

## DATA PROTECTION IMPACT ASSESSMENT

A Data Protection Impact Assessment (DPIA) is a process that helps an organisation identify and minimise the data protection risks of a project.

### Version Control

Version	Reason	Date	Author(s)
0.1	Draft - initial draft	03/04/2023	[REDACTED]
0.2	Revised draft	12/05/2023	[REDACTED]
1.0	Final version	01/06/2023	[REDACTED]
1.1	Revised 4.1 lawful bases for SHP	07/06/2023	[REDACTED]
2.0	Edits Following Review of Pilot Phase	23/10/2023	[REDACTED]

Project / Work Stream Name	Strategic Insights Tool for Rough Sleeping
<p><b>Overview:</b> (Summary of the project/work stream)</p>	<p>Homelessness in London, including rough sleeping, continues to rise. Addressing the homelessness crisis is a priority for all levels of government and for civil society.</p> <p>Collectively, London seeks to implement a data-informed policy to address homelessness, resulting in improved collective action that cuts across institutions, services, and sectors. However, Greater London lacks consistency in data collection metrics, practices, and management. London lacks a system that enables us to leverage the extensive existing data to achieve these aims. The project's aim was to develop a system which allows join-up of data, thereby delivering the required intelligence and insights. The Greater London Authority and London Councils are working with Bloomberg Associates to develop an integrated rough sleeping data solution for local government and non-government service providers across London to help better understand flow, plan services, report, and manage performance.</p> <p>This project consists of research and statistical work. The aim is not to change the delivery of support to a specific individual, but to inform and improve service delivery at a cohort, or population level.</p> <p>A key goal of the Strategic Insights Tool for Rough Sleeping (SITRS) is to better understand what happens to rough sleepers across their journeys into, during, and after they sleep rough, which will inform policy and service decisions.</p> <p>The advent of rough sleeping is often a confluence of events over many years, from losing a job, developing a health problem, and then finally, perhaps, the loss of a supportive family member or friend. Time on the street can be short, but it can also last months or years, followed by time in hostels or other temporary arrangements. This cycle can stretch over five years, or more.</p>

For most people who experience rough sleeping, the crisis is singular, short, and requires little intervention. For a few, time spent living on the streets is an intermittent occurrence amid other life challenges over a more extended period of time. These few long-term or episodic cases are the ones likely to need more services.

Current data systems allow service providers to track some of this journey for individuals they serve, at least the time spent rough sleeping and in hostels, with self-reported data for what happens in between when people return to the streets. But there is little systematic data to map out the major fault lines. Are there patterns of users who tend to come back to rough sleeping more frequently? Are there opportunities for strengthening ties between agencies, or opportunities for interventions that are lost?

For example, a major reason for rough sleeping is that people are asked to leave their last settled base. This tenancy loss could be due to a variety of factors, including some non-specific life events such as loss of income, relationship breakdown, or emergence of a mental health or substance abuse problem. The current data systems will likely track the final crisis, but they may not indicate the intermediary steps or whether there was any attempt at intervention, such as contact with a Local Authority office. If there was contact, did it occur right before the crisis, or months or years earlier? People return to sleeping rough after being away for more than a year. What happened in between? Did they have any contact with their Local Authority?

Phase 1 of the project involved designing and developing a minimum viable product (MVP), joining data together from key data sources across the rough sleeping ecosystem. This means that the tool has been built with enough features to meet the minimum set of user requirements and validate the product hypothesis. This was successfully rolled out to all pilot organisations following an intensive 13-week build period consisting of a continuous cycle of user engagement and testing and iterative development of the solution. Phase 2 of the project will focus on onboarding the remaining 29 boroughs not involved in the pilot and an additional 7 rough sleeping service providers to make this a truly Pan-London initiative to help better understand and transform rough sleeping services across London.

As part of the pilot phase of the project, the following work was conducted in order to meet the objectives outlined above:

- Work to develop a minimum dataset for the proposed solution was led by Bloomberg Associates and this was agreed among project stakeholders. The minimum dataset was iterated upon based on working with pilot partners and this is included in Appendix 1.
- A proposal for the development of a new data solution was put together and subsequently a formal procurement process was launched, inviting bids for the contract to develop the data solution. Faculty Science were successfully appointed for the contract and commenced work on 5 June 2023, with the contract due to end on 2 February 2024 with the aim of the MVP solution being rolled out to all 33 London boroughs, the GLA, London Councils, Homeless Link, and circa 11 rough sleeping service providers.
- A data sharing agreement was established and signed between participating pilot organisations to permit the sharing of CHAIN, H-CLIC and In-Form data into the system.
- A Project Steering Committee was formed, bringing together project leads and practitioners from across the stakeholder landscape, as a body for important decision-making and risk management.
- Weekly meetings have been conducted throughout the pilot phase between

	<p>Faculty, LOTI, GLA and London Councils to maintain project oversight.</p> <p>The MVP solution was developed during a 13-week period and rolled out to pilot users from 8th September 2023. Continued assessment and feedback of the solution among users is ongoing.</p> <p>The Strategic Insights Tool for Rough Sleeping (SITRS) is a new tool that will give decision-makers in GLA, London Councils, Local Authorities, and homelessness service providers, a clearer view of rough sleeping in their local area. The new tool merges and integrates multiple sources of data, meaning that for the first time, users of this tool are able to see the aggregated journeys of rough sleepers over time, as they show up through touch points in multiple systems, which include statutory homelessness applications; contacts with housing outreach officers, as they are seen bedding down; and interactions with service providers, who are commissioned to support them through various services.</p> <p>Through the use of this tool, users are able to get actionable insights on how support can be improved, to make rough sleeping rare, brief and non-recurrent. By seeing the aggregated journeys of rough sleepers through various combined systems, users of this tool can better map the history and journeys of rough sleepers through their interaction with homelessness services. For example, by seeing the different inflows and outflows of rough sleeping by different boroughs, Local Authority or pan-London commissioners can make educated decisions about the effectiveness of different services and support, while forecasting and pre-empting rough sleeping trends over time.</p> <p>Local Authorities can also see the inflows of rough sleeping into their borough, by seeing previous statutory housing applications that rough sleepers may have made, across London Local Authorities.</p> <p>Service providers can see pan-London aggregated stats to understand the comparative benchmarking of certain rough sleeping services, as well as seeing how different boroughs compare in terms of aggregated rough sleeping journeys.</p> <p>At a high-level, the tool works according to the following 4 steps:</p> <ol style="list-style-type: none"> <li>1. Users upload their data on an agreed-upon routine basis (either manually directly to the tool or via an automated extract)</li> <li>2. Data is cleaned and normalised into a structured data model ready for processing</li> <li>3. Records are matched using a probabilistic matching algorithm which aims to identify where there are matches across the different inputting datasets</li> <li>4. Different journeys are aggregated and displayed within the tool through the user interface. This means we can ensure that data is safely and securely anonymised, while providing insight into rough sleeping journeys across many different systems.</li> </ol>
--	---

<p><b>Implementation Date:</b></p>	<p>Phase 1 - June 2023</p> <p>Phase 2 - November 2023</p>
------------------------------------	---

**Environmental Scan**

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.

The Greater London Authority, LOTI, and London Councils have conducted a scan of the various disparate data sources that make up the local data collection and data sharing landscape: CHAIN, H-CLIC, and In-Form. Included in this scan are the agencies that are required to report and collect homeless information across the four pilot London boroughs.

CHAIN (Combined Homelessness and Information Network) is a multi-agency database recording information about people sleeping rough and the wider street population in London. It is commissioned and funded by the Mayor of London and managed by Homeless Link. CHAIN allows approved and trained users to share information about work done with rough sleepers and about their needs, ensuring that they receive the most appropriate support and that efforts are not duplicated.

Homelessness Case Level Information Classification (H-CLIC) is a database collecting information from Housing Options offices operated by local authorities. It collects information regarding reasons people become homeless, a person's housing history and journey, the support needs of homeless households, and how these households link with other public services and/or the benefits system. Staff from each local authorities' Housing offices use proprietary software systems to enter H-CLIC data for clients who apply for benefits at the housing options office. This data is manually uploaded quarterly to the Department for Levelling Up, Housing and Communities (DLUHC). For this project, H-CLIC information which has been submitted to DLUHC will also be uploaded to the Strategic Insights Tool for Rough Sleeping.

In-Form is a proprietary case management software system that is used by many providers in the homelessness/rough sleeping charity sector. Homeless Link, in partnership with Salesforce, offers a configurable tool for charity service providers to subscribe to and tailor according to their needs. The tool is often used when clients are linked to services, including rough sleeping hubs or temporary accommodation settings. Some charities commission similar case management tools directly from Salesforce, rather than the out-of-the-box tool offered by Homeless Link.

Within the Strategic Insights Tool for Rough Sleeping, personal data on individuals is matched across data from multiple sources, and then aggregated trends and statistics are produced.

Following a procurement process developed by LOTI and London Councils, Faculty Science were successfully appointed as technical delivery partner for this piece of work to build and manage (in the interim) the new data phase. They have developed an MVP solution, incorporating CHAIN, H-CLIC and In-Form data from all pilot organisations. This solution is live and operational for those users and a testing period was conducted over 5 weeks immediately following the launch of the solution in order to capture feedback from users.

Based on the success of phase 1, the SITRS is being rolled out to the remaining 29 London boroughs and an additional 7 service providers which will make up phase 2. Boroughs will continue to be asked to submit their H-CLIC data and service providers their In-Form (or similar) data in line with the defined minimum data set (Appendix 1). As the SITRS is a live tool, organisations are required to submit data at a regular and agreed-upon interval to ensure the utility of the tool for all users.

	<p>Boroughs must submit H-CLIC data every 3 months with data from 1st January 2022 up until their most recent H-CLIC return submitted to DLUHC. Service providers must submit their In-Form (or similar) data on a monthly basis. Organisations can opt for manual upload or set up an automated pipeline with the support of the technical delivery team. An automated pipeline is set up with CHAIN whereby data is updated every week.</p> <p>The project team are liaising with phase 2 organisations to determine which systems each organisation uses for this data to be drawn from and the preferred upload mechanism.</p> <p>The participating local authorities and housing charities are joint data controllers, along with London Councils and the GLA, with the contractor building the system (Faculty) as data processor. At the end of the contract with Faculty (2nd February 2024), the GLA will be responsible for letting and managing the contract to host, develop and maintain the SITRS system. Homeless Link will have the responsibility for the ongoing management and maintenance of the SITRS under contract with the GLA to deliver this as a service. The GLA and London Councils are also joint data controllers as they have influence over the Strategic Insights Tool for Rough Sleeping system's design and processing.</p>
--	---

<b>Step 1: Complete the Screening Questions</b>				
<b>Q 1</b>	<b>Category</b>	<b>Screening question</b>	<b>Yes/No</b>	
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?	No	
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?	Yes	
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 EEA?	No	
1.11	Exemptions and Exceptions	Does the project relate to data processing which is in any way exempt from legislative privacy protections?	No	
1.12	Exemptions and Exceptions	Does the project's justification include significant contributions to public security and measures?	Yes	
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.**

<b>Step 2: Identify the need for a DPIA</b>		
2.1	<b>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?</b>	New/Changed
		New

**2.2 What data will be processed/shared/viewed?**

Personal data will be processed and matched across data sources (CHAIN, H-CLIC and In-Form/Salesforce) and then pulled into a merged database. Data pulled from this database into the user interface (front-end) to form the visualisations is aggregated and completely anonymised. No personal data is visible through the front-end. Authorised users will only be able to view:

(1) individual level data that they contributed (because they are the one that uploaded it, it is not visible through the user interface) and

(2) view linked information (in aggregate format) which has been de-identified.

The reasons people enter rough sleeping are varied, so the treatments to address rough sleeping are equally as varied. This research project aims to better understand those reasons by mapping and connecting the continuum of rough sleeping. The aim is not to change the delivery of support to a specific individual, but service delivery may change to the cohort, or sub-cohorts, of people experiencing rough sleeping.

**Personal Data**

Forename	Y	Surname	Y	Date of Birth	Y	Age	Y	Gender	Y
Address	N	Postal address	N	Employment records	Y	Email address	Y	Postcode	N
Other unique identifier ( <i>please specify</i> )	Y (National Insurance #)	Telephone number	Y	Driving licence number	N	NHS No	Y	Hospital ID no	N

Other data  
(*Please state*):

Relationship Status, Preferred Language, Pregnancy Status, Disabilities, Armed Forces History, Nationality, Care Leaver 18+, Current Substance Use, Current alcohol misuse support needs, Domestic Violence Services Need, Immigration Status, Entitlements to Welfare Benefits, Current Mental Health Concerns, Connection to hospital or rehab.

The current dataset list is attached in Appendix 1 (Phase 2 minimum dataset / data check list), this has been updated based on work conducted throughout the pilot phase to determine what was needed to meet the core set of requirements.

**Special Categories of Personal Data**

Racial or ethnic origin			Y	Political opinion		N	Religious or philosophical beliefs		N	
Trade Union membership			N	Physical or mental health or condition					Y	
Sexual life or sexual orientation		Y	Social service records		N	Child protection records			N	
Sickness forms	N	Housing records	Y	Tax, benefit or pension records			N	Adoption records		N
DNA profile	N	Fingerprints	N	Biometrics	N	Genetic data			N	

Proceedings for any offence committed or alleged, or criminal offence record, Current Medical Needs, Current or recent hospitalisation or in inpatient service (e.g. rehab), Current Immigration Support (legal needs)		Y
Other data <i>(Please state):</i>	Connection to Housing Options Office, Date of most recent Housing Options Office contact (if any), History of statutory homelessness or prevention duty, Borough most recently engaged in rough sleeping service(s), Preferred language, Domestic Violence Services Need, Date of first documented encounter during current episode of rough sleeping, Date Last Seen Rough Sleeping, Last Accommodation/Settled Base, Main reason (cause) for loss of settled base, Date Accommodated, Connection to hospital or rehab, Arrival date, Type of Accommodation, Departure/end date from accommodation, Departure reason, Destination Upon TA Departure, Reconnection made	

Will the dataset include clinical data? (please include)	<b>Yes/No</b>
	Yes: Focus on current medical needs, mental health needs, substance misuse needs, alcohol support needs, and disabilities, connection to hospital or rehab
Will the dataset include financial data?	No
<b>Description of other data processed/shared/viewed?</b>	

<b>2.3</b>	<b><u>Business sensitive data</u></b>	
	Financial	<b><u>None</u></b>
	Local Contract conditions	None
	Operational data	None
	Notes associated with patentable inventions	None

procurement/ tendering information		None
Customer/ supplier information		None
Decisions impacting:	One or more business function	<b>No</b>
		No
	Across the organisation	No
<b>Description of other data processed/shared/viewed (if any).</b>		

<b>Step 3: Describe the sharing/processing</b>						
<b>3.1</b>	<p><b>List of organisations/partners involved in sharing or processing personal/special categories personal data? <i>If yes, list below</i></b></p> <p>Participating organisations are considered joint data controllers. All parties will together determine the data used within the system and any future development of the system. The GLA will be responsible for letting and managing the contract to host, develop and maintain the SITRS system. This includes undertaking suitable due diligence with any contractors/data processors they employ and being responsible for the security of data within the tool and that a written agreement is in place with each data processor, and that all data processors will be bound by this agreement. Homeless Link will have responsibility for the ongoing management and maintenance of the Strategic Insights Tool from 2nd February 2024 (end of contract with Faculty) under contract with the GLA to deliver this as a service where they will be joint data controllers. The GLA and London Councils are also joint data controllers as they have influence over the Strategic Insights Tool for Rough Sleeping system's design and processing.</p>	Yes/No				
		Yes				
	<table border="1"> <thead> <tr> <th>Name</th> <th>Controller or Processor?</th> </tr> </thead> <tbody> <tr> <td><b>London Local Authorities</b> For boroughs marked with a (*) it is proposed that their accommodation data will be ingested to the tool in addition to their H-CLIC data, as they provide substantial in-house managed services</td> <td>Controller</td> </tr> </tbody> </table>	Name	Controller or Processor?	<b>London Local Authorities</b> For boroughs marked with a (*) it is proposed that their accommodation data will be ingested to the tool in addition to their H-CLIC data, as they provide substantial in-house managed services	Controller	<p>Completed and compliant with the IG Toolkit or <a href="#">Data Security and Protection (DSP) Toolkit</a></p> <p>Yes / No</p>
Name	Controller or Processor?					
<b>London Local Authorities</b> For boroughs marked with a (*) it is proposed that their accommodation data will be ingested to the tool in addition to their H-CLIC data, as they provide substantial in-house managed services	Controller					

	Westminster	Controller	Yes
	Camden	Controller	Yes
	Hillingdon	Controller	Yes
	Lambeth	Controller	Yes
	City of London	Controller	Yes
	Southwark*	Controller	Yes
	Ealing	Controller	Yes
	Redbridge	Controller	Yes
	Kensington & Chelsea	Controller	Yes
	Croydon	Controller	Yes
	Hackney	Controller	Yes
	Enfield*	Controller	Yes
	Haringey*	Controller	Yes
	Brent	Controller	Yes
	Wandsworth	Controller	Yes
	Waltham Forest	Controller	Yes
	Barnet	Controller	Yes
	Barnet Homes (The Barnet Group)	Controller	Yes
	Barking & Dagenham	Controller	Yes
	Tower Hamlets	Controller	Yes

	Lewisham	Controller	Yes
	Hammersmith & Fulham	Controller	Yes
	Islington	Controller	Yes
	Greenwich	Controller	Yes
	Hounslow	Controller	Yes
	Newham	Controller	Yes
	Harrow	Controller	Yes
	Merton	Controller	Yes
	Bexley	Controller	Yes
	Bromley	Controller	Yes
	Sutton	Controller	Yes
	Havering	Controller	Yes
	Richmond	Controller	Yes
	Kingston	Controller	Yes
	<b>Other Public Sector</b>		
	GLA	Controller	No (not subject to them)
	London Councils	Controller	No (not subject to them)
	<b>Charities/service providers</b>		

	SHP	Controller	No (not subject to them)
	Homeless Link (CHAIN)	Controller	No (not subject to them)
	St. Mungo's	Controller	No (not subject to them)
	Thames Reach	Controller	No (not subject to them)
	Evolve Housing	Controller	No (not subject to them)
	Look Ahead	Controller	No (not subject to them)
	Providence Row Housing Association	Controller	No (not subject to them)
	Salvation Army	Controller	No (not subject to them)
	YMCA St Paul's	Controller	No (not subject to them)
	Your Place	Controller	No (not subject to them)
	Trinity Homeless Projects	Controller	No (not subject to them)
	Centre Point	Controller	No (not subject to them)
	<b>Contractor</b>		
	Faculty Science	Processor	Contractor not subject to these toolkits, but has suitable security standards.
	Homeless Link (post-Feb 2024)	Controller	As above
<b>3.2</b>	<b>If you have answered 'yes' to 3.1 is there an existing ' Data</b>	Yes/No	

	<p><b>Processing Contract' or 'Data Sharing Agreement' between the Controller and the Processor?</b></p>	<p><b>Yes</b></p> <p>A contract is in place with the contractor building the system (Faculty Science) which establishes the relationship with Faculty Science as a 'Data Processor' and outlines the processing activities in line with this DPIA.</p> <p>A DSA is in place to cover sharing with the listed partners where they are Controllers.</p> <p>Each Local Authority will be responsible for ensuring that they have Data Processing Contracts in place with any local Processors.</p> <p>No data will be shared by organisations joining as part of Phase 2 until they have signed the Phase 2 DSA. Once a signature has been received from an organisation, they can begin the onboarding process.</p> <p>There will be a contract in place between the GLA and Homeless Link by the end of Phase 2 as Homeless Link assumes responsibility for the ongoing management and maintenance of the tool at the end of the Faculty contract (2nd Feb 2024).</p>
--	--	--

<p><b>3.3.</b></p>	<p><b>Has a data flow mapping exercise been undertaken?</b> If yes, please provide a copy, if no, please undertake</p>	<p>Yes, data flows for the SIT have been attached as Appendix 2.</p>
--------------------	--	--

<p><b>3.4</b></p>	<p><b>Does the project involve employing contractors external to the Organisation who would have access to personal or special categories of personal data?</b></p>	<p>Yes / No</p> <p>Yes - Faculty Science have been employed as contractors to build the tool and onboard all stipulated organisations. Appropriate measures have been taken to limit access to personal data (listed below).</p>
-------------------	---	--

<p><b>3.5</b></p>	<p><b>Describe in as much detail why this information is being processed/shared/viewed?</b> <i>(For example Direct Patient Care, Statistical, Financial, Public Health Analysis, Evaluation. See NHS Confidentiality Code of Practice Annex C for examples of use)</i></p>
-------------------	--

Local authorities are subject to a public duty to reduce or end rough sleeping in their areas, as identified, for example, in the Homeless Act 2002, and the Homelessness Reduction Act 2017. However, no single organisation is fully accountable for ending rough sleeping in a community. A collaborative, strategic, data driven approach, with shared commitments, must be undertaken by partners to make rough sleeping rare, brief, and non-recurrent. London seeks to implement a data-informed policy, but this cannot be accomplished without improved collective action that cuts across institutions, services, and sectors. However, Greater London (and most of England) lacks consistency in data collection metrics, practices, and management.

**Benefits to Organisations:**

- Better data received more quickly
- Allows the discharging of local duties in respect to homelessness prevention and intervention
- Supports identification of improved interventions and more effective service delivery

**Benefits to Society:**

- Less rough sleeping overall
- Less money spent overall on rough sleeping prevention and client support

**Benefits to Individuals:**

- Better targeted support to reduce rough sleeping for individuals making it rare, brief, and nonrecurrent

**The fundamental challenges are:**

- Too little join up between systems and stakeholders in some local areas
- Too much duplication across different systems with reporting requirements
- Extracting the data needed for casework, reporting and performance management is often difficult
- The data that commissioners need for reporting, management and commissioning is not always collected

**Key strategic goals of this work are intended to:**

- Enhance the effective operation of services for those who are rough sleeping
- Improve understanding of the rough sleeping population (including complexity of needs and flow)
- Improve understanding of individual outcomes and views of rough sleepers
- Learn what strategies work and inform public policy, service delivery and commissioning
- Ease and enhance reporting and performance management
- Measure success (in making homelessness rare, brief, non-recurrent) and align with DLUHC/CHI efforts
- Ensure solutions and resources are properly understood, fully utilised, and linked to need (See User Requirements - Appendix 4)

**Step 4: Assess necessity and proportionality**

#### 4.1

**Lawfulness for Processing/sharing personal data/special categories of personal data?** Local authorities are subject to a public duty to reduce or end rough sleeping in their areas, as identified, for example, Part 7 of the Housing Act 1996, the Homeless Act 2002, and the Homelessness Reduction Act 2017. The Domestic Abuse Act 2021 amends Part 7 of the 1996 Act to strengthen the support available to victims of domestic abuse. The Act extends priority need to all eligible victims of domestic abuse who are homeless as a result of being a victim of domestic abuse.

<https://www.gov.uk/guidance/homelessness-code-of-guidance-for-local-authorities/overview-of-the-homelessness-legislation#:~:text=The%20primary%20homelessness%20legislation%20%E2%80%93%20that,threatened%20with%20or%20actually%20homeless>

Rough sleeper services commissioned by, or in a smaller number of cases, working in partnership with and endorsed by, local authorities and the GLA, exercise a function on behalf of these bodies, in order to further this public duty.

Charities and service providers in the homelessness sector deliver support services and accommodation to those experiencing, or at risk of experiencing, homelessness and rough sleeping. They use a variety of lawful bases dependent on the services offered.

It is in the public interest to work to reduce rough sleeping and homelessness, as well as being a statutory requirement for local authorities. Reducing homelessness positively impacts individuals, improves the overall wellbeing of an area, and reduces public health risks.

Homeless Link manage the CHAIN system under commission from the GLA, for the task in the public interest to achieve outcomes that benefit rough sleepers and the wider street population, and thereby to reduce or eliminate rough sleeping in London. Homeless Link will also have responsibility for the ongoing management and maintenance of the Strategic Insights Tool from 2nd February 2024 (end of contract with Faculty) under contract with the GLA to deliver this as a service, where they will be joint data controllers.

The GLA and London Councils are joint data controllers as they have influence over the Strategic Insights Tool for Rough Sleeping system's design and processing. The Digital Economy Act supports data sharing between the GLA and boroughs and vice versa under Part 5, Cpt 1 s35.

The reasons people enter rough sleeping are varied, so the treatments to address rough sleeping are equally as varied. This research aims to better understand those reasons by mapping and connecting the continuum of rough sleeping. There is no one system that exists to achieve this rough sleeping data solution, so a new solution has been developed in order to leverage existing quality data sources in the short-term and identify data gaps to be addressed in the longer term. The Strategic Insights Tool has been developed to fulfil this purpose. With the MVP now live and operational, the rollout of this tool to other organisations listed above will expand the reach of its impact as it becomes a pan-London initiative. Longer-term plans to further develop the solution are under consideration.

This pilot project is subject to a joint data controller arrangement, since more than one organisation is involved in making decisions about the purpose and means of processing on the system.

The parties use different lawful basis conditions to process the personal data. For the local authorities, their original processing of the data matches the lawful basis for all parties as joint controllers for this project, which is public task for personal data, and substantial public interest (Schedule 1, Part 1 DPA 2018 - 'safeguarding of children and of individuals at risk'), and research (Schedule 1, Part 2 DPA 2018 - 'research') for special category data.

The other parties use the following lawful bases for the original processing (details can be found in the Appendix):

- Article 6(1) (e) public task or Article 6(1) (f) legitimate interests

	<ul style="list-style-type: none"> <li>Article 9(2) (g) substantial public interest or Article 9(2) (j) archiving and research</li> </ul> <p>UK GDPR Article 5 states that, "...further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation');". Additionally, the ICO Guide to the General Data Protection Regulation (GDPR) states that, "If your new processing is for research purposes, you do not need to carry out a compatibility assessment, and in most circumstances you can be confident that your lawful basis is likely to be either public task or legitimate interests."</p> <p>This is a research project and is considered to be 'not incompatible' to the original lawful basis for processing, for any of the bases used by the parties. Processing of data relating to criminal convictions and offences or related security measures for the parties' original purposes, and this project, is in accordance with Schedule 1, Part 2, Section 6 of the Data Protection Act 2018. This allows for such processing that is 'necessary for reasons of substantial public interest' where the purpose relates to 'the exercise of a function of the Crown, a Minister of the Crown or a government department'.</p>
--	--

<b>4.2</b>	<b>Will the information be processed/shared electronically, on paper or both?</b>	Electronic	Yes
		Paper	No

<b>4.3</b>	<b>How will you ensure data quality and data minimisation?</b>
------------	--

This project consolidates data from varied sources and each partner is responsible for ensuring the accuracy and relevance of the personal data that it processes and shares and must have clear processes in place for managing data quality.

Any party learning of the inaccuracy of personal data is responsible for taking appropriate action to correct it and informing the parties with whom that data has been shared.

Client records from each database are matched and merged using probabilistic matching. A probabilistic matching algorithm, implemented by the data science and engineering team at Faculty, uses unsupervised machine learning techniques to try to identify any matches between records across systems. For example, the algorithm may detect that a person who has previously completed a statutory homelessness application within a certain Local Authority, as seen in a local authority's H-CLIC uploads, has appeared in bedded down contacts by housing outreach officers, within the CHAIN system. In this instance, records relating to this person between H-CLIC, where their statutory homelessness application would be, and CHAIN, where their rough sleeping outreach contacts are recorded, would be associated with one another, and merged into a single rough sleeping "journey".

If conflicts arise between the original records, a prioritisation mechanism selects the information from the most reliable source. However, such conflicts have little impact on the final result, because they tend to affect fields like the name or phone number which are of no use after matching. This might not be as straightforward as just matching on the person's name, as typos, fake names, and other reasons, might mean that the name values differ between systems. As a result, the complex matching algorithm considers a number of different factors, including "fuzzy matching" between names, and other columns such as contact details, to score any potential matches.

The algorithm only associates records that meet an 85% probability of being a match. This gives us a high level of confidence and includes as few false positives (cases where we've accidentally matched two different

people) in the final matched dataset as possible. The current level of recall is 91%, giving the model a high degree of accuracy. The 9% that is missing was due to quality issues e.g. we cannot confidently say that two records belong to the same person (they had different names or a lot of missing fields). Recall is a measure of how many relevant elements are detected i.e. true matches. 91% recall means that 91/100 true matches are correctly identified. Whilst this represents a high level of accuracy, it should be acknowledged that the risk here is that we miss 9/100 matches and numbers subsequently appear lower in places where they should be higher. To achieve higher recall, the data to be matched would need to be fuller and of better quality. Whilst this should be acknowledged, it's important to note that the SITRS exists to show trends on a system and population level. It is not a substitute for published data and reports. Data in the SITRS may be slightly different to that found in published, static, reports. This is due to differences in processing and timing of data preparation. The recall percentage will also vary as we ingest / match new data dependent on the quality of that data.

Measures are also taken to improve data quality as part of this process. When data quality was low, we conducted various cleaning activities to standardise the data before conducting the matching:

- We parse phone numbers (convert them to a different format to ensure consistency)
- We removed name that seemed suspicious e.g. they were too long or contained digits
- We parsed national insurance numbers (convert them to a different format to ensure consistency)
- We put all names in lowercase to ensure consistency across records
- We also remove duplicates

When organisations submit data to be ingested into the tool, these reports may look slightly different depending on the organisation and system they used. A custom mapping model is used in order to standardise all data submissions into a format that matches the minimum data set request.

An additional deliverable of this project will be some recommendations, developed by Faculty, on how organisations can improve their data quality. This will be key for maintaining a strong degree of accuracy in the SIT going forward and continually improving the quality of the data and insights available through the tool.

### **Data Minimisation**

The data being requested from participating organisations is in line with an agreed minimum data set to serve the purposes of the tool and this has been continuously iterated on throughout phase 1 to refine and clarify the required data items and ensure appropriate data minimisation. Some service provider organisations asked that their unmatched data is deleted from the database (back-end) and these requests were complied with appropriately. Subsequently it was decided that all unmatched data from all service provider organisations should be deleted and this will be the approach taken forward, both to ensure consistency of the insights, and to ensure data minimisation. This will be an automated process.

The data scope has been reviewed at the end of phase 1. Some of the data shared was not used such as:

- Alcohol use
- Substance misuse
- Relationship status
- Pregnancy status
- Prison history
- Current mental health concerns
- Medical needs
- Domestic violence
- Care Leaver history
- Entitlement to welfare benefits

This was because phase 1 was conducted over a very short period of time (13 weeks) and due to some delays experienced in receiving data from pilot users, the project team had to de-prioritise some of the original requirements.

Whilst the above data has not been used to support delivery of the MVP, it is required for planned feature developments which were not able to be incorporated into phase 1 due to the above mentioned constraints. Planned developments included in the work plan involve the development of additional filters e.g. a "Pregnancy" or "Substance misuse" filter, meaning users are able to filter to these groups of people and understand how they interact with services. Secondly, the development of visualisations specific to the data e.g. "What is the movement of those with substance misuse?". These are important enhancements to build in as we know these categories have a significant impact on homelessness and the types of interventions that would be appropriate and effective. It was determined that other features / filters were of higher priority in the MVP phase, however it would be useful for SIT users to understand this information and this does form a part of the product work plan. Furthermore, if the data were to be deleted from the SIT environment, removed from future data requests, and deemed out of scope going forward, it would require significant resource to collect this information. The delivery team would have to work with participating organisations to amend the reports they provide (both now for Phase 1 users, and then for both Phase 1 and 2 in the future). Regardless of whether this is automated or manual, it would still utilise a significant amount of both the providers' time and delivery team resource. At the very least, it would require a 60 minute 1:1 conversation with c.47 organisations in the first instance, likely needing more time with the majority to iterate. As a lot of this data is currently captured by participating organisations in free text format, we also need the data to be able to explore its usability as we remain in development phase. The tool doesn't currently support free text but we could allocate technical resource to building this functionality if we determine the data to be usable, or work with organisations to convert it into a more structured format.

The unused data from phase 1 users is currently retained in the SIT environment but is not released or visible to users in any way. Retaining this data and continuing to collect it from phase 2 organisations will enable us to conduct the above planned development work and therefore we recommend it continues to form part of the data request.

We had to expand some of the data requirements to include more granularity e.g. CHAIN - time and location information for events data where it was asking if someone was seen bedding down (i.e. when and where, enabling us to build a much richer picture of the journey of that individual). Building these journeys are dependent on knowing the start and end dates of different event types - this information was not captured to the required level of granularity in the minimum dataset. For H-CLIC data there were a few columns which weren't well-defined. We therefore requested eligibility, whether someone had priority need, which type of duty the applicant received, and which type of service they got during the duty.

The updated minimum data set is attached in Appendix 1.

During the course of the usage testing, it was discovered by a user that in some instances where multiple filters were applied to the data it was possible to get the number of individuals showing in a particular insight down to 1 or 2. Given that this tool is not supposed to present identifying information, the project team considered in depth how to mitigate possible re-identification risks. The project team sought professional expertise on the matter and considered the key question "can a user work out something new about an individual?". In summary it was determined that if a user filtered to 2 individuals, and they already knew who one of them was, there would be a 100% chance that they would learn something new about that person e.g. their sexual orientation, where they moved to after rough sleeping in X location. This puts that person's private data at risk and therefore the recommendation is that we limit the view in line with ONS best practice. In practice, this means that any visualisation where there is an output of 5 or lower, we show "equal to or less than 5".

4.4	<b>Have individuals been informed about the proposed use of their personal or special categories of personal data?</b>	
-----	--	--

	<i>For example, do the organisations/partners listed in section 3.1 have updated Fair Processing Notice available to patients on their websites?</i>	
--	--	--

	<p>Following UK GDPR Article 5, research is ‘not incompatible’ with purposes for processing so additional privacy information is not required for this project. However, all parties are recommended to review their privacy information and adapt as necessary to reference research, and in the case of housing charities and similar, to describe sharing data with local authorities for the purposes of reducing homelessness.</p> <p><a href="#">Guidance</a> was developed by LOTI’s Pan-London IG Lead to support charities and other service providers working with people experiencing homelessness, to help them better engage with the project and provide advice on providing appropriate privacy information to individuals and choosing the lawful basis for processing personal data.</p> <p>During the course of phase 1 of the project, one service provider did review and adapt their privacy information. Not only did this enable them to effectively engage with the project and be transparent with the individuals they provide services to, but they noted that they were “in a far more robust IG position for all future research partnerships and projects.”</p> <p>Subsequently, some of the service providers entering phase 2 of the project have also used the guidance to review and adjust their privacy notices.</p>	
4.5	<p><b>How will you help to support the rights of individuals?</b></p>	
	<p>Each controller remains responsible for complying with the applicable data subject rights. Each controller in this joint controller relationship has their own documented policies and procedures which detail how they will meet their own obligations in respect of data subject rights.</p>	
4.6	<p><b>Are arrangements in place for recognising and responding to Subject Access Requests (SARs)?</b></p>	
	<p>This is a joint controller relationship and the data shared into the new system is a duplicate of data held by each organisation. Each data controller already has processes to handle SARs, and each will need to ensure reference to this system is included in SARs where relevant.</p> <p>No SARs were identified during Phase 1 of the project. Participating organisations have been consulted and they will notify the project team should any SARs be received. Homeless Link will be responsible for documenting a process for handling any SARs and how to handle data subject rights that might lead to the data within the system needing to be changed, such as a rectification action.</p>	
4.7	<p><b>Will the processing of data include automated individual decision-making, including profiling?</b></p> <p><i>If yes, please outline the profiling processes, the legal basis underpinning the process, and the rights of the data subject</i></p>	
	<p>No/not applicable—this data will not be used in a way that will impact individuals, such as for determining programme, service or benefits eligibility. Users of the tool can examine aggregate data for research and planning or reporting of broad outcomes for accountability.</p>	

4.8	<p><b>Will individuals be asked for consent for their information to be processed/shared?</b></p> <p><i>If no, list the reason for not gaining consent e.g. relying on other lawful basis, consent is implied where it is informed.</i></p>	No
<p>Consent is not the lawful basis for processing this data. The data controllers use public task or legitimate interests for their lawful basis, with substantial public interest for special category data. This is a research project and research is considered 'not incompatible' with the original purposes for processing.</p> <p>UK GDPR Article 5 "...further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes ('purpose limitation')"</p>		
4.9	<p><b>As part of this work is the use of Cloud technology being considered either by your own organisation or a 3<sup>rd</sup> party supplier? If so please complete the embedded questionnaire.</b></p>	<p>Yes, details are covered in the questions below.</p>
4.10	<p><b>Where will the data be stored</b></p> <p><i>Examples of Storage include bespoke system (e.g. EPR, Emis &amp; other clinical systems, SharePoint, data repository, Network Drives, Filing cabinet (office and location), storage area/filing room (and location) etc.</i></p> <p>All data is stored within the eu-west-2 (London) AWS (Amazon Web Services) region. Two services are used for storage. RDS (Relational Database Service) is used to store the cleaned, transformed, and matched outputs. The database is deployed to a public subnet to support access from Faculty's data science workbench. All connections to the database are restricted by a network firewall, only permitting connections from whitelisted IP (Internet Protocol) addresses, and requiring connections to be encrypted. Amazon S3 (Amazon storage service) is used to store the raw datasets on their initial entry into the environment, whether from manual upload or automated ingestion.</p> <p>When the end of the Faculty contract is reached on 2nd February 2024, Homeless Link will assume responsibilities for the management and maintenance of the tool, under contract with the GLA to deliver this as a service. It is proposed that the data will then be stored in the Homeless Link AWS tenant.</p>	
4.11	<p><b>Data Retention Period</b></p> <p><i>How long will the data be kept?</i></p> <p>Organisations have different retention periods for the data they already hold.</p> <p>Personal data retention in the project will be 5 years, to allow data matching over time. A key goal of the Strategic Insights Tool for Rough Sleeping is to better understand what happens to rough sleepers across their journeys into, during, and after they sleep rough, which will inform policy and service decisions. For most people who experience rough sleeping, the crisis is singular, short, and requires little intervention. For a few, time spent living on the streets is an intermittent occurrence amid other life challenges over a more extended period of time. These few long-term or episodic cases are the ones likely to need more services.</p> <p>Being able to see the full user journey for those with higher service needs requires a longer timeline. The advent of rough sleeping is often a confluence of events over many years, from losing a job, developing a health problem, and then finally, perhaps, the loss of a supportive family member or</p>	

	<p>friend. Periods during which individuals have not been engaged with the organisations do not necessarily indicate a positive move away from the streets, and may, for example, result from time spent in prison or rough sleeping in another geographical location. Time on the street can be short, but it can also last months or years, followed by time in hostels or other temporary arrangements. This cycle can stretch over five years, or more. Being able to see journeys over a longer period of time will be immensely useful for more complex, and more service needing individuals.</p> <p>As part of the phase 1 review we have also considered what would be an appropriate ‘trigger’ for the 5 year retention period. It has been decided that this will be based on an individual rather than on events. So if an individual is still ‘live’ in the system, i.e. they have had any contact/interaction/event in the past 5 years, this data will be retained, but if there have been no contacts it will be deleted. There are some additional nuances with this in relation to accommodation types and whether there is an end date associated with them. For example if an individual is marked as in ‘settled’ accommodation and there have been no further interactions in the last 5 years, even if there is no end date, we class this as a positive long-term outcome and consider the data to have reached the disposal point.</p> <p>The 5-year retention period has been reviewed and is deemed to still be appropriate at this stage. It is also aligned with the way the DLUHC rough sleeping indicators are defined; if an individual hasn’t been seen rough sleeping for 5 years, they stop treating the individual as an existing rough sleeper, and this means that if they return to the streets 5 years and 1 day after their last rough sleeping episode, they are treated as a new rough sleeper.</p>
--	--

<b>4.12</b>	<b>Will this information be shared/processed outside the organisations listed above in question 3? If yes, describe who and why:</b>	Yes/No
	<p>Yes—the data will be extracted from each of the data controllers (local authorities and charity service providers) and then records will be matched, integrated and stored in a data warehouse, where aggregate data reports can be generated. This data warehouse is managed for the pilot by the contractor.</p> <p>The SIT is currently being managed by Faculty Science for the duration of the contract. These responsibilities will be assumed by Homeless Link, under contract with the GLA to deliver this as a service, after 2nd February 2024.</p>	Yes

**Step 5: Information Security Process**

<b>5.1</b>	<b>Is there an ability to audit access to the information?</b>	Yes/No
	<p>The only information visible through the tool to users is the aggregated outputs which are anonymised. The contractor currently audits who has logged on to the system, when, and for how long. This data is currently being pulled manually but there is a planned action to build in additional audit functionality ahead of handover to the GLA and Homeless Link.</p> <p>Access to the AWS account(s) containing data relating to this project is limited to Faculty AWS organisation administrators, and the core technical delivery team. Access to this data is managed and audited using AWS IAM (Identity and Access Management) profiles. Remove (programmatic) access to the raw data is protected by:</p> <ul style="list-style-type: none"> <li>• AWS authentication: The technical delivery team uses secret keys to access cloud services from their development environments.</li> </ul>	

	<ul style="list-style-type: none"> <li>● Connection firewall: Only certain IPs are white listed to access the data. The technical delivery team needs to use a VPN (Virtual Private Network) connection.</li> <li>● Database level credentials - a connection string to 'log in' to the RDS database</li> </ul>	
5.2	<p><b>How will access to information be controlled?</b></p> <p>The SITRS solution architecture has been designed with best practice security-by-design principles in mind.</p> <p>Currently access to the email identity and the login credentials associated with this identity (In-Form, CHAIN logins), is currently limited to Faculty's Senior Data Engineer working on the project only. In future, this may be extended to the core delivery team.</p> <p>Access to AWS account(s) containing data relating to this project are limited to Faculty AWS organisation administrators, and the core technical delivery team via Faculty's AWS SSO (single sign-on).</p> <p>Access to specific resources (such as the database) has further access restrictions, such as firewalls restricting access to only approved IP addresses and required VPN connections.</p> <p>All data is stored in Faculty owned AWS accounts. All data is stored in the eu-west-2 (London) region. As mentioned above, it is proposed for the data to be moved to Homeless Link's AWS tenant post 2nd February 2024 when they assume ongoing management and maintenance responsibilities under contract with the GLA.</p> <p>All data is encrypted whether in transit or at rest.</p> <ul style="list-style-type: none"> <li>● In Transit: When data needs to be transferred over the internet, e.g. when ingesting data from INFORM/CHAIN Salesforce applications, or serving data via the core backend API (Application Programming Interface) connections will be encrypted using Transport Layer Security (TLS) e.g. <b>https</b>.</li> <li>● At Rest: All data stored in the environment is encrypted at rest using managed encryption keys</li> </ul> <p>Users during this phase can only be added to the tool by a Faculty admin. After the planned handover to Homeless Link, they will assume this role alongside the responsibility for setting the starters/leavers process. They already do this for the CHAIN database so have a replicable process in place.</p> <p>We have opted for passwordless authentication. Passwordless logins are more secure than traditional passwords as they use a second factor of authentication that is more difficult for attackers to compromise. In the case of the SIT, users must authenticate via a 'magic link'. In practice, this means that users follow the web link to access the site where the tool is hosted, they are prompted to enter their email address (associated with being an approved user) and they then receive a 'magic link' to their email address. This link is specific to the associated account and expires after 5 minutes.</p> <p>A permissions model has also been set up such that participating organisations have access to an appropriate level of information.</p> <p>This is as follows:</p> <ul style="list-style-type: none"> <li>● <b>Admin:</b> Access to all matched data and all views (LOTI, GLA, London Councils, Homeless Link) <ul style="list-style-type: none"> <li>○ NB: This role has also been assigned to named users from Bloomberg Associates for testing purposes for a limited time in accordance with their consultancy work on behalf of the GLA.</li> </ul> </li> </ul>	

	<ul style="list-style-type: none"> <li>● <b>Local Authority:</b> Access to Greater London, the subregional view to which the LA belongs, and only the borough-level data that the LA has provided. LAs also have access to all data that is matched to data that they have provided.</li> <li>● <b>Service provider:</b> Access to Greater London view, and only data within the subregional / LA views that has been matched to data provided by the SP. No access to other service providers' data - this is all collated into an 'other' category where visualisations are broken down by service provider.</li> </ul> <p>During this early phase of the project we have chosen not to make any of the data publicly available. However a consideration for future phases is which parts of this data we may wish to publish. We also recognise that users may utilise the insights from the SIT in ways which could end up in the public domain (for example if an LA uses some of the information as part of a committee report). As the outputs are all anonymised there are no data protection concerns.</p> <p>The anonymity of outputs also allows organisations to share outputs and their use of the outputs with other partners or contractors without data protection concerns.</p>												
5.3	<p><b>What roles will have access to the information?</b> (list individuals or staff groups)</p> <p>Access to the personal data is limited to the organisation that provided it (Directors &amp; Local Authorities, Service Managers) and the contractor (Faculty). Approved users within participating organisations (e.g. Senior Officials &amp; Policy Makers, Directors &amp; Local Authorities, and Service Managers) will have access to the aggregate outputs based on the permissions model detailed above.</p>												
5.4	<p><b>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?</b></p> <table border="1" data-bbox="194 1283 1530 1489"> <tr> <td data-bbox="194 1283 639 1404">Username and password</td> <td data-bbox="639 1283 711 1404"></td> <td data-bbox="711 1283 901 1404">Smartcard</td> <td data-bbox="901 1283 968 1404"></td> <td data-bbox="968 1283 1393 1404">key to locked filing cabinet/room</td> <td data-bbox="1393 1283 1530 1404"></td> </tr> <tr> <td data-bbox="194 1404 639 1489">Secure 1x Token Access</td> <td data-bbox="639 1404 711 1489"></td> <td colspan="3" data-bbox="711 1404 1393 1489">Restricted access to Network Files</td> <td data-bbox="1393 1404 1530 1489">Y</td> </tr> </table> <p>Other: <i>Provide a Description Below.</i></p> <p>Detailed above.</p> <p>Access to the AWS account(s) containing data relating to this project is limited to Faculty AWS organisation administrators, and the core technical delivery team. Access to this data is managed and audited using AWS IAM profiles. Remove (programmatic) access to the raw data is protected by:</p> <ul style="list-style-type: none"> <li>● AWS authentication: The technical delivery team uses secret keys to access cloud services from their development environments.</li> <li>● Connection firewall: Only certain IPs are white listed to access the data. The technical delivery team needs to use a VPN connection.</li> <li>● Database level credentials - a connection string to 'log in' to the RDS database</li> </ul> <p>As part of the planned handover, a system-level security policy (SLSP) will be developed based on existing detailed security documentation, as well as a business continuity plan / disaster recovery (DR) protocol. These responsibilities sit with Homeless Link and this will be a requirement within their</p>	Username and password		Smartcard		key to locked filing cabinet/room		Secure 1x Token Access		Restricted access to Network Files			Y
Username and password		Smartcard		key to locked filing cabinet/room									
Secure 1x Token Access		Restricted access to Network Files			Y								

	<p>contract with the GLA to provide the ongoing maintenance of this tool. London Councils/LOTI and Faculty will support the development of these documents. We will also be conducting penetration testing as part of phase 2 which will help to identify and remediate any possible weak-spots in the system's defenses.</p> <p>Whilst the SIT is hosted by Faculty for the remainder of phase 2, there is reasonable DR protocol in place. This is broadly split into two buckets - recovery of data and recovery of the system/application.</p> <p>Data - Our primary data store is our database. A snapshot of this is taken on a weekly basis and retained for a month. This means at any point in time, if everything went wrong, or our data was deleted for some reason, we could recover the state of the database from the previous week's snapshot.</p> <p>System - Our entire application infrastructure is created and can be managed using an infrastructure as code tool called Terraform. This essentially means none of it is manually created, and simply put, by running a set of scripts we could recreate our entire infrastructure stack.</p> <p>Our application code, which includes the data pipeline code as well as the front end application, is all version controlled and stored separately in Faculty's Gitlab.</p> <p>Additionally, to generally support availability and recovery, Faculty host across 2 AWS availability zones, which basically means we by design can make use of 2 AWS UK locations which are distinctly isolated, so if one goes down, we pretty much automatically switch to the other.</p>	
5.5	<p><b>Is there a documented System Level Security Policy (SLSP) for this project? If yes, please embed a copy below:</b></p> <p>SLSP is required for new systems.</p> <p><i>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.</i></p>	<p>Yes/No</p> <p>Not currently but as detailed above there is a planned piece of work to develop a SLSP as part of phase 2 with Homeless Link (who will take over management and maintenance of the tool from Faculty). We do have detailed system security documentation in place which will form the basis of this.</p>
5.6	<p><b>Are there Business Continuity Plans (BCP) and Disaster Recovery Protocol for the proposed/existing system or process?</b></p> <p><i>Please explain and give reference to such plan and protocol</i></p>	<p>Yes/No</p>
		<p>Not currently but as detailed above there is a planned piece of work to develop this as part of phase 2 with Homeless Link (who will take over management and maintenance of the tool from Faculty). This will be a requirement stipulated within their contract with the GLA. Existing DR protocol within the Faculty environment is detailed above.</p>

5.7	<b>Is Mandatory Staff Training in place for the following?</b>	Yes/No	Dates
	• Data Collection:	Yes	
	• Use of the System or Service:	Yes	
	• Information Governance:	Yes	
	• Data Controllers / Suppliers	Yes	
	All data controllers and the processor have mandatory training as part of their existing processes. For use of the system, training is supplied to users from participating organisations by the contractor. A user guide has also been developed as an offline training resource, which clearly outlines how to use the system, how visualisations have been developed etc. Furthermore, prior to being onboarded to the tool, users must review and accept a Terms of Service governing acceptable use of the system.		
5.8	<b>Are there any new or additional reporting requirements for this project?</b>		
	• What roles will be able to run reports?		
	Authorised users of the tool, which includes but is not limited to: 1. Senior Officials & Policy Makers 2. Directors & Local Authorities 3. Service Managers (service providers)		
	• What roles will receive the report or where will it be published?		
	1. Senior Officials & Policy Makers 2. Directors & Local Authorities 3. Service Managers (service providers)		
	All data available within the tool is anonymised. Authorised users may wish to use it as part of existing reporting mechanisms which could lead to information being published internally or externally e.g. as part of a committee report.		
	• Will the reports be in person-identifiable, pseudonymised or anonymised format?		
	Population level–aggregate data that has been matched in the data warehouse. All data visualised in the tool is anonymised.		
	• Will the reports be in business sensitive or redacted format (removing anything which is sensitive) format?		
	No/N/A		
5.9	<b>Have any Information Governance risks been identified relating to this project?</b> (if Yes the final section will need to be completed)	Yes/No	
		Yes	

<b>Step 6: Identify and Assess Risks</b>			
<b>Describe source of risk and nature of potential impact on individuals.</b> Include associated compliance and corporate risks as necessary.	<b>Likelihood of harm</b>	<b>Severity of harm</b>	<b>Overall risk</b>
<i>Note: risks here are risks of this sharing ONLY. Signatories should have DPIAs for their own individual systems and methods, covering their local risks.</i>			
Data retention period	Medium	Low	Medium
Ensuring cybersecurity and due diligence processes	Medium	High	High
Data quality and matching	Medium	Medium	Medium
Access Controls vs Audit Controls	Medium	Medium	Medium
That suitable privacy information is provided to clients	Low	Low	Low
Organisations submit data outside the scope of the data request	Medium	Low	Medium
Re-identification	Low	Low	Low

<b>Step 7: Identify Measures to reduce risk</b>				
<b>Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in step 6</b>				
<b>Risk</b>	<b>Options to reduce or eliminate risk</b>	<b>Effect on risk</b>	<b>Residual risk</b>	<b>Measure approved</b>

Data retention period - Medium	<ol style="list-style-type: none"> <li>1. The project group have provided justification for a 5 year retention period for the personal data, to allow the project to follow journeys of individuals who experience rough sleeping on multiple occasions.</li> <li>2. The project considers this to be proportionate based on the pilot.</li> <li>3. The contractor can meet the retention period and automate retention and deletion.</li> <li>4. The retention period will continue to be considered throughout phase 2 rollout.</li> </ol>	Reduced	Low	Yes
--------------------------------	--	---------	-----	-----

Ensuring cybersecurity and due diligence processes - High	<ol style="list-style-type: none"> <li>1. Contract with the processor binding between relevant parties</li> <li>2. Cybersecurity was part of the due diligence for the procurement <ol style="list-style-type: none"> <li>a. Subprocesses for Article 28 reference</li> </ol> </li> </ol> <p>Further cyber security activities will be completed during phase two, including pen testing and establishing disaster recovery and business continuity plans.</p>	Mitigate	Medium	Yes
---	--	----------	--------	-----

Data quality and matching - Medium	<p>Client records from each database are matched, using probabilistic matching, which uses correlations between data fields to determine the statistical likelihood that multiple records are the same. The current recall rate demonstrates a high degree of accuracy at 91% with the missing 9% due to data quality. The contractor has embedded into the process approaches to address and resolve conflicts.</p>	Tolerate	Medium	Yes
------------------------------------	--	----------	--------	-----

Access Controls vs Audit Controls - Medium	We have ensured that the new system has suitable audit and access controls. These are being further considered and developed before the tool is handed over from the contractor.	Reduced	Low	Yes
That suitable privacy information is provided to clients - Medium	The use of personal data for research is considered 'not incompatible' so privacy information provided for the original processing purposes continues to apply.  Parties are likely to take different approaches to whether they amend the privacy information they provide to data subjects. During phase 1 and preparations for phase 2, several organisations have updated their privacy information to bring it in line with best practice and all parties confirm by signing the DSA that they have suitable privacy information.	Tolerated	Low	Yes
Organisations submit data outside the scope of the data request - Medium	A data checklist has been developed to ensure the data request is as specific as possible.  Conversations with phase 2 organisations have been frontloaded to ensure they understand the scope of the data request.  All data outside the scope of the request has been removed from the database. Automated process now in place to remove any data transferred in the future that is out of the scope of the request.	Reduced	Low	Yes
Re-identification - Low	Followed best practice, in line with ONS, to not disclose outputs less than 5.  "Equal or below 5" shown in visualisations where it would be possible to filter down to less than this number.	Tolerate	Low	Yes

**Step 8: Sign off and record outcomes**

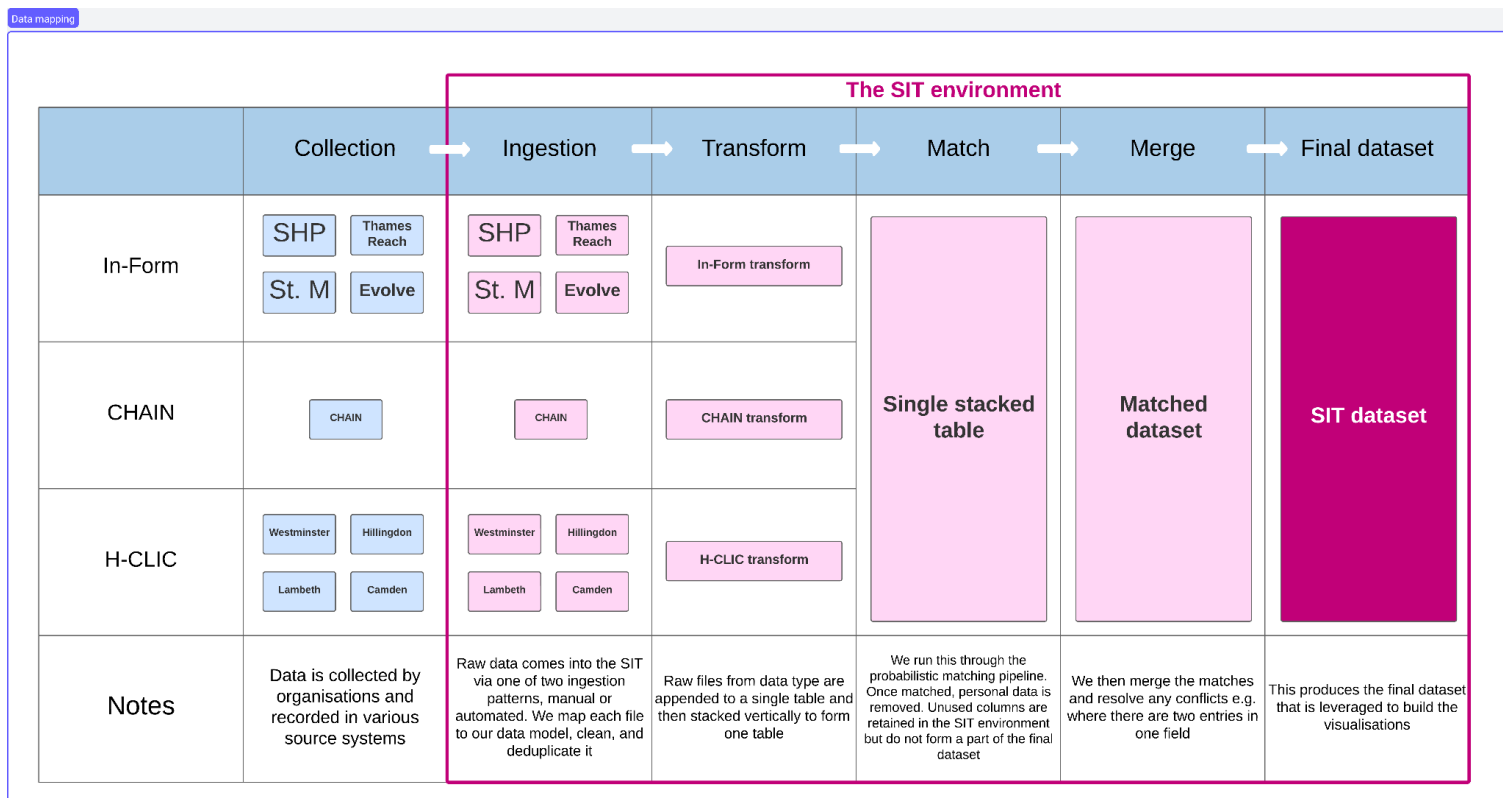
Item	Name/date	Notes
Measures approved by:	Anna Humpleby	
Residual risks approved by:	Anna Humpleby	
DPO advice provided:	Dr Fatima Zohra 7th November 2023	
<p>Summary of DPO advice:</p> <p>The IG report of phase one was successful in considering the risks and benefits of the use of the data and identifying the actions necessary for the data processing to continue into phase two of development. An IG report of phase two must be completed before a move to phase three, to ensure that proposed benefits continue to be achieved, and that any issues of concern can be mitigated. A new/revised DPIA is required before the project decides to move to roll out the project more widely (phase 3).</p> <p>London Councils is responsible for the contract with Faculty and for assuring suitable cyber security standards for the system. Prior to the handover of system hosting from Faculty to Homeless Link, London Councils and Homeless Link will need to ensure that suitable cyber security standards apply to the system, along with suitable plans for Disaster Recovery and Business Continuity. London Councils and Homeless Link must arrange for the cyber security of the system to be approved by IT security professionals during phase 2.</p> <p>All phase two data controllers are prompted to review their privacy notices.</p>		
DPO advice accepted or overruled by:		If overruled, 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

## Appendices

### Appendix 1: Updated Phase 2 Minimum Dataset / Data Checklist

Required fields from all sources (H-CLIC coding in brackets)	Required fields from H-CLIC only	Fields that were part of the original minimum dataset and are to be used for further feature developments
Unique client ID (1.1, A1.2)	Eligibility (1.3, 1.6)	Staying safe plan
Client name (A1.4, A1.5)	Application date (1.22)	Disabilities
DOB (A1.3)	Priority Need (7.5)	Relationship status
Preferred language (not in H-CLIC)	<b>Duty information (sections 4, 5, 7, 8)</b>	Immigration status
Ethnicity (1.7)	Duty type	Pregnancy status
Gender (1.24)	Duty start date	Employment / economic status
Nationality (1.9)	Duty end date	Armed forces history
Sexual orientation (1.8)	Location of duty event	Prison history
Phone number (not in H-CLIC)	Duty activity	Current mental health concerns / substance / alcohol misuse
Email (not in H-CLIC)	Reason for duty ending	Current medical needs
NI Number (A1.6)	Duty outcome / destination	Domestic violence services needs
NHS Number (not in H-CLIC)		Care leaver history
CHAIN ID (not in H-CLIC)		Entitlement to welfare benefits
Accommodation start date (9.2, 9.3)		
Accommodation end date (9.4, 9.9)		
Accommodation type (9.5)		
Accommodation location (9.8)		
Reason for leaving / accommodation end reason (7.7, 7.8)		
Departure destination / outcome		

## Appendix 2: SIT Data Flows



## Appendix 3: Service Providers lawful basis conditions

Organisation	Lawful basis
SHP	Article 6 (1)(e) public task Article 9(2)(g) substantial public interest
Homeless Link (CHAIN)	Article 6 (1)(e) public task Article 9(2)(g) substantial public interest
St. Mungo's	usually Article 6 (1)(e) public task usually Article 9(2)(g) substantial public interest
Thames Reach	usually Article 6 (1)(f) legitimate interests Article 9(2)(g) substantial public interest
Evolve	Article 6 (1)(f) legitimate interests Article 9(2)(g) substantial public interest
Look Ahead	Article 6(1) (e) public task or Article 6(1) (f) legitimate interests Article 9(2) (g) substantial public interest

Providence Row Housing Association	Article 6(1) (f) legitimate interests Article 9(2) (g) substantial public interest
Salvation Army	Article 6 (1)(e) public task Article 9(2)(g) substantial public interest
YMCA St Paul's	Currently under review by the organisation. It is expected that they will rely on Article 6(1) (e) public task or Article 6(1) (f) legitimate interests and Article 9(2) (g) substantial public interest or Article 9(2) (j) archiving and research. Consent is not the appropriate lawful basis for this research project.
Your Place	Currently under review by the organisation. It is expected that they will rely on Article 6(1) (e) public task or Article 6(1) (f) legitimate interests and Article 9(2) (g) substantial public interest or Article 9(2) (j) archiving and research. Consent is not the appropriate lawful basis for this research project.
Trinity Homeless Projects	Currently under review by the organisation. It is expected that they will rely on Article 6(1) (e) public task or Article 6(1) (f) legitimate interests and Article 9(2) (g) substantial public interest or Article 9(2) (j) archiving and research. Consent is not the appropriate lawful basis for this research project.
Centre Point	Article 6(1) (e) public task or Article 6(1) (f) legitimate interests Article 9(2) (g) substantial public interest

#### Appendix 4: User Requirements

Types of users:		
1. Senior Officials & Policy Makers		
2. Directors & Local Authorities		
3. Program Managers (service providers)		
4. Case workers, outreach (service providers)		
5. Persons sleeping rough		
Rec(s)#	User	User Requirement (Functionality)
<b>Research and Planning: Identifying trends, population needs and tracking outputs from the user journey (temporal)</b>		
1	Senior Officials & Policy Makers, Directors & Local Authorities, Service Managers	Understand the common barriers preventing those sleeping rough to improve access the support they need and move off the street
2	Senior Officials & Policy Makers, Directors & Local Authorities, Service Managers	Understand what happens to rough sleepers after they move on to <u>short-term</u> accommodation, in order to understand any needs for improvement to services
3	Senior Officials & Policy Makers, Directors & Local Authorities, Service Managers	Understand what happens to rough sleepers after they move on to <u>long-term</u> accommodation or other solutions (e.g., reconnection), in order to understand any needs for improvement to services
4	Senior Officials & Policy Makers, Directors & Local Authorities, Service Managers	Track statutory offers and outcomes of statutory interventions to understand effectiveness of supports and providers
5	Senior Officials & Policy Makers, Directors & Local Authorities, Service Managers	--Identify trends and emerging issues and promptly act, commissioning a solution --Segment the rough sleeping population into different cohorts according to their housing status/needs to better tailor services
6	Senior Officials & Policy Makers, Directors & Local Authorities, Service Managers	Understand the pathways into homelessness and take action to reduce the factors that cause and contribute to rough sleeping through preventive or diversion services
<b>Reporting of Outcomes for Accountability</b>		
7	Directors & Local Authorities, Service Managers	Generation of reports for these users. Functionality should include: --Ability to report in <u>multiple formats</u> , to ease administrative burdens of required reporting to multiple entities/funders --Allows <u>self-generation</u> of reports, rather than that having to make a data request
8	Senior Officials & Policy Makers, Directors & Local Authorities, Service Managers	Generation of reports for these users. Functionality should include: --Run reports on activity/services available outside immediate area of responsibility to support collaboration --Run reports to provide evidence of what works and examples of good practice to guide policy, funding and commissioning
9	Directors & Local Authorities, Service Managers	Helps fulfill/report on Performance (KPI) of services delivered/commissioned
<b>Case Management</b>		
10	Directors & Local Authorities, Service Managers, Case workers	Upload or link (user owned data) with other data systems or modules (e.g., other case management systems)
<b>Data Reporting Adaptability and Improvement</b>		
11	Senior Officials & Policy Makers, Directors & Local Authorities, Service Managers	Informs a process of standardized data collection across the sector to align measures, improve comparability, and reduce administrative waste
12	Senior Officials & Policy Makers, Directors & Local Authorities	Capability to easily adapt to new data requirements (e.g., changes to existing databases, add-on of new databases, dataset modifications, new policies such as GDPR update)