Data Protection Impact Assessment

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	Updated to include changes to retention of children's data and references to data platform	22/06/2023	Joe Luxton
1.4	Reviewed and approved by Islington IG Panel	29/06/2023	Joe Luxton

Project / Work Stream Name	LIIA Project: Children's Services Insights	
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)	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 based 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. After a successful pilot with five boroughs (Enfield, Islington, Merton, Wandsworth, Richmond and Kingston), the LIIA team is now expanding the project with all 32 London Boroughs and the City of London Corporation. The project went live with the following timetable: September 2022 – All LAs signed the information governance documents November 2022-April 2023 – SSDA903 returns for 6 years (2016/2017 – 2021/2022) shared by all LAs for Children's Services Insights and Pan-London Sufficiency work January – August 2023 – Proposal to develop LIIA/ALDCS secure data platform authorised and work commenced, implementation date August 2023 **Contractual Arrangements** The LIIA team developed a common Data Processing Agreement (DPA) and contract to be used between each Data Controller, and the

	Data Processor. DPAs were also agreed between the Data Processor (London Councils) and all sub-processors involved in processing. These agreements developed in consultation with the Information Governance Group for London (IGfL). DPOs should note that this is a revision of this project which was itself 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. The DPA was previously developed for a project which was recently selected as an ICO case study for good practice in sharing sensitive 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. These 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 supports the use case: **Sharing Children's Services Insights**, and corresponds to Schedule 3 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 in the LIIA data platform. • Scripts provided by the LIIA team then processes this data on the LIIA data platform 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 **<u>Use case: Sharing Children's Services Insights</u>** **Context** This use case for the LIIA Project involves aggregating and sharing Boroughs' data from three Children's Services datasets that are produced as part of Boroughs' statutory duties (Children in need 'CiN' Census, Children looked after return 'SSDA903', Ofsted Annex A). The analysis, to be conducted by LIIA analysts hosted at London Borough of Waltham Forest (LBWF), aims to: • accelerate service improvement, by enabling the identification and prioritisation of opportunities for improvement, and the identification of good practice in other Boroughs • monitor inequalities, by enabling comparative analysis of the odds of key social work interventions (e.g. Section 47 Investigations, Child Protection Plans, Care Orders) being used with families of different ethnicities Data will be aggregated and shared such that no individuals are identifiable. Information will be analysed at the Borough level, with Boroughs identified in the shared analysis. The analysis will be shared only among DCSs in London Boroughs. **Use Case Specific Data Processing** • Pan-London extracts for each of the three datasets are accessed by LIIA analysts at LBWF via a secure bearer token to Power BI hosted by LBWF

	• Individual-level data are held in cache in Power BI, accessible only by LIIA analysts at LBWF • Descriptive analysis of event frequencies and breakdown by Borough, age group, ethnicity group, with comparison by Borough conducted in Power BI report • Power BI report shared with DCSs through personal, secure link. Data in report can be accessed at Borough-level only
Implementation Date:	Estimated 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 2 – 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 and families in contact with children's social 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 at the start of the project might have been considered novel: 1. Sending their data to a third party (LBWF) to be processed instead of doing it in-house (although we note that the same data is routinely provided to the third parties DfE and to Ofsted to 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 both DfE and Ofsted are known to 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 Ofsted) and combined with data from other LAs in order to conduct very similar processing for a very similar purpose, this is not novel processing. In addition, the project has already conducted this type of processing using the SSDA903 data provided by LAs and which has started to be published for DCS & LA use. However, there is sufficient ambiguity about whether that removes novelty to warrant consideration of a DPIA and the material changes to the processing platform merit this revision and review. Given the 'once for London' approach central to the LIIA project, and the standardisation of processes and data flows that is established, we believe it is legitimate for a full DPIA to be conducted by only one Borough or a small IGfL working group, on behalf of all others, and that summary DPIAs are sufficient for all others.

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?	Yes
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?								
	Personal Data								
	Forename		Surname		Date of Birth	**X**	Age	**X**	Gender	**X**
	Address		Postal address		Employment records		Email address		Postcode	**X**
	Other unique identifier (*please specify*) Unique Pupil Number LA Child ID		Telephone number		Driving license number		NHS No		Hospital ID no	
	Other data *(Please state):*								

Data Subjects

Children and Young People who meet any of the following conditions in the eleven years prior* to the analysis being commissioned by ALDCS:

- the subject of an initial contact with Children's Services
- Referred to Children's Social Care
- Considered 'Child in Need' under the section 17 definition
- Subject to a Child Protection Plan or Care Order
- Been eligible for services by virtue of being a Care Leaver

Practically, this is intended to include children and young people included in the Cin Census or SSDA903 annual returns and those included in the 'Annex A' report (excluding children placed

		for adoption and the adults who adopt them; in Annexes 10 & 11). 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. The definitive list of fields is attached as Appendix 2 – 'The Data Extracts and Their Scope'. In summary, it covers: • Unique identifiers (e.g. unique pupil number) • Demographics (e.g. gender, age, ethnicity) • Referrals and assessments (e.g. dates, outcomes) • Safeguarding events (e.g. dates and outcomes of S47 enquiries) • Child in Need (Cin) episodes (e.g. start, end, need code) • Child Protection (CP) and Child Looked After (CLA) Episodes (e.g. start, end, categories of need or abuse, types of support provided) • CLA Placements (e.g. start and end, provider, postcodes) **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 – unique pupil number and per-LA child ID are captured to assist with checking data quality, and re-linking data across Cin Census, the SSDA903, and Annex A. **Frequency of Collection** Cin Census and SSDA903 are annual collections. Annex A will most likely be collected quarterly – this will be determined based

		on whichever is the least frequent collection suitable to answer the questions the DCSs agree.								
	Special Categories of Personal Data									
	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):*	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								
	Will the dataset include clinical data? (please include)	No								
	Will the dataset include financial data?	No								
	Description of other data processed/shared/viewed?									

2.3	Business sensitive data	Y/N	Details
	Financial	No	N/A
	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).			

Step 3: Describe the sharing/processing			
3.1	List of organisations/partners involved in sharing or processing personal/special categories personal data? *If yes, list below*		Yes/No
			Yes
	Name	Controller or Processor?	Completed and compliant with the IG Toolkit or Data Security and Protection (DSP) Toolkit
			Yes / No
	Local Authorities (Signatories to the Child Level DPA for London boroughs)	Controller	Yes (generally)
	London Councils	Processor	TBC
3.2	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?		Yes/No
			Yes. This will be covered in the Child Level DPA for London Boroughs
3.3	Has a data flow mapping exercise been undertaken? If yes, please provide a copy at Annex 2 below, if no, please undertake one		See attached Data Flow map in Appendix 1
3.4	Does the project involve employing contractors external to the Organisation who would have access to personal or special categories of personal data?		Yes / No
			No

3.5	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)*
	Sharing Children's Services Insights 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, and the identification of good practice in other boroughs;

	b. **Monitoring equalities** – by enabling comparative analysis of the odds of key social work interventions (e.g. Section 47 Investigations, Child Protection Plans, Care Orders) 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, and to benefit registered social workers by supporting their progression and better addressing their needs.

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 will be factored in to the ETL process as detailed below in step 1 (ii):

- 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

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 4.

   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. restricting the analysis to individuals who are in scope during an eleven-year period – chosen due to the limitations the initial six year period gave us particularly in our ability to 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. name, named social worker, and address (all common in Annex A returns).

2. **Protect anonymity** – Degrading indirect identifiers which have a greater level of specificity than we believe we are likely to need – e.g. postcode to postcode sector (a c. 200x 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.

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 para. 7 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 para. 7 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
No. There are prior concerns about the use of machine learning and automated decision making in Children's Services, and so these have been placed out of scope. This scope restriction will be written into the DPAs between data controllers and the data processor.		
4.8	Will individuals be asked for consent for their information to be processed/shared?	Yes/No

	If no, list the reason for not gaining consent e.g. relying on other lawful basis, consent is implied where it is informed.	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

London Data Platform infrastructure

The application is architected to be secure by design. It follows National Cyber Security Centre (NCSC) best practice for cloud applications and its use of Amazon Web Services for all core application components means that it is based on the most solid and robust of cloud platforms.

Specifically, it adheres to the NCSC guiding principles as follows:

Principle 1 - Data in transit protection

- The Data Platform will use the latest SSL certificates.
- Connections will be 128bit encrypted and authenticated using TLS 1.2. (data encrypted in transit)

Principle 2 - Asset protection and resilience

- Data will be hosted on Amazon Web Services (AWS) in a UK availability zone and a UK data centre.
- The AWS data centre is secure and highly monitored (full list of procedures in place available here).
- Data tables will be stored in AWS S3 buckets protected by a private key/user password combination. (data encrypted at rest)
- AWS is certified ISO 27001 (full list of AWS certifications available here).

Principle 3 - Separation between customers

- Robust use of Identity and Access Management (IAM) prevents users accessing or processing data or systems outside their own.

Principle 4 - Operational security

- The platform will be subject to robust and well-defined operational processes
- Releases will be timely and well managed through a deployment management framework
- Incident management will have clear procedures and communication will be timely

Principle 5 – Governance framework

- A governance framework will be established to clearly define feature extension signoff, administration oversight and eventual handoff of the system to a suitable long term owner as required.

Principle 6 - Personnel security

- AWS personnel are subject to robust vetting and security checks.
- Access to systems from inside the London Data Platform will be limited to those identified in the data flows section above.

Principle 7 –Secure development

- Engineering will follow software development best practices, including merge protections and code review
- Deployment will be automated and follow a test > release process across multiple environments before production
- Code will be subjected to multiple levels of automated and manual testing

Principle 8 - Supply chain security

- Software dependencies will be tracked and vulnerabilities automatically reported
- Dependencies in code run on the processing environment will be subject to a robust and well managed allow-list

Principles 9 and 10 - Secure user management/ Identity and authentication

- Uploads will be Private as default.
- Single-sign-on via major identity service providers (Microsoft Active Directory, Google Cloud Identity) will provide best in class password complexity enforcement and multi-factor-authentication requirements.
- A full audit log for user access and data movement will be recorded

Principle 11 - External interface protection

- Exposed user interfaces will follow best practice for user facing application security – including cross-site scripting protection and cross-site request forgery protection
- The minimum required API will be exposed to reduce the available attack surface to the smallest possible

Principle 12 – Secure Services Administration

- Administration of AWS services and infrastructure is carried out entirely through the AWS console and administration system which is subject to the resiliency and security principles laid down by Amazon themselves.

Principle13 - Resilience and auditing

- The application will be fully penetration tested before production release is granted.
- Regular security audits by qualified third-party organisations will be carried out.
- Data protection from loss and lack of availability on AWS is covered by their business continuity and disaster recovery policy.

	Principle 14 – Security by Design and Default • The design of the system is such that users are subject to guardrails which prevent accidental misuse. • Code run on the platform will be sandboxed and airgapped from the wider internet • Controls and role based access will prevent data leaking between users or clients
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 is stored and processed on a secure environment within the LIIA data platform; subject to technical, physical, and process controls as befits the sensitivity of the data. We consider this platform to be an appropriate solution for storing and processing data from all London Boroughs, with the choice to pseudonymise and minimise data from the source datasets to be conducted here, rather than at each Borough. This will ensure that standardised datasets are received and will minimise the data processing requirements of each Borough, in line with our 'once for London' approach.
4.11	Data Retention Period *How long will the data be kept?*
	Our initial analysis of the SSDA903 data received from London LAs through the CLD – Sharing Children's Services data insights project has shown that we are unable to properly provide analysis for Directors and senior leaders on a number of key elements; • what are the journeys that children are taking through the social care and care systems? • how long are children know to social care and in care for? • do children have a number of interactions with social care services and other children's services before they become looked after? This is because we have found that of the children who were looked after between 1st April 2016 and 31st March 2021, over 7,000 started to be looked after at some point before the earliest piece of data that was provided to us. This means that for those children we do not know how long they were looked after for at the point that they ceased to be looked after or to the 31st March 2021 if they continue to be looked after. We also cannot see the full extent of their journey through the care system which affects our ability to understand if they had a different experience based on some of their characteristics, eg ethnicity, gender, disability. This affects our ability to clearly see if disproportionality exists in the London care system. When we combine the SSDA903 data with the CiN census we will also be unable to see if there was social care involvement with these young people before they entered the care system, or if they had previously been in the care system.

	In the 6 years of data that we received, there were a total of 40,000 children who were looked after at any time, so this cohort of 7,000 children for whom we don't know the start of their care period represents 17.5% of the total cohort. We therefore propose to extend the period covered for the 'Sharing Children's Services data insights' part of the CLD project to eleven years to cover the journey from birth to the beginning of adolescence or the beginning of adolescence to the transition to leaving care (and all the variations in between) and then review if that is sufficient. Initially we plan to build on the existing data from 2016/17 (from 2022/2023 in the case of SEN2 which has only just become a child-level data return) and add subsequent years data as they become available to accumulate to eleven years' worth of data. Should DCS' want us to collect data from earlier years, in addition to collecting new years' as they pass, then we would review the case for this and if necessary consult with IGfL. (Adding earlier years would introduce additional technical complexities which may not be desirable.) This will enable service design, practice and resource planning to benefit from the insights gained and support DCS' to meet their statutory obligations. In a reporting year approximately 5,000 children become looked after, so over the course of five additional years, this would add approximately 25,000 children to the dataset. 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 twelve years' data) and to destroy all data at the end of the programme or on request of the data controller, unless at the point of review it is felt that there is a robust and proportionate case for extending further. The intention is not to extend indefinitely and certainly not beyond 25 years – which would cover the current statutory responsibilities for Children's Services for both Care Leavers and children with Special Educational Needs – and therefore allow the complete arc of a child's interactions with children's services in these areas to be seen. Appendix 5 gives examples of the type of analysis we are able to do and the kind that are not possible due to the limitations in the data period covered.	
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 London DataStore before integration to London DataStore processes. Social Finance are also training the LIIA team, including analysts at LBWF, to maintain and extend that code.	

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
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 LB Waltham Forest by LIIA analysts hosted at LBWF. The extract is a download of the full dataset. Data is stored in cache in Power BI, hosted by LBWF. Analysis to aggregate individual level data to Borough level is conducted in Power BI. Only LIIA analysts working on the project will have access to individual-level data.						
5.3	What roles will have access to the information? (list individuals or staff groups)						
	Power BI analysis collected in Power BI report, at a Borough level, with Boroughs identifiable. The report is shared via individual link to named individuals at all London Boroughs. Access to the report is managed by LIIA analysts at LBWF. Links shared with individuals will allow access only to that individual.						
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	**x**	Restricted access to Network Files				
	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 is required for new systems. *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
							Yes/No

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
	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 DataStore 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?		
	LIIA analysts at LBWF.		
	• What roles will receive the report or where will it be published?		
	Power BI analysis collected in Power BI report, at a Borough level, with Boroughs identifiable. The report is shared via individual link to named individuals at all London Boroughs. Access to the report is managed by LIIA analysts at LBWF. Links shared with individuals will allow access only to that individual.		
	• Will the reports be in person-identifiable, pseudonymised or anonymised format?		
	N/A		
	• 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

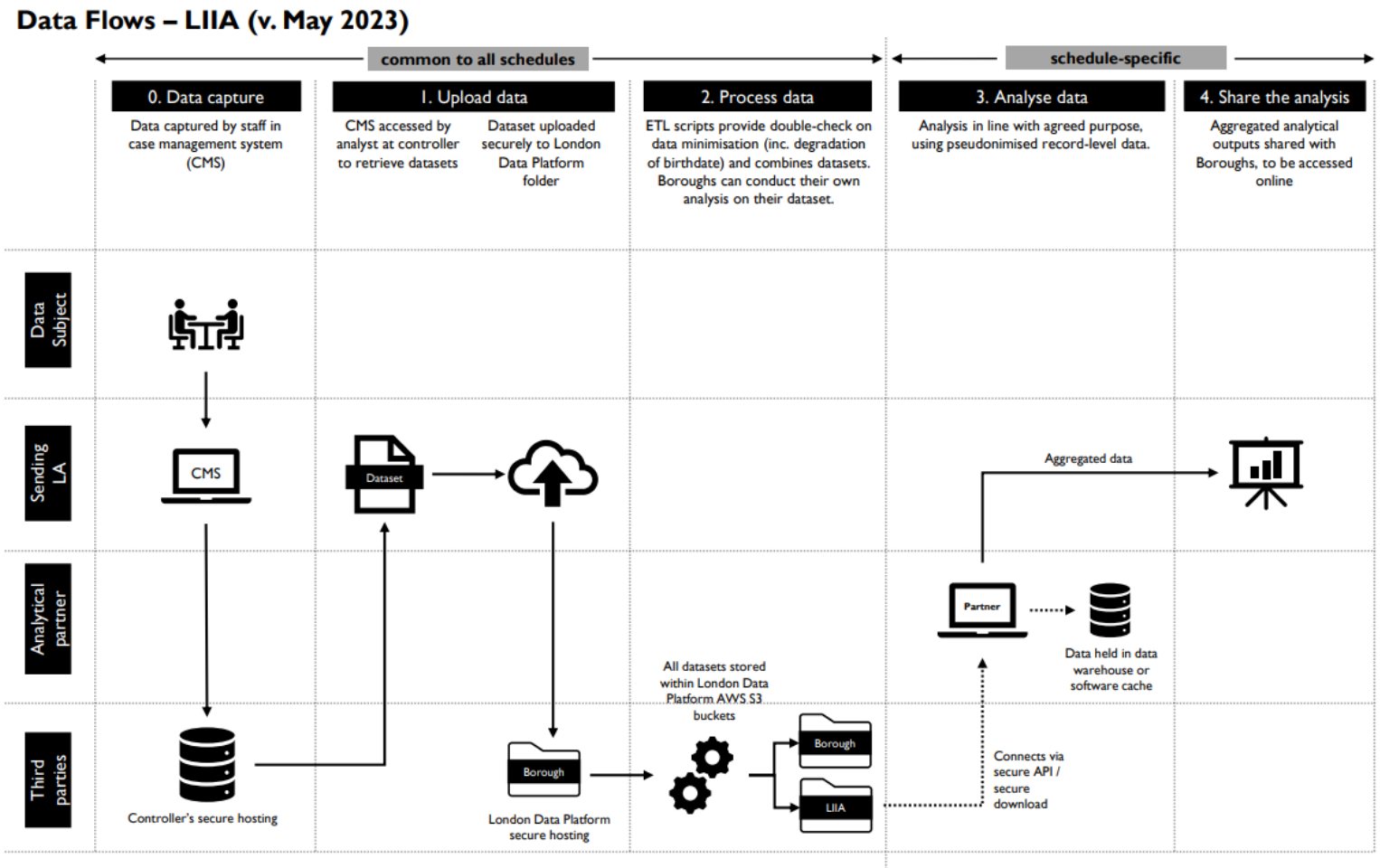
Step 7: Identify Measures to reduce risk

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 27/06/2023	
Residual risks approved by:	Joe Luxton 27/06/2023	
DPO advice provided:	Leila Ridley 04/07/2023	
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

2023 Children-s
Services Insights Appe

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

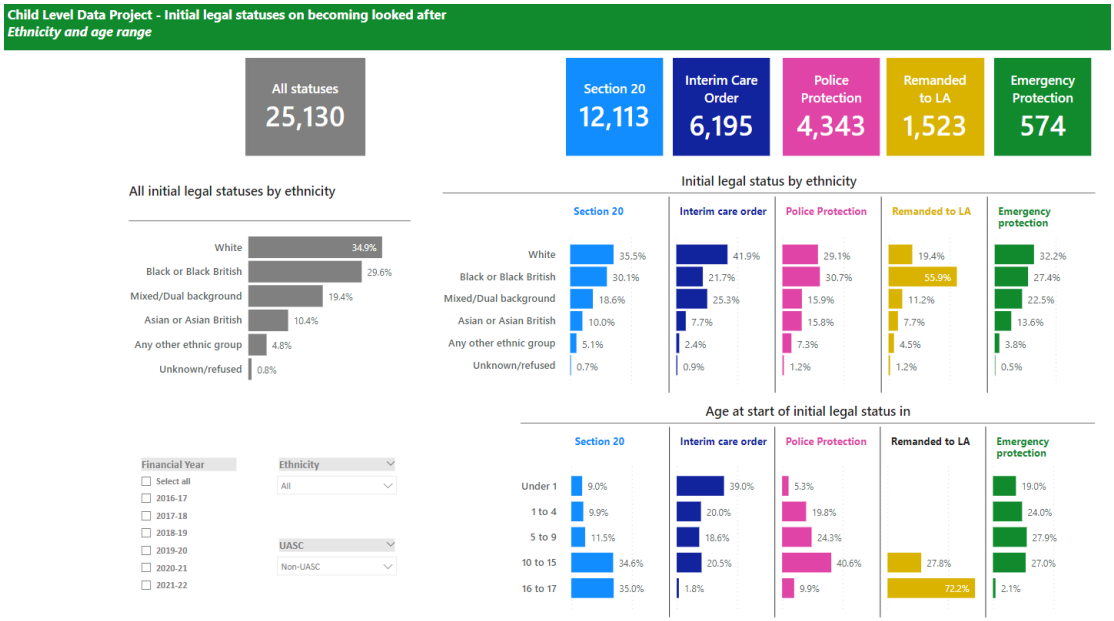
You have the right to object to your data being used this way. If you wish to exercise it then please contact <*insert details*>.

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

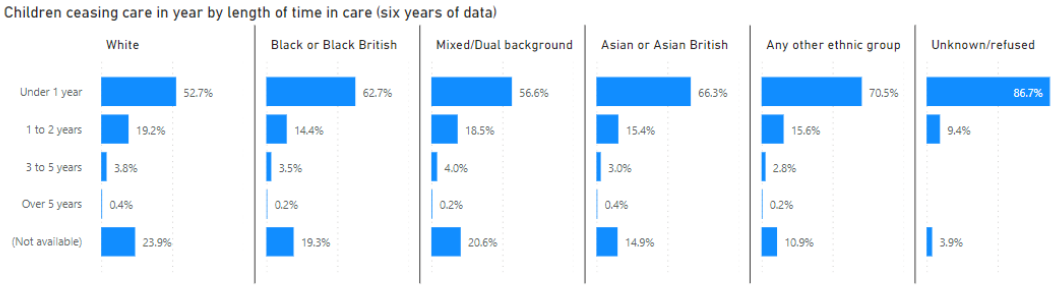
LIIA Child Level Data
DPIA - Children's Serv

Appendix 5: Example of the type of analysis that has been possible from the SSDA903 data received and where there are limitations due to the period covered being 6 years

This is an example of analysis we can be confident in – we know the start dates of children who came into care since April 2016, hence we know their initial legal status. Looking at non-UASC by ethnicity we see remand is disproportionally high for black children (these totals are for all six years):



For duration in care, there's a high proportion – 5,000 out of 24,694 non-UASC, so around 20% - where we just don't know how long they were in care at the time they ceased (shown as *(Not available)*). Could be anywhere from a month to their entire life. This restricts our ability to analyse children's journeys and discern trends over time.



Children ceasing care in year by length of time in care

Years in Care Band	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22	Total
Under 1 year	2,753	3,045	2,520	2,385	1,975	1,856	**14,534**
1 to 2 years	263	561	819	817	819	925	**4,204**
3 to 5 years	38	33	43	173	270	333	**890**
Over 5 years	10	13	4	6	19	15	**67**
(Not available)	1,894	1,020	732	531	440	383	**5,000**
Total	**4,958**	**4,672**	**4,118**	**3,912**	**3,523**	**3,512**	**24,695**