Data Protection Impact Assessment drafted for pan-London by LB Islington, and adopted by LB Camden

Version	Reason	Date	Author(s)
1.0	New	11/04/2022	Joe Luxton
1.1	Reviewed and updated following IGfL presentation from Social Finance	05/05/2022	Joe Luxton
1.2	Reviewed and approved by Islington IG Panel	23/05/2022	Joe Luxton
1.3	Reviewed and updated with 2023 amendments to the use-case	27/06/2023	Joe Luxton
1.4	Reviewed and approved by Islington IG Panel	29/06/2023	Joe Luxton
1.5	Updated to include purposes for aggregate data. Reviewed and approved.	24/08/2023	Joe Luxton

Project / Work Stream Name	LIIA Project: Pan-London Sufficiency Analysis	
Project / Work Stream Lead	Name	Joe Luxton
	Designation	Data Protection Manager– London Borough of Islington
	Telephone	020 7527 8002
	Email	Joe.luxton@islington.gov.uk
Overview: (Summary of the project/work stream)	**2023 Amendment** This DPIA has been updated to incorporate the additional data that is being requested to expand the analysis provided to DCS's beyond that which is possible using only the statutory datasets. The new dataset that is being requested (the "placement costs dataset") has been designed in collaboration with five London Boroughs as part of a project to improve use of financial data by LAs, funded by Department for Education and led by London Borough of Barnet. The dataset includes information that is already shared as part of the statutory extracts above, as well as additional data about children and the costs associated with placements.	

	Additionally, we are asking to expand the extent of the statutory data being shared to include a module related to care leavers. Field-level detail of the minimisation that will be conducted is provided in Appendix 2. One of the changes is to the existing data flow structure, with the data ingress, processing and egress role previously played by GLA (the London DataStore) now taken on by Social Finance, who are building a novel data platform for this purpose. The reason for this change is that the London Data Store, while robust at dealing with open access data, lacks some of the primary security features necessary for handling data containing personally identifiable information. The full details of this will be described in an amended Schedule 2 of the Data Processing Agreement. As this is a significant change to the existing agreement, two members of Information Governance for London are taking an active role in the scoping and development of this platform to ensure that it meets the data security standards necessary for the data being shared. **Overview and context** London has a regional approach to sector-led improvement, overseen by the Association of London Directors of Children's Services (ALDCS). Known as the 'London Innovation and Improvement Alliance (LIIA), this is a standing body for cooperating on the improvement of Children's Services through identification and sharing of best practice, including creation of shared datasets and comparative analyses. Within the LIIA structure we have a shared analytical team, currently based at London Councils and with IT hosted at LB Waltham Forest. They agree questions to be answered with the ALDCS and deliver it by taking in aggregate data from all Boroughs, producing pan-London analyses, and sharing these back to the ALDCS. As the LIIA has matured, the DSCs have begun to ask for analysis of issues which are important to improving outcomes in London, but which require boroughs to share personal data. Therefore, they have commissioned this project to establish a secure and ethical approach to conducting any pan-London analyses which rely on individual-level data. The process is being designed around three principles: • **Respect for the rights of data subjects** – data processing is proportionate to benefits, and in line with subjects' expectations about how that data should be used. • **Minimising work for Boroughs** – by using wherever possible datasets which each borough already has and relying on the

pan-London infrastructure already created for data collaborations including IGfL, the London DataStore, and the Information Sharing Gateway.

- **Focus on use cases which improve outcomes** – enabling us to maximise improvement for the resources spent, and clearly link each act of processing to a specific legitimate purpose

The LIIA team are being supported in this by Social Finance, a not-for-profit data specialist who have previously developed the information governance and technical infrastructure for multi-LA data collaborations using individual-level data from children's services data.

Contractual Arrangements

The LIIA team are developing a common Data Processing Agreement (DPA) and contract to be used between each Data Controller, and the Data Processor. DPAs are also being agreed between the Data Processor (London Councils) and all sub-processors involved in processing. These agreements are being developed in consultation with the Information Governance Group for London (IGfL).

DPOs should note that this project is a replication of a project which Social Finance ran in the South East, where four LAs approved the same processing as well as very similar data flows, DPAs, and contracts. We have permission to share those documents with you.

The DPA was originally developed for a project which has recently been selected as an ICO case study for good practice in sharing sensitive data.

This DPIA is for the use case: **Pan-London Sufficiency Analysis**, and corresponds to Schedule 5 of the DPA between LIIA and the Boroughs

Overview of Intended Processing: Common to all use cases

- Each Borough uploads data, including personal sensitive data, onto a private, borough-specific folder on the LIIA data platform.
- Scripts provided by the LIIA team then processes this data on the LIIA data platform in three ways:
    - Preparation of single Borough's data for analysis, including:
        - Checking whether agreed pseudonymisation and data minimisation has been done prior to sending, and implementing it if not (e.g.

deletion of fields not required; degrading highly disclosive data such as postcodes and dates of birth);
        - Assessment of data quality (missing values, logically inconsistent values);
        - Transformation of data to conform to a common schema.
    - Loading the prepared data for all Boroughs into a pan-London database;
    - Creating extracts from that database for analytical purposes specific to the use case.
- The single-Borough output of step 1 are made available back to the Borough, free for them to use for their own internal analysis
- The extracts created in step 3 are made available to an approved analyst (either at London Councils or a named sub-processor approved by the DPOs) to produce the pan-London analyses specific to the use case
- Where appropriate the LIIA d&i team will provide pan-London aggregate anonymous data for work commissioned or approved by ALDCS/LIIA or as agreed with partner organisations (where this data has already been collected through the LIIA CLD project) and so help reduce burdens on individual LA data & performance teams (relates to all use cases).

<u>Use case: Pan London Sufficiency Analysis</u>

Context

This use case for the LIIA Project involves aggregating and sharing Boroughs' data from Children's Services dataset that is produced as part of Boroughs' statutory duties - Children looked after return 'SSDA903'. The analysis, to be conducted by the Commissioning Alliance, based at London Borough of Ealing, aims to:

- accelerate service improvement, by enabling the identification and prioritisation of opportunities for improvement, and the identification of good practice in other Boroughs
- compare the commissioning of placements for looked after children across London Boroughs to improve the market information available to make commissioning decisions
- monitor inequalities, by enabling comparative analysis of the odds of key outcomes for children in care (e.g. the distance

between their home and their placement, placement breakdowns) for children of different ethnicities

Data will be aggregated and shared in reporting such that no individuals are identifiable, though there is a risk of re-identification of individuals due to small aggregations in some analyses. Boroughs will be identifiable in the shared analysis. The analysis will be shared among DCSs in London Boroughs and additional third parties, such as placement providers, subject to Memoranda of Understanding to be agreed between ALDCS and the LIIA Board. The reason for the additional sharing in this case is to provide the outputs of analysis to both the creators (London Boroughs) and suppliers (placement providers) of demand for care placements. As a key purpose of this use case is to facilitate change in the commissioning of placements, providing this two-way exchange of information is deemed a vital component of the analysis.

Use Case Specific Data Processing

- The pan-London extract is accessed by analysts at the Commissioning Alliance (or on their behalf by approved sub-processors) via a secure download from the LIIA data platform into a MS Azure data warehouse, hosted by Social Care Network
- Individual-level data are held in the data warehouse, accessible only by named individuals from Commissioning Alliance (and approved sub-processors)
- Data from Ofsted produced reports about placement providers, including data relating to placement cost and quality, will be joined to the SSDA903 data at the placement level.
- Descriptive analysis of event frequencies, sequences of events, and breakdown by Borough, age group, ethnicity group and placement type by geography (at the postcode district level), with comparison by Borough conducted in Power BI report, hosted by Social Care Network
- Power BI report shared with DCSs and other third parties, such as placement providers, through personal, secure link.

<u>Pan-London Sufficiency Analysis</u>

Data Source

The data is initially collected by frontline staff working for or on behalf of Children's Services as part of the exercise of the authority's statutory duties. It is initially stored in the authority's case management system.

	Extracts from the application database are then prepared for annually submission to the DfE. This extract is re-used as inputs for the LIIA pan-London analysis. The processing to produce these pan-London datasets is designed to produce an additional layer of minimisation between the full datasets provided by each Borough, and the data being analysed. The new dataset that is being requested (the “placement costs dataset”) has been designed in collaboration with five London Boroughs as part of a project to improve use of financial data by LAs, funded by Department for Education and led by London Borough of Barnet. The dataset includes information that is already shared as part of the statutory extracts above, as well as additional data about children and the costs associated with placements. Additionally, we are asking to expand the extent of the statutory data being shared to include a module related to care leavers **Nature of Processing** Descriptive analysis to understand and compare how placements for children in care are commissioning across London Boroughs. This can support London LAs in planning and improving Children’s Services. Outputs will be tables and charts showing aggregate data (no PII). However, in visualisations that display the number of children from a Borough placed in a specific placement type in a selected postcode district, there is the potential to describe very small numbers of children, potentially as low as a single child. This is because the density of placements may be very low in less populated areas and the capacity of some placement types is also low. Though a single child may be described in this way, no directly identifiable data about that child, such as an identifier or date of birth, can be accessed by the user. This functionality does increase the risk, however, that a child becomes re-identifiable due to the circumstances of their Borough of origin, their demographic profile and their placement location. There will be no machine learning, no automated decision making, and no attempts to support decision making about an individual case.

Implementation Date:	The LIIA Child Level Data Project has now launched, with all 33 London Boroughs signed up to share data and data from the SSDA903 shared by all Boroughs for one of the three starting use cases, Pan-London Sufficiency Analysis. Reporting from this use case has also gone live and has being shared with approx. 100 users across London Boroughs. Changes to be implemented by August 2023
<u>Environmental Scan</u> Describe the consultation/checks that have been carried out regarding this initiative or, project of similar nature, whether conducted within your organisation or by other organisations. *Please provide any supporting documents such as benefit study, fact sheets, white papers, reports or refereed articles published by industry associations, technology providers, and research centres.*	We do not need to consult with data subjects as the purpose is 'public task' and the data is being used in line with the purposes outlined in the data controllers' existing privacy notices (see Appendix 3 – Guidance on privacy notices). However, in light of research on public attitudes to sharing health and social care for secondary purposes we propose publishing blogs on the LIIA website to explain what we are doing, the benefits we hope to achieve for London, and how we are protecting individuals' privacy in the process. **Why We Think This May Need a DPIA** The data to be processed concerns vulnerable individuals (e.g. children in care). Data will be anonymised to the fullest extent possible, but in most cases it will retain some risk of identification by third parties in the event of a data breach. The purposes are analysis of administrative data for the purpose of delivering the LAs' statutory duties - with an explicit bar on: identification of individual data subjects, determining whether individuals do or do not get a service, automating any decision making about an individual, use of machine learning. These purposes and means are not novel and are in line with the Boroughs' existing privacy notices. However, two things might be considered novel: 1. Sending their data to a third party (Commissioning Alliance) to be processed instead of doing it in-house (although we note that the same data is routinely provided to DfE for similar processing and purpose); 2. Combining their data with that of other Boroughs to enable new questions to be answered (although we note that DfE combine the same datasets and conduct similar processing for the same purpose). There is an argument that because the same data is already transferred to third parties (DfE) and combined with data from other LAs in order to conduct very similar processing for a very similar purpose, this is not novel processing. However, there is sufficient

	ambiguity about whether that removes novelty to warrant consideration of a DPIA.

Step 1: Complete the Screening Questions

Q	Category	Screening question	Yes/No
1.1	Technology	Does the project introduce new or additional information technologies that can substantially reveal an individual's identity and has the potential to affect that person's privacy?	Yes
1.2	Technology	Does the project introduce new or additional information technologies that can substantially reveal business sensitive information, specifically: have a high impact on the business, whether within a single function or across the whole business?	No
1.3	Identity	Does the project involve new identifiers, re-use or existing identifiers e.g. NHS or NI number, Local Gov. Identifier, Hospital ID no. or, will use intrusive identification or identity management processes or, electronic linkage of personal data?	Yes
1.4	Identity	Might the project have the effect of denying anonymity and pseudonymity, or converting transactions that could previously be conducted anonymously or pseudonymously into identified transactions?	Yes (potentially)
1.5	Multiple organisations	Does the project involve multiple organisations, whether they are public sector agencies i.e. joined up government initiatives or private sector organisations e.g. outsourced service providers or business partners?	Yes
1.6	Data	Does the project involve new process or significantly change the way in which personal data/special categories of personal data and/or business sensitive data is handled?	Yes
1.7	Data	Does the project involve new or significantly changed handling of a considerable amount of personal data/special categories of personal data and/or business sensitive data about each individual in a database?	Yes
1.8	Data	Does the project involve new or significantly change handling of personal data/special categories of personal data about a large number of individuals?	No
1.9	Data	Does the project involve new or significantly changed consolidation, inter-linking, cross referencing or matching of personal data/special categories of personal data and/or business sensitive data from multiple sources?	Yes
1.10	Data	Will the personal data be processed out of the U.K?	No
1.11	Exemptions and Exceptions	Does the project relate to data processing which is in any way exempt from legislative privacy protections?	No

Q	Category	Screening question	Yes/No
1.12	Exemptions and Exceptions	Does the project's justification include significant contributions to public security and measures?	No
1.13	Exemptions and Exceptions	Does the project involve systematic disclosure of personal data to, or access by, third parties that are not subject to comparable privacy regulation?	No

The purpose of the screening questions is to confirm that the data protection laws are being complied with, or highlights problems that need to be addressed. It also aims to prevent problems arising at a later stage which might impede the progress or success of the project.

Answering "Yes" to any of the screening questions above represents a potential Information Governance (IG) risk factor, please proceed and complete the next section.

Step 2: Identify the need for a DPIA									
2.1	Is this a new or changed use of personal data/special categories of personal data and/or business sensitive data that is already processed/shared??					New/Changed			
						Changed			
2.2	What data will be processed/shared/viewed?								
	<u>Personal Data</u>								
	Forename		Surname		Date of Birth		Age	**X**	Gender	**X**
	Address		Postal address		Employment records		Email address		Postcode	**X**
	Other unique identifier (*please specify*) **LA Child ID** **Placement URN**		Telephone number		Driving license number		NHS No		Hospital ID no	
	Other data *(Please state):*								

<u>Pan-London Sufficiency Analysis</u>

Data Subjects

Children and Young People who are considered 'looked-after' within the meaning of the Children's Act 1989 in the six years prior to the analysis being commissioned by ALDCS.

Young People who were previously considered 'looked-after' within the meaning of the Children's Act 1989 who had their 17th to 25th birthday in the six years prior to the analysis being commissioned by ALDCS

What Data

The data being used is pseudonymised administrative data collected in the delivery of services, for the purposes of statutory reporting and the purposes noted above. Additionally, to expand the scope of the data available for analysis, a novel dataset, also pseudonymised, is being processed.

		The definitive list of fields is attached as Appendix 2 – 'The Data Extracts and Their Scope'. In summary, it covers:
		• Unique identifiers (e.g. LA child ID)
		• Demographics (e.g. gender, age, ethnicity)
		• Child Looked After (CLA) Episodes (e.g. start, end, categories of need or abuse)
		• CLA Placements (e.g. start and end, provider, postcodes)
		• Needs data related to CLA, including data on disabilities, special educational needs, and factors identified at assessment, including concerns relating to the child or their family about:
		  ○ substance misuse
		  ○ domestic violence
		  ○ mental health
		  ○ sexual exploitation
		  ○ trafficking
		  ○ gangs
		  ○ self-harm
		  ○ abuse
		  ○ female genital mutilation
		  ○ criminal exploitation
		• Data about the costs of placements for CLA (costs and type of contract used in procurement)
		The additional data collected relating to disabilities and factors identified at assessment is already shared by LAs for the LIIA project for another use case, "Children's Services Insights", with the analysis conducted by the LIIA team at London Councils. Documentation relating to this use case can be found in the DPIA carried out for this use case and in Schedule 3 of the DPA.
		
		Inclusion of Personal Data
		
		For at least some subjects, data will cover:
		• Gender - required for equalities monitoring

		• Location based data (degraded postcode identifying clusters of c. 3,000 households, collected for children in care placements) - required to understand links between area characteristics (e.g. inferred socio-economic status; gang territories) and needs/outcomes, to answer questions such as whether some areas might be under/over-served, and whether children placed 'out of area' have worse outcomes Other Unique Identifiers – per-LA child ID is captured to assist with checking data quality **Frequency of Collection** SSDA903 is collected annually. The placement costs dataset will be collected quarterly.

<u>Special Categories of Personal Data</u>

Racial or ethnic origin	**X**	Political opinion		Religious or philosophical beliefs							
Trade Union membership		Physical or mental health or condition									**X**
Sexual life or sexual orientation		Social service records		Child protection records							**X**
Sickness forms		Housing records		Tax, benefit or pension records		Adoption records					
DNA profile		Fingerprints		Biometrics		Genetic data					
Proceedings for any offence committed or alleged, or criminal offence record											

Other data *(Please state):*	**Inclusion of Special Category Data** Yes – for at least some data subjects, the data includes: • Racial or ethnic origin – required for equalities monitoring • Mental and physical health (via a 'need code' applied following assessment') – required to understand needs and to identify good practice in meeting them • Disabilities data • Special educational needs • factors identified at assessment, including concerns relating to the child or their family about: ○ substance misuse

		○ domestic violence ○ mental health ○ sexual exploitation ○ trafficking ○ gangs ○ self-harm ○ abuse ○ female genital mutilation ○ criminal exploitation

Will the dataset include clinical data? (please include)	No
Will the dataset include financial data?	Yes
Description of other data processed/shared/viewed?	

2.3	Business sensitive data	Y/N	Details
	Financial	Yes	Placement Cost dataset collected quarterly
	Local Contract conditions	No	N/A
	Operational data	No	N/A
	Notes associated with patentable inventions	No	N/A
	procurement/tendering information	No	N/A
	Customer/supplier information	No	N/A

Decisions impacting:	One or more business function	Y/N
		No
	Across the organisation	No
Description of other data processed/shared/viewed (if any).		

<table>
<tr><th colspan="4">Step 3: Describe the sharing/processing</th></tr>
<tr><td rowspan="5">3.1</td><td colspan="2" rowspan="2">List of organisations/partners involved in sharing or processing personal/special categories personal data? If yes, list below</td><td>Yes/No</td></tr>
<tr><td>Yes</td></tr>
<tr><td>Name</td><td>Controller or Processor?</td><td>Completed and compliant with the IG Toolkit or Data Security and Protection (DSP) Toolkit</td></tr>
<tr><td colspan="2">Yes / No</td></tr>
<tr><td>Local Authorities (Signatories to the Child Level DPA for London boroughs)</td><td>Controller</td><td>Yes (generally)</td></tr>
<tr><td></td><td>London Councils</td><td>Processor</td><td>TBC</td></tr>
<tr><td rowspan="2">3.2</td><td colspan="2" rowspan="2">If you have answered yes to 3.1 is there an existing Data Processing Contract or Data Sharing Agreement between the Controller and the Processor?</td><td>Yes/No</td></tr>
<tr><td>Yes. This will be covered in the Child Level DPA for London Boroughs</td></tr>
<tr><td>3.3</td><td colspan="2">Has a data flow mapping exercise been undertaken?

If yes, please provide a copy at Annex 2 below, if no, please undertake one</td><td>See attached Data Flow map in Appendix 1</td></tr>
<tr><td rowspan="2">3.4</td><td colspan="2" rowspan="2">Does the project involve employing contractors external to the Organisation who would have access to personal or special categories of personal data?</td><td>Yes / No</td></tr>
<tr><td>No</td></tr>
<tr><td rowspan="2">3.5</td><td colspan="3">Describe in as much detail why this information is being processed/shared/viewed?
(For example Direct Patient Care, Statistical, Financial, Public Health Analysis, Evaluation. See NHS Confidentiality Code of Practice Annex C for examples of use)</td></tr>
<tr><td colspan="3"><u>Sharing Children’s Services Insights</u>

The project exists to help the London Directors of Children’s Services to deliver their statutory obligations under section 17 of the Children’s Act 1989 “to safeguard and promote the welfare of children in need in their area” and section 149 of the Equality Act 2010 to deliver the “public sector equality duty”. It aims to do this by:

a. Accelerating service improvement – by enabling the identification and prioritisation of opportunities for improvement; the identification of good practice in other boroughs; and the increase in information exchange between London Boroughs and providers of care placements to facilitate improvement in market conditions;</td></tr>
</table>

	b. **Monitoring equalities** – by enabling comparative analysis of the odds of key outcomes for children in care (e.g. being placed out of Borough, being placed in unsuitable placements leading to placement breakdown) being used with families of different ethnicities; Through this, the project aims to benefit vulnerable children, young people, and their families by improving the quality of services which safeguard them from harm and help them to develop to their full potential.

Step 4: Assess necessity and proportionality			
4.1	Lawfulness for Processing/sharing personal data/special categories of personal data?		
	UK GDPR	DPA 2018	Other Lawful Basis
	Personally Identifiable Data		
	UK GDPR Article 6(1)(e) '…for the performance of a task carried out in the public interest or in the exercise of official authority…'	The DPA section 8(c) – "the exercise of a function conferred on a person by an enactment or rule of law", specifically the public tasks are: • "to safeguard and promote the welfare of children within their area who are in need" – a statutory duty under the Children's Act 1989 • To deliver the "public sector equality duty" outlined in the Equalities Act 2010 including the needs to "advance equality of opportunity between persons who share a relevant protected characteristic and persons who do not share it" and to "take steps to meet the needs of persons who share a relevant protected characteristic that are different from the needs of persons who do not share it"	

	Special Categories of Personally Identifiable Data			
	UK GDPR Article 9(2)(g) '…processing is necessary for reasons of substantial public interest…'	The DPA Schedule 1 Part 2 section 2 "'Safeguarding of children and individuals at risk' and 'Equality of opportunity or treatment' satisfying DPA section 10 (3)		
4.2	Will the information be processed/shared electronically, on paper or both?		Electronic	X
			Paper	
4.3	How will you ensure data quality and data minimisation?			

Data Quality

Data quality checks are factored in to the ETL process, specifically step 1 (ii) of the process which is common to all use cases as detailed below:

- Each Borough uploads data, including personal sensitive data, onto a private, borough-specific folder in the London Datastore.
- Scripts provided by the LIIA team then processes this data on the London Datastore in three ways:
    1. Preparation of single Borough's data for analysis, including:
        i. Checking whether agreed pseudonymisation and data minimisation has been done prior to sending, and implementing it if not (e.g. deletion of fields not required; degrading highly disclosive data such as postcodes and dates of birth);
        ii. Assessment of data quality (missing values, logically inconsistent values);
        iii. Transformation of data to conform to a common schema.
    2. Loading the prepared data for all Boroughs into a pan-London database;
    3. Creating extracts from that database for analytical purposes specific to the use case.
- The single-Borough output of step 1 are made available back to the Borough, free for them to use for their own internal analysis
- The extracts created in step 3 are made available to an approved analyst (either at London Councils or a named sub-processor approved by the DPOs) to produce the pan-London analyses specific to the use case

Data Minimisation

Every effort is made to restrict the sharing of sensitive data only when proportionate to the benefits of analysis. Accordingly, the following data minimisation occurs for this process:

- Removal of modules relating to adopted children from SSDA903. This eliminates all data about children who have been adopted (unless they have had other interactions with Children's Services e.g. been in care prior to adoption)

- Removal of information about the motherhood status and dates of birth of children born to looked after children
- Removal of unique pupil number, which prevents the possibility of the data extract being linked to other datasets that could lead to the identification of individuals
- Degradation of date of birth to month of birth
- Degradation of all postcodes to postcode district

In this, we are balancing the desire for *data minimisation* with the practical need not to have to ask the LAs for new data extracts each time we specify a question. This is a legitimate trade-off to consider - ICO guidance explaining the application of the Data Protection Act 2018 is clear that *"You must not collect personal data on the off-chance that it might be useful in the future. However, you may be able to hold information for a foreseeable event that may never occur if you can justify it."*

Our approach is to request a single annual data submission from each LA (Annex A may be more frequent, depending on needs communicated by ALDCS) – making working with the project viable for them in terms of workload, but to then:

1. **Apply minimisation in our specification of the data request**– removing all data which we do not believe we are likely to need for our purposes, and degrading data which is more specific than we need it to be. The precise data request we are making – including which datasets, fields, and periods, is attached as Appendix 2.

   Specifically:

   a. Removing a large number of individuals from our scope by:

      i. removing data on children who have been adopted and who have not been considered a child in need or accessed other children's social care services;

      ii. removing data fields describing adopters;

      iii. removing data fields describing the children of looked after children

      iv. restricting the analysis to individuals who are in scope during a six-year period – chosen because previous analysis has shown to be the shortest period we can use and still be able to conduct journey-based analysis and be confident in it.

   b. Removing data fields from our scope where we are unlikely to require them for the types of analyses which serve our purposes – e.g. information about reviews of looked after children, information about health checks.

2. **Protect anonymity** – Degrading indirect identifiers which have a greater level of specificity than we believe we are likely to need – e.g. removing unique pupil number, degrading postcode to postcode district (a c. 540x reduction in specificity) and date of birth to month of birth and school year (a c. 30x reduction in specificity).

3. **Incorporate Minimisation into our ETL Process** – essentially setting the code which prepares the data ready for use to check that minimisation has been applied by the sender, and then to apply it automatically if it has not – deleting and degrading data as appropriate before it is loaded into the database for analysis.

4. **Add an additional layer of minimisation between the prepared data, and the data being analysed –** by performing all individual analyses on specially created extracts which only contain the data necessary for that query, rather than on the full dataset. If the operation scales, this allows us to restrict the number of people who ever have access to the full dataset to a small number of staff in the LIIA platform team.
5. **Implement a Robust Data Registration and Destruction Process.** A register of all project data assets will be maintained. The scope of necessary data will be reviewed every six months, and any data falling outside it will be securely destroyed.

Controlling Function Creep

A key risk here is that having authorised processing for one purpose, the unit then begins to stretch and eventually break the agreed scope.

To control this:

- All lines of enquiry will need to be agreed with by the ALDCS through their regular meeting, or by their nominated representative (currently Ben Byrne, Strategic Lead for the London innovation and Improvement Alliance);
- Local Authority DPOs will have the option to subscribe to a regular update letting them know what lines of enquiry are being pursued and how they relate to the purpose, and we will maintain regular contact with IGfL to allow them to scrutinise the work.
- A summary of each enquiry (although not the outputs) will be publicly logged on the LIIA website, with the purpose it relates to.

  All new use cases, that change the nature of the data being shared and/or the purpose of its processing will be subject to additional approvals from DPOs through the additions of new Schedules to the DPA between London Councils and London Boroughs and the creation of new guides for DPIAs, such as this one

4.4	Have individuals been informed about the proposed use of their personal or special categories of personal data? *For example, do the organisations/partners listed in section 3.1 have updated Fair Processing Notice available to patients on their websites?*	Yes/No TBC
	Participating boroughs will need to review their fair processing notices as per the guidance in Appendix 3	
4.5	How will you help to support the rights of individuals?	
	Processor obligations are addressed in 7.5 of the DSA	
4.6	Are arrangements in place for recognising and responding to Subject Access Requests (SARs)? *If no, please describe how rights are exercised. If Yes, please detail.*	Yes/No Yes

	Each Local Authority (Controller) will be responsible for managing Subject Access Request through their internal corporate procedures. Processor responsibilities to assist with Data Subject Rights requests is addressed in 7.5 of the DSA.	
4.7	Will the processing of data include automated individual decision-making, including profiling? *If yes, please outline the profiling processes, the legal basis underpinning the process, and the rights of the data subject*	Yes/No No
There will be no machine learning, no automated decision making, and no attempts to support decision making about an individual case.		
4.8	Will individuals be asked for consent for their information to be processed/shared? *If no, list the reason for not gaining consent e.g. relying on other lawful basis, consent is implied where it is informed.*	Yes/No No
	Relying on other lawful basis	
4.9	As part of this work is the use of Cloud technology being considered either by your own organisation or a 3rd party supplier? If so please complete the cloud security questionnaire and add as an annex or state below why it is not required.	Yes/No Yes
	Social Care Network data security protocols **Penetration Tests** The most recent, independent Pen Test was completed in 2022. The test has proven more than satisfactory. **2 Factor Authentication** By default, SCN's CHARMS application uses a 2-step process to authenticate users, involving a username/password combo followed by selected characters from a passphrase. These mimic the way banks in the UK allow access to online applications. **Software Development Lifecycle** SCN's software is developed using C# and ASP.Net and runs on Windows Servers using Microsoft SQL Server as the data store. As code is written it is checked by VeraCode, a static code analysis tool which identifies any vulnerabilities that may have been written into the codebase by developers. Security Testing of beta releases are undertaken by the security Architect. Internal Pen Testing is undertaken at every major release by SCN. **Defence in Depth** SCN's philosophy is defence in depth. All data is encrypted using TLS 1.2 to servers, a Web Application Firewall analyses the requests to reject any injection or client-side attacks, and IIS is set to implement the strongest security available. Code is scanned by VeraCode, ASP.Net security is enabled, all internal traffic is sent over HTTPS, and all the data in the database is	

	encrypted, both in transit and at rest. Transparent Data Security, TDS, in SQL Server is used to achieve this. **Backups** Backups are taken every day and managed by the cloud provider. This ensures that there is no member of staff at SCN who could delete backups. Backups are available for 6 months. Transaction logging is used to enable any problems with data after the last backup and before the next. **Multi-Tenanted Solution** SCN's applications are delivered as off the shelf, Software as a Service solutions, SAAS. Customers have their own Website and Database implementation on our infrastructure which is provided by UKFast. Data is stored in two datacentres, on either side of the city of Manchester, to ensure availability. All hardware infrastructure is mirrored in each datacentre. One datacentre acts as the failover - all activity in the prime datacentre is immediately updated to the failover datacentre in real time. **Certifications** SCN is Cyber Essentials Plus certified and is starting the ISO 27001 certification process. The infrastructure provider UKFAST is ISO 27001 certified and also ISO 27018, ISO 9001 and ISO 22301. **Data Storage** All data is stored in the UK and backed up in UK. **Availability and resilience** 100% Connectivity Availability - This is access to the infrastructure 99.5 Application Availability - This is access to the application.
4.10	Where will the data will be stored? *Examples of Storage include bespoke system (e.g. EPR, Emis & other clinical systems, SharePoint, data repository, Network Drives, Filing cabinet (office and location), storage area/filing room (and location) etc.*
	Data Storage Pan-London extracts are accessed by secure download from the LIIA data platform to a MS Azure data warehouse hosted by Social Care Network (SCN). The extract is a download of the full dataset. Data is stored in a warehouse independent of any other data sources, with access controlled by Commissioning Alliance and managed by Azure Active Directory. Data stored in SCN's data warehouse is encrypted using TLS 1.2, while in transit and at rest. All data is stored and backed up in the UK with accessed managed through 2-factor authentication. SCN are Cyber Essentials Plus certified. -
4.11	Data Retention Period *How long will the data be kept?*

	Retention and Deletion The intention is to store data going back up to six years from the point of analysis. Previous analysis finds that six years is the minimum for good quality journey analysis (an analytical approach we expect to employ), but that having more than six years' data does not materially help to answer new questions or answer existing questions with greater certainty or granularity. The DPA places a duty on the processor (and any sub-processors) to securely destroy any data outside this scope (e.g. if over time we come to have five years' data) and to destroy all data at the end of the programme or on request of the data controller. The processor is also obliged to evidence this destruction to the controller if requested.	
4.12	Will this information being shared/processed outside the organisations listed above in question 3? If yes, describe who and why:	Yes/No Yes
	The DPAs between the Controllers and the Processor will contain a schedule listing approved sub-processors, and a stipulation that approval has to be sought from the controllers to add further sub-processors. **Additional Sub-Processors** Social Finance Ltd, a not-for-profit data and strategy specialist is providing Python code to prepare the data for analysis. This code is QAd and tested by the LIIA data platform before integration to platform processes. Social Finance are also training the LIIA team, including analysts at LBWF, to maintain and extend that code. Social Care Network are providing the MS Azure data warehouse that will host the pan-London data extract (individual-level data) and the Power BI report (aggregated to Borough level). Simpsons Associates will be assisting Commissioning Alliance with the analysis of the data extract and creation of the Power BI report. Access to the MS Azure warehouse by staff of Simpsons Associates will be managed by Commissioning Alliance. All Simpson Associates consultants have Security Check (SC) clearance. Simpsons Associates hold ISO 270001 and ISO 9001, Cyber Essentials and Cyber Essentials Plus certificates and regularly work with LAs, police forces and healthcare trusts. **Ensuring the Processor Applies the Agreed Controls** The DPAs between Controllers and the Processor give the Controller right to audit the Processor's compliance with conditions for processing. The DPAs also require the Processor to agree equivalent protections and audit rights from any sub-processors. The DPAs between the Controllers and the Processor will contain a schedule listing approved sub-processors, and a stipulation that approval has to be sought from the controllers to add further sub-processors.	

	Access to Analysis Power BI analysis conducted in Power BI report, with Boroughs identifiable. The report is shared via individual link to named individuals at all London Boroughs. Access to the report is managed by analysts at Commissioning Alliance. Links shared with individuals will allow access only to that individual. Dissemination of these aggregate analyses to individuals from organisations other than the London Boroughs, such as with placement providers, will be managed in line with Memoranda of Understanding between the ACDCS and LIIA Board.

Step 5: Information Security Process							
5.1	Is there an ability to audit access to the information? *If no, please provide a reason why this is not required. If yes, please describe auditing.*						Yes/No TBC
	Regular security audits by qualified third-party organisations will be carried out.						
5.2	How will access to information be controlled?						
	Pan-London extracts are accessed by secure download from the LIIA data platform to a MS Azure data warehouse hosted by Social Care Network (SCN). The extract is a download of the full dataset. Data is stored in a warehouse independent of any other data sources, with access controlled by Commissioning Alliance and managed by Azure Active Directory. Data stored in SCN's data warehouse is encrypted using TLS 1.2, while in transit and at rest. All data is stored and backed up in the UK with accessed managed through 2-factor authentication. SCN are Cyber Essentials Plus certified. Data in London Borough of Ealing's data warehouse will be joined to data from Ofsted produced reports about placement providers, including data relating to placement cost and quality. This joining will be performed on information relating to care placements, and not on information relating to individual children. All joining will be performed in the Azure data warehouse hosted by SCN. The joining will take place using the following fields from the children's data (SSDA903 or placement costs dataset): - placement URN - placement postcode (at postcode district level) The purpose for adding this additional data to the SSDA903 data is to allow analysis of the outcomes of episodes of care using child-level information from the SSDA903 alongside the quality and cost of placement providers using placement-level information from the Ofsted reports						
5.3	What roles will have access to the information? (list individuals or staff groups)						
	Analysis to aggregate individual level data is conducted in Power BI, hosted by Social Care Network. Only analysts from the Commissioning Alliance, Social Care Network and Simpsons Associates working on the project will have access to individual-level data.						
5.4	What security and audit measures have been implemented to secure access to and limit use of personal data/special categories of personal data and/or business sensitive data?						
	Username and password		Smartcard		key to locked filing cabinet/room		
	Secure 1x Token Access		Restricted access to Network Files				**x**
	Other: *Provide a Description Below*:						

5.5	Is there a documented System Level Security Policy (SLSP) for this project? If yes, please add a copy as an annex. *SLSP refers to the architecture, policy and processes that ensure data and system security on individual computer systems. It facilitates the security of standalone and/or network computer systems/servers from events and processes that can exploit or violate its security or stature.*	Yes/No TBC
5.6	Are there Business Continuity Plans (BCP) and Disaster Recovery Protocol for the proposed/existing system or process? *Please explain and give reference to such plan and protocol*	Yes/No Yes
	Several safeguards are in place to ensure resilience of the data storage, leading to the repellence of previous denial-of-service (DoS) attacks. These include annual penetration tests. Data protection from loss and lack of availability on AWS is covered by their business continuity and disaster recovery policy.	

5.7	Is Mandatory Staff Training in place for the following?	Yes/No	Dates
	• Data Collection:	London Data Platform staff with access to the systems are all accredited under the ONS Secure Researcher training. LIIA to confirm re: sub processors.	
	• Use of the System or Service:		
	• Information Governance:		

5.8	Are there any new or additional reporting requirements for this project? *If no, skip to 5.9. If yes, provide details below.*	Yes/No No
	• What roles will be able to run reports?	
	Power BI analysts at London Borough of Ealing	
	• What roles will receive the report or where will it be published?	
	Power BI analysis conducted in Power BI report, with Boroughs identifiable. The report is shared via individual link to named individuals at all London Boroughs. Access to the report is managed by analysts at Commissioning Alliance. Links shared with individuals will allow access only to that individual. Dissemination of these aggregate analyses to individuals from organisations other than the London Boroughs, such as with placement providers, will be managed in line with Memoranda of Understanding between the ACDCS and LIIA Board.	
	• Will the reports be in person-identifiable, pseudonymised or anonymised format?	

	Data will be aggregated and shared in reporting such that no individuals are identifiable, though there is a risk of re-identification of individuals due to small aggregations in some analyses. Boroughs will be identifiable in the shared analysis. The analysis will be shared among DCSs in London Boroughs and additional third parties, such as placement providers, subject to Memoranda of Understanding to be agreed between ALDCS and the LIIA Board. The reason for the additional sharing in this case is to provide the outputs of analysis to both the creators (London Boroughs) and suppliers (placement providers) of demand for care placements. As a key purpose of this use case is to facilitate change in the commissioning of placements, providing this two-way exchange of information is deemed a vital component of the analysis.	
	• Will the reports be in business sensitive or redacted format (removing anything which is sensitive) format?	
	N/A	
5.9	Have any Information Governance risks been identified relating to this project? If yes, the final section must be completed.	Yes/No Yes

Step 6: Identify and Assess Risks

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary.	Likelihood of harm	Severity of harm	Overall risk
Data Breach	High	Low	Medium
Data Subjects Unaware of or Not Understanding Processing	Low	High	Medium
Scope Creep takes analysis beyond legitimate purpose	Medium	Medium	Medium
Reduced Trust in Data Controllers if Project is Misconstrued as involving automated decision making or facilitating new level of surveillance of individuals	Medium	Low	Low

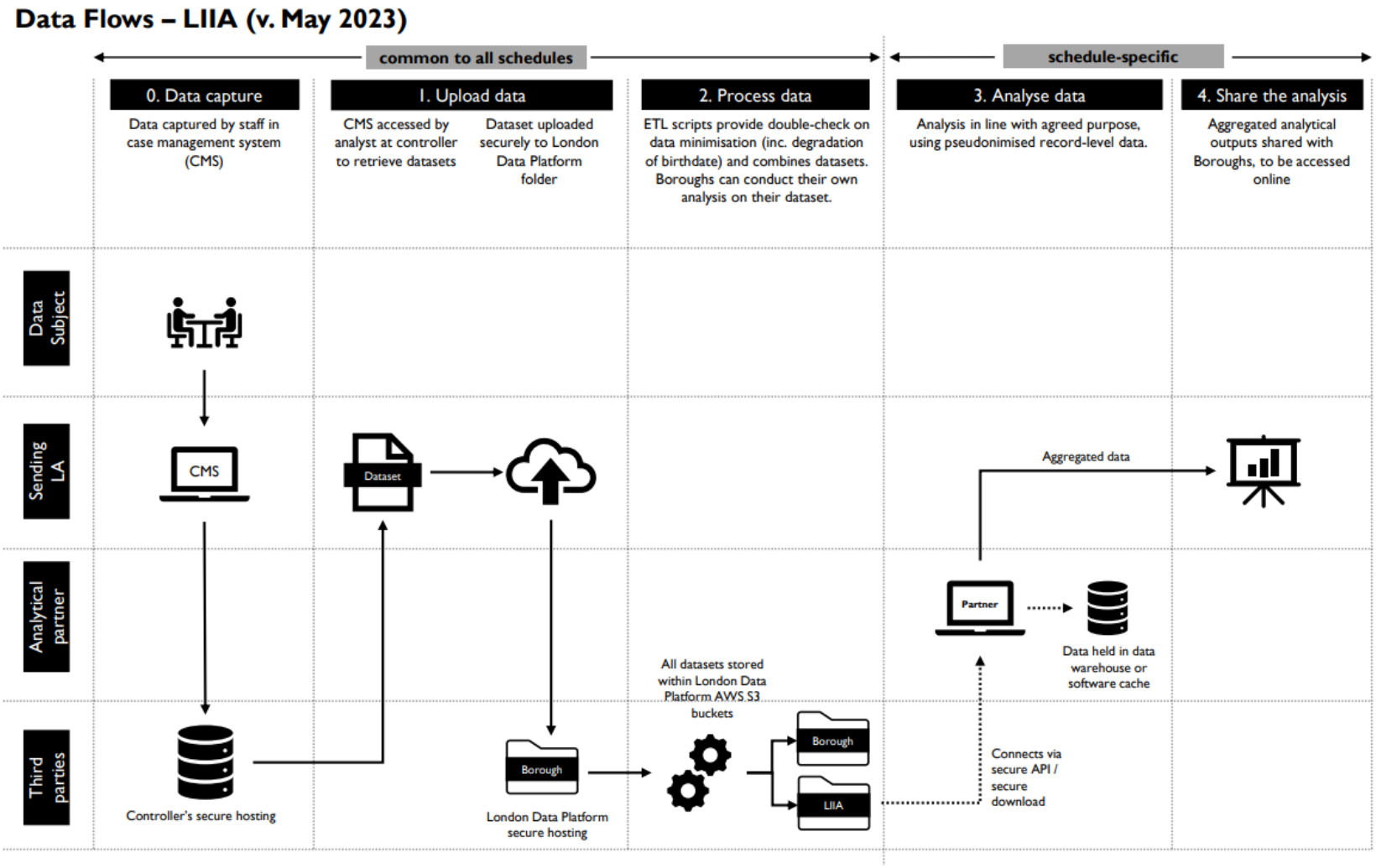
Step 7: Identify Measures to reduce risk

Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in step 6

Risk	Options to reduce or eliminate risk	Effect on risk	Residual risk	Measure approved
Data Breach	Data minimisation as outlined above to reduce impact. Technical, physical and process protections legally mandated and auditable – to reduce probability	Reduced	Low-Medium	Yes
Data Subjects Unaware of or Not Understanding Processing	Review privacy notices prior to going live and amend if required Public communication about the project – specifically addressing this.	Reduced	Low	Yes
Scope Creep takes analysis beyond legitimate purpose	Enhanced governance and transparency as outlined above	Reduced	Low	Yes
Reduced Trust in Data Controllers if Project is Misconstrued as involving automated decision making or facilitating new level of surveillance of individuals	Public communication about the project – specifically addressing this. Reduces likelihood that one person misconstruing the purpose spreads.	Reduced	Low	Yes

Step 8: Sign off and record outcomes		
Item	**Name/date**	**Notes**
Measures approved by:	Joe Luxton 24/08/2023	
Residual risks approved by:	Joe Luxton 24/08/2023	
DPO advice provided:	Leila Ridley 24/08/2023	
Summary of DPO advice: I am happy to approve this processing – Leila Ridley		
DPO advice accepted or overruled by:	**N/A**	If overruled, you must explain your reasons
Comments:		
Consultation responses reviewed by:	**N/A**	If your decision departs from individuals' views, you must explain your reasons
Comments:		
This DPIA will kept under review by:	The DPIA will be reviewed by the respective DPOs of each organisation when required	The DPO should also review ongoing compliance with DPIA

Appendix 1: Data Flow



Appendix 2: Data Extracts and their Scope



Pan London
Sufficiency Analysis D

Appendix 3: Note on privacy notices

Most Boroughs will already have privacy notices that provide sufficient information about the processes described here. However, for Boroughs that wish to provide specific information about the project in their Children's Services privacy notice, we recommend the following wording to be added:

London Innovation and Improvement Alliance

The LIIA project is a pan-London initiative to address important issues for children in London that can only be answered by examining London's data as a whole. By creating a secure platform where local authorities can share data with each other and other analysts, the project will improve the breadth and quality of data analysis available to local authorities in London.

Data agreements are in place to ensure that:

- data is pseudonymised to reduce the risk of individuals being identified e.g. "Tim Smith, DOB 17th Jan 2000, postcode SW14 2JU" becomes "ID 58095927, DOB Jan 2000, postcode SW14"
- under no circumstances will the data be used for any automated decision making
- all data is transferred, handled or stored in accordance with the Data Protection Act
- access to the data is confined to the smallest possible number of people to produce the analysis
- all data is destroyed after six years
-

You have the right to object to your data being used this way. If you wish to exercise it then please contact <*insert details*>.

Appendix 4: DPO’s guide to Data Protection Impact Assessment (supporting documentation used to complete this DPIA)



LIIA DPO's guide to
DPIA - Pan-London Su

Appendix 5: Social Care Network data security protocols

Penetration Tests

The most recent, independent Pen Test was completed in 2022. The test has proven more than satisfactory.

2 Factor Authentication

By default, SCN's CHARMS application uses a 2-step process to authenticate users, involving a username/password combo followed by selected characters from a passphrase. These mimic the way banks in the UK allow access to online applications.

Software Development Lifecycle

SCN's software is developed using C# and ASP.Net and runs on Windows Servers using Microsoft SQL Server as the data store. As code is written it is checked by VeraCode, a static code analysis tool which identifies any vulnerabilities that may have been written into the codebase by developers. Security Testing of beta releases are undertaken by the security Architect. Internal Pen Testing is undertaken at every major release by SCN.

Defence in Depth

SCN's philosophy is defence in depth. All data is encrypted using TLS 1.2 to servers, a Web Application Firewall analyses the requests to reject any injection or client-side attacks, and IIS is set to implement the strongest security available. Code is scanned by VeraCode, ASP.Net security is enabled, all internal traffic is sent over HTTPS, and all the data in the database is encrypted, both in transit and at rest. Transparent Data Security, TDS, in SQL Server is used to achieve this.

Backups

Backups are taken every day and managed by the cloud provider. This ensures that there is no member of staff at SCN who could delete backups. Backups are available for 6 months. Transaction logging is used to enable any problems with data after the last backup and before the next.

Multi-Tenanted Solution

SCN's applications are delivered as off the shelf, Software as a Service solutions, SAAS. Customers have their own Website and Database implementation on our infrastructure which is provided by UKFast. Data is stored in two datacentres, on either side of the city of Manchester, to ensure availability. All hardware infrastructure is mirrored in each datacentre. One datacentre acts as the failover - all activity in the prime datacentre is immediately updated to the failover datacentre in real time.

Certifications

SCN is Cyber Essentials Plus certified and is starting the ISO 27001 certification process. The infrastructure provider UKFAST is ISO 27001 certified and also ISO 27018, ISO 9001 and ISO 22301.

Data Storage

All data is stored in the UK and backed up in UK.

Availability and resilience

100% Connectivity Availability - This is access to the infrastructure
99.5 Application Availability - This is access to the application.