data.	✓ With more resources, it would be possible to improve the process for social workers	✗ Implies significant investment in resources e.g. hiring, technology purchases	✗ Resources are council-specific, better resources at scale would require significant political buy-in
4. Enhanced auto-filling	✓ Less incorrect + missing data, as well as a slight gain in time for workers	✓ Applied to one council on a few error types, it is quite straightforward	✓ Low costs on human resources (IT team mainly)	✗ Systems, workflows, and use of auto-filling are highly council-specific

3a CONCLUSION: CAN WE HELP AVOID DATA ERRORS?

- We have identified four options to help avoid data errors:
  1. Simpler processes
  2. Data validation
  3. Better resources
  4. Auto-filling
- Options 1, 2 and 4 have the most potential, but **cannot be rolled out nationally** given variation in data recording processes.
- However, based on our learnings from user research, we could start a **common guidance document** sharing data recording best practices and recommendations, focusing on the 903 top errors



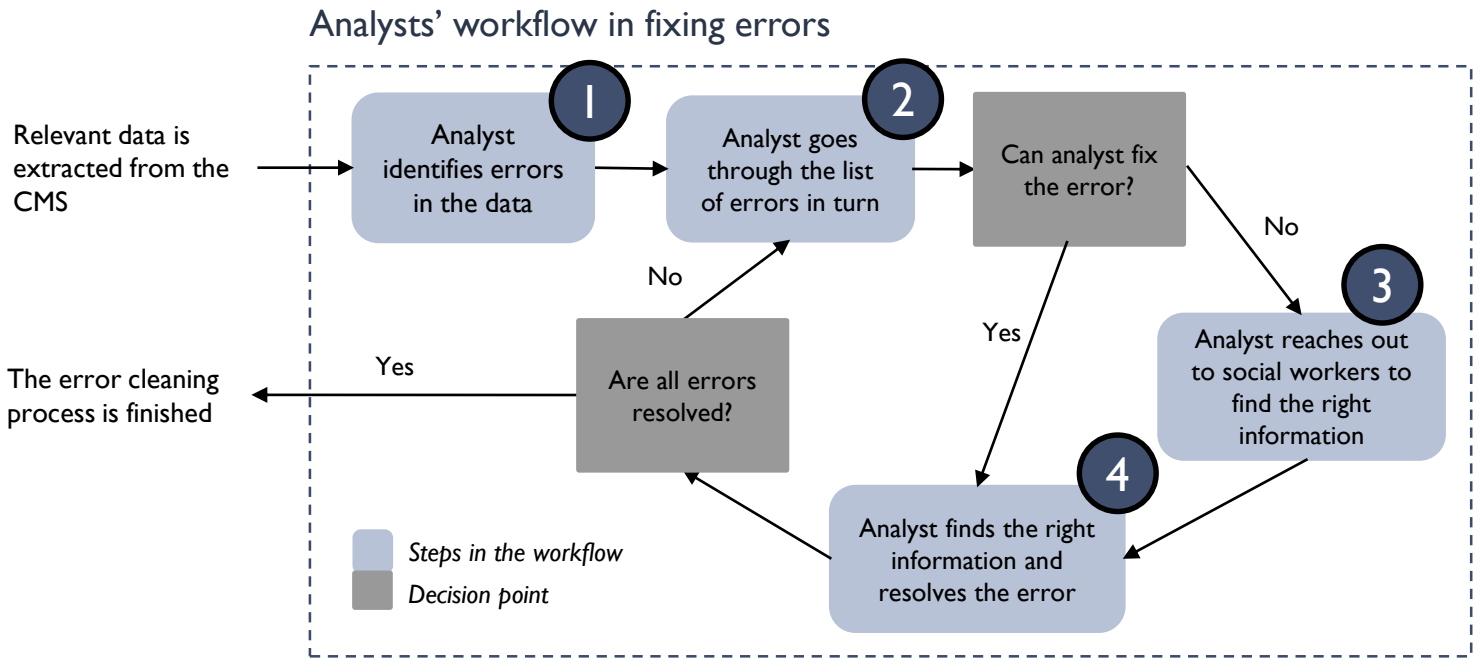
NEXT WE EXPLORED IF WE COULD HELP IMPROVE CLEANING

3a | Can we help avoid data errors?

3b | **Can we help improve error cleaning?**

3b OUR APPROACH TO: "CAN WE HELP IMPROVE ERROR CLEANING?"

WE INVESTIGATED POTENTIAL SOLUTIONS TO IMPROVE EACH STEP OF THE ERROR CLEANING PROCESS.

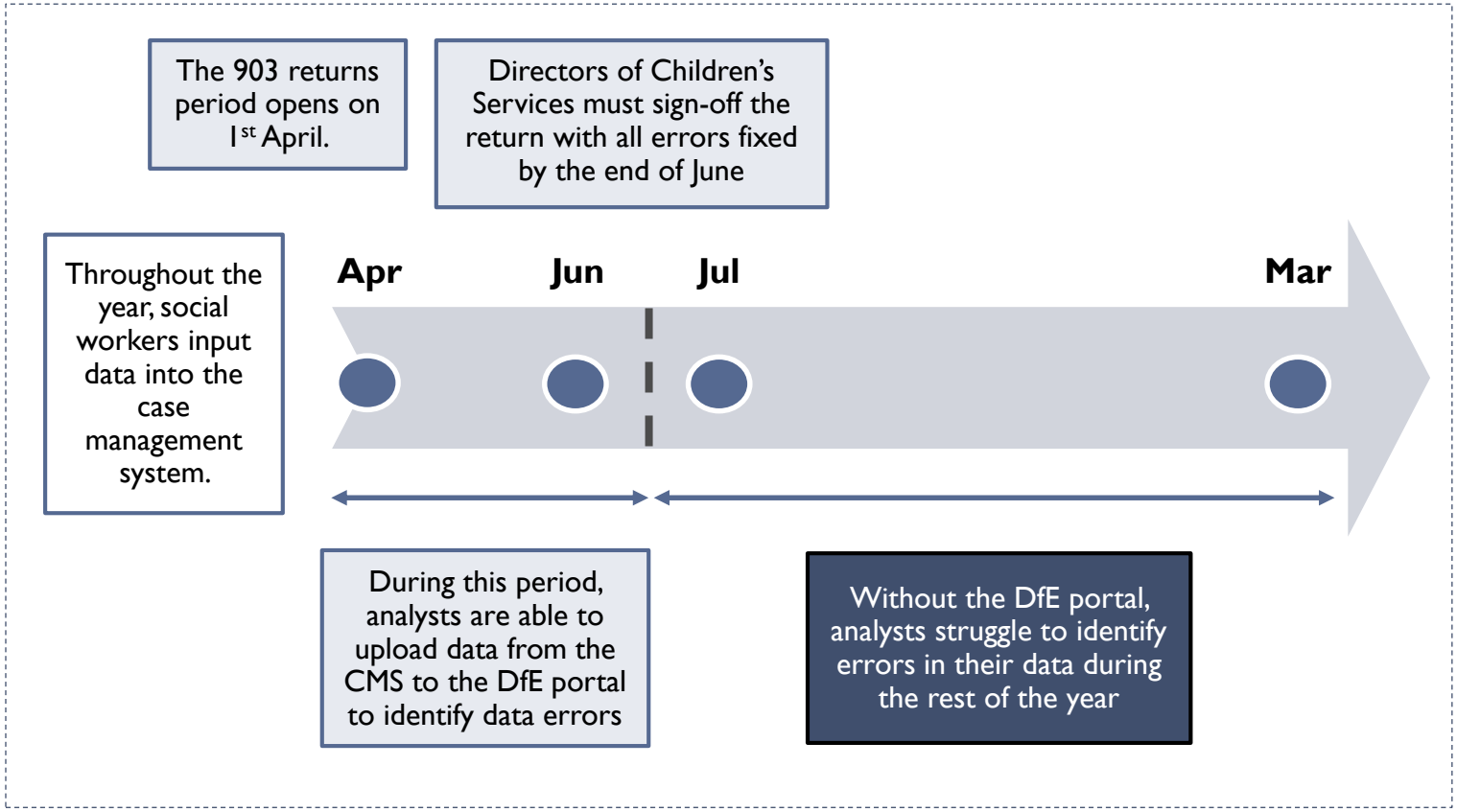


1	2	3	4
Can we improve error identification?	Can we improve process management?	Can we improve communication?	Can we improve information finding?

3b 1. SHOULD WE IMPROVE ERROR IDENTIFICATION?

Analysts across local councils **do not** currently have a comprehensive way of identifying data errors when the DfE portal is not accessible

The SSDA903 statutory return timeline:



3b 1. SHOULD WE IMPROVE ERROR IDENTIFICATION?

Analysts in a few local councils have attempted to **create their own error identification tools**, but have had **limited success** due to resource and technical constraints

Time constraint

- *"Creating data quality reports is an area that we are lacking. It's a strain on our time and resources."*
- *"There are some things that we have to improve, but where on the list does that come?"*

Technical barrier

- *"We tried to write some validation rules for simple errors using Business Objects, but we are still a long way from covering those written by DfE."*
- *"The reports we have created are separate and stand-alone reports. The DfE platform aggregates everything together and captures more."*

Inadequate capability

- *"We do what we can. We are not IT basically. There's limitations to our ability to write these error reports. There have been errors where we have attempted to write error reports for, we haven't been able to one reason or another."*

Yes, there is a strong and common need for a way to identify errors year round

3b 2. SHOULD WE IMPROVE PROCESS MANAGEMENT?

Analysts have developed a **structured way of resolving data errors** once they have been identified

Step 1: Analysts look at the overall picture and focus on errors that can be cleaned en masse

"We would see if we can amend something in the export out of our system e.g. is there a logical error that's easy to fix"

Step 2: Analysts then fix errors that are easy and make sense. These are usually the common errors.

"We start cleaning with the easiest and progress from there"

"We start with trivial thing, such as dates aren't coherent or missing information"

Step 3: Analysts finally resolve a long tail of more complicated errors.

"It's the long tail where there's only a few that are slow. These are the ones where the error messages are very misleading"

"The tail is harder to fix"

No, analysts have established ways of managing the error cleaning process and don't see an issue here

3b 3. SHOULD WE IMPROVE COMMUNICATIONS?

Analysts clean most errors by relying on supporting evidence in the case management system or other databases and only occasionally need to contact social workers for additional information

ANALYST

In some cases, analysts do reach out to social workers for information required to resolve data errors.

- *"I sometimes have to contact the social worker to find out more information"*
- *"Some errors are more complicated, requires more digging and a longer conversation with SWs"*

However, analysts and social workers believe that it is often not a problem…

- *"Don't think communicating with social workers is the issue. It's more troublesome to find the right information"*
- *"If things are missing, we send out a spreadsheet to another team and get them to fill in the missing cells. They usually respond quite quickly"*

… and they can mostly find evidence needed in the CMS or other databases

- *"I carry out a lot of the actual data cleansing for the Stat Return myself from the evidence in ICS and only occasionally need to contact the social worker"*
- *"I tend to have the information available in a different format, either on an excel spreadsheet or I'll find the information myself on our CMS"*

No, communication with social workers is a not a major common issue across councils

3b 4. SHOULD WE IMPROVE INFORMATION FINDING?

Analysts look for information in different places depending on the specific errors that they are trying to resolve. However, where the source of information is centrally maintained and updated, the process of information finding could be automated to save analysts time.

Analyst

In most cases, information needed can be found without having to consult social workers

> *"I tend to have the information available in a different format, e.g. on a spreadsheet or our CMS system."*

	Where do analysts look for info?	Potential for common improvements?
1	**In the CMS** – often information is recorded in the wrong place or in unstructured case notes	No – there is potential for improvement but a common solution isn't feasible. The process of recording data and compiling information differs significantly across local councils, so any solution would need to be highly locally customised
2	**Other databases** within the local council, such as health records	
3	**Ofsted's database,** which contains placement information such as URN, placement postcode, placement provider etc.	Yes – If either placement postcode or URN is known, analysts can use this to resolve placement related errors, which account for almost 40% of total errors.

Yes, we could improve the process by automatically searching Ofsted data and suggesting corrections for placement errors

3b CONCLUSION: CAN WE HELP CLEAN DATA ERRORS?

	Each step within the error fixing process		Potential for improvement?
1	Can we improve error identification?		Yes, there is a strong and common need for a way to identify errors year round
2	Can we improve process management?		No, analysts have established ways of managing the error cleaning process and don't see an issue here
3	Can we improve communication?		No, messaging social workers is a not a major common issue across councils
4	Can we improve information finding?		Yes, we could improve the process by automatically searching Ofsted data and suggesting corrections for placement errors

3 SOLUTION EXPLORATION CONCLUSION (1/2)

We've explored a broad range of approaches to improving data quality which aim to either:

a) Avoid data errors occurring
b) Improve cleaning of errors

We tested each of these potential approaches through user research and analysis to assess if they are:

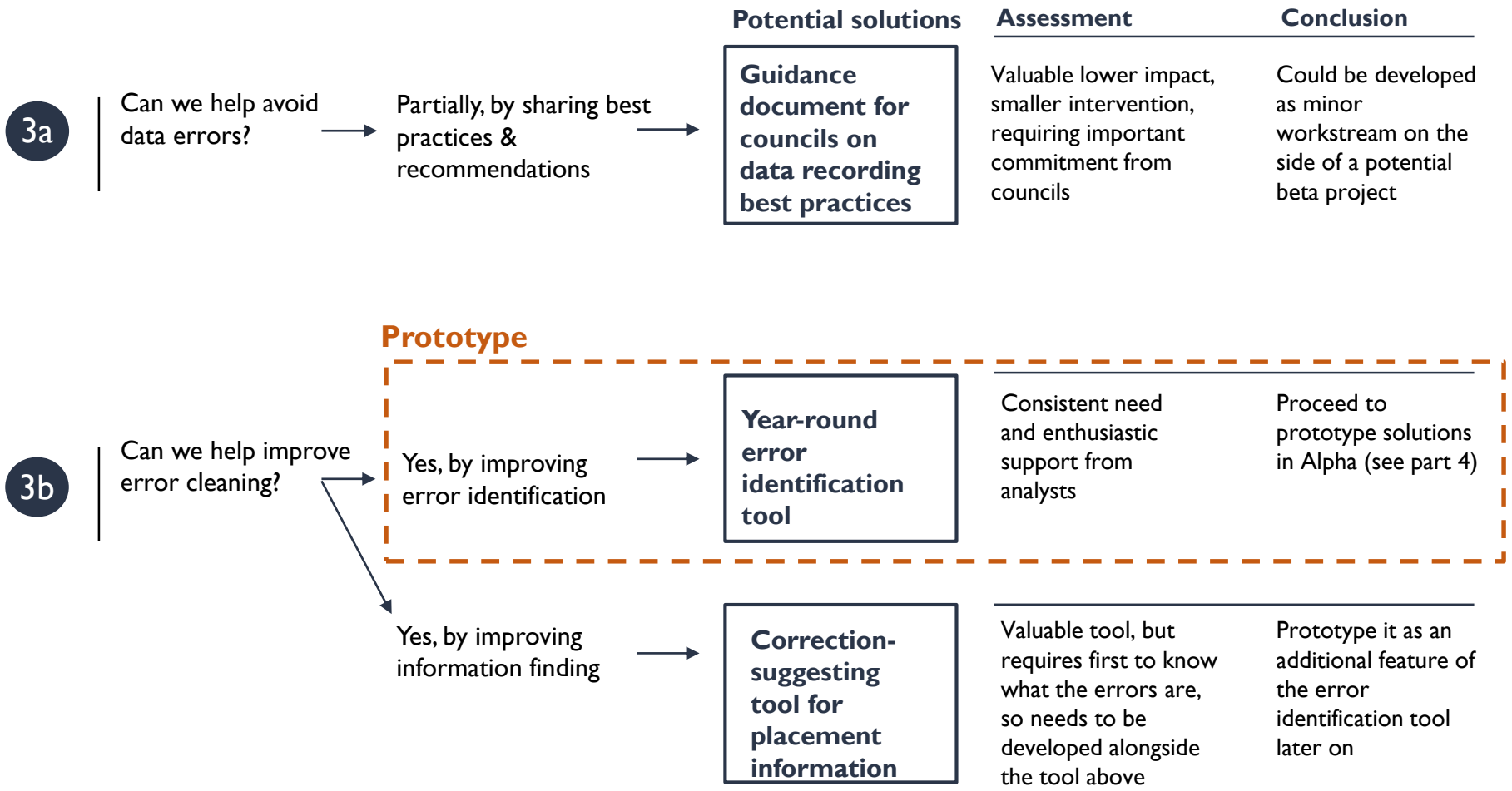
1. Impactful
2. Feasible
3. Cost-effective
4. and if a Common solution can be developed across councils

Three of the ideas developed passed these filters:

1. Helping analysts identify errors year round
2. Automatically identifying information
3. Developing a shared guidance on data recording process

3 SOLUTION EXPLORATION CONCLUSION (2/2)

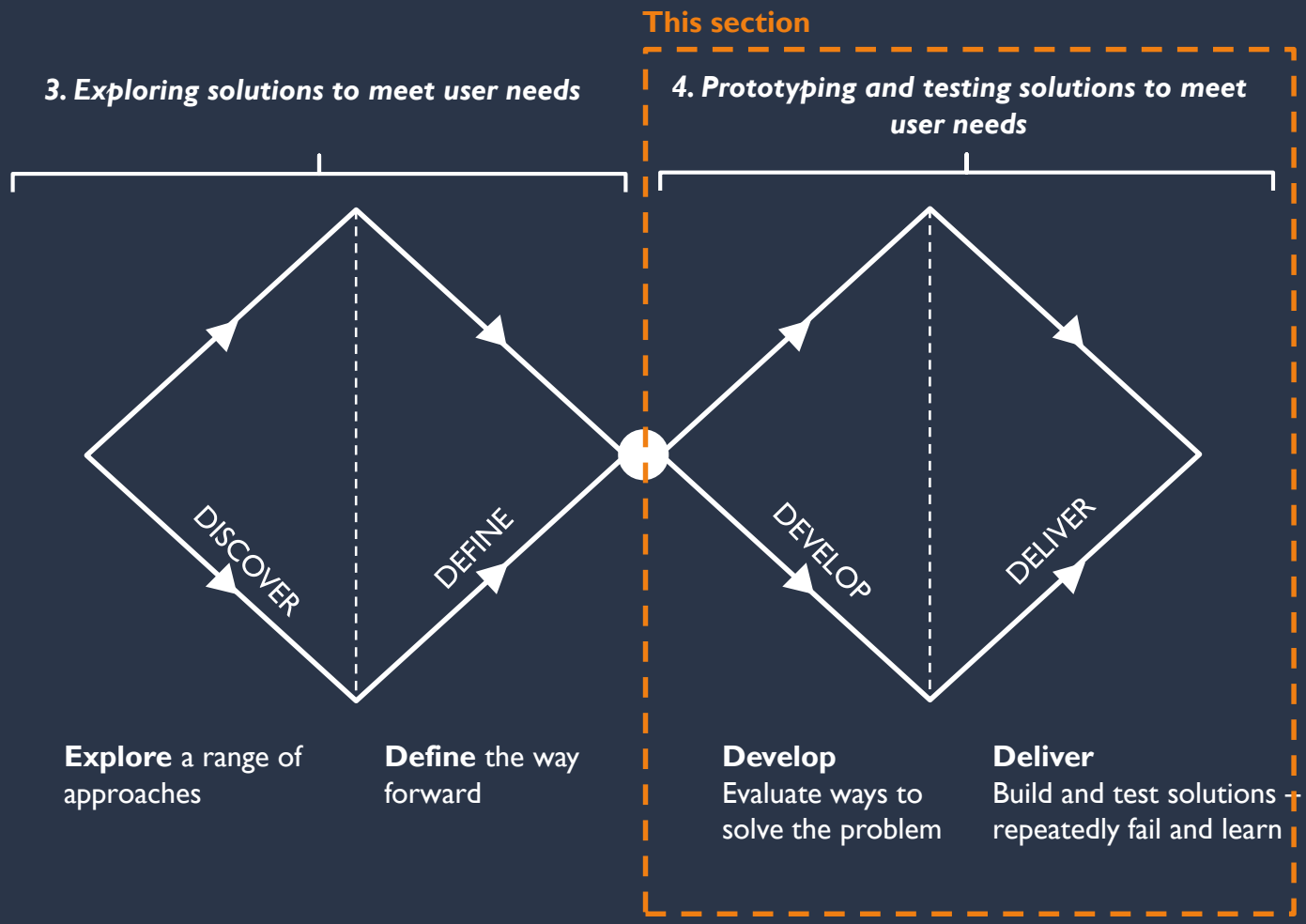
In our Show and Tells with the council partners, MHCLG, DfE and Social Finance we assessed the merits of the three potential approaches and concluded:



We should prototype the year-round identification tool as part of this alpha project

1. Test how common the user needs identified in Discovery are across other councils
2. Understand the impact that meeting these needs would have
3. Explore solutions to meet user needs
4. Prototype and test solutions to meet user needs

WE USED THE DOUBLE DIAMOND APPROACH TO EXLPORE POTENTIAL SOLUTIONS



WE PROCEEDED TO DEVELOP A PROTOTYPE THAT COULD HELP ANALYSTS IDENTIFY AND RESOLVE DATA ERRORS

As an analyst cleaning data,

I need: to identify errors and effectively resolve them throughout the year

So I can: keep the quality of data on children in care high for analysis

In order to meet this user need, we need to test:

4a. Application functionality and design

What are the functionalities and design that analysts would find useful in an application to identify errors year round and effectively resolve them?

4b. Technical set up

What technical set up would enable the application to be most rapidly implemented and scaled across local councils?

OUR APPROACH TO TESTING APPLICATION FUNCTIONALITY, DESIGN AND TECHNICAL SET UP

What we need to test	Our approach in testing
4a. Application functionality and design What are the functionalities and design that analysts would find useful in an application to identify errors year round and effectively resolve them?	We produced designs using the prototyping tool Figma and iterated the features of the applications with analysts through remote moderated usability tests
4b. Technical set up What technical set up would enable the application to be most rapidly implemented and scaled across local councils?	We built a working prototype and shared it with analysts to test if they are able to use it successfully in their local environment

OUR APPROACH IN TESTING APPLICATION FUNCTIONALITY AND DESIGN AND TECHNICAL SET UP

What we need to test	Our approach in testing
4a. Application functionality and design What are the functionalities and design that analysts would find useful in an application to identify errors year round and effectively resolve them?	We produced designs using Figma and iterated the features of the applications with analysts through remote moderated usability tests
4b. Technical set up What technical set up would enable the application to be most rapidly implemented and scaled across local councils?	We built a working prototype and shared it with analysts to test if they are able to use it successfully in their local environment

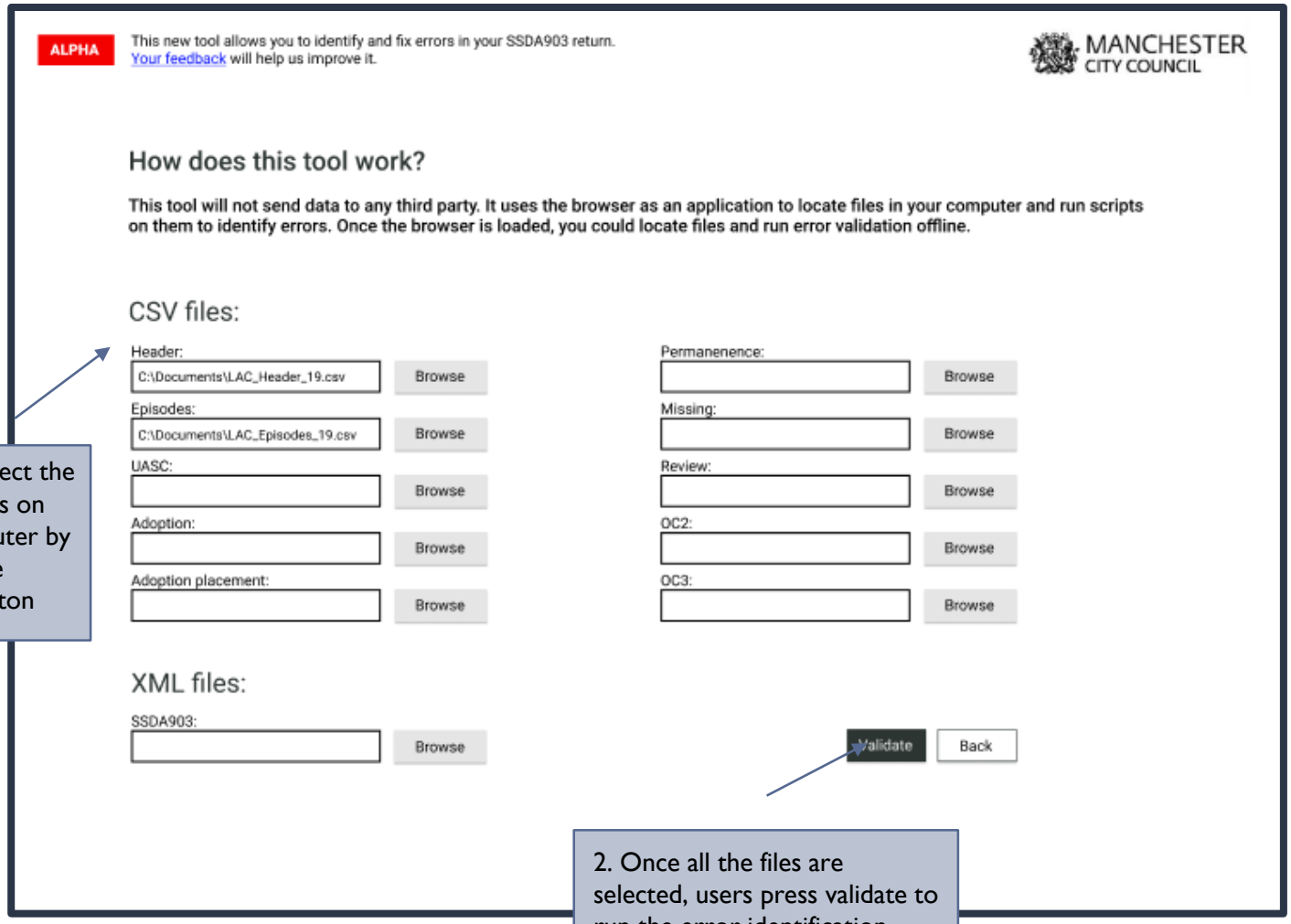
OVERVIEW ON TESTING PROTOTYPE FUNCTIONALITIES AND DESIGN

	Activity	Partners involved
1	We produced **an initial design** based on our understanding of analysts' workflow and using reusable components and patterns from Gov.uk Design System	
2	We **iterated our design** based on feedback and suggestion and further tested with other analysts	
3	We **iterated our design a second time** and further tested across councils. We have considered the accessibility of our application and will review it further in Beta.	
4	We then **presented our design** to collect any outstanding feedback during a "Show and Tell" session attended by local councils, GMCA and MHCLG	
5	All the feedback and learning are compiled into **a product specification**. It contains detailed descriptions of all user needs that are broken down into workflow	

Stockport Metropolitan Borough Council
Wigan Council
milton keynes council
East Sussex County Council
Isle of Wight Council
West Berkshire Council
Manchester City Council
GMCA Greater Manchester Combined Authority
Ministry of Housing, Communities & Local Government

INITIAL APPLICATION DESIGN (1/2)

WE BASED A FIRST DESIGN ON OUR KNOWLEDGE OF ANALYSTS' WORKFLOW, THE CURRENT DFE PORTAL AND GOV.UK DESIGN SYSTEM

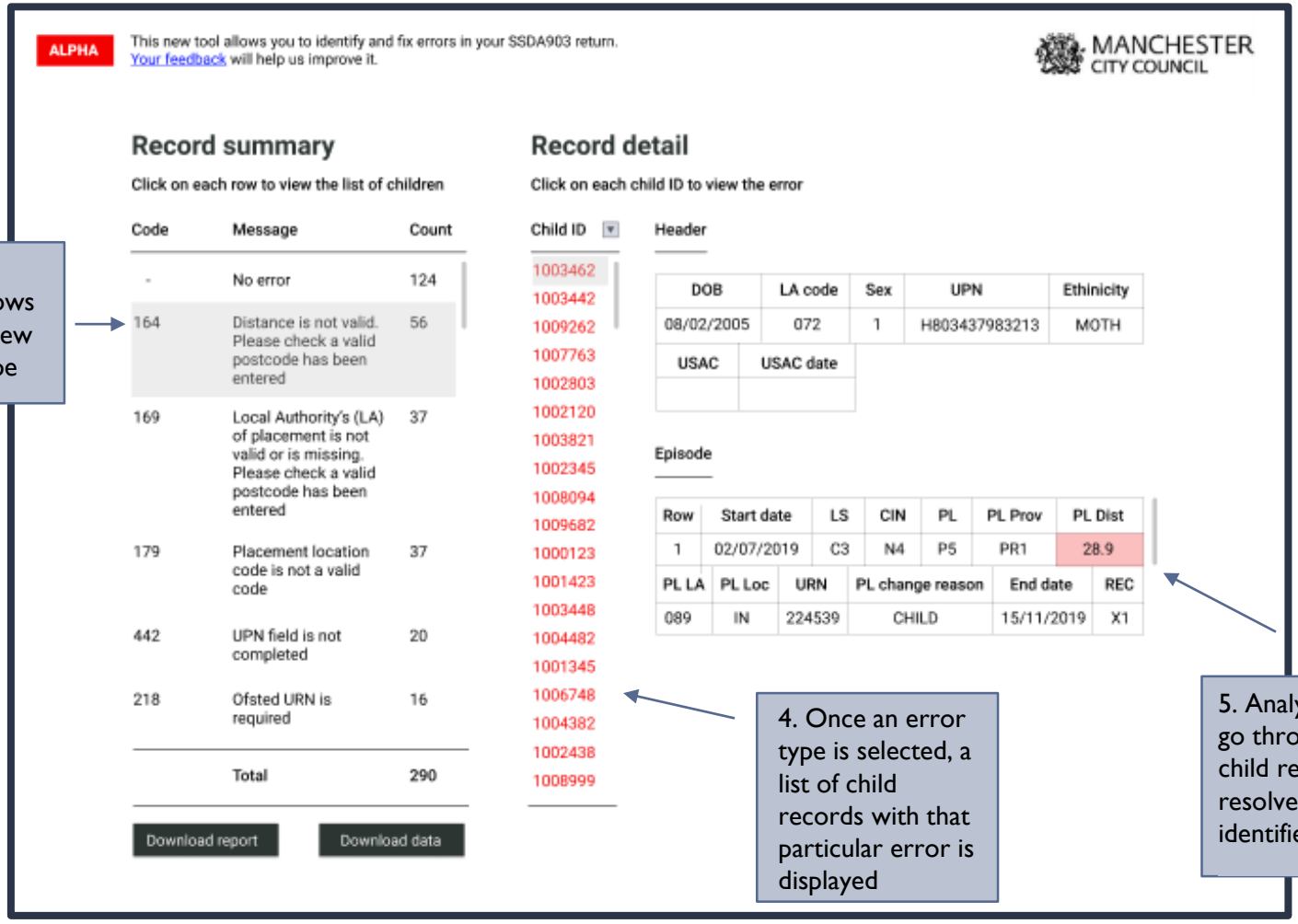


1. Users select the relevant files on their computer by pressing the browse button

2. Once all the files are selected, users press validate to run the error identification

INITIAL APPLICATION DESIGN (2/2)

WE BASED A FIRST DESIGN ON OUR KNOWLEDGE OF ANALYSTS' WORKFLOW, THE CURRENT DFE PORTAL AND GOV.UK DESIGN SYSTEM



WE COLLECTED FEEDBACK ON THE INITIAL APPLICATION DESIGN FROM ANALYSTS…

Learning	Supporting quote
Analysts liked the design and found the navigation intuitive	*"It looks very self-explanatory. It is quite similar to the DfE system but condensed – which is good!"* *"It is looking great, very easy to understand"*
Analysts felt that the display of information could be refined further	*"There is an awful lot of information on the right hand side. Formatting wise, it could be a bit more user-friendly"*
Analysts liked the option of viewing a list of records based on error type. This is not something they can currently do on the DfE portal.	*"It's looking great, I can see this feature being very useful"* *"I like the ability to group errors by type. This is not something that we can easily do within the current DfE system"*
Analysts wanted an easy way to view all the errors associated with each child record	*"In this way, you see errors one at a time. Even though it's useful to group errors, I need to see all the errors associated with a Child ID"*
Analysts wanted to be able to re-validate their data once they have made a correction	*"For instance, I would enter the URN to get rid of URN error, but then I wouldn't know if the URN matches the placement provider"* *"Sometimes we will clear the top layer errors without knowing what's hidden underneath"*

… AND IMPROVED IT BY IMPLEMENTING A SERIES OF CHANGES

Learning	Changes implemented
Analysts liked the design and found the navigation intuitive	We kept the broad layout but refined the display so that analysts can more easily view information, particularly if a child has been through many care episodes
Analysts felt that the display of information could be refined further	
Analysts liked the option of viewing a list of records based on an error type. This is not something they can currently do on the DfE portal.	We condensed the record summary section into the filter to keep the option for analysts to view errors by type
Analysts wanted an easy way to view all the errors associated with each child record	We anchored the display around child IDs on the left so that analysts can navigate through each child record and view all the associated error codes and descriptions
Analysts wanted to be able to re-validate their data once they have made a correction	We created a button for analysts to re-validate for errors while making corrections so that they can see whether errors have been resolved or if new ones have appeared

WE CAREFULLY CONSIDERED ACCESSIBILITY AND HOW WE WOULD FURTHER EXPLORE THIS IN BETA

Alpha accessibility considerations

In alpha we:

1. **Ensured accessibility for current users** – we know all analysts use the DfE SSDA903 portal, so by starting with this design we ensure it is accessible to all. However, we think we should improve upon this, as it will not be accessible to all
2. **User research with people with disabilities** – we explored whether we could research with analysts with relevant disabilities. However none of our users within the 9 councils had any disabilities hindering their use
3. **Initial assessment of accessibility** – we did an initial assessment of our tool against the Web Content Accessibility Guidelines (2.1) and assessed that it was feasible for this prototype to meet these in beta and live. To confirm this assessment, we tested example features e.g. highlight errors with text as well as colour (as the DfE portal does), to aid those with colour vision deficiency

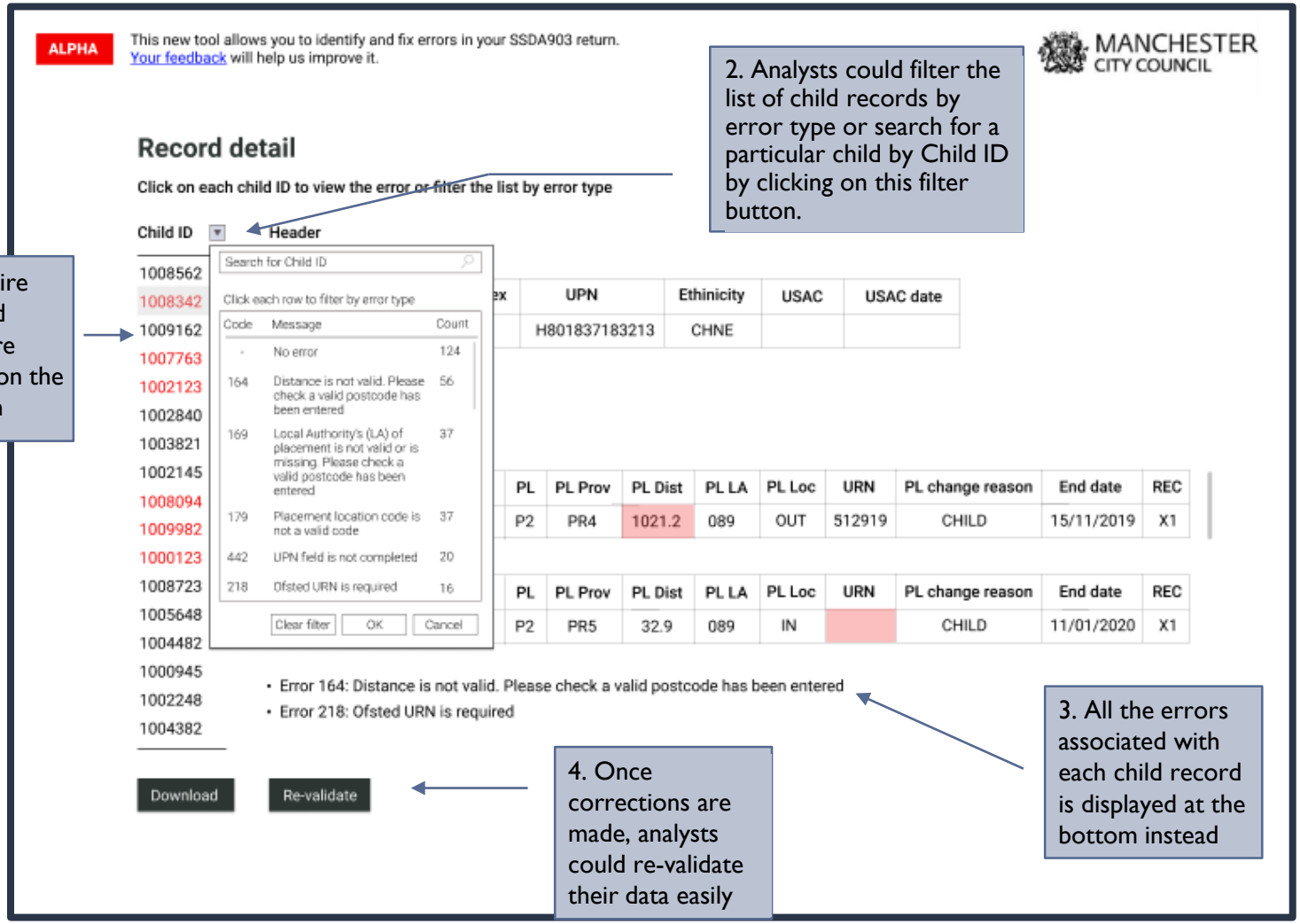
Beta accessibility considerations

In beta we would:

1. **Research with users with disabilities** – we need to reach out to other councils to find and test with analysts with disabilities impacting their use of technology. If we can't get coverage of a range of relevant disabilities, then we will need to research with people with these disabilities who aren't our users
2. **Assistive technologies** – we will test the tool with assistive technologies (e.g. readers, magnifiers, speech recognition)
3. **Accessibility** – we will implement accessibility guidelines (following WCAG 2.1) to ensure the tool is accessible
4. **Accessibility audit** – we will externally audit the accessibility of the tool
5. **Accessibility statement –** we'll publish a plain English accessibility statement

SECOND VERSION OF APPLICATION DESIGN

WE ITERATED OUR INITIAL APPLICATION DESIGN WITH FEEDBACK AND SUGGESTIONS FROM ANALYSTS



WE COLLECTED ANOTHER ROUND OF FEEDBACK FROM ANALYSTS…

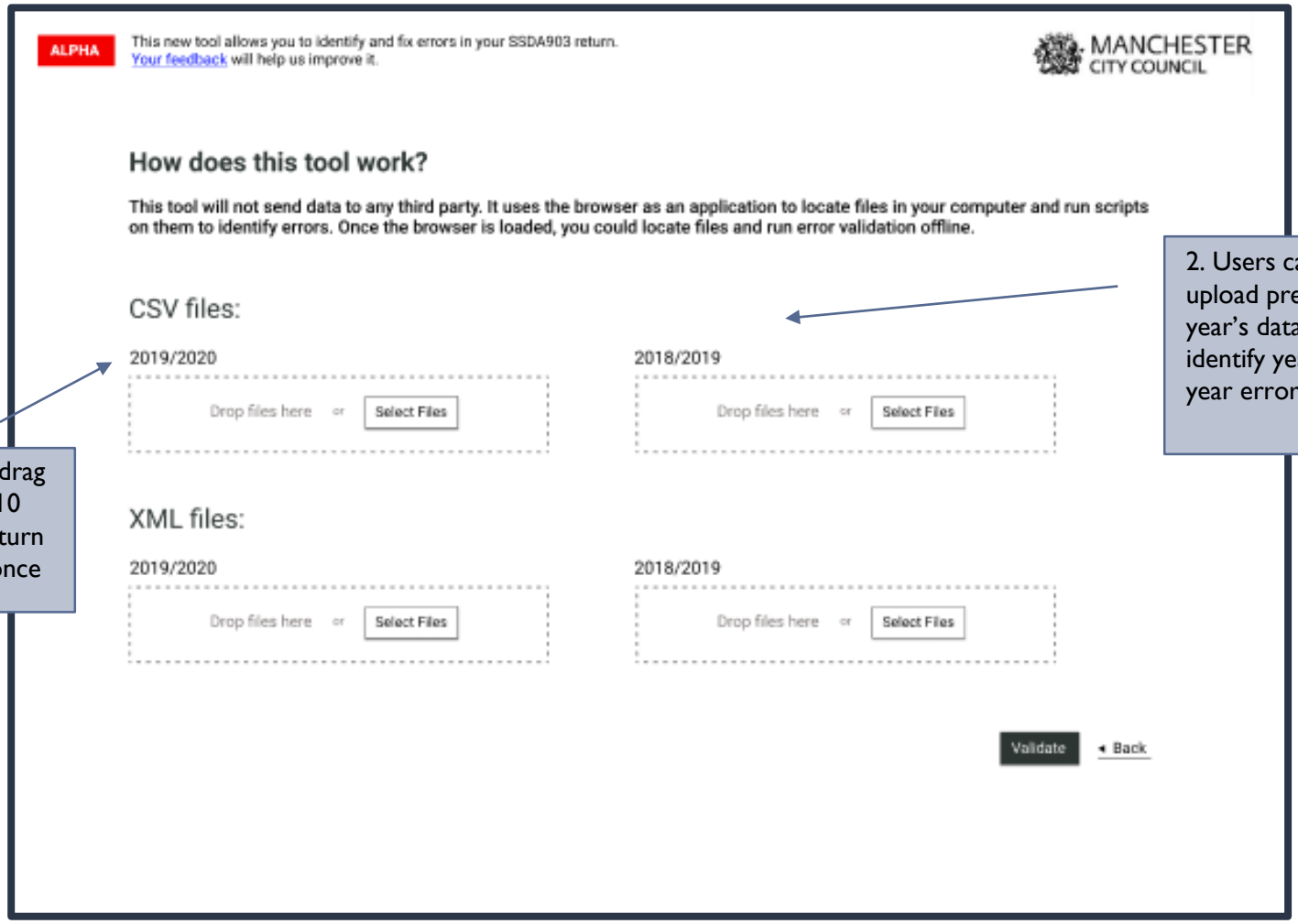
Learning	Supporting quote
Analysts found the design visually appealing and liked the flexibility of viewing errors based on Child ID and error types	*"It's looking very good."* *"I like that this gives you the flexibility to group by errors to understand the big picture and then go through one by one for data cleaning."*
Analysts thought that other info is best displayed as tabs next to the episode tab	*"I don't need to see other modules when I am focused on cleaning specific errors related to one."*
Analysts said possible answers should be displayed near error description	*"If the application automatically scans through Ofsted record and display info near the errors, it'd easily save me a week of work!"* *"Problems with placement reference numbers form a significant part of our total error count, so this will be useful."*
Analysts would find it useful to be able to view and amend records across years	*"It's be useful to have the ability to check for year-on-year errors"* *"We do get quite a bit of year-on-year errors and currently it's difficult to identify them"*
Analysts would like to be able to download the source data and specific error reports	*"I amend the errors on the portal to avoid reuploading files, so it'd be great if I can do that here"* *"I find the summary and list of Child ID reports extremely useful for us to have"*
Analysts would like to be able to drag and drop files into the application instead of manually selecting each	*"Including last year's file, there could be 20 that need to be uploaded. It'd be great if we could drag and drop those into a box"*

… AND IMPROVED THE SECOND VERSION OF THE APPLICATION WITH FURTHER CHANGES

Learning	Changes implemented
Analysts found the design more visually appealing and liked the flexibility of viewing errors based on Child ID and error types	No further action taken
Analysts thought that additional info are best displayed as tabs next to the episode section	We created tabs next to the episode section that allows analysts to flick between different pages of information
Analysts said possible answers should be displayed near error description	We added suggestions for possible answers that is drawn from Ofsted placement database near the error description
Analysts would find it useful to be able to view and amend records across years	We added a section for analysts to upload files of the preceding years and a year filter so that analysts can view information across two years easily
Analysts would like to be able to download the source data and error reports	We added download button that allows analysts to download two useful reports and the source data
Analysts would like to be able to drag and drop files into the application instead of manually selecting each	We created areas where analysts can drop a set of files into instead of having them manually uploading CSV files

FINAL VERSION OF THE APPLICATION DESIGN (1/2)

WE FINALISED THE DESIGN AFTER RECEIVING LARGELY POSITIVE FEEDBACK DURING A SHOW AND TELL SESSION ATTENDED BY LOCAL COUNCILS, GMCA AND MHCLG

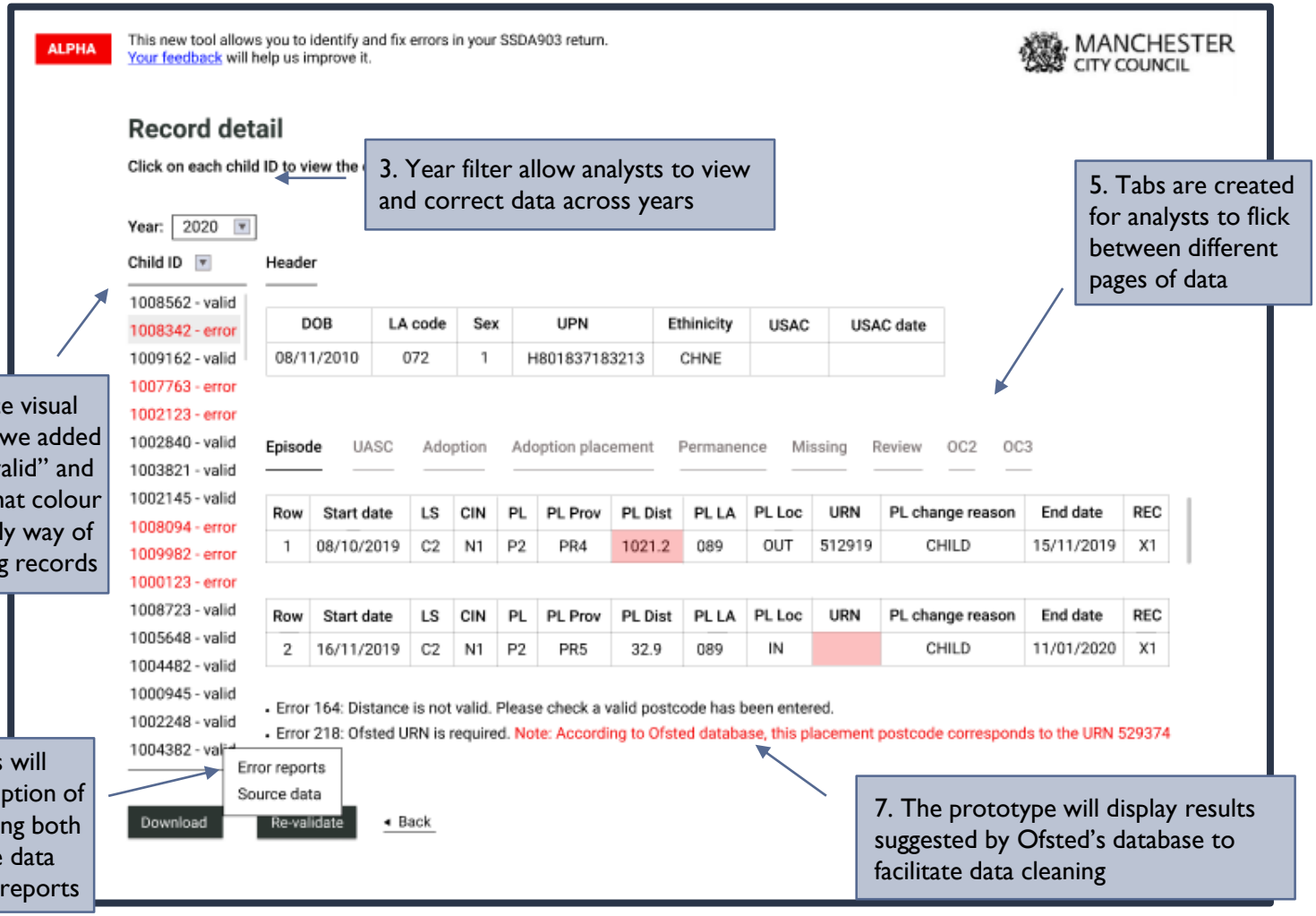


1. Users can drag and drop all 10 SSDA 903 return CSV files at once

2. Users can upload previous year's data to identify year-on-year errors

FINAL VERSION OF THE APPLICATION DESIGN (2/2)

WE FINALISED THE DESIGN AFTER RECEIVING LARGELY POSITIVE FEEDBACK DURING A SHOW AND TELL SESSION ATTENDED BY LOCAL COUNCILS, GMCA AND MHCLG



WE COMPILED A DETAILED LIST OF USER SPECIFICATIONS FOR BETA

Epic	User Needs	Priority	Essential?
Upload data	Can upload episode & header files	Alpha	Essential
	Can upload CSV files	Alpha	Essential
	Can upload all files	Beta	Essential
	Can upload files from the previous year	Beta	Essential
	Can upload xml file	Beta	Essential
	Can upload a number of files at once by dragging and dropping, instead of manually browsing each	Beta	Essential
Validate errors	Checks for errors 218 (URN required), 164 (distance invalid) and 178 (placement provider code invalid) - see Alpha errors tab	Alpha	Essential
	Checks for all pre-validation errors (formatting, date mismatch etc.) - see Prevalidation errors tab	Beta	Essential
	Checks for all the latest version of DfE data validation errors - see DfE data validation tab	Beta	Essential
	Checks for errors against Ofsted provider list - see Ofsted placement data tab	Beta	Essential
Error display	Errors are displayed by ChildID and ChildIDs with errors highlighted red and labelled "error" - see Figma design tab for this Epic	Alpha	Essential
	Clicking on each ChildID reveals the header and episodes information	Alpha	Essential
	Error codes and description are written below episodes	Alpha	Essential
	A count of total error is shown below the list of ChildID	Beta	Essential
	Analysts can search for a record based on ChildID, filter all errors or based on error types	Beta	Essential
	There is a tab for each of the additional module for analysts to select and view data	Beta	Essential
	Error descriptions are clickable. Once clicked, the shaded cell becomes darker to differentiate the cell that corresponds to a particular error code	Beta	Nice-to-have
Data cleaning	Analysts can edit data directly on the browser	Beta	Essential
	Analysts can refresh by clicking the validate button once changes are made	Beta	Essential
Download data	Analysts can download updated data, with an additional tab detailing all the changes compared to uploaded data	Beta	Essential
	Analysts can download two error reports, summary of validation check failures and list of children with error codes - see Error report tab	Beta	Essential
	Analysts can download CSV files	Beta	Essential
	Analysts can download XML files	Beta	Nice-to-have
Tech infrastructure	Tool doesn't require DPA	Alpha	Essential
	Tool doesn't require IT involvement	Alpha	Essential
	Tool tracks who has used it and how many errors they have	Alpha	Nice-to-have
	Tool comes with a technical document that describes how it's been developed (browser, pandas etc)	Alpha	Essential

Feature list | Alpha errors | Prevalidation errors | DfE data validation | Ofsted placement data | Figma design | Error report

APPLICATION DESIGN AND FUNCTIONALITY CONCLUSION

From our research on user needs, we found that analysts need an application which could help them:

1. Identify errors within their 903 data year round
2. Automate some cases of information finding

Through remote usability testing, we captured a detailed specification of the functionality and design that the application should have. This will serve as guidelines for us as we continue to develop the application.

We had to be mindful about aligning our design with the existing DfE portal as it is something that our users are familiar with. However, we were able to make a few valuable improvements:

a. We created an option of grouping child records by error type, which simplifies the process of resolving errors.
b. For about 40% of errors, we will be automatically suggesting possible answers based on Ofsted's placement database

OUR APPROACH IN TESTING APPLICATION FUNCTIONALITY AND DESIGN AND TECHNICAL SET UP

What we need to test	Our approach in testing
4a. Application functionality and design What are the functionalities and design that analysts would find useful in an application to identify errors year round and effectively resolve them?	We produced designs using Figma and iterated the features of the applications with analysts through remote moderated usability tests
4b. Technical set up What technical set up would enable the application to be most rapidly implemented and scaled across local councils?	We built a working prototype and shared it with analysts to test if they are able to use it successfully in their local environment

OVERVIEW ON TESTING TECHNICAL SET UP

	Activity	Partners involved
1	We had **an initial brainstorming session** to identify a set of major technical requirements and prioritised the ones that are most important to test	
2	We **developed** three possible solutions and assessed them against the technical requirements	
3	We **built and tested** the prototype internally with reference to the validation codes shared by the DfE. We focused on a minimum set of features sufficient to test the concept	
4	We then **tested the prototype externally** with analysts by asking them to run it in their local environment	

Stockport Metropolitan Borough Council
GMCA Greater Manchester Combined Authority
milton keynes council
East Sussex County Council
Isle of Wight Council
West Berkshire Council
Ministry of Housing, Communities & Local Government
Wigan Council
Manchester City Council

WE IDENTIFIED A SET OF MAJOR TECHNICAL REQUIREMENTS FOR THIS PROTOTYPE…

#	Technical requirement	Rationale
1	The tool shouldn't involve the sharing of personal data	Sharing personal data requires data processing agreements to be entered with each local council and a secure hosting environment for storing and processing data. This will incur significant investment in time and resources.
2	The tool shouldn't need software installation	Local software installation will require approval by the IT team of each local council. This will meaningfully delay the adoption of the application.
3	The tool needs to be able to read 903 data in CSV and XML formats	Local councils have developed an easy way of extracting 903 data from their CMS in CSV or XML as these are the formats accepted by the DfE portal.
4	The tool needs to apply DfE validation checks to the data	The DfE's validation checks are comprehensive and standardized across all local councils.
5	The tool needs to display info as the mock-up prototype does	The mock-up design has been robustly tested with analysts in local councils to ensure that it displays information in an intuitive manner
6	The tool needs to allow analysts to edit data	Resolving data errors on the application itself speeds up the process as it avoids correcting the underlying CSV or XML files and then reuploading them
7	The tool needs to be able to download reports and data	Reports are useful for local councils to review aggregate errors within their data, and analysts need to be able to download data that reflects changes they made

…AND PRIORITISED THE ONES THAT ARE ESSENTIAL TO TEST DURING THIS ALPHA PHASE

#	Technical requirement	Priority	Testing phase
1	The tool shouldn't involve the sharing of personal data	**High priority:** These are the most technically challenging constraints. We identified this as our riskiest technical assumption as the two requirements appear to be contradictory to each other.	**Alpha**
2	The tool shouldn't need software installation		
3	The tool needs to be able to read 903 data in CSV and XML formats	**High priority:** These are a fundamental part of analysts' workflow in resolving data errors. From our user research, it is clear that the biggest challenge that analysts face is in identifying the errors that exist within their data.	**Alpha**
4	The tool needs to apply DfE validation checks to the data		
5	The tool needs to display info as the mock-up prototype does		
6	The tool needs to allow analysts to edit data	**Medium priority:** Being able to edit data and download the corrections made will speed up the data cleaning process, but are not essential. As long as analysts are able to identify errors within their data, they can still resolve the errors directly in their case management system.	**Beta**
7	The tool needs to be able to download reports and data		

WE GENERATED AND ASSESSED THREE OPTIONS AGAINST OUR PRIORITISED TECHNICAL REQUIREMENTS

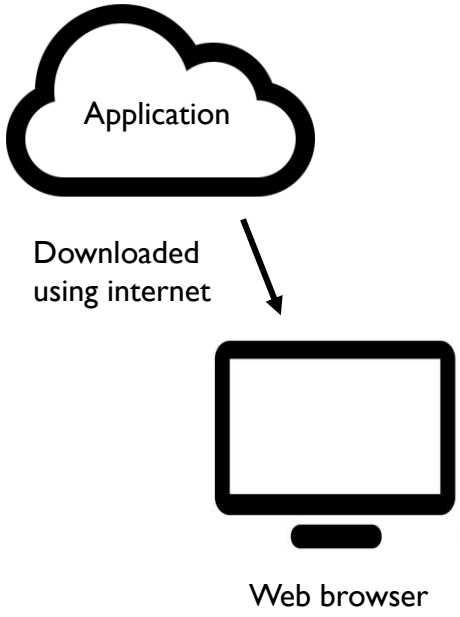
	#	Options	Analysis	Requirement
	1	**Installable software:** A software that analysts could download, install and run on their own equipment	Require the software to be packaged in such a way that it can be easily approved by local IT teams and added to the whitelist	✓ *1. No data sharing* ✗ *2. No installation* ✓ *3. Read data files* ✓ *4. Validate errors* ✓ *5. Display info*
	2	**Hosted application:** A web-based service that analysts could upload data to	Requires a central organization to be able to safely process personal data, data sharing agreements with local councils and a secure hosting environment with appropriate controls	✗ *1. No data sharing* ✓ *2. No installation* ✓ *3. Read data files* ✓ *4. Validate errors* ✓ *5. Display info*
	3	**In-browser application:** an application that would run in the browser	Does not require data to be sent off to a server managed by a central organization or local installation beyond what most users have available. The trade off is speed as the amount of processing in the browser is limited.	✓ *1. No data sharing* ✓ *2. No installation* ✓ *3. Read data files* ✓ *4. Validate errors* ✓ *5. Display info*

We decided to test an in-browser application because it best meets the technical requirements set out by the councils

HOW DOES AN IN-BROWSER APPLICATION WORK?

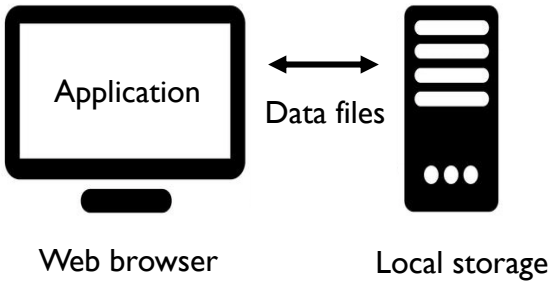
Step 1:

By visiting the URL, a copy of the application is downloaded to user's computer, as all websites work. Once loaded, the computer can be disconnected from the internet as the application can run its test locally.



Step 2:

The application will then read the data files in the local storage and display results directly. Throughout the process, the data files are kept in the local storage and not sent off to the internet.



WE THEN BUILT A PROTOTYPE WITH A SET OF MINIMUM FEATURES THAT MEETS OUR ACCEPTANCE CRITERIA

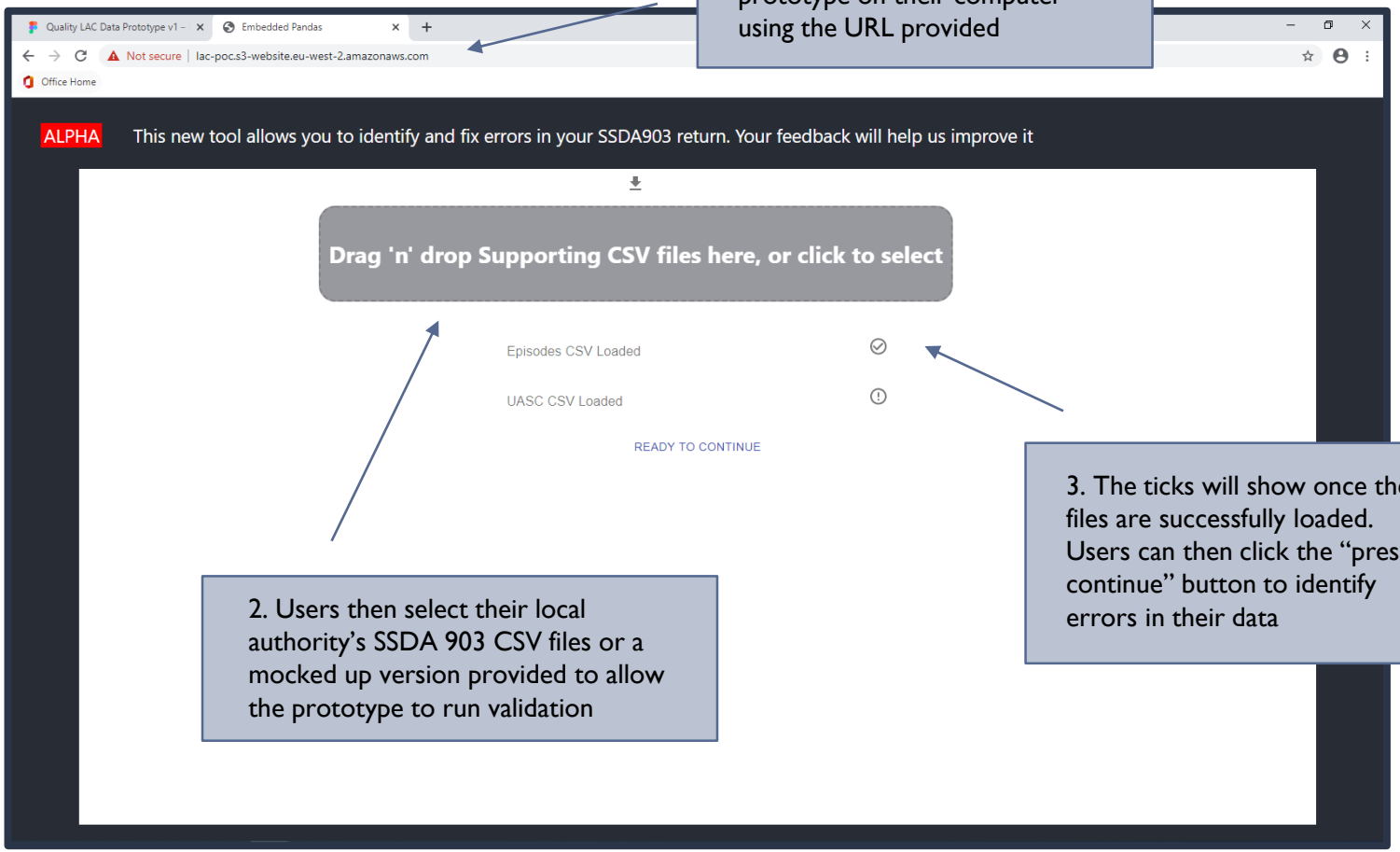
#	Technical requirement	Acceptance criteria
1	The tool shouldn't involve the sharing of personal data	Using an in-browser application, data files will be kept in user's local storage and not sent anywhere through the internet. We will test that users are able to select data files for validation in an offline setting once the browser is loaded,
2	The tool shouldn't need software installation	We will test that users are able to access the prototype in their local environment through the URL provided.
3	The tool needs to be able to read 903 data in their usual formats (CSV or XML)	We will test that the tool is able to read files in CSV format, leaving XML format for the subsequent phase. Data on looked-after children are divided into 10 discrete CSV files. Here, we will focus on implementing the first 2.
4	The tool needs to apply DfE validation checks to the data	We will test that the tool is able to validate data for 2 error types among all those that are currently implemented by the DfE system.
5	The tool needs to display info as the mock-up prototype does	We will test that the tool is able to allow users to see which child record suffers from the 2 error types

Our prototype is therefore an in-browser application that is accessible through a URL and capable of identifying 2 error types in an offline setting

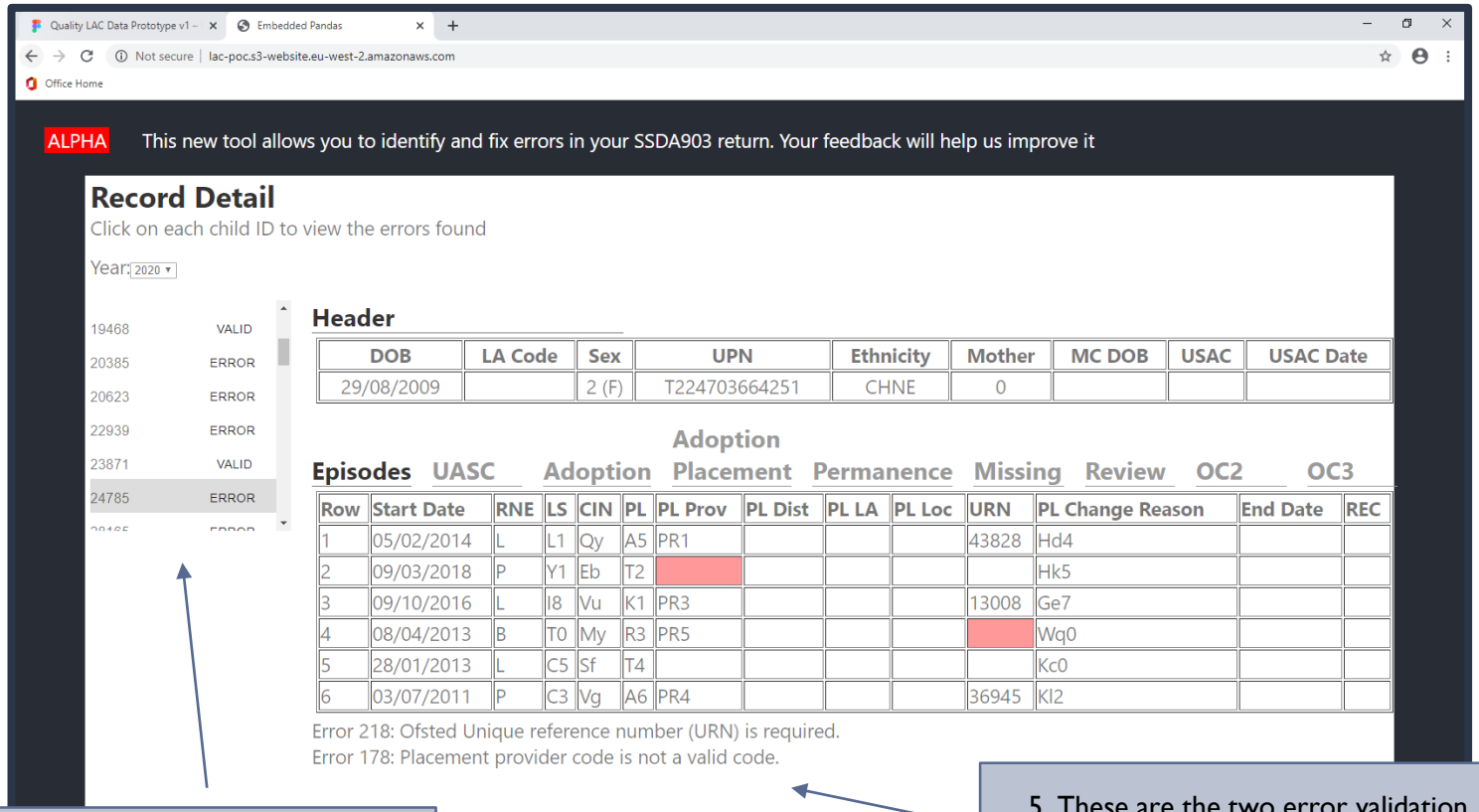
WHAT TECHNOLOGY DID WE USE TO BUILD THE PROTOTYPE?

#	Technology	Description	Function
1	Web browser	A software application, such as Google Chrome or Internet Explorer, for accessing information from a particular website.	Retrieve the prototype from a web server and then display the content on the user's device
2	Python	An open-source programming language with a large library of tools that can be used for data science	One of the tools, Pandas, allows us to work programmatically with data. It is currently used to run data validation in our application.
3	Pyodide	Pyodide is a tool that brings Python into the browser via WebAssembly	Traditionally, Python is not a language that can be run in a browser and requires a remote server to run on. Pyodide removes this requirement.
4	React	React is a JavaScript library for building user interfaces	React is used to speed up development and provide the components to display the results of data validation on the screen

WE TESTED OUR PROTOTYPE WITH ANALYSTS IN THEIR LOCAL ENVIRONMENT (1/2)



WE TESTED OUR PROTOTYPE WITH ANALYSTS IN THEIR LOCAL ENVIRONMENT (2/2)



4. Users can navigate the records using the scroll bar and click on each to display errors

5. These are the two error validation types that the prototype currently supports. Users could see which records in their data suffers from these errors

THE PROTOTYPE PASSED THE ACCEPTANCE CRITERIA BASED ON THE FEEDBACK FROM ANALYSTS

#	Acceptance criteria	Learning	Supporting quotes
1	Select data files for validation in an offline setting once the browser is loaded.	Analysts were able to run the application offline	*"I tried switching off my wifi after the browser loaded and it worked just fine."*
2	Access the prototype in their local environment through the URL provided.	Analysts could access the application locally with just the URL, even though the loading time varied	*"It loaded [in] less than a minute for me"* *"I've loaded the browser (which did take a little while, but nothing too horrendous)."* *"To confirm, the prototype successfully loaded…"*
3	Read 2 files in CSV format	Analysts could use the application to read their actual 903 CSV files	*We only have XML files, but I have used your mocked up CSV files. They worked with the application"* *"Yes I used our own last year's 903 CSV files with the application"*
4	Validate data for 2 error types	Analysts could identify errors in their data for those that were implemented	*"I tried the application using our own data files. The errors all showed up ok"* *"I used your CSV files and the error validation looks accurate"*
5	Display which child record suffers from the 2 error types	Analysts are able to navigate through all the records in their data and view those with errors	*"Very impressed with the prototype that's been developed!"* *"The application is very visual without being too much and the errors flag as soon as you click the error next to the Child's ID instead of having to ask for the error to be shown"*

TECHNICAL SET UP TESTING CONCLUSION

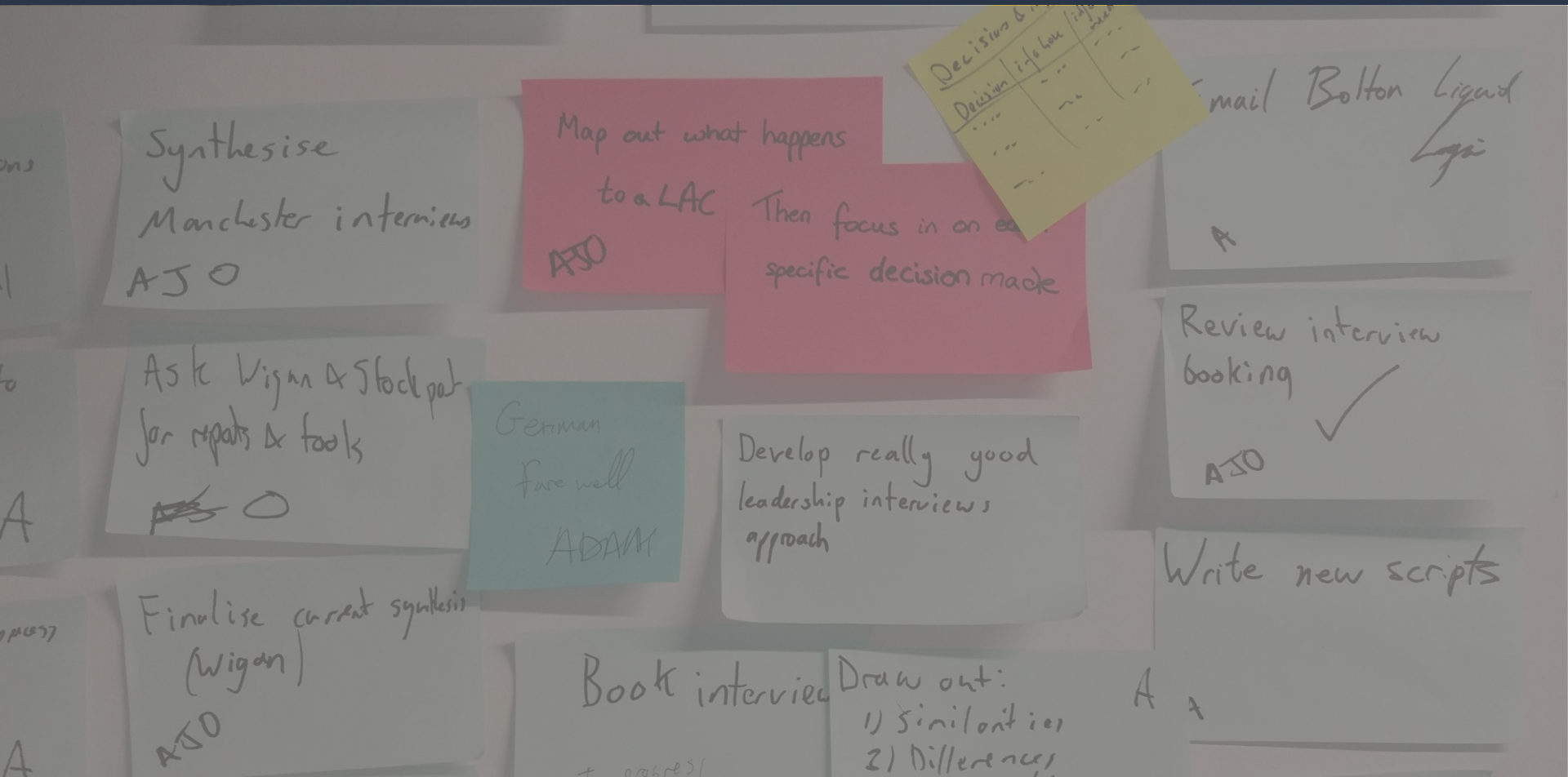
We explored three solutions and decided to build a prototype for an in-browser application. We developed acceptance criteria based on a set of technical requirements and tested the prototype with our users in their local environment.

The feedback we collected from technical set up testing shows that our prototype:

1. Is capable of handling all the fundamental parts of analysts' workflow in resolving data errors, including reading data files, running validation and displaying errors.
2. Could be accessed by a user through a simple URL. There's no need for putting data processing agreements in place or installing a software locally, both of which will incur significant investment in time and resources.

We are confident that is a feasible solution that both solves a real need and can be scaled rapidly to all 152 local councils.

3. RECOMMENDATIONS



USER RESEARCH SUMMARY

Our alpha stage user research has shown that:

1. The user needs identified in discovery are common across a further 6 councils
2. The impact of poor data quality is significant: it meaningfully impacts analysis and erodes leadership trust, stopping them from using evidence to improve services
3. Whilst there are a range of potential solutions for addressing data quality, the most impactful, feasible and common across councils would be to support analysts to identify data errors and surface the correct information where possible
4. We can create a design for an error identification tool that meets analysts' needs in a common way, with enhanced usability beyond the DfE SSDA903 portal
5. It is technically feasible to implement this design in a common and accessible way across councils

Our prototype testing has resulted in a detailed specification for an error identification and cleaning tool that meets analysts' needs. We have successfully built a minimum viable version of this that tests a subset of the SSDA903 data for a small subset of errors, using just one upload method, and meeting acceptance criteria based on the key technical challenges

CORE RECOMMENDATION

Our recommendation, based on our user research and analysis of the business case (see following section), is to progress to beta to develop an error identification tool for Analysts to:

1. Identify errors in children in care data
2. Highlight the correct information for placement errors (37% of errors)

The tool should follow the detailed feature specification developed through our prototype testing and be implemented as a browser-based tool, building upon our minimum viable product

WIDER RECOMMENDATIONS

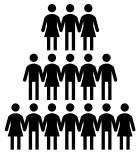
This tool would be valuable for and usable by all 152 children's services departments – following a private beta, it should be shared widely rapidly

1. This tool could easily be expanded to cover other key children's services datasets, starting with the SSDA903's sister-dataset, the Children in Need Census

2. The beta phase could also be accompanied by a small thread of work to initiate a shared-standard for improving data quality, based on the learnings of this work

WE THINK OUR SOLUTION SITS AT THE HEART OF THE LOCAL DIGITAL DECLARATION PRINCIPLES

We're working on a common problem
A range of national evidence, and our research across 9 councils shows that our problem is common to all. There's the potential to create something that benefits every council in the country



We've found a shared solution
By focusing on a common statutory process and a solution not requiring data sharing and independent local systems, we've developed a prototype tool that can be used by every council. Our research shows new councils can adopt the tool in a few minutes

We've designed around user needs
We've gone further to design around user needs, meaning we've identified exactly what analysts need to most effectively clean data on children in care, including a number of improvements beyond the existing SSDA903 submission portal (e.g. auto-identification of placement info, grouping child-records by error type)

We're fixing the plumbing
We've created a modular, flexible solution. By putting error-identification in the control of councils and open-sourcing the code, we have a modular building block to fit with other data processes. Anyone can expand the tool to cover other statutory data, to check for other error types, to check non-statutory data, to define other common data models, or to integrate into local analysis and data management

BETA PLAN OVERVIEW

There are three key things we need to do in beta:

Team needed:

Tool development

Build upon the alpha MVP to include the full features identified in our alpha research, building, testing and releasing iteratively. Key steps:

- Create synthetic test data to ensure validation rule fidelity
- Implement the full set of DfE validation rules
- Implement the UI
- Add xml upload
- Add upload of full SSDA903 dataset & historic data
- Add placement information identification feature
- Add functionality to edit and download data
- Explore approaches to speed up alpha prototype
- Document and open-source the code

Technical architect

Senior dev

Junior devs (low-cost outsourcing for value-for-money)

Product management, design and user research

Iteratively test and refine the tool with users at each release, particularly considering accessibility and incorporating GDS Design system components, accessibility audit

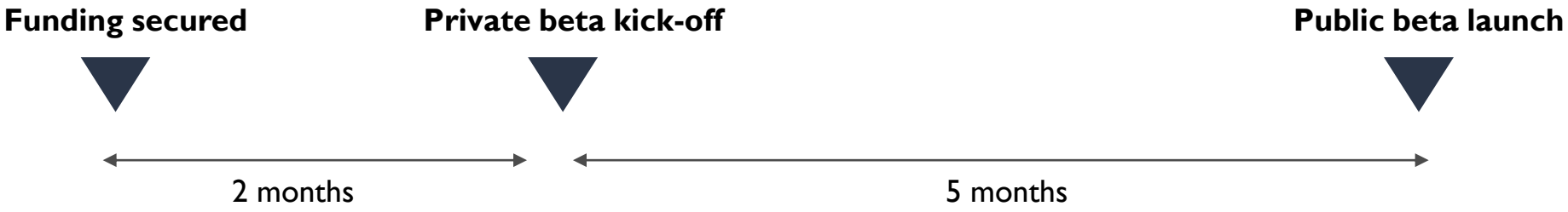
Product manager
Delivery manager
User researcher
Interaction designer

Network engagement and support

Create and deliver a communications plan to make analysts across all councils aware of the tool. Support the expanding user group of analysts

Communications & network lead

BETA TIMELINES



Project set-up

Activities:

- Procurement for technical support and accessibility audit
- Contracting
- Data access
- Engagement with wider council network to bring in further testing partners

Team: GMCA lead, LDCU lead, SF lead, GMCA procurement and IG

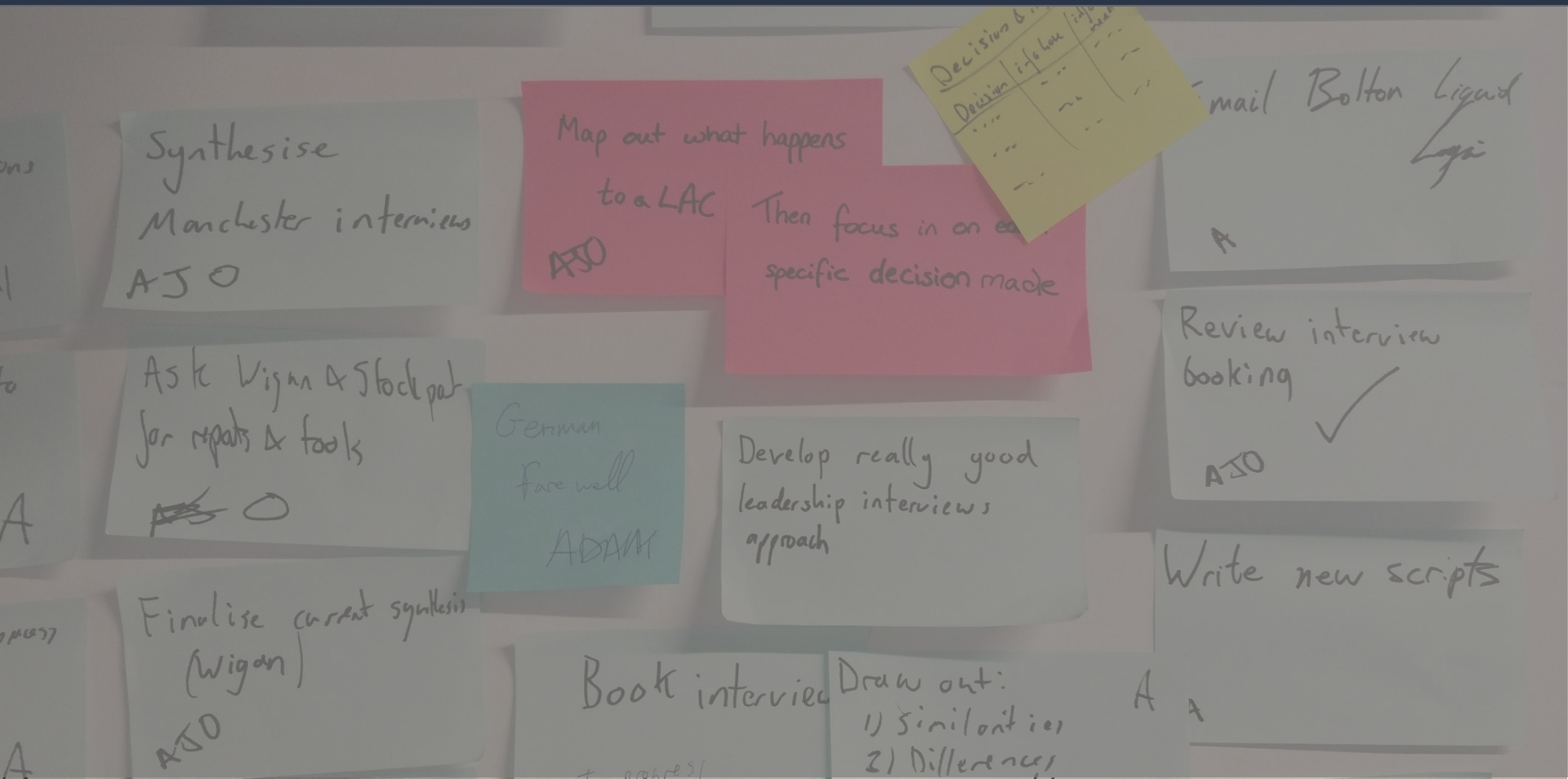
Private beta

Activities:

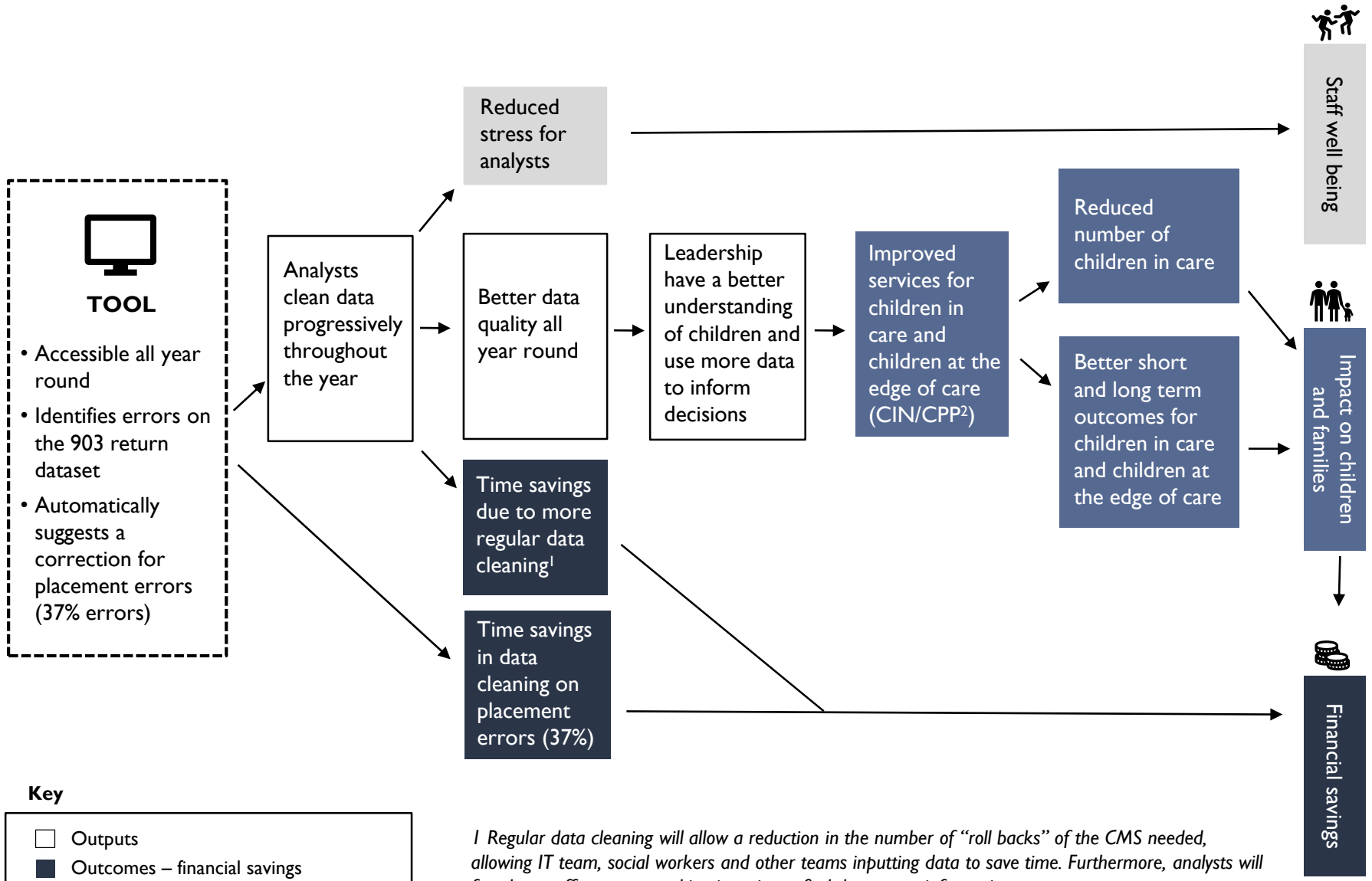
- Tool development
- User research and design
- Network engagement
- Accessibility audit
- Planning for public beta
- Write up user research report
- Open-sourcing of code and reusable templates (e.g. for browser-based deployment)

Team: Full partners i.e. all councils, MHCLG, DfE. Full beta team i.e. tech architect, developer, user researcher, designer, product manager, delivery manager, outsource developers, network engagement lead

4. BUSINESS CASE



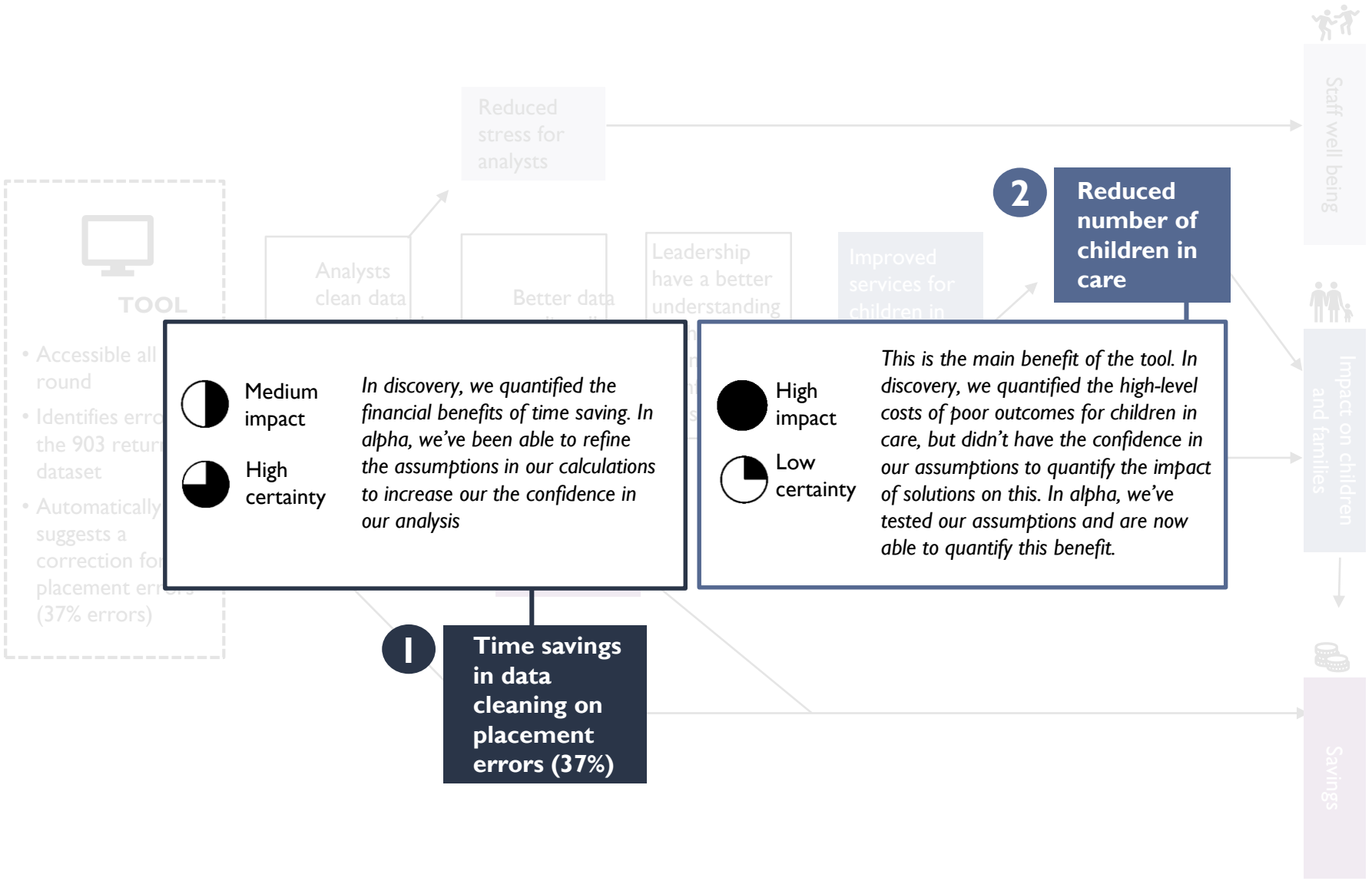
THEORY OF CHANGE FOR THE SOLUTION



1 Regular data cleaning will allow a reduction in the number of "roll backs" of the CMS needed, allowing IT team, social workers and other teams inputting data to save time. Furthermore, analysts will face less staff turnover, making it easier to find the correct information.

2. Children and young people under a Children in Need Plan or a Child Protection Plan

THEORY OF CHANGE FOR THE SOLUTION



BENEFITS OVERVIEW

THIS BUSINESS CASE FOCUSES ON TWO KEY BENEFITS OF THE TOOL:

1

Time savings for analysts

The tool enables analysts to automatically identify the correct placement information, saving them time searching for this information

This is a small benefit. In discovery, we quantified the financial benefits of this time saving. In alpha, we've been able to refine the assumptions in our calculations to increase our the confidence in our analysis

Benefit size: medium | *Confidence Factor: high*

2

Improving support for children

The tool helps authorities keep data quality high all year round rather than just at one point in the year. Our user research suggests that this would increase leadership trust in data. This is a key enabler of using analysis and evidence to improve support for children

This is the main benefit of the tool. In discovery, we quantified the high-level costs of poor outcomes for children in care, but didn't have the confidence in our assumptions to quantify the impact of solutions on this. In alpha, we've tested our assumptions and are now able to quantify this benefit.

Benefit size: large | *Confidence Factor: low*

1 TIME SAVINGS FOR ANALYSTS – EXPLANATION

AUTOMATING THE IDENTIFICATION OF PLACEMENT INFORMATION WOULD SAVE ANALYSTS TIME

Problem: Analysts spend several months each year working to submit the SSDA903 return on children in care to the DfE. This time-intensive and stressful process centres around cleaning this data so it passes a set of validation rules. The majority of analysts time is spent finding the correct information where data is either missing or invalid. The cost of this process for councils is over £22,500 / year[1]

Solution impact: The tool developed helps analysts identify errors year-round, reducing the stress of the intensive submission process. The tool shows analysts where data errors are. For errors related to placements (37% of total errors) the tool automatically shows analysts what the correct information is, based on records from Ofsted. This eliminates the majority of the time-consuming process for these errors

Benefits: The tool would save the majority of analysts' time for placement errors. We quantify this as a proportion of the total cost of the return. This free time can be more productively used on developing analysis on the effectiveness of children's services, to help leadership improve support to children in care. There are also a set of wider time-saving and stress-reduction benefits. Here we have focused on quantifying the benefits we can be most confident of

1. The best quantified estimated of the cost of the children's services returns comes from Gloucestershire County Council, who have outsourced the production of both their SSDA903 and CIN Census (the sister-return to the SSDA903, covering children in need) to a provider for £45,000 / year. This £45,000 figure is an underestimate of the true cost of these returns as A: Gloucestershire consider that this is cheaper than producing them in-house, B: Gloucestershire will still have to deploy some staff time to support the contractors in producing the returns. Analysts assess that the workload required for the SSDA903 and CIN Census returns are approximately equal, so we estimate the cost of the SSDA903 return as being £22,500 / council. These costs are consistent with our estimates across Stockport, Wigan and Manchester in Discovery

TIME SAVINGS FOR ANALYSTS – CALCULATION

AUTOMATING THE IDENTIFICATION OF PLACEMENT INFORMATION WOULD SAVE ANALYSTS TIME

£22,500[1]	x	37%[2]	x	90%[3]	x	85%[4]
Cost of SSDA903 return per council		Proportion of errors which could be automatically fixed		Time saving on each error where information is automatically identified		Confidence Factors

Benefits not quantified:

- Wider time-savings e.g. due to reduction in rollbacks, cleaning errors before social workers change roles, cleaning errors soon after events rather than later
- Time saving on data cleaning whilst doing analysis
- Stress reduction for analysts

= **£6,370**
Average saving per council per year

Calculation details

1. See previous slide for explanation of average cost of the SSDA903 return per council
2. Proportion of errors which related to placements: see error analysis in section 2a
3. The majority of analyst time on error cleaning is spent finding the correct information (see analysis in section 2b). Here we quantify that as 90%
4. Using GDS benefits case confidence factor data we rated the data out of 5 on if it is current (5), relevant (5), range (3), quality (5), consistent (4). The additional research in alpha has enabled us to increase our confidence from ~65% in Discovery

2 IMPROVING SUPPORT FOR CHILDREN – EXPLANATION

GOOD DATA QUALITY IS REQUIRED TO IMPROVE SUPPORT FOR CHILDREN IN CARE THROUGH EVIDENCE AND ANALYSIS

Problem: Councils need evidence and analysis to understand the most effective ways to transform support for children in care and to improve the very poor outcomes currently seen.

For councils to do this they need a strong data culture. Data quality, and the trust in data that comes from this, is one essential part of this. Although data quality alone will not improve outcomes for children in care, it is one of the key building blocks alongside e.g. data infrastructure, analysis tools, analysis skills and evidence-driven culture.

Solution impact: The tool will enable councils to keep the quality of their data on children in care high throughout the year, rather than at just one point. Our user research suggests that this will increases leadership's trust in data, removing a blocker for them using analysis and evidence to transform services.

Benefits: From councils that have used data and evidence to transform services (usually through large amounts of manual data cleaning), we can see the potential scale of benefits (see case studies on the following pages).

Data quality alone doesn't provide all of these benefits. To estimate the impact that data quality has in these transformations, we've looked at the total cost of transformation and estimated the proportion of this cost that was due to data quality. This gives an estimate of the scale of the data quality problem. We assess that improving data quality would on average lead to benefits of approximately this proportion of what other councils have achieved through the use of data.

2 IMPROVING SUPPORT FOR CHILDREN – CALCULATION

GOOD DATA QUALITY IS REQUIRED TO IMPROVE SUPPORT FOR CHILDREN IN CARE THROUGH EVIDENCE AND ANALYSIS

£23.5m[1]
Average spend on support for children in care per council

x

10%[2]
Reduction in children in care that councils have achieved through use of analysis and evidence

x

10%[3]
Estimate of the role of good data quality in enabling these transformations

x

30%
Confidence Factor

These assumptions are based on case studies of councils that have used data to transform outcomes for children in care, presented in the following two pages

Benefits not quantified:

- The benefits to children and families of improved support
- The long-term financial benefit to government of better outcomes
- The benefits based on improvements to services for children in need and on protection plans

=

£70,500
Average saving per council per year

Calculation details

1. There are 78,150 children in care across 152 children's services departments, each costing councils an average of £45,647/year. Source: DfE Children's Services Spending and Delivery statistics
2. Case studies of councils using data analysis and evidence to improve services (see following slides) show that savings of 4-7% are achievable. If a council were to make just two evidence-led transformations they could reduce costs by >10%
3. To estimate the role that data quality has in data transformation we asked the project teams in the two case studies what proportion of the project work was data quality related. They estimated 10-15% so we take the lower bound for conservatism
4. Using GDS benefits case confidence factor data we rated the data out of 5 on if it is current (4), relevant (2), range (2), quality (1), consistent (3) to give a confidence factor of 60%. We then halve this to account for the uncertainty around the estimate of the role of data quality in enabling analysis to give an overall 30% confidence factor

2 IMPROVING SUPPORT FOR CHILDREN – CASE STUDY 1

NEWCASTLE TRANSFORMED CHILDREN'S SERVICES BASED ON DATA ANALYSIS TO ENABLE 43 CHILDREN IN CARE TO SAFELY RETURN TO THEIR FAMILIES

Case study:

Newcastle City Council Family Insights Programme

Newcastle City Council

Description

NCC transformed children's services through data analysis by identifying needs profiles of children. With this insight they were able to restructure their services into specialist teams around these needs profiles, offering better support. DfE/Kantar evaluation[1] showed a range of benefits, including 87 children in care safely returned to their families vs 44 in a baseline case

Impact - 43 children in care safely returned to their families[1]

Net saving on children in care - 7%[2]

Proportion of work driven by data quality - 15%[3]

1. DfE/Kantar Public evaluation of the Newcastle City Council Family Insights Programme
   Newcastle City Council had 340 children in care before the programme and spent £45,000/child/year on average. the reduction in children in care has a £3.7m saving vs a £2.7m total investment. This £1m net saving (7%) is a significant underestimate as a) its only the first years benefits and b) it only accounts for benefits from children in care de-esscalating – DfE children looked after statistics, DfE Children's Services Spending and Delivery statistics
2. The total transformation funding from DfE was £2.7m
3. The project team estimate that at least 15% of the work of the programme was getting data quality sufficient for the analysis

2 IMPROVING SUPPORT FOR CHILDREN – CASE STUDY 2

ESSEX USED DATA ANALYSIS AND EVIDENCE TO MAKE THE BUSINESS CASE FOR EARLY INTERVENTION, HELPING KEEP 132 CHILDREN OUT OF CARE

Case study:

Essex County Council edge of care intervention

Description

Essex County Council analysed the needs of the population of children on the "edge of care" (those at high-risk of entering care) vs those in care. They assessed the potential for Multi-Systemic Therapy to help these children avoid care and identified a positive business case, helping them secure impact investment of £3.1m for this intervention. Evaluation by Oxford University showed savings of £17.9m[1]

Impact - 132 children kept out of care[1]

Net saving on children in care - 4%[2]

Proportion of work driven by data quality - 10%[3]

1. Evaluation by Oxford University showed a reduction of 96,000 care days vs a control group
2. Essex Country Council had 1,480 children in care before the intervention and spend £55,000 / child / year on average. The £14.8m net saving is 4% of this budget. – DfE children looked after statistics, DfE Children's Services Spending and Delivery statistics
3. The project team conservatively estimate that 10% of the work involved was related to data quality

THE LONG TERM SAVINGS FROM IMPROVING OUTCOMES FOR CHILDREN IN CARE ARE SIGNIFICANT

Beyond the initial savings to councils on placement costs etc. from improving support, there would also be a large set of lifetime benefits for these children. These lifetime benefits would save money for a variety of central government departments. Below we have quantified the overall scale of these benefits. However, due to their long-term, non-cashable nature, we haven't included them in our benefits calculations

Area[1]	% of Care Leavers that experience outcome	% of non-Looked After Children that experience outcome	Cost of outcome per year	Additional cost of outcome due to Care Leavers	Organisation bearing costs
Criminal justice[2]	**5.6%** *of care leavers are in custody at any point*	**0.13%** *of the population are in custody at any point*	**£34,840** *Annual cost per offender in prison*	**£754m**	**MoJ**
Employment Welfare benefits[3]	**11%** *Estimated care leaver unemployment rate*	**4%** *Unemployment rate for the general population*	**£3,063** *52 weeks of Job Seeker's Allowance*	**£81m**	**DWP**
Employment Foregone tax[4]	**68%** *Estimated employment rate for care leavers*	**75%** *Employment rate for the general population*	**£3,020** *Annual tax on average UK salary*	**£80m**	**DWP**
Health[5]	**46%** *of care leavers with mental health needs*	**13%** *of children with mental health needs*	**£2,338** *England average community expenditure per referral*	**£307m**	**DHSC**
Homelessness Homeless services[6]	**2%** *of care leavers homeless at any point*	**0.5%** *of the population homeless at any point*	**£14,808** *Estimated average cost of homeless services per person over one year*	**£88m**	**MHCLG**
Homelessness NHS, mental health, drug & alcohol[6]	**2%** *of care leavers homeless at any point*	**0.5%** *of the population homeless at any point*	**£7,717** *Estimated average cost of NHS / mental health / drug & alcohol over one year*	**£46m**	**DHSC**

1. All numbers here are rounded. Calculations are based on non-rounded figures, with any small differences in calculated figures due to this. A detailed breakdown of the calculations can be found in the supporting business case model.
2. 30% of YP in custody are CLs (Report by HM Inspectorate of Prisons); there are currently 83,000 prisoners in the UK (MoJ) with an adult population of 52,4M (ONS); cost per prisoner including overhead is based on NEF unit cost database.
3. NEETs data was used to estimate CL unemployment rate: CL are 3x more likely to be NEET vs their peers (Support for Care Leavers, Briefing Paper 08429, House of Commons Library), we therefore estimate a correspondingly higher unemployment rate for CL.
4. CL employment rate was estimated using the difference between general employment rate of 75% (ONS) and estimated LAC employment rate of 68%. LAC employment rate based on LAC unemployment rate of 11% and conservatively assuming that the same proportion of LAC as of the general population are out of the labour force (21%). Average UK salary in 2015 was 27,600 (ONS), giving £3,020 per person at a 20% tax rate on salary above personal allowance.
5. Based on relative proportions of CL vs. non-CL having MH needs, and the average community expenditure per referral (source: NHSE).
6. Based on relative proportions of CL vs. non-CL experiencing homelessness, and the estimated average cost of homelessness in relation to direct services and associated NHS, mental health, drug & alcohol costs (Source: Crisis, Better Than Cure report, 2016). The percentage of care leavers homeless at any point only includes those cases known to the council, so we expect this could be an underestimate of the true figures.

BENEFITS SUMMARY

Total savings per council

Time savings for analysts:		Improving support for children:		Total benefits
£6,400 Average saving per council / year	+	**£70,500** Average saving per council / year	=	**£76,900** Average saving per council / year

Total common savings

The tool could easily be used by all children's services departments. Given enthusiasm, we expect take-up to be high, so have presented scenarios for 50 councils using, every council using and a downside where only the project partners use the tool

Yearly benefits across 9 partner councils	**£692,000** Savings / year
Yearly benefits across 50 councils	**£3,840,000** Savings / year
Yearly benefits across 152 councils	**£11,680,000** Savings / year

COSTS OVERVIEW

Discovery	12 week discovery project Team: product management, 2 x user researcher, business analyst, senior oversight, council staff time	£110,000
+		
Alpha	16 week alpha project Team: product management, delivery manager, 2 x user researcher, business analyst, tech architect, developer	£100,000
+		
Beta	20 week beta project to develop and share tool Team: product management, delivery manager, tech architect, senior developer, outsourced developers, user researcher, designer, network engagement	£230,000
=		£440,000 Investment

Live maintenance	Ongoing network engagement, rule updating, tool maintenance, product management and user research	£20,000 / year ongoing

COST-BENEFIT ASSESSMENT

Costs and benefits under different scale scenarios

	1 council	9 councils (Alpha partners only)	50 councils (Partners, northwest southeast groups and contacts)	152 councils (All councils in England)
Investment	£440k	£440k	£440k	£440k
Net Annual Benefits	£57k	£672k	£3.8m	£11.7m
5-year ROI	0.4x	5x	27x	81x
5-year NPV (3.55% discount rate)	£37k	£4.6m	£28m	£85m
	This would not be viable for one council	**Downside** – fail to scale beyond project partners	**Base case** – scale across the North West, South East and partner networks in the first year	**Upside** – Our research suggests the tool would be both beneficial and usable for every council

THE TOOL WOULD REPAY INVESTMENT 27-FOLD OVER 5 YEARS ACROSS 50 COUNCILS

If you want to find out more, then please get in touch!

Paul Holme
Strategic Lead for Intelligence and Data Science
Manchester City Council

p.holme@manchester.gov.uk
@holme_paul