



## WP4. Pilots, citizen involvement, integration and validation.

### D4.1 Pilot preparation and citizen involvement

Grant Agreement N°723139  
NICT management number: 18301

# BIGCLOUT

*Big data meeting Cloud and IoT  
for empowering the citizen ClouT in smart cities*

H2020-EUJ-2016 EU-Japan Joint Call

EU Editor: **LANC**      JP Editor: **NTTE**      Nature: Report

Dissemination: Public

Contractual delivery date: 01/01/2017

Submission Date: 22/03/2017



Co-funded by the EU H2020 GA. 723139 and NICT GA. 18301

---

## ABSTRACT

This document sets out the guidelines for setting up and running pilot trials. It offers guidelines for the type of trial, citizen recruitment, data management and ethics. Even if the trials will take place at later phases of the project, this deliverable will help trial owners to get prepared for the execution and evaluation of trials from the early stages of the project.

---

## Disclaimer

*This document has been produced in the context of the BigClouT Project which is jointly funded by the European Commission (grant agreement n° 723139) and NICT from Japan (management number 183). All information provided in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. This document contains material, which is the copyright of certain BigClouT partners, and may not be reproduced or copied without permission. All BigClouT consortium partners have agreed to the full publication of this document. The commercial use of any information contained in this document may require a license from the owner of that information.*

*For the avoidance of all doubts, the European Commission and NICT have no liability in respect of this document, which is merely representing the view of the project consortium. This document is subject to change without notice.*

## Revision history

Version	Who	When	Notes
V0.1	LANC	2016-09-15	Initial draft
V0.2	LANC	2016-09-24	Added key messages, tidied sections
V0.3	LANC	2016-11-06	Added field trial design tables, and links to ethics process
V0.4	NTTE	2016-11-11	Updates based on NTTE feedback
V0.5	LANC	2016-11-16	Added references to ethics, data handling and informed consent appendices. Formatted into BigClouT template
V0.6	ALL	2016-12-10	Partners feedback
V1.0	LANC	2016-12-27	First draft complete version
V1.2	BRI/NTTE	2017-01-30	Input from Bristol/NTTE
V1.3	GRE	2017-02-07	Input from Grenoble
V1.4	KEIO/NTTE	2017-02-16	Updates from Keio/NTTE
V1.5	ENG	2017-02-18	Input of Idea management tool
V1.6	LANC	2017-03-05	Last review
V1.7	AK	2017-03-08	Quality check
V1.8	CEA	2017-03-21	Update of Grenoble section, final review and submission



## TABLE OF CONTENT

<b>1</b>	<b>BACKGROUND &amp; PURPOSE .....</b>	<b>6</b>
<b>2</b>	<b>EFFECTIVE REAL-WORLD TRIALS .....</b>	<b>8</b>
2.1	UNDERSTANDING THE TRIAL RESEARCH GOALS .....	8
<b>3</b>	<b>CORE GUIDELINES .....</b>	<b>10</b>
3.1	ETHICS .....	10
3.2	STAKEHOLDER ENGAGEMENT.....	11
	<i>OpenIDEAS: an online ideation and engagement platform .....</i>	<i>12</i>
3.3	EVALUATION METHODOLOGY .....	15
3.4	DATA GATHERING (TECHNICAL) .....	17
3.5	TECHNOLOGY COMPONENTS AND RESEARCH OUTCOMES.....	18
<b>4</b>	<b>PLANNED TRIALS – USE OF GUIDELINES.....</b>	<b>19</b>
4.1	BRISTOL .....	19
1.	<i><b>Trial 1: Smart Energy Trial.....</b></i>	<i><b>19</b></i>
2.	<i><b>Trial 2: Mobility Prediction.....</b></i>	<i><b>22</b></i>
4.2	GRENOBLE .....	24
3.	<i><b>Trial 1: Impact of business events to city local economy .....</b></i>	<i><b>24</b></i>
4.	<i><b>Trial 2: Management of Industrial Estates .....</b></i>	<i><b>27</b></i>
4.3	FUJISAWA.....	31
5.	<i><b>Trial 1: MinaRepo - Leveraging Participatory Sensing Technology for Daily City Operation ..</b></i>	<i><b>31</b></i>
6.	<i><b>Trial 2: LokeMon- Location-aware Virtual Character as a Sensor .....</b></i>	<i><b>34</b></i>
4.4	TSUKUBA CITY .....	36
7.	<i><b>Trial 1: Tourist trip planning .....</b></i>	<i><b>36</b></i>
8.	<i><b>Trial 2: Personalized tourism information for overseas visitors .....</b></i>	<i><b>37</b></i>
	<b>SUMMARY/CONCLUSION.....</b>	<b>40</b>
<b>5</b>	<b>APPENDIX 1 BIGCLOUT ETHICS PRIMER .....</b>	<b>41</b>
o	INTRODUCTION .....	41
o	ETHICS BACKGROUND .....	41
o	ETHICAL ASSESSMENT PROCESS.....	43
o	ETHICAL PRINCIPLES .....	44
▪	<i>Overview .....</i>	<i>44</i>
▪	<i>Definitions .....</i>	<i>44</i>
▪	<i>Principles Explained.....</i>	<i>46</i>
o	BIBLIOGRAPHY .....	49
<b>6</b>	<b>APPENDIX 2 – PROCEDURES FOR REAL-WORLD TRIALS.....</b>	<b>51</b>
o	DEFINITION OF REAL-WORLD TRIALS.....	51
o	NATURE OF STUDIES .....	51
o	PARTICIPANTS.....	52
o	DATA .....	52
o	DOCUMENTATION .....	52
<b>7</b>	<b>APPENDIX 3 – GUIDE TO OBTAINING INFORMED CONSENT.....</b>	<b>53</b>
o	ENABLING INFORMED CONSENT .....	53



○	GENERAL POINTS .....	54
○	INTRODUCTORY REMARKS .....	54
○	CONDITIONS FOR PARTICIPATING .....	54
○	RISKS/BENEFITS .....	54
○	ACCESS TO INFORMATION, CONFIDENTIALITY, AND PUBLICATION OF RESULTS.....	55
○	BIBLIOGRAPHY .....	55
<b>8</b>	<b>APPENDIX 4- GUIDE TO SECURE DATA STORAGE .....</b>	<b>56</b>
○	SECURE DATA STORAGE – CORE CONCEPTS .....	56
▪	<i>Pseudonymous Identifiers</i> .....	56
▪	<i>Anonymization</i> .....	57
▪	<i>Adequate Protection</i> .....	57
○	SECURE DATA PROCESSING PRINCIPLES.....	58
▪	<i>General Principles</i> .....	58
▪	<i>Electronic Data Safekeeping</i> .....	58
▪	<i>Physical Storage</i> .....	58
▪	<i>Data Processing</i> .....	59
▪	<i>Data Retention and Deletion</i> .....	59
○	USEFUL SOFTWARE.....	59
▪	<i>Encryption Tools</i> .....	59
▪	<i>Secure Deletion Tools</i> .....	60
▪	<i>Bibliography</i> .....	60
<b>9</b>	<b>APPENDIX 5 THE OPENIDEAS PLATFORM.....</b>	<b>61</b>
<b>10</b>	<b>APPENDIX 6 GRENOBLE QUESTIONNAIRE FOR BUSINESS EVENTS PARTICIPANTS .....</b>	<b>62</b>
<b>11</b>	<b>APPENDIX 7 QUESTIONNAIRE FOR INNOVALLÉE EMPLOYEES IN GRENOBLE .....</b>	<b>65</b>

## LIST OF FIGURES

Figure 1.	BigClouT participating cities.....	6
Figure 2	Stakeholder engagement from ideation to sustainable project. ....	12
Figure 3:	OpenIDEAS - needs, challenges and ideas.....	13
Figure 4.	Bristol Engagement process: Trial 1.....	21
Figure 5.	Bristol engagement process: Trial 2 .....	23
Figure 6.	Grenoble engagement process: Trial 1 on business events .....	27
Figure 7.	Grenoble engagement process: Trial 2 on Industrial Zones.....	30
Figure 8.	Fujisawa: Problem of current information sharing among city staffs.....	31
Figure 9.	Fujisawa: process of field trial preparation.....	32
Figure 10.	Fujisawa: overview of Minarepo .....	33
Figure 11.	Fujisawa: dissemination of field trial .....	34
Figure 12.	Fujisawa: Concept of lokemon.....	35
Figure 13.	Tsukuba Use case Scenario (Tentative).....	37
Figure 14.	Tsukuba City Stakeholder engagement process .....	39



## LIST OF TABLES

---

Table 1. Performance indicators for Trials .....	7
Table 2. Summary of research methods.....	15
Table 3. List of potential stakeholders for Grenoble trial 1 .....	25
Table 4. Innovallee industrial zone as trial space .....	29
Table 5. ACM code of ethicS – Moral Imperatives (excerpt from <a href="http://www.acm.org/constitution/code.html">www.acm.org/constitution/code.html</a> ).....	42



## 1 Background & Purpose

A core aspect of the BigClouT project is a set of real-world trials that will be run in the participant cities, i.e. Bristol, Fujisawa, Grenoble and Tsukuba City. These trials are designed to test the BigClouT architecture and platforms to ensure that technological developments meet the needs of cities allowing them to exploit BigClouT results to develop and deliver new Smart City applications and services.

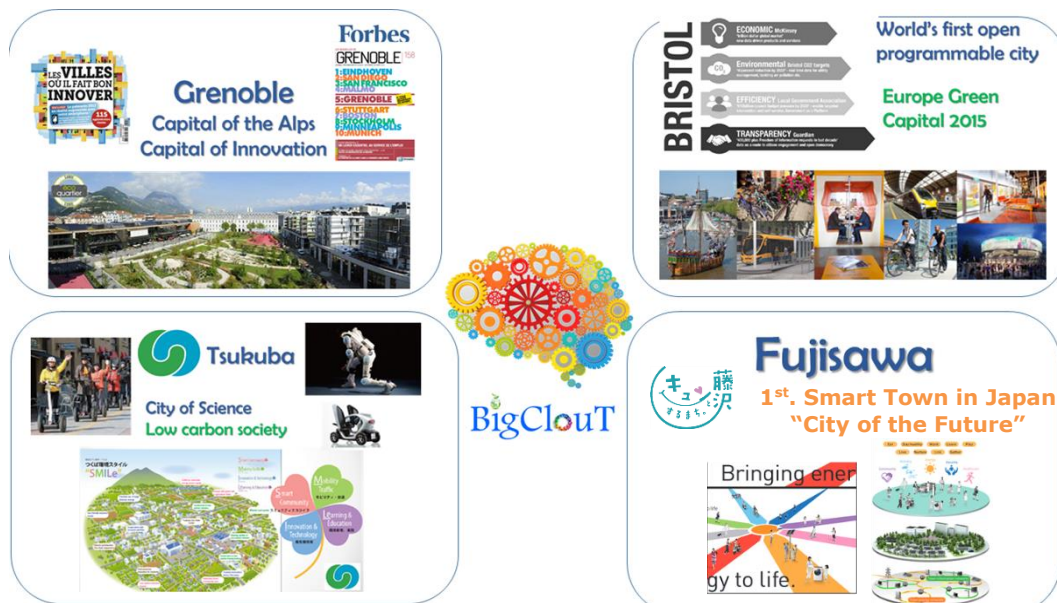


FIGURE 1. BIGCLOUT PARTICIPATING CITIES

These trials will be used to validate aspects of the BigClouT architecture and ensure the project meets its main objectives. In particular the trials will help validate objectives 1-3 which focus on technology development and objective 5 which focuses on long term sustainability, ie legacy. However, the trials are most directly related to ensuring the project meets objectives 4.

- ☐ OBJ1. To build an interoperable architecture enabling data-driven IoT applications
- ☐ OBJ2. To enable self-awareness in smart city platform with programmability and dependability properties
- ☐ OBJ3. To provide libraries and tools for scalable knowledge extraction
- ☐ OBJ4. To design and assess, with citizen and end-user involvement, attractive smart city services

- OBJ5. Propose sustainable dissemination and exploitation plans and create an ecosystem of innovators (SMEs, startups, citizens, etc.) with realistic win-win business models

To ensure Obj4 is met, the project has provided a set of performance indicators which will be used to measure its progress. These are detailed in the table below:

TABLE 1. PERFORMANCE INDICATORS FOR TRIALS

Obj4. BigClouT Trials with citizen involvement		
<b>4.1 Develop a number of viable smart city services and applications with all relevant city stakeholders</b>	Number of stakeholders (e.g., local authority representatives, developers, citizens, SMEs and industrials) involved in service or application definition	5 per city
	Number of citizens ranking use cases by e-consultation	200
<b>4.2 Actively involve end-users in the trial execution and evaluation</b>	Number of trials organized per city	2
	Gather feedback from specific trials and assess sustainability	On 60% of trials

It is important to recognize that the trials are part of an overall demonstration component that begins with use case definition (WP1) and carries through to the core demonstration activities running from M8 to M36. These will take the form of end-to-end technical component integration demonstrations and also field trials based on the use cases allowing us to both monitor progress and evaluate progress against the defined KPIs. Starting from the use case definition phase, the trials will follow 3 main phases: prototype, large-scale deployment and validation. They will take place from M8 to M24 (prototyping) and from M24 to M36 (deployment and validation). To ensure the best results, cities will be actively involved in the coordination of the trials. In the EU, because the cities are full members of the project, they will lead the trials. In Japan, because the cities are not project members, they will cooperate in the coordination, fully supporting the deployment phases.

A special case exists for the planned EU-JP field trial - this is a project goal and will require a more detailed level of coordination. It is expected that the approach to this field trial will be developed during year 2 as the project progresses its planning.

As a final note, this document represents the initial deliverable from the task, a second deliverable is planned for M18 (D4.2). Since the trial planning process is still underway, i.e. the trials are as yet, not well defined, this document has focused on guidelines for the process, data management and ethics. It is expected that D4.2 will contain more details on procedures for running the trials and for analysing results as the actual trials become clearer during M7-18.





## 2 Effective real-world trials

---

Developing an effective trial will involve, in part, a compromise between the needs of the project, city stakeholders and potential end-users. Especially in the case where we are not paying participants to use our systems and where trial systems will be competing for attention with everyday life and its demands. By developing a formative understanding of the needs of the stakeholders, we may be more effective in shaping a trial that can be more successfully adopted.

- A trial will likely have more engagement and higher impact if it meets a genuine stakeholder need, or promises some benefit in return (e.g. an improvement in a particular service they use frequently). There is a question over the length and number of participants in each trial. Longer trials in more naturalistic experimental conditions will require greater engagement from participants, and thus more intrinsic motivation from, or extrinsic compensation (e.g. incentives).
- All participants do not have to be city residents with all their diversity and interests, we may focus on meeting a specific need of a particular demographic or group. The population might be easier to reach and impacts easier to assess and interact with (note that face to face interactions and interviews are expensive in terms of time and staff requirement).
- If the trial also meets the needs of an existing organisational stakeholder (for example), then it is also likely to piggyback on existing efforts and work in harmony rather than in competition with their otherwise potentially busy lives.

Examples of such everyday needs might be:

- Improving an everyday activity like transportation to work, or making payment simpler or less complicated for frequently used services
- Offering new smart city services based on maintaining connections from city dwellers to their remote loved ones outside the city

Fieldwork is key to understanding these needs and alignment between the trials and stakeholder interests. We could consider co-design mechanisms and focus groups to uncover the issues faced by target groups. There will already be government services and charities who are targeting known important city problems, and by working with these we will benefit from their knowledge and experience.

A brief overview of some of the key issues when considering trials is available in the document Procedure for Real World Trials (Appendix 2).

### 2.1 Understanding the trial research goals

While it is clear that understanding end users and stakeholder needs will be important to ensure the trial meets stakeholder goals, it is also important to recognize that the BigClouT trials have a research goal. The research goal dictates the technology chosen to implement the trial, the design of the trial so that it generates the data needed to validate the hypothesis





and the evaluation of the trial. It is critical that all trials designed and developed clearly identify their research goals, their methodology and their evaluation criteria. This may be simply to understand better if a BigClouT technology piece meets a certain objective, e.g. how can edge processing be used to support a particular smart city service scenario, or it could be more user focused, e.g., does the use of virtual gaming characters lead to better citizen engagement with city services.

In all cases, the trial should clearly document what the trial's research goals are, how those will be explored via the trial, how the goals will be evaluated and how the results will be communicated.



### 3 Core Guidelines

---

Core guidelines are those that we think all trials *should follow*, i.e. a trial should have appropriate documentation referring to the guideline. We contrast these with *additional guidelines* (see later) that describe guidelines that projects may *optionally* follow.

It is important to note at this stage that these are guidelines and not mandated. All cities are different and the trials we plan to run have to work within a complex city ecosystem. As such, cities are the ultimate decision makers on whether they will follow a guideline or not.

#### 3.1 Ethics

**Key message:** *all trials should develop an **ethics plan**, addressing the guidelines discussed below and in particular, ensuring they meet the requirements laid out in internal ethics process described in D7.1*

Each trial will raise a number of ethical issues as stakeholders are engaged, data is gathered, opinions sought and trials deployed. It is important that each trial develops an ethics process at an early stage and uses the ethics process to guide the trial and its engagement with stakeholders and end users.

Ethically, BigClouT trials should be considered '**low risk**', since we are unlikely to be working with at risk groups, employing deceit, or conducting any form of experiment that could harm participants. Participants will be healthy consenting adults. Risk will be largely confined to potential privacy threats to data concerning participants ('personal data') and their potential identification, and or involve 'tracking location or observation of people').

WP7 within BigClouT has already laid out initial ethics requirements - these mainly focus on the protection of personal data (PoPD). The project has an ethics committee (EC) in place that will review both the use cases and the trial plans to ensure that any issues are identified and handled correctly. The deliverable D7.1 mandates some core requirements and (will) include a discussion of data protection issues. The NICT handbook, "Private data Protection" is recommended as one of the guidelines for handling personal data.

We recommend the adoption where possible of the following set of ethical principles as previously enumerated in the FP7 PD-NET FET-Open grant number: 244011 ethical handbook (<http://pd-net.org/ethics/>). Specifically, that BigClouT trials:

1. Maximize Possible Benefits and Minimize Possible Harms
2. Obtaining Voluntary Informed Consent
3. Ensuring Right to Withdraw
4. Disclosing Detriment Arising from Participation in Research
5. Providing Data Protection and Privacy
6. Limiting Disclosure
7. Following Minimal Intrusion Principle
8. Offering Adequate Incentives
9. Special Provisions for Experiments Involving Children and other Vulnerable People
10. Avoiding Deception



These are explained in the associated project ethics primer (Appendix 1).

A key requirement is to follow the internal ethics process detailed in Deliverable D7.1. In particular, each trials should provide details of:

1. *Synopsis of trial*
2. *Type of participant expected*
3. *Type of data to be collected*
4. *Ethical approvals*
  1. *Copies of ethical approvals*
  2. *Existing overall process to get these approvals*
5. *Detailed information on the procedures that will be implemented for data collection, storage, protection, retention and destruction:*
6. *Details on the procedures and criteria that will be used to identify/recruit research participants*
7. *Procedures for participant information*
8. *BigClouT project information sheet*
9. *Participant informed consent forms. Including details on:*
  1. *Details on the procedures and criteria that will be used to identify/recruit research participants must be provided.*
  2. *Detailed information must be provided on the procedures that will be implemented for data collection, storage, protection, retention and destruction and confirmation that they comply with national and EU legislation.*
  3. *Confirmation: the applicant must explicitly confirm that the data used are publicly available.*

In addition to the core ethics guidelines on data protection and participant consent, other ethics issues that should be included in the ethics process include:

- Accessibility and the digital divide. How to ensure all citizens have access to the trial, not just ‘young, digital literate adults’

Note. The ethics guidelines discussed in this document will need to adapt to Japan side situations. Where applicable, Japan side processes, agreements and norms take precedence.

### 3.2 Stakeholder engagement

**Key message:** *all trials should develop a **stakeholder engagement plan**, addressing the guidelines discussed below and in particular, containing a participant recruitment and management section.*

A critical aspect of all trials that will be conducted by the BigClouT project is the need to engage stakeholders at all stages of the trial. It is important for both the validity of the trial and for the overall goals of the project that trials identify the key stakeholders at each stage of the trial. While this document is focused primarily on the trial phase, it is important that a trial considers the definition stage (use case from WP1) and ensure continuity of stakeholder engagement from initial engagement through use case definition, project proposal and into the project implementation and trial.



As a general guideline, BigClouT trials should follow the general process laid out in Figure 2.

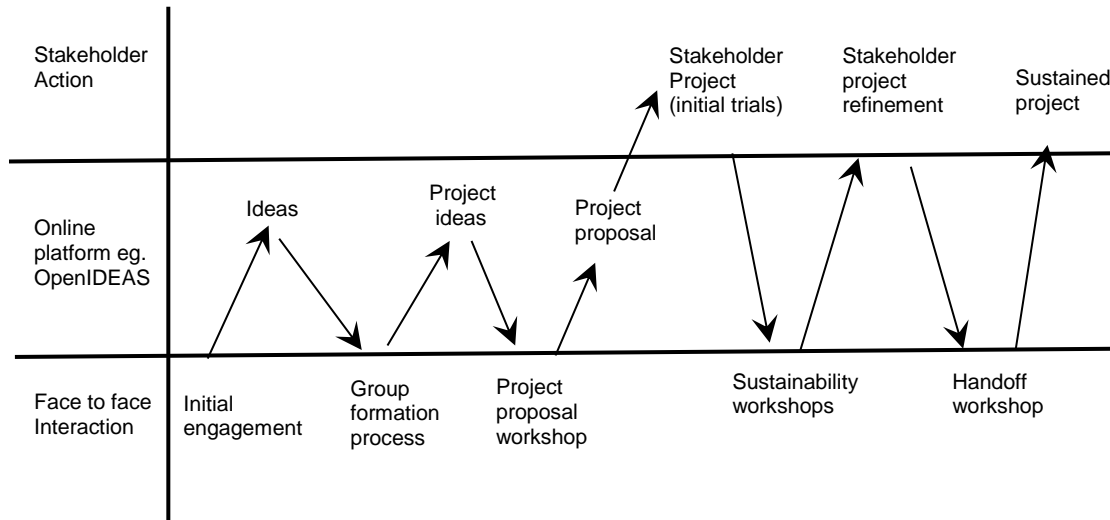


FIGURE 2 STAKEHOLDER ENGAGEMENT FROM IDEATION TO SUSTAINABLE PROJECT.

As Figure 2 shows, the process from idea to project begins with initial face to face discussion groups involving stakeholders such as users or businesses, results in ideas that are posted to an online discussion group. These ideas are then refined in workshop that aims to have groups form around idea. Once formed, groups prepare project ideas for online discussion, and then a 2<sup>nd</sup> workshop is used to prepare project proposals. These are again posted to the online forum which then form the basis of an actual project (or trial). Once a project trial is underway, sustainability workshops should be run with stakeholders to monitor progress and to begin to generate ideas with the stakeholder community of how to ensure the project continues after the trial phase is over.

BigClouT project will make use of the OpenIDEAS tool from the partner Engineering for the idea generation phase. The tool is described in the section below.

#### *OpenIDEAS: an online ideation and engagement platform*

OpenIDEAS is an Idea Management System providing a social and collaborative environment to express, identify and discuss needs, problems, ideas and possible solutions; in particular its main aim is to close the gap between citizens and public administrations.

OpenIDEAS establishes a co-definition and co-creations environment providing tools to identify and solve problems through collaboration between different stakeholders such as Public Administrations, citizens, businesses, academics, associations, etc, and to promote their participation in the co-definition of ideas and solutions.

OpenIDEAS provides functionalities to manage three main concepts: need, challenges and ideas.

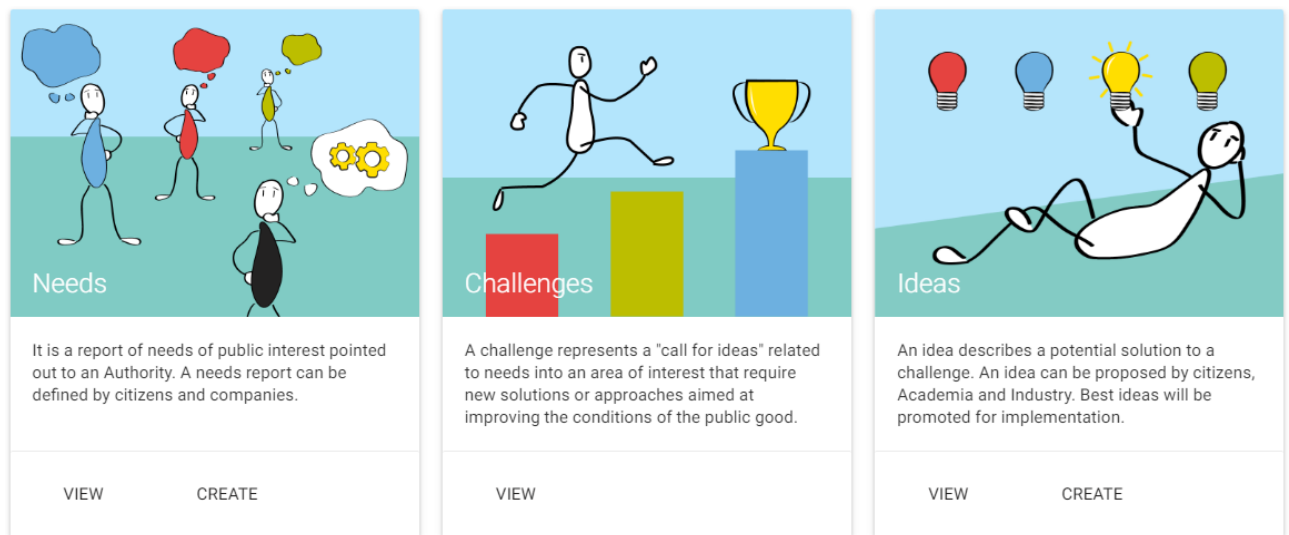


FIGURE 3: OPENIDEAS - NEEDS, CHALLENGES AND IDEAS

A **need** represents a report about something of public interest (for instance a problem, or a necessity) pointed out to an authority, for example a Public Administration.

**Challenges** represent "calls for ideas" to identify solutions to solve a problem; challenges can be created only by authorities (e.g. Public Administrations) and they can be derived from a need or can be created from scratch.

An **idea** represents a possible solution for a problem. An idea can be proposed for a specific challenge or can be submitted to an authority.

OpenIDEAS enables user to discuss about needs, challenges and ideas in order to allow authorities to identify the most relevant problems to solve (i.e. needs) and the best solutions (i.e. ideas).

Furthermore, description of needs, challenges and ideas can be enriched with documents (e.g. images, text document, etc. that can be attached), geographic information (e.g. point of interests), keywords and topics (such as: environment, energy, economy, society, social services, transportation, mobility, education, culture, city government, business, tourism, leisure, public safety, quality of life, etc.).

Ideas follow a precise lifecycle composed of five steps:

1. Idea generation: the idea is submitted.
2. Evaluation and selection: users discuss about the idea, provide feedback and evaluate; in this phase the authority is able to evaluate the idea and to select it; when the idea is selected, it moves on the next step.
3. Refinement: in this phase author of the idea collaborates with the authority in order to improve the idea; when the refinement of the idea is completed, the idea can move on the implementation phase.
4. Implementation: in this phase the idea is realized.
5. Monitoring: in this final phase, the implementation of the idea is monitored.

Using this process and the OpenIdeas platform as a general framework for stakeholder engagement, we can identify the following guidelines:

A **stakeholder engagement plan** should be developed that addresses the following issues:

**Identify stakeholders.** Who are the stakeholders that will be involved in the trial? - stakeholders may include government (Federal, Municipal, Local, etc.), regulators, land & property developers, ICT service providers, systems integrators, utility providers, transport operators, citizens, etc. For each stakeholder describe their involvement, goals and activities during the process leading up to the trial, the trial itself and the evaluation/sustainability phase of the trial. BigClouT has committed to engaging at least 5 stakeholder groups for each trial.

**Identify trials user and develop recruitment plan:** As part of the stakeholder engagement plan, it is important that all trials have a clearly defined recruitment plan that lays out who are the target users, how they will be recruited to the trial (including promotion/marketing etc.) how they are engaged/motivated during the trial etc. Obviously each trial will have different requirements depending on the nature of the plan, but the following points should be considered.

- Specify the minimum number of participants required for the trial (our proposal states we will engage 200 users in each use-case evaluation)
- Specify the target mix of participants (male/female, age, demographics etc)
- Identify mechanisms for recruiting participants. Discuss the recruitment channels and partners e.g. city stakeholders? Do we go via particular interest groups? What population or demographic make up are we interested in recruiting?
- Specify the expectations for the participants, i.e. How many engagements, how often and for how long?
  - Normally there is attrition in trials, especially over longer periods, so significance of losing participants should be considered.
- Identify the incentives for the users to participate. Is there a reward for engaging with the project, e.g. a bursary or fee/competition entry?
- How does the city promote and publicise Smart City trials to stakeholders?
- How do we continuously keep in touch with the participants, mainly in order to get feedback (any kinds: needs, feedback to the app/service, etc. etc.)
- Additionally, we should ensure that the results of the on-going evaluation and the final evaluation are fed back to stakeholders so they can see progress. A useful part of this feedback would be appropriate visualizations to communicate results.
- Resourcing plan for the trial. Outline the resource needs of the trial covering the resources needed to recruit and interact with end users, fixing bugs, gathering data (sensing, questionnaires etc), management and sharing of data etc.

Recruiting participants and obtaining informed consent is covered by the [obtaining informed consent primer \(Appendix 3\)](#).





### 3.3 Evaluation methodology

**Key message:** All trials should develop an **evaluation and goals** statement that outlines the goals of the trial, the project KPIs that will be met, and the evaluation methodology that will be used for the trial. A key question is whether to design an experiment or assess user experience.

The design of the user trial and its evaluation methodology are critical to the research goals. Of particular note is the type of trial that is undertaken. For example the goal of the trial may be primarily to assess a user experience to provide feedback on a particular service or application idea, or to test a hypothesis about user behaviour. Alternatively it may be an experiment to measure the performance of a particular piece of BigClouT software, for example the performance of the edge computing capability of the D-NR software component. These different trials could be carried out in different ways, for example understanding user feedback may primarily come from questionnaires and surveys based on a lightweight or artificial trial, or they could come from experimentation in the lab which is extrapolated into a real world trial, or from real world trials using a natural setting ie real world city deployments. Each type of trial has different strengths and weaknesses. (A useful introduction can be found in Wynekoop and Conger Wynekoop and Conger<sup>1</sup> See also<sup>2</sup>)

In the table below, we summarize a number of different approaches to running trials (Enquiry type) and highlight the strengths and weaknesses of the different approaches.

TABLE 2. SUMMARY OF RESEARCH METHODS

Enquiry Type	Purpose	Strengths	Weaknesses
Interaction logging	Tracking interaction frequency/ time	Scales to large number of participants. Invisible to participants. Shows patterns of use.	Motivation for engagement or disengagement not captured (need observation/interviews). Privacy invasive.
Experience sampling	Samples non-functional or motivation 'in context'	Provides data from the field without need for direct observation	Participant inconvenience/ fatigue. Reduced return rates.
Follow up questionnaires	Sample subjective user experience	Scales to large number of	Low completion rates (10% typical),

<sup>1</sup> Wynekoop, J.L. and Conger, S.A.: A Review of Computer Aided Software Engineering Research Methods. In Proceedings of the IFIP TC8 WG 8.2 Working Conference on The Information Systems Research Arena of The 90's, Copenhagen, Denmark (1990)

<sup>2</sup> Jesper Kjeldskov, Connor Graham. A Review of Mobile HCI Research Methods [Human-Computer Interaction with Mobile Devices and Services](#) Volume 2795 of the series [Lecture Notes in Computer Science](#) pp 317-335



		participants.	depending on incentives. Self-report rather than objective measures. Needs careful design to balance.
Ethnographic observation/interviews	Understand how technology fits with everyday life	Rich qualitative data source. Insight into appropriation and adoption of technologies.	Small 'n' due to resource limitations. Requires skilled practitioners. Participants are aware they're observed.
Focus groups	Engage with specific stakeholder groups	Lots of information in a short and cost effective way.	Small 'n'. May have 'group think' effects. Subjective and based on opinion rather than field observation or objective measures.
Measuring application/service-specific quantitative performance indicators	Measuring how the evaluated software/application/service is performing towards the designed specific goal.	Direct data gathering	Motivation and engagement subject to end user vagaries

Questions that should be considered in the ***evaluation and goals statement***, which have a bearing on study, software and evaluation design include:

- What experience are we trying to measure? For which stakeholders?
- How do we go about surveying or measuring this?
- Short or longitudinal evaluation/experiences?
- Do we need to track engagement? Qualitative/quantitative metrics of engagement?
- What do we need to know about participants for our analysis/conclusions and how do we protect their privacy?
- Which parties are conducting the evaluation (are we relying on 3rd parties or self-reporting)?
- Are we looking to measure statistically significant effects or improvements?
- Technology side evaluation. What aspect of the BigClouT technology platform is being evaluated and what is the methodology used for that evaluation. See section "Technology Components and research outcomes"?
- When do we measure user's evaluation (e.g., satisfaction)? (probably both before and after the experiment, in order to compare the satisfaction of the end users)



- Also, related to this, 2 evaluations “before” and “after” need to be the same evaluation in order to compare the results. Thus, we need to carefully design the evaluation before the actual experiment starts.

Careful thought needs to be given to the evaluation methodology as it drives the overall way that the experiment will be designed and carried out. This requires that early planning is needed, even at the use-case stage so that the role of the use-case is clear in the overall evaluation. Failure to carry this out in the early phase of the project is likely to result in use-cases and therefore experimental trials that are interesting in their own right, but provide no useful data on the value of the BigClouT technologies as a framework for smart city services and applications.

### 3.4 Data gathering (technical)

**Key message:** *All trials should develop a **data management plan** describing data that will be captured, its format and how it will be managed. Additionally, details of what data sets will be shared using the BigClouT data repository should be provided.*

The ethical issues of gathering data from and about users is discussed in the Ethics section (Above). This section is focused on the technical aspects of data gathering with a goal of ensuring that all project partners can access and use data. Since the exact nature of the trials is still under definition, this section provides some general guidelines. Once the trials have been specified, D4.2 will provide more concrete statements about the trial data gathering.

We should take care to understand the “story behind the data” that we use. Data is often partial and may be intentionally or unintentionally biased or selective.

What data do we need in order to capture the effects of the trials or interventions?

- Types of data
  - Is this qualitative or quantitative data?
  - What scale of effect are we trying to observe and what is the size of the effect we are trying to measure?
  - Careful consideration needs to be made of the challenges of collecting data (especially qualitative interview or experience data) at scale?
- Format of data
  - Guidelines on types of data to be gathered, e.g. user data, usage data, technology performance data etc.
  - How to share data between partners/sites – all projects should identify the data sets they will collect at the use-case and trial stage and should indicate how they will make those available to other project partners.
  - Data formats – do we want to adopt some common data formats so we can easily share data, eg XML, JSON, others? A challenge in smart city projects are the bespoke formats of data, which makes comparability across trials and cities difficult. Open data formats should be adopted, if possible. We should be clear on what we need to compare across trials/cities.
- Management of data
  - Roles, keys and credential management for accessing data without violating ethics

- Suggest we create a data management plan (if that's not in this document) so there is clarity on how we handle data/sensitive data for the project.
- Curating data after project end
  - The data management plan for each trial should specify how data is treated at the end of the project, including which data is disposed of, and which retained as a project output.
  - Data retention should be observant of recruitment protocol agreements, and the sensitivity of the data (e.g. only anonymised data is suitable for publication that does not reveal personally identifiable information)
  - Should we include a plan for after the project?

The secure handling of data to minimise risk and privacy violations is also discussed in the document “Guide to secure data storage” provided as Appendix 4.

### 3.5 Technology components and Research Outcomes

**Key Message:** *All projects should provide an indication of BigClouT technology components and infrastructure they plan to exercise in the trial and how this helps meet the 5 core objectives of the project.*

As discussed earlier (Sec 2.1), it is important to recognize that the BigClouT trials have a research goal. The research goal dictates the technology chosen to implement the trial, the design of the trial so that it generates the data needed to validate the hypothesis and the evaluation of the trial. As discussed in the evaluation methodology section (Sec 3.3) it is critical that all trials clearly identify their research goals, their methodology and their evaluation criteria. Part of that evaluation criteria is an evaluation of the use of the core BigClouT platform and its underlying technology component.

Returning to the core objectives of the project:

- OBJ1. To build an interoperable architecture enabling data-driven IoT applications
- OBJ2. To enable self-awareness in smart city platform with programmability and dependability properties
- OBJ3. To provide libraries and tools for scalable knowledge extraction
- OBJ4. To design and assess, with citizen and end-user involvement, attractive smart city services
- OBJ5. Propose sustainable dissemination and exploitation plans and create an ecosystem of innovators (SMEs, startups, citizens, etc.) with realistic win-win business models

Each trial should indicate which of core technology objectives it is exploring (Obj1-3) and which specific technologies it is using.

- During use case development it would be helpful to identify potential components exercised
- Projects should report on the experiences using technology components as part of the final trial analysis/report



## 4 Planned trials – use of guidelines

In this section we provide details of the initial planned trials of the four pilot cities of the project: Grenoble, Bristol, Fujisawa and Tsukuba, We outline how they have used the trial guidelines in this document for initial planning.

### 4.1 Bristol

#### 1. *Trial 1: Smart Energy Trial*

This trial is about exploiting BigClouT's novel data-adaptive machine learning techniques for predictive analysis and the power consumption of users. The trial will reuse the infrastructure installed by the European project REPLICATE<sup>3</sup>.

The objective of the project is to make householders aware about different phenomenon, that otherwise would have a very difficult detection like 'the phantom load' also known as 'vampire power'. This is the electricity consumed by electronic and electrical appliances while they are switched off (but are designed to draw some power) or in a standby mode. This consumption may be of the order of 10% of the electrical energy used by a typical household.

Saving electricity not only will affect the house holder pocket, but electricity is very often generated by combustion of hydrocarbons (oil, coal, gas) or other substances, which releases substantial amounts of carbon dioxide, implicated in global warming, and other pollutants such as sulphur dioxide, which produces acid rain, so at the same time the user is helping to take care of the planet.

#### **Ethics plan**

- **Type of participants expected**

It is only available to owner occupier and privately rented properties within the Bristol City Council area and moreover the precondition is they had to join the Warm Up Bristol program.

- **Type of data to be collected**

Mainly Electric consumption and environmental data will be collected from the homes involved. No other type of Personal identifiable information will be collected. (TBC)

- **Ethical approvals**

The University's Ethics Policy and Procedure<sup>4</sup> governs the ethics of research across the University of Bristol. It applies to all staff, students and anyone else carrying out research

---

<sup>3</sup> <http://replicate-project.eu/>

<sup>4</sup> [http://www.bristol.ac.uk/media-library/sites/red/documents/research-governance/research\\_ethics\\_policy\\_v6\\_220515\\_FINAL.pdf](http://www.bristol.ac.uk/media-library/sites/red/documents/research-governance/research_ethics_policy_v6_220515_FINAL.pdf)



under the auspices of the University. All research that has ethical implications or involves human participants, their tissue and/or data must have an ethical review.

- **Details on the procedures and criteria that will be used to identify/recruit research participants.**

For recruitment of participant households we will utilise the existing Warm Up Bristol recruitment campaign with Involvement of KWMC (Knowle West Media Centre) to carry out targeted recruitment of one hundred and fifty residential participant households clustered in the specified demonstration district. This will link into the citizen engagement and involvement actions. The process will consist of:

- Survey of properties. (Bristol Energy Service)
- Calculating cost of installations (Bristol Energy Service)
- KWMC Producing combined collaterals (subject to workshop)

### **Stakeholder engagement plan (Smart Energy)**

- **Specify the minimum number of participants required for the trial**

REPLICATE will install at least 150 Home automation units in different houses around the city, so it can be said that 150 will be the minimum number

- **Specify the target mix of participants (male/female, age, demographics)**

According to the REPLICATE project, Bristol will contract works via its Warm Up Bristol contractor framework in the retrofitting of 150 residential buildings in Easton and Lawrence Hill Neighbourhood.

It is not an individual personal experiment, because it is based on the occupants of the whole house and based on the demographic information provided by the city council<sup>5</sup>, we can estimate that:

- Age from 0 to 15: 25%
- Age from 16 to 64: 67%
- Age from 65 or older: 8%

Related to the gender it can be said that the male/female distribution is around 50%

- **Identify mechanisms for recruiting participants. Discuss the recruitment channels and partners**

We are going to reuse the sensors from the REPLICATE project, so the participants will be the residents who have already signed up for the scheme and have collected relevant household information<sup>6</sup>

- The process will consist of:
  - Survey of properties (Bristol Energy Service)
  - Calculating cost of installations (Bristol Energy Service)

<sup>5</sup> <https://www.bristol.gov.uk/documents/20182/33904/Mid-2015+Population+Estimates+for+Bristol+Local+Authority/7a8232da-baa5-4be7-913b-ba2560f2b459>

<sup>6</sup> <https://warmupbristol.co.uk>





- KWMC (replicate specific products) Producing combined collaterals (subject to workshop)
- **Specify expectations for the participants, How many engagements, how often and for how long?**

Once the sensors have been deployed in the selected homes, no more interactions would be needed from the participants, because this research has been designed for to be a non-intrusive pilot.

- **Identify the incentives for the users to participate**

Using REPLICATE's infrastructure, exploit BigCloudT's novel data-adaptive machine learning techniques for predictive analysis to save money on their electric bill.

### Smart Energy trial is at the project proposal phase

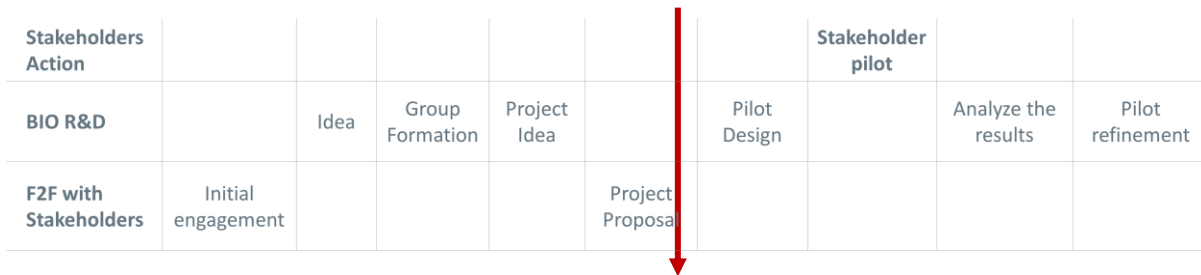


FIGURE 4. BRISTOL ENGAGEMENT PROCESS: TRIAL 1

The initial engagement was during the ICT and energy demand management workshop involving Bristol City Council, University of Bristol and Bristol is Open, and thus we get the idea of merging Smart Energy with Smart Homes.

The following steps were about involving KWMC, who had previously experience deploying sensors inside homes.

During the following step, we were developing deeper the idea, until we arrived to the project proposal that it is still open, and it will grow with the project.

### Data Management plan (Smart Energy)

We are starting the pilot design, but it is planned that the data will be stored in a dedicated FIWARE instance deployed inside the Bristol is Open cloud using different platform assets like short term historic<sup>7</sup> so potentially it will be protected by OpenStack Keystone<sup>8</sup>. Data will be isolated from the internet by two firewalls and Public key cryptography<sup>9</sup>.

Data will be available through the FIWARE version of the OMA NGSI 10 interface. It is a RESTful API via HTTP. Its purpose is to exchange context information<sup>10</sup>.

<sup>7</sup> <http://fiware-iot-stack.readthedocs.io/en/latest/sth/>

<sup>8</sup> [http://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/OpenStack\\_Keystone](http://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/OpenStack_Keystone)

<sup>9</sup> [https://en.wikipedia.org/wiki/Public-key\\_cryptography](https://en.wikipedia.org/wiki/Public-key_cryptography)

<sup>10</sup> [https://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/FI-WARE\\_NGSI-10\\_Open\\_RESTful\\_API\\_Specification](https://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/FI-WARE_NGSI-10_Open_RESTful_API_Specification)

## 2. Trial 2: Mobility Prediction

- **Synopsys of trial**

Bristol Is Open (BIO) together with University of Bristol High Performance Network Group will deploy the use case of using Data Analysis to extract the citizen mobility pattern in a smart city environment. Flows of monitoring data will be processed in a real-time, streaming manner to ensure low latency analytics.

- **Type of participants expected**

Mainly, the participants that we expect are citizens which usually walk around the harbour.

- **Type of data to be collected**

We are going to collect MAC address, the position of the participants and the timestamp of each one of these measures. We are aware that it is confidential information, so the mac addresses will be hashed directly at the sniffer device.

- **Ethical approvals**

The University's Ethics Policy and Procedure governs the ethics of research across the University of Bristol. It applies to all staff, students and anyone else carrying out research under the auspices of the University. All research that has ethical implications or involves human participants, their tissue and/or data must have an ethical review.

- **Procedures for participant information**

We have been checking deeper into the use case. According to several reference sources<sup>111213</sup>, MAC address is not considered as Personal identifiable information provided we don't link it with any other information (phone, name, sex, age ....). In that case, we are not required to ask for approval to the participants to get involved in.

From Bristol is Open and University perspective, we want to be transparent with the citizens and explain clearly what is the proposal for retrieving data and how we are going to use it. We are still working on the best way of communicating to the citizens.

### Required Data for Stakeholder plan (Mobility Prediction)

- **Specify the minimum number of participants required for the trial**

Around 1000 MAC address should be enough in order to characterize the "Hot Paths"

- **Specify the target mix of participants (male/female, age, demographics)**

According to the Bristol City Council demographic research about the neighbourhood of BS1 (Bristol city centre and Redcliffe) the whole population of BS1 is 2545 persons<sup>14</sup>:

---

<sup>11</sup> <https://fpf.org/2014/03/27/mac-addresses-and-de-identification/>

<sup>12</sup> <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6525606&tag=1>

<sup>13</sup> <https://7suite.com/2016/09/what-is-pii-personal-data>

<sup>14</sup> <https://www.bristol.gov.uk/statistics-census-information/the-population-of-bristol>



- Less than 18 years old: 12%
- Between 18 and 65 years old: 81%
- More than 65: 7%

The female population is around 47% and the male is 53%

- **Identify the incentives for the users to participate**

Users will benefit if they know the real time and the historical records about the number of people that use their route. Users could for example, decide the place to have lunch if they don't have time or choose the shorter route in their way to the office.

Institutions based on serving the citizens, like the city council (i.e. planning the road maintenance in off peak times) or the waste collection service, could use this data for improving their services (i.e. adapting the street sweepers and cleaners hours), which will improve the citizen life.

### Mobility Prediction trial is at the project proposal phase:

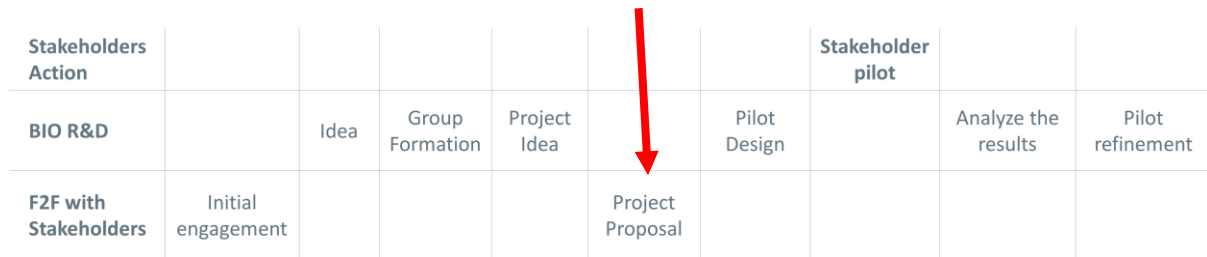


FIGURE 5. BRISTOL ENGAGEMENT PROCESS: TRIAL 2

### Data Management plan (Mobility Prediction)

We are starting the pilot design, it is planned that the last value of the data will be stored in a dedicated FIWARE instance deployed inside the Bristol is Open cloud, potentially it will be secured by OpenStack Keystone<sup>15</sup>. Anonymization (i.e. hashing) will be done during data acquisition, to protect it. Data will be isolated from the internet by two firewalls and Public key cryptography<sup>16</sup>.

Data will be available through a FI-WARE version of the OMA NGSI 10 interface. It is a RESTful API via HTTP. Its purpose is to exchange context information<sup>17</sup>.

<sup>15</sup> [http://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/OpenStack\\_Keystone](http://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/OpenStack_Keystone)

<sup>16</sup> [https://en.wikipedia.org/wiki/Public-key\\_cryptography](https://en.wikipedia.org/wiki/Public-key_cryptography)

<sup>17</sup> [https://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/FI-WARE\\_NGSI-10\\_Open\\_RESTful\\_API\\_Specification](https://forge.fiware.org/plugins/mediawiki/wiki/fiware/index.php/FI-WARE_NGSI-10_Open_RESTful_API_Specification)

## 4.2 Grenoble

### 3. *Trial 1: Impact of business events to city local economy*

#### **Synopsis of Trial**

Grenoble-Alpes Métropole hosts several large events, trade shows and fairs every year in its Alpeexpo exhibition centre. In the future, it would like to host an increasing number of events in order to boost the attractiveness of the area and in turn to boost economic development.

Currently, there is no way for the Métropole to measure the economic impact of these events - for example - the use of hotels, shops, restaurants & transport – by the people attending these events.

The Métropole would like to develop a tool that allows this monitoring to take place.

In turn, these results would be used to better attribute public resources to improve public services for the visitors (transportation, tourism, etc.).

#### **Ethics Plan**

Grenoble-Alpes Métropole has been consulting with legal experts, subcontracted to assist the European partners of this project in order to ensure that the ethics requirements are adhered to. **Two meetings** have already been taken place to exchange about the planned use cases with the legal experts in order to adequately cover requirements in terms of ethics. Details will be presented in the Deliverable 7.1.

#### **Type of participant**

This use case is particularly targeting the visitors coming to participate a specific event or fair organised in the city. He/she expects to have a good experience during his/her stay in the city with a maximum amount of services, information and recommendation provided in terms of transportation, tourism, social events, etc.

#### **Type of data to be collected**

In return of provided services, we expect to be able to monitor some behaviour from the participant, such as the transportation mode used, restaurants and shops visited, amount of money spent during the stay, etc. in a complete anonymised or user controlled way. The user will be using a mobile applications, which will be in interaction with its surrounding environment composed of sensors, actuators and other devices.

The need to for anonymity is very important concerning information regarding individuals' movements and spending. It is therefore paramount to have a critical mass of participants in order for the results to be anonymised.


## **Stakeholder involvement**

A number of stakeholders have been identified and contacted regarding this use case as listed in the table below:

TABLE 3. LIST OF POTENTIAL STAKEHOLDERS FOR GRENOBLE TRIAL 1

 <p>Alpexpo Exhibition Center</p>	<p>Composed of modular imbricated and interconnected spaces, 2 congress centres, 1 concert hall and 2 exhibition halls, Alpexpo is the main event complex in Grenoble. Approximately 70 different events (of various sizes, attracting varying different attendees) are organised in Alpexpo every year. Of these, 10 can be classified as international events. A meeting already organised with Alpexpo on December 2016, in which Alpexpo expressed its high interest for the project use cases.</p>
 <p>Grenoble World Trade Center</p>	<p>Part of the World Trade Center Association (WTCA), Grenoble's WTC has 2 500 m<sup>2</sup> of space for hosting international events just next to the Grenoble main train station.</p> <p>WTC Grenoble organises many international events along the year and ready to collaborate. WTC has an event management system that can already provide us useful information about the congress participants.</p>
 <p>Maison Minatec exhibition center &amp; Insight Outside (events management company)</p>	<p>Maison MINATEC hosts 1,000 sq. m of dedicated special-events space, including an amphitheater, meeting rooms, and reception areas. Located near the railway station and tram stop, Maison MINATEC has the necessary facilities for accommodating some 40,000 annual visitors at the heart of its high-tech MINATEC innovation campus.</p> <p>Maison Minatec is located inside of the facilities of CEA and managed by Insight Outside, company in charge of organising events. Several meetings took place with those stakeholders that show high interest to project results.</p>



	<p>Grenoble Tourism office is in charge of tourism for pleasure and business tourism in the area. The tourism office has been contacted and is interested in helping us to develop this use case. For example, a tool is currently in development to regroup all tourist activities in the area (museums, visits to the bastille fortifications, walking tours...) which could be integrated in the app that we want to create for Big ClouT</p> <p>In addition to the tourism office, we are planning to contact the CCI, club des hoteliers and réseau Label Ville, which can provide us interesting feedbacks from their experience of interaction with local shops, restaurants, hotels, etc, in order to give better directions to our use cases.</p>
<p>Grenoble Tourism Office</p> <p>Chamber of Commerce &amp; Industry (CCI)</p> <p>Club des hoteliers (Group of hotels)</p> <p>Réseau Label Villes (Group of shopkeepers)</p>	

The actors listed above have already expressed a clear interest in the results of the project. The fact that we are currently missing an evaluation tool for the economic impact of events is very clearly a problem for event organisers.

In addition to the meetings organised, during the international Semicon tradeshow, that took place in Grenoble in October in Alpexpo Exhibition Center, a survey was carried out with participants to see whether they would be interested in using this app. The short survey can and the results can be found as Annex in the Section 10. The survey was done to 16 person. 15 over 16 mentioned that they would be interested in a mobile application to guide him/her around Grenoble, and 14 of them agreeing to have it in a combined way with a professional usage. 10 over 16 agrees on giving some anonymised personalised information for the sake of having better service, while 2 of them were not sure and 3 of them refused. The responses were therefore overwhelmingly positive, which validated the project idea and encouraged us to continue progressing on this use case and proposing a project proposal as a next step





**Business events use case is at the project idea phase:**

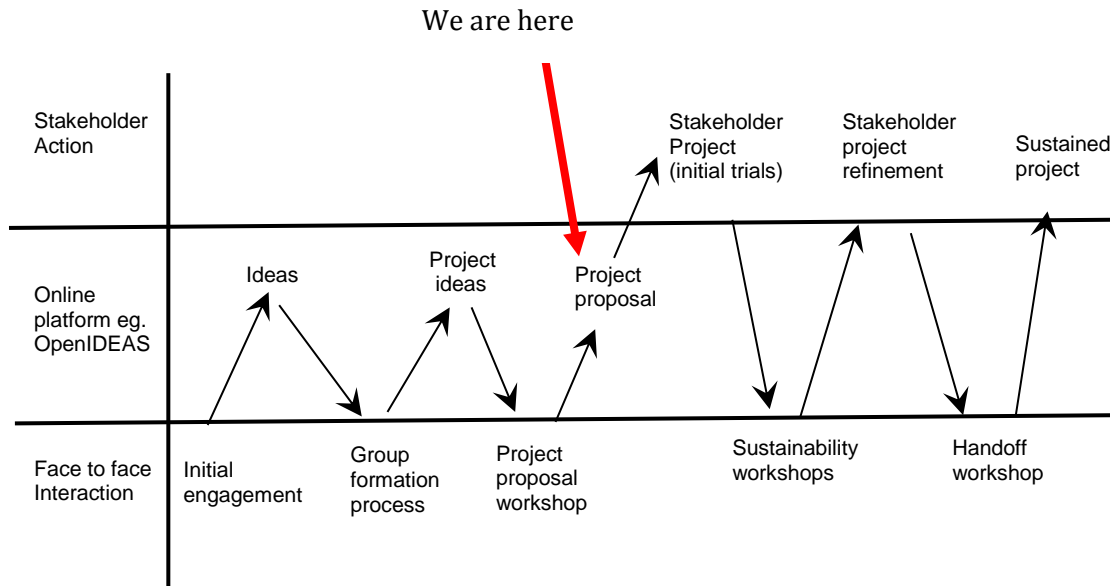


FIGURE 6. GRENOBLE ENGAGEMENT PROCESS: TRIAL 1 ON BUSINESS EVENTS

#### 4. Trial 2: Management of Industrial Estates

##### Synopsis of Trial

Grenoble-Alpes Métropole owns or manages several different industrial estates on its territory. These estates range in size and in the areas of activities, from high tech companies to artisans to shops and services.

The Métropole wants to be able to know who is using these estates and how. For example, how many people arrive each day by different modes of transport, how many people leave the sites for lunch, how many deliveries are received on the site, etc.

The goal is to be able to improve the services proposed by the Métropole in these zones, (for example, is there a need for improved public transport? Is there a need to put in place a canteen?) and also to create a social network enabling for grouped orders and potentially car-sharing solutions.

The goal is to be able to better attribute public resources, to improve the working conditions on these zones and also to improve their environmental impact.

Three principal stakeholders have been identified and contacted regarding this use case:

- Inovallée Industrial Estate
- Espace Comboire Industrial Estate
- Presqu'île group of employers

### **Ethics Plan**

Grenoble-Alpes Métropole has been consulting with legal experts, subcontracted to assist the European partners of this project in order to ensure that the ethics requirements are adhered to. **Two meetings** have already been taken place to exchange about the planned use cases with the legal experts in order to adequately cover requirements in terms of ethics. Details will be presented in the Deliverable 7.1.

### **Type of participant**

This use case is mainly targeting the Employees working in the industrial zones in order to improve their daily life at their working place environment. Visitors to the zone (for meetings, events, etc.) are also other type of participant that can benefit from the target application.

### **Type of data to be collected**

The application is planned to be in form of a mobile application interacting with the user and the BigClouT platfor, that will centralise various types of data sources such as types of transportation, traffic jams, number of vehicles in the zone, number of people visiting the area, environmental data, status of restaurants (menu, crowdness, offers, etc.), event information, participatory sensing data (employees providing data with crowd sourcing techniques), etc.

### **Stakeholder involvement**

As the most organized industrial estate in the city area, it has been decided to use the Inovallée area as a first test case. It is therefore with the management of this estate that discussions have progressed most.



TABLE 4. INNOVALLÉE INDUSTRIAL ZONE AS TRIAL SPACE

	<p>Innovallée zone groups together 380 companies and over 12 000 jobs. 70% of the jobs are in the digital technologies sector. There are also 2 incubators on the zone housing 30 start-ups.</p> <p>Innovallée association, with 7 staff members, is dedicated to managing the zone by providing grouped services, organising events, helping with human resources issues and communication. The association has direct and continues contacts with the employees of the zone</p>
	
<p>One of the 3 physical meetings organised with the Innovallée association</p>	

3 physical meetings organised with the management of the Innovallée zone. During the discussions we have steered the project in a direction in which we will mainly focus to put in place a real tool which could be used by the people working in the zone.

Several different uses have been identified:

- Plan of the zone with contact details and descriptions of all the companies (which already exists in a paper format)
- Information about the transport options, the location of cars available in the car-sharing programme “Cité Lib”, link with the “Métromobilité” app which gives real time information about public transport

- Information about building works/ traffic disruptions in the area
- Information about restaurant options, times and menus
- Information about sporting and cultural activities available
- Events, training, workshops, general communication that could interest employees in the area
- Etc.

In the middle of March, a survey will be carried out with the users of the site in order to validate the contents of the Inovallée app. A questionnaire has been prepared and will be used to consult the employees of the zone on March 21<sup>st</sup>. The questionnaire will contain questions about the expectations of the employees about their working environment such as transportation modes, information on restaurants, environmental information, cultural and sportive events information, etc. Section 10 gives the English translation of the survey,

An important question mark remains concerning the long-term prospects of this app – it is necessary to find financing that will allow the app to be maintained in terms of technical updates and content, on a long-term basis and not just for the duration of the BigClouT project.

There has been recently a discussion concerning the monitoring of another industrial estate “Espace Comboire”. This estate includes mostly shops and has a lot of daily traffic flow. An ambitious plan, involving the shops and businesses on the site as well as actors in the tech field, to create an internationally recognised cross-channel retail hub is currently under discussion and could be integrated into the Big ClouT project. This will be clarified in the coming months of the project.

**Industrial zone monitoring use case is at the project idea phase:**

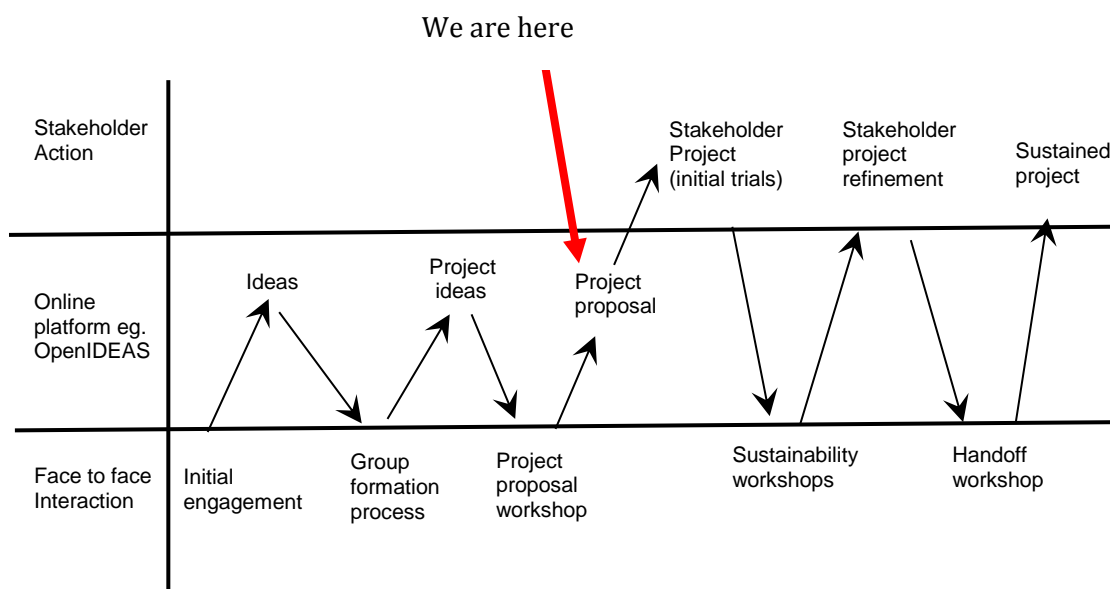


FIGURE 7. GRENOBLE ENGAGEMENT PROCESS: TRIAL 2 ON INDUSTRIAL ZONES

### 4.3 Fujisawa

#### 5. Trial 1: MinaRepo - Leveraging Participatory Sensing Technology for Daily City Operation

The purpose of this field trial is to make daily city operation smoothly and effectively, and to get big data for further trials in Fujisawa city. To achieve the goal, we designed and implemented a prototype of participatory sensing system for city operation – called MinaRepo. This section describes current status of MinaRepo and how the MinaRepo follows ethics guideline described in above document.

##### - Background

There is much information to be shared among city staffs for city operation. For example, in environmental section of Fujisawa city, information such as illegal garbage, dead animals, graffiti, road conditions are reported and shared by city staffs. However, this information is so far shared in an analog way – using telephone and FAX to report and share such information. The problems of such sharing way is followings – slow reporting turns to slow reaction of city management, and it is difficult to store and analyse such data because such information is basically managed as paper documents (see Figure 8. Fujisawa: Problem of current information sharing among city staffs Figure 8 below). Therefore, it is necessary to change analog and old way of sharing to more efficient way by using smart city technologies.

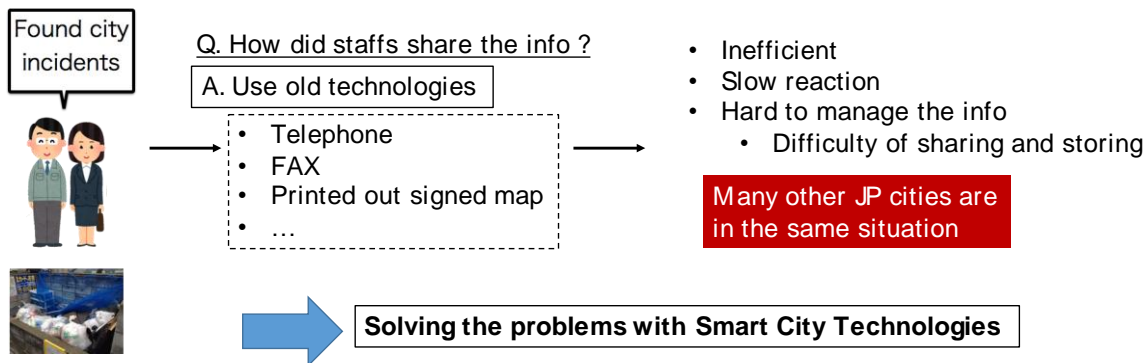


FIGURE 8. FUJUSAWA: PROBLEM OF CURRENT INFORMATION SHARING AMONG CITY STAFFS

##### - Trial Preparation

Following the guidelines proposed in this document, we had several meetings with stakeholders – Fujisawa city staffs and related company who works for Fujisawa city. Especially, we focus the environmental section of Fujisawa city to cope for the first field trial. Figure 9. shows our process of the field trial preparation. In the continuous meetings with stakeholders, we identified several details of problems and approach to solve the problem with considering ethics guideline.

**First Process:** finding problems of stakeholders, and proposing solutions

As the first steps of the preparation, we interviewed stakeholders to identify their problems for daily city operation. At the same time, we analyse the problems and provide ideas to not

only solve the problems but also provide additional advantages of the field trial. This interaction with stakeholders fulfilled the following ethics guideline.

- Maximize possible benefits and possible harms
- Offering adequate incentives

### **Second Process:** cycle of prototype creation and reviews by stakeholders

As the second process of the preparation, we designed and developed prototype of the field trial system, and let stakeholders review our prototype. With concrete working prototype, stakeholders can understand what kind of data can be collected and shared with easy usage. This process contributed to fulfill following ethics guideline.

- Ensuring right to withdraw
- Providing data protection and privacy
- Limiting disclosure

In addition to above guideline, we shared basic information of field trial and discuss its pros and cons in the meetings of all processes with stakeholders. This ensures to fulfill the following guideline:

- Obtaining voluntary informed consent
- Disclosing detriment arising from participation in research
- Avoid deception

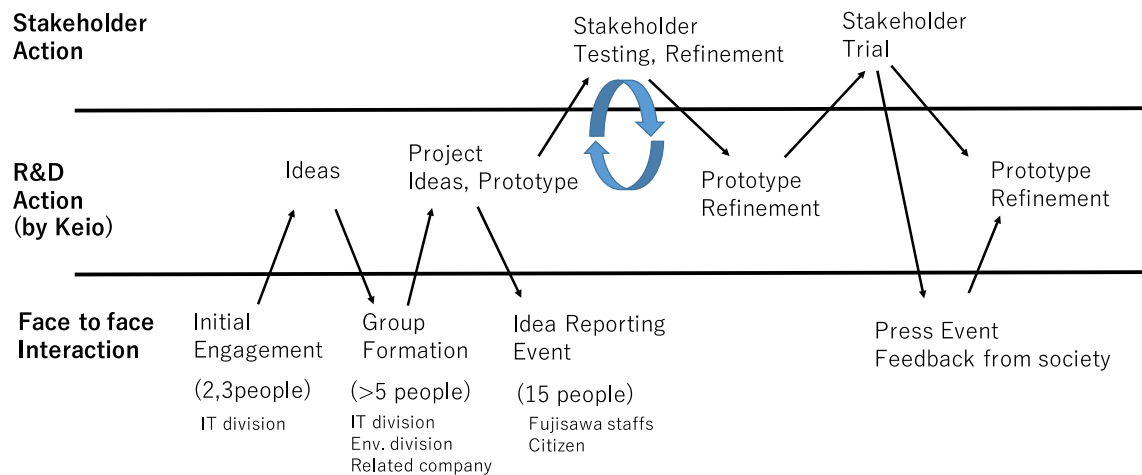


FIGURE 9. FUJISAWA: PROCESS OF FIELD TRIAL PREPARATION

Through above interaction with stakeholders, we already released first prototype of system, and start field trial with environmental section and other stakeholders from October 2016.



- MinaRepo as technical solution

Figure 10. shows overview of MinaRepo, a participatory sensing system for the field trial. MinaRepo are mainly composed by two components: client application and server-side application. As the client application, city staffs can use smartphone to collect city incidents by leveraging camera, GPS location information. Those collected information can be labelled easily by selecting the types of city incidents, and additional comment by city staffs. The information is shared by using BigClouT information dissemination platform, called SOXFire by Keio University. At the server-side application, collected information is visualized as map-based and list-based interface. As such, city staffs can review latest city information intuitively, and operate immediate action for city operation.

During 3 months' field trial so far, we collected more than 1,000 reports from city staffs and stakeholders. The collected information has correct labels for further analysis, we will start to analyse city's patterns with machine learning technique for more efficient city management.



FIGURE 10. FUJISAWA: OVERVIEW OF MINAREPO

- Dissemination

Through the active and periodic meetings with stakeholders, our purpose for the field trial and its advantages are highly appreciated among various section of Fujisawa city. The Mayor of Fujisawa city also appreciated the importance of our field trial, so that we had a press release event with the mayor of Fujisawa city. As a result of the press release, our trial has widely appeared in several medias (both offline and online medias) and also TV news at national broadcaster (see Figure 11. ). This smooth dissemination should be the result of the merit of following ethics guideline.



- Press release with a mayor of Fujisawa



- Online media, Newspapers
- More than 5 media



- National TV program

FIGURE 11. FUJISAWA: DISSEMINATION OF FIELD TRIAL

## 6. Trial 2: LokeMon- Location-aware Virtual Character as a Sensor

While target of MinaRepo is city staff, the target of second trial is to engage citizen for collecting massive city information by leveraging participatory sensing. This section describes the purpose of the trial and ethics plan for the field trial.

### - Background

Recent progress of mobile devices such as smartphones enables human to leverage their perceptive faculties as a part of sensing framework. This sensing framework, so called participatory sensing, distributes various sensing tasks (e.g., weather report, waiting time in a queue, traffic conditions etc.) to possible participants. By people sending a text, photo, sound data and so on, we can get subjective and qualitative data that it has been hard for physical sensors to get so far, such as an atmosphere of a place. It would make wide and high density sensing possible. However, there are some following problems of existing participatory sensing systems.

**1: Privacy issues** - Users who are participating in sensing send information related to each location spot. Therefore, it leads to exposure of both temporal and spatial privacy information.

**2: Information quality issues** - Focusing on the user when answers questions, generally one's real name or "user name" (pseudonym) have been used in existing crowd sensing systems. In such approaches, they cause problems about invasion of privacy and inflexible control quality of contributed data such as false information, ravage and slander battle.

**3: User motivation issues** - It is important to give incentive because it is a burden on users to get involved in crowd sensing. Without adequate incentive mechanisms, most users will not be willing to participate.



## Synopsis of the field trial

To solve the above problems we adopt an approach of personifying information transmitting entity, called “Lokemon”, which means location monster, characterized by each sensing space as a personifying target.

Lokemon is a brand new way of motivating citizen to participate crowd sensing without any privacy issues. Lokemon ask users to pretend themselves to be cute monsters associated to location spots when communicating with various people. As shown in Figure 12. , any users currently located near the spot can be a Lokemon and answer questions from other remote users. Remote users can ask questions related to the location wherever they are. “How many people are lining at the bus stop now?” “What's the mood of the restaurant right now?” The “Lokemon-ized” user will answer to questions asked to the Lokemon by the remote users.

Moreover, by designing monsters we can control quality of contributed data: In our daily life, we control our behavior to try to fit in. Our hypothesis is that, similarly on the Internet, we will change our behavior if we act as one of a group. That is why we use monsters, which we use each monster as kind of people's common understanding.

Pretending to be a Lokemon is a completely new and easy way to communicate with other people. Through Lokemon, we also aim at improving a chance for enjoying sensing as well as feeling attachment to a spot or a community by gamification such as character collection.

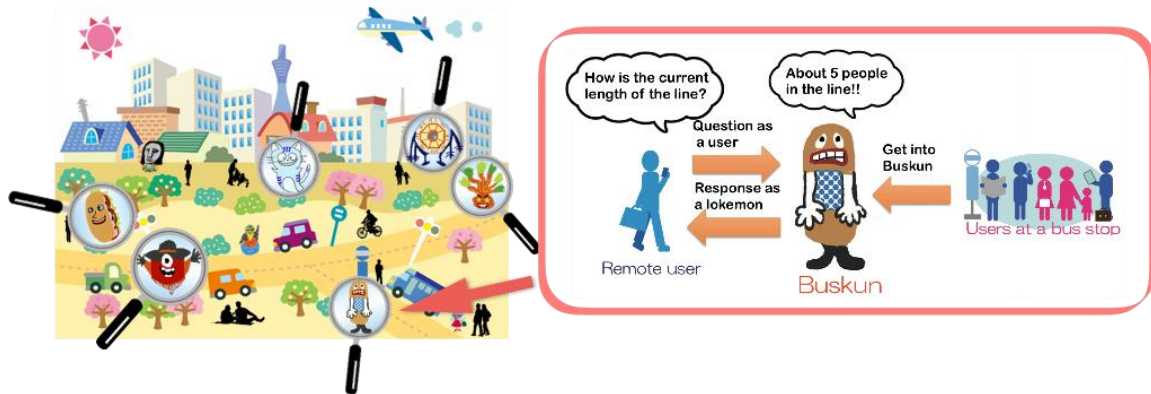


FIGURE 12. FUJISAWA: CONCEPT OF LOKEMON

### - Plan for stakeholder engagement

For the field trial of Lokemon, there are many things to be concerned. By following ethics guideline, we make our field trial plan clearer for creating effective and useful application.

- Maximize Possible Benefits and Minimize Possible Harms
- Disclosing Detriment Arising from Participation in Research

Firstly, Lokemon has many possibilities to collect data, and enhance attraction of the location. However, there are also some possible harm from LokeMon-ized location. To maximize the possible benefits and minimize possible harms, we will have several meeting with stakeholders.

- Obtaining Voluntary Informed Consent
- Ensuring Right to Withdraw
- Limiting Disclosure

Secondly, we will make clear of LokeMon's types of collecting information to stakeholders and citizen, and offering right to withdraw of the information or LokeMon-ized location also. Not only provide explanation in LokeMon application itself, we will provide details of information at our field trial website linked from Fujisawa city official web page.

- Offering Adequate Incentives

For deploying our application into Fujisawa city as sustainable application, we also would like to explain the benefit of field trial, and also ensure adequate incentives for stakeholders including citizens. This design of incentives must be discussed with Fujisawa city and location-owners.

#### 4.4 Tsukuba city

Tsukuba city is still in the planning stages for the trials it will run. Currently two trial cases have been identified:

##### 7. *Trial 1: Tourist trip planning*

- Problems/Needs
  - Improve visitors' satisfaction by recognizing their behaviour. Especially visitors who cannot achieve their purpose to visit to Tsukuba such as climbing Tsukuba Mountain because of the bad weather, or other interesting sightseeing spots around.
- Proposed solution
  - Provide real time and optimal tourism information or tour route by recognising their behaviour and predicting environmental information (weather, traffic).
  - In addition to the information, provide the city's attractive and convenient information to visitors in Tsukuba by push based delivery, so that visitors will stay longer to stop by sightseeing points or restaurants and revisit.



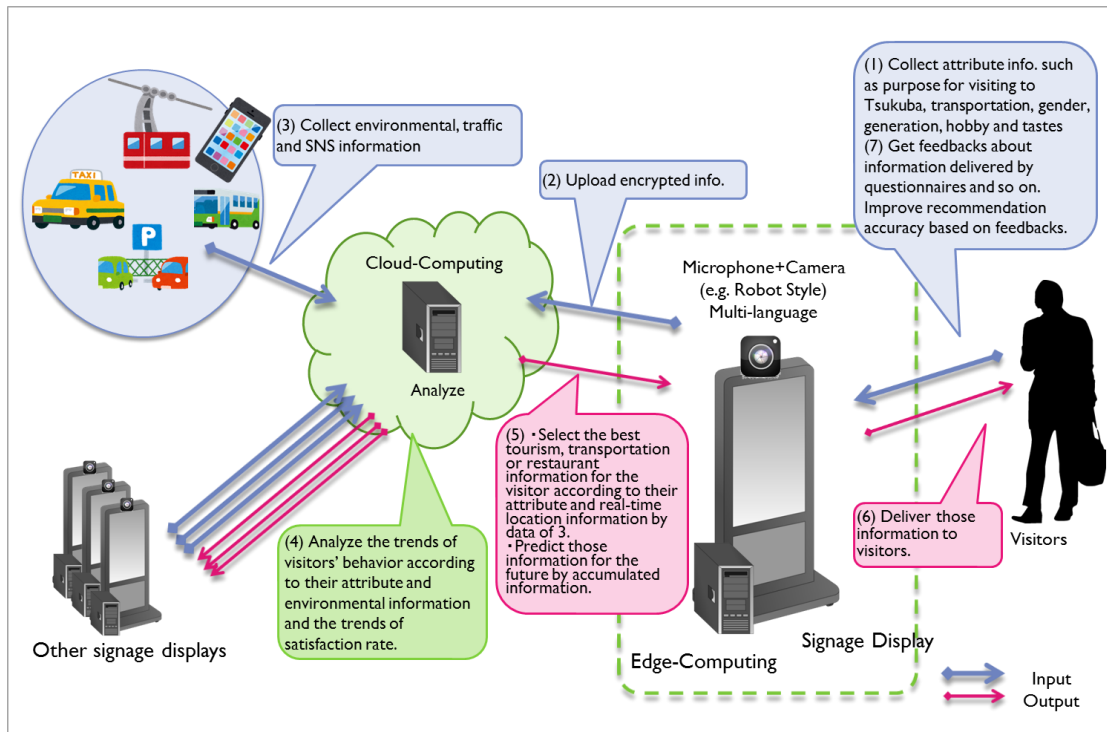


FIGURE 13. TSUKUBA USE CASE SCENARIO (TENTATIVE)

## 8. Trial 2: Personalized tourism information for overseas visitors

This use case is an extension of the first one, which is targeting overseas visitors.

- Problems/Needs
  - Effectively use visitors' information for tourism promotion by recognizing visitors' attribute such as their countries, numbers of visits, purposes (and so on).
- Providing information such as
  - Instruction to the local destination (for any vague needs such as "I want wine!").
  - Guide the instruction to the area including downtown, not a specific shop (e.g.: An area with many sushi restaurants).
  - Place (or direction) that is easy to get a taxi.
  - Emergency contacts (hospital, police).
  - Restroom map (degree of cleanliness).
  - Getting information to avoid dangerous areas.
  - Waiting time of restaurants and List up of the restaurants by waiting time.



- Proposed solution
  - Provide useful and suited information and city services to visitors from overseas by recognizing their attribute such as number of visits or their countries. In addition to that, provide multilingual concierge service for foreign visitors to Tsukuba to solve their needs and problems such as tourism, emergency, disaster prevention and transportation since Tsukuba city has many visitors from the other countries for international conferences.

Given the similarities with Grenoble city (which is also the twin town of Tsukuba), the trials of Grenoble have the potential of being replicated in Tsukuba.

### **Stakeholder engagement plan (for both trials)**

Field trial plan for Tsukuba city is now in the process of selecting stakeholders after with scenario idea according to the use case scenario. Seven meetings with IT division of Tsukuba city were held so far. In addition to F2F meeting, email and documentations are used for a main communication tool to clarify things to do, schedule and so on. Here are the points from guideline that are concerned during the discussions.

- Maximize Possible Benefits and Minimize Possible Harms
- Obtaining Voluntary Informed Consent
- Disclosing Detriment Arising from Participation in Research
- Limiting Disclosure

Tsukuba use case scenarios are targeting visitors to Tsukuba city. So candidates of stakeholders listed up are organizations in Tsukuba city who are related to tourism or international conference and would cooperate on the project.

Candidates of stakeholders:

- Tsukuba city IT division
- Tsukuba city tourism section
- Tsukuba city traffic section
- Tsukuba city international division
- Kanto railroad bus corporation
- Center for Computational Science
- Tsukuba Express
- Tsukuba Tourism Convention Association
- Tsukuba Science Tour Office
- Tsukuba Product Association

Here are the points from guideline that are concerned during the discussions.

- Maximize Possible Benefits and Minimize Possible Harms
- Obtaining Voluntary Informed Consent
- Ensuring Right to Withdraw
- Disclosing Detriment Arising from Participation in Research





- Limiting Disclosure
- Offering Adequate Incentives

## Stakeholder meetings

Here is the process of Tsukuba field trial discussion so far. As noted above, these meetings are for both trials as the participants are the same.

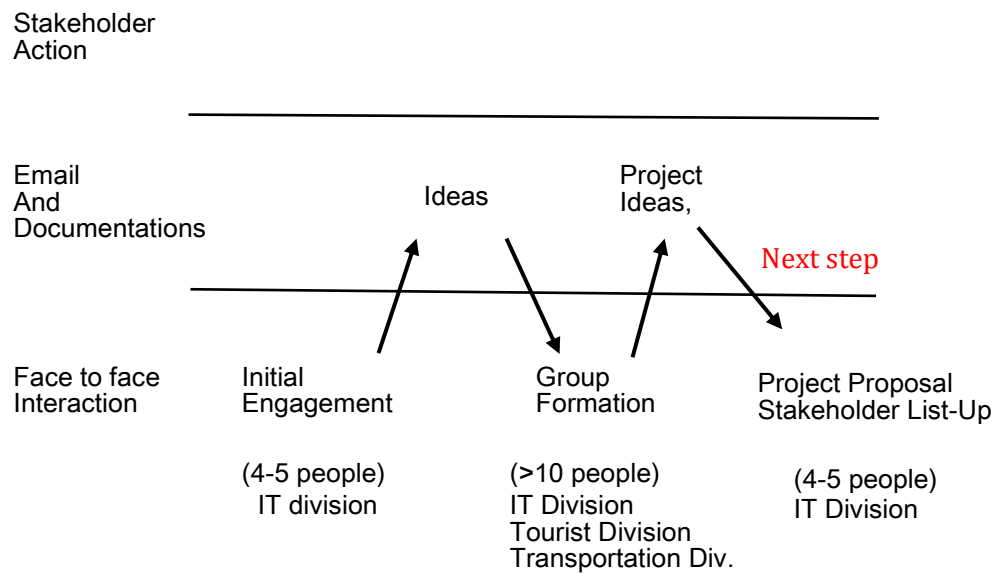


FIGURE 14. TSUKUBA CITY STAKEHOLDER ENGAGEMENT PROCESS

## Summary/Conclusion

---

This document is the first of two deliverables describing procedures for the trials including participant recruitment, data gathering and ethics. The core aim of the document is to develop some initial guidelines to help guide the initial planning phases of the trials. Since the use case development is still underway in WP1 and so trials plans are, at this stage, still early stage, this document has focused on a set of generic guidelines for how trial partners should address some of the key issues such as Ethics, stakeholder engagement, data management, trial evaluation etc.

Following the guidelines, each trial partner has provided information about how they followed those guidelines. Since we are at the early phase of trial planning, most of the trials are at the stakeholder engagement and project idea phase. As the use cases are finalised, and the initial trial planning begins, then D4.2 will update this document with details on trial specific issues.

Although the BigClouT project has already identified the initial city partners, it might be worthwhile including a section in the guidelines on how best to identify potential future city partners. These could join later in the project or for follow-on projects, by taking into consideration criteria such as population distribution, geography, city issues, collaborative framework for FT, data to be provided, infrastructure, sensors, etc.

# APPENDIX 1: BIGCLOUT ETHICS PRIMER

### ○ Introduction

This ethics primer is intended for BigClouT project members and provides an introduction to ethical issues in research involving studies and experiments with humans (e.g. for the city based field trials). It also outlines the ethical assessment process adopted by the project.

Ethical guidelines on human subject experiments provide guidance on how to properly treat human subjects and their data. While computer science typically does not concern itself with experiments in the same way that, say, medicine or psychology does (i.e., directly experimenting *upon* individuals), many of our studies will ultimately collect and store information that may or may not be associated with individuals. This information may inconvenience or even threaten the physical and psychological well-being of test subjects, should it be used for unforeseen purposes or be shared with unintended recipients. For example, if usage measurements of how an employee uses work-related software would leak, his or her employer might learn of substandard performance, hidden attitudes, or detect obvious errors in conduct. By following the ethical principles set forth in this handbook, work in PD-Net should minimize potential threats to human subjects stemming from project-related user studies.

Use the table of contents below to locate relevant information in this document. Of particular interest is the PD-Net Ethical Assessment Process and the fundamental Ethical Principles that research within the project follows.

### ○ Ethics Background

There is a vast amount of literature on ethics, ranging back to Plato and Aristotle. However, Mackay [2] points out that modern professional ethics codes in the field of computer science are mostly influenced by two relatively modern perspectives: Mill's *utilitarianism*, and Kant's *deontology*.

**Deontological Ethics** (deon is Greek for “duty”, logos means “science”) infers moral obligations from the characteristics of a certain action, without regard for its consequences. Thus, an action that is morally good might still have serious negative consequences. One of the most prominent proponents of deontological ethics was Immanuel Kant, who formulated a “Golden Rule,” his categorical imperative, for determining the morality of an action:

Act only on that maxim whereby thou canst at the same time will that it should become a universal law.



Note that Kant's Golden Rule is not just a reformulation of the Biblical Golden Rule "All things whatsoever you would have men do unto you, do you ever so to them,"<sup>18</sup> as it explicitly requires moral principles to be universally applicable, to become a universal law of nature. It is thus a categorical imperative, not just a hypothetical imperative, which only applies conditionally (e.g., only if you want people to do A to you, do A to them).

**Teleological Ethics** on the other hand, derives morality not from the intentions, but from the consequences of actions, e.g., whether it leads to "desirable" effects (telos is Greek for "goal" or "end"). In the context of a research study, this could for example be taken to allow for the deception of study subjects if it would lead to more relevant results, while not negatively affecting the subjects. The exact nature of these effects, i.e., what exactly constitutes a desirable effect, is of course no less debated than the moral truths of the deontologists.

The most prominent teleological ethical theory is that of *utilitarianism*. Its main proponents were the late 18th- and 19th-century English philosophers Jeremy Bentham and John Stuart Mill. John S. Mill was an ardent proponent of the freedom of individuals from government interference. In his 1859 essay *On Liberty*, Mill proposed as the proper balance between individual liberty and governmental authority the "harm principle:"

[T]he only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others [241].

**Professional Codes** from, e.g., the *National Society of Professional Engineers* or the *Association for Computing Machinery (ACM)*<sup>19</sup> try to give practical guidance to their members through publishing a Code of Ethics for their respective fields (see Table 1 on the next page). They are typically a mixture of deontological ("be honest and trustworthy") and teleological ("contribute to society and human well-being") approaches. Table 1 below lists the "moral imperatives" from the ACM Code of ethics.

Professional associations such as the ACM Codes of Ethics to provide practical ethical guidance to their members. They are typically a mixture of deontological and teleological approaches.

TABLE 5. ACM CODE OF ETHICS – MORAL IMPERATIVES (EXCERPT FROM [WWW.ACM.ORG/CONSTITUTION/CODE.HTML](http://WWW.ACM.ORG/CONSTITUTION/CODE.HTML)).

- **Contribute to society and human well-being.** When designing or implementing systems, computing professionals must attempt to ensure that the products of their efforts will be used in socially responsible ways, will meet social needs, and will avoid harmful effects to health and welfare
- **Avoid harm to others.** To minimize the possibility of indirectly harming others, computing professionals must minimize malfunctions by following generally accepted standards for system design and testing. Furthermore, it is often necessary to assess

<sup>18</sup> See (Matthew 7:12)

<sup>19</sup> See [www.nspe.org](http://www.nspe.org) and [www.acm.org](http://www.acm.org), respectively.



the social consequences of systems to project the likelihood of any serious harm to others.

- **Be honest and trustworthy.** Honesty is an essential component of trust. Without trust an organization cannot function effectively. The honest computing professional will not make deliberately false or deceptive claims about a system or system design, but will instead provide full disclosure of all pertinent system limitations and problems.
- **Be fair and take action not to discriminate.** Discrimination on the basis of race, sex, religion, age, disability, national origin, or other such factors is an explicit violation of ACM policy and will not be tolerated.
- **Honor property rights including copyrights and patent.** Violation of copyrights, patents, trade secrets and the terms of license agreements is prohibited by law in most circumstances. Even when software is not so protected, such violations are contrary to professional behavior.
- **Give proper credit for intellectual property.** Computing professionals are obligated to protect the integrity of intellectual property. Specifically, one must not take credit for other's ideas or work, even in cases where the work has not been explicitly protected by copyright, patent, etc.
- **Respect the privacy of others.** This imperative implies that only the necessary amount of personal information be collected in a system, that retention and disposal periods for that information be clearly defined and enforced, and that personal information gathered for a specific purpose not be used for other purposes without consent of the individual(s).
- **Honor confidentiality.** The principle of honesty extends to issues of confidentiality of information whenever one has made an explicit promise to honor confidentiality or, implicitly, when private information not directly related to the performance of one's duties becomes available.

While professional code offer a more practical approach to ethical theory, they still fail to prescribe practical steps to be taken when planning and conducting a user study or performing an observation in the field. As a consequence, this primer prescribes a concrete set of principles to follow when designing, conducting, and analyzing field trials, user studies, and other experiments involving human study subjects. However, the above principles (see Table 1) should be taken as overarching principles that should guide all actions of researchers during all stages of the project.

### ○ Ethical Assessment Process

The list below outlines the steps that all user studies, observations, field trials, and interviews should follow.

1. Fill out *BigClouT Ethical Worksheet* **prior** to planned begin of study
  - a. Prepare Consent Form if needed (see **Guide to Informed Consent**)



2. If needed, seek local approval from local *Institutional Review Board* (IRB) and regulatory bodies
  - a. If IRB assessment required, prepare necessary documents and submit
  - b. If regulatory approval required, prepare necessary documents and submit
  - c. Incorporate any feedback, resubmit if necessary
3. If new IRB approval and/or ERB assessment, submit results to project Coordinator prior to planned begin of study

Proceed with the planned study only if all relevant studies have been approved by the ERB and all local IRB issues (if applicable) have been addressed.

### ○ Ethical Principles

Studies and observations in BigCouT follow the 10 basic principles outlined below.<sup>20</sup> Individual types of studies and how these principles apply are described in a separate set of guides, the so-called “Study Process Templates” (SPT). There is, e.g., the *Procedures Public Trials* or the *PD-Net Guide to Volunteer Studies*. There are also guides to general principles such as the *Guide to Obtaining Informed Consent* and the *Guide to Secure Data Storage*.

See also the section describing the Ethical Assessment Process on page for details on how these guidelines are used in practice, in particular the role of the Ethical Review Board.

#### ▪ *Overview*

1. Maximize Possible Benefits and Minimize Possible Harms
2. Obtaining Voluntary Informed Consent
3. Ensuring Right to Withdraw
4. Disclosing Detriment Arising from Participation in Research
5. Providing Data Protection and Privacy
6. Limiting Disclosure
7. Following Minimal Intrusion Principle
8. Offering Adequate Incentives
9. Special Provisions for Experiments Involving Children and other Vulnerable People
10. Avoiding Deception

#### ▪ *Definitions*

- The **participants** in research may be the active or passive subjects of such processes as observation, inquiry, experiment or test. They may be collaborators or colleagues in the research process or they may simply be part of the context e.g. where passers-by are part of the context but not the subjects of an on-campus study on (adapted from: [1])

---

<sup>20</sup> Principles taken from [1],





- **Personally identifiable information (PII)** is data that can – with reasonable effort – be *linked* to an individual. Linking requires that a person can be distinguished – with more or less certainty – from all other persons. Examples of PII are, e.g., the first and last name; a home or other physical address including street name and name of city or town; an email address; and a telephone number. Also a seemingly random identifier, such as an IP address or a **Bluetooth MAC** address, can become PII, if these can in turn be linked to any other PII such as a physical address. The information necessary to perform the linking may be in another (local) database, in the public domain (e.g., white pages), or can be obtained with reasonable effort from another (remote) database. Note that much depends on the size of the potential user population: if an experiment involves only members of a particular department, then knowing even a relatively innocuous data point such as “height” may already constitute PII.
- **Anonymisation, pseudonymisation and identifiability** (from [4]): ‘Anonymous’ often means data which does not identify an individual; ‘anonymised’ means data which has been rendered anonymous; ‘pseudonymised’ and ‘coded’ means data where obvious identifiers (e.g. names and addresses) have been replaced with indirect identifiers (e.g. numbers) in the main data set and the indirect identifiers are then held with the obvious identifiers in a separate data set (known as the ‘key’). The key term underlying all of the above definitions in the context of European data protection law is the ‘identifiability’ of an individual from the data. For European data protection law to bind research on personal and sensitive personal data one must ask: is the individual identified either immediately from the data or when that data are combined with other data in the hands of another person. This combination extends only to *reasonably foreseeable* linkings of data. Therefore, data which is gathered anonymously without any identifiers will be outside the scope of European data protection law; data which is pseudonymised or coded will be within the scope of the law as it is possible to reintroduce the two separate data sets and identify individuals; data which was gathered as identifiable data and then anonymised is subject to the data protection legislation when it contains identifiable data (most importantly at the point of gathering the data, requiring the disclosure by the researcher to the research participant of information including the purpose of the processing and contact details).
- **Sensitive information** (from [5]). Sensitive data include data “revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life” (Article 10 of Regulation 45/2001; Article 8 of Directive 95/46/EC). The processing of such information is in principle prohibited, except with the explicit consent of the data subject. Processing and storing sensitive information requires a significantly increased level of security (adequacy principle – see the *Guide to Secure Storage*).



## ■ Principles Explained

### **Do No Harm**

One of the most fundamental principles of ethical research is the harm principle. It stems from the 1979 Belmont report [7], which identifies basic ethical principles that should underlie all types of behavioural research involving human subjects. The report lists three basic principles: Autonomy, Beneficence, and Justice. The last principle (justice) simply states that both the risks and the benefits of research should be distributed equally. This may become relevant, e.g., when selecting participants. The Beneficence principle then obligates the researcher to secure the well-being of all selected participants, i.e., to “maximize possible benefits and minimize possible harms.” In practice, each trial must explicitly list possible risks to participants and explain how their adverse effects are mitigated (e.g., involuntary disclosure of personal data). This information must also be communicated to potential participants *prior to enrolment*, which supports the first of the three principles, Autonomy, and which is explained in the next item below.

### **Informed Consent**

The first of the three Belmont principles, Autonomy, stipulates that each participant should be given the respect, time, and opportunity necessary to make his or her own decisions, in order to “make sure they undertake activities freely and with awareness of possible adverse consequence.” This principle is implemented through a so-called *Informed Consent* process. Consent to participate in research is a process, rather than an event [3]. Researchers should plan for and outline how consent is initially obtained and how it is reviewed throughout the study. Also, in order to give meaningful consent, participants must understand the goals of the research, the study/experiment to be performed, the data that is collected, and the uses of this data. While written consent is preferred, oral consent might be more appropriate in some situations, e.g., walk-up interviews on public places. Note that the principle of informed consent includes the principle of disclosing detriment and ensuring the right to withdraw (see below). A detailed description of the informed consent process can be found in the *Guide for Obtaining Informed Consent*.

### **Right to Withdrawal**

As part of the informed consent procedure outlined above, potential participants must be informed of the right to refuse to participate in the study, and that they can withdraw their consent to participate at any time without reprisal [8]. This should include the right to withdraw retrospectively, i.e., in the light of experience of the investigation, or as a result of debriefing. Any such request should result in the destruction of the participants own data. Note that the unconditional right to withdrawal has been questioned repeatedly by researchers [9][10], as it might not only prevent researchers from encouraging study subjects to continue participation, but also lead to a prior dismissal of potential drop-outs from the pool of study participants. As described in the informed consent principle above, researchers should seek an on-going dialog with participants throughout the study, in order to ensure that while proper information is given regarding withdrawal, participants are adequately encouraged to continue.



## Disclosing Detriment

Disclosing possible detriments arising from participation is an integral part of obtaining informed consent from potential study participants. As part of the *Ethical Worksheet*, BigClouT researchers will have to explicitly list the risks to study subjects that could stem from participation. This information must explicitly be disclosed to participants in the informed consent documents.

## Privacy

Data collection, storage, and use of personally identifiable information (PII) in general must follow the EU legal framework (i.e., the 1995/46/EC Directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data), as well as individual national legislation on data protection. Legal requirements on processing PII typically comprise the following aspects, roughly modelled after the OECD Guidelines of 1980 [11]:

Notice—data subjects should be given notice when their data is being collected;

Purpose—data should only be used for the purpose stated and not for any other purposes;

Consent—data should not be disclosed without the data subject’s consent;

Security—collected data should be kept secure from any potential abuses;

Disclosure—data subjects should be informed as to who is collecting their data;

Access—data subjects should be allowed to access their data and make corrections to any inaccurate data; and

Accountability—data subjects should have a method available to them to hold data collectors accountable for following the above principles

The principles of notice, consent, and disclosure are covered by the project’s informed consent process (see the *Guide for Obtaining Informed Consent*). The purpose principle is addressed through the rigorous use of *Ethical Worksheets* explicitly describing the individual study goals. Security is ensured by following the *Guide for Secure Storage*, and by explicitly stating the data storage and processing circumstances for each individual study in the *Ethical Worksheet*. Access and accountability are given by including access and inspection methods into the information sheet administered as part of the informed consent (e.g., the contact information of the Principle Investigator responsible for the study, as well as the contact details of the Ethical Review Board and – if applicable – any local institutional review board).

In addition, European data privacy laws require the proportionality/data minimization principle – see the “Minimal Intrusion Principle” listed below.

## Limiting Disclosure

Personally identifiable data collected as part of a study will only be made available to researchers directly involved with the research, on a “need to know” basis. The *Ethical*



*Worksheet* asks researchers to explicitly list all members of the consortium who will participate in the administration of a particular study, as well as any outside researchers.

### **Minimal Intrusion Principle**

The principle of “minimal intrusion” or “data minimization” means that a one should limit the collection of personal information to what is directly relevant and necessary to accomplish a specified purpose. Data should also be retained only for as long as is necessary to fulfill that purpose. This principle derives from Article 6.1(b) and (c) of Directive 95/46/EC , which provide that personal data must be “collected for specified, explicit and legitimate purposes” and must be “adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed”. This also often referred to as the proportionality principle, i.e., “whether the means employed by the measure to be evaluated are suitable and reasonably likely to achieve its objectives” [12]. In the context of BigClouT research, this means that researchers should only collect information that provides the data required to address current research questions. The *Ethical Worksheet* asks researchers to explicitly state what data should be collected in a study (question 3.6), and requires a justification on the suitability of this data (and the study methods employed) for the stated research purpose (question 3.4).

### **Adequate Incentives**

The use of incentives to recruit and retain research subjects is typically rather innocuous. However, when study subjects are in a dependency relationship with the researcher (e.g., students in a course), or when the participant’s aversion to the study may be strong (e.g., high risk, degrading research, aversion on principle) [13]. In the context of this project, incentives should only comprise adequate compensation for the participants’ time and efforts, e.g., for coming to weekly meetings. If necessary, small incentives in the form of vouchers for online stores such as Amazon.com or iTunes may be given, or all participants enter into a drawing of a small set of prizes.

### **Avoiding Deception**

The act of deceiving study participants is often used in psychological experimentation, in order to ensure that study subjects do not inadvertently change their “natural” behavior in order to please the experimenter and/or to make themselves appear in a better light. Infamous examples of deceptive experiments are the Stanley Milgram experiment of 1974 [14] or Zimbardo’s Stanford Prison experiment of 1971 [15]. Today, as a general rule, deception is not acceptable when doing research with humans. Using deception jeopardizes the integrity of the informed consent process and can potentially harm participants. Should the use of deception become necessary in any field study within the scope of the project, explicit approval by an institutional review board (if available) should be sought. If possible, the ethical review board should also be asked for comments (e.g., using the *Ethical Worksheet* for describing the experiment).

The use of deception requires an in-depth justification of why the deception is necessary for the study, and the steps taken to safeguard study participants.



## Vulnerable Participants

Vulnerable participants are those unable to give their unambiguous informed consent, such as children, people with cognitive disorders, or those with cultural or intellectual difficulties in speech and understanding. Experiments seeking the participation of vulnerable participants must be justified and get prior approval of an appropriate institutional review board, and, if possible, obtain feedback from the ethical review board. Informed consent must be obtained from parents or other appropriate legal guardians.

### ○ Bibliography

- [1] British Educational Research Association (bera): **Revised Ethical Guidelines for Educational Research**, 2004. Available from <http://www.bera.ac.uk/files/guidelines/ethica1.pdf>
- [2] Mackay, Wendy E.: **Ethics, Lies and Videotape...** In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Denver, Colorado, United States, May 07 - 11, 1995. ACM Press/Addison-Wesley Publishing Co., New York, NY, pp.138-145. DOI: [10.1145/223904.223922](https://doi.org/10.1145/223904.223922)
- [3] UofT Research Ethics Board (REB): **Guide for Informed Consent**. University of Toronto, Canada, April 2010.
- [4] Caroline Gans-Combe (ed.): **Data Protection and Privacy Ethical Guidelines** (Version 5). European Commission, September 18, 2009
- [5] The European Data Protection Supervisor (EDPS): Data Protection Glossary. The EDPS Website. See <http://www.edps.europa.eu/EDPSWEB/edps/EDPS/Dataprotection/Glossary>
- [6] Eléonore Pauwels: **Ethics for Researchers. Facilitating Research Excellence in FP7**. See <ftp://ftp.cordis.europa.eu/pub/fp7/docs/ethics-for-researchers.pdf>
- [7] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research: **The Belmont Report - Ethical Principles and Guidelines for the Protection of Human Subjects of Research**. U.S. Department of Health, Education, and Welfare, 1979. See <http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.htm>
- [8] World Medical Association: **WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects**. 18th WMA General Assembly, Helsinki, Finland, 1964. See <http://www.wma.net/en/30publications/10policies/b3/index.html>
- [9] Sarah J. L. Edwards: **Research participation and the right to withdraw**. Bioethics. 2005 April; 19(2):112-30. DOI: [10.1111/j.1467-8519.2005.00429.x](https://doi.org/10.1111/j.1467-8519.2005.00429.x)<http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8519.2005.00429.x/full>
- [10] Stefan Eriksson and Gert Helgesson: **Potential harms, anonymization, and the right to withdraw consent to biobank research**. European Journal of Human Genetics



- (2005) 13, 1071–1076. doi:10.1038/sj.ejhg.5201458; published online 29 June 2005.  
See <http://www.nature.com/ejhg/journal/v13/n9/abs/5201458a.html>
- [11] Organization for Economic Co-Operation and Development: **Guidelines on the Protection of Privacy and Transborder Flows of Personal Data**. See [http://www.oecd.org/document/18/0,2340,en\\_2649\\_34255\\_1815186\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html)
- [12] Christopher Kuner: **Proportionality in European Data Protection Law And Its Importance for Data Processing by Companies**. Privacy & Security Law Report, Vol. 07, No. 44, 11/10/2008, pp. 1615ff.
- [13] Ruth W. Grant and Jeremy Sugarman: **Ethics in Human Subjects Research: Do Incentives Matter?** Journal of Medicine and Philosophy, 29(6): 717–738, 2004. See <http://www.waisman.wisc.edu/EVENTS/ethics/sprin06-sem2-incentives-compensation.pdf>
- [14] Stanley Milgram: **Obedience to Authority**. New York: Harper & Row, 1974
- [15] Philip G. Zimbardo: **The power and pathology of imprisonment**. Congressional Record. (Serial No. 15, 1971-10-25). Hearings before Subcommittee No. 3, of the Committee on the Judiciary, House of Representatives, Ninety-Second Congress. Washington, DC: U.S. Government Printing Office.





# PROCEDURES FOR REAL-WORLD TRIALS

## A BIGCLOUT REAL-WORLD TRIAL PROCESS TEMPLATE

BigClouT will conduct a number of studies to explore how members of the public interact with BigClouT systems in the wild, i.e., in a public or semi-public setting. In this document we describe our operating guidelines for such studies. All such studies will follow local guidelines and additionally conform to the rules and procedures set out in this document.

### ○ Definition of Real-World Trials

Demonstrator trials allow researchers to investigate actual use of BigClouT systems in people's everyday lives. Trials may not necessarily involve an explicit user recruitment step but instead observe the interaction of a potentially large set of opportunistic users that happen to encounter a specifically fielded system by chance. However, such trials might trigger (or run concurrently to) a volunteer study, where people can personally register as volunteer participants. Public trials can be conducted in a range of locations including public places, offices, or universities. In all cases explicit permission of the corresponding authorities (e.g., university or city council) will be obtained.

### ○ Nature of Studies

Real world trials in BigClouT typically involve observing participants interacting with digital services possibly via mobile, wearable or situated technologies. Observations of such interactions can be direct, i.e., using a researcher on-site taking written notes, making audio-recordings, or capturing video footage (in accordance with local rules and regulations), or indirect, i.e., based on interaction logs, click-stream information, recorded contextual information such as location or sensor data, and network traffic data. Information about both direct and indirect observations would be available at or through the installation, e.g., in the form of an information brochure or an online notice. BigClouT trials never involve activities that place members of the public at risk of bodily harm. The experimental procedure for the studies varies according to the nature of the research but will always be overseen by a local PI.



### ○ Participants

Careful thought will always be given to the location of our trials in order to ensure that applications and content are appropriate for those who become part of the study. No vulnerable participants (e.g., children or people with cognitive disorders) will be targeted. A brief description of BigClouT, as well as a comprehensive description of the study and contact details on appropriate headed paper will be available to members of the public on request.

### ○ Data

Data collection, storage, and use of personally identifiable information (PII) in general will follow the EU legal framework, as well as individual national legislation on data protection. The project has produced information sheets to help ensure researchers are familiar with these requirements. Where local legislation requires it information will be posted notifying members of the public about the ongoing study. This information might, e.g., describe the data being collected, the purpose of this data collection (i.e., the particular research aspect under investigation), the recipients of this data (i.e., only project researchers), the use of the data (i.e., that data will only be published in anonymous form), the name and full contact information of the PI responsible for the data collection, access information (i.e., how to get a copy of the data collected about oneself), and how long data will be retained.

All collected data will use *pseudonymous identifiers* for all subjects whenever possible, or use appropriate secure storage procedures to safeguard PII where such anonymisation at collection time is not possible (such as photographs). PII collected through indirect observations (e.g., clickstream data, form filling data) that pertain to non-recruited participants will be deleted or anonymized prior to archival. We have produced an information sheet to provide researchers with guidance on secure data storage. Collected data will normally be deleted at most 3 months after BigClouT ends, though earlier times are possible, unless clearly stated otherwise. All publications will only use fully anonymized data when reporting qualitative and quantitative data.

### ○ Documentation

Details of each public trial will be recorded on an ethics worksheet and these worksheets will be uploaded to and stored on the project's secure datastore.



## GUIDE TO OBTAINING INFORMED CONSENT

A BIGCLOUT STUDY DESIGN BRIEF

Consent to participate in a research study should be understood as a process rather than an event [1]. Researchers should plan for and articulate the steps by which consent is **initially obtained** and the steps by which it is **reviewed** throughout the study. In order for participants to give meaningful consent, they should be able to understand the **intent of the research**, be clear about **what they are being asked to do** and if any risks are involved, and know **how their information will be used**.

### ○ Enabling Informed Consent

Consent may be documented in many ways. Oral or implied consent are as legitimate as written consent, and in some contexts may even be more appropriate. The key idea is to go over the information verbally and **document the process** of gaining consent in field notes so as to **leave a written trail**. Even with oral consent, however, is still reasonable to leave written material with the participant (e.g., an information letter).

Consent must always be in **language that is understandable** and not legalistic or too scientific, and the consent process should make **room for questions**, as appropriate to the research context. When a written-and-signed approach to consent is used, the **information letter** and **consent form** are best presented as one document.

**The Information Letter** should begin with an invitation to potential participants and should explain why they have been asked to participate. The body should provide a brief (i.e., a paragraph or two), plain-language **description of the BigClouT project**, the particular **study that participation is sought for**, and the **nature of participation**. An explanation of how key ethics issues—such as **consent and confidentiality**—will be handled, along with a discussion of **risks and benefits**, and compensation if any, should follow. The information letter should be written as if it was being sent **from the researcher to the participant**, that is, in the 2nd person. It should include an introduction of the researchers and their affiliations.

**The Consent Form** should include a brief summary of what will happen from the participant's perspective—without redundancy. It should note that the study has been **explained** to the participant, and the participant has had a chance to have his or her **questions answered**. The basic elements of consent, bulleted below, should be taken into account relevant regardless of process – whether written in hard copy, via e-mail, on the web, or presented verbally in person or over the phone. However, not all items are appropriate for



all protocols, and some additional items may be useful on a case by case basis. Appendix A contains an example form.

#### ○ General Points

- Use letterhead of the department/organization undertaking the research
- The language level is appropriate to the age and reading level of the participant population
- Affiliation and contact information for the investigators and (where appropriate) research coordinator is included
- Participants are given a copy of the information letter to keep for their own

#### ○ Introductory Remarks

- Introductory information on BigClouT
- An invitation to participate should be worded in a **professional and respectful** manner
- The **time commitment and the location** of where the study will be conducted should be clarified.
- The reason why the potential participant is being approached should be explained, and a list of relevant **inclusion and exclusion criteria**, should be provided.
- If relevant, the **number of participants** who will be involved should be mentioned (e.g., if this could affect confidentiality - see below)

#### ○ Conditions for participating

- There must be an explicit statement that the individual's participation is **voluntary**, and that he/she may refuse to participate, may **withdraw at any time**, and may decline to answer any question or participate in any parts of the tasks – all **without negative consequences**
- Any **conditions on withdrawal** of data if the participant chooses to withdraw from the study should be clarified (e.g., if data are anonymized or de-linked, they cannot be withdrawn; similarly, it is almost impossible to withdraw data from a focus group discussion)
- Information regarding use of **audio and video recordings** (including potential use for presentation purposes) should be broken out as separate options, to which participants can consent (or not).

#### ○ Risks/Benefits

- Reasonably **foreseeable risks, harms or inconveniences**, and how they will be managed should be clearly explained in lay terms



- **Potential benefits**—including information that there is no direct benefit—should be mentioned, as appropriate
- Information about any **payment or compensation** for participation or expense reimbursement should be mentioned (but not over-emphasized)

#### ○ Access to information, confidentiality, and publication of results

- Information regarding **who will have access to the data** should be clarified, including sharing data among project partners
- Information regarding **retention and disposition** (e.g., deletion) of the data during and after completion of research is relevant. Note: destruction of data is not the only acceptable method of disposition. Methods will depend on the identifiability, sensitivity, and richness of the data.
- If applicable, different **degrees of confidentiality** should be presented as options
- The **procedures for maintaining confidentiality** should be described, if relevant—e.g., use of study-specific ID numbers, pseudonyms, generic descriptors, composites, or aggregates
- Any foreseeable **limits to confidentiality**—e.g., for participation in focus groups, research involving key informants or duty to report—should be mentioned
- The researcher's **intent to publish** or make public presentations based on the research should be made explicit
- A **summary of the research results**, and a mechanism to provide the summary, should be offered

See also the definition of Informed Consent given in the Directive 2001/20/EC [2] relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human, which also applies to other (i.e., non-clinical) EU-funded research. See also Directive 95/46/EC [3] for general guidelines regarding the role of informed consent.

#### ○ Bibliography

- [1] Informed Consent Guide, University of Toronto, Research Ethics Board (REB), April 2010
- [2] Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. OJ L 121, 1.5.2001.
- [3] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data OJ 23 November 1995 No L. 281 pp. 0031-0050.



# GUIDE TO SECURE DATA STORAGE

---

## A BIGCLOUD TRIAL TEMPLATE

Data collected in the context of BigClouT should be protected from loss, corruption, and unauthorized access. Of particular importance is the confidentiality of **personally identifiable information (PII)**.

### ○ Secure Data Storage – Core Concepts

The concept of personally identifiable information (PII) expresses the fact that we want to prevent collected information to be *linked* to an individual. Linking requires that a person can be distinguished – with more or less certainty – from all other persons. PII is what allows us to perform this linking. Examples of PII are, e.g., the first and last name; a home or other physical address including street name and name of city or town; an email address; and a telephone number. Also a seemingly random identifier, such as an IP address or a **Bluetooth MAC** address, can become PII, if these can in turn be linked to any other PII such as a physical address.<sup>21</sup>

When storing data related to user studies and experiments, we thus need to either protect access to all and any data collected from participants, or alternatively ensure that this data remains *unlinkable* to the participants, by separating and safeguarding all PII. Note, however, that the concept of privacy is much broader than just ensuring the secure storage of PII. While this document describes only the secure data storage principles adopted in BigClouT, please refer to the Ethics Primer for a comprehensive overview of privacy requirements in the project.

### ▪ *Pseudonymous Identifiers*

A typical approach in user studies is to use *pseudonymous identifiers* for all participants, and to store the data that links the participants' identity to those identifiers separately from the collected data (also called “coded identification”). The researcher maintains a list that contains a link between the subject's name and a random code number or pseudonym, and then uses only the code number/pseudonym to mark the data and responses from that

---

<sup>21</sup> For example, many legislators today consider IP addresses as PII, since providers (e.g., ISPs) can typically associate an IP address with a physical address for any point in time, even if these are being repeatedly changed (e.g., for DSL access).



subject. The list that contains the coded link should be kept secure, and separate from the data and responses – if stored in electronic form it must be encrypted, if on paper it should be kept in a safe place (e.g., locked office). Obviously, care must be taken so that the remaining data does not involuntarily contain other PII, such as the above-mentioned IP addresses or Bluetooth MAC addresses, or demographic information from a small group that implicitly allows identification. Also, some study data such as video and audio recordings might inherently allow for the identification of a participant, and thus may need to be protected in their entirety.

If numeric identifiers such as IP addresses or Bluetooth MAC addresses are to be stored in system logs, these should ideally be anonymized before storage, so that the logs themselves can be stored without additional security precautions (i.e., without having any PII embedded). Note that simple hashing of such identifiers does *not* provide adequate anonymization of PII, as long as the set of potential matches can easily be enumerated (e.g., a set of students in a class). If later re-identification of such values is needed, use an encrypted hash (as “message authentication code”, MAC) and store the key with other PII data in a separate, encrypted file.

- *Anonymization*

If you want to fully anonymize collected identifiers such as phone numbers or Bluetooth MAC addresses, simply use a one-time key for computing an encrypted hash, then destroy the key. You might also want to consider hashing only part of the identifier, e.g., the last 6 digits of a Bluetooth MAC address, but leave the first 6 digits in the clear. This allows, e.g., to identify devices and classes of devices. See also [1] for a more detailed discussion of anonymization techniques of common PII. Note that even where only anonymised data are used, *adequate* security for storage and handling of such data must be ensured [6].

- *Adequate Protection*

The level of protection for all stored data should be commensurate to the expected *impact level* of any inadvertent disclosure of the data. The recommendations in this brief describe the baseline for secure data storage to be adopted in BigClouT – refer to the risk assessment portion (section 4) of the *BigClouT Ethics Worksheet* for a particular user study in order to see whether your study might warrant higher levels of protection. In general, AES encryption should be employed with key strength of at least 64 bits, which corresponds to passwords of at least 10 characters using a mix of uppercase and lowercase, numbers, and special characters. See the useful software links at the end of this brief for links to password generators. If sensitive data (in the sense of directive 1995/46/EC) is recorded, higher protection levels (up to 128 bits) might be necessary. Data may be individually encrypted using application-specific encryption formats (e.g., Word documents) and archive storage formats (e.g., ZIP files), though a transparent directory-based or partition-based disk-encryption approach may offer both a higher level of protection and simplification of access.



## ○ Secure Data Processing Principles

The list below outlines the core principles of secure data storage in BigClouT.<sup>22</sup> They need to be taken into account for *all* data collections in the project, i.e., both for records of user studies and interviews, as well as for system logs that record PII (e.g., Bluetooth MAC addresses).

### ▪ *General Principles*

1. Code data as early as possible, i.e., replace all PII with anonymous identifiers.
2. Keep the coding key<sup>23</sup> separate (in a physically separate space or in a separate electronic file) from the data.
3. Work with de-identified data at all times, unless this is not possible for your work.
4. Never write down passwords. Use a password-storage application if needed (see the Useful Software links below), together with a strong Master Password.

### ▪ *Electronic Data Safekeeping*

5. Store all files containing PII on encrypted drives, or use encrypted files. Ensure that no temporary unencrypted copies are left (e.g., trash).
6. Ensure that you always have a backup of your (encrypted) files in another location, e.g., on another server of your university, in order to avoid data loss.
7. If you need to remotely access data, always use a VPN connection or an encrypted Remote Desktop Session.

### ▪ *Physical Storage*

8. Avoid hard copy media (e.g., video tapes, DVDs, printed documents) for storing PII as much as possible.
9. Store any such media in an institutional environment with restricted access and lockup capability (e.g., in your office in a desk that is locked at all times).
10. If the contents of the media is crucial to your research, ensure that a backup copy is (equally safely) stored in another physical location on campus to avoid data loss from, e.g., fire.
11. Only take such information off-site if absolutely necessary. If you must do so, take all reasonable security precautions consistent with protecting a high-value asset.

---

<sup>22</sup> These principles are modeled after [2]

<sup>23</sup> In this case, the term “key” refers to the index that links the anonymous identifier to PII of the individual study subject, not to an „encryption key“ per se. However, if it is impossible to remove PII from the actual data collection, the entirety of the data must be encrypted with a sufficiently long key and the key must be kept separately as described in principle 2.



- *Data Processing*

12. Do not store or disclose personally identifiable or confidential data other than as necessary for your research
13. Keep an accurate and up-to-date log, detailing your use of personally identifiable and/or confidential data and the specific security and privacy protection measures that you apply.
14. Immediately report privacy concerns (like possible data loss) to the PI and/or your local ERB contact.

- *Data Retention and Deletion*

15. Ensure that records are retained only as long as is required to accomplish research purposes and satisfy legal and policy retention requirements. Note that all PII collected in BigClouT must be deleted at the end of the project, though local retention requirements might take precedence.
16. Ensure the secure destruction of all personally identifiable or confidential information at the end of applicable retention periods. Documents must be shredded, video tapes re-recorded, DVDs and CDs securely disposed, and electronic files must be securely deleted using repeated overwriting.
17. If you made additional copies of your data for backup purposes, (securely) delete all on-line copies (e.g., on servers). If additional off-line copies remain due to institutional backup procedures (e.g., on tape or WORM media), you *must* ensure that all *access keys* are securely deleted (see points 4 and 15 above).

- Useful Software

Below you will find some recommendations for secure data storage and processing software. These may help in implementing some of the guidelines mentioned in this document.

- *Encryption Tools*

**Truecrypt** (MacOSX, Linux, Windows): A free (i.e., without cost) on-the-fly encryption (OTFE) tool. Can encrypt entire partitions transparently, or create new virtual disks that can be mounted as if they were an external disk. Cross-platform.

**FileVault, DiskUtility** (MacOSX): Both tools are part of recent versions of MacOSX. FileVault allows users to encrypt the entire home directory, DiskUtility can be used to encrypt individual directories. To avoid accidental loss of sensitive data it is recommended that laptops have disk encryption facilities (e.g. requiring a BIOS password, turned on by default). For individual files, a simple password-protected zip-file may also be sufficient. The free cross-platform compression utility **7-zip** offers AES-based file encryption. This might also be sufficient if sharing files with other project members via email. Obviously, the password must not be shared via email or instant messaging apps, but only in person or via a phone call.



A *Password Safe* application such as **KeePass** (Windows) or the cross-platform equivalent **KeePassX** (Windows, Linux, MacOSX) can be used to safely store a large number of passwords. Many other (free and non-free) alternatives exist. These applications also allow the creation of random passwords of arbitrary length using a built-in *Password Generator*, which greatly improves the entropy of the used passwords. See [1],[4] for hints on how to come up with (and remember!) a good master password. 1Password provides similar functionality with a more streamlined user interface and cross-device support at modest cost. If you regularly exchange such files via email, you should consider the use of **encrypted email**. Free PGP-based extensions are available for most modern email programs (e.g., the Enigma plug-in for Mozilla Thunderbird). Note that the recipient must use the same encryption system (either PGP/GnuPG or S/MIME based).

#### ▪ *Secure Deletion Tools*

MacOSX already comes with a secure delete option with its standard Trash. “**Secure Empty Trash**” is an option that allows one to securely wipe the entire trash (File > Secure Empty Trash). The system will then perform a 35-pass over the files contained within the trash. Linux systems (e.g., Ubuntu) can use the built-in **shred** command to securely overwrite a file. You might also want to use the **scrub** command to wipe *unused* space periodically. Also, the open source “**Wipe**” utility (<http://wipe.sourceforge.net/>) can be used. See also [5] for further information.

#### ▪ *Bibliography*

- [1] Aad, Imad, and Valtteri Niemi, “NRC Data Collection and the Privacy by Design Principles”. Proceedings of the 1<sup>st</sup> Workshop Sensing for App Phones (PhoneSense 2010), at SenSys 2010, Zurich, November 2010
- [2] Data Security Standards for Personally Identifiable and Other Confidential Data in Research. University of Toronto, Research Ethics Board (REB), April 2010
- [3] “Lockdown Password Guidelines.” Lockdown.co.uk - The Home Computer Security Centre. July 2009. Available from [www.lockdown.co.uk/?pg=password\\_guide](http://www.lockdown.co.uk/?pg=password_guide)
- [4] “Microsoft Advisory: Create Strong Passwords”. Microsoft Online Safety Blog, Microsoft. Available from [www.microsoft.com/protect/fraud/passwords/create.aspx](http://www.microsoft.com/protect/fraud/passwords/create.aspx)
- [5] Peter Gutmann, „Secure Deletion of Data from Magnetic and Solid-State Memory”, Sixth Usenix Security Symposium, 1996. Extended version at [www.cs.auckland.ac.nz/~pgut001/pubs/secure\\_del.html](http://www.cs.auckland.ac.nz/~pgut001/pubs/secure_del.html)
- [6] Eléonore Pauwels: “Ethics for Researchers. Facilitating Research Excellence in FP7”. See <http://ftp.cordis.europa.eu/pub/fp7/docs/ethics-for-researchers.pdf>

Many of the appendices are derived from the PD-NET project. Here is a link to the main ethics document and the PD-NET microsite.

[PD-Net Ethics handbook in PDF format.](#)

Micro-site describing the full process: <http://pd-net.org/ethics/>.



## 9 Appendix 5 The OpenIdeas platform

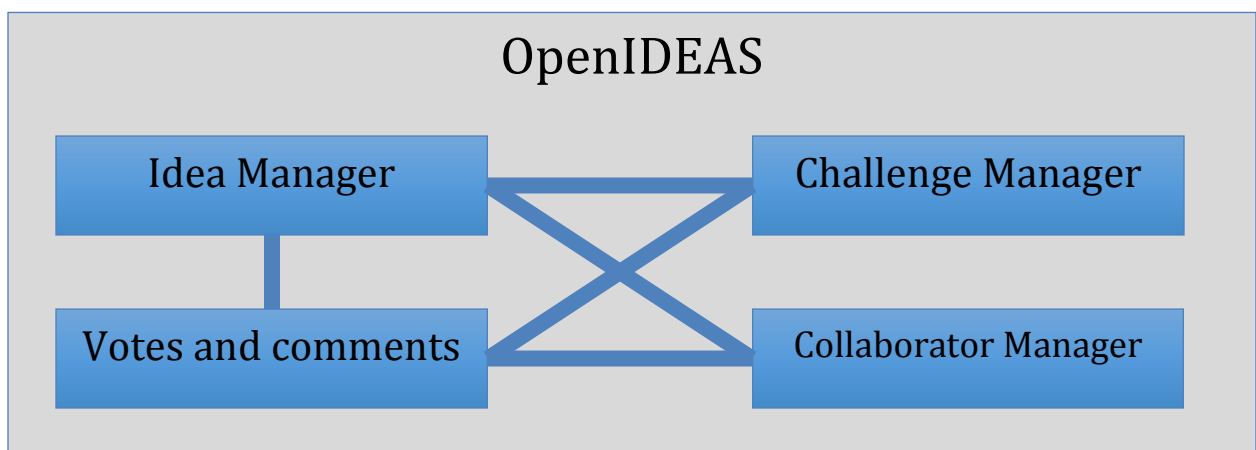
From a technological point of view, OpenIDEAS is based on Liferay [https://www.liferay.com/], an enterprise portal providing a large set of features for the development of websites and portals, such as

- User registration and authentication;
- Role and permissions management;
- A Web content management;
- A documents and media management system;
- Etc.

In particular, Liferay provides an open source implementation of the Java Portlet Specifications (JSR168)[Java Community Process, “JSR 168: Portlet Specification v.2.0”, https://www.jcp.org/en/jsr/detail?id=168, 2003] and it is possible to extend its functionalities using different programming languages (e.g. Java, PHP, Ruby, etc.); portlets are reusable modules that can be executed in a specific container.

OpenIDEAS is composed by four modules:

- “Ideas Manager” is in charge of managing needs, ideas and their idea life cycle and the co-creation activities related to them.
- “Challenges Manager” is in charge of managing challenges and all activities related to them.
- “Votes and comments” provides functionalities allowing users to express opinions on ideas or and to rate them.
- “Collaborator Manager” applies the user permissions to the collaborative activities during the Ideas and Challenges life cycle.



## 10 Appendix 6 Grenoble Questionnaire for business events participants

---

SEMICON EUROPA  
25-27 OCTOBER 2016  
GRENOBLE

# BIGCLOUT

*Big data meeting Cloud and IoT  
for empowering the citizen ClouT in smart cities*

The world is facing a number of critical challenges such as global warming, economic crisis, security threats, inequality, and natural disasters and ageing society. ICT solutions have the potential to change the world and improve the **quality of life and security of its citizens**.

BigClouT project will in particular make use of today's three key technologic enablers, **Internet of Things (IoT), cloud computing and big data**, for the objective of increasing the efficiency in using urban infrastructure, economic and natural resources shared by the increasing population.

BigClouT will offer an analytic mind to the city by creating **distributed intelligence** that can be implanted throughout the whole city network either for large or smaller urban areas.

BigClouT will leverage the results of the EU-Japan [\*ClouT\*](#) project and will bring them beyond, by adding for instance distributed intelligence with edge computing principles, big data analytics capability, in addition to self-awareness and dependability properties towards a **programmable smart city platform**.

Resources and knowledge from prestigious European and Japanese institutions are now together to keep on creating a long-lasting synergy **between EU and Japan** for tackling future city challenges.



## QUESTIONNAIRE

---

- 1 Where are you from?
- 2 Is this your first time in Grenoble?
- 3 Do you regularly attend similar events abroad (exhibitions, conferences, etc.)?
- 4 Do you use smart phone applications for your business trips?
- 5 Would you be interested in a phone application to guide you around Grenoble (transportation from hotel to conference center, souvenir shops, restaurants, etc.) during your time in the city?
- 6 Would you be interested in phone application that combined professional usage (eg. A networking tool for exhibitions) and a tool to guide you around the city?
- 7 Would you be interested in receiving offers giving you discounts in local shops & restaurants?
- 8 Would you give some “anonymized” personalized information (amount you spend during your stay in Grenoble at restaurants, hotels, shops, etc., access to your calendar) for the benefit of receiving further better services during your stay in Grenoble such as personalized offers, transportation service, etc.



Question	Answers																
Where are you from?	France	Italy	Italy	Taiwan	France	France	France	France	France	France	France	France	France	France	Germany	Germany	Germany
Is this your first time in Grenoble?	No	No	No	Second	No	No	No	No	No	No	No	No	No	No	No	Yes	Yes
Do you regularly attend similar events abroad?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes	
Do you use smart phone applications for your business trips?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	
Would you be interested in a phone app to guide you around Grenoble?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Would you be interested in a phone app that combined professional usage and a tool to guide you around the city?	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Would you be interested in receiving offers?	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	
Would you provide anonymised personal information?	Yes	?	?	Yes	No	No	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	

## 11 Appendix 7 Questionnaire for Innovallée employees in Grenoble

---

We are considering the creation of a mobile phone application for employees of Innovallée, which would gather all the practical information of the technopole (catering, travel, events, etc.). In order to best meet your needs, could you take a few minutes to answer this questionnaire?

1. Do you have a smart phone? ☐ yes / ☐ no

2. Do you use mobile apps? ☐ yes / ☐ no

3. If yes, on which areas? : transportation, informations, ... :  
.....

4. Would you be interested in an application proposing specific practical information about Innovallée? ☐ yes / ☐ no

4. Which type of information would you like to find in this application?

☐ Restaurant:

- ☐ Location of restaurants and food trucks
- ☐ Opening/Closing hours
- ☐ Menu & Price
- ☐ Alerts in case of problems (Exceptional closing, etc.)

☐ Transportation:

- ☐ Transportation options (bus, tram, car sharing, etc.) → timetables, travel time, car renting, location of available shared cars, route conditions, alerts in case of problems, etc.

☐ Events and animations

☐ Training

☐ Sportive activities

☐ Cultural activities

☐ Construction work planned in the zone

☐ Plan of Innovallée

☐ Brief presentation of enterprises of the zone

☐ Shops in the zone

☐ Nearest nurseries and maternal assistants

5. Which other information would you like to find in this kind of application? .....  
.....

