Citation for published version:

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Data Management Plan

**Core Principles underpinning this plan**
1. Publicly-funded research data are a public good, produced in the public interest.
2. Publicly-funded research data should be openly available to the maximum extent possible.
3. Researchers have a duty to give close attention to the research process, to uphold the dignity and rights of research participants, and seek to generate new knowledge that can be of use, and (re)used by others.

**Existing Data**
The research objectives require data not available from other sources. A search of the Discover UK Data Service database using the terms “refugees in Glasgow” found 5 entries, four of which focused on UK policy and were not approaching the topic comparatively. Some data that exists (Moving on: Integration and onward migration of dispersed refugees in the UK, Persistent identifier 10.5255/UKDA-SN-851857) can be brought into situate and triangulate the findings of the proposed research which will supplement data collected as part of the proposed research. However, existing data are not sufficient to address the objectives of this study across four cases and so new data are needed.

**New Data**
The project aims to interview two hundred stakeholders in four urban locations across four substantive workpackages (50 each in Glasgow, Nicosia, Malmo and Calabria). The main datasets created will include:

- The interview transcripts with stakeholders in PDF format; these will be transcribed professionally, anonymised at point of transcript but retain broad information on the informants’ sector and service represented, expertise etc.
- Videos and photographs generated in the Urban Living Labs, including photographs of their local area, maps and photographs of their everyday activities will be used for analysis; these will be provided in JPEG for photographs and MPEG.4 for videos, provided they are relevant to the theme of the research and young people consent to sharing them for other researchers’ purpose;
- Materials produced mobile exhibition will also consist of photographs and perhaps texts (visual maps, collage, objects etc.), and we plan to digitise these by photographs in JPEG format and make these available online, but also deposit them with the other datasets.

Following the guidance from the UK Data Service on documentation to be provided for archiving, the protocol supplied with the above data will include information on:

- the context of data collection: this will take the form of the project proposal, including the aims, objectives, research themes and questions and copies of research instruments;
- data collection methods: this will include a position paper on the methodology for the study prepared in advance of the fieldwork, including data collection protocols, sampling design, instruments used, hardware and software used, temporal and geographic coverage, and transcription methods
- the protocol will also provide information on the structure of data files (which geographical area covered, number of cases and records, relationships between files) and information on processes of data validation, quality assurance procedures carried out (including checking, proofing, cleaning), or editing, proofing or quality control of materials (for example, conventions in transcribing, translation if required etc.)
- we will also provide information on data confidentiality, access and use conditions, where applicable, and records produced during the fieldwork but not available, especially when people do not consent to the sharing of the data for other research.

**Quality Assurance**
In order to maintain high quality assurance during data collection, entry and analysis:
1. Any open-ended interviews will contain a semi-structured, standard set of questions. While it is crucial for the interviewees to narrate their own insights and experiences, standardization will aid later comparative analysis.
2. Data will be collected and stored using digital audio recording (e.g., MP3) where interviewees and permit. In case they do not, interviews will be undertaken with note-taking. All notes will be typed up according to agreed formats and standards.

3. All interviews and focus groups will be transcribed verbatim by an experienced transcriber, making sure that summarizing is avoided.

4. Procedures for analysing the interview and survey data will be refined and tested during each stage of interview/surveys and with each set of participants.

5. All interviewees and subjects of visual materials will be provided with a copy of their transcript and/or visual materials to check for accuracy and confirmation that the material can be treated as data.

6. Analysis of transcripts will be conducted with the use of NVivo for the data from the scoping interviews, focus groups and case studies, based on an agreed coding framework developed jointly by the team. We will provide the raw data (anonymised at point of transcript), coding tree, coded data, and associated memos and notes, as developed for the purposes of the research.

Back-up and security of data
All data collected for the project will be securely stored on institutional servers as soon as possible after the data collection has taken place. Researchers will follow a protocol by which data is collected in audio format with digital recorders and then downloaded into the university systems under a secure space created for the project. The University of Strathclyde’s provided storage is dual sited and replicated between two data centres which are physically separated by several hundred metres. Data links between datacentres are provided by dual disparate fabrics, providing added resilience and the central I.T. service provides tape based backup to a third and fourth site, to back up data from each of the two data centres. The University also implements file system snapshots, which allow users a certain degree of self-service recovery for recently deleted data an enhanced data security interface is available (cloud-type ‘Strath-cloud’) which is backed up several times during the day and will reduce risk of data corruption or loss.

Data security and storage
Most data produced will be in a digital format and stored through the secure data management systems mentioned above. All data collected in Scotland will be stored at the University of Strathclyde, Data security is provided by access controls defined either at a user level or group based where the data is being stored in shared data space. These databases will only be accessed by the team and by the vetted transcribers. Hard copies of transcripts, if produced, are likely to be anonymised, as we will implement a system of anonymising data at transcription stage. Other data which are sensitive include the participants’ personal data, collected through the surveys and consent forms- these will be stored in a locked filing cabinet in the PI’s.

Data encryption and file sharing
Management of data sharing will be an important consideration in the project, given the location of the two research teams involved. We propose to use as a collaborative virtual environment provided by the University of Strathclyde’s system, DMPOnline, providing a central, up-to-date and secure location to store research data and share them across the team. The data management plan for the project has been created in DMPOnline and will be further developed and shared via DMPOnline. The team will use this as a collaborative space with full access to all data for researchers and controlled levels of access to data for transcribers, with project-wide access to documentation and research outputs. Access will be controlled through individual accounts, making remote access easy, with no risk of transferring data files or loss of data.

Expected difficulties in data sharing
We do not anticipate major obstacles in sharing the newly generated data, other than the cases in which participants may deny consent for data sharing beyond the purposes of the project. In order to anticipate these concerns, we will explain to the participants the purposes of sharing the data, the steps we will take in fully anonymising the accounts and the specific sets of data we will want to use for depositing. We will respect the participants’ rights to give consent to data being used in the project, but to withhold consent for granting open access to data.

Copyright/Intellectual Property Right
For data collected via the scoping interviews, focus groups and case studies, which will be recorded and/or transcribed, the PI’s and first Co-I’s Universities will hold the copyright of recordings and transcripts, but each participant is an author of his or her recorded words in the interview (Padfield, 2010). In order to facilitate the use of transcripts for publication and data repository, we will aim to obtain transfer of copyright from interviewees to the research team, in writing, through an ‘assignment’ (transfer document), signed in conjunction with the consent form to be interviewed. Similarly, for visual data produced during the case studies, we will ask participants for transfer of copyright and permission to use the data in dissemination, art exhibition and for depositing the data from the project with the UK Data Service.

Responsibilities for data management
Responsibility for data management, preparation of data for sharing and archiving, safe storage and use of data and dissemination will reside with all team members; however, the PI will take the responsibility of establishing from the early stages a data management plan and ensuring that protocols for data collection and analysis are in place early on for quality assurance of data.

Role of PI: develop data management plan; ensure team agreement of use of data management plan; liaise with IT team at Strathclyde to ensure secure space organised for the project; manage and check transcribers and their access to the database; ensure discussion of data management in regular team meetings; main role for managing data collected in Scotland (collection, analysis, safe storage, preparation for archiving).

Role of the Co-Is: lead on designing templates for transcripts and templates for information collection on participants; oversee the preparation of data for archiving (anonymising, cleaning of data); main role for managing data collected in England (collection, analysis, safe storage, preparation for archiving).

Preparation of data for sharing and archiving
As detailed above, we will see data preparation for sharing and archiving as an ongoing activity throughout the project, to ensure data are cleaned and prepared for archiving at the end and costs are minimised later on. Qualitative data- transcripts of the qualitative data will follow a protocol agreed for the whole team (described above under ‘Quality assurance’). The coding tree, coded data, and associated memos and notes, as developed for the purposes of the research also will be provided in an NVivo format.