

WORK PACKAGE (WP) 6 RESPONSIBLE INNOVATION OF THE SYNTHETIC BIOLOGY APPLICATIONS TO WATER SUPPLY AND REMEDIATION PROJECT

DATA COLLECTION

What data will you collect or create?

We will be collecting a range of quantitative and qualitative data. These include:

1. One on One Interviews with stakeholders in the water industry.(n=6) in order to explore their perceptions of innovation and the role and potential of new technologies in the water sector.
2. Interviews with Principle investigators (PI's) on the Frontier Engineering Project (n = 5). These are conducted at the beginning of the project in order to assess the PI's perceptions of responsibility and interdisciplinary working.
3. Reconvened Focus groups (RFG) with members of the public (n = 6). We will use these RFG's to probe current public understandings of syn bio and present participants with research conducted in WP 1- 5 to gauge initial reactions. we will also use these to introduce some groups to syn bio in the lab in order to assess impacts on perceptions and understandings
4. We will administer questionnaires to participants in the RFG so that we can gauge any changes in perceptions and understandings prior to and post the RFG.
5. Ethnographic notes based on laboratory based research and observations undertaken by [Researcher1], written up as 'Fieldnotes'. (Currently n = 27).
6. We will write 'Briefing Notes' and 'Research Summaries' for the research scientists working on WP 1 - 5 that explore the political, economic and ethical issues of the laboratory research and in relation to the wider research and innovation process.
7. We will administer questionnaires to the research scientists in order to explore perceptions of their research and to explore the underlying values that drive research.
8. We will use open source software such as Tweetcatcher and NodeXL - as well as writing code in R - to search and store tweets from the social networking site Twitter. Our focus is tweets about synthetic biology and responsible innovation. We will analyse the changing language used to talk of both issues over the course of the project and undertake social network analysis to explore the stability or otherwise of the networks that discuss these issues.
9. We will create an agent based simulation of the SynBio innovation system using the open source program NetLogo. We will model the operation of the syn bio eco system as well as introduce responsible innovation as a factor in that model and explore the intended and also any unintended consequences that emerge from modelling a complex, adaptive system. This has the potential to provide useful data on the ways in which RI can impact research and innovation and what kinds of RI are most effective.

10. We will use QSR NVivo 10 to store videos (from sources such as YouTube) and audio recordings of public conferences and workshops on synthetic biology and responsible innovation.
11. We are using Evernote as a means to store websites that are of interest to the project (free to use software)

How will the data be collected or created?

For all interviews and reconvened focus groups (RFG) we will audio record them using the 'Easy Voice Recorder' software and they will be stored as .wav files. Both one on one interviews and RFG's will be undertaken using a semi structured interview protocol. The FRG's will be transcribed and the interviews partially transcribed. We will create a master Folder 'Frontier Engineering. Workpackages 6' and a number of sub folders, so for these data we will create three separate folders labelled as:

- Reconvened Focus Groups.
- Stakeholder interviews.
- PI Interviews.
- Postdoc Questionnaires.
- FieldNotes
- Briefing Notes
- Research Summaries
- SynBio Tweets
- Responsible Innovation Tweets
- RRI Tweets
- Agent Based SynBio Innovation Ecosystem

[Researcher2] and [Researcher1] hold regular meetings and discuss data storage and data capture, with both reading through and exploring the data.

The questionnaires to capture perceptions of the laboratory researchers will be held initially in word (.docx) format in a Folder entitled:

- Reconvened Focus Groups.
- Stakeholder interviews.
- PI Interviews.
- Postdoc Questionnaires.
- FieldNotes
- Briefing Notes
- Research Summaries
- SynBio Tweets
- Responsible Innovation Tweets
- RRI Tweets
- Agent Based SynBio Innovation Ecosystem

The Ethnographic fieldnotes will be kept in a folder:

- Reconvened Focus Groups.
- Stakeholder interviews.
- PI Interviews.

- Postdoc Questionnaires.
- FieldNotes
- Briefing Notes
- Research Summaries
- SynBio Tweets
- Responsible Innovation Tweets
- RRI Tweets
- Agent Based SynBio Innovation Ecosystem

The Briefing Notes and Research Summaries will be kept in two folders:

- Reconvened Focus Groups.
- Stakeholder interviews.
- PI Interviews.
- Postdoc Questionnaires.
- FieldNotes
- Briefing Notes
- Research Summaries
- SynBio Tweets
- Responsible Innovation Tweets
- RRI Tweets
- Agent Based SynBio Innovation Ecosystem

The tweets returned from searching using open source software will be kept in a excel file with three folders

- Reconvened Focus Groups.
- Stakeholder interviews.
- PI Interviews.
- Postdoc Questionnaires.
- FieldNotes
- Briefing Notes
- Research Summaries
- SynBio Tweets
- Responsible Innovation Tweets
- RRI Tweets
- Agent Based SynBio Innovation Ecosystem

The agent based simulation will be kept in a folder

- Reconvened Focus Groups.
- Stakeholder interviews.
- PI Interviews.
- Postdoc Questionnaires.
- FieldNotes
- Briefing Notes
- Research Summaries
- SynBio Tweets

- Responsible Innovation Tweets
- RRI Tweets
- Agent Based SynBio Innovation Ecosystem

We will also put the simulation on [Researcher1]'s GitHub page so others can fork to it and contribute, which will also act as a form of peer review.

QSR Nvivo data re stored in the 'NVivo' folder.

Evernote data is stored on cloud based service not on desktop/laptop.

DOCUMENTATION AND METADATA

What documentation and metadata will accompany the data?

We will record metadata in a .txt file that accompanies the data collection. We will include information that will describe key aspects of the data, using DDI best practise, including:

- Principal investigator
- Funding sources
- Data collector/producer
- Project description
- Sample and sampling procedures (if appropriate)
- Substantive, temporal, and geographic coverage of the data collection
- Data source(s)
- Variables
- Technical information on files
- Data collection instruments
- Citations to related publications
- Technical information on files, e.g., information on file formats, file linking
- Data collection instruments
- Interviewer guide
- Coding instrument

Archives also value receiving the following information

- Principal investigator
- Funding sources
- Data collector/producer
- Project description
- Sample and sampling procedures (if appropriate)
- Substantive, temporal, and geographic coverage of the data collection
- Data source(s)
- Variables
- Technical information on files
- Data collection instruments
- Citations to related publications
- Technical information on files, e.g., information on file formats, file linking

- Data collection instruments
- Interviewer guide
- Coding instrument

ETHICS AND LEGAL COMPLIANCE

How will you manage any ethical issues?

All interviewees and focus group participants will sign a Consent form agreed to by the School ethics committee.

We have guaranteed anonymity to our interviewees and focus group participants. Therefore we will not be depositing .wav files as this would compromise that guarantee. However, anonymised transcripts of the interviews and focus groups will be deposited.

We will make sure consent forms make provision for future sharing of data.

All identifying information will be kept in a locked filing cabinet and not stored with electronic files.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

The research is not expected to lead to patents.

IPR issues will be dealt with in line with University of Glasgow policy and necessary guidance will be sought from the IPR and Commercialisation Team

STORAGE AND BACKUP

How will the data be stored and backed up during the research?

Storage and back up will be in three places:

- On Laptop of [Researcher1]
- On a portable storage device (hard drive)
- On the University of Glasgow server.

[Researcher1] will be responsible for the storage and back up of data. This will be done weekly.

How will you manage access and security?

Both the laptop and external storage device will be password protected. The risks are that the computer will be hacked and the external drive stolen. The laptop has anti-virus software installed which is updated daily. The external device is in a locked cabinet. Moreover the data files will also be password protected. The paper consent forms from interviewees/focus group participants will be kept in a locked cabinet in the university office of [Researcher1].

SELECTION AND PRESERVATION

Which data are of long-term value and should be retained, shared, and/or preserved?

We will preserve the data for 10 years on university servers and also deposit the data in a UK data archive

What is the long-term preservation plan for the dataset?

We will deposit our data at the end of the project in an appropriate data depository e.g. UK Data Archive and/or the Consortium of European Social Science Data Archives. There is no additional cost.

Where possible we will store files in open archival formats e.g. word files converted to .txt files and excel files converted to .csv. Where this is not possible we will include information on software used and version number.

DATA SHARING

How will you share the data?

Some of the ongoing data will be shared on [Researcher1]'s Github page (results and code from NetLogo, data from twitter searches)

For all other data we will use the University of Glasgow Institutional Data Repository, 'Enlighten: Research Data', where we can direct interested parties, rather than handling each request individually.

Are any restrictions on data sharing required?

Data which underpins any publication will be made available at the time of publication.

All unpublished data will be deposited in a data repository 12 months after the end of the award.

RESPONSIBILITIES AND RESOURCES

Who will be responsible for data management?

[Researcher1] will be responsible for

- data capture
- data management
- metadata production
- storage and backup
- data archiving
- data sharing

What resources will you require to deliver your plan?

We have costed in time and effort within the Frontier Engineering grant to prepare the data for sharing / preservation. We have sufficient storage and equipment to undertake these tasks.