



Project no. 732027

**VIRT-EU**

**Values and ethics in Innovation for Responsible Technology in EUrope**

Horizon 2020

ICT-35-2016

Enabling responsible ICT-related research and innovation

Start date: 1 January 2017 – Duration: 36 months

**D1.6**

## **Open Research Data Management Plan**

Due date: 30 April 2017

Revision submission date: 10 April 2019

Number of pages: 27

Lead beneficiary: ITU

Author(s): all partners

## Project Consortium

Beneficiary no.	Beneficiary name	Short name
<b>1 (Coordinator)</b>	IT University of Copenhagen	ITU
<b>2</b>	London School of Economics	LSE
<b>3</b>	Uppsala Universitet	UU
<b>4</b>	Politecnico Di Torino	POLITO
<b>5</b>	Copenhagen Institute of Interaction Design	CIID
<b>6</b>	Open Rights Group	ORG

## Dissemination Level

<b>PU</b>	Public	<b>X</b>
<b>CO</b>	Confidential, only for members of the consortium (including the Commission Services)	
<b>EU-RES</b>	Classified Information: RESTREINT UE (Commission Decision 2005/444/EC)	
<b>EU-CON</b>	Classified Information: CONFIDENTIEL UE (Commission Decision 2005/444/EC)	
<b>EU-SEC</b>	Classified Information: SECRET UE (Commission Decision 2005/444/EC)	

## Dissemination Type

<b>R</b>	Document, report	
<b>DEM</b>	Demonstrator, pilot, prototype	
<b>DEC</b>	Websites, patent filling, videos, etc.	
<b>O</b>	Other	
<b>ETHICS</b>	Ethics requirement	<b>X</b>

## Contents

Executive Summary .....	4
Ethical Considerations – data types and their vulnerabilities .....	6
Data Sharing and Analysis within VIRT-EU .....	7
GDPR Considerations.....	8
Table 1: Legal ground for different data collection activities .....	8
Quantitative Data Management .....	10
Data collection activities .....	11
Table 2. Quantitative Data Collection Activities overview .....	12
Data storage approaches.....	12
Data processing activities .....	13
Table 3. Expected VIRT-EU data processing information.....	13
Qualitative Data Management.....	13
Data collection activities and oral consent.....	14
GDPR considerations for oral consent.....	15
Written consent.....	16
Data storage approaches.....	16
Data Processing Activities .....	17
Design Research Data Management.....	17
Data collection activities and consent .....	17
Data storage approaches.....	18
Data processing activities .....	18
Open Research Data Requirements.....	18
ITU Research Ethics Committee Statement .....	20
Appendix 1: Research Information Sheet .....	21
Appendix 2: Consent Form.....	22
Appendix 3: Addition to consent form .....	23

## Executive Summary

This document is a revision of the Deliverable 1.6 (submitted on 28 April 2017). This deliverable has been reviewed to describe the procedures adopted to comply with the Regulation (EU) 2016/679, which is applicable since May 25. Universities carry out processing operations that are often (like in our case) strictly related to specific and temporary projects and have a limited legal staff and DPO staff that should cover a variety of research initiatives being also compliant with administrative law requirements, in terms of internal procedures and due administrative process. Against this background, the different obligations imposed by the GDPR (e.g. records of processing activities, exercise of data subject's rights, assessment operations) concern controllers (i.e. the university). We, as researchers, are persons "acting under the authority of the controller or of the processor" (Art. 29 GDPR) and, according to our role and the existing administrative and GDPR framework, we cannot directly assume any deliberation on behalf of our institutions, since we do not have any power either to legally represent them or to assume decisions on behalf of the data controller especially without any previous decisions adopted by the designated bodies. As such, the text presented in this deliverable is as complete as it can be prior to the relevant administrative units completing their review and finalizing the necessary agreements.

The VIRT-EU Open Research Data Management Plan (DMP) outlines how research data will be handled during and after the project duration. It describes what types of data will be collected, processed or generated with which methodologies and standards, whether and how these data will be shared or made open, and how these data will be curated and preserved. The DMP describes the data management life cycle for all data sets. The purpose of the DMP is to provide an analysis of the main elements of the data management policy that will be used in VIRT-EU with regard to all data sets that will be generated by the project. The DMP will be regularly revisited and updated if necessary to ensure consistency and compliance of project data practices. The DMP is established by the Coordinator with contributions from all partners involved in data collection, processing and management and agreed on by the participants in order to detail explicitly and formally all data management aspects. VIRT-EU is committed to ensuring compliance with the EU GDPR, the Data Protection Act and the National codes of conduct such as, for example, the Danish Code of Conduct for Research Integrity<sup>1</sup>.

For the purposes of this project, personal data of natural persons different from the people directly involved in the project (i.e. researchers, administrative team, and technical staff) are processed by LSE, ITU, UU, ORG and CIID only. Data are processed by each partner autonomously and they act as data controllers with regard to the processing activities described below. All partners that act as data controllers have adopted appropriate data protection policies concerning data processed for research purposes (see below) and data are processed for the purposes of the VIRT-EU project according to these policies.

---

<sup>1</sup> Danish Code of Conduct for Research Integrity (2014): <http://ufm.dk/publikationer/2014/the-danish-code-of-conduct-for-research-integrity>

<sup>2</sup> Sharing collected Twitter data – even if public – is not allowed by Twitter TOS. Nevertheless it is possible to share a list of TweetIDs that can be used by other researchers to re-collect the same dataset of tweets and verify the results. This ensures 4

The partners, in their role of controllers, adopt all the measures necessary to comply with the obligations of the Regulation (EU) 2016/679 (e.g. records of processing activities, data protection impact assessment, etc.) within the context of all of their research activities.

In order to facilitate appropriate data management, storage and processing activities and ensure the highest level of scientific output, the VIRT-EU partners are in the process of negotiating a joint-controller agreement facilitated by the DPOs at each institution.

The purpose of the empirical components of VIRT-EU research activities is:

First, to achieve objective 1 of the project:

**Empirically identify how local culture and network society influence the understanding and movement of particular social values among *technology developers*** and how local difference and networked commonalities can influence the development of ethical subjects from a virtue ethics perspective, using data mining, social network analysis (SNA), qualitative inquiry and design methods.

Second, empirical data collection and analysis are intended to lay the groundwork for accomplishing objectives 2 and 3 of VIRT-EU:

**Develop a Privacy, Ethical and Social Impact Assessment (PESIA) framework** shaped by state of the art legal research and empirical data, to enable developers and other societal stakeholders to reflect upon, evaluate and take into account not only the data protection, security and privacy aspects of new technologies but also the ethical and social concerns embedded within that challenge autonomy and freedom.

**Systematically consider and implement the PESIA framework by co-designing self-assessment tools with technology developers**, who may not be able to anticipate the future use of their projects and their clients and partners, grounded in existing developer practices and based on quantitative, case study and design research that identifies how ethics operate as process.

To achieve these three core objectives of the project, VIRT-EU requires the collection of three different types of data. First is quantitative data about developers, collected from online social media and digital repositories, the second is qualitative data collected through the various field sites and the third is data collected in the course of co-design workshops and stakeholder workshops as well as during the final Design Challenge. These data collection practices involve the collection of personal information, and require careful handling for storage and processing as well as decision-making with respect to preparing these data for participation in the EU Open Research Data Pilot program. The different nature of these data requires different data collection and handling protocols. All approved data collection and handling protocols will follow the general guidelines at both national and EU level (<http://ec.europa.eu/justice/data-protection>). No vulnerable population will be involved in our study and the data will represent no more than minimal risks to participants beyond what is normally entailed in the participation in developer communication and collaboration events and networks. The risks and discomforts entailed by participation in this study are minimal, as detailed in the core parts of

the DMP. Data collection consists of participant observation online and offline, interviews, collection of publicly available online data from online forums and social networks, anonymous surveys, co-design and stakeholder workshops, demonstrations and exhibits. Participants may withdraw from the study at any point. All requests for data deletion or redaction will be immediately honored to the best of our ability.

The IT University of Copenhagen Research Ethics committee has committed to evaluating and approving the VIRT-EU personal data management plan prior to submission of the official deliverable to the EU and should major changes become necessary. Further, the VIRT-EU consortium includes Associate Professor Rachel Douglas-Jones who is an expert in research ethics as this is her primary area of research. She will conduct research ethics oversight for all data collection activities, development of information and consent forms, requirements for the handling of consent forms and data and safe custody of research outcomes. Finally, project members Javier Ruiz (ORG) and Associate Professor Alessandro Mantelero (POLITO) are experts in the relevant European legislation as it applies to data protection law. They will help us ensure compliance with the Regulation (EU) 2016/679.

The VIRT-EU Open Research Data Management Plan provides information about the following points in detail for each of the three different types of data collection (quantitative, qualitative and design):

- Data collection activities (including selection of legal ground and approaches to managing consent)
- Data storage approaches
- Data processing activities

The DMP gives a first overview on the diversity, scale and amount of data, which will be handled during the VIRT-EU project. Further, this document includes a brief discussion of research ethics as it pertains to VIRT-EU data management activities. Finally, we include a section that details requirements to be fulfilled for data sets to be determined as sharable through the open research data pilot program.

The DMP is not a fixed document, but evolves during the lifespan of the project. This is the second iteration of the DMP developed to address project reviewer commentary and suggestions. As institutional-level agreements can often take a significant amount of time, especially given the very new and evolving set of institutional practices with respect to GDPR compliance, this DMP still does not include the finalized joint controller agreement. However, this DMP represents the particulars that will be included into said agreement.

## **Ethical Considerations – data types and their vulnerabilities**

Data management is an ethical issue. The types of data collected here carry different vulnerabilities, as the sections below indicate: some are more easily linked to personally identifying information than others. By dividing the DMP according to data type, we acknowledge and address this ethical dimension of data management. However, there are also different ethical

issues at the different stages of research, from collection to processing and storage. The DMP addresses key ethical issues common to the project as a whole but we have tailored these issues to be appropriate and relevant to each type of data collected. Working from the recommendations combined from the Danish Code of Conduct for Research Integrity and the research ethics guidelines of our multiple professional organisations, we have integrated ethical considerations into our plan for how data in this project will be collected, processed and stored. These include consideration of when and where consent processes take place, how data is anonymized and what happens to re-identification keys, the technicalities of storage of data over time, how data will be shared within and beyond the project, and takes into consideration the representation of participants at the point of publication.

Participants in VIRT-EU data collection and processing activities are not definitionally vulnerable: they are not dependent on VIRT-EU researchers, they do not lack decision-making capacity nor do they lack an ability to protect their own interests; they are not historically vulnerable on grounds of inequality or exclusion, nor vulnerable to coercion or undue influence. Through the consent procedures, recommended by professional organizations and followed within VIRT-EU, they are provided with the opportunity to autonomously decide, without coercion and with the option to withdraw, that they wish to participate in the project. However, researchers in the project will be attentive during data collection to the potential for vulnerabilities to arise due to participants engagement with the research project itself, given the networked professional nature of the field into which we intervene, and the potential power asymmetries that may come to exist between researchers and participants. These risks will be minimized by VIRT-EU's anonymization practices and reflexive publication strategies, as mentioned below.

## **Data Sharing and Analysis within VIRT-EU**

VIRT-EU is a research project with inter-connected work packages and researchers working within and across institutions. This raises a challenge for generating an efficacious work environment for researchers, without unduly burdensome or restrictive blocks on mutual work. VIRT-EU addresses this challenge through a careful use of university, commercial and open source infrastructures available to the project members. Our quantitative data will be securely stored on university-managed servers in Sweden, allowing all relevant consortium members controlled and limited access to the data via a secure web interface detailed below. Unprocessed social media data listed below are shared only between ITU and UU, by granting to the project members operating at ITU controlled access to the secure servers from UU.

Unprocessed observational and interview data are not shared among the research partners. Our qualitative and design data will be stored in secure folders on the servers controlled by the relevant partner who is in charge of collecting data. Collaborative coding and data analysis tools will only be used on de-identified field notes and transcripts. The de-identification of these data subjects is the responsibility of the relevant data controller with the collaborating partners setting requirements for the type of data they need for their scientific activities. Access to physical artifacts produced through co-design will be managed via digital images stored in shared folders on an encrypted cloud server,

ensuring appropriate levels of pseudonymization in place. All participant lists will be stored by the relevant data controllers and access will be made to collaborating partners in a controlled fashion. No personal data are included in the deliverables or in any other documents made public by the VIRT-EU project.

These interactions will be driven by the requirements of the project and governed by the terms set out in the joint controller agreement. Consortium partners are separately responsible for observing the principles for processing of personal data, to the extent that the regulations apply to the responsibilities of the party concerned, pursuant to the joint-controller agreement currently under negotiation.

## GDPR Considerations

### Legal ground for data collection

The VIRT-EU project requires heterogeneous data collection mechanisms and thus different legal ground for ensuring ethical conduct. See Table 1 for details.

Form of data collection	Legal ground	Other comments
Qualitative ethnographic participant observation and informal interaction	Oral consent	Although the ethnographic research is performed by university researchers, the requirements of research ethics mandate ensuring consent for participation. Following the best practices in research ethics as set out by professional societies such as American Anthropological Association (US) and the Association of Social Anthropology (UK), oral consent is recorded in researcher field notes
Qualitative formal interviews, design and stakeholder workshops, Design challenge participation	Written consent	Consent forms are securely stored by the responsible partner conducting data collection
Automated Twitter and MeetUp data collection based on hashtags and observation of Twitter and MeetUp activity on the topic of IoT	Public interest	These data are collected and processed only by public universities and not SME and non-profit partners. However, significant effort will be made to ensure data minimization and providing facilities for data subjects to request data deletion

**Table 1: Legal ground for different data collection activities**

The collection of qualitative data by our ethnographers entails both written and oral consent (for a more specific discussion of oral consent see section on Qualitative Research below). Oral consent is registered by the researcher in their field notes. This makes it possible for the researcher in question to comply with any later data deletion requests from the data subject.

The collection of qualitative data in all other research engagements such as: formal interviews, co-design and stakeholder workshops, demonstrations and exhibits as well as during the design challenge will be conducted with explicit written consent.



Online data collected by the project does not rely on consent as a legal basis, but on the performance of a task in the public interest: scientific research in publicly funded institutions. We will however, comply with the information requirements, whenever possible.

### **Information transparency**

The VIRT-EU project makes study goals and data management information prominently available on the project website. This ensures that all participants in the project research activities as well as those concerned or interested in our efforts can gain detailed information at any time.

Data subjects whose information is directly collected by consortium members are provided with information in accordance with Article 13 of GDPR, using information that is similar to the online resource but using printed out hard copies.

### **Purpose specification**

We explain to our research subjects in the information sheet that the main purpose of the processing is scientific research but we will also need to communicate with them and comply with legal requirements.

In this plan and our information sheet we provide as much detail as possible about purposes within the limitations inherent in research. As explained in recital 33 of GDPR:

*“It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. <sup>2</sup>Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. <sup>3</sup>Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose.”*

### **Data minimization**

As explained throughout this document we apply data minimization principles at every stage of the research.

### **Research exemption in GDPR**

The implementation of the research exemption across the EU is being developed at the time of writing, and we are in the process of consulting with our respective institutions how to best implement national legislation. As a matter of principle we have agreed to comply with any requests for access, rectification or erasure, independently of the legal rights of the subject in the country.

### **Fulfilling obligations to the Data Subjects**

Consortium members are responsible for ensuring the rights of the Data Subject through careful observation of the General Data Protection Regulation rules. Moreover, as data controllers, all consortium members are responsible for assisting each other, to the extent that this is deemed relevant and necessary, in order for all parties to be able to observe the obligations required in relation to the Data Subjects.

## Quantitative Data Management

The digital data sources are a selection of online social media from which we will extract information about: which event individuals and groups under analysis attend – MeetUp data – and how they communicate about IoT related topics – Twitter. Thus VIRT-EU data collection will mainly take place on social media platforms that are used within a public and professional context. Platforms such as MeetUP and Twitter, are usually perceived as public spaces where public visibility is often not only expected but actively sought. Despite the general perception of these digital platforms as public spaces, VIRT-EU data collection will actively pursue data minimization by not collecting additional information beside publicly available usernames, lists of publicly displayed connections and publicly shared content. Usernames, when belonging to companies, corporate, civil society or media actors (e.g. @ericsson or @wired), will be retained as is. Usernames, when belonging to identifiable individuals, will not be used in any public communication (e.g. social media posts, articles, public presentations) without explicit consent. Existing tools to collect online data (including the one we have re-engineered and integrated with our system) make it difficult to strictly enforce the principles contained in the GDPR. In particular, the tools act as black boxes and do not provide the possibility of customizing the data collection process without modifying their code, but only provide the collected data as output. As a consequence, pseudo-anonymization of the data is performed as soon as technically possible: after the data collection process is terminated. This is a concern despite the fact that the collected data is from public sources and is password-protected during storage at all times (following the results of a Privacy Impact assessment analysis we have not identified any significant risks in temporarily storing public data inside a password-protected database).

To address this concern, we devoted some project resources to design a prototype tool for data collection from social media, where the data can be pseudo-anonymised at collection time and a setting is added so that a notice can be automatically sent to the user accounts present in the data. Unfortunately our effort to automatically inform the users of their data being collected was stopped by Twitter itself, because such form of the automatic notification system was interpreted as an infringement of its terms of services.

Even though we target publicly available data (e.g., MeetUp data, tweets) for our data collection activities, in case of releasing these networks, or of aggregated multilayer networks, as part of the Open Research Data Pilot we will remove all references to the identity of the individuals replacing the user identifiers with randomly generated placeholders. In cases where data collection will occur on forums and discussion spaces where login requirements can create an expectation of privacy, the research team will announce our presence and intentions to the community on the forum, inform them about the study and follow the requirements of the community in our data collection practices. Due to the nature of the project, focusing mainly on developers and IoT companies, monitoring and tracking activities will be based on specific keywords or groups that minimize the risk of accidentally collect data from vulnerable groups or minors. However, if VIRT-EU researchers come to know that we have accidentally collected social media data that is potentially able to identify the author as a member of a vulnerable group or a minor, the data will be immediately removed from the project's data archive. As it is often impossible to assess vulnerable group membership quantitatively, researchers will perform random qualitative checks on our data periodically to guard against such a possibility.

For our algorithmic studies we do not need to retain information about the source of each network, making identity identification practically impossible. During the research phases all project data will be securely stored inside our single handling sites at Uppsala University (in Sweden), where only the portion of the data required to perform the analysis of the network will be replicated in encrypted form. Even if the project will collect publicly available data emergent information might be obtained thanks to the multiplex nature of the data we are collecting. Therefore we will act according to a very clear policy for accessing the full dataset: Access to the server will only be granted to researchers officially involved in the project, or associated researchers (e.g., master thesis students) after signing a non-disclosure agreement. Anonymised versions of specific portions of data will be made available for validation and re-use through the Open Access to Research Data Pilot which will be negotiated and decided by consortium members.

## Data collection activities

The quantitative digital data collection approved when this document is being finalized lists four specific data sources: Twitter, MeetUp, ad-hoc websites. Given the existing diversity in available data and in technological solutions it is required a platform-specific description of the data collection procedures. An overview is provided in Table 1.

### *Twitter:*

Using both REST and STREAM API Twitter will be queried in order to collect: a) tweets produced using a pre-defined topical hashtag (e.g. #iot, #iotdays) b) list of followers/following connections between publicly available key twitter users identified during the ethnographic phase.

### *MeetUp:*

Using the available public API we will collect attendance data of specific iot-related public events. Attendance data will then be used to construct an attendance network where people attending the same events will appear to be connected.

Throughout the project we are adopting a data minimization strategy by collecting only the data strictly necessary to achieve the goals of the project:

Platform	Data collection	Data type	Anonymized	Pseudo-anonymization	Explicit user authorization	Information	Shared
Twitter	API	Tweets	NO	YES	NO	NO	NO <sup>2</sup>
Twitter	API	Following networks	NO	YES	NO	NO	NO
MeetUP	API	Attendance	IF released	YES	NO	NO	YES

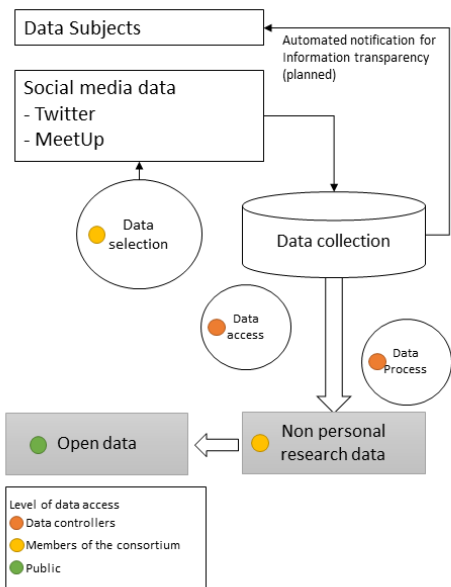
<sup>2</sup> Sharing collected Twitter data – even if public – is not allowed by Twitter TOS. Nevertheless it is possible to share a list of TweetIDs that can be used by other researchers to re-collect the same dataset of tweets and verify the results. This ensures replicability, which is one of the cornerstones of scientific practice.

**Table 2. Quantitative Data Collection Activities overview**

While previous authorization from users will not be obtained, being a task in the public interest, and information has proven to be impossible, we will however put in place a secure web page where the users will be able to check what information has been collected (e.g. what messages and what set of connections) and he/she will have the opportunity to withdraw his/her consent and remove the data from the dataset.

**Data storage approaches**

Data will be stored in password-protected virtual machines not directly accessible from outside a VPN at Uppsala University. Both raw and curated relational data will be stored. In addition, a pool of monitoring systems will be deployed. The systems will be controlled through a web interface that will also be used to inspect the database. Access to the web interface will be granted only to the data controller for Quantitative Data: Uppsala University and IT University of Copenhagen. Backups will happen at virtual-machine level.



*Figure 1: Overview of the data flow and data access for the quantitative data.*

A relational database will be used to store the curated information extracted from the raw data. Every time an actor is identified as relevant, s/he will be inserted in the database and annotated with properties and connections to other actors. The role of the database is to keep a high data quality, by avoiding duplicate information and simplifying the identification of missing information. A conceptual schema of the database will be prepared. The data will be stored for up to 12 months after the end of the project to ensure completion of the final analyses and publication and then deleted.

*Web interface*

The web interface, built with Django over a relational database management system, will be used mainly to allow all researchers involved in the project to provide information about the data to be extracted in a controlled way and to give information about the data already stored in the system. This site will be password-protected. That is, all researchers in

the consortium are able to contribute particular relevant hashtags to focus Twitter data collection. For example, several researchers have recently initiated a Twitter data collection activity on the hashtag #FUD to address the emergent relevant discussions around a particular concept that the qualitative researchers have identified as relevant through a recent analysis of manifestos and other public documents produced by the communities of practice under study. Further, qualitative researchers contribute Twitter handles for particular well known public figures in the IoT developer community to help focus data collection. The data collected in this fashion allows the quantitative analysis to produce an overview of IoT-related activity in Europe by analyzing follower, retweet and public conversation network patterns.

### *Data retrieval process specification*

The units responsible for the data collection are UU and ITU. All the parties will receive a process specification describing the steps to submit new requests for data collection to the units at Uppsala University and ITU Copenhagen. All interactions from partners other than Uppsala University will happen through the Web interface. While adding to the monitoring system a new user or a new hashtag is open to all the member of the consortium, the actual access to the collected data is only possible to the UU and ITU, while processed and pseudonymized data will be made available to the other consortium members.

### **Data processing activities**

Different types of digital data will be processed in different ways. Tweets obtained from hashtag driven collection will be processed in order to obtain communication networks (identifying patterns of interaction between Twitter users) and topic analysis in order to identify the main topics of the online conversations. MeetUp co-attendance data will be transformed into a bipartite network and then projected onto a network showing the co-attendance as a relation between users. Network data from Twitter, MeetUp, will be used to build a multilayer network model as described in the research proposal. All the necessary details of data processing are provided in Table 3. The strike-through formatting indicates the alterations to originally expected data collection activities that the project partners have made in the course of the project so far.

<b>Data</b>	<b>Original format</b>	<b>Processed format</b>	<b>Analysis</b>
Tweets	Text (.csv)	Network (.graphml)	Communication Networks
Tweets	Text (.csv)	Annotated text (.csv)	Topic identification
Twitter following	Network (.csv)	Network (.graphml)	Cluster analysis
MeetUp	Attendance (.csv)	Network (.graphml)	Cluster analysis
Multilayer network	Network (.csv)	Network (.mpx)	Cluster & centrality analysis

**Table 3. Expected VIRT-EU data processing information**

### **Qualitative Data Management**

VIRT-EU will analyze and map the ethical practices of European hardware and software entrepreneurs, maker and hacker spaces, and community innovators. Leveraging state of the art collaborative SSH and ICT methodological innovations, data collection tools will include data mining, quantitative and qualitative social science and design research practices. Specifically, data will be collected through ethnographic observations, surveys, questionnaires, structured and/or semi-structured interviews, co-design and stakeholder workshops, and a Design Challenge activity.

The VIRT-EU consortium is comprised of experienced anthropologists and ethnographers who will follow the ethical guidelines of the Anthropological American Association (US) and the Association of Social Anthropology (UK) – considered the gold standard for ethical research practice in the field (where these guidelines diverge, the more stringent will be followed.) Consortium members involved

in this form of data collection are trained to address ethical issues reflectively throughout data collection and in the course of analysis.

### **Data collection activities and oral consent**

For our qualitative and ethnographic studies of online and offline communities of practice, we will collect screenshots, audio/visual recordings, field notes and observations. We recognize that these forms of data collection may identify individuals. Key pillars of ethics in ethnographic and qualitative studies are reflexivity and accountability. Researchers at the various research sites are responsible for identifying and adhering to local community norms and practices about data and information. Any recording of interviews, collection of photographic evidence and the like, will be conducted with the explicit consent of participants (see below). Participants will be informed that they are welcome to request removal and destruction of any collected data at any time, and that the point of contact is the principal investigator at the university who employs the researcher who they meet. Information about the research project will be provided in written form and will include the contact information for the principal investigator. We also provide details of the Data Protection Officer where applicable in accordance with Article 13 of GDPR.

By agreeing to take part in this study participants will be contributing to research related to identifying how local culture and network society influence the understanding and movement of particular social values among *technology developers*, and how local difference and networked commonalities can influence the development of ethical subjects from a virtue ethics perspective. There is little or no risk, inconvenience or discomfort to those who participate in this study, assured through the extensive use of pseudonymization and anonymization in the research process and reporting. Participants will be able to request that any data relating to their participation is destroyed up until project completion or until the relevant data have been published in project reports and other output. A practice of anonymization or pseudonymization is practiced by default unless a research participant specifically requests to be identified in research findings. The practice of oral consent enables the researchers to ensure that participants are not directly identified in the data should they prefer not to be identified. However, qualitative data is situational and complex and, as such, removal of all identifying data is impossible. In order to ensure confidentiality throughout, qualitative researchers will store digital data on secure encrypted hardware and physical materials under lock and key (see below for details).

The research team will anonymise all participants in the research, acknowledging that full anonymisation is not always possible within small communities of practice, and respect any requests for information to be left off the record. The qualitative researchers will continually reflect on their own role in constructing the narrative of their research site, and acknowledge any power differentials that might result. Results of qualitative fieldwork will be communicated back to participants through co-design exercises and stakeholder meetings throughout the project in WP3, WP5 and during the final Design Challenge event in WP6.

Throughout all qualitative fieldwork will be taking informed consent as a basic ethical tenet of ethnographic and design research practice. However, the use of the written consent form in the process of ethnographic research is counterproductive. Thus we will primarily rely on iterative oral consent in interviewing and participant observation, with provision of written information including contact

details of researchers who can be contacted with questions or in the event that the participant withdraws consent to participate. Ethnography is a long-term engagement with others, and the up-front nature of a signed document means that what is being consented to will change over time. As the Association of Social Anthropology puts it, “Given the open-ended and often long-term nature of fieldwork, ethical decision-making has to be undertaken repeatedly throughout the research and in response to specific circumstances” (ASA 2011:2). Further, the use of the written consent form requires participants to identify themselves in writing, while the use of oral consent allows participants to remain anonymous if they so choose.

Oral consent entails making sure that interlocutors understand they are taking part in a research project and are free to refuse participation; understand that they can opt out of the study at any time; know if and when we are collecting information that could identify them, and that we will respect their wishes to have their identity acknowledged or kept confidential; they understand what kinds of information we are collecting and what we are carrying from our interactions with them. The risks and benefits of participating in the research will be discussed, along with options to minimize risk (e.g. anonymity). Oral consent/Verbal consent is a recognized form of alternate consent by the two main bodies of anthropology, the American Anthropological Association (USA) and the Association of Social Anthropologists (UK), most often used in ethnographic practice in settings where written consent would be disruptive, unwelcome, or inappropriate. In engaging with oral consent, researchers will pay particular attention to explicit opt out requests that will be noted down in field-notes and then transferred into the participant consent and opt-out information document maintained by the project.

As a backup precaution all data will be reviewed immediately after collection and cross-checked with participant consent and opt-out information collected in prior engagements. If any data has been inadvertently collected relating to a non-consenting individual, this material will be immediately deleted or destroyed. The documentation accompanying oral consent is typically a Research Information Sheet as mentioned above (see Appendix 1). In our case, a jargon free study information sheet with further information will be available online, which will cover project goals, data treatment and participant rights, as well as the contact details of the researchers involved including the local principal investigator, as well as the ethical oversight committee at ITU. This information will be presented to all participants via our website with the link available on our business cards.

### **GDPR considerations for oral consent**

According to the Guidelines on Consent under Regulation 2016/679: *“The GDPR does not prescribe the form or shape in which information must be provided in order to fulfill the requirement of informed consent. This means valid information may be presented in various ways, such as written or oral statements, or audio or video messages”*. In fact, Recital 32 sets out specific guidance on this: *“Consent should be given by a clear affirmative act ... such as by a written statement, including by electronic means or an oral statement.”* As such oral statements of consent are in fact compliant with GDPR and, given the field-note style of ethnographic research, these are recorded in field-notes. Moreover, Recital 33 with respect to research states: *“It is often not possible to fully identify the purposes of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognized*



*ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose.”* These statements recognize the contingent nature of research and the requirements of flexibility inherent in conducting certain types of scientific work.

The ethical research practice guidelines provided by the professional associations for ethnographers and anthropologists recognize oral consent as a vitally necessary approach in such work. Ethnographers and anthropologists are specifically trained in ethical conduct through reflexive praxis in such situations.

## **Written consent**

For more structured data collection such as formal interviews, surveys or questionnaires, all participants will be presented with a Research Information Sheet prior to entering the survey (<https://virteuproject.eu/researchinfo>). The purpose, procedures, and risks of participation will be conveyed to participants in keeping with the processes of informed consent, and participation will proceed only if they agree without coercion. We will present potential participants with an overview of the research and accept their active selection to continue with the survey as an indication of consent. No personally identifying information will be collected throughout the survey unless participants actively choose to identify themselves in their answers or to contact us directly. Such identifying information will be carefully removed from the data prior to analysis and retained for reporting only if participants explicitly wish to be identified, as is sometimes the case with this form of data collection.

## **Data storage approaches**

Audio recordings, transcriptions, online observation data and field notes will be manually scrubbed of all identifying information and assigned a code as soon as data collection is complete and the lead researchers return from the field. The code key linking the data with identifying information will be kept separately in a secured folder in a shared folder repository maintained by the project. Access to the repository is made available only to project personnel involved in the empirical data collection and processing. Once coded, all data will be retained indefinitely for future study. The key will be destroyed upon completion of the project. Oftentimes qualitative studies can yield important insights much later, after similarly themed studies have been conducted at different sites, and some national research bodies require retention of research findings generated in publicly funded projects. However, the insight potentials of such qualitative data do not rely on the identification key. Thus, upon completion of the project the de-identified data will be stored securely as encrypted files, maintained by the coordinating institution.

Data collected through surveys or structured questionnaires will be carefully scrubbed of identifying information if any should be entered. Such information will be kept for reference only in accordance with explicit wishes of the participants.

In case during the project it is considered relevant to acquire more information about some of the individuals involved in the public online discussions, we will first contact these individuals, ask for an explicit consent in retaining their data, ensure that they are informed about our usage of the data, retention and disposal periods, that they can at any time remove or update the information in our



servers, and we will follow the general guidelines at both national and EU level (<http://ec.europa.eu/justice/data-protection>).

With explicit consent of participants we may make anonymised qualitative data available for validation and re-use through the Open Access to Research Data Pilot, which will be negotiated and decided by consortium members.

## **Data Processing Activities**

Given the diversity of data sources and data activities, personally identifying information will at times be necessary to connect different data streams and to interpret these. The use of such information will be done with utmost care for data analysis only and never for reporting or publication. For example, personally identifying information may be necessary to connect quantitative data collected via Twitter with ethnographic observations in order to aid interpretation of both types of data. Such analyses, however, will be kept securely in shared project folders and all partners will agree to ensure that no identifying information is ever revealed in published form.

Qualitative data processing includes a range of activities such as coding, semantic analysis, memoing, situational maps and thematic analysis among others. VIRT-EU researchers will ensure that all analysis output is scrubbed of identifying information when such output is shared with parties outside of the project members such as research participants or members of the advisory board. Any publications that include quotes or pictures will be anonymized. All anonymous survey data will be analyzed and published at the population and not at the individual level.

## **Design Research Data Management**

The design research component of VIRT-EU will pay special attention to research ethics practices ensuring that all participants in co-design workshops and stakeholder meetings are provided with study information sheets and are given the opportunity to provide written consent for participation. Participants will retain the right to revoke consent and to stop participating at any time. Written consent will also be collected for official structured or semi-structured interviews conducted by the consortium members during the evaluation phase at the conclusion of the Design Challenge as personal information will be collected during those interviews to assist participation in the Design Challenge itself. Design research in practice holds to the same research integrity tenets that govern all of qualitative research activities of VIRT-EU.

## **Data collection activities and consent**

Throughout all qualitative fieldwork researchers will be taking informed consent as a basic ethical tenet of research practice. The design research component of VIRT-EU will primarily focus on conducting co-design and stakeholder workshops throughout years 2 and 3 of the project. Design research is not constrained to workshops alone, thus a research Information Sheet (see Appendix 1) will be provided to participants for those instances when data is gathered via structured surveys and/or questionnaires, and structured/semi-structured interviews where no personal information is collected or where personal identifiers are removed from the data. The Research Information Sheet will establish clearly that consent is implied for participating in the survey or questionnaire, and

participating in the interviews and co-design workshops. If, however, a research engagement extends beyond asking questions to include activities such as discussing incidents and activities that could potentially identify the individual, then written consent is required.

As co-design and stakeholder workshops entail prolonged and intensive face-to-face engagement with participants, however, additional care will be taken to obtain consent for participation and to respect the wishes of participants with regards to the management of research data and the physical artifacts produced. To ensure the integrity of research practice and research content gathered during co-design or stakeholder workshops, written consent forms (see Appendix 2) will be used. These consent forms will be stored in locked cabinets at the Copenhagen Institute of Interaction Design separately from the data to ensure that the data remain confidential. Upon completion of the project, consent forms will be stored by the coordinator and retained for five years for audit purposes.

### **Data storage approaches**

All data collected in the course of co-design workshops and stakeholder workshops will be carefully scrubbed of all identifying information. Participant lists will be kept separately for the duration of the project to ensure a longitudinal aspect to the co-design process that requires repeated engagements with the same participants in co-design workshops and during and after the Design Challenge held at the end of the project to enable an evaluation of effectiveness of project outcomes. Participant lists will be destroyed at the completion of the project. All physical artifacts produced in the course of the workshops will be held under lock and key in the offices of the Copenhagen Institute of Interaction Design for the duration of the project and destroyed upon completion of the project unless permission is obtained directly from the participants to display these artifacts.

Anonymized and processed data will be kept separately and retained after project completion. All participants will be notified of the possibility of their data being kept after project completion prior to initial participation and consent will be obtained via consent form.

### **Data processing activities**

We will use the processes described above for all qualitative data for all design-specific data collected during co-design and stakeholder workshops.

## **Open Research Data Requirements**

Selected data will be shared publicly during or after the lifetime of the project. Which data will be shared will depend on the scientific relevance of the collected data, nevertheless personal data and sensitive data will never be released or shared. All long term data collected or generated will be deposited in a repository. If required, the entire information content of the VIRT-EU project can be stored on disk for archiving. This functionality can also be used to transfer VIRT-EU content to another system. The final repository has not been chosen yet but the consortium is currently testing the opportunities offered by Dataverse. Dataverse satisfy most of the necessary requirements:

- Offers an open source option that will be hosted at ITU.
- Data will be physically located in EU (Denmark)

- Supports a large variety of data types, from network data to qualitative observations.
- Supports differential access to data allowing for public or private data access.
- Supports several types of license for shared data, including CC0 (CC Zero Waiver) or ODBL (Open Database License).

## ITU Research Ethics Committee Statement

The ITU Research Ethics Committee has reviewed the original VIRT-EU Open Research Data Management Plan and confirms its compliance with the Danish Code of Conduct on Research Integrity. Given the 2-week deadline to complete these revisions there was not enough time to obtain this review approval. We will include the review approval in our next amendment to this document at the time when institutional permissions from all of the relevant DPOs will be received (currently in progress).

---

Espen Arseth, Professor  
Chair, ITU Research Ethics Committee

## **Appendix 1: Research Information Sheet**

VIRT-EU recognizes that the particulars of the information contained in this research information sheet – the identity of the data controller, the specific purposes of data processing and the exercise of the rights of access to data – will differ on the basis of the specific circumstances in which data are collected. The information below will be augmented as applicable based on the research context and prior experience. The information presented in this version of the Research Information Sheet is an exemplar of information to be provided to all research participants. This research information sheet will be paired with the “GDPR information form” (see Appendix 3) as well as the consent form whenever the consent form is necessary, such as when organizing co-design or stakeholder workshops.

## RESEARCH INFORMATION SHEET

### Project: VIRT-EU: Values and Ethics in Innovation for Responsible Technology in Europe

**Principal Investigator's names:**

1. Irina Shklovski – ITU Copenhagen
2. Javier Ruiz – Open Rights Group
3. Matteo Magnani – Uppsala University

4. Alison Powell – London School of Economics
5. Alessandro Mantelero – Politecnico Di Torino
6. Annelie Berner – Copenhagen Institute of Interaction Design

You are invited to take part in a research study about values and ethics in innovation practices. If you would like more information about the project please see <http://virteuproject.eu>

**What does the research involve?**

The aim of this research is to empirically identify how local culture and network society influence the understanding and movement of particular social values among *technology developers*, and how local difference and networked commonalities can influence the development of ethical subjects from an ethical perspective. Empirical data collection and analysis are intended to lay the groundwork for accomplishing the following objectives:

**Develop a Privacy, Ethical and Social Impact Assessment (PESIA) framework** shaped by state of the art legal research and empirical data, to enable developers and other societal stakeholders to reflect upon, evaluate, and take into account not only the data protection, security and privacy aspects of new technologies but also the ethical and social concerns embedded within that challenge autonomy and freedom.

**Systematically co-design with technology developers self-assessment tools and tools supporting ethical discussions in IoT design.** As IoT developers may not be able to anticipate the future use of their projects and their clients and partners, we seek to support decision-making with tools grounded in existing developer practices and based on quantitative, case study and design research that identifies how ethics operate as process.

**What does participation entail?**

The researcher/s will conduct formal and informal group or individual interviews with participants who agree to be interviewed. The formal interviews will be audio taped and transcribed. If at any point you are uncomfortable with the audiotaping, please let us know and we will turn off all recording devices.

You may also be invited to participate in any of the following: a short online survey, a series of co-design workshops, an online discussion via an email list, and/or a Design Challenge focused on developing new approaches to ethical innovation practices. The workshops will be conducted at a time convenient for those involved. Participation in the project and in any of these research activities is optional and voluntary.

There is no cost to you for participating. You may refuse to participate or discontinue your involvement at any time without penalty. You can choose to skip a question at any time. You can request that your data are removed from the project data corpus at any time.

**Source of funding**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727040.

**Consenting to participate in the project and withdrawing from the research**

You can consent to participating in the project and in any of the activities outlined above by completing the consent form attached to this Explanatory Statement and returning it to the researcher conducting the particular study and acting as the representative of the data controller. You can also withdraw from the research at anytime. Up until the project results are reported in publications you can request that any data relating to your participation be completely destroyed. Upon publication of results, the data the project holds will still be destroyed but the publications can not be altered.

- (a) If you participate in the observational and informal interview activities, we may collect anonymised data about your experiences. For instance, the researchers may collect data about your opinions on responsible innovation, data protection, issues of privacy, and security. We will rely on oral consent for this kind of data collection that will occur as part of your interaction with our researchers, trained in ethnographic research methods. Should you express any reservations in the course of the interaction with the researchers, they will ensure that your data are treated in accordance with your wishes. We will process any information lawfully and securely in accordance with data protection legislation.
- (b) Survey, workshops, Design Challenge, structured and semi-structured interviews: if you wish to participate in the survey, workshops, Design Challenge, and/or structured and semi-structured interview please sign and return the consent form to the researchers.

**Possible benefits and risks to participants**

By participating in this study you will be contributing to research related to understanding how European IoT innovators and developers make ethically consequential decisions – about code, hardware and data – for new connective devices, and the assumptions about human behaviour, privacy and freedom that underpin European cultures of IoT innovation. There is little or no risk, inconvenience or discomfort to you in participating in the study.

**Confidentiality and storage of data**

All research data will be de-identified and stored securely and confidentially. This means that your name and identifying details will be removed prior to reporting and publication and will not be shared with others. However, we cannot guarantee complete anonymity since the content of what you say, despite our best intentions, may be identifiable to people who work/communicate closely with you. Only the researchers involved in the data collection will have direct access to the data. The data will not be used for purposes other than the purposes of this research project.

**Results**

Results and project outputs will be published on the project website, reports, conference proceedings, and in journal articles. You are welcome to email the researchers to request findings, which will be emailed to you in summary form.

**Additional information**

If you have any comments, concerns, or questions regarding the conduct of this research in general or questions about the VIRT-EU project coordination activities, please contact the project administrator – Jonas Langeland Pedersen by phone +45 7218 5227; by email [jolp@itu.dk](mailto:jolp@itu.dk); or by mail: IT University of Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen S, Denmark.

## **Appendix 2: Consent Form**

While data processing will be governed by a project-wide joint controller agreement, data collection will typically be performed by a single partner or at most two partners in collaboration. As such, in order to avoid confusion and to ensure that data subjects know exactly who is responsible for data collection activities each partner has edited the consent form in order to add their relevant information for identifying data protection officer/information commissioner for the relevant institutions. The form presented in this appendix is in its generic form.





<<DATA CONTROLLER LOGO>>

virtueproject.eu

## CONSENT FORM FOR INTERVIEWS, SURVEY, AND WORKSHOP RESEARCH PARTICIPANTS

**Project: VIRT-EU: Values and Ethics in Innovation for Responsible Technology in Europe**

**Data Controller:**

<<Controller physical address and contact  
information>>

**Principal Investigator(s):**

<<Partner PI>>

**Researchers**

<<Project research team involved in the  
research>>

I have been invited to participate in the VIRT-EU research project specified above. I have read and understood the attached Explanatory Statement and I hereby provide consent to take part in this project:

1. I confirm that I received information about the project. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. ☐
3. I understand that any information given by me may be used in future reports, articles or presentations by the research team and their collaborators in the VIRT-EU project. ☐
4. Select one of the following:
  - a. I **do not** want my name used in any publications of reports in this project ☐
  - b. Please use an alternative name for me (write name here) \_\_\_\_\_ ☐
  - c. I would like my name used and understand that what I have said or written as part of this study will be used in reports, publications and other research outputs so that anything I have contributed to the VIRT-EU project can be recognized as my contribution. ☐
5. I agree to the interview / working group / consultation being photographed, audio and/or video recorded ☐
6. I agree to the use of quotes, pictures, video and audio in publications and presentations (In the case of withdrawal of consent all records of you will be deleted, or in case where you appear with other participants who have given their consent, you will be anonymized to make sure your presence in the dataset is unrecognizable) ☐

Name of Participant: \_\_\_\_\_

Signature \_\_\_\_\_ Date \_\_\_\_\_

### **Appendix 3: Addition to consent form**

The addition to the consent form will also be edited to ensure that the relevant data controller information is easily available to the data subjects. This form will accompany both the study information sheet and the consent form in all our data collection activities.



<<DATA CONTROLLER LOGO>>

**virt:eu**

### **Data Handling Information Notice**

We primarily process your personal information with your consent for scientific research, as explained in the research information sheet in detail. You may have signed a consent form at one of our workshops or provided oral consent to one of our researchers, who has kept a written record of this. As part of the research we may have to contact you from time to time but we will not be making any marketing offers. If you have specifically consented we may use your information in dissemination materials and presentations. You can revoke your consent at any time.

Your information will also be shared with other VIRT-EU consortium members under strictly controlled circumstances as detailed in the project's Data Management Plan, which can be found online [PUBLIC URL of D1.6]. Your information may also be accessed if we are forced to comply with any legal obligations, including audits by the project funders.

We will keep your information in our systems for the duration of the project plus the legally required periods for auditing purposes, currently five years. In addition, your information may be sent to a long-term research repository with your name and any personal identifiers removed and the dataset appropriately anonymized in compliance with the requirements of the GDPR.

Your information will not be transferred outside the European Union at the time of writing. We must make you aware that two of the project partners are located in the United Kingdom, which has started proceedings to exit the European Union in March 2019. The future data protection arrangements are unknown but we shall ensure compliance at all times. If you have specific concerns please let us know.

European legislation on the protection of personal data (GDPR) creates a general right for individuals to request access, rectification or erasure of data. However, the law also allows for certain exceptions for scientific research, which change from country to country and in some cases have not been fully developed at the time of writing. For avoidance of doubt we want to assure you that we will always comply with your requests independently of the legal situation.

**If you have any issues you have a right to lodge a complaint with the Information Commissioner's Office (ICO)/Data Protection Officer (DPO), their contact details are as follows:**

<<Relevant ICO or DPO contact information>>

---