
Plan Overview

A Data Management Plan created using DMPonline

Title: Citizen Science To Achieve Coproduction at Scale (C-STACS)

Creator: Olamide Todowede

Affiliation: University of Nottingham

Funder: UK Research and Innovation (UKRI)

Template: UKRI Template

Project abstract:

In recent years there has been the emergence of citizen science - active public involvement in science. Citizen science has primarily been used in environmental research and more recently is starting to be used in health research. The need for innovative mental health research has been identified by mental health service users, with the call for recovery and a person-centred approach. This study aims to apply citizen science approaches in mental health, to improve the lives of the 16 million UK people currently living with mental health problems. This study comprises three work packages (WPs). In WP1 we will conduct a systematic review and wide stakeholder consultation in order to develop a conceptual framework for citizen science in mental health. We will then conduct two proof-of-concept studies. In WP2 we will run a citizen science project investigating self-management approaches that people living with mental health problems find useful in their lives. WP2 will supplement current clinical expertise by characterising coping and wellbeing strategies. In WP3 we will run a citizen science project using envisioning to develop new knowledge from citizen scientists about what a recovery-supporting mental health system looks like. WP3 will inform mental health system transformation. Overall, the goal of C-STACS is to develop a new approach to creating ecologically valid knowledge from the expertise of mental health stakeholders, especially from people with lived experience of mental health issues.

ID: 95290

Start date: 31-01-2022

End date: 31-07-2024

Last modified: 02-07-2022

Grant number / URL: <https://gtr.ukri.org/projects?ref=BB%2FV011707%2F1>

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Citizen Science To Achieve Coproduction at Scale (C-STACS)

0. Proposal name

0. Enter the proposal name

Citizen Science To Achieve Coproduction at Scale (C-STACS)

1. Description of the data

1.1 Type of study

C-STACS is a mixed-methods study that is divided into three (3) work packages.

1.2 Types of data

This project will collect the following types of data.

Data set 1: Semi-structured interviews

This is qualitative data consisting of participants demographic, interview transcripts and field notes.

Data set 2: Citizen science interactions

This is qualitative and quantitative data (non-personal and non-identifiable details) consisting of participant information, self-management ideas, ratings, and comments provided through an online citizen science platform. We will be using the citizen science methodology in the collection of data. In WP2, we will generate new knowledge from citizens about what a recovery-supporting mental health system looks like; and in WP3, we will generate data by investigating the self-management approaches those citizens living with mental health problems actually find useful in their lives. In this work package, the public will contribute as citizen scientists, completing and rating stories (WP2) and proposing and rating self-management strategies (WP3). Data will be provided by citizen science contributors who take part in a WP2 or WP3 project through an online citizen science platform.

1.3 Format and scale of the data

Data set 1: Semi-structured interviews

Semi-structured interviews with around 100 participants will be recorded using Microsoft Teams. These will be saved on OneDrive as audiovisual files (.mp4). These will be transcribed using the MS Teams transcription and anonymised and then saved as Microsoft Word documents (.docx) in UoN Onedrive. Data will be grouped into themes of the stakeholders using Microsoft Word documents (.docx) and uploaded into NVIVO software for analysis. Analysed NVIVO files (.sav) will be saved to OneDrive. Teams recordings (.mp4) will be deleted once transcribed. Field notes will be taken by hand during interviews, then typed up into Word file (.docx) and saved to Onedrive. Hardcopies will be shredded after typing out the field notes"

The volume will be up to 100. The transcript will be comprehensively anonymised (e.g. using pseudonymisation and data masking strategies) following UK Data Service guidance (www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation/qualitative) to make re-identification impossible. The demographic information will be labelled with alphanumeric and re-identification will be impossible. Optional consent will be requested from participants to allow for the re-use of anonymised data, and all data receiving such consent will be openly available and licensed under the creative commons license.

Data set 2: Citizen science interactions

This dataset will be generated from participants completing tasks and rating on an online platform and this will be downloaded to synthesise and analyse. This is the WP 2 and 3 of the C-STACS study.

The format data will be downloaded as Project Data (CSV format) and Talk Data (JSON format) exports will be generated using the Data Exports tab of the Zooniverse Project Builder. The volume of this dataset is difficult to estimate as this is first-in-field research, but probably between 250 and 100,000. All data collected will be openly available, and provided by people using a personal log-in so the study team will not collect or have access to personal information.

2. Data collection / generation

2.1 Methodologies for data collection / generation

In work package 1 (WP1) -We will be using a semi-structured interview and expert consultation with mental health and citizen science stakeholders to develop a conceptual framework. We will be generating data by capturing the perspectives of key stakeholder groups (citizen science academics/contributors, mental health service users, informal carers, and professionals). The study participants will be anonymized

In Work packages 2 and 3 - We will be using the citizen science methodology in the collection of data. **In WP2**, we will generate new knowledge from citizens about what a recovery-supporting mental health system looks like; and in **WP3**, we will generate data by investigating the self-management approaches those citizens living with mental health problems actually find useful in their lives. In this work package, the public will contribute as citizen scientists, completing and rating stories (WP2) and proposing and rating self-management strategies (WP3). Data will be provided by citizen science contributors who take part in a WP2 or WP3 project through an online citizen science platform.

2.2 Data quality and standards

Data set 1: Semi-structured interviews

The interview will be conducted by well-trained personnel (Olamide Todowede) and supervised by the Principal investigator (Mike Slade). The topic guide will be refined among the study team over time and will include a distress protocol. The University's automated transcription will be used to transcribe data it is potentially 90% accurate and secure. The interview will be scheduled at the convenient time of the participants and well planned ahead of time.

Data set 2: Citizen science interactions

The quality of data contributed on the citizen science platform will be validated and checked through the online platform.

3. Data management, documentation and curation

3.1 Managing, storing and curating data

We will use UoN-provided storage for all working data. UoN licenses Microsoft OneDrive, an ISO 27001 information security management compliant service that allows secure and controlled sharing of data amongst the research team (managed access). University of Nottingham Teams encrypt data both in transit and at rest and is approved against the University's Handling Restricted Data Policy. The service provides continual failover support. All project data will be kept within an O365 Team available to project members. No data will be saved outside of UoN managed systems, although curated data will be available on the citizen science platform which will be encrypted.

Data will be preserved and archived in the Research Data Management Repository and used for sharing data when required. PI Slade is the data asset owner for all data sets and has overall responsibility for all data management operations. Strategic decisions such as when to share anonymised interviews transcripts from data asset 1 outside of the project team will be taken by and enacted by the data asset owner. No personal data will be shared outside the study team. Resources are included for preparing data for archiving and ensuring access is in line with FAIR (Findable, Accessible, Interoperable, Reusable) principles.

The backup of data will be the responsibility of the data asset owner (Mike Slade), and effective backup procedures will be put in place for all data assets using OneDrive (with managed access) and/or MS Teams (with private channels), both of which are secure with routine and version-controlled backup functionality. Backup procedures will maintain equivalent levels of security as for primary copies of data. Confidential electronic data (i.e., data asset 1) will be stored in a password-protected location. Each interview participant will be allocated a unique identifier generated by the data asset owner before data collection, which will be used to index their data. It is envisaged that consent forms will be managed electronically, either recorded at the start of the videocall or completed using e.g., MS Forms or completed on paper and scanned before destroying the paper copy. Any identifying personal data will be stored separately from other data.

3.2 Metadata standards and data documentation

Data in O365 Teams will be stored in folders that will be labelled according to the data sources e.g. separate folders will be used for

data generated during the semi-structured interview sessions and saved according to the stakeholder's group. Also, data generated from the citizen science platforms will be saved in a separate folder and Separate folders will contain non-research project data. Metadata including protocols for data collection and analysis will be kept in separate folders. Participant-level metadata will be stored in addition to participant demographics: name, all known participants' contact details, date and time of the interview, and name of the researcher.

3.3 Data preservation strategy and standards

All anonymised research data and transcripts created by the project will be deposited in the UoN research data archive (<https://rdmc.nottingham.ac.uk/>) once the study has ended. The UoN data archive is underpinned by commercial digital storage which is audited on a twice-yearly basis for compliance with the ISO 27001 standard. UoN will retain and preserve research data for a minimum of 7 years, but data will be retained for longer periods of time where it is of continual value to users. No cost has been charged to this project for data archiving as we anticipate that the amount of data generated for long-term retention will not exceed 50GB (the capacity provided free by the University).

4. Data security and confidentiality of potentially disclosive information

4.1 Formal information/data security standards

The UoN data archive is underpinned by commercial digital storage which is audited on a twice-yearly basis for compliance with the ISO 27001 standard.

4.2 Main risks to data security

There is some risk that data will be lost but this is minimised by using the UoN systems. The use of UoN systems does protect the data and enable to limit of access to data. We will use UoN-provided storage for our working data. UoN licenses Microsoft Teams, allowing for secure and controlled sharing of data among the research team. Microsoft Teams encrypts data both in transit and at rest and is approved against the University's Handling Restricted Data Policy. The service provides several layers of automatic back up and, in a disaster scenario, files can be recovered. Access to data stored in MS Teams is via secure log-in with multi-factor authentication.

In the citizen science methodology of the study, if necessary, e.g. if specific patterns of contribution run the risk of de-identification, further anonymisation approaches will be used to ensure contributors cannot be identified.

5. Data sharing and access

5.1 Suitability for sharing

All anonymised data (from Data set 1 with consent for re-use will be available at University of Nottingham Research Data Management Repository (RDMR) (rdmc.nottingham.ac.uk) which allocates a citable DOI and URL for deposited data.

Data set 1: Records of each interview will aggregate the following data items: interview audio/video recording, verbatim transcript obtained using the secure and GDPR-compliant University of Nottingham automated transcription service with manual checking by the Research Fellow, an anonymised transcript, field notes, anonymous demographics information. All items will be linked by a unique identifier.

Retention periods for identifying data items, metadata and paper records will be decided by the data asset owner, informed by relevant local protocols.

Anonymised transcripts, demographic information and field notes will be shared with the broader project team during the project, transmitted where necessary in an encrypted form, with a password required for decryption.

The project team reserves a right to a period of privileged use of interview data up until the final publication has been produced, and all project investigators have the right to retain a copy of anonymised transcripts, fields notes and demographic data beyond the end of the project. Demographic information will only be released in summary form to avoid identifying participants. Some detail may be redacted to further protect participant identity and access can be withheld if the project team lack capacity to redact details that

might identify participants.

Data set 2: Project Data (CSV format) and Talk Data (JSON format) exports will be generated using the Data Exports tab of the Zooniverse Project Builder, during and at the end of the project, and made available in line with the Zooniverse User Agreement and Privacy Policy.

5.2 Discovery by potential users of the research/innovation data

Data will be published in open access peer-reviewed journals with raw data made available via the UoN Research Data Management Repository once published. Data will also be made available by request to the principal investigator (Mike Slade).

5.3 Governance of access

The PI (Mike Slade) of the project is the data owner and will have overall responsibility for the research data sharing. He will specify complexity rules and approve whether data should be supplied to a potential new user.

5.4 The study team's exclusive use of the data

Data will remain exclusive to the study team until the point of publication/patent filing.

5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions

All data is anonymised and personal information will be de-identified.

5.6 Regulation of responsibilities of users

This is not applicable to this study.

6. Responsibilities

6. Responsibilities

- The PI of the project is the data owner and will have overall responsibility for the research data.
- The Research Fellow will act as the Data Steward having day-to-day responsibility for the management of the data, for adding and removing access to the data, and for ensuring the DMP and relevant policies are followed. They will also ensure that data is archived and shared at the end of the project although the PI will be the direct contact for enquiries after this point.
- All project members are required to follow the DMP and to have read the relevant UoN policies. All project members are responsible for their own use and management of data.

7. Relevant policies

7. Relevant institutional, departmental or study policies on data sharing and data security

Policy	URL or Reference
UoN Code of Research Conduct and Research Ethics	(code-of-research-conduct-and-research-ethics.pdf (nottingham.ac.uk))
UoN Research Data management Policy	University of Nottingham Research Data Management Policy (sharepoint.com)
UoN Information Security Policy	
UKRI	https://www.ukri.org/publications/mrc-data-sharing-policy/
UK Data Service guidance	www.ukdataservice.ac.uk/manage-data/legal-ethical/anonymisation/qualitative

8. Author and contact details

8. Author of this Data Management Plan (Name) and, if different to that of the Principal Investigator, their telephone & email contact details

Dr Olamide Todowede
Olamide.todowede@nottingham.ac.uk