



The DELTA project has received funding from the EU's Horizon 2020 research and innovation programme under grant agreement No 773960



# DELTA

Project Acronym: **DELTA**

Project Full Title: **Future tamper-proof Demand rEsponse framework through seLf-configured, self-opTimized and collAborative virtual distributed energy nodes**

Grant Agreement: **773960**

Project Duration: **36 months (01/05/2018 – 30/04/2021)**

## DELIVERABLE D8.2

### DELTA Data Management Plan v1.0

Work Package: **WP8 – Dissemination, Communication & Exploitation activities**

Task: **T8.2 – Exploitation Plan & Policy Guidelines**

Document Status: **Final**

File Name: **[DELTA]\_D8.2\_v1.0**

Due Date: **31.10.2018**

Submission Date: **October 2018**

Lead Beneficiary: **CERTH**

#### Dissemination Level

Public

X

Confidential, only for members of the Consortium (including the Commission Services)

## Authors List

Leading Author				
First Name		Last Name	Beneficiary	Contact e-mail
Dimosthenis		Ioannidis	CERTH	<a href="mailto:djoannid@iti.gr">djoannid@iti.gr</a>
Co-Author(s)				
#	First Name	Last Name	Beneficiary	Contact e-mail
1	Dimitris	Papageorgiou	CERTH	<a href="mailto:dimpapageorg@iti.gr">dimpapageorg@iti.gr</a>
2	Aris	Apostolidis	CERTH	<a href="mailto:arapostol@iti.gr">arapostol@iti.gr</a>
3	Georgios	Karagianopoulos	HIT	<a href="mailto:g.karagiannopoulos@hit-innovations.com">g.karagiannopoulos@hit-innovations.com</a>
4	Raúl García	Castro	UPM	<a href="mailto:rgarcia@fi.upm.es">rgarcia@fi.upm.es</a>
5	Venizelos	Efthymiou	UCY	<a href="mailto:efthymiou.venizelos@ucy.ac.cy">efthymiou.venizelos@ucy.ac.cy</a>
6	Christof	Amann	E7	<a href="mailto:christof.amann@e-sieben.at">christof.amann@e-sieben.at</a>
7	Nikoleta	Andreadou	JRC	<a href="mailto:Nikoleta.ANDREADOU@ec.europa.eu">Nikoleta.ANDREADOU@ec.europa.eu</a>
8	Alexandre	Lucas	JRC	<a href="mailto:Alexandre.Lucas@ec.europa.eu">Alexandre.Lucas@ec.europa.eu</a>
9	Alessio	Baiocco	NTNU	<a href="mailto:alessio.baiocco@ntnu.no">alessio.baiocco@ntnu.no</a>

## Reviewers List

Reviewers			
First Name	Last Name	Beneficiary	Contact e-mail
Ioannis	Moschos	CERTH	<a href="mailto:imoschos@iti.gr">imoschos@iti.gr</a>
Christof	Amann	E7	<a href="mailto:christof.amann@e-sieben.at">christof.amann@e-sieben.at</a>
Alessio	Baiocco	NTNU	<a href="mailto:alessio.baiocco@ntnu.no">alessio.baiocco@ntnu.no</a>

## Legal Disclaimer

The DELTA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 773960. The sole responsibility for the content of this publication lies with the authors. It does not necessarily reflect the opinion of the Innovation and Networks Executive Agency (INEA) or the European Commission (EC). INEA or the EC are not responsible for any use that may be made of the information contained therein.

## Copyright

© CERTH. Copies of this publication – also of extracts thereof – may only be made with reference to the publisher.

## Executive Summary

DELTA is a H2020 research and innovation action funded by the European Commission's Directorate-General for Research and Innovation (DG RTD). The main objective of DELTA is to introduce a DR management platform that distributes parts of the Aggregator's intelligence into lower layers of a novel architecture, based on Virtual Power Plant VPP principles, in order to establish a more easily manageable and computationally efficient DR solution with the ultimate aim to introduce scalability and adaptiveness into the Aggregator's DR toolkits.

The present report is the deliverable D8.2 corresponding to the first version of the project's Data Management Plan (DMP). Throughout the document, the initial versions of the produced datasets are presented along with the following information:

- Type and short description of the datasets to be produced, collected and processed;
- Management processes of research data during and after the completion of the project;
- Standards and formats concerning the metadata; and
- Information about data sharing, exploitation and preservation.

In the light of the required information and taking into account the guidelines for reports related to "Pilot on Open Research Data in Horizon 2020", an appropriate dataset template was created and sent to partners who are responsible for specific components and their corresponding produced datasets. The information of the collected datasets can provide feedback for the specifications of the project's Data Management Portal that will be maintained by CERTH.

As the corresponding tasks and pilot activities progress, the datasets will be elaborated and the information on the templates will be updated. The next version of the deliverable is due to be delivered on M18 including the updated information.

## Table of Contents

<b>1. Introduction.....</b>	<b>7</b>
1.1 Scope and objectives of the deliverable .....	7
1.2 Structure of the deliverable .....	7
1.3 Relation to other activities in the project.....	7
<b>2. General Principles.....</b>	<b>8</b>
2.1 Participation in the Pilot on Open Research Data.....	8
2.1.1 Data Availability and Handling.....	8
2.1.2 Open Access to Scientific Publications.....	9
2.1.3 Open Access to Research Data .....	10
2.2 IPR management and security.....	10
2.3 Data Protection, Ethics and Security .....	10
<b>3. Data Management Framework .....</b>	<b>12</b>
3.1 Format of datasets .....	12
3.2 Data modeling, Standards and meta-data .....	12
3.3 Data Sharing .....	13
3.4 Archiving and preservation (including storage and backup) .....	14
3.5 Datasets List.....	14
<b>4. Description of Datasets .....</b>	<b>15</b>
4.1 Datasets for Aggregator's Operations.....	15
4.2 Datasets for Virtual DELTA Nodes' Operations .....	20
4.3 Datasets from Pilot sites' telemetry devices .....	25
4.4 Collaborative datasets.....	29
4.5 Datasets related to the DELTA Blockchain .....	31
<b>5. Conclusions.....</b>	<b>35</b>
<b>6. References.....</b>	<b>36</b>

---

## List of Figures

<b>Figure 1 DELTA Data Handling .....</b>	<b>9</b>
<b>Figure 2 DELTA Ontology Development Process .....</b>	<b>13</b>

## List of Acronyms and Abbreviations

Term	Description
BEMS	Building Energy Management System
DMP	Data Management Plan
DPO	Data Protection Officer
DS	Data Set
DVN	DELTA Virtual Node
FEID	Fog-Enabled Intelligent Device
IPR	Intellectual Property Rights
LV	Low Voltage
MV	Medium Voltage
NDA	Non-Disclosure Agreement
OA	Open Access
VPP	Virtual Power Plant
KPI	Key Performance Indicator
API	Application Programming Interface
DR	Demand Response
DSS	Decision Support System
WP	Work Package
EC	European Commission
EU	European Union
CA	Consortium Agreement
GDPR	General Data Protection Regulation
UML	Unified Modelling Language
ICT	Information Communication Technology
RAID	Redundant Array of Independent Disks
TCP	Tranmission Control Protocol
Q&A	Quality & Assurance
TRL	Technology Readiness Level
HTTPS	Hypertext Transfer Protocol Secure
SSL	Secure Sockets Layer
CAPTCHA	Completely Automated Public Turing test to tell Computers and Humans Apart
M2M	Machine to Machine
ETSI	European Telecommunications Standards Institute
SAREF	Smart Appliances REference
CoAP	Constrained Application Protocol
MQTT	Message Queuing Telemetry Transport
OpenADR	Open Automated Demand Response
OWL	Ontology Web Language

## 1. Introduction

### 1.1 Scope and objectives of the deliverable

---

This report aims to introduce the general principles and the necessary information for outlining the Data Management Plan (DMP) within DELTA project. The mentioned principles and standards should be considered by DELTA partners during the collection, organization, storage and sharing of the generated datasets. These datasets will be produced within different project's activities.

The document generally presents two types of datasets. The one is the datasets from the pilot sites' meters (raw data) and the second one is the (processed) datasets from the system's developed components. After proper evaluation and post-processing of the datasets (i.e. anonymization if needed), such data can be used for further research by third parties. In this context, the concept of Open Access is described briefly, as it is a quite important concept in research projects.

### 1.2 Structure of the deliverable

---

The document covers the following topics:

- Section 2 presents the general principles for Data Management Plan;
- Section 3 refers to the necessary information for the description of datasets according to Data Management Guidelines for H2020 programs and the corresponding template;
- Section 4 includes the initial version of the identified datasets within DELTA project; and
- Finally, some general conclusions and closing remarks are provided in Section 5.

### 1.3 Relation to other activities in the project

---

The deliverable D8.2 is an outcome of Task 8.2 that is part of WP8 Dissemination, Communication and Exploitation activities. The provided information can be used as input for the respective tasks concerning the exploitation activities and the work packages related to components' development activities. The templates present various exploitable data that should be handled properly. In addition, the dataset type, along with the format, standards and data models can be used as input for the activities and deliverables related to development of specific components during WP3, WP4 and WP5.

## 2. General Principles

### 2.1 Participation in the Pilot on Open Research Data

DELTA participates in the Pilot on Open Research Data launched by the European Commission along with the Horizon2020 program. The members of the DELTA consortium embrace the concepts and the principles of open science and acknowledge the benefits of reusing and evaluating already produced data for promoting and supporting research and innovation projects at European level. The data generated during the project activities may be available in open access for further analysis and exploitation. Towards ensuring the proper usage of data, all the relevant principles for data handling should be described.

#### 2.1.1 Data Availability and Handling

The term of **open access (OA)** concerns the free, online provision of re-useable scientific information to other users. The scope of publically funded research and innovation projects and incentives is to contribute to the improvement of different sectors within society serving environmental, economic and social needs and objectives. The benefits of open access to scientific outcomes are outlined in the “Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020” [1] and are described briefly below:

- Accessible existing scientific publications and results form the basis for efficient peer-to-peer knowledge exchange and improvement of the quality of research reducing the effort and eliminating the duplicated results;
- Connection between market and innovation actions becomes faster and seamless thanks to information flow; and
- Transparency during research is ensured, while facilitating the progress and the dissemination of ideas.

The processes for the identification and management of the exploitable results concerning IPR and patent issues and the individual exploitation plans of each partner will be presented and analyzed in the first version of the deliverable DELTA Exploitation Plan in M24.

Data availability can be divided in three categories, as mentioned below:

- **Open Data:** Data that are publicly shared for re-use and exploitation
- **Consortium:** Confidential data that are available only to the members of the consortium and the EU Commission services and subject to the project Non-Disclosure Agreement (NDA)
- **Private:** Data that are retained by individual partners for their own processes and tests

Within the DELTA project, datasets will be subdivided as follows:

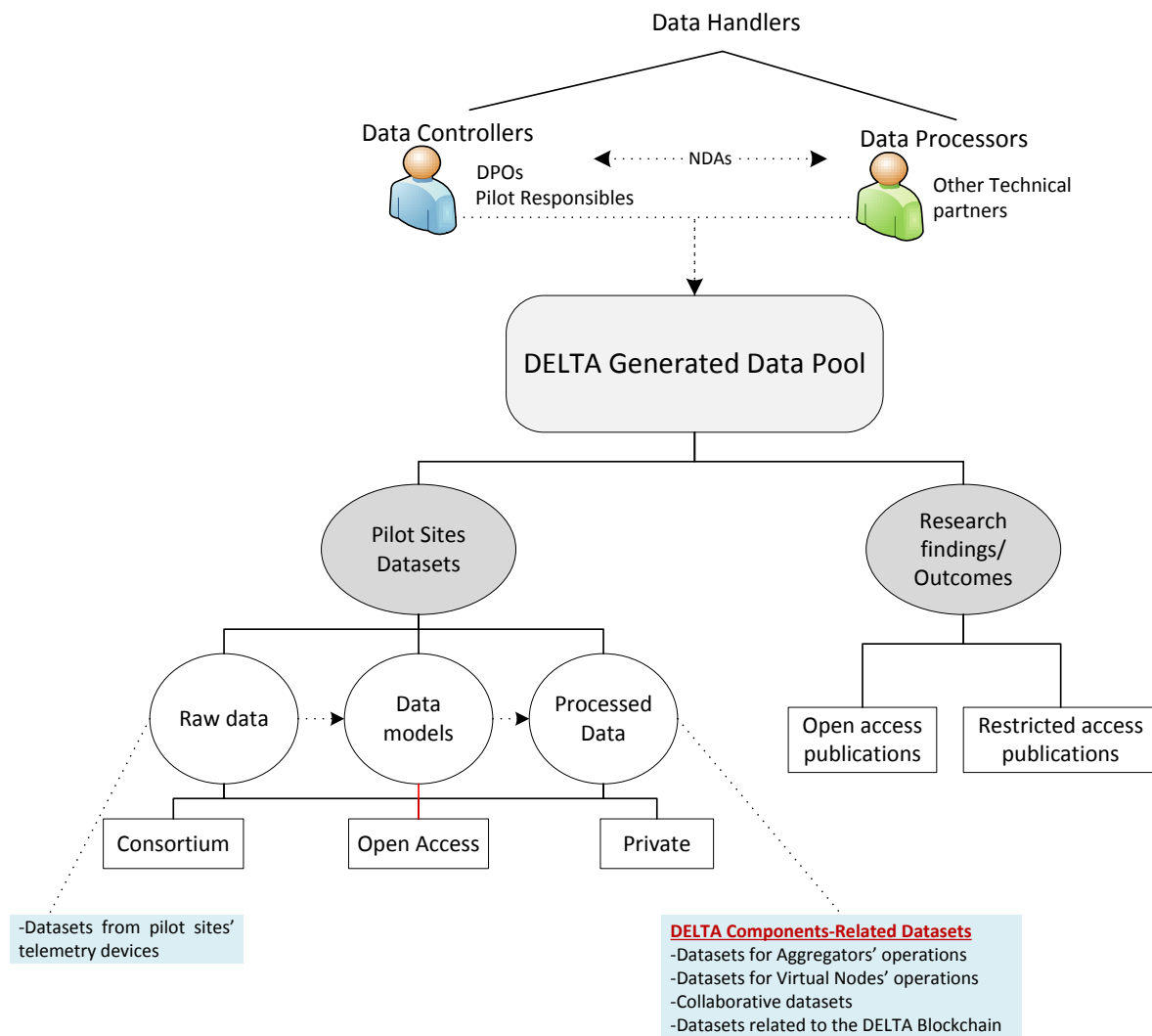
- Pilot sites generated datasets shared between the Consortium partners
- Pilot sites generated datasets that are used for individual partner purposes (Private)
- Pilot sites generated datasets shared to the public (Open Data)
- Research findings and outcomes that should be publicly disseminated (Open Data)

Shared datasets will be allowed between the consortium members for the purpose of fulfilling the project’s objectives, For this reason, and under specified conditions, NDAs would be introduced and signed among involved data processors and controllers before distributing these data within the consortium members. On the other hand, open access to the public may be granted on a fully anonymized dataset, under specified conditions by the consortium members.

DELTA has introduced two main roles concerning data handling activities; the data controllers and the data processors, already described in [4]. The data controllers are the ones imposing the purposes and rules of data processing, while the processors are the ones processing the personal data on behalf of



the controller. In DELTA, the data controllers will be the Pilot Responsibles (KiWi, UCY and CERTH for the Lab –pre-pilot- deployment and tests) and the Data Protection Officers (DPOs), whereas the data processors will be all the technical partners who will need to analyze data from the pilot use cases.



**Figure 1 DELTA Data Handling**

### 2.1.2 Open Access to Scientific Publications

Open access is defined so that each beneficiary must ensure free of charge online access, for any user, to all peer-reviewed scientific publications related to the results extracted from its scientific research. Partners may define the background needed in any manner, and may exclude specific background (not necessarily prior to signature of EC grant agreement). It is possible to grant exclusive licenses to background and foreground if the other partners waive their access rights and depending on previous agreements. The EC may object to exclusive licenses being granted to third parties established in non-associated third countries for ethical, competitiveness or security reasons (where appropriate, a requirement to notify the EC will apply). Partners may agree to additional or more favorable access rights than those provided for in the consortium agreement. At a preliminary stage, partners agreed on open access publishing. However, in the future, partners may opt for gold or green access to peer-reviewed scientific publications, which might result from the project, depending on the type of information to be published.

### 2.1.3 Open Access to Research Data

There are four main aspects of open data summarized in the acronym FAIR [2]:

- **Findable:** Data has a unique, persistent ID, located in a searchable resource, and documented with meaningful metadata.
- **Accessible:** Data are readily and freely retrievable using common methods and protocols, metadata are accessible even if the data are not.
- **Interoperable:** Data are presented in broadly recognized standard formats, vocabularies, and languages.
- **Re-useable:** Data has clear licenses, and accurate meaningful metadata conform to relevant community standards and identifying its content and provenance.

As the project's pilots also concern residential buildings, the datasets from users should be checked if they require aggregation or anonymization for security or commercial reasons prior to release.

## 2.2 IPR management and security

Upon the completion of DELTA project, along with the integrated platform (WP6), a number of software and hardware technological components (WP3-5) will be extracted. During the development of these components, Intellectual Property will be generated that has to be protected through patents, yet made available for other partners for their own work in the project, and exploited outside of the project by appropriate licensing. The management of the project knowledge and IPRs are specified in the Consortium Agreement (CA) that is signed by partners. Its content reflects in some cases complement the terms and conditions defined in the Commission Contractual Rules. More specifically, the CA covers topics such as: Individual and Joint Ownership of the knowledge, Protection of knowledge, Publication of results, Use and dissemination of knowledge arising from the project, access rights, Open Source and Standards, etc.

DELTA projects, as a research and innovation action, addresses low-TRL technologies with the aim to provide market oriented solutions. The project consortium includes many partners from the academic and institutional sector (UCY, JRC, UPM, NTNU and CERTH), as well as private companies (HIT, KiWi, e7 and EAC). These partners will obviously have Intellectual Property Rights on their technologies and data. Therefore, the DELTA consortium should crosscheck with the concerned partners before every publication of data.

A holistic security approach will be followed, in order to protect the pillars of information security (confidentiality, integrity, availability). The security approach will consist of a methodical assessment of security risks followed by their impact analysis. This analysis will be performed on the personal information and data processed by the proposed system, their flows and any risk associated to their processing.

Security measures may include secure protocols (HTTPS and SSL), login procedures, as well as protection about bots and other malicious attacks, as for example CAPTCHA technologies. Moreover, the demonstration pilot sites apply monitored and controlled procedures related to the data collection, their integrity and protection. The data protection and privacy of personal information will include protective measures against infiltration as well as physical protection of ore parts of the systems and access control measures.

## 2.3 Data Protection, Ethics and Security

All the issues related to data protection, data management procedures, ethics and security are described in [4]. During the pilot site activities, all the required procedures will be followed in compliance with National and EU regulations and laws, including but not limited to the GDPR guidelines. For some of the activities to be carried out, there may be the need to collect basic personal

data (e.g. name, contact details etc.). Such type of data will be protected in accordance with the EU Data Protection Directive 95/46/EC [5] concerning the protection of individuals with regard to the processing of personal data and on the free movement of such data. Several EU Member States implemented their national legislation on data protection transposing the EU Data Protection Directive 95/46/EC. The UK (<https://ico.org.uk>) and Cyprus (<http://www.dataprotection.gov.cy>) members have also transposed this directive on their corresponding national legislation.

### 3. Data Management Framework

#### 3.1 Format of datasets

For each dataset the following characteristics will be specified [3]:

DS_XX_ Component	
Data Identification	
Dataset Reference / Name	<Summarized description of the dataset>
Dataset Description	<Mention the produced datasets with a brief description and if they contain future sub-datasets>
Source (e.g. which device?)	<From which device and how the dataset will be collected. Mention also the position of installation>
Relevant Architectural Component	<Mention the name of the architectural component which is linked with the metering device>
Partners services and responsibilities	
Partner in charge of the data collection (if different)	<Partner Name>
Partner in charge of the data analysis (if different)	<Partner Name>
Partner in charge of the data storage (if different)	<Partner Name>
WPs and tasks	<e.g. WP3, T3.4>
Metadata and Expected Size	
Info about metadata (Production and storage dates, places) and documentation	<Provide the status of the metadata, if they are defined and their content>
Estimated volume of data	<Mention the potential data volume>
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	<Purpose of the data collection/generation and its relation to the objectives of the project>

#### 3.2 Data modeling, Standards and meta-data

A semantic analysis of data from metering sensors and field devices will be applied to deliver an ontology, which will be implemented using the OWL ontology language (standardized by the W3C). During the development activities of the project, other relevant standard and generic ontologies will be taken into account, such as the W3C Semantic Sensor Network ontology, the ETSI SAREF ontology or the oneM2M ontology. They will also be made available in XML and JSON format, along with their relevant meta-data.

The modelling language Unified Modelling Language (UML) will be used for the collection, analysis and processing of requirements as well as for the specification message exchanges and overviews of architecture and behavior specifications.

The data models/ontologies created within DELTA will be retained also after the project lifetime for the research community (open-access data) to be able to further extend them in different or similar business scenarios.

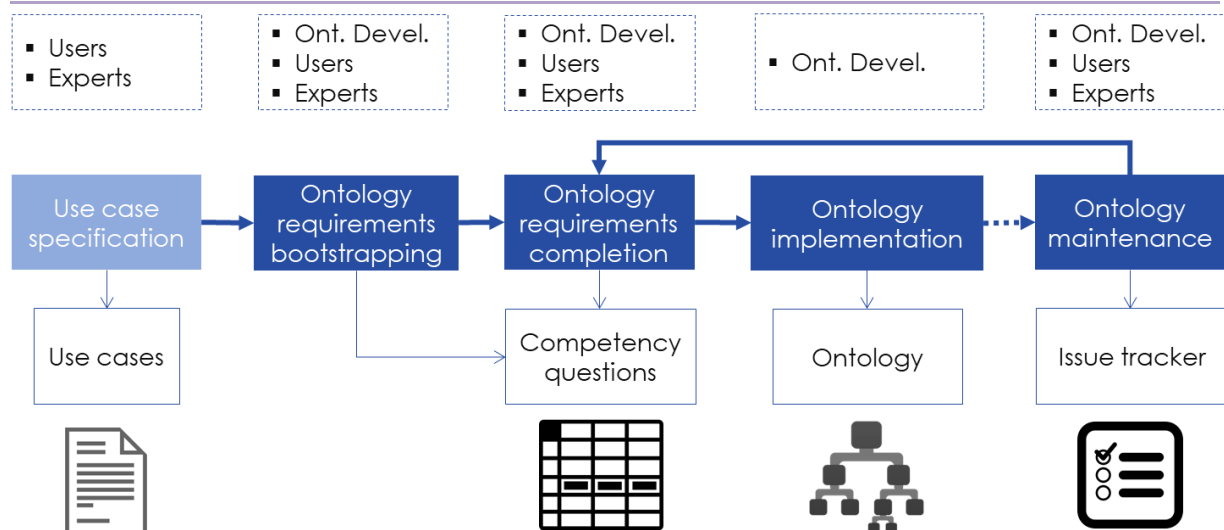


Figure 2 DELTA Ontology Development Process

### 3.3 Data Sharing

The DELTA partners can use a variety of methods for exploitation and dissemination of the data including:

- Using them in further research activities (outside the action)
- Developing, creating or marketing a product or process
- Creating and providing a service
- Using them in standardization activities

The main sharing enabler of the datasets to be produced within the pilot activities will be the DELTA Data Management Portal, created and maintained by CERTH, aligned with Europe's Digital Agenda for better exploiting the potential of Information and Communication Technologies (ICTs) to foster innovation, economic growth and progress<sup>1</sup>. Different access levels within the Data Management Portal will provide the necessary access to both consortium (e.g. internal datasets) and public (open-access) datasets. In particular, data will be made available for researchers in a controlled environment, where authorization and authentication mechanisms will be used to get access to the data in the appropriate format (e.g. XML and/or JSON). Such exchange format will be defined in DELTA T1.3.

Secure application programming interfaces (APIs) will be implemented for delivering a robust information flow among the various DELTA components, through common industry protocols such as CoAP, MQTT and HTTP(S). In case of public datasets, these APIs may also be provided to third parties.

Towards data interoperability, the DELTA information exchange and specifically Demand Response related information will be based on well-known DR-related protocols such as OpenADR (resulting in openADR-compliant data), or other –more general- energy standards such as ETSI SAREF and SAREF4ENER (an extension of SAREF that was created in collaboration with Energy@Home and EEBus).

As for the data models/ontologies which are developed in DELTA, these would be maintained and found through <http://delta.iot.linkeddata.es/>.

<sup>1</sup> <https://ec.europa.eu/digital-single-market/en/europe-2020-strategy>

### 3.4 Archiving and preservation (including storage and backup)

Data will be stored in a secured form (e.g. data encrypted with a strong cryptographic protocol) in servers (via a local instance of the DELTA Repository T1.3) indicated by the pilots or the technology providers. Descriptive metadata will be also stored/provided by such repository. To ensure data reliability, RAID and other common backup technologies may be employed. Local backups of the data, if needed, could be also established by the pilot responsables (e.g. using their SharePoint infrastructure, managed by the Microsoft exchange and 365 servers).

### 3.5 Datasets List

The produced datasets are defined, presuming certain interactions between the various modules of the DELTA architecture. These interactions are bound to change in a later stage of the project, when the tools will be developed and a clear description of the DELTA components interrelations will be available. More specifically, the datasets described in this section are subdivided, in accordance with Figure 1, into: i) datasets belonging to the architecture components of DELTA (DS\_01 - DS\_13 & DS\_18 - DS\_23), ii) Raw datasets from the pilot sensors/meters (DS\_14 – DS\_17) , iii) data related with the data modelling activities of DELTA (DS\_24).

<i><b>DELTA Architectural Element</b></i>	<i><b>Related Task</b></i>	<i><b>Task Leader</b></i>
DS_01_Energy Market Price Forecasting	T4.3	JRC
DS_02_Node Flexibility Data Monitoring and Profiling	T4.4	CERTH
DS_03_Energy Portfolio Segmentation & Classification (DELTA Nodes)	T4.2	HIT
DS_04_DR & Flexibility Forecasting	T4.3	JRC
DS_05_Grid Stability Simulation Engine	T4.1	UCY
DS_06_Self Portfolio Energy Balancing	T4.1	UCY
DS_07_Asset Handling Optimization	T4.4	CERTH
DS_08_Consumer/Prosumer Flexibility Data Monitoring and Profiling	T3.2	HIT
DS_09_Load Forecasting	T3.2	HIT
DS_10_Consumer/Prosumer Clustering	T3.3	CERTH
DS_11_Generation/Consumption Optimal Dispatch	T3.2	HIT
DS_12_Inter/Intra Node Energy Matchmaking	T3.2	HIT
DS_13_Lightweight Toolkit (FEID)	T3.4	CERTH
DS_14_UCY_Pilot_BEMS	T7.3, T7.4	UCY
DS_15_UCY_Pilot_Smart Meters	T7.3, T7.4	UCY
DS_16_KiWi_Pilot_IndoorQualitySensors	T7.3, T7.4	KiWi
DS_17_KiWi_Pilot_Energy Smart Meters	T7.3, T7.4	KiWi
DS_18_Aware-enabled Energy Behavioral Platform	T6.4	CERTH
DS_19_Social Interaction and Cooperation	T6.4	CERTH
DS_20_Demand Response Visualisation Kit	T6.3	CERTH
DS_21_DELTA Blockchain	T5.2	NTNU
DS_22_Smart Contracts	T5.2	NTNU
DS_23_Threat Mitigation	T5.3	NTNU
DS_24_DELTA Information Modeling	T1.3	UPM

## 4. Description of Datasets

### 4.1 Datasets for Aggregator's Operations

In this Section, the dataset comprising the information exchanged at the Aggregator's level is described. The Aggregator would have to communicate with market and grid stakeholders for certain business operations and long-term decision planning, as well as with its customers/prosumers, through the Virtual Nodes and FEIDs. Such information would be expressed through the Aggregator's toolkit components seen below.

<b>DS_01_Energy Market Price Forecasting</b>	
<b>Data Identification</b>	
Dataset Reference / Name	Sell price of the DR bids to the markets
Dataset Description	<u>Input datasets:</u> Demand historical data; season/day; weather forecast; fuel price; price elasticity; history of transactions <u>Output datasets:</u> Forecasted market price (day-ahead, intra-day, weekly)
Source (e.g. which device?)	In: To forecast the market price for a future time frame (out), historical and current data would be extracted from the database (Repository), Aggregator, web services, Market stakeholders
Relevant Architectural Component	Energy Market Price Forecasting
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilots Responsible Partners (KiWi, UCY, CERTH)
Partner in charge of the data analysis (if different)	JRC, Technical Partners
Partner in charge of the data storage (if different)	Pilots Responsible Partners (KiWi, UCY, CERTH)
WPs and tasks	The data are going to be collected within activities of WP4 and more specifically within activities of T4.3
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	A detailed description of the dataset contents will be included (e.g. regarding the process for energy market price evaluation).
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be less than 10 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	Dataset will be used for defining the DR capacity to be sold to the market (through the Self portfolio balancing module T4.1)

<b>DS_02_Node Flexibility Data Monitoring and Profiling</b>	
<b>Data Identification</b>	
Dataset Reference / Name	Energy flexibility-related data coming from Nodes and MV Assets & profiles' generation
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets:</u> Available energy flexibility data ranges from each Node and consumption/generation</li> </ul>



	<p>data from MV Assets</p> <ul style="list-style-type: none"> <li>• <u>Output Datasets</u>: Profiling of Node clusters from the Aggregator's perspective.</li> </ul>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>• In: To evaluate and align the profiles of the Nodes and MV Assets from the Aggregator's perspective (out), data would be extracted from the Virtual Delta Nodes and the MV Assets (smart meters) as well (via implemented communication links/APIs)</li> </ul>
Relevant Architectural Component	Node Flexibility Data Monitoring and Profiling
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
Partner in charge of the data analysis (if different)	CERTH, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of WP4 and more specifically within activities of T4.4
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Dataset will be accompanied with a detailed documentation of its contents. Indicative metadata may include: (a) description of the pilot setup (e.g. location, date, etc.) and procedure that led to the generation of the dataset, (b) specific event and state of this event, etc.
Estimated volume of data	<ul style="list-style-type: none"> <li>• Volume of data will depend on the: <ul style="list-style-type: none"> <li>○ Number of Nodes participating in DR</li> <li>○ data exchange intervals</li> <li>○ meta-data</li> </ul> </li> <li>• Such data are expected to be in the range of 10-70 kb per transmission</li> </ul>
<b>Data exploitation and sharing</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used for profiling purposes of the Aggregator.

### DS\_03\_Energy Portfolio Segmentation & Classification (DELTA Nodes)

<b>Data Identification</b>	
Dataset Reference / Name	Clusters of Aggregator's portfolio into Nodes and groups of large customers
Dataset Description	<ul style="list-style-type: none"> <li>• <u>Input Datasets</u>: Energy and Flexibility profiles from the Nodes and FEIDs</li> <li>• <u>Output Datasets</u>: Initial allocation of FEIDs into Nodes</li> </ul> <p>Dataset for clustering that refers to the initial allocation of FEIDs into Nodes based on a simulation performed by the Aggregator at their premises. This allocation would be re-calculated at the Aggregator level, when a new customer is introduced into the system (addition of new customers and removal of customers that leave its portfolio (allocation of smart contracts etc.))</p>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>• In: To evaluate the portfolio segmentation (out),</li> </ul>



	data would be extracted from the Nodes & FEIDs (energy patterns/profiles), Aggregator (KPIs-related)
Relevant Architectural Component	Energy Portfolio Segmentation & Classification
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
Partner in charge of the data analysis (if different)	HIT, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of WP4 and more specifically within activities of T4.2
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	A detailed description of the dataset contents will be included (e.g. what is the process for allocation of FEIDs into Nodes)
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of FEIDs</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be less than 5 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used for grouping existing and new customers into Nodes (large clusters of prosumers).

## DS\_04\_DR & Flexibility Forecasting

<b>Data Identification</b>	
Dataset Reference / Name	DR & Flexibility Forecasting for all Assets into the Aggregator's portfolio
Dataset Description	<p><u>Input Datasets:</u> MV Assets generation/consumption data, Nodes flexibility data</p> <p><u>Output Datasets:</u> flexibility capacity range for Nodes and MV Assets, seen from the Aggregator's perspective</p> <p>Produced dataset resulting from computing the MV Assets (existing portfolio) flexibility capacity range and the sum of them calculated by the Aggregator.</p>
Source (e.g. which device?)	In: To evaluate the flexibility range for all Aggregator's portfolio (out), data would be extracted from the MV Assets (via smart meters' APIs) and Nodes
Relevant Architectural Component	DR & Flexibility Forecasting
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
Partner in charge of the data analysis (if different)	JRC, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of

	WP4 and more specifically within activities of T4.3.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	A detailed description of the dataset contents will be included (e.g. what is the process for prediction of the portfolio's available flexibility capacity range).
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of Nodes and MV Assets</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be in the range of 1-10 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	Dataset will be used for the self-portfolio balancing (nodes and FEIDs unit commitment)

## DS\_05\_Grid Stability Simulation Engine

<b>Data Identification</b>	
Dataset Reference / Name	Grid Stability Data
Dataset Description	<ul style="list-style-type: none"> <li>In: Smart meters and FEIDs</li> <li>Out: probability of grid instability based on flexibility capacity exchange on the grid bus</li> </ul> Responsible for monitoring the stability (e.g. voltage fluctuations) within the available portfolio and run background scenarios simulating issues that could occur based on current status.
Source (e.g. which device?)	In: Smart meters & FEIDs Data will be accessed from within the DELTA framework. Data from all components in the Aggregator's toolkit will be used for analysis.
Relevant Architectural Component	Grid Stability Simulation Engine
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/KiWi, CETH)
Partner in charge of the data analysis (if different)	UCY, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CETH)
WPs and tasks	The data are going to be collected within activities of WP4 and more specifically within activities of T4.1.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	A detailed description of the dataset contents will be included (e.g. on which bus we perform the measurements, etc.).
Estimated volume of data	Assuming the data are collected at an aggregated and half-hourly timestamped level the data volume will be less than 1 MB per day, although this will scale linearly with the number of individual nodes if individual targets are specified as a requirement of the aggregator's communication protocol.
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data)	Optimum load profiles at the aggregator level will be

analysis)	used to inform aggregator pricing and signaling strategies in response to market forces. Probabilistic predictions of grid level events that can influence market forecasts will also be provided and limitations of local participation from the flexibility and physical perspective defined.
-----------	---

## DS\_06\_Self Portfolio Energy Balancing

### Data Identification

Dataset Reference / Name	Energy flexibility capacity and Nodes commitment allocation
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets</u>: Flexibility capacity range vs cost range per capacity &amp; Smart Contracts' conditions</li> <li><u>Output Datasets</u>: Total flexibility capacity bids, sell price and Nodes commitment allocation, to be sent to the Market Stakeholders. Also, price incentives, rewards and penalties per Customer for capacity delivered</li> </ul>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li><u>In</u>: to evaluate the Nodes' flexibility capacity allocation, data will be extracted from the Virtual Delta Nodes, MV Assets included in Aggregator's portfolio &amp; FEIDs</li> </ul> <p>The dataset will be collected from the services provided at Virtual Delta Node level and/or FEID-level information (e.g. smart contracts' constraints).</p>
Relevant Architectural Component	Self-Portfolio Energy Balancing

### Partners services and responsibilities

Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
Partner in charge of the data analysis (if different)	CERTH, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of WP4 and more specifically within activities of T4.1

### Metadata and Expected Size

Info about metadata (Production and storage dates, places) and documentation?	Dataset will be accompanied with a detailed documentation of its contents. Indicative metadata may include: (a) description of the pilot setup (e.g. location, date, etc.) and procedure that led to the generation of the dataset, (b) specific event and state of this event, etc.
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of Nodes and FEIDs participating in DR</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be in the range of 5-30 kb per transmission</li> </ul>

### Data exploitation

Data exploitation (purpose/use of the data analysis)	The datasets will be used for the evaluation of the energy assets' flexibility provision in the pilot sites.
--	--

<b>DS_07_Asset Handling Optimization</b>	
<b>Data Identification</b>	
Dataset Reference / Name	Balancing result for excess or shortage of energy for large customers and Nodes
Dataset Description	<ul style="list-style-type: none"> <li>○ <u>Input Datasets:</u> Energy flow from the Nodes and the MV Assets</li> <li>○ <u>Output Datasets:</u> Energy flow re-allocation</li> </ul> <p>It will provide the DSS with assessments of the available energy (excess or shortage) within the Aggregator portfolio, targeting a balanced state where the use of available flexibility is maximized.</p>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>• In: MV Assets (e.g. smart meters via implemented APIs) and Nodes</li> </ul> <p>Data will be accessed from within the DELTA framework. Data from all components in the Aggregator's toolkit will be used for analysis.</p>
Relevant Architectural Component	Asset Handling Optimization
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	CERTH
Partner in charge of the data analysis (if different)	UCY, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of WP4 and more specifically within activities of T4.4
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Timestamped projected standard and optimized flexibility profiles from analysis of T3.2 data. Presented at the aggregated level.
Estimated volume of data	This is dependent on the level of information available from T3.2. The format will be timestamped data on demand and flexibility projections as is with enhanced flexibility limits presented from the optimization procedure.
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The data will be used to identify asset-handling strategies that are appropriate given the projected profiles and flexibility calculations.

#### 4.2 Datasets for Virtual DELTA Nodes' Operations

In this Section, the dataset comprising the information exchanged at the Nodes' (and FEIDs) level is described. The Node will be the intermediary virtual actor between the Aggregator and the customer, therefore each Node should handle information regarding Aggregator's requests and energy-related data from each prosumer, via the FEIDs. Such information would be expressed through the Nodes' toolkit components seen below.

<b>DS_08_Consumer/Prosumer Flexibility Data Monitoring &amp; Profiling</b>
<b>Data Identification</b>

Dataset Reference / Name	Energy profiles' dataset at the Node level
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets</u>: Available energy flexibility data &amp; profiles from each FEID</li> <li><u>Output Datasets</u>: Virtual Node energy profiles</li> </ul> <p>Monitor data coming from FEIDs and create the Node-level profiles.</p>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>In: Fog-Enabled Devices (via implemented APIs) at prosumers' side</li> </ul> <p>The dataset will be collected using FEIDs at prosumers' side.</p>
Relevant Architectural Component	Consumer/Prosumer Flexibility Data Monitoring & Profiling
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/ KiWi, CERTH)
Partner in charge of the data analysis (if different)	HIT, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of WP3 and more specifically within activities of T3.2
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Metadata are yet to be defined by the solution provider.
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of Nodes and FEIDs participating in DR</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be in the range of 10-70 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used to later-on create the Nodes' clusters

## DS\_09\_Load Forecasting

### Data Identification

Dataset Reference / Name	Dataset for day-ahead and intra-day energy consumption for each Node
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Dataset</u>: energy consumption data from FEIDs</li> <li><u>Output Dataset</u>: forecasted energy consumption per Node cluster</li> </ul>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>In: Fog-Enabled Device (via implemented APIs) at prosumers'-side</li> </ul>
Relevant Architectural Component	Load Forecasting
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partner e.g. UCY/KIWI
Partner in charge of the data analysis (if different)	

different)	HIT, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	WP3, T3.2
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	A detailed description of the dataset contents will be included (e.g. different granularities, accuracy etc.)
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of FEIDs exchanging data</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be less than 100 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets can be used for identifying power excess or shortage per Node and examine Node stability.

## DS\_10\_Consumer/Prosumer Clustering

<b>Data Identification</b>	
Dataset Reference / Name	Prosumers' clusters
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets</u>: KPIs from the Aggregator &amp; Energy patterns/profiles from the FEIDs</li> <li><u>Output Datasets</u>: Prosumers' clusters</li> </ul> <p>Collections of prosumers will be formed into Nodes, according to their energy patterns and other similar characteristics (i.e. KPIs)</p>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>In: FEIDs and Aggregator's KPIs</li> </ul> <p>The data will be collected from FEIDs at the prosumers' premises and also from Aggregators.</p>
Relevant Architectural Component	Consumer/Prosumer Clustering
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/ KiWi, CERTH)
Partner in charge of the data analysis (if different)	CERTH, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of WP3 and more specifically within activities of T3.3.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Metadata are yet to be defined by the solution provider.
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of FEIDs exchanging data</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be less than 5 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data)	The datasets will be used for grouping existing and new

analysis)	customers into Nodes.
-----------	-----------------------

## DS\_11\_Generation/Consumption Optimal Dispatch

### Data Identification

Dataset Reference / Name	Dataset for next-day and intra-day flexibility capacity range for each Node
Dataset Description	<ul style="list-style-type: none"> <li>○ <u>Input Dataset:</u> Energy profiles and flexibility capacity range from FEIDs</li> <li>○ <u>Output Dataset:</u> Flexibility plan for next-day and intra-day per Node cluster</li> </ul>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>• In: Fog-Enabled Device (via implemented APIs) at prosumers'-side</li> </ul>
Relevant Architectural Component	Generation/Consumption Optimal Dispatch

### Partners services and responsibilities

Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
Partner in charge of the data analysis (if different)	HIT, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)
WPs and tasks	WP3, T3.2

### Metadata and Expected Size

Info about metadata (Production and storage dates, places) and documentation?	A detailed description of the dataset contents will be included (e.g. which clusters of FEIDs will follow the specific plan etc.).
Estimated volume of data	<ul style="list-style-type: none"> <li>• Volume of data will depend on the: <ul style="list-style-type: none"> <li>○ Number of customers/FEIDs</li> <li>○ data exchange intervals</li> <li>○ included meta-data</li> </ul> </li> <li>• Such data would be on average ~5-30 kb per transmission</li> </ul>

### Data exploitation

Data exploitation (purpose/use of the data analysis)	The datasets will be used for computing the future state flexibility capacity range
--	---

## DS\_12\_Inter/Intra Node Energy Matchmaking

### Data Identification

Dataset Reference / Name	Dataset for Energy excess / shortage per each Node
Dataset Description	<ul style="list-style-type: none"> <li>• <u>Input Dataset:</u> Energy measurements from FEIDs / FEID clusters</li> <li>• <u>Output Dataset:</u> Energy excess / shortage &amp; assignment of them to another Node cluster / FEID</li> </ul>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>• In: Fog-Enabled Device (via implemented APIs) at prosumers'-side</li> </ul>
Relevant Architectural Component	Inter/Intra Node Energy Matchmaking

### Partners services and responsibilities

Partner in charge of the data collection (if different)	Pilot Responsible Partner e.g. UCY/KIWI
Partner in charge of the data analysis (if different)	HIT, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partner (UCY/KiWi, CERTH)



WPs and tasks	WP3, T3.2
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Dataset will be accompanied with a detailed documentation of its contents. Indicative metadata may include: (a) description of the experimental setup (e.g. location, date, etc.) and procedure that led to the generation of the dataset, (b) specific event and state of this event, involved humans per time interval, etc.
Estimated volume of data	<ul style="list-style-type: none"> <li>Format not yet defined (e.g. csv, XML, JSON etc.): to be delivered in DELTA Information Modelling T1.3</li> <li>openADR-based or compliant schema for data transfer</li> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of FEIDs, Nodes exchanging data</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be less than 100 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets can be used for facilitating self-balancing process within the Node.

## DS\_13\_FEID

### Data Identification

Dataset Reference / Name	Energy flexibility dataset at the prosumer side & actuation signals
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Dataset</u>: close to real-time LV Assets' (connected to FEIDs) energy measurements, prosumers' KPIs-related data (e.g. comfort) and actuation signals from the Nodes/Aggregator</li> <li><u>Output Dataset</u>: Available energy flexibility data &amp; energy patterns/profiles to be delivered to each Virtual Node. Also, Actuation signals to prosumer's Assets</li> </ul> <p>Produced dataset will be used to support optimal DR requests' distribution to each Prosumer's Assets.</p>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>In: LV Assets, Smart meters, sensors, BMS (via implemented APIs) at prosumers'-side</li> </ul> <p>The FEIDs will collect energy and related information from their linked Assets and systems at the building/facility level.</p>
Relevant Architectural Component	Lightweight Toolkit per Customer and DELTA Fog Enabled Agent
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partner (UCY/ KiWi, CERTH)
Partner in charge of the data analysis (if different)	CERTH, Technical Partners
Partner in charge of the data storage (if different)	CERTH, Pilot Responsible Partner



WPs and tasks	The data are going to be collected within activities of WP3 and more specifically within activities of T3.4.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	A detailed description of the dataset contents will be included (e.g. what is the process of energy flexibility evaluation).
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of Assets participating in each FEID-enabled cluster</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be less than 10 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used for the evaluation of the energy assets' flexibility provision in the pilot sites.

#### 4.3 Datasets from Pilot sites' telemetry devices

In this Section, the dataset comprising the information stemming from the Pilot sites is described. Such data will be made available from smart metering equipment and intelligent monitoring equipment (such as BEMS), seen below.

<b>DS_14_UCY_Pilot_BEMS</b>	
<b>Data Identification</b>	
Dataset Reference / Name	UCY BEMS Data
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets</u>: FEID actuation signals &amp; requests for data extraction</li> <li><u>Output Datasets</u>: Building monitoring data</li> </ul> <p>The UCY campus operates Building Energy Management Systems that automate the controls of building specific variable loads. The parameters from these systems will be extracted and stored as datasets for use in the DELTA project</p>
Source (e.g. which device?)	Installed BEMS systems from Siemens, Honeywell and/or Johnson Controls, the actual BEMS candidates are yet to be finalized in the DELTA project. Source of data requests will be the FEIDs.
Relevant Architectural Component	Fog Enabled Agent
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	UCY
Partner in charge of the data analysis (if different)	UCY, Technical Partners
Partner in charge of the data storage (if different)	UCY
WPs and tasks	The data are going to be collected within activities of WP7 and more specifically within activities of T7.3, T7.4.
<b>Metadata and Expected Size</b>	

Info about metadata (Production and storage dates, places) and documentation?	Metadata will include building use type and occupancy information. Ongoing discussions are being held with the BEMS manufacturers for the exact specification of the available monitoring data. Time dependent set point and actual temperatures for heating and cooling control for user comfort requirements are the principle parameters.
Estimated volume of data	This will be confirmed once the discussions with the BEMS manufacturers are complete. It is an ongoing process.
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The collected data will be used to document historic occupancy and comfort parameters, set point and actual temperatures. Associated load profiles will also be recorded. The data will be used to assert historic and extrapolate future demand and flexibility profiles of the buildings in question.

## DS\_15\_UCY\_Pilot\_Energy Smart Meters

<b>Data Identification</b>	
Dataset Reference / Name	Building Power Generation, Usage & Quality
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets</u>: FEID requests for data extraction</li> <li><u>Output Datasets</u>: monitoring data</li> </ul> <p>Dataset for energy profiling and power quality assessment. Timestamped current, voltage, frequency, real and reactive power and/or power factor measurements will be measured at various points in the pilot site (yet to be determined) will one meter installed per building. Data will be stored for analysis and communication with the DVN.</p>
Source (e.g. which device?)	The data will be collected using Schneider Electric PowerLogic ION power quality meters communicating over Modbus TCP. Source of data requests will be the FEIDs.
Relevant Architectural Component	-
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	UCY
Partner in charge of the data analysis (if different)	UCY, Technical Partners
Partner in charge of the data storage (if different)	UCY
WPs and tasks	The data are going to be collected within activities of WP7 and more specifically within activities of T7.3, T7.4.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Each entry will be recorded minutely using average values determined from one second spot value measurements. The meters will be distributed throughout buildings the UCY campus that form part of the UCY pilot site.

Estimated volume of data	This will be confirmed once the full dataset, data format and communication protocols are asserted. At first approximation, it will be 256-bits per row, one row per minute, some overheads for metadata. Approximately 0.5 MB per day per meter. This will be confirmed once the meters have been installed and tested.
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The collected data will be used to document historic generation, consumption and quality profiles and to extrapolate from these future profile predictions.

## DS\_16\_KiWi\_Pilot\_IndoorQualitySensors

<b>Data Identification</b>	
Dataset Reference / Name	KiWi indoor quality data set
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets:</u> FEID or DVN actuation signals &amp; requests for data extraction</li> <li><u>NOD devices:</u> indoor comfort monitoring devices, providing data on Temperature, relative humidity, CO2, VOX and room occupancy</li> <li><u>Output Datasets:</u> Building monitoring data</li> </ul>
Source (e.g. which device?)	Installed NOD devices (Moor House, Ernest Dence and other residential estates in Greenwich)
Relevant Architectural Component	Fog Enabled Agent
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	KiWi
Partner in charge of the data analysis (if different)	KiWi, Technical Partners
Partner in charge of the data storage (if different)	KiWi
WPs and tasks	The data are going to be collected within activities of WP7 and more specifically within activities of T7.3, T7.4.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Metadata will include building use type and occupancy information. The area monitored via NOD equipment will be mapped out on building plans.
Estimated volume of data	As an estimate, a CSV file export containing energy consumption data on a minute by minute level for 365 days on a single metering point is around 9MB
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The collected data will be used to document historic occupancy and comfort parameters, set point and actual temperatures. Associated load profiles will also be recorded. The data will be used to assert historic and extrapolate future demand and flexibility profiles of the buildings in question.

## DS\_17\_KiWi\_Pilot\_Energy Smart Meters

Data Identification	
Dataset Reference / Name	Building Power Generation, Usage & Quality
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets:</u> FEID or DVN requests for data extraction</li> <li><u>Output Datasets:</u> monitoring data</li> </ul> <p>Dataset for energy profiling and power quality assessment. Timestamped current, voltage, frequency, will be measured at various points in the pilot site (Moor House) for relevant energy intensive assets (e.g. 6 x chillers). Data will be stored for analysis and communication with the DVN.</p>
Source (e.g. which device?)	<p>Metering data sources will be:</p> <ul style="list-style-type: none"> <li>KiWi Power Fruit – metering and communication device for all commercial and industrial assets (eg. Chillers @ Moor House)</li> <li>Schneider sub metering system for HH data (entire building and sub-zones @ Moor House)</li> <li>Hildebrand Glow CT clamps and data hubs for residential data (via Hildebrand &amp; KiWi back end systems through dedicated APIs)</li> </ul> <p>Where feasible, source of data requests will be the FEIDs OR Virtual Nodes interfacing with KiWi Operations Management Platform via APIs</p>
Relevant Architectural Component	-
Partners services and responsibilities	
Partner in charge of the data collection (if different)	KiWi
Partner in charge of the data analysis (if different)	KiWi, Technical Partners
Partner in charge of the data storage (if different)	KiWi
WPs and tasks	The data are going to be collected within activities of WP7 and more specifically within activities of T7.3, T7.4.
Metadata and Expected Size	
Info about metadata (Production and storage dates, places) and documentation?	Each entry will be recorded minutely using average values determined from one second spot value measurements. Each metering point will have a unique identifier and the impact area mapped out on building plans
Estimated volume of data	As an estimate, a CSV file export containing energy consumption data on a minute by minute level for 365 days on a single metering point is around 9MB
Data exploitation	
Data exploitation (purpose/use of the data analysis)	The collected data will be used to document historic generation, consumption and quality profiles and to extrapolate from these future profile predictions.

#### 4.4 Collaborative datasets

In this Section, the dataset regarding the information exchanged between customers on a daily basis, through the DELTA collaboration services, as well as the demand response rewards exchange mechanisms are described.

<b>DS_18_Award-enabled Energy Behavioral Platform</b>	
<b>Data Identification</b>	
Dataset Reference / Name	Assign rewards to customers
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets:</u> Demand flexibility profiles &amp; Smart contracts' conditions and fulfillment rate</li> <li><u>Output Datasets:</u> Respective reward points to the customers (monetary or behavioral based)</li> </ul> <p>Datasets will be produced by the reward point system that will be integrated into the smart contracts framework. The indicators included in each smart contract will be the points earned/kW of DR enabled capacity.</p>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li><u>In:</u> Customers' requested flexibility rate that has been fulfilled (gathered through FEIDs)</li> </ul>
Relevant Architectural Component	Award-enabled energy behavioral platform
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
Partner in charge of the data analysis (if different)	CERTH, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of WP6 and more specifically within activities of T6.4.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	The reward points will be translated in social collaboration and awards-based gamification schemes.
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of customers</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be less than 10 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used for enhancing key characteristics for identification of customers and improving end-user engagement.

<b>DS_19_Social Interaction and Cooperation</b>	
<b>Data Identification</b>	
Dataset Reference / Name	Dataset for knowledge diffusion among customers (consumers/prosumers)

Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets:</u> Customers' related data (e.g. UserID) &amp; Awards</li> <li><u>Output Datasets:</u> Exchange of relevant information between customers that are actively being engaged to DR services (including multimedia content/Q&amp;A/chat/notifications etc.)</li> </ul> <p>Produced datasets will concern suggestions and incentives, timeline of customers activities, social connections among customers etc.</p>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li><u>In:</u> Customers</li> </ul>
Relevant Architectural Component	Social Interaction and Cooperation
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partners (UCY, KiWi, CETH)
Partner in charge of the data analysis (if different)	CETH, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partners (UCY, KiWi, CETH)
WPs and tasks	The data are going to be collected within activities of WP6 and more specifically within activities of T6.4.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Profiles of customers and collaborative models will be generated indicating the dependencies from energy related events.
Estimated volume of data	<ul style="list-style-type: none"> <li>Format not yet defined (e.g. csv, XML, JSON etc.): to be delivered in DELTA Information Modelling T1.3</li> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of customers, activities and DR requests</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> <li>Such data would be less than 10 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used for promoting best practices from customers and healthy competition among customers (that will be extensively researched regarding the impact to energy markets).

## DS\_20\_Demand Response Visualization Kit

### Data Identification

Dataset Reference / Name	Visualized dataset of real-time and historical energy information
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets:</u> Energy flexibility and profile data from FEIDs</li> <li><u>Output Datasets:</u> Real-time and historical energy information</li> </ul> <p>Produced dataset will include information for consumption, generation, flexibility, emissions and they will be accompanied by Demand Response information (e.g. available DR strategies, DR cost and energy</p>

	estimations, rewards etc.)
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>• <u>In</u>: FEIDs at prosumer's side</li> </ul> <p>The dataset will be collected through FEIDs.</p>
Relevant Architectural Component	Demand Response Visualization Kit
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
Partner in charge of the data analysis (if different)	CERTH, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
WPs and tasks	The data are going to be collected within activities of WP6 and more specifically within activities of T6.3.
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Visualized dataset will be accompanied with detailed documentation (e.g. procedure that led to the generation of dataset, specific event and state of this event etc.) in order to inform properly the involved stakeholders.
Estimated volume of data	<ul style="list-style-type: none"> <li>• Volume of data will depend on the: <ul style="list-style-type: none"> <li>○ Number of customers</li> <li>○ data exchange intervals</li> <li>○ included meta-data</li> </ul> </li> <li>• Such data would be less than 100 kb per transmission</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used to present to each actor the respective information of DR events (e.g. customers will be able to see information about the DR applied on them, but not on the DVN etc.)

#### 4.5 Datasets related to the DELTA Blockchain

In this Section, the data that is sent/logged via the DELTA blockchain network are described.

<b>DS_21_DELTA Blockchain</b>	
<b>Data Identification</b>	
Dataset Reference / Name	Verified energy network data after blockchain implementation
Dataset Description	<ul style="list-style-type: none"> <li>• <u>Input Datasets</u>: updated distributed ledger data created from DELTA stakeholders and DELTA blockchain-based transactions (log-oriented architecture)</li> <li>• <u>Output Datasets</u>: requested distributed ledger data created from DELTA stakeholders and DELTA blockchain-based transactions (log-oriented architecture)</li> </ul> <p>Produced verified dataset will enable energy data traceability and secure access for stakeholders.</p>



Source (e.g. which device?)	<ul style="list-style-type: none"> <li>In/Out: Nodes creating the ledgers (data logging) via Apps (e.g. Demand Response Visualization Kit) and visualizing the information included in the distributed ledger</li> </ul> <p>The dataset will be stored in the DELTA distributed ledger.</p>
Relevant Architectural Component	DELTA Cyber Security Services – Smart Contracts
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
Partner in charge of the data analysis (if different)	NTNU, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
WPs and tasks	WP5, T5.2
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Metadata are yet to be defined by the solution provider.
Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of customers/FEIDs</li> <li>data exchange intervals</li> </ul> </li> <li>included meta – data</li> </ul>
<b>Data exploitation</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used by users for usage tracking and aggregation of data related to him/her. In addition, the produced dataset will be used for decision-making and risk assessment among stakeholders.

## DS\_22\_Smart Contracts

<b>Data Identification</b>	
Dataset Reference / Name	Dataset related to smart contract implementation
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets</u>: behavioral rulesets for all participants of the smart contract,</li> <li><u>Output Datasets</u>: contract checking output (e.g. promised vs delivered flexibility) and rewards/penalties execution.</li> </ul>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>In: DELTA Blockchain</li> </ul>
Relevant Architectural Component	DELTA Cyber Security Services – Smart Contracts
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
Partner in charge of the data analysis (if different)	NTNU, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
WPs and tasks	WP5, T5.2
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	Metadata are yet to be defined by the solution provider.



Estimated volume of data	<ul style="list-style-type: none"> <li>Volume of data will depend on the: <ul style="list-style-type: none"> <li>Number of customers/FEIDs/rules</li> <li>data exchange intervals</li> <li>included meta-data</li> </ul> </li> </ul>
<b>Data exploitation and sharing</b>	
Data exploitation (purpose/use of the data analysis)	The datasets will be used for improving DR strategies and enhancing end-user engagement.

## DS\_23\_Threat Mitigation

### Data Identification

Dataset Reference / Name	Dataset for cyber security in energy networks
Dataset Description	<ul style="list-style-type: none"> <li><u>Input Datasets</u>: history of attacks, network vulnerabilities</li> <li><u>Output Datasets</u>: threat mitigation rules</li> </ul>
Source (e.g. which device?)	<ul style="list-style-type: none"> <li>In: FEIDs and monitoring devices</li> </ul> <p>The dataset will be collected through FEIDs and monitoring devices.</p>
Relevant Architectural Component	DELTA Cyber Security Services – Threat Mitigation

### Partners services and responsibilities

Partner in charge of the data collection (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
Partner in charge of the data analysis (if different)	NTNU, Technical Partners
Partner in charge of the data storage (if different)	Pilot Responsible Partners (UCY, KiWi, CERTH)
WPs and tasks	WP5, T5.3

### Metadata and Expected Size

Info about metadata (Production and storage dates, places) and documentation?	Dataset will be accompanied with a detailed documentation of its contents. Indicative metadata may include: (a) description of the experimental setup (e.g. location, date, etc.) and procedure that led to the generation of the dataset, (b) specific event and state of this event, involved humans per time interval, etc.
Estimated volume of data	<ul style="list-style-type: none"> <li>Data volume In = Energy network conditions related data</li> <li>Data volume Out = Risk Assessment (= Threat x Vulnerability x Impact)</li> </ul>

### Data exploitation

Data exploitation (purpose/use of the data analysis)	The produced dataset can be used for supporting life-cycle components requirements in terms of updates etc. In addition, they will enable the safety- oriented management of the entire system.
--	---

## DS\_24\_DELTA Information Modeling

### Data Identification

Dataset Reference / Name	DELTA Information Model
Dataset Description	This dataset contains the data required to model the information that is exchanged among components and

	abstract concepts used in the DELTA project.
Source (e.g. which device?)	Repositories that are present in the DELTA architecture and data sent among components.
Linked Architectural Component	DELTA Repository
<b>Partners services and responsibilities</b>	
Partner in charge of the data collection (if different)	Partners that need to model information
Partner in charge of the data analysis (if different)	UPM, Technical Partners
Partner in charge of the data storage (if different)	UPM
WPs and tasks	WP1, T1.3
<b>Metadata and Expected Size</b>	
Info about metadata (Production and storage dates, places) and documentation?	not yet available.
Estimated volume of data	several hundreds of bytes
<b>Data exploitation and sharing</b>	
Data exploitation (purpose/use of the data analysis)	Standardization and research purposes.

## 5. Conclusions

The DELTA DMP is based upon the datasets for procedures and infrastructure that are anticipated at this point in the project. The first version of this deliverable aims at providing the first identified datasets and outlining a draft version of the components specifications, so as to implement the DELTA Data Management Portal.

The next actions will be to focus on semantics and further clarification of procedures, participant and stakeholder engagement and identifying areas that need special attention. Activities for a Data Management Portal will be initiated and changes to the datasets may occur after systematic studies on pilots' activities.

An important conclusion is that all the partners are responsible for different kind of datasets. At this phase of the project, there are many difficulties to specify exactly all the relevant datasets for the projects' activities. This process will be supported from the studies and the tests on the pilot sites.

The DELTA Data Management Plan will focus on providing the necessary infrastructure for the appropriate collection, publication and storage of metadata. This metadata will be managed by each data producer and will be integrated in the Data Management Portal.

## 6. References

- [1]. **European Commission, 2017:** Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, version 3.2.
- [2]. **European Commission, 2016:** H2020 Programme – Guidelines on FAIR Data Management in Horizon 2020, version 3.0.
- [3]. **European Commission 2015:** Guidelines on Data Management in Horizon 2020, version 2.0.
- [4]. **D10.1 NEC – Requirement No. 1**, DELTA project, Confidential Deliverable, WP10, 2018
- [5]. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=en>