COMPACT

CYBERSECURITY FOR LOCAL ADMINISTRATIONS

D2.6 Data Management Plan (v1)

Work Package:	WP2
Lead partner:	KUL
Author(s):	Danaja Fabčič Povše (KUL), Mariacarla Staffa (CINI), Daniela Messina (CINI), Filipe Apolinário (INOV), Cornelia Gerdenitsch (AIT)
Due date:	October 2017
Version number:	0.1 **Status:** Draft

Grant Agreement N°:	740712
Project Acronym:	COMPACT
Project Title:	COmpetitive Methods to protect local Public Administration from Cyber security Threats
Call identifier:	H2020-DS-2016-2017
Instrument:	IA
Thematic Priority:	Secure societies – Protecting freedom and security of Europe and its citizens
Start date of the project:	May 1st, 2017
Duration:	30 months

Dissemination Level	
PU: Public	✓
PP: Restricted to other programme participants (including the Commission)	
RE: Restricted to a group specified by the consortium (including the Commission)	
CO: Confidential, only for members of the consortium (including the Commission)	

Revision History

Revision	Date	Who	Description
Internal	01/08/2017	Danaja Fabčič Povše	Template and input for KUL
Internal	13/10/2017	Filipe Apolinário	Input for INOV
Internal	16/10/2017	Mariacarla Staffa, Daniela Messina	Input for CINI
Internal	17/10/2017	Cornelia Gerdenitsch	Input for AIT
Internal	20/10/2017	Danaja Fabčič Povše	Integration and editorial review

Quality Control

Role	Date	Who	Approved/Comment
Internal reviewer	27/09/2017	Ioana Cotoi (ENG)	Approved
Internal reviewer	25/10/2017	Alexander Krock (BIT)	Approved

Disclaimer:

This document has been produced in the context of the COMPACT Project. The COMPACT project is part of the European Community's Horizon 2020 Program for research and development and is as such funded by the European Commission. All information 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. For the avoidance of all doubts, the European Commission has no liability with respect to this document, which is merely representing the authors' view.

Table of Contents

Definitions and acronyms

COMPACT	COmpetitive Methods to protect local Public Administration from Cyber security Threats
DMP	Data Management Plan
DPD	Data Protection Directive
FAIR data	Findable, accessible, interoperable and re-usable data
GDPR	General Data Protection Regulation
IPR	Intellectual property rights
S.E.L.P.	Security, ethics, legal and privacy

1. Introduction

This deliverable will set out the first version of the data management plan (DMP) for the COMPACT project. A DMP is a key element of good data management, which is especially important in the COMPACT context, as all Horizon 2020-funded projects from 2017 onward are required to contain a DMP.[1]

This DMP is based on the European Commission's Guidelines on FAIR Data Management in Horizon 2020[2] and the COMPACT Grant Agreement.[3]

It reflects the consortium's comprehensive approach towards data management. It is a living document, which will be updated in months 18 and 24 (DMP version 2, and the final version, respectively), due to the possible significant changes, including but not limited to:

- Use of new data,
- Changes in consortium policies (e.g. new innovation potential, decision to file for a patent, etc.),
- Changes in consortium composition and external factors (e.g. new consortium members joining or existing members leaving).

This deliverable will contribute towards legal and ethical compliance regarding data protection, alongside the Deliverable 2.5 'S.E.L.P. Framework'. While the latter focuses specifically on legal and ethical aspects of principles and minimum requirements of procedures, necessary for proper data collection, this document will serve as a project management tool, implementing those requirements in terms of data management.

In order to implement the open data principle, the DMP sets out the following information:

- The handling of research data during and after the end of the project,
- What data will be collected, processed and/or generated,
- Which methodology and standards will be applied,
- Whether data will be shared/made open access and
- How data will be curated and preserved (including after the end of the project).

Sections 2 to 7 of this document will cover the different DMP components, based on the outline suggested in the Guidelines. They are based on input from the following partners: AIT, CINI, INOV and KUL, as indicated in the relevant sections.

---

[1] European Commission, H2020 Programme Guidelines on FAIR Data Management in Horizon 2020 Version 3.0, July 2016, p. 3.
[2] European Commission, H2020 Programme Guidelines on FAIR Data Management in Horizon 2020 Version 3.0, July 2016, p. 4.
[3] Grant agreement no. 740712 – COMPACT, Part B – Section 2.2.5.

2. Data summary

2.1. AIT

What is the purpose of the data collection/generation and its relation to the objectives of the project?

For AIT the purpose of collecting user data is to understand user behaviour on an analytical basis.

What types and formats of data will the project generate/collect?

AIT will record audio and video of test participants. In addition we will save log-files within the prototypes. We will also collect data via online surveys.

Will you re-use any existing data and how?

AIT will not re-use any existing data.

What is the origin of the data?

AIT will collect data by observing users during technology interaction and asking them (either in real time or via online surveys).

What is the expected size of the data?

1 TB (which will mainly be video recordings of end-user interaction behaviour)

To whom might it be useful ('data utility')?

Recordings of end-users (besides being the basis for end-user-studies in the project) are – due to their heavy context dependence – not useful to third parties. It would also create a privacy problem for end-users if the recordings would be public. Hence they are closed.

2.2. CINI

What is the purpose of the data collection/generation and its relation to the objectives of the project?

Within COMPACT Project, CINI is in charge of developing an advanced Security Information and Event Management (SIEM) system endowing LPAs' organisation with real-time monitoring capabilities. SIEM services receive log files, which represent records

of the events occurring within an organization's systems and networks when a user attempts to authenticate into the system or a system event occurs (such as starting a service or shutting down the system, etc.). The content/records of these log files related to computer security information are then analysed for investigating malicious activities. A particular alarm or event is generated in relation to the particular detected attack.

What types and formats of data will the project generate/collect?

SIEM systems come with a number of adapters for receiving data/events from a wide variety of sources, such as Operating System (OS) log files (in proprietary or open formats) or Commercial Off The Shelf (COTS) products for logical and physical security monitoring, including: Wireshark, Nessus, Nikto, Snort, Ossec, Argus, Cain & Abel, OpenNMS, Nagios, CENTEROS, Ganglia, Milestone, openHAB, IDenticard, FieldAware, and CIMPLICITY.
In terms of data generated, a format has not been defined yet. However, any XML-compliant format (such as for example the "Json" (JavaScript Object Notation) format) can represent a valuable solution for alarms/events generated by the SIEM system.

Will you re-use any existing data and how?

It is not definitive in this phase, since the datasets are not currently specified. However, in a first phase of the SIEM service implementation, we foresee to use existing anonymized data present within the archives of some of the LPAs involved in the project.

What is the origin of the data?

Data collected and analysed by the SIEM system will be originated by testing activities carried out at a pilot sites (LPAs involved within the project) and will be relied on the collection and analysis of log files of the LPAs participating in the COMPACT project.

What is the expected size of the data?

At this stage of the project, it is not possible to predict the size of the data that will be processed, but we can hypothesize that they will be more than a terabyte.

To whom might it be useful ('data utility')?

Data collected by the SIEM system can be useful to the other technical partners in charge of developing COMPACT's tools and services, such as risk assessment tool or personalization of training courses for LPAs' employees, etc.. They could be useful to other research groups working on similar research, as well as for testing alternative SIEM solutions.

2.3. INOV

What is the purpose of the data collection/generation and its relation to the objectives of the project?

> During the COMPACT project, INOV will collect data to test and demonstrate its Business process intrusion detection system (BP-IDS). This data collection is related with the project objective "SO3: Lower the entry barrier to timely detection and reaction to cyber-threats", and may occur during the tasks: "Task 4.3 Threat intelligence and monitoring Component"; "Task 4.5 Integration of solutions in a unified platform"; "Task 5.1 Validation and Demonstration scenarios"; "Task 5.2 Trials Setup"; and "Task 5.3 Pilot execution and demonstration".

What types and formats of data will the project generate/collect?

> It is not definitive in this phase, since the datasets are not currently specified. However, it is expected that BP-IDS collects data from multiple sources of data, such as: network traffic generated during the communications of the monitored hosts; or by inspecting specific files present in the file-system of the monitored hosts.

Will you re-use any existing data and how?

> It is not definitive in this phase, since the datasets are not currently specified. But it is expected that all the data collected is self-contained in the dataset used, and not re-used existing data.

What is the origin of the data?

> It has not been decided yet, the datasets need to be specified first in order to respond to this question.

What is the expected size of the data?

> It is difficult to estimate the size of the data at this stage, because the size of the data highly varies on the network protocols or the files monitored.

To whom might it be useful ('data utility')?

> The principal benefactor of this dataset will be CMA that will use the tools developed to monitor threats against their infrastructure. INOV will use it for adapting BP-IDS for LPAs. Besides INOV this dataset might be useful to technical partners in the COMPACT project, that require live data to adapt their technical solutions to LPA environments.

3. FAIR data

Under Horizon 2020's principle of open access to data, research data must be FAIR: findable, accessible, interoperable and reusable. This will contribute to the use of data in future research.[4]

In order to be **Findable**:

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specify the data identifier.

In order to be **Accessible**:

- A.1. (meta)data are retrievable by their identifier using a standardized communications protocol.
- A1.1. the protocol is open, free, and universally implementable.
- A1.2. the protocol allows for an authentication and authorization procedure, where necessary.
- A2. metadata are accessible, even when the data are no longer available.

In order to be **Interoperable**:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles.
- I3. (meta)data include qualified references to other (meta)data.

In order to be **Re-usable**:

- R1. meta(data) have a plurality of accurate and relevant attributes.
- R1.1. (meta)data are released with a clear and accessible data usage license.
- R1.2. (meta)data are associated with their provenance.
- R1.3. (meta)data meet domain-relevant community standards.

Answering the following questions will contribute towards compliance with the FAIR data standards. The answers are provided in a comprehensive manner, not on a yes/no basis.

---

[4] Mark D. Wilkinson et al., The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data* **3**, Article number: 160018 (2016).

3.1. Making data findable, including provisions for metadata

3.1.1. AIT

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

> The scientific publications from AIT will end up with a DOI when accepted at a conference.

What naming conventions do you follow?

> None.

Will search keywords be provided that optimize possibilities for re-use?

> Yes.

Do you provide clear version numbers?

> Yes, versioning is already implemented in the document templates.

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

> AIT TX will use standard HCI classifiers from ACM.

3.1.2. CINI

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

> Not defined yet. However, if the Zenodo Repository will be adopted for data storing and sharing, the persistent identification through DOIs for sharing research result will be adopted.

What naming conventions do you follow?

We refer to the "Glossary of Key Information Security Terms" provided by NIST[5] or, in turn, to the SANS Glossary of Security Terms[6].

Will search keywords be provided that optimize possibilities for re-use?

At this stage we have not yet planned to provide keywords for optimizing re-use.

Do you provide clear version numbers?

Not defined yet

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Not defined yet

3.1.3. INOV

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

Not defined yet

What naming conventions do you follow?

Not defined yet

Will search keywords be provided that optimize possibilities for re-use?

Not defined yet

Do you provide clear version numbers?

Not defined yet

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Not defined yet

---

[5] R. Kissel. Glossary of key information security terms. NIST Interagency Reports NIST IR 7298 Revision 1, National Institute of Standards and Technology, February 2011.
[6] http://www.sans.org/security-resources/glossary-of-terms/.

3.2. Making data openly accessible

3.2.1. AIT

What methods or software tools are needed to access the data?

> The COMPACT project strives to make data available in a format, which can be read by free tools (also) to not force people to buy software only to read through the COMPACT outcomes.

Is documentation about the software needed to access the data included?

> No.

3.2.2. CINI

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.
Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out. There has been no opt-out in the COMPACT project yet.

> Data produced and/or used in the project will be made openly available by default only after a pseudonymisation and/or anonymization process in order to prevent that the data can be attributed to a specific person.

How will the data be made accessible (e.g. by deposition in a repository)?

> Data will be made accessible through a research data repository. The consortium will take measures to enable third parties to access, mine, exploit, reproduce, and disseminate the data free of charge.

What methods or software tools are needed to access the data?

> The best candidate tool for data sharing – at the time of this writing – is ZENODO, an OpenAIRE/CERN compliant repository. Zenodo builds and operates a simple and innovative service that enables researchers, scientists, EU projects and institutions to share, preserve and showcase multidisciplinary research results (data and publications), that are not part of the existing institutional or subject-based repositories of the research communities.
> Zenodo enables researchers, scientists, EU projects and institutions to:
> - easily share the long tail of small research results in a wide variety of formats,

> including text, spreadsheets, audio, video, and images across all fields of science.
> - display the research results and receive credit by making the research results citable and integrating them into existing reporting lines to funding agencies like the European Commission.
> - easily access and reuse shared research results.

Is documentation about the software needed to access the data included?

> Yes it is.

Is it possible to include the relevant software (e.g. in open source code)?

> It is possible, but not decided yet if open source code will be included.

Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

> The consortium plans to deposit data in an OpenAIRE compliant research data repository.

Have you explored appropriate arrangements with the identified repository?

> Not defined yet

If there are restrictions on use, how will access be provided?

> Not defined yet

Is there a need for a data access committee?

> Not defined yet

Are there well described conditions for access (i.e. a machine-readable license)?

> Not defined yet

How will the identity of the person accessing the data be ascertained?

> Not defined yet

3.2.3. INOV

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out. There has been no opt-out in the COMPACT project yet.

Not defined yet

How will the data be made accessible (e.g. by deposition in a repository)?

Not defined yet

What methods or software tools are needed to access the data?

Not defined yet

Is documentation about the software needed to access the data included?

Not defined yet

Is it possible to include the relevant software (e.g. in open source code)?

Not defined yet

Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

Not defined yet

Have you explored appropriate arrangements with the identified repository?

Not defined yet

If there are restrictions on use, how will access be provided?

Not defined yet

Is there a need for a data access committee?

Not defined yet

Are there well described conditions for access (i.e. a machine-readable license)?

Not defined yet

How will the identity of the person accessing the data be ascertained?

Not defined yet

3.3. Making data interoperable

3.3.1. AIT

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

AIT will rely on XML when publishing HCI-patterns, which guarantees data exchange with existing HCI-pattern providers.

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

PLML (pattern language mark-up language)

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

No.

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Yes.

3.3.2. CINI

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available

(open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

Yes

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Not defined yet

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

Not defined yet

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Not defined yet

3.3.3. INOV

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

Not defined yet

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Not defined yet

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

Not defined yet

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Not defined yet

3.4. Increase data re-use (through clarifying licenses)

3.4.1. CINI

How will the data be licensed to permit the widest re-use possible?

Not defined yet

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

At this stage of the project it is not possible to predict the date for making re-use available.

Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

Anonymised data will be usable by third parties.

How long is it intended that the data remains re-usable?

We intend to store data and make it re-usable for an appropriate period of time according to the key guidelines established by the following regulations: Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union (NIS Directive)   Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

Are data quality assurance processes described?

Not defined yet

3.4.2. INOV

How will the data be licensed to permit the widest re-use possible?

> Not defined yet

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

> Not defined yet

Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

> Not foreseeable at this stage

How long is it intended that the data remains re-usable?

> For the duration of the project.

Are data quality assurance processes described?

> Not defined yet

4. Allocation of resources – the whole consortium

According to the Horizon 2020 rules, costs related to open access to research data are eligible for reimbursement during the duration of the project under the conditions defined in the COMPACT Grant Agreement, in particular Articles 6 and 6.2.D.3.[7] These are direct costs, related to subcontracting of project tasks, such as subcontracting the open access to data.

What are the costs for making data FAIR in your project?

> According to the budget, AIT has planned to address 10000 EUR for making data FAIR in the project.

---

[7] See the Annotated Model Grant Agreement, available at:
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf#page=83

How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

Not defined yet

Who will be responsible for data management in your project?

A specific role has been foreseen in the project, the Data Controller (DC). Salvatore D'Antonio from CINI has been appointed as Data Controller and will be responsible for data management.

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

Not defined yet

5.   Data security

5.1.1. AIT

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Standard data at AIT is stored on a state-of-the-art secured storage. For sensible data encrypted file storages are created on demand with extremely restricted access.

Is the data safely stored in certified repositories for long term preservation and curation?

AIT runs regular backups on all data. This ensures preservation.

5.1.2. CINI

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Regarding security, all the data collected will be stored on a database only accessible to authenticated users on the partner premises. Regarding the data recovery, database backups will be stored on premises and only accessible to CINI.

Is the data safely stored in certified repositories for long term preservation and curation?

> It is not definitive in this phase.

5.1.3. INOV

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

> Regarding security, all the data collected will be stored on a database only accessible to authenticated users on the partner premises. Regarding the data recovery, database backups will be stored on premises and only accessible to INOV.

Is the data safely stored in certified repositories for long term preservation and curation?

> It is not definitive in this phase, but it is not expected to store collected data in a repository.

6.   Ethical and legal aspects

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

> Two types of data will be used in the COMPACT research: personal data and anonymised data.
>
> Personal data is defined in the General Data Protection Regulation (GDPR) as 'any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person'.

When processing personal data, data protection legislation applies. Until May 25th 2018, this is the Data Protection Directive (DPD)[8] and the relevant legislation, transposing the DPD into national law, and after that date, the General Data Protection Regulation (GDPR).[9] Deliverable D1.2 'S.E.L.P. Management Plan' sets out the management procedures, enabling the consortium to comply with legal requirements. Ethics requirements are addressed in WP8, Deliverables D8.1-8.3.

Anonymised data, on the other hand, are not subject to such requirements.[10] This is because once data have been successfully anonymised, their subjects can no longer be (re-)identified. Therefore, there are no specific data protection-relevant provisions in EU law, which hinder dissemination or further use of anonymised data. Data must be anonymised in a manner that absolutely prevents the data subject from being re-identified.

While there are no specific provisions in the Open Data Research Pilot requiring the participants to anonymise data, the open research data from the COMPACT project will be anonymised before it is made publically available.

Anonymisation is a data processing operation, so the GDPR requirements apply before and while it is being carried out,[11] especially the basic principles such as data minimisation and purpose limitation. The procedure for carrying out a GDPR-compliant anonymisation procedure is described in Deliverable D2.5, 'S.E.L.P. Framework'.

Regarding possible intellectual property (IP) restrictions on the use of research data, these are dealt with in the Consortium Agreement. Research data qualifies as 'results', which are defined as any (tangible or intangible) output of the action such as data, knowledge or information – whatever its form or nature, whether it can be protected or not – that is generated in the action, as well as any rights attached to it, including intellectual property rights. Research results are owned by the partner which produced them. Regarding access to such results, partners will conclude individual agreements with end-users.

Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

---

[8] 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.

[9] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[10] See Recital 26 of the GDPR: The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.

[11] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques, p. 7.

All participants will give their free and informed consent before any personal data is obtained from them. In order to do so, they will be provided with Informed Consent Forms and Information Sheets, which set out the purposes and means of data collection and their relevance in the COMPACT project. They take into account the principles of data minimisation and purpose limitation. Data minimisation refers to processing on the data, which are adequate, limited and necessary for the research purposes, including time limitation on storage and amount of data. Purpose limitation means that processing will be carried out for a specific, explicit and legitimate purpose, i.e. research for the purposes of the COMPACT project, as well as storage for potential further research, which is explained in the Informed Consent Form and Information Sheet.

7.   Other

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

Not defined yet (AIT, CINI, INOV).