Health, demographic change and wellbeing
Personalising health and care: Advancing active and healthy ageing
H2020-PHC-19-2014
Research and Innovation Action

ACANTO

A CyberphysicAl social NeTwOrk using robot friends

Deliverable 10.1.3

Data management plan

Deliverable due date: 09.2015	**Actual submission date: 10.2015**
Start date of project: February 1, 2015	**Duration: 42 months**
Lead beneficiary for this deliverable: UNITN	**Revision: 1.0**
Authors: Luigi Palopoli	
Internal reviewer: All the node leaders	

The research leading to these results has received funding from the European Union's H2020 Research and Innovation Programme - Societal Challenge 1 (DG CONNECT/H) under grant agreement n° 643644.

Dissemination Level		
PU	Public	
CO	Confidential, only for members of the consortium (including the Commission Services)	**X**

Contents

Executive Summary

This deliverables reports the policies adopted by the ACANTO consortium to manage sensitive user data with high privacy and security standards. Such policies will set the stage for a proper management of sensitive information for the entire duration of the project.
In the same document, we also define the specific procedures whereby the open access policy adopted by the EC in the H2020 framework programme and specified in the Grant Agreement will be implemented in ACANTO.

Chapter 1

Introduction

The ACANTO project operates with sensitive information on the participants (which in some cases are clinical patients) and aims to produce a portfolio of technologies that could be commercially exploitable in a time horizon of a few years. Both this aspect mandate the adoption of high security and confidentiality standards for the data.
On the other hand, at the heart of modern research is an extensive scientific dialogue, with a timely sharing of data and experiences. This accelerates innovation, allows researchers to build on previous work improving the quality of the results, fosters collaboration and avoid duplication efforts (with obvious efficiency gains). These requirements have gained prominence and have brought about significant innovation in the political choices of all the main public agencies funding and sponsoring research. The European Commission is no exception to this general internation trend. The commitment of the EC toward open access of the research results is reflected in official guidelines [1] and in the wording of the Grant Agreement (e.g., art. 29.2). In addition, from the specific nature of ACANTO and from its being part of the "societal challenges" programme, we derive a particular emphasis on the involvement of citizens, economic stakeholders, governmental agencies and charities. All these considerations require the adoption of liberal standards for the scientific dissemination of information, in accordance with the mandate in art. 29.2 of the ACANTO grant agreement.
Under these premises, the ACANTO researchers could be in a quandary over whether to enforce highly conservative standards for confidentiality or to share as much as possible their scientific findings. In order to avoid problems and misunderstanding and to streamline the whole process of data collection and of dissemination of results, we need to define clear guidelines on how to treat data and on how to disseminate the results. This is the purpose of this document.

Chapter 2

Security Model of the ACANTO project

2.1 Scope of the security model

The activities of the ACANTO project will strongly rely on active involvement of older adults (with impairments of various kind) during all the phases of the project. During the research the following information will be collected from participants:

1. Classification of the different older adults involved according to their:
   (a) Gender,
   (b) Age,
   (c) Impairments.
2. Collection of data concerning the expectations of the users, their impression on the use of the tested devices and an assessment of the effectiveness of their use of the devices in several benchmarks,
3. Collection of clinical data as clinical status (e.g. Charlson index), main disease, falls, ability to perform daily activities, mental status (e.g. Minimental test), and functional status including frailty (e.g. Fried index). As stated in the deliverable for Baseline measurements D1.1.
4. Recording of videos and still images.

The ethical issues arising from the interaction of the researchers with the users will be managed in accordance with the guidelines reported in the project DoW, by always complying with the national legislation and with the internal regulations of the partners involved in the project. In particular, the European Directive 95/46/EC defined by the European Parliament and Council on 10/24/1995 regarding the data protection for individuals and the data manipulation and distribution will be strictly followed by the project security model. In the Spanish case, the Data Protection Act number 15/1999 (December 13th) adapted to the European directives by the Royal Decree 1720/2007 promulgated on 21st December 2007 will be applicable to the project. The specific goal of this document is to present and discuss the issues related to the treatment of the collected data in electronic form, their storage on different media (CD/DVD/USB PENS), and their distribution using network connections.

2.2 General Principles

The security model of ACANTO rests on three pillars:

- Data anonymity,

- Informed consent
- Limited circulation of the information.

Data anonymity will be guaranteed whenever possible. The only exemption to anonymity is for the researcher directly interacting with the participants. Data anonymity will be controlled by creating two lists. The first (index) list allocates a unique code to each participant (participant code). Any other lists, where data is associated to a participant will be indexed via the participant code. The creation of an index list is to allow participants to withdraw their data as required by ethics policy. The index list will be stored electronically by the lead researcher responsible for data collection. This will not be shared with any other researchers. It will be stored in a password file on a secure computer. The files will be destroyed five years after the end of the project. Any participant can require the cancellation of the entry related to his name up to one month after the data collection and this request will be immediately granted.
The *informed consent* policy will require that each participant will provide his/her informed consent prior to the start of any activity involving him/her. A specific section will specify the consent to treatment and storage of electronic data and of recordings. For illustrative purposes we report, as an appendix, a template of an informed consent form that will be completed by participants. Public distribution of elements of information that can reveal the identity of the users (e.g., videos or pictures) for scientific dissemination purposes will be explicitly authorised by the participant as part of this process.
To achieve a *limited circulation* of the information, the database containing in anonymous form the data collected from the users (e.g., the results of questionnaires and of laboratory experiments) will be distributed to the partners through protected and encrypted Internet connections. The raw data will only be shared if it is required for the development. In most cases summary data will be provided to research partners. The researchers will never pass on or publish the data without first protecting participants' identities. No irrelevant information will be collected and the participants' well-being will always be considered a priority in the study. At all times, the gathering of private information will follow the principle of proportionality by which only the information strictly required to achieve the project objectives will be collected. In all cases, the right of data cancellation will allow all users to request the removal of their data from the project repository at any time.
More details on our security framework and on our specific technological solutions are offered in the next sections.

2.3 Security Framework

In order to accomplish the creation of a security framework it is essential to focus on the issues of access and identity authentication, authorisation and auditing (AAA). Therefore, our main objective is to develop a base security system that standardises the processes of Authentication, Authorisation and Auditing of the various information sources involved.

2.3.1 Authentication

The Data Protection Act requires that any operator who is granted access to sensitive data with be authenticated. Authentication technology should be strict when dealing with sensitive and confidential data available to the users of the platform. To do this, a username and a password will be used so that the person who wants to access to the health data of a citizen confirms that he has access to the system. In addition, we will use a RSA encryption mechanism, with each operator receiving a personal private key.

2.3.2 Authorisation

The objective of the authorisation is to determine the rights of a user of an information system. In our specific context, the access rights will enable the researcher to access health data from a citizen. This compels us to have an access policy to the information. For each group of researchers, we will specify which content can be accessed based on functionality, security and confidentiality criteria. Technological or hierarchical criteria should not be part of the criteria to decide the access rights. For instance, the simple fact that an operator is the general manager or the health information system administrator does not per se grant him/her access to all information, including health data of the patients.

2.3.3 Audit

The Data Protection act mandates a regular of the health data access of all citizens by authorised operators. If, in the execution of the project, the consortium needs to store data related to the participant health, then we will use a Database supporting the auditing functionality. We will audit queries to the Database containing sensitive data, as well as the modifications of the stored data (new inclusions, changes or deletions). The implementation of the audit functions requires implementing appropriate services in the existing infrastructure. Since the tool that is preferred to be used to access data recorded in the DB of the system to develop is Hibernate, it will allow us to use the tool Envers that performs the audit of the access to the data automatically, being this recorded in specific audit tables.

2.4 Considerations on indirect human presence

The activities of the ACANTO project consider a second possible source of sensitive information. Indeed, our activities in the user centres could indirectly reveal the presence of other people than the observed ones, who could or could not participate in the project. Examples of this information are the following: 1) Videos from security cameras used for model behaviour observation or another kind of studies. 2) Images of the user's centre, surroundings or use case locations, with the objective of academic or commercial dissemination of the project. The users centre will yield the rights of these contents to the ACANTO Consortium through appropriate private contracts. The petitions for withdrawing personal information will be processed by the users centre.
The data containing such information will be treated in a same way as the ones regarding the users directly involved in the project. Specifically, we will apply the same policy in terms of security, complying in all cases with the ethical issues, general principles and security framework detailed in this chapter in case of risk of possible identification of individuals. In case of difficult identification of subjects, we will guarantee the data anonymity, erasing all characteristic feature of people present on the contents, using the needed instruments or techniques with the goal of avoiding in any case the identification of them. This process will be done before circulation between partners or using the contents in the project.

2.5 Summary of Technological solutions

We report below a table of the main technological solutions used for the different security issues mentioned in the section above.

GOAL	Technological Solution
Guaranteeing complete anonymity where required.	The collected data will be labelled with participant codes. Participant consent forms will be held separately and will not reference the participant code. These will be paper based and held in a locked filing cabinet on the researchers site.

GOAL	Technological Solution
Safe keeping of the documentation on informed consent.	The informed consent will be provided by the interested subject by filling in an appropriate form. This document will be generated by the information system and it will contain an explanation of what the subjects allow the operator to do with his/her private data, along with all data identifying the citizen and the user. In the appendix, we report a possible template of informed consent form that we use. The authorised personnel must keep this physical document under lock and key. Information on the interested person can also be stored in electronic form in an database or in a spreadsheet. The spreadsheet or the database will be encrypted and its access will be password-protected and granted only to a few selected operators. The password will be changed on a regular temporal basis. When this document is registered, a code will be generated that anonymously identifies the subject in the Data base of the project. Only after a successful registration of this data can the entry related to the subject be used for the project purposes. In case the database is stored in a laptop computer, it will have to be endowed with one or more of the following features: 1) biometric authentication of the user, 2) smart-card based authentication, 3) remote destruction of the data. Moreover, the notebook will support hardware encryption of the data.
Remote access	In the general case, the ”raw” data related to the participant to the project, will be handled only by the researchers interacting with the participant and made available to the rest of the consortium only in aggregate form. If, for special cases, some other researchers should need to access to the ”raw” data, the interested participants will be informed. Only after their consent is extended to the requiring researcher, can he/she have access to the anonymised data. In this case, if the access is remote, the system has to have the following Researchers in the consortium can have access to the anonymised data through an internet connection using a web interface with the following features: 1) the data will be stored on a server whose only access to the internet is via an SSL connection (no other ports will be enabled), 2) the connection will be through an https protocol with both password and key based interconnection. Each user will receive his/her own private key and will be required to sign a privacy statement whereby he/she commits not to share the key with anybody within or without the research unit.
Auditing	If the execution of the project requires storing health sensitive data, we will use the Hibernate tool. This will allow us to use the tool Envers that performs the audit of the access to the data automatically, being this recorded in specific audit tables.

Chapter 3

The ACANTO Open Access Policy

One of the most important challenges of research projects like ACANTO is to produce scientific knowledge for the benefit of the individuals, of the communities and of European economy. This goal requires a thorough dissemination activity of the research results toward the scientific community, the care–givers communities and the members of the public.
One of the cornerstones of our dissemination strategy is to secure a timely and regular publication of the scientific findings of ACANTO in peer–reviewed, reputable (high impact) journal and conferences. This will ensure a proper consideration of ACANTO's result in the scientific communities of interest. On the same level of importance is the implementation of a consistent and far reaching *open access publishing policy*. As specified in the H2020 Guidelines on open access publishing [1], by this term we mean the practise of providing free and unrestricted access to scientific publications to read, download and reuse.
In our context, scientific information has two distinct facets:

1. Open access to scientific publications,

2. Open access to research data that underlie the scientific publications.

In the following text, we specify the ACANTO policy for each of these two.

3.1 Open access to scientific publications

According to the contractual obligations specified in ACANTO's grant agreement art. 29.2, "Each beneficiary must ensure open access (free of charge online access for any user) to all peer-reviewed scientific publications relating to its results." We obviously will comply with this obligation and develop a specific policy that suits the specific needs and requirements of the work developed in ACANTO.

3.1.1 What is covered by our open access policy and what is not

The open access policy for scientific publications *applies whenever a partner of the project or a group of partners decides to produce a scientific publication* containing the results of a research activity. This decision is taken on the following grounds:

- the publication is scientifically relevant and brings forth significant advances in the state of the art of the interested discipline,

- (if applicable) the data contained in the publications fulfil the requirements specified by the Ethical Committees of the partner/partners that collected the data,

- the publication of the data does not compromise the possibility of taking on actions for the protection of the intellectual property and for their commercial exploitation.

The open access policy does not apply to partial results which are produced at intermediate steps of the project and are not deemed scientifically relevant. On the other hand, the most important goal of ACANTO's dissemination policy is to maximise the presence of the attendance and the visibility of the ACANTO results in all the events related to active aging, robotics, embedded systems and computer vision. Hence, the general philosophy is to avoid being overly conservative and publish the results even in workshops or work in progress sessions in main conferences even before the research has come to a full completion, whenever this choice does not come at the expense of the scientific credibility of ACANTO's results.
As regards the publication of results that require protection of intellectual property, the Consortium Agreement has set up the necessary procedures to obtain a timely approval from the competent offices of the interested partner in case the work is done in cooperation between multiple partners. This will arguably reduce unnecessary delays in the publication of results that could be exploited in commercial products. When the scientific publication is not possible (e.g., because of a pending patent registration), the vision of the ACANTO project and the commitment of its partners is to publish the results as soon as IP protection procedure permit it, even after the end of the project.

3.2 Procedures adopted to implement the open access policy to scientific publications

The open access policy will be widely applied both to peer-reviewed publications (i.e., publications that are evaluated by "peers") and to other types of publications such as books, unreviewed reports and other type of "grey" publication (i.e., publication that are informally published without the supervision of scientific publishers). We we will henceforth refer to the first type of publications as "peer-reviewed" and to the second as "informal".

3.2.1 Definitions

Green and gold open access routes. The H2020 guidelines [1] provide two main routes to the implementation of an open access policy:

The green open access: this route is based on self-archiving meaning that the published article or the final peer–reviewed manuscript is archived into a repository by the authors (or by a representative of the authors); some publishers could require an *embargo* period of time before the paper is made concretely available to the public ;

The gold open access: in this model the article is immediately provided in open access mode by the publisher (typically after a payment of the publication costs by the authors).

We will use both of these options as the specific situation (and the choice of the authors) demand.

Repositories A repository for scientific publications is generally defined as an online archive. The H2020 guidelines give full freedom on the choice of the repository: it can be an Institutional Repository or a subject–based centralised repository. If the Institution the authors belong to do not have a specific infrastructure of this kind, the EU offers a facility (`http://www.openaire.eu`), with a comprehensive list of available repositories. Other lists are available at `http://roar.eprints.org` and `http://www.openoard.org`.

Bibliographic Metadata The ACANTO grant agreement requires that along with each publication stored in a repository its bibliographic metadata be available in open access. Such data have to be in a standard form and have to include:

- The terms "European Union (EU)" and "Horizon 2020";
- The name of the action ("Research and innovation action"), acronym ("ACANTO") and grant number ("643644");
- The publication date, the length of the embargo period (if applicable), and a persistent identifier.

Accepted version and published version An *accepted paper* is s a version which has been revised by the author to incorporate review suggestions, and which has been accepted by the publisher for publication.
The final, *published version* is the reviewed and accepted article, with copy-editing, proofreading and formatting added by the publisher.

3.2.2 Procedures for peer-reviewed publications

The authors of ACANTO publication have the freedom to opt for either for a green or for a gold policy. In case of *a green policy* the procedure is as follows:

1. As soon as the paper is accepted, the draft of the accepted manuscripts is stored in a repository of the authors' choice along with bibliographic metadata,
2. If requested by the publisher, the paper is left unpublished for the duration of the embargo period; such period cannot exceed 6 months;
3. After the embargo period expires, the open access is granted to every one via the repository;
4. The paper publication is notified to the project coordinator and to the exploitation and dissemination list (acanto_wp9@list.disi.unitn.it).

In case of a *gold access policy* the procedure is:

1. As soon as the paper is accepted, the draft of the accepted manuscripts is stored in a repository of the authors' choice along with bibliographic metadata,
2. The open access is granted to every one via the repository;
3. The paper publication is notified to the project coordinator and to the exploitation and dissemination list (acanto_wp9@list.disi.unitn.it).

The costs incurred for publication are eligible for reimbursement as long they are incurred before the end of the project.
In both caes the prublication will be reported in the ACANTO website, along with bibliographical metadata and with a pointer to the repository where the paper is stored. A similar information will be reported in the annual dissemination report.
Clearly, the choice of whether to take a green or a gold policy is also determined by the specific publisher. For the authors' convenience, we report below the policy contained in the copyright agreement of some of the most relevant publishers at moment of this writing:

IEEE: The IEEE specifies its policy in a document that can be found in the association website[2]. In summary:

- IEEE recognises to the authors the right to post on the authors website on the accepted version but not on the published version
- IEEE also has an open access program for *gold access policy*, which at the moment is limited to the societies journals.
- IEEE allows its authors to follow mandates of funding agencies and post the accepted version into publicly available repositories.

Clearly, the IEEE requires to explicitly mention that the document is IEEE copyrighted, reporting the publication data and the DOI as soon as they are available.

ACM: The ACM policy can be found in the website https://www.acm.org/publications/policies/copyright_policy. At point 2.5 the document discusses the author rights in the following therms `The original Owner/Author permanently holds these rights:`

`... * Post the Accepted Version of the Work on (1) the Author's home page, (2) the Owner's institutional repository, or (3) any repository legally mandated by an agency funding the research on which the Work is based.` Thereby there is no apparent problem in the adoption of a green open access policy for the *accepted versions* of the manuscripts. The ACM also offers the possibility to opt for a gold open access policy, as detailed in the website http://authors.acm.org/main.html: `Authors have the option to choose the level of rights management they prefer. ACM offers three different options for authors to manage the publication rights to their work. .... Authors who wish to retain all rights to their work can choose ACM's author-pays option, which allows for perpetual open access through the ACM Digital Library. Authors choosing the author-pays option can give ACM non-exclusive permission to publish, sign ACM's exclusive licensing agreement or sign ACM's traditional copyright transfer agreement`

Elsevier: The Elsevier policy on authors right can be found in the website http://www.elsevier.com/about/company-information/policies/sharing. Regarding the author rights on the *accepted versions* of the manuscripts, we find the following wording `Authors can share their accepted manuscript:`

`Immediately`

- `via their non-commercial personal homepage or blog by updating a preprint in arXiv or RePEc with the accepted manuscript`
- `via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group ...`

`After the embargo period`

- `via non-commercial hosting platforms such as their institutional repository`
- `via commercial sites with which Elsevier has an agreement`

`In all cases accepted manuscripts should:`

- `link to the formal publication via its DOI bear a CC-BY-NC-ND license...`

`...` The CC-BY-NC-ND license can easily be obtained through the website http://creativecommons.org/licenses/ and is explicitly recommended by the EC to *enable open access in its broadest sense*.

Springer: An excerpt from the copyright transfer agreement reads: `Authors may self-archive the author's accepted manuscript of their articles on their own websites. Authors may also deposit this version of the article in any repository, provided it is only made publicly available 12 months after official publication or later. He/ she may not use the publisher's version (the final article), which is posted on SpringerLink and other Springer websites, for the purpose of self-archiving or deposit. Furthermore, the author may only post his/her version provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be provided by inserting the DOI number of the article in the following sentence...` The only problem is with the duration of the embargo (12 months), which is not compatible with the EC requirement of a 6 month maximum duration. To stay on the safe side, it could be recommendable to use a golden approach along the lines in the website `http://www.springer.com/open+access?SGWID=0-169302-0-0-0`.

In conclusion the self archiving policy seems to be compatible with the most important publishers, as far as it is limited to the *accepted version* of the paper. With other publishers, the evaluation should be made on a case by case basis. In the extreme case in which self archiving is prohibited and commercial open access options are not available, the authors should avoid the journal.

3.2.3 Procedures for informal publications

The researchers in ACANTO are strongly encouraged to adopt an open access policy also for informal publications such as technical reports and white papers. The procedures is in this case is very simple:

1. When a technical report is published (e.g., on an institutional website), the authors stores a version of the paper, along with the available metadata, in a repository of her/his choice.
2. The publication is notified to the project coordinator and to the exploitation and dissemination list (acanto_wp9@list.disi.unitn.it).

The coordinator will take carte of a timely publication of the paper in the ACANTO website.

3.3 Open Research Data

An interesting novelty of H2020 is the platform known as Open Research Data Pilot for the dissemination of the data that could be used by different researchers to replicate the experiments or the analysis presented in the scientific publications. Given that most of the data generated for ACANTO research either contain sensitive information on the participants (in some case clinical data) or could be useful for a possible commercial exploitation, the consortium opts out of the Pilot. However, the consortium will consider the possibility of a participation limited to the data that can be safely disclosed.

Bibliography

[1] The European Commission. Guidelines on open access to scientific publications and reearch data in horizon 2020 – version 1.0. http://ec.europa.eu, December 2013.

[2] IEEE. An faq on ieee policy regarding authors rights to post accepted versions of their articles. https://www.ieee.org/documents/author_version_faq.pdf.