PRIVACY IMPACT ASSESSMENT

PROJECT NAME:	**Edinburgh CitySounds**
PROCESS DESCRIPTION:	Initial draft, SC: 7th February 2018
	v0.1c, SC: 22nd February 2018, following review and discussions
	v0.1d, EK: 23rd Feb 2018, edits
	V0.2a, EK: 26th Feb 2018, incorporating input from CM
	V0.3a SC, ST: 26th February 2018, completed missing sections.
	V0.3c SC: 27th February 2018, added GDPR compliance information following discussion with RG (DPO).
	V0.3d SC: 28th February 2018, updated box safety following review by TW.
	V1.0A SC: 28th February 2018, risks table updated by ST.
	V1.0b EK: 06th March 2018. Corrected Informatics Ethical review wording.
	V1.0c EK&SC: 9th March 2018, version for public release
	V1.0d SC: 11th March 2018, minor tweaks to correct grammar errors and to clarify use of manual screening and voice scrambling.

Date PIA started:	7th February 2018

Section 1- Preparation for Screening

1.1 Project Outline – what and why

Note: Explain the scope of the project to ensure you know its aims and its potential impact, explain what the project consists of and why it is undertaken. Map the data flows – where do you obtain the data, how are they processed, where are they stored.

Scope

Edinburgh CitySounds is one of 17 experiments selected for the second phase of OrganiCity (http://organicity.eu/meet-new-experimenters). OrganiCity is an EU funded service for experimentation that explores how citizens, businesses and city authorities can work together to create digital solutions to urban challenges.

Edinburgh CitySounds will explore and celebrate the richness of sounds in the city, benefiting from recent innovations in digital technology and network infrastructure. It will focus on how biotic (wildlife based), abiotic (weather-related) and anthropogenic (human activity based) sounds captured in a central urban greenspace can inform community groups and citizens about biodiversity and health and well-being, as well as provide a unique resource for artists and data scientists.

Please see https://citysounds.eu for additional background information.

The project includes several community engagement workshops at which participants will contribute views on how the soundscape data we collect could be used to address issues within the community. As well as an interest in urban biodiversity, noise pollution is one obvious example that may be raised through this consultation – the data we collect would enable the times and periods where excessive noise may be present within the Meadows area to be identified.

Purpose

This is an experimental research project whose primary purpose is the long-term collection of soundscape data 24/7 from the Meadows greenspace area in the centre of Edinburgh, to facilitate research in audio-based data science and with an initial focus on bioacoustics and biodiversity monitoring.

The audio samples that we would collect through a year-long natural cycle will constitute a unique and substantial soundscape dataset. With further work, this large body of audio data could provide valuable training input for machine-learning. This in turn will allow us to develop robust audio classifiers capable of recognising different categories of events within a noisy city soundscape. Such machine learning models would be capable of detecting the presence of birds, bats and other wildlife, and potentially distinguishing individual species on the basis of their sound signatures. Collecting data over this 12-month period will also enable us to demonstrate how biodiversity varies throughout the seasons.

Data Flows

The data flow and data processing scheme is described in detail in Appendix C. In this section of the PIA we provide high level overview.

The CitySounds system will only process newly collected audio data. Six Audio Capture Devices (ACDs) will be built and deployed at different locations within the Meadows area. These devices will capture wide spectrum audio in the range 0Hz to 96KHz, enabling audio traits both within and beyond normal human hearing range to be captured.

Each ACD will capture a 10 second sample in rotation, thus leading to six successive samples across the course of 60 seconds. The ACDs will continuously deliver their 10-second audio sample files via a dedicated WiFi network created for the project, using a secure file transfer protocol. These files will be sent to a dedicated Data Collector server using the University's IT infrastructure. The Data Collector carries out further analysis and processing, including frequency-based voice scrambling.

Once an appropriate subset of privacy-preserving audio samples has been selected by the project team, these will be published to a separate server operating as an Edinburgh city OrganiCity node. Users registered on the OrganiCity platform (see https://docs.organicity.eu/#accounts-and-registration) will be able to access these audio samples. The privacy-preserving steps we will take are discussed in the next section.

Protecting Privacy — Handling of Voice Traces

Potentially, some of the audio data we will capture may contain traces of voice from passers-by. If the content of such spoken utterances is intelligible, it might contain information that will identify specific individuals. As a result, we may incidentally be capturing personal data. However, we will not have any direct knowledge of who these voice traces belong to since we are not connecting to any other system or data source which would allow us to carry out data linkage and thereby determine the identity of the speaker or the identity of persons possibly mentioned by the speaker.

Whilst the number of occurrences of intelligible speech being recorded by the ACDs will be very small in practice, we are implementing a two-fold approach to protect against any potential privacy intrusions that could arise from voice traces that we may incidentally have picked up with the ACDs' microphones. This approach is motivated by the goal of being able to make some of the audio data available to third parties via the OrganiCity platform.

1. Most audio samples that we collect will contain data in which no speech at all is present. A subset of these samples will be of sufficient interest to a wider audience to be worth sharing, and would not require further processing. In such cases, we will manually verify that no voice is present, before deciding whether they should in fact be shared.

2. In other cases, it may not be feasible to manually confirmed that voice traces are not present, or we may not have been able to conclusively confirm absence of voice traces. In these cases, prior to any publication we will apply a voice scrambling process to the audio

sample that renders speech unintelligible. The voice scrambling algorithm is described in detail in Appendix D.

Only a relatively small number of audio samples that are to be published will be reviewed manually. To confirm that no voice traces are present in an audio sample, two members of the project team will independently review the sample. This will involve listening to the sample and viewing the sample's frequency spectrogram. The reviewer will record their judgements in a digital document stored in a secure location such as the project's SharePoint folder within the University's Office365 system.

1.2 List of stakeholders

Note: This should cover all individuals involved in the project and those that may be affected by it – internal and external stakeholders. At this stage you want to have as broad a list of groups as possible- this can be edited down at a later stage for more focused consultation.

- **Sponsor**: Ewan Klein (*PI for the project*, Informatics)
- Tony Weir (Director, IS-Infrastructure / ISG, *and nominated* ***Data Owner***)
- **Named Individual**: Simon Chapple (Senior Data Technologist, IS-Infrastructure: ACD design, build, operation. Data Collector, Reporting and Output Feeds)
- **Operational Systems Support**: CIS, EDINA
- Cat Magill (Informatics Researcher): project community liaison
- Stephen Taylor (IoT Programme Manager)
- Renate Gertz (University of Edinburgh DPO)
- Other University of Edinburgh project collaborators (Martin Parker, Jonathan Silvertown, Graham Stone)
- Peter Davidson, Park Ranger, CEC
- Sarah Hughes-Jones, CEC DPO
- External project partners (Scottish Wildlife Trust/Edinburgh Living Landscape, Friends of the Meadows, New Media Scotland)
- OrganiCity Administration

1.3 External context

Note: This involves conducting a search for prior projects of a similar nature, from both inside and outside the organisation. This may reveal design features that have been created by other project teams in order to address much the same categories of problem confronted by your project. Note any lessons that can be learned.

Sounds of New York City (SONYC)
https://wp.nyu.edu/sonyc/
https://www.6sqft.com/smart-microphones-are-recording-city-sounds-to-help-create-a-quieter-new-york/
https://www.raspberrypi.org/magpi/sounds-new-york-city/
November 2016: "Researchers at New York University and Ohio State University have installed microphones at points throughout New York city that will learn to recognize the pneumatic drills, bizarrely noisy Fresh Direct trucks and other street sounds that form our familiar daily cacophony. The recording

devices use technology that was developed to identify migrating birds, the way the Shazam app records and identifies song snippets. The study will begin collecting 10-second bits of audio at random intervals, then begin labelling the urban din using UrbanEars, a machine-listening engine. The sensors are being trained to identify the many 'sonic irritants' that plague city life, including the seasonal (snow plows, air conditioners) and the maddeningly ceaseless (garbage trucks, construction). The project, called Sounds of New York City (SONYC) has the goal of creating an aural map that could help the city track and control noise pollution in addition to empowering residents to get involved."

Quotes regarding privacy:

> In case you're worried about the sensors picking up bits of private conversation, Dr. Bello said conversations "heard" by the microphones "could not be reconstructed from the recordings," with assurance from an independent acoustical consultant hired to address this concern.
>
> SONYC has worked hard to ensure that the project doesn't encroach on privacy. "The audio data is collected in ten-second snippets," says Charlie.
> Furthermore, recordings are randomly separated in time to ensure privacy is maintained.
>
> "We have done a lot of work to maintain privacy on the project, and have had an external consultant confirm that street-level, intelligible speech at conversational levels cannot be picked up," insists Charlie. "A person would have to shout at the sensor for the speech to be intelligible, and that wouldn't constitute a private conversation."
> The team also deploy signs below each node to inform people what they are doing.

Lessons:
Precedent – What we are seeking to do has been achieved in a different context previously in a major urban conurbation, New York City. Therefore, with the addition of appropriate privacy-preserving techniques employed (we are not assuming that we will simply not pick up low level conversational speech; we are going the extra step), we should be able to achieve something similar here in Edinburgh. We will also be recording only in ten-second snippets at each Audio Capture Device location. We will ensure that appropriate information sheets are displayed in already established signage points in the Meadows.

Sound Shredding: Privacy Preserved Audio Sensing
Sumeet Kumar, Le T. Nguyen, Ming Zeng, Kate Liu, Joy Zhang
Conclusion in the paper: "Audio is a valuable source of contextual information, which is crucial for many context-aware mobile applications. However, beside context information the captured audio signals often contain sensitive speech content. In this work, we show that sound shredding and subsampling are effective means for making speech not recognizable, while preserving sufficient information for context, gender and speaker recognition.

Through the experiments, we showed that no speech content could be recognized from the processed signal by either human or automated computer techniques."

Lessons:
'Sound shredding' is an alternative term for the technique we have described as voice scrambling – this will ensure the content of any conversation incidentally captured is rendered unintelligible.

Stage 1 completed by:	Simon Chapple	Date:	8th February 2018
Revisions	Ewan Klein		26th Feb 2018

Section 2- Screening

Note: The information you have gathered in Stage 1 should assist you in addressing the screening questions, and if you find difficulty answering any you should consult the University's Governance Services.

2.1 Technology

2.1.1 Will there be new or additional information technologies that have substantial potential for privacy intrusion?
Yes. We have highlighted the potential for the Audio Capture Devices deployed in the public Meadows area to incidentally capture traces of individuals' voices. We have described earlier in this document (see also Appendix D) how the audio data will either be manually screened to ensure voice traces are not present, or scrambled to render individual's voices and the words they have spoken to be unintelligible in any of the sound samples that we may subsequently publish.

2.2 Data collection

2.2.1 Will the project involve the collection of new information about individuals?
No – we are not seeking to identify individuals or collect specific information about them.

2.2.2 Will the project compel individuals to provide information about themselves?
No.

2.3 Identification methods

2.3.1 Will there be new or substantially changed identity authentication requirements that may be intrusive or onerous?
No.

2.4 Involvement of multiple organisations

2.4.1 Will the initiative involve multiple organisations, whether they are public service partners, voluntary sector organisations or private sector companies?

Yes. The ACDs are deployed into the public Meadows greenspace area, and so we have engaged both with City of Edinburgh Council and local community groups such as Friends of the Meadows.

2.5 Changes to the way data is handled – considering the actual processing

2.5.1 Will there be new or significant changes to the handling of types of personal data that might be of particular concern to individuals? This could include information about racial and ethnic origin, political opinions, health, sexual life, offences and court proceedings, finances and information that could enable identity theft.

No.

2.5.2 Will the personal details about each individual in an existing database be subject to new or changed handling?

No.

2.5.3 Will there be new or significant changes to the handling of personal data about a large number of individuals?

No.

2.5.4 Will there be new or significantly changed consolidation, inter-linking, cross-referencing or matching of personal data from multiple sources.

No.

2.6 Changes to data handling procedures – considering policy documents and standards

2.6.1 Will there be new or changed data collection policies or practices that may be unclear or intrusive?

No.

2.6.2 Will there be changes to data quality assurance or processes and standards that may be unclear or unsatisfactory?

No.

2.6.3 Will there be new or changed data security arrangements that may be unclear or unsatisfactory?

No.

2.6.4 Will there be new or changed data security access or disclosure arrangements which many be unclear or permissive?
No.

2.6.5 Will there be new or changed data retention arrangements that may be unclear or extensive?
No.

2.6.6 Will there be changes to the medium of disclosure for publicly available information in such a way that the data becomes more readily accessible than before?
No.

2.7 Justification

2.7.1 Does the project's justification include significant contributions to public security measures?
No.

2.7.2 Is there to be public consultation?
Yes, through the community groups associated with the Meadows.

2.7.3 Is the justification for the new data handling unclear or unpublished?
No.

Stage 2 completed by:	Simon Chapple	Date:	8th February 2018

Stage 3- outcome of screening

3.1 Preliminary identification of risks

The table below lists the key privacy risks that have been identified by the screening process.

Note: You do not need to do a detailed assessment of the risks at this at this stage, but before proceeding with the PIA it is important to be reasonably clear about what the main risks are.

The Risk table present in the full PIA has been redacted from here for security reasons.

3.2 Decision on how to proceed

Note: From the work you have done above, you should now be in a position to determine whether you need to do a PIA, or whether a privacy law compliance check is sufficient. Record your conclusion below.

Due to the potential to incidentally record traces of a person's voice through our Audio Capture Devices, we consider it necessary to complete a PIA.

Name of decision officer	N/A Project team carried out a risk review which identified the need for a PIA.
Name of Information Management Team Member who agreed this decision	N/A

Stage 3 completed by	Simon Chapple	Date	8th February 2018

If you have decided a PIA is not necessary, go straight to **Stage 6** to complete the privacy law compliance check.

Stage 4 — Preparation for consultation and analysis

4.1 Governance arrangements

Note: Select the appropriate text below and complete the table. Add other names/roles as appropriate.

The PIA will be managed as a separate project within the IoT Programme. The following individuals will be on the project team.

Name	Their role
Steve Taylor	PIA and Project Manager
Simon Chapple	Technical Lead, device build, system build
Mark Strevens	Back-end System build
Enterprise Services	Back-end System build
CIS	Wireless receiver install and config
Ewan Klein	PI
Cat Magill	Community Liaison

4.2 Internal stakeholders to be consulted

Note: This builds on the work you did at 1.3 to identify the stakeholders. You should now consider in more detail what the interests of the various internal stakeholders are and the level of involvement they will have in the PIA.
.

Internal Stakeholders		
Governance and/or legal team (Involve where there are complex legal compliance issues)	Ensuring compliance with any relevant legislation	
Rena Gertz – DPO	Ensuring compliance with any relevant data protection legislation.	Will be asked to review and comment on the PIA.
Tony Weir – Director of IT Infrastructure Division.	Confirmation that the project and PIA approach is acceptable to ITI.	Will be asked to review and comment on the PIA.
IoT Governance and Ethics Action Group	Provide advice and guidance regarding governance and ethical	Will be asked to review and comment on the PIA.

	approach for IoT projects.	
Convenor of the Ethics Panel of the School of Informatics	Ensuring the project complies with the School of Informatics Ethics Code.	Will be asked to verify that the proposal adheres to the School's Ethics Procedure.
David McClelland, Snr IT Security Consultant	Provide advice and guidance regarding IT security matters.	Review the solution architecture and risk register and comment and advise as needed.

4.3 Consultation Plan

Note: If there is already a public consultation strategy or plan in place for the project, you don't need to have a separate one for the PIA. You do however need to make sure that it encompasses all the privacy aspects of the project.

Explain below what approach you are taking to the consultation plan.

Our consultation plan centres on a series of meetings and workshops involving partners and members of the public. As well as receiving verbal and written feedback from these events, we will establish a public website which explains the project and provides a facility for anyone to make comments or ask questions (website: citysounds.eu).
Our internal consultation plan consists of meetings with internal stakeholders and sharing the PIA for review and comment,

4.4 Resources

Note: Consider whether you will need additional resources to carry out the PIA effectively. This is most likely to be the case if your stakeholder consultation on privacy issues is going to be wide-ranging. Note the requirement, if any, below.

The project budget provides funding to support the level of community engagement we believe we require.

Stage 4 completed by	Steve Taylor	Date	26/02/2018

Stage 5- Consultation

Note: For Large PIAs, where there has been extensive consultation, you may wish to produce a separate consultation report, which should then feed into stage 7. Always complete Stage 6 to consider compliance with the Data Protection Act and other privacy laws.

5.1 External consultation

Note: Decide what type of external consultation will be most appropriate and will give you the best and most complete results – focus groups, mail shots, …

The project kick-off meeting involved representatives of all the external partners, including representatives of subsidiary or allied organisations.

Two public workshops took place on 19th February, following extensive advertising through a variety of channels, including social media and email. Representatives of the following organisations attended: UoE, Glasgow University, Sustrans, Scottish Wildlife Trust, Northumbria University, Royal Botanic Gardens of Edinburgh, Friends of the Meadows and Bruntsfield Links, Historic Environment Scotland, Greening our street, Upmo.

A further community engagement workshop will take place before the end of March.

Written information about the project has been disseminated to Friends of the Meadows, Scottish Wildlife Trust, and via them to their broader network of contacts.

5.2 External stakeholders

Stakeholder name	The privacy issues they raised
Scottish Wildlife Trust	None
Friends of the Meadows and Bruntsfield Links	None
New Media Scotland	None
City of Edinburgh Council DPO	None

5.3 Internal stakeholders

Note: Do not include members of the PIA team.

Stakeholder name	The privacy issues they raised
DPO	None.
Director of IT Infrastructure Division.	None.

IoT Governance and Ethics Action Group	None.
Convenor of the Ethics Panel of the School of Informatics	None.
Snr IT Security Consultant	One additional risk regarding brute force password attacks identified which is now included in the risk table in Appendix B.

Stage 6- Compliance with privacy laws

Note: The Data Protection Act (DPA) is relevant to any PIA, and a DPA compliance check should always be carried out. The Data Protection Officer will be able to advise you on the relevance of other privacy laws.

6.1 Data Protection Act 1998 (DPA)

Note: The template to use for the DPA compliance check is based on the one in Appendix 2 of the ICO's PIA Handbook v2, and can be found in Appendix A of this document. The Governance Officer (Data Protection & Legal) will be able to assist with this.

A Data Protection compliance check has been carried out as part of this PIA, the details of which are in Appendix A.

From this we have concluded that we are operating under Part IV Exemptions, section 33 of the DPA -research purposes. We are not directly processing personal data but have the potential to incidentally capture traces of voice. We are applying a strict manual screening process combined with voice scrambling technology to ensure privacy of individuals is preserved in any data we may share more widely.

From May 25th 2018 the GDPR will be in effect. Then, as things currently stand, we rely on Article 6 1.(f) [**legitimate interests**] for processing personal data. Should against all expectations a special category of personal data be captured, then Article 9 2.(g) [**reasons of substantial public interest**] would apply.

6.3 Human Rights Act (HRA) (Article 8)

Note: In most cases HRA considerations will be covered by the other work on this PIA, including the DPA compliance check. If that is the case, you can simply record here that there are no special considerations that are not covered by other aspects of the PIA.

HRA Article 8: Right to respect for private and family life
There are no special considerations that are not covered by other aspects of the PIA.

6.4 Privacy and Electronic Communications Regulations 2003 (amended 2011) (PECRs)

Note: See Appendix 3 of the ICO's Handbook for a PECR compliance check template.

We are not subject to the PECRs as we are not carrying out electronic communications with individuals as part of this project, and neither are we running a public access website that tracks user access through browser cookies.

6.5 Regulatory and Investigatory Powers Act 2000 (RIPA)

This project does not involve individuals' telecoms devices neither fixed nor mobile and neither does it involve the postal service. RIPA therefore does not apply.

6.6 Common Law duty of confidence

We are taking all steps to protect the privacy of individuals whose voice may incidentally be captured by Audio Capture Devices used in this system. Traces of voice will be scrambled by a robust algorithm rendering them unintelligible prior to any level of wider sharing. ACDs are not placed inside private dwellings.

6.7 Others

Please see GDPR referenced in section 6.1 which will apply from May 25th 2018.

Stage 7- Risk analysis

Note: You should carry out the risk analysis using exactly the same methodology as you do for other project risks. The table in Appendix B is provided as a guide only and should be adapted to conform to your project risk register.

The Guide to PIAs provides some useful pointers on the types of solutions to privacy risks that can be explored (see section 7).

The table in Appendix B shows the key risks that have been identified, and the mitigations applied to reduce those risks.

Stage 8- Approval

8.1 Recommendation

Note: Drawing on your analysis of the privacy risks and other project risks, explain which option presents the best way forward. If significant risk remains, you should explain what the problem is and why the stakeholder consultation failed to resolve this. Your recommendation may then be that the project needs to be re-thought.

We recommend that the project progresses as described in this PIA, including the mitigations identified for the risks.

8.2 Approval

Note: For large projects, this stage should align with the Full Business Case and approval should be given by the relevant budget holder. All you need to record below is who has approved the recommendation at 8.1 and the terms of that approval.

Ewan Klein approved.

Stages 5-8 completed by:	Steve Taylor & Simon Chapple	Date:	26/02/2018

Stage 9- Readiness for service

Note: Explain below what checks were carried out before the service went live to ensure that the privacy solutions approved as part of this PIA are working, and that the system or process is still legally compliant.

The technical design has been peer reviewed within the project team.

As of 26/02: A test plan is being devised that will include tests for all processes within the solution, along with expected results for each individual test. Each test will be carried out and either approved as successfully completed, or the issue addressed as needed.

This PIA has been carried out, and this document details the engagements and risk reviews that have been carried out regarding the production, review and approval of the PIA.

The Convenor of the Ethics Panel of the School of Informatics has reviewed the PIA and has concluded that the proposal does not need to be discussed at the School's Research Committee; within the framework of the School's Research Ethics Procedure, Level 2 self certification is sufficient.

Note: As this deployment does not include any changes to existing services, and does not require a technical cutover for go-live, CAB approval is not required, and therefore no Change ticket has been raised.

A number of pre-requisite actions needed before the first live ACD is deployed have been identified, and the project team will review and confirm that all of these are in place prior to the first live deployment in the Meadows.

Stage 9 completed by:	Steve Taylor	Date:	26/02/2018

Stage 10- Review or audit

Note: Indicate below how and when the post-implementation audit or review will be carried out.

Formal monthly reviews for first quarter, followed by quarterly reviews by SC, EK and ST. The frequency of reviews can be modified as a result of the outcomes of preceding reviews.

These reviews will consider:

- General overview
- Use of data by OrganiCity
- Risks and Issues

Stage 10 completed by:	Steve Taylor	Date:	26/02/2018

Appendix A

Data Protection Compliance Check

Note: completion of this template requires knowledge of the Data Protection Act 1998. It should be completed with assistance from the Governance Officer (Data Protection & Legal).

Where you have already provided the information at Stage 1 of the main PIA Template, simply cross-refer to the relevant answer.

	Question	Answer
1.	What type of personal data is going to be processed?	We may incidentally capture traces of people's voices within our audio samples. We will not know who these people are. We will manually screen a subset of audio samples to verify absence of voice traces and apply a voice scrambling algorithm to unscreened audio samples to ensure that any voice traces that may be present are rendered unintelligible.
2.	Which of the grounds in schedule 2 of the DPA will provide a legitimate basis for the processing?	This is an experimental research project. We are the "data controller". We are not seeking to directly capture personal data in a structured form or to gather data about people's identities. Data is screened to ensure voice is not present and where screening has not been applied the data is scrambled in the voice spectrum ensuring if any voice trace is present that it is rendered unintelligible. Please see Part IV Exemptions, section 33 of the DPA where conditions 1(a) and 1(b) hold for this project: https://www.legislation.gov.uk/ukpga/1998/29/section/33 Where it is stated: (1) In this section— "research purposes" includes statistical or historical purposes; "the relevant conditions", in relation to any processing of personal data, means the conditions— (a)that the data are not processed to support measures or decisions with respect to particular individuals, and (b)that the data are not processed in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject.

3.	If sensitive personal data is going to be processed, which of the grounds in schedule 3 (in addition to the schedule 2 grounds) will provide a legitimate basis for that processing? Note – Sensitive personal data is personal data consisting of information as to (a) the racial or ethnic origin of the data subject, (b) their political opinions, (c) their religious beliefs, (d) whether they are a member of a Trade Union, (e) their physical or mental health, (f) their sexual life, (g) the commission or alleged commission by them of any offence and (h) any proceedings for any offence committed or alleged to have been committed by them.	Not applicable. Sensitive personal data is not processed.
4.	Are there any special considerations relating to Article 8 of the Human Rights Act that will not be covered by the PIA? Note – This Article provides that everyone has the right to respect for his private and family life, his home and correspondence. It is subject to qualifications relating to national security, crime etc.	No.
5.	Will any of the personal data be processed under a duty of confidentiality? If yes, how is that confidentiality being maintained?	Yes. Confidentiality will be maintained by voice scrambling unscreened audio samples. A manual screening process will be performed to determine if audio samples may be shared without voice scrambling being applied. Only specific named individuals whom are members of staff on the project will carry out the screening process. The secure system on which this audio data is held in its raw form is only accessible to these named individuals, and the raw personal data is held encrypted within this system.
6.	How are individuals being made aware of how their personal data will be used?	Through additional notices placed at specific locations in the Meadows, and communications and workshop engagement with the community representatives and interest groups about this project.

7.	Does the project involve the use of existing personal data for new purposes?	No.
8.	What procedures will be in place for checking that the data collection procedures are adequate, relevant and not excessive in relation to the purpose for which the data will be processed?	Through PIA review and University roles of the PIA reviewers, e.g. Data Protection Officer, and security experts. We will also apply a range of testing and independent review of our voice scrambling algorithm. The voice scrambling algorithm we are using is based on previously validated privacy techniques in peer-reviewed research.
9.	How will the personal data be checked for accuracy?	Not applicable.
10.	Has the personal data been evaluated to determine whether its processing could cause damage or distress to data subjects?	Not applicable. We are not directly processing people's personal data.
11.	Will there be set retention periods in place in relation to the storage of the personal data?	Not applicable. We have no identified personal data recorded in the system.
12.	What technical and organisational security measures will be in place to prevent any unauthorised or unlawful processing of the personal data?	The Audio Capture Devices operate without retaining any data internally, and have a number of physical security features – please see the detailed description in Appendix C. The Data Collection system is internal to the University of Edinburgh network and is firewalled. Only critical services and related ports will be running on the system. It is only accessible to a small number of prescribed individuals, each of whom have separate authentication credentials for accessing the system (and therefore their individual accesses can be logged), which must conform to "strong password" requirements. A publishing process will be applied, including voice scrambling, before audio samples can be made public on the OrganiCity Node, which is a wholly separate and essentially isolated server from the Data Collection system, and within its own network DMZ.

13.	Will you be transferring personal data to a country outside of the European Economic Area? If so where, and what arrangements will be in place to ensure that there are adequate safeguards over the data?	No.

Summary of Security Measures

Appendices B, C and D have been omitted in the version of this document intended for public distribution since some of the details contained in these appendices could potentially represent a security risk.

However, we provide here a summary of the measures we have taken to ensure security of the overall system.

Discontinuous Recording

The six Audio Capture Devices (ACDs) operate as a collective continuous sound recording system. Each one will capture a 10-second sample of audio per each minute, interleaved with one another in sequence so that a full 60 seconds of sound per minute is captured across all the ACDs. This means that in each specific ACD location only a 10-second sound sample is recorded, then a gap of 50 seconds before the next 10-second sample is recorded. This localised discontinuous recording is advantageous from the perspective of privacy, meaning that only intermittent snippets of conversation could be captured by the system.

No Data on Device

The ACDs do not persist nor cache any sound recordings on their local filesystem (an SD card); they are held purely in volatile memory until transferred successfully (or transfer ultimately fails), at which point they are immediately deleted from the device. This means that if someone were to obtain access to the device they would not have access to any sound recordings that have been made by that device.

Data Encryption

The audio file itself is first encrypted on the ACD and then transferred from the ACD to a central server via encrypted communications. The data is therefore always in an encrypted form both in transit and at rest.

Voice scrambling

We have implemented an audio processing algorithm that transforms and scrambles the frequency band associated with human voice, rendering it unintelligible such that no spoken words remain discernible to the listener. This scrambling process will be applied automatically to any audio we choose to share with experimenters within the OrganiCity programme and for which we have not separately verified that voice traces are not present through manual screening. The speech-scrambling algorithm we are using is based on previous best practice that has been validated in this domain. We will also be randomising the parameter settings applied by the algorithm to each ten-second time sample of the audio.