HORIZON 2020 - ICT-14-2016-1

AEGIS

Advanced Big Data Value Chains for Public Safety and Personal Security

WP1 - AEGIS Data Value Chain Definition and Project Methodology

D1.2 – The AEGIS Methodology and High Level Usage Scenarios

Version 1.0

Due date: 31.05.2017	**Delivery Date**: 15.06.2017
Author(s): Spiros Mouzakitis, Evmorfia Biliri, John Tsapelas (NTUA), Cinzia Rubattino, Elisa Rossi (GFT), Michele Caira(GFT), Alexander Stocker, Christian Kaiser (ViF), Sotiris Koussouris, Fenareti Lampathaki (SUITE5), Dimitrios Miltiadou (UBITECH), Alessandro Testa(HDI)	
Editor: Spiros Mouzakitis (NTUA) **Lead Beneficiary of Deliverable**: NTUA	
Dissemination level: Public	**Nature of the Deliverable:** Report
Internal Reviewers: Andreas Schramm, Yury Glikman (Fraunhofer), Konstantinos Perakis (UBITECH)	

EXPLANATIONS FOR FRONTPAGE

Author(s): Name(s) of the person(s) having generated the Foreground respectively having written the content of the report/document. In case the report is a summary of Foreground generated by other individuals, the latter have to be indicated by name and partner whose employees he/she is. List them alphabetically.

Editor: Only one. As formal editorial name only one main author as responsible quality manager in case of written reports: Name the person and the name of the partner whose employee the Editor is. For the avoidance of doubt, editing only does not qualify for generating Foreground; however, an individual may be an Author – if he has generated the Foreground - as well as an Editor – if he also edits the report on its own Foreground.

Lead Beneficiary of Deliverable: Only one. Identifies name of the partner that is responsible for the Deliverable according to the AEGIS DOW. The lead beneficiary partner should be listed on the frontpage as Authors and Partner. If not, that would require an explanation.

Internal Reviewers: These should be a minimum of two persons. They should not belong to the authors. They should be any employees of the remaining partners of the consortium, not directly involved in that deliverable, but should be competent in reviewing the content of the deliverable. Typically this review includes: Identifying typos, Identifying syntax & other grammatical errors, Altering content, Adding or deleting content.

AEGIS KEY FACTS

Topic:	ICT-14-2016 - Big Data PPP: cross-sectorial and cross-lingual data integration and experimentation
Type of Action:	Innovation Action
Project start:	1 January 2017
Duration:	30 months from **01.01.2017** to **30.06.2019** (Article 3 GA)
Project Coordinator:	Fraunhofer
Consortium:	10 organizations from 8 EU member states

AEGIS PARTNERS

Fraunhofer	Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.
GFT	GFT Italia SRL
KTH	Kungliga Tekniska högskolan
UBITECH	UBITECH Limited
VIF	Kompetenzzentrum - Das virtuelle Fahrzeug , Forschungsgesellschaft-GmbH
NTUA	National Technical University of Athens – NTUA
EPFL	École polytechnique fédérale de Lausanne
SUITE5	SUITE5 Limited
HYPERTECH	HYPERTECH (CHAIPERTEK) ANONYMOS VIOMICHANIKI EMPORIKI ETAIREIA PLIROFORIKIS KAI NEON TECHNOLOGION
HDIA	HDI Assicurazioni S.P.A

Disclaimer: AEGIS is a project co-funded by the European Commission under the Horizon 2020 Programme (H2020-ICT-2016) under Grant Agreement No. 732189 and is contributing to the BDV-PPP of the European Commission.

EXECUTIVE SUMMARY

The document at hand, entitled "The AEGIS Methodology and High Level Usage Scenarios", constitutes a report of the performed work and the produced results of Task 1.4 "Regulatory Framework for Data Protection, IPR and Ethical Issues" and T1.5 "Methodology Elaboration and High Level Usage Scenarios". The scope of the current report can be described in the following axes:

- Detailed high-level usage scenarios are drafted to outline the workflows and functionalities that the project is expected to support, highlighting the perspective of the end-users.
- The first version of the AEGIS methodology towards data-driven innovation in the domains of Public Safety and Personal Security is defined, describing the user interactions and workflows to be supported by the AEGIS system.
- An initial definition of the Minimum Viable Product (MVP) to be developed during the project is provided based on the available usage scenarios and the defined methodology.
- The project's strategy towards ethical and data privacy and IPR considerations is defined in detail, taking into consideration all data privacy requirements and existing regulatory instruments, concluding with the definition of the AEGIS Ethical, Privacy and Data Protection Strategy.

The results of the current deliverable, including the integrated project methodology, the analysis leading to MVP definition, the extracted features and functionalities and the data protection requirements and strategy will be leveraged as input to the technical tasks of the project and will offer guidance towards selecting the most appropriate technologies for the AEGIS solution. Updates to the work and results described here will be presented in D1.3 entitled "Final AEGIS Methodology".

Table of Contents

LIST OF FIGURES

ABBREVIATIONS

AAL	Active and Assisted Living
AEGIS	Advanced Big Data Value Chains for Public Safety and Personal Security
ALLDS	Aggregated Local Linked Data Space
API	Application Programming Interface
CEO	Chief Executive Officer
COPD	Chronic Obstructive Pulmonary Disease
D	Deliverable
DoA	Description of Actions
EAB	Ethics Advisory Board
HVAC	Heating, Ventilation and Air Conditioning
IAQ	Indoor Air Quality
ICT	Information and Communication Technology
IP	Intellectual Property
IPR	Intellectual Property Rights
IT	Information Technology
MVP	Minimum Viable Product
OBD	On Board Diagnostics
PAYG	Pay As You Go
PSPS	Public Safety and Personal Security
REST	REpresentational State Transfer
SLOD	Security Linked Open Data
SME	Small and Medium-sized Enterprise
T	Task
WP	Work Package

1. INTRODUCTION

1.1. Objectives of the deliverable

This deliverable is related to the activities performed during the first iterations of Task 1.4 and 1.5, regarding the integrated AEGIS methodology towards data-driven innovation in Public Safety and Personal Security (hereinafter PSPS), the refinement of the project's concept and supported workflows, as well as the legal evaluation and assessment of the AEGIS technologies and the integrated systems. More specifically, the main objectives of the deliverable are to:

- Provide high-level usage scenarios of AEGIS that showcase how stakeholders will interact with AEGIS and with each other through AEGIS, which are their expected outcomes from using the project's offerings and which are the envisioned workflows enabling them to achieve their goal.
- "Translate" the envisioned usage scenarios into initial insights for the required flow of information, data value chain functionalities and all core operations to be supported by the project's platform and model all envisioned user interactions with AEGIS in an integrated methodology that clearly defines the workflows to be supported. These insights will be leveraged also to provide a first definition of the AEGIS Minimum Viable Product (MVP).
- Identify, monitor and analyse relevant legal and regulatory legislation relevant to AEGIS innovations and implementation. This work includes the identification of relevant national, regional, legislation and regulatory instruments, the extraction of data protection requirements and the initial definition of the AEGIS Data Policy Framework.

1.2. Insights from other tasks and deliverables

The current deliverable is strongly related to D1.1 "Domain Landscape Review and Data Value Chain Definition", as it builds on top of the work reported there in order to further refine the AEGIS concept, outline usage scenarios and identify data requirements in terms of privacy and security. More specifically, D1.1 (a) identified and described in detail 11 AEGIS stakeholder groups and extracted their initial Big Data needs, hence outlining the expected interactions, (b) provided a large volume of indicative data sources per PSPS stakeholder group that manifest all 4 Vs of Big Data (Volume, Velocity, Variety, Veracity), highlighting the need for common standards and improved semantics, as well as the importance of handling geo-referred data and time-series, (c) sketched the AEGIS data value chain and stakeholders value chains that form the complete PSPS ecosystem and (d) presented an extensive literature review for Big, Linked Data tools and methods which offers actionable insights into how certain user needs can be mapped to and realised by specific technological artefacts. The outcomes from all these processes evidently served as input for the work being reported in D1.2.

Furthermore, D6.1 "Plan for Dissemination, Communication and Stakeholder Engagement" and D7.1 "Project Exploitation Plan" present a grouping of the 11 identified stakeholder groups in terms of shared high-level reasons for utilising AEGIS, which was leveraged here in order to efficiently select representative usage scenarios. D7.1 further provided some more detailed insights into the expected outcomes and benefits of the project, including an initial list of both

tangible and intangible expected assets (semantic models and linked data schemas, data handling algorithms, micro-services and services management, cross-sector analytics and visualisation, business brokerage) that were used as consulting material in the current work.

Finally, the work presented here is closely related to the work performed in T3.1 "Technology Requirements and Integration Components Analysis" and T3.2 "Demonstrator User Requirements" and close collaboration was pursued in order to ensure that the high-level usage scenarios, the integrated methodology and the AEGIS MVP are aligned with and can act as input for the project's user stories and, subsequently, functional and technical requirements. Towards this goal, input by the pilots was not limited to their individual cases, but contributed to the understanding of the wider PSPS ecosystem needs.

1.3. Structure of the deliverable

The deliverable is organised in four sections. Following the first introductory section, each subsequent section corresponds to a different objective from the ones described above. More specifically, section 2 provides a short but descriptive overview of the project's concept and presents detailed usage scenarios that clearly show the envisioned stakeholder interactions and outline the required data value chain processes, as seen from the end-user's perspective. Section 3 extracts from the envisioned scenarios some high-level, but more concrete, functionalities to be considered by AEGIS and further organises them to facilitate the definition of the project's MVP. At the same time the section provides the initial definition of the integrated AEGIS methodology for data-enabled innovation in the PSPS domains. Finally, section 4 describes the AEGIS ethical, privacy, data protection and IPR strategy.

2. AEGIS CONCEPT AND HIGH-LEVEL USAGE SCENARIOS

2.1. Concept in a nutshell

AEGIS aspires to realise the integration and combined analysis of a plethora of diverse data sources, cross-domain and cross-lingual, having different formats and conforming to various standards, having in common their connection to PSPS, which may correspond to the way they were produced and/or the way they are to be consumed. Such data sources range from various sensor data, to company databases and unstructured data from Web 2.0 sources. AEGIS will support the complete range of data processes that comprise the Big Data Value Chain, which at a high-level include data acquisition, analysis, curation, storage and usage and at a lower level span from semantic annotation to interlinking and visualisation. To effectively handle the diversity of the data, AEGIS will provide models and vocabularies that capture the underlying semantic connections of the data when seen in a PSPS context.

The identified stakeholders of the project are categorised in eleven discrete groups, namely Smart Insurance, Smart Home, Smart Automotive, Health, Public Safety/Law Enforcement, Research Communities, Road Construction Companies, Public Sector, IT Industry, Smart City and End Users. At the same time, the expected users of the AEGIS offerings could be classified into different categories depending on their knowledge background (e.g. familiarity with technology and data processing) and their primary interest in using AEGIS (e.g. data provision, service consumption, service creation, exploratory data analysis etc.). The current deliverable provides a first, high-level grouping of the users, however discussions are ongoing and the final user categories will be documented in D3.1.

To fulfil the needs of all stakeholders and user roles, AEGIS will develop a central platform and a set of additional tools to facilitate (a) big data interlinking under a common PSPS context through commonly agreed upon semantics, (b) big data discoverability and consumption through intuitive and configurable services, (c) big data and intelligence sharing through clearly defined interactions and data privacy and security preserving mechanisms, (d) advanced sensory data manipulation, (e) easy application of big data analysis and visualisation and, ultimately, (f) quicker and easier roll-out of data-enabled applications related to the PSPS domains.

2.2. High-level usage scenarios

This section provides five high-level usage scenarios of AEGIS, which have been provided by the partners based on the brainstorming activities that were conducted to refine the project's concept.

Scenarios were initially drafted as workflows, i.e. as clear sequences of steps that a stakeholder would follow in order to achieve her goal. The stakeholders, their reasons for using AEGIS and their expected outcomes were selected based on the insights provided in D1.1 "Domain Landscape Review and Data Value Chain Definition", D6.1 "Plan for Dissemination, Communication and Stakeholder Engagement" and D7.1 "Project Exploitation Plan", as well as initially shared input from the AEGIS pilot partners. After collecting feedback on the initial drafts, the outlines of five scenarios were chosen as representative of all core differentiated AEGIS users, in terms of user needs to be covered. The concrete, descriptive scenarios presented

here, are the result of extensive discussions and consultations among partners to ensure that the scenarios correspond to real life problems in data analysis related to PSPS.

For each of the scenarios, apart from the actor, the overview and the detailed description, the partners have provided possible alternative actors and actors interested in the outcome of the scenario, in order to highlight the inherent need for data sharing and the high potential in data-enabled collaborations among PSPS stakeholders.

2.2.1. Scenario 1: Advanced time-series analytics in the automotive sector

Actor
Vehicle research centre
Alternative Actors
Company creating digital car services
Actors interested in the outcome
Municipalities, Traffic Police, Road construction and maintenance companies, Drivers, Insurance Companies
Overview
A research team creates services on top of streaming vehicle data to (a) identify unsafe driving patterns and correlate them with external conditions and (b) to timely detect damages in the road network.
Scenario
Alex works in Auto4All, a research centre focusing on environmentally friendly vehicles. His team mainly uses vehicle specifications coming from car manufacturers in order to evaluate their environmental footprint but, until recently, did not have any field data to work with. They have now been given access to real time streaming data from 500 taxis, installed by the taxi company that owns them, whose CEO is interested in helping the drivers reduce fuel consumption and improve their driving patterns. For each of the taxis, the installed sensor generates an entry every 10-30 milliseconds (depending on the sensor) with the following values: *trip ID,* vehicle sensor data like *calculated engine load, engine coolant temperature, intake manifold, engine speed, vehicle speed, intake air temperature, air flow rate, throttle position, barometric pressure, ambient air temperature, throttle actuator;* GPS data *latitude, longitude, altitude*; Gyroscope sensor data: *gyro_x, gyro_y, gyro_z*; Acceleration sensor data: *acc_x, acc_y, acc_z*; and *time*. The taxis perform transfers mostly in the city of Vienna and each taxi is used by three different drivers, in 8-hour shifts, however there is no entry to indicate when the driver change happens.

Alex is looking for ways to explore and extract insights from the newly acquired data, however he and his team are struggling since they do not have the knowledge or the resources to perform the required time-series analysis. Based on their expertise on the subject, they are certain that these data can reveal interesting driving patterns and already have some initial ideas regarding the type of algorithms that could help identify them. They would also like to be able to visualise them on a map in a way that also reflects the time progress to design the most appropriate analysis strategy. At the same time, they are certain that these data would be much more valuable if combined with weather data, but are unaware of a way to make such a combination and, in any case, the free weather service they know does not provide new values in such short time intervals.

Alex proposes to use AEGIS in order to perform the initial data experimentation. He visits AEGIS web platform and explores the functionalities offered through his personal experimentation space and proceeds to upload, for private usage, a small subset of the trip data that correspond to 7 days for 50 taxis. Prior to uploading, Alex is prompted to consider anonymising the data or removing certain columns, using tools provided by AEGIS, in case there are sensitive information he does not wish to upload, but he continues without performing such a process. In the data that are given to Alex's team, the sensor id has been replaced by random ids that cannot be linked to the specific taxi. Furthermore, these data correspond to registered taxi transfers and do not hold any information that could help identify the driver or the passengers.

Having uploaded the data, Alex now proceeds to perform a reduction, through an easy to use AEGIS service, to keep one entry every 100 milliseconds, since he thinks that at this point he does not need so dense entries. As a start, he wants to visualise all 50 time-series at once, so he chooses the option to combine multiple time-series in one, by simply adding one extra column to hold the time-series id. From the available visualisation options that AEGIS provides, Alex chooses the map and is able to easily picture the safety-relevant events (braking and acceleration) and notices that hard brakes are already evident. Filtering the results to show one id per time and selecting the time evolution mode he notices slight changes that could correspond to different driving patterns from distinct drivers.

More confident about the value of the data, Alex now decides to combine them with more information. Navigating to the AEGIS services, he finds two distinct services that are very relevant to his target:

1. a weather service providing precipitation, clouds, pressure, temperature and wind measurements, which can be configured based on the desired granularity in terms of location and time intervals
2. a configurable service providing annotated car accident data from various selected sources and regarding various locations

He imports the two services as additional input sources for his visualisation and finds the accident indicators on the map in some cases correspond to the locations with hard braking. At the same time, filtering by weather also visually reveals some patterns. In

order to further improve the visualisation, Alex chooses to apply the AprioriAll algorithm, which is provided by AEGIS, on the combined (driving and weather) time-series and creates a new visualisation, this time to show only the extracted sequential patterns.

Alex is now certain that AEGIS can help his team focus on the parts of the data analysis that they are mostly interested in, removing the burden of configuring technically challenging solutions from scratch, so he decides to proceed to leverage AEGIS capabilities for the heavier data processing tasks and computations.

After connecting the sensor data streams to the AEGIS cloud, Alex's team proceeds to add the weather data-as-a-service and accident-data-as-a-service (that Alex had already used) as input in three algorithms, which have been configured to suit their needs. More specifically, the team schedules a braking detection algorithm, an acceleration detection algorithm and a road damage detection algorithm to run every 10 minutes. All three are configurations of specific pattern mining algorithms (namely AprioriAll and AprioriSome) that are by default provided by AEGIS, with some use-case specific rules applied, which have been defined based on their domain expertise. The team then refines some of the visualisations that were initially created by Alex and extends them to provide configurable interactive dashboards that highlight the driving patterns and indicate specific locations and combinations of circumstances that are related to unsafe driving behaviour.

Alex proceeds to further configure his personal instances of the dashboards according to his needs and asks that every member of the team does the same and freely run individual experiments, whilst keeping the "master" dashboards intact.

After a week of experimentation, Alex and his team finalise the analysis that the taxi company was interested in and proceed to send them a link to the AEGIS interactive maps that show under which circumstances the drivers are not driving optimally.

However, Alex and his team have also identified another useful way to leverage the sensor data: after a month of experimenting with the visualisation of the road damage detection algorithm, performing also on-the-field validation of results, the team realises that it provides clear insights into parts of the road network that need repair. Their analysis is so reliable they can even provide alerts when the weather conditions increase the risk of accident above a certain threshold. Although this visualisation no longer holds any driving data, Alex asks the taxi company for consent in order to monetise his team's finding. Since the company is already an AEGIS service consumer, they happily agree to act as a formal data provider as well for this specific service.

Through AEGIS, Alex and his team publish this visualisation in the form of a weekly and a monthly report and create an alert service where interested parties can subscribe for road risk notifications. These are now available as AEGIS data-as-a-service and visualisation-as-a-service and Alex is confident that the Municipality of Vienna, as well as the road construction companies will be interested in using them.

Benefits
• Reduced time required for tedious time-series manipulation processes (e.g. reduction, replacement of values) • Higher level of automation in several parts of the business workflow, leading to significantly reduced time to produce results • Availability of a wide variety of algorithms and easy experimentation and comparison of results • Provision of innovative services that reach a wider audience and monetisation of results • Access to a wide variety of useful data sources to mesh-up with imported data (e.g. weather, accidents, …) • Easy provision of appealing web-based reports from aggregated data for customers
Challenges
• Seamless combination and usage as algorithm input of various diverse time-series • Connection of streaming data in a way that supports various processing tasks to be applied in (almost) real-time • Configurable visualisations and provision of visual analytics functionalities for Big Data • Terms of usage and licensing schemes for data and services coming from multiple stakeholders • Combination of time series data with other structured as well as non-structured data

2.2.2. Scenario 2: Data-enabled services for enriched real-time navigation system

Actor
Software company creating navigation systems
Alternative Actors
Smart city planners
Actors interested in the outcome
Drivers, Car insurance companies, Municipalities, Traffic Police
Overview

A for-profit organisation uses various AEGIS data-as-a-service services to implement an enhanced navigation system for the city of Vienna that shows traffic, accident statistics, road accident risk indicators and events that may affect traffic.
Scenario
Elias is a senior developer in Auto4Techs, a software company based in Austria, and his team has developed a land-based navigation system, which entered the market two years ago and has managed to gain a market share about 10% so far. The team is now in the process of brainstorming for ways to become more competitive; however, there are many established systems in the market and they cannot find a promising way to differentiate their core offering. Elias states that it would be interesting to extend the navigation system with additional layers of information, providing services on top of the core product. He suggests enriching the system with accident risk annotations and real-time alerts for events that could affect traffic. However, although all team members find the ideas interesting, nobody is aware of a way to acquire the information required for the first one. At the same time, the second one would require increased effort in terms of identifying possible input sources, retrieving and processing data and nobody in the team has experience on natural language processing to do that. They bring their ideas to the management level to ask for time and resources to work on these new services, but they are informed that the company has decided to reduce the effort devoted to the navigation system and the team must come up with solutions that do not require significant in-house development effort. Elias proposes that all team members search for existing solutions that could be leveraged towards implementing the new services more easily and quickly, both internally in the company and externally and report back their findings. Through this process, one of his colleagues comes across AEGIS and sends the link to Elias. Elias opens the interactive map that shows the AEGIS services by region and navigates to Austria. He discovers that AEGIS offers several data-as-a-service services with information about Austria. More specifically, he finds six that are extremely interesting: 1. Weather service providing precipitation, clouds, pressure, temperature and wind measurements, which can be configured based on the desired granularity in terms of location and time intervals 2. Crime service providing geo-located crime events based on open data 3. Car accident service providing geo-located car accidents with severity annotation based on open data 4. Events service providing a configurable real-time stream of a wide-variety of events, including riots, concerts, football matches etc. 5. Road damage indicator service, which provides geo-located indicators of possible road damage. The service is provided by a research centre for environmentally friendly cars based in Austria and is only available for the city of Vienna (ref Scenario 1).

6. Car accident service, similar to the third one, but provided by an international auto insurance company which also operates in Austria, based on their internal data.

The first four services are tagged as AEGIS offerings, whereas the other two are proprietary services provided by other organisations. Elias reviews the services and examines their expected output, the number of entities already using them, their maintenance information, terms of use and price and believes they will significantly facilitate the development of the enriched navigation system. He brings them to the manager who decides to use the four services provided by AEGIS and the one from the research centre regarding road damage since it is highly important information that they could not acquire elsewhere. He proposes, however, to skip the last one from the insurance company as the alternative offered by AEGIS is less expensive, since it is based on open data.

Elias and his team proceed to gain access to the five services which they can then easily consume through the provided RESTful APIs. Since all services provide latitude and longitude for each entry, the team can very easily integrate them into their navigation system. The only addition required is to design intuitive signs to represent each of the newly acquired pieces of information, e.g. road damage sign, large number of accidents in the area, large concert starting within the hour, etc. Furthermore, aside from the original two ideas (accident risk estimations and event notifications), Elias and his team are able to leverage the acquired services towards developing an additional service for their navigation system, which they call “ParkSafe”, which provides indications on the map to avoid specific regions for parking due to extreme weather conditions, high crime rates or local events like protestations, strikes etc.

In the end, the team releases the updated navigation system with various subscription modes, to let every driver decide which services to include in her system. Their navigation system is now the most inclusive navigator for the city of Vienna and the responses from the local drivers are extremely positive, so the manager has already asked them to examine the availability of AEGIS services for other locations as well.

Benefits

- Reduced in-house development time and effort
- Higher level of automation in several parts of the business workflow, leading to significantly reduced time to produce results
- Provision of innovative services
- Easier access to a variety of data sources
- Discoverability of and easier collaboration with other companies

Challenges

- Level of localisation for the provided services
- Public data sources can change suddenly in both structure and content without any notice

• Terms of usage and licensing schemes for data and services coming from multiple stakeholders

2.2.3. Scenario 3: Data-Driven Monitoring and Alert Services for Impaired or High Risk Groups Individuals

Actor
Social Care Service Provider
Alternative Actors
Hospitals, Doctors, Physicians, Rehabilitation Centres, Formal Carers Organisation
Actors interested in the outcome
Insurance Companies, Hospitals, Doctors, Physicians, Rehabilitation Centres, Formal Carers, Informal Carers, Impaired People
Overview
A social care service provider is developing an infrastructure relying on domain specific data (coming from individuals, own knowledge and external data sources) that will allow to serve impaired or under risk individuals with improved protection, recommendation and notification offerings to avoid high risk situations.
Scenario
Amanda is the head data analyst of SameHealthForAll Inc, a private-public social care service provider that is co-funded by the government and some private funds in order to provide health services to individuals and to other relevant organisations as well, acting both as an intermediary expert and service provider to primary and secondary health institutions and professionals, but also as an organisation that has direct interface with individuals. Amongst the key strategic SameHealthForAll Inc goals, is the design and roll-out of specific health related services addressed to individuals that have certain diseases or conditions that classify them to high risk groups (like for example Alzheimer or heart condition patients, disabled people and pregnant women), to which also elderly people belong. The services offered to this group by the organisation vary from formal recommendation and notes that are issued towards the health care providers, to informal or simplified guidelines and notifications issued to the general public via the press, flyers and media channels. Those are mostly based on previous experience, and one of Amanda's daily tasks is to go through the news coming from the press, other health agencies and related organisations (such as environmental agencies) and indicate

incidents above certain thresholds that are connected with past experiences. Once such an incident is identified, Amanda gathers her team and, after consultation, the risk factor of an incident is determined. If this is characterised as "high", then the appropriate procedures for raising awareness are triggered, with announcements being issued, updates of protection guidelines being drafted and communication to the relevant stakeholders being performed.

The identification of such incidents has been made lately more straightforward thanks to the introduction of novel ICTs that are able to grasp data from external resources, whether these are of public and open nature, or private, belonging to organisations with which SameHealthForAll Inc. collaborates and has access to notification and data subscription services. However, the organisation is still constrained to the provision of recommendations and services of generic nature, which although valuable, are not that intelligent and their potential is limited by various constraints. The main constraints of those have to do with the inability to tackle needs which are way more complicated and highly dependent on the context of each individual, as well as locality, as most data acquired refer to large (often nation-wide) geographical areas.

In the quest of finding ways to offer advanced services, which could be directed both to other organisations, but also directly to high risk groups individuals, Amanda has forwarded a proposal to the board of directors to establish a strategic partnership called "Health2home2020" with specific healthcare providers, smart home and automation service providers and other organisations, who are able to provide the missing pieces to complete the envisaged services offering, by utilising their resources and infrastructure to serve SameHealthForAll Inc. with data coming from individuals and from their environment, as well as data coming from other known clinical data sources that are essential for conducting the analyses. Of course, apart from the willingness of organisations to share data and value, other issues have to be resolved, ranging from the absence of infrastructure for analytics, to common data schemas and data confidentiality.

During the latest discussions in a cross-sector meeting of the "Health2home2020" alliance, Amanda suggested to the partnership to base their operations on the offerings of the AEGIS platform, as it promises many solutions that solve most of the challenges that have been identified.

Visiting the AEGIS platform, Amanda creates an organisation profile for the company and then creates a project called "Assisted ALL Analytics" marking it as a "private" project. In the project, Amanda invites some of her colleagues to have access to the same project, giving them full access to the data. At the same time, she also sends invitations to external collaborators with specific rights to upload data to the project's repository, indicating also the expected data structures and schemas.

The other members of the "Health2home2020" alliance, having received the invitation from the platform, join the project and are presented with the list of data that they have been asked to upload. At their disposal are a number of tools that could anonymise, clean and transform data, in order to meet the expectations of Amanda. After specifying what kind of data they will upload, specific dialogues guide them through the process,

prompting them to take actions (possibly leveraging AEGIS tools) towards ensuring compliance with data privacy terms, in case the nature of the data to be uploaded imposes such requirements. Once they do this, they are presented with their data, while Amanda is able to cross-check that these data comply with the necessary requirements, and if not transform them accordingly. This step is done multiple times, as some data are initially uploaded only for experimentation purposes (e.g. some samples of datasets), while there is also the option to connect data to the platform through a project specific API endpoint.

Having all the data in one place, Amanda is now able to invoke several analyses, choosing which data to combine as well as the algorithms to utilise. Those come out of a predefined algorithms library, while it is also possible for Amanda to conduct an analysis using her own algorithms, which, however, requires an additional step in order to pass an approval process by the system's administrator, to guarantee proper resources utilisation in AEGIS, following the request to include a new algorithm that has to trigger a consultation and testing process led by the AEGIS team. The overall results are then presented in a dashboard that visualises the outputs of the analyses, where access can be provided to any member of the project, while the results can be also exported in various formats. What is especially interesting for Amanda, is the option to export the data through an API. This can come either from an external stimulus, such as a weekly call from an external system residing in the "SameHealthForAll Inc" datacentre, or from triggers specified and enabled in the AEGIS platform, such as the updating of a dataset or the occurrence of an event.

At the moment, she has saved two analyses, which are called "IndoorConditionsRisks" and "TemperatureRiskPlots". The first one is able to relate indoor air quality conditions, as retrieved from IAQ sensors installed in premises, further correlated with HVAC devices operational status, IAQ data from external weather stations and IAQ constrains/regulations as defined by national and international health organisations. This correlated analysis will further enable:

1) home automation services to trigger the appropriate control strategies on HVAC devices towards addressing IAQ requirements
2) social care service providers to trigger notifications about IAQ conditions.

The second one is correlating positioning and health information gathered from wearable devices, location and weather conditions along with Public Health Information Sources towards the identification of:

1) possible environmental conditions that pose risk for the health and well-being of individuals (e.g. alert individuals suffering from COPD in case of increased humidity outdoors),
2) pattern irregularity which could signify cognitive deterioration (e.g. wandering off without any apparent reason),
3) physical wellbeing deterioration and frailty status (e.g. detection of falls or taking the individual more time to complete typical physical routines).

In all those analyses, specific user groups are formed based on various aspects such as their demographics, their conditions etc.

At the same time, all involved stakeholders of the alliance are also permitted to upload their data publicly, setting specific licence requirements for reuse, and can also explore data and services already uploaded in the platform that will help them to further improve the value of their in-house information. In the same manner, Amanda is also able to identify in the platform a social media mining service that sends notifications based on specified keywords specified. After she accepts the license agreement she starts experimenting with it by asking the service to return alerts coming out of Twitter stream analysis regarding ice incidents at the greater district of Cambridge. She also decides to utilise these notifications as a trigger to run again the analysis called "TemperatureRiskPlots" that has been part of the project she initiated. Furthermore, she also finds a free service about weather conditions that offers more detailed information than the one she is currently using, and decides to replace this data stream in her analysis.

Having conducted the analyses, Amanda presents to the alliance the benefits of AEGIS and persuades them to utilise the service at production level, as it fully covers the needs of the whole alliance, while it does not require high investments costs, or the presence of a data analyst in each organisation. By doing this, Amanda is able to gather data for the platform to build the services she has envisaged based on the outcomes of the analyses. These enable her organisation to develop application services that offer personalised information and recommendations to stakeholders. These applications utilise both the data streams of the AEGIS analytics, as well as the intelligence coming out of them, to define specific rules and triggers for sending out information. At the same time, some of these services (depending on data nature and strategic value) provide data back to the AEGIS platform, either directly, through API calls to the platform's endpoints, or indirectly, as in the case of healthcare institutions who have stated that they prefer to first retrieve and review data and then re-upload selected parts of them to the platform.

At last, the services are deployed, and the organisation, together with all the value chain collaborators are able to offer added value services to their end-users, as well as to each other. Furthermore, new data and service exploitation models are introduced through the platform, as essential information such as indoor conditions from sensors, as well as activity tracking information of anonymised personas are offered online, and other interested stakeholders can experiment with them and then come to an agreement for their reuse with the data owners.

Benefits

- Secure online collaboration in data collection and management
- Improved offerings in services, data discovery and utilisation
- Less investments in local IT infrastructure and experts
- Decrease data analysis times
- Data analysis complexity reduction

• High level of automation through API calls and analysis triggering methods • Availability of reports and visualisation for all stakeholders in the value chain • Integration with third party services for analysed data export through APIs • Availability of a wide variety of algorithms • Monetisation/Value Generation potential through the exposure of own data and service offerings
Challenges
• Seamless combination and usage as algorithm input of various diverse time-series • Discovery of location specific data set, especially from external resources • Connection of streaming data in a way that supports various processing tasks to be applied in (almost) real-time • Configurable visualisations and provision of visual analytics functionalities for Big Data

2.2.4. Scenario 4: Personalised early warning system for asset protection and commercial offering

Actor
Insurance Company
Alternative Actors
-
Actors interested in the outcome
Insurance Company, Insurance Brokers/Agents, Insurance Company Customer
Overview
The following scenario presents how AEGIS solution can be used in order to improve the quality of services an Insurance Company provides to its customers and thus the image of the Company towards them. Amongst the key options to be used by the

Insurance Company is the monitoring near real time of the potential impact of a risk or threat on its business and its client portfolio.

Scenario

Sylvia is a data analyst in the Insurance company "ASInsurance4lifes" with the main task to analyse emerging threats and evaluate their impact on the company's portfolio and subsequently to the insured assets of the company's clients.

So far, Sylvia's daily work relied on the debriefing of press releases and media broadcasts to note down potential risk situations, and using some simple risk detection models she was able to identify emerging threats for clients. However, the models used, due to high data granularity, are only able to provide generalised assumptions and cannot be directly related to the detailed needs of clients, while at the same time data protection issues and the inability to update each client's contextual data hinder Sylvia and her department from delving into more details for clients as to provide smarter and more personalised services.

Sylvia has been using the AEGIS solution to monitor the potential impact that a natural disaster can have on her company's business. In particular, when a potential risk or threat is identified (e.g. a natural disaster like a hailstorm or a flood), Sylvia's department is tasked to identify as soon as possible insured assets and customers that are potentially affected, towards providing them with an efficient, personalised assistance service.

In order to perform this task, Sylvia has rolled-out some event detection services, available on the public "on-cloud" AEGIS solution. These services have been configured in order to gather data from publicly available sources, such as weather conditions and forecasts and web news as well as social media data. These services are able to, first, identify a potential risk or threat and, second, to classify the type of risk. Moreover, AEGIS services offer the options to associate a geo referred information to collected data, while at the same time other historical datasets that have to do with past weather conditions, flooding incidents etc. are consulted to build a more factual forecasting model that can better predict risk occurrence and impact. Furthermore, Sylvia has also found that other forecasting models are already available in the platform and provided as services by third parties, concerning one specific region of interest, Lombardia, where data acquisition by the usual open data sources is limited. Sylvia has already tried out those models over the platform, and to her surprise, she found out that one of them provides very accurate predictions. Therefore, she has agreed to its terms of usage and is utilising this one for forecasting conditions in that specific region, and complement the other forecasts she has already selected.

By consuming these services, the company is now able to detect the risk as soon as it occurs and to know exactly the place where it is occurring.

Since company's internal information, like customer details and portfolio, cannot be shared with a public on-cloud infrastructure, the company decides to use an encrypted cloud space and data undergo specific pre-processing prior to being uploaded to ensure that sensitive information will not be compromised.

Let's consider the case in which an intense hailstorm takes place in a specific geographical area. By consuming the AEGIS on-cloud services that Sylvia has already set up, the company detects in (close to) real time the fact that a potential risk can affect insured assets. It also identifies the kind of the risk (hailstorm) and the geographical area.

This information is combined by Sylvia with internal enterprise databases of the company:

- The kind of risk identified by AEGIS services is combined with internal portfolio system in order to extract the contracts whose guarantees cover the identified risk;
- The geo referenced information contained in AEGIS data is combined with the previous dataset to filter only the contracts associated with assets in the geographical area interested by the hailstorm;
- The previous dataset is combined with internal customer details in order to extract customers that are potentially affected by the identified risk. To this, a specific mobile App offered by the company is also used to retrieve the latest location of individuals, and allow for better prediction of the customers affected;
- Final outputs are the list of insured assets, and thus company's exposure to the risk, and a list of customers potentially affected by the identified risk.

Both lists are sent to the AEGIS advanced analytics suite and allow business users to monitor and forecast the potential impact of the threat on Company's business. At the same time, the company is able to evaluate the offering of micro-insurance contracts to potentially affected customers, after conducting analyses that identify the risk exposure and threat level for each type of asset, and the optimum pricing strategy, taking into consideration the number of such contracts offered, the already accepted contracts and financial exposure of the company as well.

The list of the customers is used to contact them and to provide an efficient and personalised assistance service. Thus, depending on customer's preferences, the company can

- Send the registry to the Call Centre which will contact the customer
- Contact directly the customer through the dedicated mobile app / chatbot, thus avoiding the Call Centre

By contacting customers as soon as possible, the company is now in a position to improve its brand image and, if necessary, also can provide customers with the most appropriate indications. For example, in the case that the hailstorm damages a customer's vehicle, the company can immediately put him in contact with one of the associated

bodyshops.
Benefits
• To provide a better service to the company's customers. • Minimise reaction time before handling the risk • Offer value added services to customers • Improve company's image and brand
Challenges
• Public data sources can change suddenly in both structure and content without any notice • Risk type identification is not straight-forward • Level of data and service localisation • Real-time requirements for data analysis (e.g. identification of potentially affected insured assets and prompt notifications).

2.2.5. Scenario 5: Open Innovation platform for Data Experimentation and Service Offering

Actor
A data analyst
Alternative Actors
A consultant, a data engineer, a data-analysis SME, a citizen, a student
Actors interested in the outcome
All actors using AEGIS
Overview
A data analyst is using the AEGIS platform to experiment with data owned by him or discovered on the platform, resulting in the generation of analyses, reports, visualisations as well as data streams which he can then offer as a service through the AEGIS infrastructure to any interested stakeholder.
Scenario

Vicente is a data analyst that has recently completed his postgraduate studies in data science and is working as an independent contractor/consultant in different firms and organisations. He is an open source and open data enthusiast and highly believes in the data sharing economy, trying to contribute back to all communities he is engaged in. His business and research interests focus on big data and data analytics technologies. He usually works from home to deliver various reports to the companies he has contracts with.

Growth of business seems to be in a good track, but Vicente is struggling with the increasing demand for data processing infrastructure he needs. It is not only his client base that is growing, but also the volume and variety of data itself are growing as well. For this reason, he has recently conducted an analysis of available solutions that could meet his needs for both experimentation and production purposes, and at the moment he has concluded that the AEGIS platform is in a position to cover his demands. What he finds especially interesting is that AEGIS inherently provides elevated support for data analytics related to public and personal safety and security applications, since he has a rich client base interested in these topics.

Vicente has already created an account in AEGIS. Since he is going to use it not only for small scale experimentation (but also for storing his data and conducting large scale analyses) he has decided to select a PAYG plan that takes into consideration the data volume retained in the platform and the computing requirements for each analysis (CPU time and Memory). Currently, he has three different repositories in his account, two that deal with data analytics services he is offering to two of his clients, and one he is using for experimentation purposes.

In the first two repositories, Vicente has set up a chain of services which conduct analyses at prescribed time intervals and deliver both data streams (via a dedicated API) and reports (through a dashboard, that is based on the default visualisation features offered by AEGIS), which are both currently set to “private”, not allowing external entities to access them. During the set-up of those repositories, Vicente has utilised the standard analytics features offered by the platform, and has built a chain of algorithms where output of one analysis if being fed into another. He has also found online in AEGIS platform some datasets which are more detailed than the ones he has used, and after a short experimentation with them in the third repository (see below), he decided that their inclusion in his business cases is quite beneficial. The first one came for free, while the second has a monthly basis usage license, which Vicente decided to acquire. However, as the data delivered by that dataset comes only in .csv format with lots of unnecessary (for Vicente) data that increase the storage needs unnecessarily, he has combined some existing AEGIS services into one larger preprocessing service that cleans the data and transforms it into JSON. Vicente has decided to make his new data manipulation service (i.e. chain of AEGIS services) available for free to everybody in AEGIS (although initially he was thinking of putting a price on it, to balance the costs he pays for the data licenses).

The third repository that Vicente has set up is his main "playground". There he is able to upload different datasets, retrieve data from APIs, utilise the services provided by AEGIS to combine, clean and transform his own and other third-party data sources which he discovers through the platform, and is also able to conduct small scale experiments. The main objective of this is to constantly evaluate the offering of the platform (in terms of data, services and algorithms) and port artefacts that seem beneficial to his "production" repositories. For this to happen, he uses some features that allow him to replicate the attributes and environmental parameters of his "production" repositories and install them into the experimentation repository, and then try to integrate his ideas there, without compromising his running production instance.

At the end of the day, Vicente is happy for having found a platform that saves him a lot of time and of course cost for his job. Although he understands that there exist some limitations in the AEGIS platform due to its nature as a centrally offered PaaS, he thinks that the trade-off of customisation for improved ease of use is completely worth it for his own case.

Benefits

- Utilisation of online resources on a PAYG scheme, lowering in-house investments
- Improved reliability
- Improved performance and scaling opportunities
- Online collaboration in service delivery, data collection and management
- Improved offerings in services, data discovery and utilisation
- Availability of a wide variety of algorithms
- Monetisation/Value Generation potential through the exposure of own data and service offerings

Challenges

- Fair usage policies and pricing strategy
- Combination of existing methods and services
- Combination and configuration of AEGIS services chains to form integrated "user-authored" data manipulation services

3. AEGIS METHODOLOGY AND MVP DEFINITION - (FIRST)

The approach followed during the creation of the high-level AEGIS scenarios encouraged the description of workflows as perceived and envisioned by the end users. Thus, the scenarios reveal the reasons why users would consider adopting the project's offerings, the needs they would like to cover and the steps they would expect to go through in order to achieve this, deliberately "ignoring" any technical aspects and difficulties, as well as implications related to data/service availability and business strategy. The scenarios are the core input for the definition of the project's integrated methodology and of its MVP.

3.1. Feature Extraction from Scenarios

Each of the detailed scenario descriptions presented in section 3.2 has been abstracted to a sequence of steps, each of which is related to a number of implied features in order for the described functionality to be provided by the AEGIS platform and tools.

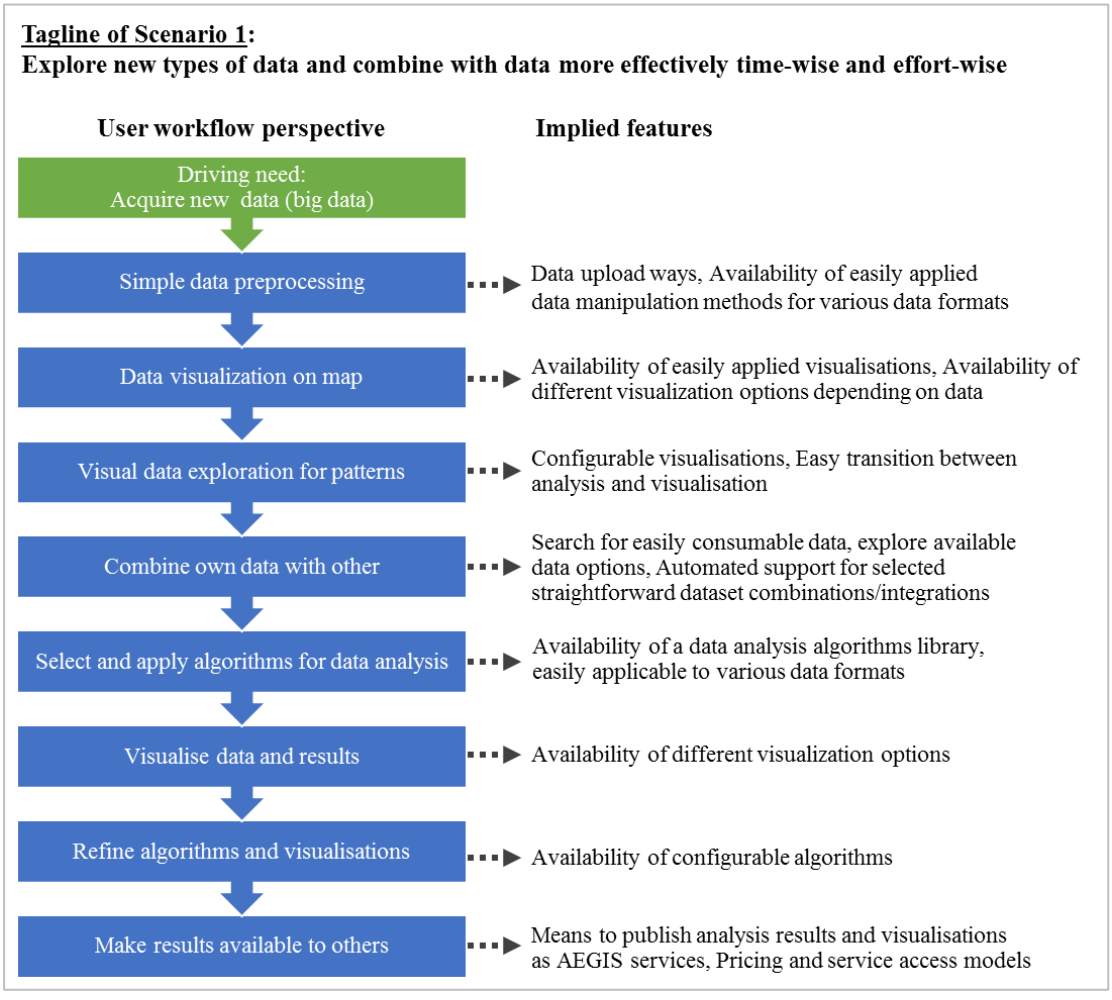

Figure 1: Features extracted from Scenario 1

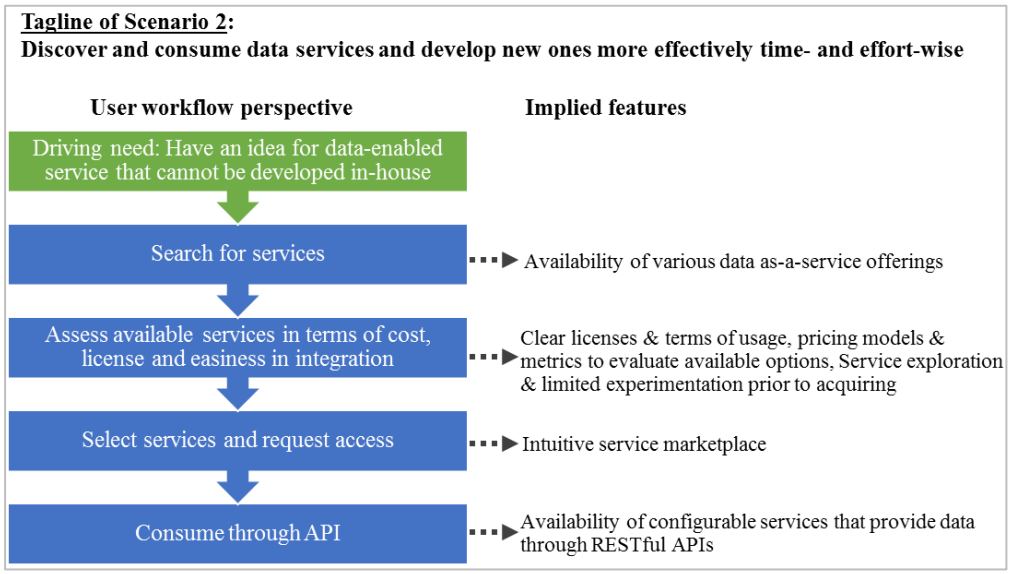

Figure 2: Features extracted from Scenario 2

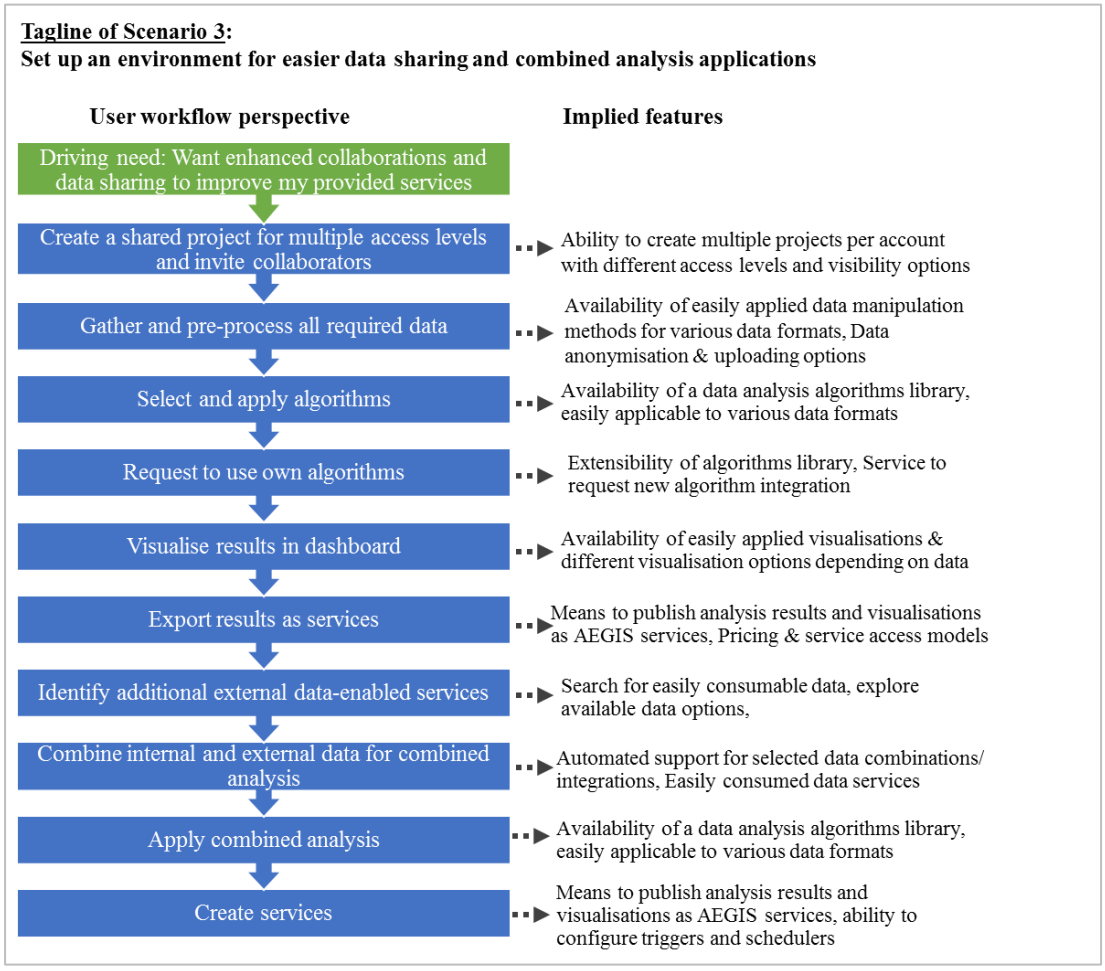

Figure 3: Features extracted from Scenario 3

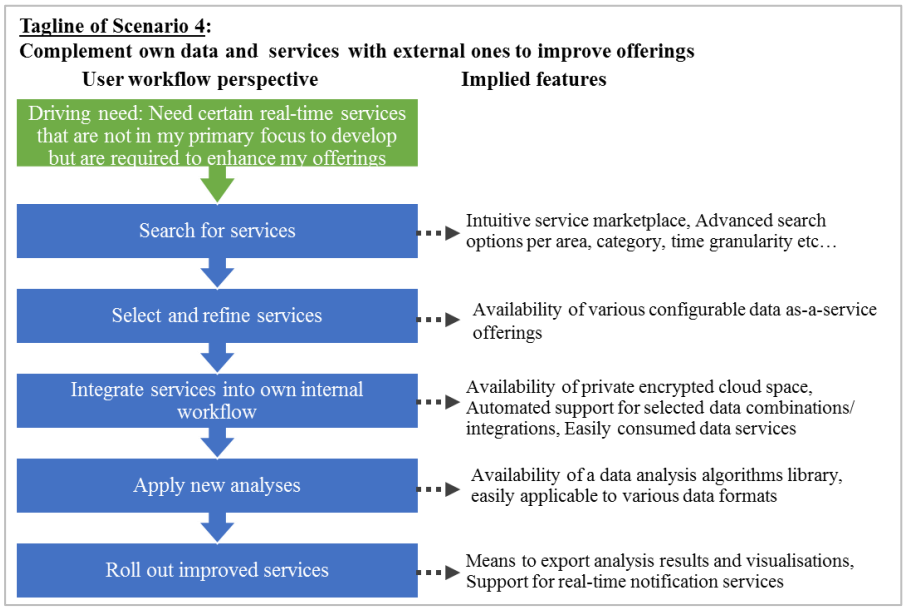

Figure 4: Features extracted from Scenario 4

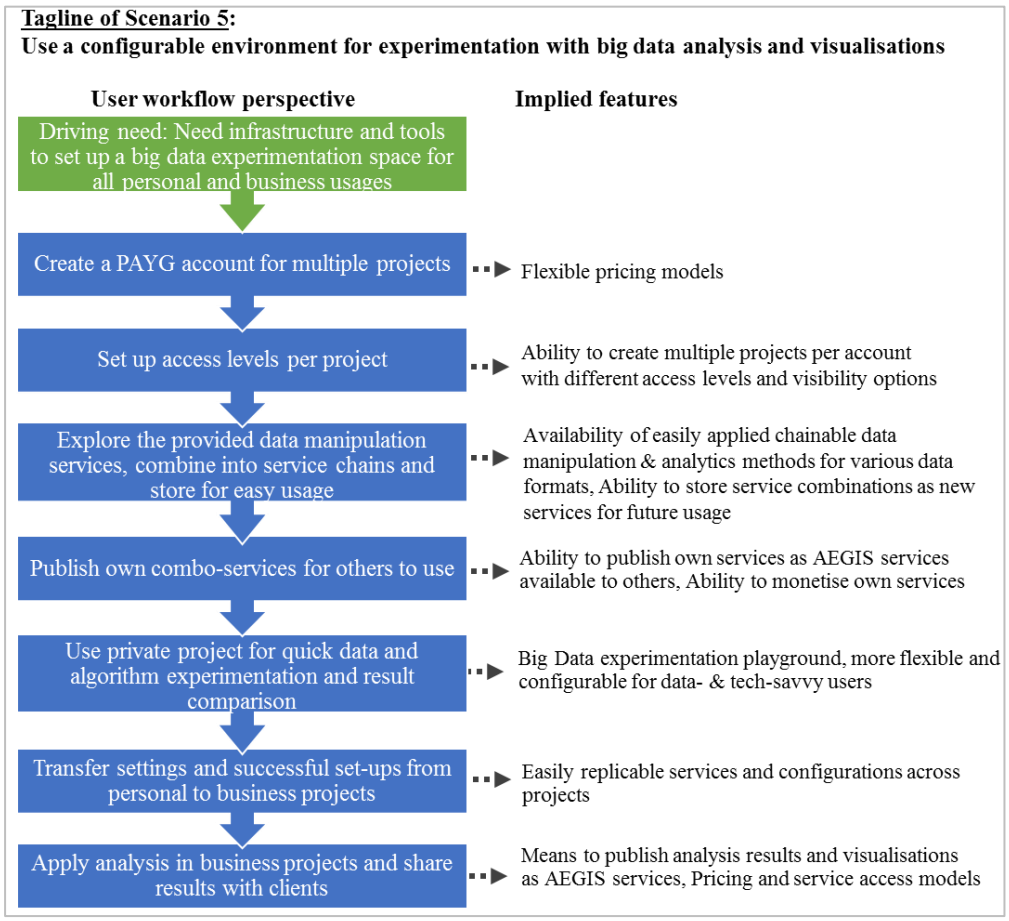

Figure 5: Features extracted from Scenario 5

Supplementary to the above functionalities, the described scenarios unveil an outlined need for discovering and consuming specific data-enabled services through AEGIS. Indicative examples of such services mentioned are: weather services, crime data services, event identification services and social media monitoring services. Furthermore, in order to achieve the envisioned level of automation in data processing methods and especially the need for seamless data combination identified in all scenarios, strong semantics need to be defined and applied in the background.

3.2. Integrated Features Diagram

Although, in order to allow out-of-the-box and productive thinking, no restrictions were imposed to ensure cohesion and compatibility of the scenarios in terms of common descriptions in the workflows, the above analysis shows that the scenarios do hint to a large number of shared functionalities. It should be noted that the extracted features do not substitute the user stories and detailed functional and technical requirements to be defined later in the project, but will serve as input to these processes. In order to highlight the relations among the extracted features and leverage them towards identifying the MVP, a grouping of the features and the supported processes by the AEGIS offerings is required. **Figure 6** presents an initial high-level grouping of the identified features based on whether they correspond to (a) the way projects and accounts in AEGIS platform are configured, (b) the core big data processing functionalities that are related to processes from the Big Data Value Chain described in D1.1 and (c) the way produced results (reports, visualisations, analytics, services etc.) are consumed.

Figure 6: Grouping of scenarios' features

Since most of the features belong to the Core Big Data Value Chain processes, features in this category have been further grouped (using colour-coding) based on their more specific goal. The four identified internal groups can be conceptually described as:

- Related to data and results visualisation
- Related to multi-source and multi-format configurable big data analysis
- Related to data-as-a-service discovery, exploration and acquiring
- Related to more advanced experimentation and configuration of the provided by AEGIS building blocks that address the needs of more tech-savvy and data-savvy users

Obviously, the above feature list is neither exhaustive nor, necessarily, accurate, but provides important insights into the workflows that drive the integrated AEGIS methodology for PSPS innovation and, ultimately, the features that the MVP should support.

3.3. Integrated Methodology

The previous analysis identified the expected interactions of the users with AEGIS, in various settings and for various purposes, and outlined a set of features and functionalities enabling them. Leveraging these insights, the current section presents the AEGIS generic information workflows, modelling interactions of the users with the AEGIS system in a unified way which constitutes the integrated methodology to be followed towards achieving PSPS data-enabled innovation.

The AEGIS users have been grouped under the following high-level categories:

Data provider: The user's main objective is to make her/his data available for processing or consuming in the AEGIS system.

Service provider: The user's main objective is to create a service on top of PSPS data that is available through the AEGIS system, leveraging the set of data processing, analysis, visualisation etc tools provided by the system. In this context, a service may be data (to be consumed as-a-service), visualisations, reports, dashboards, RESTful API endpoints etc.

Service Consumer: The user's main objective is to consume a service offered through AEGIS. In this context, this includes: accessing a visualisation through a link to AEGIS, downloading a report from AEGIS, performing requests to an AEGIS API endpoint etc.

Administrator (AEGIS): This user has advanced capabilities in the AEGIS system and may perform certain jobs that are not offered by the core platform (e.g. through advanced data curation tools that enable more fine-grained data manipulation and/or schema updates), that require extensions of the current system (e.g. adding a new custom algorithm) etc. This is an auxiliary role to highlight the need for non-automated functionalities in certain tasks.

It should be stressed that the categories are not mutually exclusive, but are used to better separate and describe the various workflows enabled in the AEGIS system. In an end-to-end usage of the AEGIS system, a user may transparently transit among the categories of service provider, service consumer and data provider.

Below, in **Figure 7** is presented the integrated AEGIS methodology diagram envisioning the high-level workflows. There are three more diagrams, each one featuring a specific workflow corresponding to one of the three main user roles.

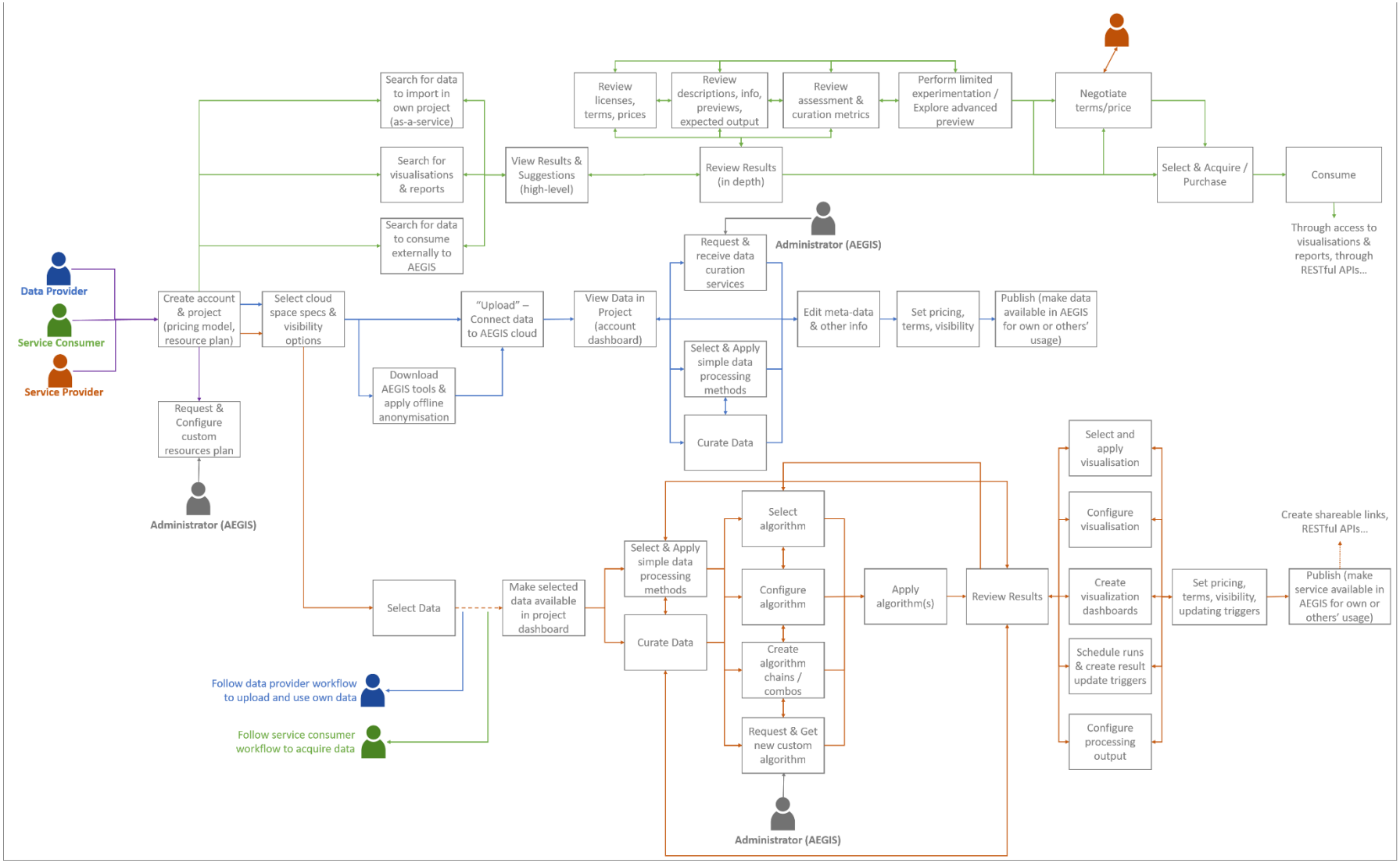

Figure 7: Integrated AEGIS Methodology

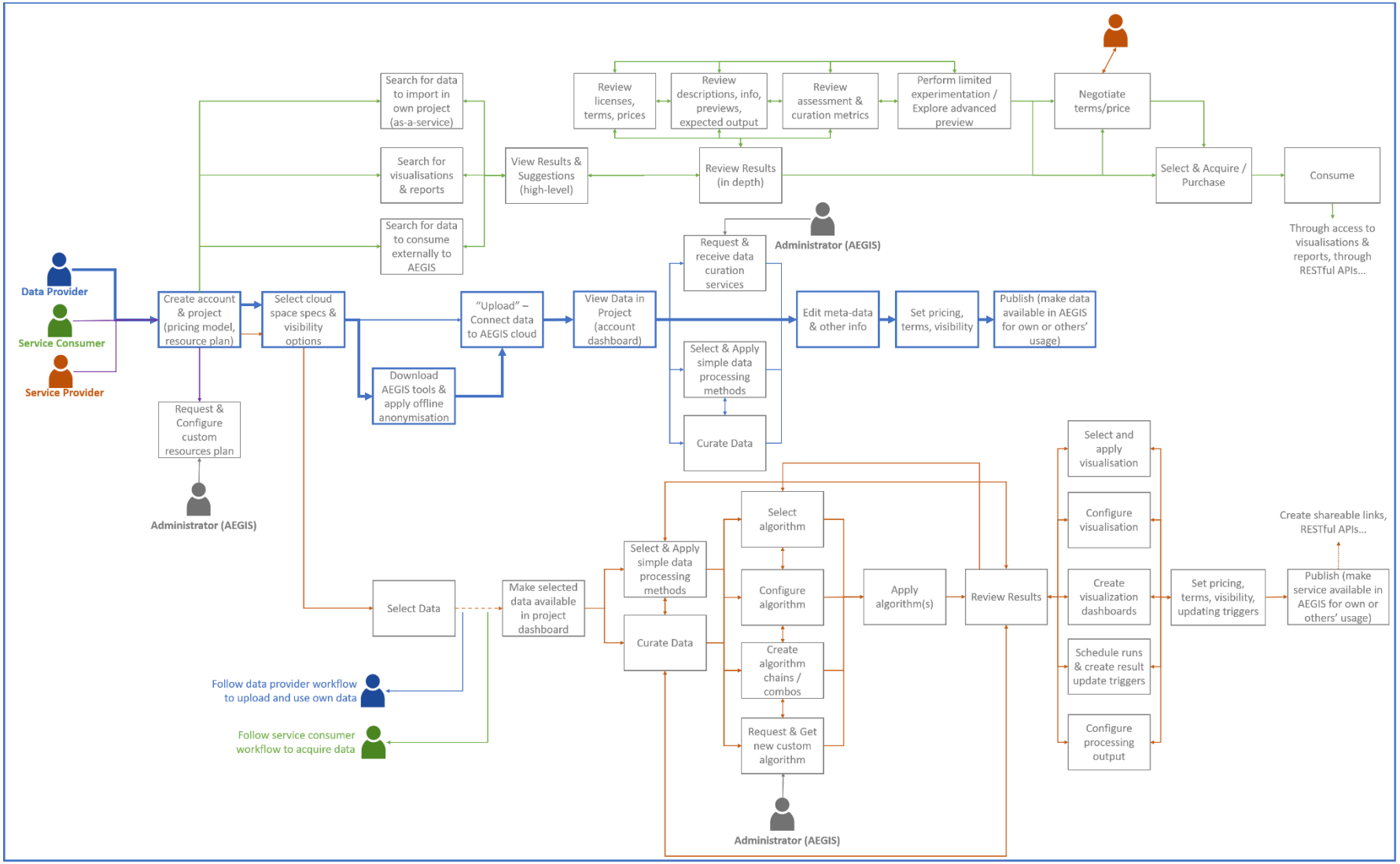

Figure 8: Indicative basic data provider workflow

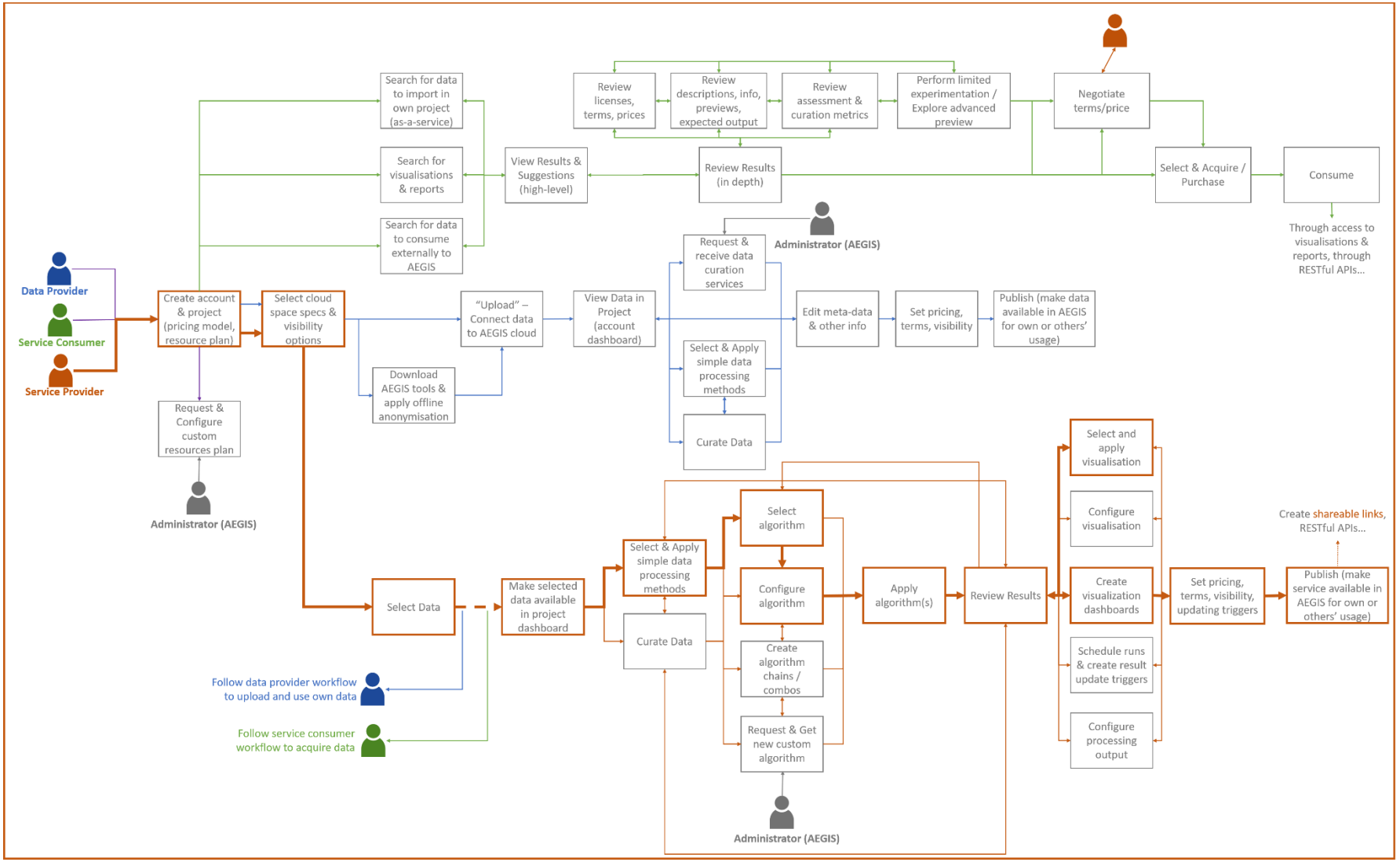

Figure 9: Indicative basic service provider workflow

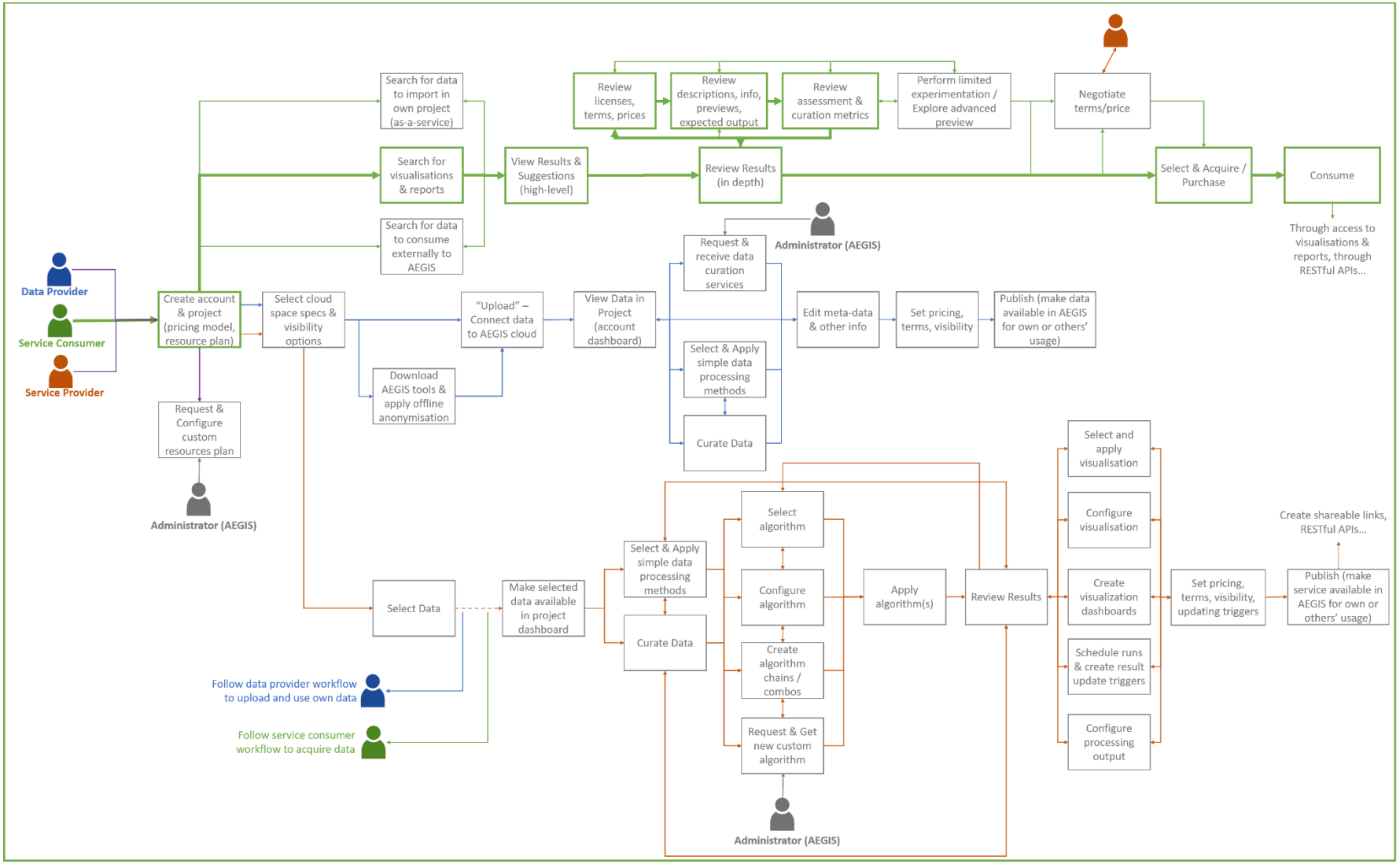

Figure 10: Indicative basic data consumer workflow

The previous diagrams correspond to the initial definition of the AEGIS integrated methodology. They will be further refined and extended as required for the final definition of the methodology, which will be presented in D1.3, marking the completion of WP1 activities.

3.4. MVP Features

Figure 11 presents a more detailed grouping of the identified features ((previously presented in Figure 6) under the specific processes they are related to, based on: the scenarios, the data value chain described in D1.1 and the envisioned workflows of an AEGIS platform user, presented in the previous section (**Figure 7**, **Figure 8**, **Figure 9**, **Figure 10**). The features shown in red font are the ones that have been primarily identified as parts of the MVP in order to ensure a balance between the ability of the platform to support an end-to-end workflow for data-enabled innovation in a PSPS application and the complexity of the concept to be used as a validation of the MVP hypothesis (i.e. whether the envisioned platform succeeds in addressing real stakeholders' problems).

More specifically, a set of qualitative criteria was applied, including importance for the AEGIS concept, foreseen implementation effort, dependency on other features, dependency for other features and expected usability for non-familiar users (i.e. all initial users who, by definition, cannot be familiar to the AEGIS offering prior to the MVP roll-out). Moreover, the integrated AEGIS methodology diagram was consulted to ensure that the selected features form non-interrupted workflows and that all essential components have been included. Hence, the current feature selection is expected to provide useful insights to the forthcoming tasks so as to ensure that the MVP serves its purpose in being both minimum but also viable, offering a complete experience of the AEGIS concept to the users.

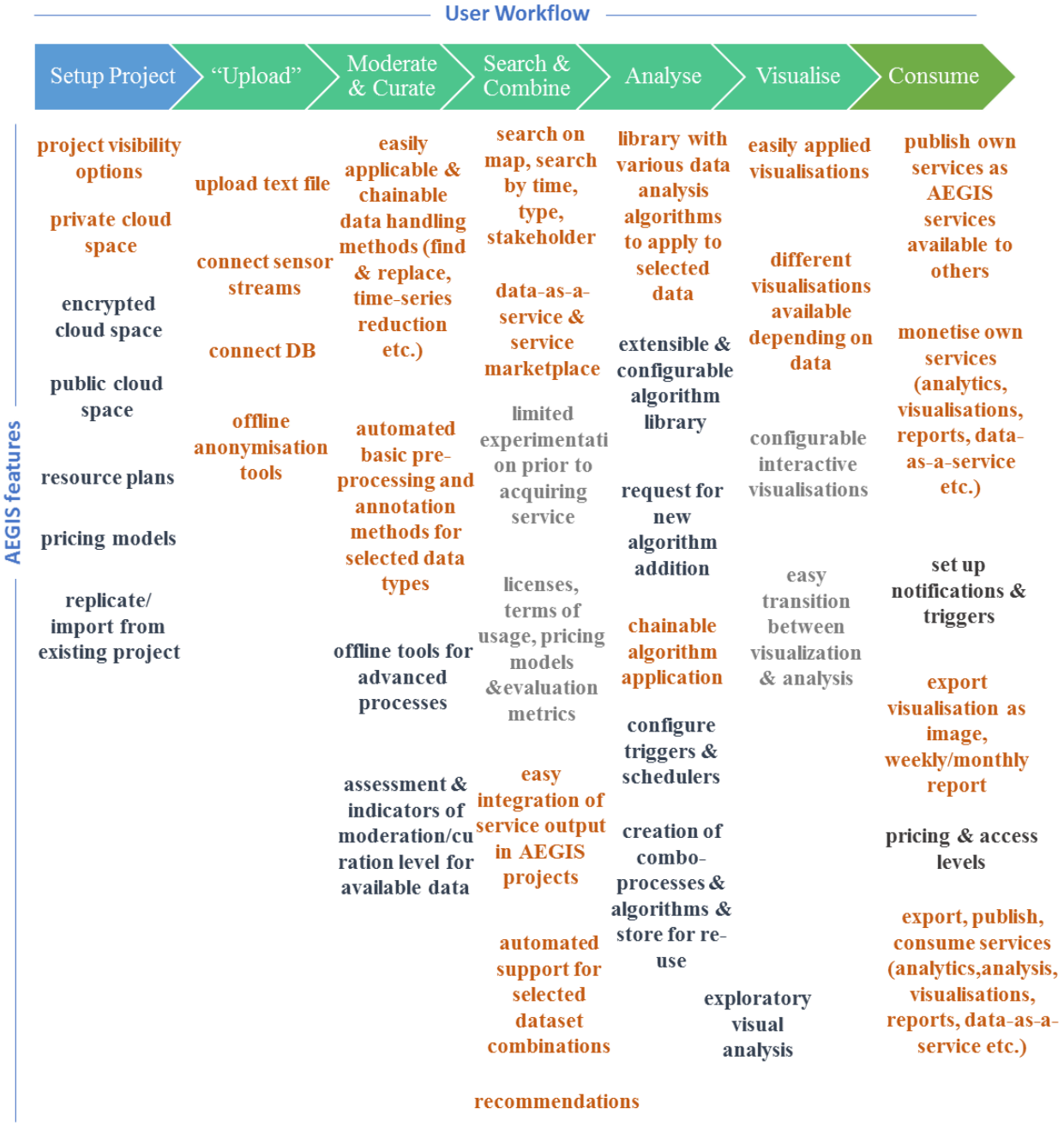

Figure 11: MVP features

Based on the above analysis, the AEGIS platform will mainly serve as a service marketplace and big data-enabled business intelligence creation space for all stakeholders across the PSPS value chain. The importance of providing flexible account/project configurations and collaboration and monetisation means cannot be overlooked, however the differentiating factor that will determine the success of the solution seems to lie in the ability to extract insights from available data in a PSPS context and roll-out innovative data-enabled services. Towards that end, AEGIS needs to mainly focus on how to:

(a) Offer flexibility in terms of data formats it can handle, pointing to an enhanced data import engine to support all commonly used data formats in PSPS applications (e.g. sensory data).

(b) Facilitate discoverability, acquiring and consumption of interesting data services and seamless combination under a PSPS semantic context. Towards this goal, AEGIS should provide a service marketplace for data, analytics and visualisation sharing and acquisition. AEGIS should also implement and provide several configurable data-as-a-service services to cover common stakeholder needs (e.g. for crime, news and weather data) enabling their easy consumption and ensuring they are compatible to be integrated in all analysis processes running on the platform.

(c) Enable the selection from predefined options and the application of various algorithms on the cloud targeting both generic and more specific domain needs. This entails the implementation of a powerful big data analytics engine, leveraging semantics and linked data in the background to deliver the required intelligence capabilities customised to the PSPS domain.

(d) Provide intuitive easy to create visualisations, through a set of available visualisation options, configurable to an extent and easy to combine in user created dashboards.

(e) Export the visualisations and analysis results for easier consumption and sharing with others. In order to achieve this, AEGIS should provide various export options, indicatively including creation of links to share/embed reports/visualisations and publishing analysis results through RESTful APIs.

The work presented here constitutes the first approach towards the definition of the AEGIS MVP. It should be stressed that the MVP is meant to capture the added value offering of AEGIS as a product and is, by design, abstracting the technical implications in order to highlight how the end user will utilise the project's offerings. That said, the above definition lacks certain features that need to be developed in order for the envisioned platform to function properly. Indicatively, semantic enrichment of the data is evidently an important step towards interlinking and analysis. However, since semantic enrichment is a crucial, yet background process (in the sense that it is not the goal of the end-user but a necessary step) it is not clearly described as part of the MVP. The process of technical requirements elicitation will highlight and address all such gaps in order to complement and connect the MVP components, paving the way from conceptualisation to realisation.
The current MVP definition will be further refined based on the user stories and collected user requirements (to be reported in WP3 deliverables), as well as the feedback to be retrieved both from the project's pilots, but also from the external to the consortium stakeholders. The final AEGIS MVP will be reported in D1.3.

4. AEGIS ETHICAL, PRIVACY, DATA PROTECTION AND IPR STRATEGY

4.1. Objectives

AEGIS Ethical, Privacy, Data Protection and IPR Strategy (in brief "EP Strategy"), outlined in this section, will serve:

i) to define the regulatory framework for data protection, IPR and Ethical Issues that will drive the Data Policy framework of the AEGIS platform and comply with EU directives on data safety and privacy;
ii) to illustrate an overview of AEGIS platform and components, focusing on portions of the system processing personal data, as well as representing the purpose of the processing of personal data and describing the origin of personal data and its collection method;
iii) to elicit the legal, data protection and ethical requirements (legal, technical, organisational, personnel and material requirements), providing input to the use cases, the architecture and specification task and specifying the measures to cover these requirements for data protection, and
iv) to assess to what extent they have been taken into account during project implementation and within the final AEGIS system.
v) to define ethics roles, procedures and roadmap.

4.2. Relations to internal AEGIS environment

AEGIS EP Strategy is strictly interrelated to the overall project implementation and final achievements, being aimed at providing the basis for the main guidelines that AEGIS Consortium will have to respect towards ethics, privacy and data protection, to be constantly updated during project's lifecycle. Its final release, notably regarding the Data Policy framework, will be delivered in D1.3 "Final AEGIS Methodology"

Given this, AEGIS EP overall Policy is particularly interconnected with:

- T2.2 "Data Policy and Business Brokerage Frameworks", because this is devoted to the design of the core methods for powering both the Data Policy Framework and the Business Brokerage Framework, including categories and predefined lists to describe data IPR, security, trust and quality features, as well as extra tag for the classification of personal and sensitive data, IPR annotations, and methods to cross-check IPRs and allow a semi-automatic negotiation;
- WP4 "AEGIS Infrastructure Implementation and Rollout" and WP5 "AEGIS Data Value Chain Early Community Demonstrators", because AEGIS EP Strategy supplies key input to the use cases, the architecture and the specification task, thus representing the reference point for assessing to what extent the legal and ethical requirements have been taken into account;

- T6.4 "Project Data Management Handling", since this task is expected to work in synergy with T1.4 and with WP9, though from different and complementary perspectives, in view of continuously monitoring the data protection and ethical issues of the project, as well as the IPR issues of the data to be contributed to the platform;
- T7.1 "Project and Demonstrators Exploitation Planning and Data Sharing IPR Definition", where a special focus should be given to the IPRs not only of the technology but also of the data to be exchanged over the AEGIS platform;
- WP9 "Ethics Requirements", pursuing the compliance with the listed set of "ethics requirements" that the project must comply with. A close connection is established particularly with D9.1 "OEI – Requirement N° 1", to be delivered at M18, where the EP Strategy described in this deliverable will be updated (if needed) and integrated with the overall Data Protection Impact Assessment methodology. Also, the other requirements set out in WP9 pertain to the ethical and legal concept are tackled in this document:
    - opinion or confirmation by the competent Institutional Data Protection Officer and/or authorisation or notification by the National Data Protection Authority, to be submitted where applicable (D9.2);
    - Ethics Advisory Board's periodic reports to the Commission on the implementation of the ethical concerns (issues) in project and on compliance with applicable national and EU regulations (D9.3).

4.3. Regulatory Framework

4.3.1. Introduction

AEGIS' use of technologies could potentially interfere with the right to privacy and the protection of personal data. It is therefore important to analyse the regulatory framework concerned and thus providing safeguards against the potential pervasiveness of AEGIS solutions, in order to design and develop them in a privacy-friendly fashion.

The main aim of the regulatory framework is to guarantee the individuals' sphere of autonomy within which to operate. The main legal instruments relevant to AEGIS pertain to privacy and data protection and contain a set of substantial safeguards and countermeasures against the spread of technologies resulting in an unfettered surveillance: the following chapters outline the key aspects of such regulations, relevant to project's progress and results.

Furthermore, in addition to legal provisions and principles, we will refer also to ethical, social and political oriented values applicable to AEGIS results and activities, being the "privacy in law" concept strongly interconnected not only with a number of legal values and principles -foreseeability, accountability, legality, necessity, proportionality and transparency, etc.-, but also with principles and values of ethical, cultural, social and political nature.

Within AEGIS EP Framework, and in AEGIS requirements' definition, it is therefore imperative to take all this set of variables into account in a systematic way.

The consideration of these principles will let us answer the questions why privacy matters in AEGIS R&D implementation and final system, how it should be safeguarded.

Before starting the overview, a remark has to be borne in mind: **this chapter has the ambition to look at the AEGIS project from a legal perspective**, and not to present a comprehensive analysis of the European regulatory framework of privacy and data protection - that would fall outside the scope of this deliverable.

The main documents that will be addressed are:

- European Convention of Human Rights
- Charter of Fundamental Rights of the European Union
- Regulation 2016/679/EU, repealing Directive 95/46/EC ("Data Protection Directive")
- Directive 2002/58/EC "ePrivacy Directive"
- Regulatory Framework in the selected jurisdictions, stating how privacy and data protection norms and principles are implemented in each country where the demonstrators will operate.

This composite regulatory system applicable to AEGIS is completed by European Courts' case law: though the legal system may appear somehow vague and fragmented, such a jurisprudence is very helpful for partially filling the gaps and pitfalls that can be found in legislation.

This overall framing represents the basis for setting the AEGIS ethical, privacy and data protection requirements, which will emphasise existing legal and ethical safeguards, boundaries and obligations to ensure the legitimacy and fairness of AEGIS final solutions and actions.

4.3.2. Privacy Concept and Data Protection Concept within the European regulatory system

As a starting point, it is useful to briefly examine the right to privacy and right to data protection concepts:

1. **Privacy concept**. Privacy is an ambiguous and contentious concept, varying according to time, space and peoples. It shifted from the "right to be let alone[1]", referring to the realm of intimacy and wish for solitude, as a concept hinging on physical privacy, to a broader notion of privacy, referring to the relationship between the individual and other individuals, based on "the claim of individuals, groups, or institutions to determine for themselves, when, how and to what extent information about them is communicated to others[2]". In this renovated meaning, the privacy concept encompassed several other aspects and embraces several rights, ranging from the right to be left alone and to enjoy solitude, to the right to individual autonomy, the right to control information about oneself, the right to a private life, the right to limit accessibility, the right to minimise intrusiveness, the right to exclusive control of access to private realms, the right to expect confidentiality, to the right to enjoy intimacy, reserve and anonymity and the right to

secrecy[3].

Both the European Convention of Human Rights and the European Court of Human Rights' consolidated jurisprudence recognise the right to privacy, promoting a living interpretation of the same, in the light of existing conditions.

2. **Data protection concept**. It was considered for a long time as a corollary of the right to privacy and is a relatively new autonomous human right in European legislation. This right had a new legal source of legitimacy in European legislation since the entry into force of the Lisbon Treaty. As recognised by the jurisprudence of the European Court of Human Rights and of the European Court of Justice, there is a tight relationship between privacy and data protection: the protection of personal data is functional to the enforcement of the right to privacy and, subsequently, the infringement of the individual's right to data protection leads to a violation of the right to privacy. Even so, these two rights don't totally correspond: not every privacy infringement results in a violation of the right to data protection. Data protection is more specific than privacy and is applicable every time personal data are processed.

4.3.3. European Convention of Human Rights and Charter of Fundamental Rights of the European Union

The recognition of privacy and data protection as fundamental human rights in Europe relies on the European Convention of Human Rights and the Charter of Fundamental Rights of the European Union, whilst at an international level, the Universal Declaration of Human Rights (1948) recognised the privacy as a fundamental human right by protecting territorial and communications privacy.

I. European Convention for the protection of human rights and fundamental freedoms

The European Convention of Human Rights (1950), in particular its Article 8, deals with private and family life, home and correspondence of the citizen. Since then, more enforceable European tools surpassed its value in the field of data privacy.

Article 8 recognises the privacy as one of the human rights and fundamental freedoms. It states as follows:

"1. Everyone has the right to respect for his private and family life, his home and his correspondence.

2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others."

The European Court of Human Rights' jurisprudence pointed out that private life concept extends to aspects relating to personal identity (e.g. an individual's name or picture, but also other means of personal identification and of linking to a family) and that therefore, the right to privacy established by this provision refers also to identity and personal development, also within interaction with other individuals, even in a public space, as well as to the right to establish, maintain and develop relationships with other human beings in general. This Court's case law also confronted with situations involving new technologies and its interpretation has to be taken into account in future AEGIS progress.

Article 8.2 states the lawfulness criterion, in the meaning of rule of law: it states a negative obligation for public authorities whilst allowing exceptions for interferences that are "in accordance with the law". Such rule of law is very important to ascertain the boundaries between the use of technologies (like AEGIS solutions) and democracy. The lawfulness criterion is the first step in assessing whether technological solutions are in line with Article 8.1, and it has to be applied on a case-by-case basis.

II. Charter of Fundamental Rights of the European Union

The Charter of Fundamental Rights of the European Union was proclaimed and published in December 2000 and then became legally binding in the EU Member States since the adoption of the Treaty of Lisbon on 1 December 2009.

The Charter refers to both the right to privacy and the right to data protection, containing an explicit right to respect for privacy (Article 7), as well as an explicit right to protection in case of personal data processing (Article 8). Both of these provisions have to be applied in coherence with European Court of Human Rights' interpretation of Article 8 of the European Convention on Human Rights.

Article 7 reads as follows:

"Everyone has the right to respect for his or her private and family life, home and communications".

Article 8 reads as follows:

1. "Everyone has the right to the protection of personal data concerning him or her.
2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.
3. Compliance with these rules shall be subject to control by an independent authority".

4.3.4. Regulation 2016/679/EU repealing Directive 95/46/EC "Data Protection Directive"

4.3.4.1. Reform of European Data Protection rules

The European Commission set forth a comprehensive reform of data protection rules in the EU, establishing common European rules to ensure that personal data enjoys a high standard of protection everywhere in the EU. The reform, published in the EU Official Journal on 4 May 2016, has resulted in two key pieces of legislation, of which especially the first is relevant to AEGIS:

- A **general Regulation on data protection (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. This regulation repeals the Directive 95/46/EC (General Data Protection Regulation) and, though entered into force on 24 May 2016, it shall apply as of 25 May 2018 (when it shall be binding in its entirety and directly applicable in all Member States);
- A specific **Directive on data protection in the area of police and justice (2016/680)**, 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 by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA. This Directive enters into force on 5 May 2016 and EU Member States have to transpose it into their national law by 6 May 2018.

In the meantime, the existing legislation (Directive 95/46/EC and Council Framework Decision 2008/977/JHA) remains applicable across the EU.

One of the main objectives of the reform is to give individuals back control over of their personal data and to act as key enabler of the Digital Single Market: personal data can only be gathered legally under strict conditions, for a legitimate purpose. Similarly to the repealed Directive, individuals or organisations collecting or managing personal information have to protect it from misuse and have to respect data subject's rights and the data subject is enabled to complain and obtain redress if his/her data is misused.

4.3.4.2. Definitions

Hereinafter, we will outline the concepts and definitions of personal data and data processing relevant to AEGIS, which remained substantially unchanged, and then consider the relevant Articles of the Data Protection Directive, comparing them from time to time to the new Regulation.

<u>Personal data and processing definitions</u>

Article 4 provides the following broad definitions of "personal data" and of "processing of personal data":

- "**Personal data** means 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". Therefore, the Regulation is applicable only to data subjects as natural persons, notably as human beings. According to this definition, personal data may be:
  - Identification data, which directly identifies the data subject, being pieces of information acting as identifying factors and able to distinguish a data subject from all the others;
  - Indirect identification data, which makes possible only an indirect identification of the data subject, through an association with other available information. The wording "other information available" refers both to other information available to the data controller (entity primarily in charge of the data processing) and to any information that may be possessed by any third party. It is considered sufficient, in view of the application of the regulation (as well as of the Directive), the potentiality of identification. Anonymous data, though not directly referring to a specific data subject, may keep this potentiality of identification.
- "**processing of personal data"** ("processing") means "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction".

Additional concepts relevant to AEGIS are as follows:

- "**anonymous data**", consisting of data that do not allow neither directly, nor indirectly, the identification of the data subject. As specified by Article 29 Data Protection Working Party, it is "any information relating to a natural person where the person cannot be identified, whether by the data controller or by any other person, taking account of all the means likely reasonably to be used either by the controller or by any other person to identify that individual. Anonymised data is anonymous data, which previously referred to an identifiable person, in case such an identification is no longer possible, usually thanks to processing and elaboration activities. This data does not fall within the EU data protection legislation. However, its first gathering, elaboration and processing were performed on personal data: therefore, data protection legislation has to be applied in such

activities, until data is made anonymous. It may also happen that, under certain circumstances, anonymised data receives protection in European member states' national data protection legislations or through Article 8 of the European Convention on Human Rights.

- "**Pseudonymised data**", consisting in personal data that, after its processing, become quasi-anonymous data: after such a processing, though there is the possibility to identify the data subject, the data Controller, according to the lawful data processing and data quality principles, makes the identification more difficult after their collection. In particular, the Regulation states that "pseudonymisation means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person". Therefore, though the use of pseudonymised and key-coded data is fostered by the European legislation to protect personal data (since it lowers the possible risks for the data subject), in case the data subject remains indirectly identifiable, this kind of data too is subject to application of the European regulatory instruments (Data Protection Directive and then Regulation). It should be noted, in fact, that in relation to key-coded data, Article 29 Data Protection Working Party followed the rule that, if the data subjects may be identified starting from the such data, "taking into account all the means likely reasonably to be used by the controller or any other person", it is personal data and therefore Data Protection Directive is applicable. The assessment has to be done on a case-by-case basis, considering all the specific circumstances concerned.

4.3.4.3. Main articles relevant to AEGIS, in a comparative perspective between the Data Protection Directive and the new Regulation

This paragraph depicts a survey of the main provisions of the Directive 95/46/EC, representing (till 25 May 2018) the main legal instrument relevant to AEGIS results and project implementation. This Directive, together with other two directives 2002/58/EC and 2006/24/EC and the corresponding provisions of the new Regulation, is the key for identifying the legal constraints that have to comply with and for examining their content and impact on the AEGIS project with a practical perspective. Article 89 of the Regulation provides that the "processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place…".

The following overview of the provisions of the Data Protection Directive highlights, when relevant, the main changes set for by the Regulation.

I. Acknowledgement of the data protection right

The European legislation provides the formal acknowledgement of the right to data protection as a fundamental right (Directive, Article 1, par. 1; Regulation, Recitals n. 1).

II. Key principles

Section I clarifies the key principles relating to data quality that have to drive data processing: they have to be taken into account in AEGIS too. According to Article 6, they are as follows:

- The fairness and lawfulness principles: this implies that, on the one hand, data processing has to be conducted with good faith, and malicious intents and behaviours are forbidden and, on the other hand, that it must comply with any applicable laws and regulations (including other than applicable privacy/data protection law).
- The purpose principle: "personal data must be… (b) collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes".
- The data quality principles:
    - Data relevancy principle: personal data processed have to be adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed.
    - Data accuracy principle: personal data must be… "d) accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that data which are inaccurate or incomplete, having regard to the purposes for which they were collected or for which they are further processed, are erased or rectified";
    - The limited retention of data principle: personal data must be… "e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed. Member States shall lay down appropriate safeguards for personal data stored for longer periods for historical, statistical or scientific use".

The Regulation, in Article 5, letter a), expressly refers to transparency as a key guiding principle for the processing of personal data, in addition to lawfulness and fairness. Moreover, it expressly refers to "data minimisation" (letter c). The letter d) mentions also the "accuracy" and letter f) explicitly establishes the "integrity and confidentiality" principle, in the meaning that personal data shall be "processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures". Article 25 of the Regulation refers to the principles of "Data protection by design and by default", expressly stating that, considering the set of circumstances, the controller shall implement appropriate technical and organisational measures:

- "such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects";
- "for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility".

Finally, it is important to mention here also the principle of accountability, which, besides requiring the active implementation of measures by controllers to promote and safeguard data protection in their processing activities, requires that the data controllers should be able at any time to demonstrate compliance with data protection provisions to data subjects, to the general public and to supervisory authorities.

III. Key figures of data processing: the data Controller and the Processor

Article 2 of the Directive lingers over two key figures of data processing:

- Data Controller: "natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data…". Therefore the data Controller may be a natural person or a legal entity, of both public and private nature. With regard to the same data processing, it is possible to have one or more data controllers.
- Processor: "natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller".

The whole chapter IV of the Regulation pertains to "Controller and processor" (Article 24 ss.), regulating:

- general obligations concerning, among other, the responsibility of the controller and the role of the processor, data protection by design and by default principle and related controller's duty, the case of joint controllers, the authority of the controller and of the processor, the record of processing and the cooperation with the supervisory authority (Section 1);
- the security of personal data, including the security of the processing, the notification of breach to the supervising authority and the communication of the same to the data subject (Section 2);
- data protection impact assessment and prior consultation (Section 3);
- the figure of the data protection officer, including his designation, position and tasks (Section 4);
- the codes of conduct and certification (Section 5).

IV. National law applicable

According to Article 4, "Each Member State shall apply the national provisions it adopts pursuant to this Directive to the processing of personal data where: a) the processing is carried out in the context of the activities of an establishment of the controller on the territory of the Member State; when the same controller is established on the territory of several Member States, he must take the necessary measures to ensure that each of these establishments complies with the obligations laid down by the national law applicable…". As regards AEGIS, par. 5.3.6 will provide an overview of the regulatory framework implementing European privacy and data protection legislations respectively in Italy, Austria and Greece, where the use cases and the demonstrator will be located.

The Regulation, unlike the Directive, "shall be binding in its entirety and directly applicable in all Member States".

V. The notification to the National Data Protection Body (NDPB)

Section IX regulates the notification to the competent national data protection authority, consisting in the formal communication from the Controller to such authority, in which the former provides specific and detailed information on the processing of personal data to be performed.

The implementation modalities of this notification requirement vary from country to country, for instance as to the means through which the notification has to be filed, the cases in which the notification is due and the amount of information to be provided. As regard AEGIS demonstrators, details on the notification procedures and bodies will be detailed in par. 5.4.2.

In relation to AGIS exploitation phase, is important to pay specific attention to Recital 89 of the Regulation. It states that the indiscriminate general notification obligations, provided by the Directive 95/46/EC, should be "replaced by effective procedures and mechanisms which focus instead on those types of processing operations which are likely to result in a high risk to the rights and freedoms of natural persons by virtue of their nature, scope, context and purposes. Such types of processing operations may be those which in, particular, involve using new technologies, or are of a new kind and where no data protection impact assessment has been carried out before by the controller, or where they become necessary in the light of the time that has elapsed since the initial processing".

The Chapter VI of such a Regulation is dedicated to the Independent Supervisory Authorities (Article 51 ss.), whilst art 68 regulates the "European Data Protection Board".

VI. Information to be provided by the Controller to the data subject

Transparency of the data processing towards the data subject is one of the most important principle to be fulfilled when collecting and processing personal data, in AEGIS too. The corresponding obligation to inform the data subjects cannot be exempted under national legislation, save for very limited circumstances, including the case that compliance with this information obligation results is impossible or requires a disproportionate effort for the Controller.

There is a list of minimum mandatory information to be given to the data subject, including: the purposes of the data processing; the categories of the data involved in the processing; the list of recipients (or of the categories of recipients) of data communications; data subject's right to access his/her personal data and to rectify them; the identity of the Controller and, if applicable, of his representative. In case the Controller intends to share personal data with third parties, the mandatory information must be given to the data subject no later than when such communication occurs.

The Regulation dedicates Chapter III to the "Rights of the data subjects", describing in detail transparency and its modalities, information and access to personal data. Article 13 lists the information to be provided where personal data are collected from the data subject.

VII. Criteria for data processing legitimacy

According to Article 7, the lawful performance of processing activities relies on i) the data subject's consent, thus authorising the data processing or ii) other external grounds imposing or requiring, ranging from law or contract provisions, to the need to protect data subject's vital interests or a public interest through the processing. When a legitimate interest is pursued by the Controller or a third party, data processing may be realised without consent, provided that this Controller's or third party's interest is not in contrast with other law provisions (there should be the acknowledgement by applicable law provisions that it earns safeguard). In any case, this legitimate interest should not override the right to data protection acknowledged under Article 1.

The Regulation (Article 6) is very similar to this provision and expressly establishes the conditions for consent (Article 7).

VIII. Special categories of processing: sensitive data and judicial data

The "special categories" of data, earning a higher degree of protection are:

- **Sensitive data**: "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life";
- **Judicial data**: data related to "offences, criminal convictions or security measures".

The lists are mandatory and closed: a personal data may not be considered as sensitive or judicial if it is not comprised within them. In case data processing performed on this kind of data, stricter requirements have to be fulfilled by the Controller and specific precautions are established. The analysis of them is outside the scope of this deliverable.

The issue is addressed also by the Regulation (Articles 9 and 10).

IX. Data subject's rights

The Directive describes the main privacy rights categories as follows:

- **The rights of information**, consisting in the data subject's right to be informed by the Controller on the purposes and conditions of the processing activities to be carried out on his personal data.

- **The rights of intervention**, allowing the data subject to ask that certain actions are performed on his data and also to interfere in the data processing. The list of exemptions and restrictions is provided by Article 13, whilst Article 14 pertains to the data subject's right of objection to the processing of his personal data and Article 15 refers to deployment of automated decisions processes, representing a threat for the data subject.

Chapter 3 of the Regulation disciplines the rights of data subjects, including transparency and its modalities (Section 1), information and access to personal data (Section 2), rectification and erasure, including the right to data portability (Section 3), the right to object and automated individual decision-making (Section 4) and applicable restrictions (Section 5).

X. Confidentiality and security of data processing

The confidentiality and security of the processing are key issues for personal data protection, to be tackled with high precaution in AEGIS: the risks and threats to which personal data undergoing processing activities are exposed are becoming higher (both in number and danger), notably with regard to Internet and automated data processing activities.

Security and confidentiality precautions aim at protecting personal data both in the static and in the dynamic moment of the data processing, including their storage in databases and their transfer to third parties,

Security measure may be technical (e.g. anti-virus, firewalls, authentication and authorisation systems), organisational (e.g. internal privacy policies, instructions or guidelines, internal procedures) or physical (e.g. measures to control access and ensure security of the Controller's premises).

According to Article 17, the security measures have to be adopted: "to protect personal data

against accidental or unlawful destruction or accidental loss, alteration, unauthorised disclosure or access, in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing.

Having regard to the state of the art and the cost of their implementation, such measures shall ensure a level of security appropriate to the risks represented by the processing and the nature of the data to be protected".

The Regulation addresses the issue in Articles 32, 33 and 34, stating that "taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk, including inter alia as appropriate… pseudonymisation and encryption of personal data, the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident" and other measures.

XI: procedure of "prior checking"

The Directive states that, before starting the data processing, the Controller must revert to the competent national data protection authority, which may authorise the processing and provide for specific safeguards and requirements. The modalities of such procedures are laid down by the national legislations.

In accordance with the aim of suppressing the "indiscriminate general notification obligations" (Recitals 89 mentioned hereabove), the Regulation replaces this obligation with this provision (Article 36): "the controller shall consult the supervisory authority prior to processing where a data protection impact assessment… indicates that the processing would result in a high risk in the absence of measures taken by the controller to mitigate the risk".

4.3.5. Directive 2002/58/EC "ePrivacy Directive"

The "ePrivacy Directive" (Directive 2002/58/EC on privacy and electronic communications) replaced the Directive 97/66/EC and was partially amended by Directive 2009/136/EC. The "ePrivacy Directive" pertains to the processing of personal data and the protection of privacy in the sector of electronic communications and transposes in the telecommunications sector, which is a "sensitive" area from a privacy perspective, the main principles and rules of the Data Protection Directive. Article 2 of the ePrivacy Directive expressly states to be aimed at particularising and complementing the Data Protection Directive.

Before the reform is applicable, the ePrivacy Directive, providing additional data protection rules for telecommunications networks and internet services, is expected to be repealed: the European

Commission adopted a proposal for a Regulation on 10 January 2017, which is currently under discussion in the European Parliament and the Council of the European Union. However, Article 95 of the Regulation 2016/679 states that there will not be additional obligations on natural or legal persons in relation to processing in connection with the provision of publicly available electronic communications services in public communication networks.

The main relevant provisions in relation to AEGIS are as outlined hereunder.

I. Security

Article 4. par. 1 sets forth to the obligation of adopting security measures: "the provider of a publicly available electronic communications service must take appropriate technical and organisational measures to safeguard security of its services, if necessary in conjunction with the provider of the public communications network with respect to network security. Having regard to the state of the art and the cost of their implementation, these measures shall ensure a level of security appropriate to the risk presented". The mandatory minimum precautions to be adopted were specified by the Directive 2009/136/EC, which amended Article 4.

The appropriateness of the security measures has to be assessed on a case by case basis, by making reference to the specific factual circumstances and conditions of the processing of personal data, to the state of the art technologies and to implementation costs.

In addition to these security obligations, the Controller, in case particular threats may occur for the network security, has to inform the users and also indicate possible remedies.

Article 4, as amended by the Directive 2009/136/EC, specifies the mandatory minimum precautions to be adopted. The security requirements should at least: i) "ensure that personal data can be accessed only by authorised personnel for legally authorised purposes; ii) protect personal data stored or transmitted against accidental or unlawful destruction, accidental loss or alteration, and unauthorised or unlawful storage, processing, access or disclosure, and, iii) ensure the implementation of a security policy with respect to the processing of personal data".

The Directive 2009/136/EC introduced also the definition of data breach, as follows: "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed in connection with the provision of a publicly available electronic communications service in the Community." In case of data breach, there is the 'duty to warn', consisting in the Controller's obligation to notify security breaches occurred in the course of the processing, detailing the procedures and rules for such a notification, towards both the competent national data protection authority and the interested data subject. The latter does not apply if the Controller adopted appropriate technological security measures that make data unintelligible to anyone who has no access authorisation.

II. Protection to confidentiality of the communications among individuals

According to Article 5, adequate protection has to be devoted to confidentiality in the communications. It may be limited only in case of specific circumstances.

An exemption to the prohibition of interception of communications (e.g. storing or other kinds of surveillance of communications and the related traffic data) occurs when such an interception is performed with specific precautions or by specifically authorised subjects (e.g. when users provided their consent, or applicable law provisions authorise it, or storage is functional to conveying the communications and without prejudice to the confidentiality principle).

The use of a deployment of electronic communications networks with the aim to store or have access to information kept in the user's terminal equipment is legitimate, provided that such user receives the mandatory information established by the Data Protection Directive and that he/she can oppose this data processing. The exemption is when this kind of activity is necessary from a technical point of view.

The protection of confidentiality of communications covers the communication itself, the users' terminal equipment (or other tool used by user to communicate electronically) and the information and data stored in such equipment and tools.

In case of deployment of invasive and tracking technologies (e.g. tags, spy wares, hidden identifiers and cookies), stringent provisions are set forth, considering the serious threat for users' privacy and confidentiality (e.g. it is possible to map and track users' online activities, to collect data from the technical equipment deployed).

As regards cookies, the user has to be provided with the mandatory information required under the Data Protection Directive and to be allowed to intervene on cookies, turning them down.

Similar provisions also apply to the other tracking technologies. Furthermore, it is necessary to have the data subject's consent for storage and gathering of information that is in turn stored on his/her terminal equipment (e.g. cookies and other tracking technologies), and user-friendly information has to be given to the data subject in order to enable him to willingly express his preferences, including his right to refusal.

III. Traffic data and location data

Traffic data is "any data processed in an electronic communications network, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service" (Article 2 letter C): therefore, it is personal information linked to communications and use of the Internet.

Given that traffic data poses serious concerns from a data protection standpoint and that possible potential threats regard concern surveillance, misuse and the pervasive encroaching into an individual's private sphere, its legitimate processing is subject to strict requirements.

Though some exceptions are indicated, "Traffic data relating to subscribers and users processed and stored by the provider of a public communications network or publicly available electronic communications service must be erased or made anonymous when it is no longer needed for the purpose of the transmission of a communication" (Article 6).

User' consent is considered as a tool for the protection of data subject's freedom of expression and rights to data protection and confidentiality in the communications. However, the set of mandatory information to be provided to the same, is larger than that identified under the Data Protection Directive (e.g. additional details on the types of traffic data collected and processed and on the specific time length of the processing activities). In addition, "processing of traffic data… must be restricted to persons acting under the authority of providers of the public communications networks and publicly available electronic communications services handling billing or traffic management, customer enquiries, fraud detection, marketing electronic communications services or providing a value added service, and must be restricted to what is necessary for the purposes of such activities" (Article 6).

The legitimate traffic data processing activities are only those strictly necessary and functional to achieve the specific legitimate purposes, and the traffic data may be kept and processed only for the time strictly necessary and functional to such purposes (this derives from the necessity, proportionality and time storage principles).

Location data are a type of traffic data. They "may refer to the latitude, longitude and altitude of the user's terminal equipment, to the direction of travel, to the level of accuracy of the location information, to the identification of the network cell in which the terminal equipment is located at a certain point in time and to the time the location information was recorded" (Recital 14). The concept was extended by the Directive 2009/136/EC, thus including also personal data processed by an electronic communications service.

Location data may be lawfully processed only "when they are made anonymous, or with the consent of the users or subscribers to the extent and for the duration necessary for the provision of a value added service" (Article 9). Informative requirement has to be followed as well (e.g. type of data, time length, extent, etc.).

A data subject can withdraw his/her consent at any time, and this kind of data may be accessed and processed only by persons under the authority of the Controller (or the third party providing value added services), whilst data collection and processing activities have to be limited to what is strictly necessary.

IV data retention

Article 15 refers to data retention. It assumed a key role since 2014, when the "Data Retention Directive" (Directive 2006/24/EC) was declared invalid by the Court of Justice because it did not meet the principle of proportionality and entailed a wide-ranging and particularly serious interference with fundamental rights. In fact, the retained data could provide a clear insight of data subject's private lives (e.g. his/her habits of everyday life, daily movements, frequent activities, social relationships, etc).

The annulment of Directive 2006/24/EC implied the need to refer to both ePrivacy Directive and to the guarantees of the European Convection on Human Rights and its interpretation.

The latter set forth the following principles:

- need to strict necessity and proportionality of collection, retention and transfer of data;
- rejection of the blanket data retention of unsuspicious persons and indefinite or even lengthy retention period of data retained;
- need of link between a threat to public security and the data retained for such purposes;
- need for effective procedural rules, like independent oversight and access control;
- need to address the risk of stigmatisation stemming from the inclusion of data in law enforcement databases

The ePrivacy Directive, in Article 15, par. 1, though gives Member States the possibility to exceptionally introduce data retention schemes deviating from the general prohibition to collect and store data, underlines the need to have a very strict and detailed measure of compatibility with fundamental rights standards, taking into account the formulation of Article 8 of the European Convention on Human Rights.

4.3.6. Regulatory Framework in the selected jurisdictions

The following paragraphs provide a concise overview of the regulatory framework implementing European privacy and data protection legislations respectively in Italy, Austria and Greece, the countries where the use cases and demonstrators will be located. Some of the information reported here corresponds with that inserted in D9.2, where a first snapshot of demonstrators' ethical, privacy and data protection concepts was provided.

4.3.6.1. Demonstrator 1: Road Safety Indicator

The automotive and road safety demonstrator will be developed in three versions, Broken Road Indicator, Safe Driving Indicator, and Regional Driving Style Risk Estimator. The three versions of the automotive demonstrator are aimed to provide the following benefits:

- Provide insights into road conditions based on exploiting individual vehicle sensor data, traffic data, and map data.

- Infer the driver's safety style and then calculate a safety index, through utilising vehicle sensor data along with environmental information and other content.
- Calculate a regional driving safety risk metric for certain regions including intersections, streets, cities or countries.

During the project runtime the automotive and road safety demonstrator will involve human participants as volunteers for

(a) generating driving data in the field (vehicle usage data) as well as in laboratory settings using a driving simulator (vehicle simulator data), and

(b) evaluating usefulness and usability of the developed services & applications running in a browser and/or on a mobile phone.

Before the experiments begin, an informed consent procedure will be applied. All participants who want to volunteer in the experiments of the automotive demonstrator have to sign a declaration of consent. Study participants will be made aware on of the project goals as well as of their role in the experiments. Each volunteer will be clearly informed on of the possibility to refuse to enter or to retract at any times with no consequences. All experiments will be designed and implemented according to the Data protection and privacy ethical guidelines from the European Commission[4] and to the main sources of national legislation relevant to AEGIS in Austria, in particular the "Datenschutzgesetz 2000 - DSG 2000" (Federal Act concerning the Protection of Personal Data), which is the current data protection act and the foundation of data protection law, the Telecommunications Act 2003 (TKG 2003) and Austrian Federal Constitutional Law.

4.3.6.2. Demonstrator 2: Smart Home and Assisted Living

Considering the specificities of the 2nd project demonstrator, the first step was to investigate and study the laws which are associated with the activities of the project. Beside the directives of the EU, the legislation of the countries where the demonstrator will be established (Greece) has been taken into consideration. Concisely, the legislation with which the AEGIS framework has to conform includes:

Greece – Law 2472/97 (amendments: 3471/06 & 3917/11)

The AEGIS project has to abide by the national laws of the countries that are involved in the pilots or in other activities of the project. In this section, **some key articles** will be mentioned underlying the legal and ethical scope of the AEGIS framework in the Smart Home and Assisted Living demonstrator.

1. An Authority (NDPA) has been created, as described in the following article, in order to enforce it.

Chapter D – Article 15

1. A Personal Data Protection Authority (hereinafter: the Authority) is hereby created with the task to supervise the implementation of this law and all other regulations pertaining to the protection of individuals from the processing of personal data as well as to the exercise of the duties assigned to it each time.

2. The Authority constitutes an independent public authority and will be assisted by its own Secretariat. The Authority shall not be subject to any administrative control. In the course of their duties the members of the Authority shall enjoy personal and functional independence. The Authority reports to the Minister of Justice and its seat is in Athens.

3. All necessary appropriations for the operation of the Authority shall be entered in a special code which shall be integrated in the annual Budget of the Ministry of Justice. The authorising officer for the expenditure is the President or his substitute.

2. Data Controllers must respect the provisions of Law 2472/1997 (and 3471/2006 regarding electronic communications) and more specifically:

They must collect personal data fairly and lawfully.

They must process only the data which are necessary for one or more specified purposes.

They must make sure that they keep data accurate and up to date.

They must retain data only for as long as is deemed necessary for the purpose of the collection and process thereof.

In order to carry out the data processing, the Controller must choose employees with relevant professional qualifications providing sufficient guarantees in terms of technical expertise and personal integrity to ensure such confidentiality.

The Controller must implement appropriate organisational and technical measures to secure data and protect them against accidental or unlawful destruction, accidental loss, alteration, unauthorised disclosure or access as well as any other form of unlawful processing.

If the data processing is carried out on behalf of the controller, by a person not dependent upon him, the relevant assignment must necessarily be in writing.

The controller must respect the data subject's rights to information, access and objection.

They must meet their obligations vis-a-vis the DPA (notification, granting of permit).

They must be kept informed on any Decisions, Directives or Recommendations issued by the DPA that may be important to them.

3. More specifically and based on Article 4 - Law 2472/97 (Characteristics of personal data):

1. Personal data, in order to be lawfully processed, must be: a) collected fairly and lawfully for specific, explicit and legitimate purposes and fairly and lawfully processed in view of such purposes. b) **adequate, relevant and not excessive** in relation to the purposes for which they are processed at any given time. c) **accurate and, where necessary, kept up to date**. d) kept in a **form which permits identification** of data subjects for no longer than the period required, according to the Authority, for the purposes for which such data were collected or processed.

Once this period of time is lapsed, the Authority may, by means of a reasoned decision, allow the maintenance of personal data for historical, scientific or statistical purposes, provided that it considers that the rights of the data subjects or even third parties are not violated in any given case.

2. It shall be for the Controller to ensure compliance with the provisions of the previous paragraph. Personal data, which have been collected or are being processed in breach of the previous paragraph, shall be destroyed, such destruction being the Controller's responsibility. The Authority, once such a breach is established, either ex officio or upon submission of a relevant complaint, shall order any such collection or processing ceased and the destruction of the personal data already collected or processed.

4. Article 6 defines the notification process towards contacting the NDPA for getting full consent about exploiting datasets.

The Controller must notify the Authority in writing about the establishment and operation of a file or the commencement of data processing.

In the course of the aforementioned notification, the Controller must necessarily declare the following:

a) his/her name, trade name or distinctive title, as well as his/her address. (The second item is deleted, as it is no longer valid)

b) the address where the file or the main hardware supporting the data processing are established.

c) the description of the purpose of the processing of personal data included or about to be included in the file.

d) the category of personal data that are being processed or about to be processed or included or about to be included in the file.

e) the time period during which s/he intends to carry out data processing or preserve the file.

f) the recipients or the categories of recipients to whom such personal data are or may be communicated.

g) any transfer and the purpose of such transfer of personal data to third countries.

h) the basic characteristics of the system and the safety measures taken for the protection of the file or data processing.

i) (The item was deleted pursuant to paragraph 2 of article 8 of Law 2819/2000, Official Gazette A/84)

3. The data referred to in the preceding paragraph will be registered with the Files and Data Processing Register kept by the Authority.

4. Any modification of the data referred to in paragraph 2 must be communicated in writing and without any undue delay by the Controller to the Authority'.

5. Article 7a- Exemption from the obligation to notify and receive a permit

1. The Controller is exempted from the obligation of notification, according to Article 6, and the obligation to receive a permit, according to Article 7 of the present Law in the following cases:

a. When the processing is carried out **exclusively for purposes relating directly to an employment** or project relationship or to the provision of services to the public sector and is necessary for the fulfilment of an obligation imposed by law or for the accomplishment of obligations arising from the aforementioned relationships, and upon prior announcement to the data subject.

b. When the processing involves clients' or suppliers' personal data, provided that such data are **neither transferred nor disclosed to third parties**. In order that this provision may be applied courts of justice and public authorities are not considered to be third parties, provided that such a transfer or disclosure is imposed by law or a judicial decision. Insurance companies, for all types of insurance, pharmaceutical companies, companies whose main activities involve trading of data, credit and financial institutions, such as banks and institutions issuing credit cards are not exempted from the obligation of notification.

c. When the processing is carried out by societies, enterprises, associations and political parties and relates to personal data of their members or companies, provided that the latter have given their consent and that such data are neither transferred nor disclosed to third parties. Members and partners are not considered to be third parties, provided that said transfer is carried out

among said members and partners for the purposes of the aforementioned legal entities or associations. Courts of justice and public authorities are not considered to be third parties, provided that such a transfer is imposed by law or a judicial decision.

d. When the processing involves medical data and is carried out by doctors or other persons rendering medical services a, provided that the Controller is bound by medical confidentiality or other obligation of professional secrecy, provided for in Law or code of practice, and data are neither transferred nor disclosed to third parties. In order for this provision to be applied, courts of justice and public authorities are not considered to be third parties, provided that such a transfer or disclosure is imposed by law or judicial decision.

e. When the processing is carried out by lawyers, notaries, unpaid land registrars and court officers or companies formed by the aforementioned and involves the provision of legal services to their clients, provided that the Controller and the members of the companies are bound by an obligation of confidentiality imposed by Law and that data are neither transferred nor disclosed to third parties, except for those cases where this is necessary and is directly related to the fulfilment of a client's mandate.

f. When the processing is carried out by judicial authorities or services, with the exception of the authorities referred to under item b of paragraph 2 of Article 3, in the framework of attributing justice or for their proper operation needs.

For further information please visit the Hellenic Data Protection Authority (www.dpa.gr).

While the laws establish some core principles both at European and National level, they do not establish clear lines for the field of research. The AEGIS consortium will abide by the above-mentioned legislation and will act with respect to the rights of any human being that is involved in the project either as a participant or not, according to the "*Data Protection and Privacy Ethical Guidelines*" of the Ethical Review in HORIZON 2020.

4.3.6.3. Demonstrator 3: Insurance Sector. Personalised Early Warning System for Asset Protection

The main source of regulation relevant for the Personalised Early Warning System for Asset Protection Demonstrator is the Italian Data Protection Code or Privacy Code (Legislative Decree n. 196/2003). It came into force on 1 January 2004 and superseded previous laws, in particular Data Protection Act 1996 n. 675/1996. In respect of this, the Privacy Code adopted a more practical approach, especially by removing all the previous requirements that resulted in mere formalities. The Data Protection Code, which is still in effect, was amended by a series of subsequent instruments.

The code, which is mainly applicable to all processing within the State and its territories, consists of three parts, respectively setting forth:

- the general data protection principles, applying to all organisations;
- additional measures that will need to be undertaken by organisations in certain areas (e.g. healthcare, telecommunications);
- sanctions and remedies.

The first Article of the Code expressly acknowledges that: "Everyone has the right to protection of personal data concerning himself".

The key guiding principles behind such Code are simplification, harmonisation, and effectiveness. Other important points are as follows:

- The codes encompasses the element of data minimisation and boosts organisations in making use of non-personal data whenever possible;
- Data subjects are allowed to exercise their rights and instigate proceedings in an easier manner, so that to better safeguard and promote their data protection rights. In relation to compliance and enforcement, in case data subject have been prevented from exercising his/her rights, he/she can settle disputes either through the courts or by lodging a complaint with the Garante;
- International data transfers (outside the EU), according to Article 42-45, on the one hand, businesses have to provide notification only when such a transfer is able to prejudice data subjects' rights, and, on the other hand, notifications have need not to be yearly resubmitted yearly. The transfer of processed personal data to a non-EU Member State shall also be permitted if it is authorised by the Garante on the basis of adequate safeguards for data subjects' rights;
- In case of processing of personal data, Article 26 of the Codes provides the need of the Garante's authorisation. "General Authorisations", targeted to industry sectors and/or specific categories of data, were issued by the Garante, in compliance of Article 40, to prevent private-sector data controllers from having to apply for ad-hoc authorisations;
- The processing operation related to electronic communication data is addressed in Title X "Electronic Communication". Here we can mention only some of its provisions:
    - Article 121 clearly defines the extent of application of the title: "processing of personal data in connection with the provision of publicly accessible electronic communication services on public communications networks".
    - Section 122 states that:
        - "1. Subject to paragraph 2, it shall be prohibited to use an electronic communication network to gain access to information stored in the terminal equipment of a subscriber or user, to store information or monitor operations performed by an user.
        - 2. The Code of conduct referred to in Article 133 shall lay down prerequisites and limitations for a provider of an electronic communication service to use the network in the manner described in paragraph 1 for specific, legitimate purposes related to technical storage for no longer than is strictly necessary to transmit a communication or provide a specific service as requested by a subscriber or user that has given his/her consent based on prior information as per Artcile Article 13, whereby purposes and duration of the processing shall have to be referred to in detail, clearly and accurately.
    - Section 123, in relation to traffic data, states that:

- "1. Traffic data relating to subscribers and users that are processed by the provider of a public communications network or publicly available electronic communications service shall be erased or made anonymous when they are no longer necessary for the purpose of transmitting the electronic communication, subject to paragraphs 2, 3 and 5.
- 2. Providers shall be allowed to process traffic data that are strictly necessary for subscriber billing and interconnection payments for a period not in excess of six months in order to provide evidence in case the bill is challenged or payment is to be pursued, subject to such additional retention as may be specifically necessary on account of a claim also lodged with judicial authorities.
- 3. For the purpose of marketing electronic communications services or for the provision of value added services, the provider of a publicly available electronic communications service may process the data referred to in paragraph 2 to the extent and for the duration necessary for such services or marketing, on condition that the subscriber or user to whom the data relate has given his/her consent. Such consent may be withdrawn at any time.
- 4. In providing the information referred to in Article 13, the service provider shall inform a subscriber or user on the nature of the traffic data processed as well as on duration of the processing for the purposes referred to in paragraphs 2 and 3.
- 5. Processing of traffic data shall be restricted to persons in charge of the processing who act — pursuant to Article 30 — directly under the authority of the provider of a publicly available electronic communications service or, where applicable, the provider of a public communications network and deal with billing or traffic management, customer enquiries, fraud detection, marketing of electronic communications or the provision of value-added services. Processing shall be restricted to what is absolutely necessary for the purposes of such activities and must allow identification of the person in charge of the processing who accesses the data, also by means of automated interrogation procedures…".

- Section 126, in relation to location data states that:
  - "1. Location data other than traffic data, relating to users or subscribers of public communications networks or publicly available electronic communications services, may only be processed when they are made anonymous, or with the prior consent of the users or subscribers, which may be withdrawn at any time, to the extent and for the duration necessary for the provision of a value added service.

- 2. The service provider must inform the users or subscribers, prior to obtaining their consent, of the type of location data other than traffic data which will be processed, of the purposes and duration of the processing and whether the data will be transmitted to a third party for the purpose of providing the value added service.
- 3. Where consent of the users or subscribers has been obtained for the processing of location data other than traffic data, the user or subscriber shall continue to have the possibility, using a simple means and free of charge, of requesting to temporarily refuse the processing of such data for each connection to the network or for each transmission of a communication.
- 4. Processing of location data other than traffic data in accordance with paragraphs 1, 2 and 3 shall be restricted to persons in charge of the processing acting pursuant to Section 30 under the authority of the provider of the publicly available communications service or, as the case may be, the public communications network or of the third party providing the value added service. Processing shall be restricted to what is necessary for the purposes of providing the value added service and must ensure identification of the persons in charge of the processing that access the data also by means of automated interrogation operations".

  - As regards traffic data retention other than for purposes of dealing with disputes over billing and subscriber services, according to Article 132 it is possible for communications service providers (CSPs) to retain traffic data for thirty months;
  - Article 133 and 134 deal with the codes of conduct and professional practice and enhance their importance in respect of the protection of personal data: their adoption is encouraged in highly significant sectors such as processing of data via the Internet.

- Title IV provides the definitions of the actors that perform the processing: data processor, controller and persons in charge of processing: Article 28. 29, 30;
- The security measures are set forth in Annex B;
- Article 13 refers to the set of information to be given to the data subject, orally or in writing. The usual practice is to provide him with a written information statement. Besides this, for traffic data (Article 123) and location data additional (Article 126), further information must be given. Only in restricted exemptions the Controller is exempted from the obligation of giving the information to the data subject (Article 13, par. 4).
- Article 23 and Article 24 respectively linger over the data subject's consent and excemptions. A data subject's consent has to be: express, free, specific, informed, given in advance, documented in writing in case of processing of personal data (the consent for sensitive data must be given through written instrument). In case of network

monitoring, it is relevant the specific purpose for which it is performed, to determini if there is or not the necessity to obtain the data subject' s consent. According to Articles 123 and 126, for the processing of traffic data and of location data, usually consent usually i s necessary, also for performance of value added services. As to sensitive data processing, it is necessary an authorisation issued by the Garante and data subject's written consent (save for limited exemptions).

- Title VII, in Article 42 – 45, deepens the transborder data flow and, in general, the transfer of data.

4.4. Project implementation phase

The AEGIS EP Strategy, based on the aforementioned regulatory framework, is structured into two main parts. The first moves around the project's implementation phase and refers to all the issues relevant during project's development, including ethics processes, Ethics Advisory Board's set-up and operations, AEGIS demonstrators, as well as an overview of Ethics procedures and Roadmap and hints for data protection impact assessment methodology. The second part refers mainly refers to AEGIS solutions and requirements to be complied with.

4.4.1. Ethics Advisory Board

The Ethics Advisory Board (EAB) will work closely with AEGIS Consortium during the course of the project on tackling ethical and data privacy issues that will have to do with the retrieval, the processing, and the retaining of these data. The EAB's role is directed to evaluate the AEGIS's progress and the results generated and supervise the operation of the project, in order to ensure that European and national regulations regarding data protection are fully observed and that the framework and its implementation adhere to a minimum set of ethical and legal requirements. At the same time EAB will advise the Project Partners how to proceed with the research activities in an ethically correct way and in compliance with the applicable legislations.

The EAB will be coordinated by Dr Maurizio Ferraris, as EAB Coordinator, who will be responsible for interfacing with it.

Upon demand of GFT, the EAB will perform the following activities:

a. provide expertise in specific ethics and privacy areas (as instructed by the Consortium and the EC) during the whole duration of the project and contribute to provide independent opinions and thoughts and to advise both the technical and the research partners on issues regarding the AEGIS methodology, the development of the platform and its components and the piloting operation.
b. contribute to propose the Assessment Methodology to be described in D9.1 and followed in WP1 and WP5, including, if opportune, the provision of templates at

an early stage and the coherence with the Ethical Risk Table already named in the AEGIS Annex I;

c. participate and/or contribute to AEGIS workshops or meetings, which will be conducted during the project;
d. co-create and/or review selected parts of the ethics and privacy related deliverables (e.g. Deliverable D1.2 - Aegis Methodology and High Level Usage Scenarios Aegis Methodology and High Level Usage Scenarios, Deliverable D6.3 - Data Management Handling Plan);
e. periodically report to the commission on the implementation of the ethical issues in project and compliance with applicable national and EU regulations. The Ethics Advisory Board's Report will summarise the evaluation activities of the Ethics Advisory Board and will contain the Ethics Advisory Board's recommendations. The reports will be based on a common assessment methodology as introduced in D9.1 and will be submitted as AEGIS Deliverable 9.3, as attachment to the AEGIS Periodical Reporting in Project Month 18 and, at the end of the Project, in an updated version as attachment to the AEGIS Periodical Reporting to be submitted in Project Month 30.

4.4.2. Demonstrators/use cases: initial ethics and data protection remarks

4.4.2.1. Demonstrator 1: Road Safety Indicator

The automotive and road safety demonstrator will be located in Austria. Hence the majority of vehicle usage data as well as simulation data will be collected involving participants in Austria, too. The responsible national data protection authority in Austria is Austrian Data Protection Authority (in German: 'Datenschutzbehörde'[5]), a governmental authority charged with data protection. The data protection authority is the Austrian supervisory authority for data protection, the equivalent of a national data protection commissioner in other countries.

Despite the automotive and road safety demonstrator in the AEGIS project will not involve processing any personal data, according to the corresponding business scenarios and business models developed in the project and aiming to scale these applications to the market, a future collection of personal data might be taken into account. A collection of personal data for establishing novel data-driven services in the automotive domain applies e.g. if a future user of one of these applications might link the data he or she generates during the operation of a vehicle with his or her social media accounts, e.g. to inform his social network about how he attained a safe driving style. A user might for instance use his or her Facebook or Twitter account to log in or to share information with peers, which requires a professional data protection concept to safeguard ethics and privacy for future exploitation. However, this only affects the post-project exploitation phase.

Nevertheless, in parallel to the activities conducted during the project runtime, Virtual Vehicle will therefore approach the Austrian National Data Protection Authority to discuss the requirements for data protection, if Virtual Vehicles foresees any linkage of personal data in the post-exploitation phase of the AEGIS project for services related to automotive and road safety building on the results of the AEGIS project. This will ensure that services developed in the post-project exploitation phase will be developed according to '**privacy by design**'.

Data to be collected during the experiments is **sensor data** and/or **simulation data**. Sensor data is generated through connecting a device developed at VIF 'termed vehicle data logger' to the onboard diagnostic (OBD2) interface of a car. Sensor data will include for instance vehicle speed, vehicle rpm, or vehicle acceleration to name a few types. Simulation data is generated by study participants using a driving simulator developed at VIF and may include many additional values. Both sensor data and simulation data has to be stored on a research server at VIF to allow the development of algorithms for inferring events including broken roads, patterns of safe and unsafe driving, or driving risks. Sensor and simulation data will be kept on this server till the end of the project.

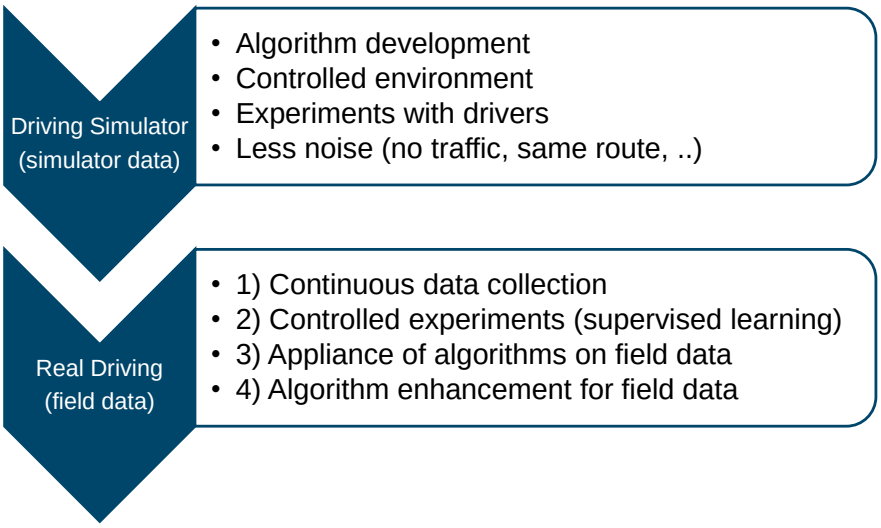

Figure 12: Simulator data and field data

During the AEGIS project, the automotive and road demonstrator involves the development and evaluation of applications running in a browser and/or on a mobile phone together with volunteers. During these automotive and road safety data related experiments, no identification data will be electronically stored on a server. Furthermore, no sensible personal data on health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction, etc. will be collected at all. The figure below shows data sources related for the automotive and road safety demonstrator.

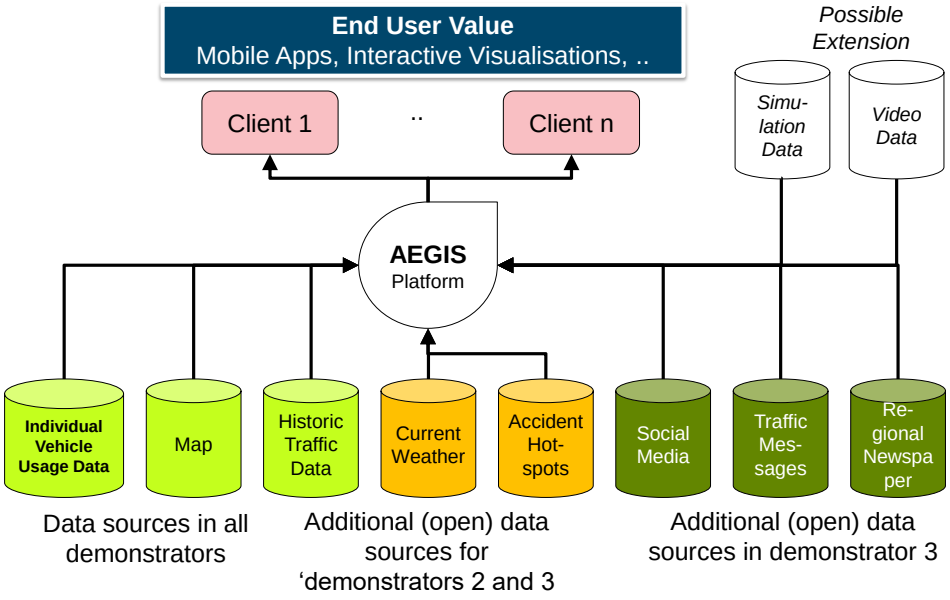

Figure 13: Data sources relevant to the automotive demonstrator

4.4.2.2. Demonstrator 2: Smart Home and Assisted Living

Towards the demonstration of Smart Home and Assisted Living services in AEGIS project, we are highlighting the list of ethics and data protection remarks by taking into account the preliminary list of use cases to be examined in the project. As part of the Smart Home and Ambient Assisted Living Demonstrator, three different applications are defined and developed:

- A mobile application acting as the **personalised guidance for elderly people**, leveraging multi-source and multi-lingual datasets to support elderly people in their daily routines and activities. The application will utilise and fuse data from multiple sources to safeguard elderly people's exposure to risks and threats.
- **Smart Home Automation for Security and Well-being** enhancement. Toward the deployment of human-centric, personalised smart automation strategies over their heating/ cooling and lighting devices, to ensure optimal comfort levels and compliance with special, ambience-related and health requirements. Sensing data (temperature, humidity, luminance, CO2, VOC, PIR) will be processed in combination with outdoor environment data to ensure (through automated control) that indoor ambient conditions always fit the personal preferences (visual and thermal comfort) of the elderly occupants.
- **Monitoring and Alert Services for 3rd Parties**. Apart from on-the-field services, AEGIS will develop and offer appropriate data-driven HMIs to social care services providers (public and private) to enable accurate monitoring of elderly activities and identification of critical incidents that may require on-the-spot physical interventions and assistance. Such 3rd party

services will utilise a variety of data, from smart -home sensors to wearable devices and from smart phone sensors (accelerometer) to social media information to ensure early and valid identification of frailty incidents, Alzheimer signs, signs of vision deteriorating based on changing light settings, etc..

Towards the demonstration of Smart Home and Assisted Living services, the following data types as retrieved from sensors and metering devices will be considered.

Occupancy (PIR)
Luminance
Indoor Air Quality
Indoor temperature and humidity
Control actions over lighting and HVAC
In-home Energy Footprint
Wearable Sensor Data
Smartphone Sensors (Accelerometer/ Gyro/ GPS)
Personal Health Data (Dummy data)

Figure 14: List of Datasets - Smart Home and Assisted Living Demonstrator

Concerning ethics and protection of personal data within the AEGIS project, we have defined two different approaches as specified also in D9.2:

1. Data protection measurements for Smart Home Demonstrator where non-personal (considering the anonymisation process as defined in this document) and not sensitive data are retrieved
2. Data protection measurements for Ambient Assisted Living Demonstrator where some of the data may be considered as sensitive data.

Although AEGIS will not have a direct interaction with human individuals (end users are the demonstrators - the demonstrator phase will deliver applications but will not be pushed to a group of selected users rather the tests will be performed as part of the research activities of the demonstrator), the consortium considers ethical issues a major topic of importance. The following bullet points have to be considered as the list of requirements, data protection requirements specific for the Demonstrator.

1. **It needs to be mentioned, that all of the ethics issues identified are already being handled by the demonstrator organisations during their daily operation activities,** as they confront with **national laws and EU directives** regarding the use of information in their daily services, as clearance for the processing, storing methods, data destruction, etc. has been provided to such organisation a priori and is not case specific. More specifically, the typical process for data handling in the specific demonstrator about Personal Data is that as follows:
   - processed fairly and lawfully;
   - collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes. Further processing of data for historical, statistical or scientific purposes is not considered as incompatible provided that Government provides appropriate safeguards;
   - adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed;
   - kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed.

   Therefore, the research to be done during AEGIS is not differentiated from the typical process in the Demonstrator organisation and thus does not raise any other ethical issues. More specifically, the legislation that the Smart Home and AAL demonstrator framework has to conform with is:
   a. European Union – Directives 95/46/EC & 2002/58/EC
   b. Greece – Law 2472/97 (amendments: 3471/06 & 3917/11)
2. It should be made explicitly clear that these data are going to be used solely for the **specific case (research activity)**, and will be completely destructed and removed from the AEGIS system after the case's finalisation. As such, the provision of such information to be entirely **voluntary**, and the provider to be **fully aware** of the purpose of the research of the data and the way it will be handled during and after the investigation. It needs to be noted that any data collection involving humans will be strictly held confidential at any time of the research. This means in detail that:
   - All the test subjects will be informed and given the opportunity to provide their consent to any monitoring and data acquisition process. The pilot tests supervisor will inform the participants with clarity about the procedure of the pilot tests, the system operation and the objectives, the data retrieval and storage and the exact dates the tests will be running.
   - All the subjects will be strictly volunteers and all test volunteers receive detailed oral information.
   - No data will be collected without the explicit informed consent of the individuals under observation and their legal guardian where applicable. This involves being open with participants about what they are involving themselves in and ensuring that they have agreed fully to the procedures/research being undertaken by giving their explicit consent.

Before the experiments start, an informed consent procedure will be applied. The respective declaration of consent is attached at the end of this document.

- No data collected will be sold or used for any purposes other than the current project.
- A data minimisation policy will be adopted at all levels of the project. This will ensure that no data which is not strictly necessary to the completion of the current study will be collected.
- If any shadow (ancillary) personal data obtained during the course of the research will be immediately cancelled.
- Specific measures will be in place in order to protect the pupils from a breach of privacy/confidentiality and any potential discrimination; In particular their names will not be made public and their participation will not be communicated to. Any incidental findings will be kept strictly confidential and erased from files under request from the enrolled subject.
- Participants will be able to quit the experiment at any point, if they wish, without any consequences. Participants will have the right to access their personal data as well as their extracted profiling parameters. He/she can exercise his/her right to access, correct and delete his/her data at any moment.

3. Moreover, every participant has the Right to obtain from the pilot controller
   - without constraint at reasonable intervals and without excessive delay or expense:
     - confirmation as to whether or not data relating to him are being processed and information at least as to the purposes of the processing, the categories of data concerned, and the recipients or categories of recipients to whom the data are disclosed,
     - communication to him in an intelligible form of the data undergoing processing and of any available information as to their source,
     - knowledge of the logic involved in any automatic processing of data concerning him;
   - as appropriate the rectification, erasure or blocking of data the processing of which does not comply with the provisions of the ethical manual, in particular because of the incomplete or inaccurate nature of the data;
   - notification to third parties to whom the data have been disclosed of any rectification, erasure or blocking, unless this proves impossible or involves a disproportionate effort.
4. During the Smart Home and AAL demonstrator, no identification data will be electronically stored on a server. This will be accomplished through an **anonymisation** of the datasets right at their **source for sensitive information** streams used within the AEGIS demonstrator, by removing the direct identifiers (e.g., name, address etc.), obfuscating sensitive information, etc. Furthermore, additional personal data (esp. secondary data, feedback, questionnaire/interview responses etc..) will be anonymised by means of aggregation and de-attribution as soon as they are not required in pseudonymised form anymore and processing and will be made on an anonymous basis.

- The data to be stored in the platform will anonymised and held securely using state of the art encryption methods. This applies for both the end users as well as for the stakeholders, hence the double pseudonymisation, so that a direct correlation between the individual and the stakeholder can be prevented.
- In addition, data will be **scrambled and abstracted** in a way that will not affect the final project outcome.
- No personal or sensitive data will be centrally stored. In contrary, a distributed database approach will be established where specific types of data will remain (still anonymised) in data source premises.
- The controller must implement appropriate technical and organisational measures (e.g. PET technologies) to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorised disclosure or access, in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing.
- After the end of the project, all collected data that can related to individuals will be deleted from the Platform

5. As part of the demonstrator, we may consider the integration of sensitive data (as defined in Section 4 of H2020 Guidance —How to complete your ethics self-assessment in line with EU Directive 95/46/EC). With regards to safeguarding privacy for this specific dataset, the demonstrator has considered along with the double pseudonymisation approach and the **generation of personas**. More specifically, personas will be created for the "grouping" of individuals with similar profiles (e.g. belonging to a similar age group and having correlatable medical profiles). This approach is not affecting at all the project outcomes, as the objective of the demonstrator is to test the technical implementation (at a research level) of the different services and applications mentioned above. Therefore, non-actual but "fake" pseudonymised individuals (thus **no actual sensitive data**) will be considered for testing the specific business functionalities of the project.

Smart Home and assisted living demonstrator evaluation requires the installation of equipment and usage of wearable devices. By taking into account the national legislation about the installation of sensors, we are presenting indicative guidelines in the field:

- All sensors utilised during the demonstrator should be privacy-preserving and should neither acquire sensitive personal data nor violate personnel's privacy.

- The controller of the study or his representative, if any, must notify the supervisory authority (Ethical Advisory Board) before carrying out any data collection process. The information to be given in the notification shall include at least:

    - the name and address of the controller and of his representative, if any;

- the purpose or purposes of the processing;
- a description of the category or categories of data subject and of the data or categories of data relating to them;
- the recipients or categories of recipient to whom the data might be disclosed;
- proposed transfers of data to third countries;
- a general description allowing a preliminary assessment to be made of the appropriateness of the measures taken to ensure security of processing.

- All offices/areas that will be monitored and controlled with any type of sensors and equipment should be appropriately marked with **Notification Posters**, describing in detail equipment used and monitoring procedures taking place towards project's objectives.
- All occupants, whose working offices/areas will be monitored during the pilot, should be thoroughly informed and their informed consent should be requested as specified above.

As a general remark, all experiments will be designed according to the Data protection and privacy ethical guidelines from the European Commission as defined in "H2020 Guidance — How to complete your ethics self-assessment". In addition, considering the need to have a clearance about any possible ethical concerns in the project, HYPERTECH (leader of Smart Home and AAL demonstrator) has contacted the national data protection authority in Greece to get a full commitment from HDPA about the AEGIS project activities. The sign from HDPA about the full clearance for AEGIS project activities will be available once received from the National Data Protection Authority.

4.4.2.3. Demonstrator 3: Insurance Sector. Personalised Early Warning System for Asset Protection

As outlined in D9.2, the Insurance sector demonstrator, through based on AEGIS technologies and the collection, knowledge of customer data and related real-time risk analysis, will lead to a more personalised mode of calculation of the risk associated with each customer, the provision of an alert system and new insurance models.

In this demonstrator, as in the others, volunteers will be involved and no identification data will be electronically stored on a server, thanks to the **anonymisation** of the datasets right at their source for sensitive information streams used within the AEGIS demonstrator, by removing the direct identifiers (e.g., name, address etc.), obfuscating sensitive information, etc. Furthermore, additional personal data will be anonymised by means of aggregation and de-attribution as soon

as they are not required in pseudonymised form anymore, and processing and will be made on an anonymous basis.

Information will come from diverse and heterogeneous data sources (geospatial information, photos, videos, social media, broadcasted news, etc.) and will be combined with the in-house big data platform of the insurance company. The information collected includes the passive collection of the data from the car itself (speed, driving style, fuel consumption, preferred roads, times and places of use, etc.).

In coherence with the project-level ethical, privacy and data protection overall strategy, a fine-tuning policy was elaborated for the Insurance Sector Demonstrator's Application by taking into account Italian regulatory system. It is fully described in D9.2.

Here it is important to remark the key requirements that have to be complied with, thus setting the frontiers of legally acceptable or affordable AEGIS measures and tools in the insurance sector demonstrator, with a particular focus on data processing.

- The Italian Informed Consent Procedure for gathering the volunteers' consent will meet the specific requirements set forth by the Italian Privacy Code. In particular:
    - Article 13 refers to the set of information to be given to the data subject, orally or in writing. The usual practice is to provide him with a written information statement. Besides this, for traffic data (Article 123) and location data additional (Article 126), further information must be given. Only in restricted exemptions the Controller is exempted from the obligation of giving the information to the data subject (Article 13, par. 4);
    - Article 23 and Article 24 respectively linger over the data subject' consent and excemptions. The data subject's consent has to be: express, free, specific, informed, given in advance, documented in writing in case of processing of personal data (the consent for sensitive data must be given in writing). In case of network monitoring, it is relevant the specific purpose for which it is performed, to determine if there is or not the necessity to obtain the data subject' s consent. According to Articles 123 and 126, for the processing of traffic data and of location data usually consent is necessary, also for performance establishing of value added services. As to sensitive data processing, it is necessary to have an authorisation issued by the Garante and data subject's written consent (save for limited exemptions).
- The security measures, as "Technical specifications on minimum data security measures", indicated by Annex B of the Privacy Code can be split in minimum and adequate measures. The first former represent the minimum standard to be adopted to have a lawful processing, while the otherslatter, though not not specifically defined by

the Code, are those considered suitable by the same Controller in relation to the specific processing having regard to the goal of minimising any possible risk that may jeopardise the personal data or that may harm the data subject. The general criteria to be followed by the Controller, according to the Code, is that, taking into consideration technological innovations, their nature and the specific features of the processing, personal data shall be kept and controlled in such a way as to minimise, by means of suitable preventative security measures, the risk of their destruction or loss (whether by accident or not), of unauthorised access to the data or of processing operations that are either unlawful or inconsistent with the processing purposes. For the processing of traffic data and location data, as written hereabove, stricter measures are compulsory (Article 123 and Article 126)[1]. These technical and organisational measures are also functionale to ensure anonymity.

- the Data Controller and Data Processors (and, in case, sub-processors, if any) will be appointed and the set of responsibilities set for by the legislation will be assigned to them.
- The notification procedure to the National Data Protection Body (NDPB) will be completed. The Italian NDPB is the so-named "Garante per la protezione dei dati personali". It is an independent Authorityauthority set up in 1997, with the function to ensure respect for individuals' dignity and to safeguard fundamental rights and freedoms in connection with the processing of personal data. The Garante is very active in this role and promotes a set of initiatives aimed at fostering the correct enforcement of the Privacy Code. Article 37 of the Code requires the notification to the Garante only in case of processing of higher-risk categories of data, by stating as follows: "1. A data controller shall notify the processing of personal data he/she intends to perform exclusively if said processing concerns:

  a) genetic data, biometric data, or other data disclosing geographic location of individuals or objects by means of an electronic communications network, …

  d) data processed with the help of electronic means aimed at profiling the data subject and/or his/her personality, analysing consumption patterns and/or choices, or monitoring use of electronic communications services except for such processing operations as are technically indispensable to deliver said services to users,…

[1] In 2008 the Garante issued a General Regulation on Security In Telephone And Internet Traffic Data, containing details on the physical, organizational and technical data security measures that have to be implemented with regard to the processing and storage of personal data

f) data stored in ad-hoc data banks managed by electronic means in connection with creditworthiness, assets and liabilities, appropriate performance of obligations, and unlawful and/or fraudulent conduct".

4.4.3. Ethics Procedures, Roadmap and Data Protection Impact Assessment Methodology

I. Composition, selection and appointment process status

The EAB will include relevant external, independent experts and practitioners with knowledge and experience regarding ethical and privacy issues. In particular, the following candidates have been selected and the procedure for their formal appointment and engagement in on-going:

- Prof. Gert G. Wagner
- Avv. Marina Da Bormida, PhD
- Ing. George D. Karagiannopoylos

The EAB would have been formed during the first month (M1) of the project. Though the candidates have been identified and preliminary interaction with them has already started, their appointment has not yet been done due to administrative and financial reasons. They have been properly addressed by the Consortium, in particular by the Coordinator in conjunction with GFT, and the formal appointment is expected within 2 months.

The EAB is formed through unanimous approval of all invited members by the Consortium partners and in agreement with the EC.

The Ethics Advisory Board will be coordinated by the EABC, who will be responsible for interfacing with the Ethics Advisory Board. This role has been assigned to Mr. Maurizio Ferraris (GFT).

At first, the Ethical Advisory Board Experts have to sign a Non-Disclosure Agreement (NDA), which was prepared by Fraunhofer, and immediately after the Expert Agreement (EA), which was prepared by GFT.

The template of both of them is included in this deliverable, respectively as Appendix B and C.

II. Reporting activities

The EAB will periodically report to the Commission on the implementation of the ethical issues in the project and on the compliance with applicable national and EU regulations. The EAB will submit the reports along with the periodic activity reports of WP8 (Coordination and Project Management), namely D8.2 at M18 and D8.3 at M30: the EAB reporting at M18 will timely ensure that the project is on the right tracks just before the completion of WP1 (AEGIS Data Value Chain Definition and Project Methodology), while at M30 just before the project end.

III. Ethics peer-review activities

If opportune, key project deliverables will be evaluated by the EAB's experts in the framework of their oversight activities. In this case, in order to gather diversified and balanced viewpoints, GFT will circulate the document to each of them separately and will collect their feedback. After this initial phase, the experts will be encouraged to discuss the concerns eventually identified and to collectively propose recommendations.

Deliverables in which ethical issues are involved and/or relevant from a privacy, data protection and ethics perspective can be identified as follows:

- Deliverable D1.2 "Aegis Methodology and High Level Usage Scenarios" (M6) and its updated release in D1.3 "Final AEGIS Methodology" (M15)
- D2.1 "Semantic Representations and Data Policy and Business Mediator Conventions" (M8)
- D2.3 Update on Semantic Representation and Data handling and Analytics Methods (M18)
- D5.2 Demonstrators Readiness Documentation and Execution Scenarios (M14)
- D5.6: Final Evaluation, Impact Assessment and Adoption Guidelines (M30)
- Deliverable D6.3 - Data Management Handling Plan (M6)

In case of need, also some parts of WP3-4-5 deliverables can be subjected to EAB's ethics peer-review.

IV. Extraordinary procedures (in case of ethical issues)

According to the DoA, in case of ethical issues partners consult:

1) at first, their own ethics departments

2) in a second time, the Ethics Advisory Board

The AEGIS partners will adhere to the recommendations of ethics departments and/or of the EAB and will implement the adequate mitigating actions, countermeasures necessary in order to reinforce ethical safeguards and fully comply with both ethical standards/best practices and regulatory obligations or constraints.

V. Ethics workshops

AEGIS consortium is considering to organise some public discussion of the privacy issues arising from the project research as part of the dissemination and public outreach activities

VI. Data Protection Impact Assessment

A comprehensive Data Protection Impact Assessment (DPIA) methodology will be elaborated in the framework of WP9, in particular in D9.1. This methodology will be strongly based on the **ethical, privacy and data protection requirements** as set forth in this deliverable and will contribute to reinforce the ethical safeguards, as well as to provide an in-depth exploration of the **societal consequences** (positive or negative) of the introduction of AEGIS system, as well as to approach data protection and ethical issues in a more comprehensive manner (going beyond the use of high-level data security solutions, as appropriately proposed by the project). As regards the exploration of the societal consequences (positive or negative) of the introduction of AEGIS system, it is important to bear in mind that trust reflects the sense of a general acceptance, in the meaning that the societal affirmation that in AEGIS a good equilibrium has been found between, on the one hand, individual privacy and ethical values and, on the other hand, interests as security, safety, economic growth. Otherwise, mistrust reflects exactly the opposite: the sense of a general unease and potential renunciation implying societal objection.

The DPIA should assess the particular **likelihood and severity of each risk** to data protection, taking into account "the nature, scope, context and purposes of the processing and the sources of the risk". The starting point will be the ethical risk table inserted into the DoA and referred to in D9.2. The impact assessment will also include "the measures, safeguards and mechanisms envisaged for mitigating each risk, ensuring the protection of personal data". The key questions driving the DPIA will include the following: what is gained, what is lost, by whom, how is this framed and measured and shared, by whom, and how is this articulated to decision-making processes related to AEGIS technologies?

The DPIA Framework is going to comprise the **assessment of the pros as well as the cons** (lock-ins, limits and constraints, SWOT analysis) of AEGIS technologies in general and of demonstrators' applications. The impact assessment will conduct **balancing assessment**, between, on the one hand, privacy/data protection tensions and, on the other hand, societal expectations and public interests related to PSPS solutions.

In relation to DPIA, it is useful to mention Article 35 of the new Regulation. It indicates that "where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data".

Lingering over such a methodology, the **mid-term and final assessment of AEGIS operations, framework and architecture** will be elaborated, in particular within the ethics report, assessing to what extent the legal requirements have been taken into account and offering recommendations

where appropriate.

4.5. Overall AEGIS platform and components

4.5.1. Methodology

Privacy-awareness and ethical compliance is one of the main objectives in designing, developing, and using AEGIS system: the system design takes privacy issues into appropriate account, bearing in mind that, from a wider perspective, a balancing operation has to be conducted between this kind of requirements and other kind of requirements (e.g. usability requirements, economic requirements). Consequently, the selection of ethical, privacy and data protection requirements and the assessment of their implementation play a pivotal role.

The approach taken for the identification and analysis of such requirements in AEGIS was not tackled from a purely legal perspective, but also rotates on the underlying ethical values, like individual's self determination, which implies both the possibility for individuals to be in control and data mininisation.

AEGIS approach was based on the combination of Privacy-by-Design method and Privacy Protection Goals method:

1. **Privacy by Design**: it addresses the design of the technical system as well as the business processes and relies on the idea that there is the need of putting privacy principles into the design process of data processing systems since the very beginning. The seven principles to be considered in the design process, as conceived by Cavoukian[6] are: "1. Proactive not reactive – preventative not remedial 2. Privacy as the default setting 3. Privacy embedded into design 4. Full functionality – positive-sum, not zero-sum 5. End-to-end security – full lifecycle protection 6. Visibility and transparency – keep it open 7. Respect for user privacy – keep it individual and user-centric".
2. **Privacy Protection Goal**: this approach, in which the private individual's point of view play a key role, considers the protection goals as central element for deriving requirements to be complied with in system design, as well as for identifying risks, countermeasures and in an evaluation perspective. Besides the well-known security protection goals named "Classic CIA Triad" (consisting of confidentiality, integrity, and availability), three further specific privacy protection goals are encompassed: unlinkability, transparency and intervenability. Protection goals promote the balance of the following privacy and security requirements against other protection goals:
    - **Confidentiality,** which refers to the protection of the information from disclosure to unauthorised parties. It can be ensured by measures/tools like encryption, enforcing file permissions and access control list to restrict access to sensitive information;

- **Integrity**, which lingers over the protection of the information from being modified by unauthorised parties. Commonly used methods to protect data integrity includes cryptography, hashing the data received and comparing it with the hash of the original message, use existing schemes such as GPG (GNU Privacy Guard) to digitally sign the data;
- **Availability** of information, which dwells upon the need to ensure the access to information by authorised parties when needed, at the right times. Data availability may be ensured, for instance, by backup, redundancy, off-site location ready to restore services relate to data centre;
- **Unlinkability**, aiming at separating data and processes, in order that processes are operated in such a way that the privacy-relevant data may not be linked across privacy domains or used for a different purpose than originally intended. The minimisation of possible infringements to the individual's privacy is connected to the minimisation of processing of personal data or, in case that data processing takes place, to the minimisation of the possible linkability and actual linkages. It may be obtained for instance by applying effective anonymisation, by separating data that are processed for different purposes, avoiding central points where personal data are or could be collected. This is coherent with the protection of the available data against misuse, where the focus is on the security protection goals.
- **Transparency**, directed to grant an adequate level of clarity of the personal data processes, including all privacy-relevant properties and actions, so that at any time it is possible to understand and reconstruct the collection, processing, and use of the information, both actual and planned. A sufficient level of transparency is a prerequisite for all kinds of control and intervention. Information has to be provided in form and extent adequate to the recipient of the information: in relation to different user groups, different ways of information concerning channels, granularity, language, etc., can be opportune.
- **Intervenability**, functional to assure that parties involved (in particular, data subjects, operators, and supervisory authorities) are able to interfere with the ongoing or planned data processing, including, if necessary, putting in place corrective measures and counterbalances, like data erasure, blocking or destruction, shutting off the system. Data subject's intervenability implies also the right to: i) withdraw consent, ii) obtain rectification and erasure of data; iii) lodge a claim or to raise a dispute to achieve remedy.

4.5.2. Key principles, legal evaluation and assessment of technologies in AEGIS

The legal evaluation of AEGIS technologies has to start with reflecting on their aim, being the presence of a legitimate aim the first requirement for the lawful data processing.

As indicated by the DoA and reminded in D9.2, data processing in AEGIS is functional to create a curated, semantically enhanced, interlinked & multilingual repository for public & personal safety-related big data, delivering a data-driven innovation that expands over multiple business sectors and considers structured, unstructured & multilingual datasets, rejuvenates existing models and facilitates organisations in the PSPS linked sectors to provide better and personalised services to their users.

By delivering services addressing the main challenges of cross-domain & multilingual applications through data identification, collection, harmonisation, storage & utilisation, the project aims to generate value and renovate PSPS sector. AEGIS technologies positively influence the welfare and protection of the general public and of individuals through prevention and protection from dangers affecting safety such as accidents or disasters. In this perspective, AEGIS solutions are aligned with the general interest and common good.

In fact, the project contributes to face some of the main PSPS' challenges, consisting i) in the lack of data discoverability and on the lack of a common structure and semantic model even for data that bear the same information type and come from similar sources and ii) in the lack of data and knowledge sharing mechanisms that in the case of safety issues are important to be properly exploited in order to timely disseminate key findings and promote the adoption of validated solutions. Project solutions, by introducing new business models through the breed of an open ecosystem of innovation & data sharing principles, will enable the creation of value chains towards more accurate risk models and proactive thinking and will revolutionise semantic technologies in big data, big data analytics & visualisations as well as security & privacy frameworks.

This aim is not only lawful, but also implies a set of positive impacts both for the society (both in terms of economic growth and of enhanced public security) and for the individual (mainly in terms of improved safety and well-being). The set of benefits derived from AEGIS data collection and processing will strengthen value generation for PSPS sector and includes, as reported in D9.2:

- Unified representation of knowledge;
- Accelerated, more effective & value packed cycles of intelligence extraction & of services & applications development;
- Introduction of novel business models for the data sharing economy & establishment of AEGIS as a prominent big data hub, utilising cryptocurrency algorithms to validate transactions & handle effectively IPRs, data quality & data privacy issues though a business brokerage framework. Besides capturing a portion of the total addressable market, AEGISis also expected to enlarge it by creating additional uncaptured value based on small data integration in typical big data repositories & algorithms.

In addition to such range of benefits, other advantages arise, in particular by facilitating and

promoting the collaboration in PSPS related domains, including public sector, insurance, environment, health, automotive, smart home, etc. AEGIS, in fact, will facilitate all companies and organisations in the PSPS linked sectors to provide better and personalised innovative services to their users, eventually of cross-domain nature and leveraging the plethora of data sources (from other domains) which, by adequate processing and combination, could further enhance and add value in the baseline services, and thus will allow smart collaborations for maximising the value offered to the end users.

In defining and, at a later stage, assessing and certifying the privacy-friendliness of AEGIS, the findings expressed by the European Group on Ethics in Science and New Technologies[7] play an important role. Such findings suggest to go beyond the traditional drastic trade-off between two goals, security/safety and freedom (including the right to privacy).

Analysing the trade-off narratives, we can see that some doctrine considers security as requirements of the state to protect the lives, welfare and basic freedoms of all citizens, and requires some trade-off between such rights to be protected and the freedom rights (including the right to privacy). Another doctrine argues that, being new technologies connected to competitiveness, jobs and economic growth, this requires to 'trade' away freedom rights, both at the policy level, for removing hindrances to the success of particular enterprises (premised on certain uses of big data like in AEGIS) and at the individual level, for exploiting the opportunities provided by such companies, especially online services.

The EGE Group believes that these framings underestimate the difficulty associated with the sensitive equilibrium between freedom and security/safety and constrain the reasoning, corralling it towards limited options and avenues, whereas it is necessary to open up new possibilities for thought as well as for individual and collective actions.

First of all, the EGE Group remarks that human dignity, which is intimately associated with freedom and responsibility "is the core principle of the European moral framework, and as such it cannot be traded off".

Given this, **the right to privacy and the right to data protection**, or the right to information and transparency, are not absolute rights. Therefore, such rights **must be balanced against other rights** and balanced against the rights of other persons or groups. Some kind of balancing, weighing, or choice between priorities is always necessary, in the meaning of need to find an equilibrium between rights of persons, on the one hand, and rights among persons, on the other hand.

In this regard, a rich jurisprudence of the European Court on Human Rights and the Court of Justice of the European Union[2] (ECJ/CJEU) has repeatedly stated that a balancing exercise with other rights is required when applying and interpreting Article 8 of the Charter of Fundamental Rights, setting forth the right to the protection of personal data.

This need for equilibrium and balancing is important for the legal evaluation and assessment of AEGIS technologies. As regards the opposite interests relevant for AEGIS technologies, they are, on the one hand, economic growth, in conjunction with public safety, well-being and personal security, and, on the other hand, the right to privacy and the right to data protection. Personal data collected within the three AEGIS demonstrators (e.g. by the use of tracking technologies) will be immediately anonymised through local dedicated services for anonymisation and filtering of data. These services will allow to process, anonymise the data and strip them of any private or sensitive information, on a local environment before uploading to private containers in order to avoid communication of any personal data outside of the data provider infrastructure. Therefore, AEGIS platform will not collect any personal data.

The AEGIS system overview and description of technologies was inserted in D9.2, Chapter 2, to which reference is recommended. Here it is important to start a preliminary assessment of AEGIS solutions from a legal, privacy, data protection and ethics viewpoint.

Besides the local dedicated services for anonymisation and filtering of data, here it is important to remark that in the framework of Data Aggregation and Harmonisation Layer, and in particular of AEGIS Data Value Chain Bus and of its annotations, specific micro-services will serve not only for delivering robust, flexible and tailor-made data handling operations and semantic tagging, but also for tagging data with different policies: the data policy library will be used to specify the visibility in terms also of security and privacy/trust levels, as well as of IPR clearance, of each dataset, or even dataset element (e.g. a field). This data tagging will be based on the Data Policy Framework.

Furthermore, in the Open Linked Big Data Space Layer, the produced output of the Data Aggregation and Harmonisation Layer is going to be stored in a public or private repository, depending on the type of the data and the policies of the corresponding SME/enterprise/organisation (in particular, during project implementation, this regards the demonstrators).

This is very important both from a privacy perspective and from an ethical perspective, taking into account also IPR issues: decisions on where to store the output/ harmonised data (within the

[2] E.g. CJEU, Joined cases C-92/09 and C-93/09, Volker and Markus Schecke GbR and Hartmut Eifert v. Land Hessen, 9 November 2010

public repository, namely the Security Linked Open Data - SLOD space, or private repository, in the meaning of internal repositories Aggregated Local Linked Data Space - ALLDS) will rely also upon security and privacy/trust level and on IPR issues of each dataset (or even dataset element). In other terms, the selection of the repository for the storage depends on the applied disclosure and IPR policy.

Considering that in AEGIS, information exchange among the SLOD space and the private repositories of each SME/enterprise/organisation is going to be supported and that publication/consumption of the produced linked data is going to be realised in real time, AEGIS Consortium's efforts in designing AEGIS platform has to be directed to make these operations compliant with the applied disclosure and IPR policy and to take safeguard measures to avoid or minimise any privacy risk or IPR infringement regarding the data stored in the internal repositories. This safeguard measures will be facilitated by the fact that this layer is interconnected with the microservices repository and the Data Policy Repository.

Finally, as regards the Business Intelligence and Analytics Layer, AEGIS Consortium will ensure that privacy-friendly and IPR-preserving modalities and tools are adopted when applying to private harmonised data stored into each private repository ALLDS and, notably, to produced linked data resulting from the information exchange among ALLDS and the SLOD space. The Partners will put great attention in defining how and where algorithms (such as Classification and Text Analysis Algorithms), data analysis tecniques and big data solutions (like Hadoop, Spark, Storm, Flink) can be applied for the extraction of linked data analytics from such data stored in private repositories or produced linked data with them.

Another consideration concerns the propagation to the SLOD space of the achievements generated by the analysis, which makes this knowledge re-usable in the future and leading to the design of advanced customised solutions. Such a propagation has to be aligned with the Data Policy Framework's discosure and IPR policies, considering the type of the data and the policies of each SME/enterprise/organisation whose internal dataset was used for the extraction of linked data analytics. This notably applies in case of linked data resulting from the information exchange among the SLOD space and the private repositories of an SME/enterprise/organisation. On this, special attention should be paid when the core offerings, which will be made available by the top level of the layer, will be offered to the various stakeholders.

In this perspective, the following elements are relevant. First of all, a key role will be especially played by the mechanism to be implemented by Business Broker Ontology, for analysing the Data Policies of each dataset under request and determining how these can be exchanged between different organisations. AEGIS will follow the notion of a virtual currency that will be used to safeguard the proper data sharing principles of the platform and will make use of blockchain technology, extending it with certificates that will be validated by the AEGIS ontology. Secondly, also the Open API Communication sub-layer will be useful for the aforementioned

purpose, being responsible for monitoring the usage and verifying that each transaction with the Business Access Layer is verified and thus accepted or rejected.

As underlined here above, it is evident that the set of services comprised by big data research in AEGIS opens promising avenues in terms of competitiveness, jobs and growth. AEGIS technologies (technologies of traceability, on-line applications, machine to machine communication, cross-correlation data analytics, predictive analytics and algorithms, etc.) touch not only on new ways to produce growth but also on new ways to produce knowledge, notably "intelligence" and scientific knowledge, as well as opportunities for the individuals. AEGIS intelligence-driven solutions fuelled by big data analytics represent a powerful tool for identifying trends, patterns, or relationships among data, for improving the individuals' quality of life safety and well-being, as well as for strengthening public security.

Nevertheless, as mentioned, they may give rise to ethical and privacy dilemmas. What is more important? Competitiveness, growth and jobs, public safety and personal security or privacy, data protection, informational self-determination, and individual freedoms?

However, rather than reasoning through a drastic trade-off paradigm, AEGIS partners prefer to concretely operate in line with the prioritisation approach fostered by the EGE's Group, based on the prioritisation of rights and interests, not giving up on any of the rights and interests and, finally, acknowledging that priorities may differ in different contexts (in particular the different sectors addressed by AEGIS). Following this prioritisation paradigm, AEGIS Ethical, Privacy and Data Protection Strategy, including requirements, has been conceived and will be implemented during project life and in the post-project phase in a way able to guarantee the proper handling of any ethical and privacy issues and the adherence to national, EU wide and international law and directives.

The pillars of it consist of:

1. focusing on notice and on choice (consent) of the data subject prior to data collection;
2. regulatory compliance and continuous legitimate ground of data processing;
3. setting ethical, privacy and data protection requirements to be complied with, elicitated through the Privacy Protection Goal approach, combined with the Privacy-by-Design approach;
4. elaboration of the Data Protection Impact Assessment Methodology, to be delivered in D9.1, including risk analysis and assessment scheme for evaluating the different proposed uses of AEGIS technologies, as well as a set of measures to minimise the privacy and ethics risks;
5. Technological fixes, including deidentification/anonymisation/ pseudonymisation of personal data: the AEGIS project is going to resort to privacy enhancing technologies, like this CloudTeams Anonymiser (developed by NTUA), and to design and develop its solutions relying upon the "Privacy by Design and by Default" approach.

4.5.3. Ethical, Privacy, Data Protection and IPR Requirements list

The handling and use of personal data is mainly regulated by the Data Protection Directive (that will be repealed by the new Regulation), setting out data subjects' rights and providing general rules on the lawfulness and fairness of the processing of personal data. Therefore, in the elicitation of AEGIS ethical, privacy and data protection requirements, references to it will be made. Nevertheless, considering the chosen holistic approach in setting these requirements, we considered other legal instruments applicable, such as the European fundamental rights framework and the national legislations applicable on a case-by-case basis, as well as ethics.

This requirements list clearly lays out a first guideline on how to conceive, develop and use AEGIS architecture and tools, without forgetting checkpoints. Anyhow, it only reflects an initial insight, which may be updated as the AEGIS architectural design develops: as AEGIS solutions and demonstrators are not yet shaped in their final fashion, the requirements list needs to be at a higher level of abstraction to cover various possible future technological choices. So far, input has been mainly taken from the DOW, literature and deliverable already produced. In a later stage of the project, for instance in D9.1, if the case we will refine or revise it.

The main AEGIS ethical, privacy and data protection requirements are as follows.

Number	Short name	Description	Assessment method	Phase	Notes
EPR.1	**Legitimate aim & purpose limitation**	This requirement implies that: i) AEGIS system and technologies have to serve a specific, explicit and legitimate aim; ii) the data have to be collected for such a purpose and not further processed in a way incompatible with that purpose; iii) adequate safeguards against misuse have to be taken. This requirement is also quoted by the DoA (Section 5.1.1): "No data collected will be sold or used for any purposes other than the current project".	DoA, D1.2 itself and D1.3, where AEGIS Methodology is respectively defined and updated, D2.1 and D2.3, where Data Policy, Data handling and Analytics Methods are respectively elaborated and refined, WP4 deliverables (D4.1-D4.4), which refer to AEGIS Platform in each of its improved releases; D9.2, Mid-term and Final Ethics Reports	All	Also the new Regulation's provisions refer to the legitimate purpose with substantially unchanged formulation. An extended analysis of the purpose of data processing in AEGIS is reported in D9.2
EPR.2	**Propotionality and data minimisation, including anonymisation**	The data minimisation principle is set forth by Article 6 of Data Protection Directive, as follows: "Personal data must be… adequate, relevant and not excessive in relation to the purpose for which they are collected and/or further processed". This requirement is also quoted by the DoA (Section 5.1.1): "A data minimisation policy will be adopted at all levels of the project and will be supervised by the Ethics Panel. This will ensure that no data which is not strictly necessary to the completion of the current study will be collected". The benefit potentially resulting from the use of that kind of data has to be clear. This requirement also implies adopting anonymisation as much as possible. The de-identification of datasets has to occur since the beginning of the processing: AEGIS datasets have to be stripped of any direct identifiers and, in addition, adequate technical and organisational safeguards have to be taken for mitigating the risks of re-identifying the individuals.	D2.2, D2.3, D3.1-D3.5, D4.1-D4.4, D5.6, D6.3, D6.5, mid-term and final Ethics Reports, D9.1	All	The principle of proportionality is expressly recognised by the Recital 4 of the new Regulation: "The right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality". AEGIS Data Policy framework will address also the issue of privacy and data anonymisation through specific micro-services to be developed.

		In the same perspective, this requirement implies minimising linkability and linkage: efforts have to be done to minimise possible linkability and actual linkages. Fostering unlikability in this way will reduce the risk of data breach and allow to safeguard the securing of the anonymity of the datasets. Regarding anonymisation, it is necessary to comply with what the DoA states: "The data to be stored in the platform will anonymised and held securely using state o the art encryption methods". Tools like the CloudTeams Anonymiser (developed by NTUA), allowing real-time efficient data anonymisation with cross domain scalability, have to be widely and timely adopted and used.			
EPR.3	**Data storage/retention minimisation**	The key rule is that "Personal data must be… kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed". After the Court of Justice' annulment of the Data Retention Directive, reference has to be made to each legal system concerned, which has its own rules on data retention. Therefore, data retention period relevant for AEGIS' demonstrators are those respectively stated by the legal system coming into relevance (e.g., for the insurance demonstrator, Italian regulatory framework). Access to the database has to be allowed only to authorised personnel, whose access is controlled through secure authentication techniques.	D2.2, D2.3, D5.6, D6.3, D6.5, Mid-term and Final Ethics Report	All	It is necessary to comply with what the DoA states: "After the end of the project, all collected data that can related to individuals will be deleted from the platform". Moreover, as regards demonstrators, the DoA specifies that "personal data will be used solely for the specific case, and will be completely destructed and removed from the AEGIS system after the case's finalisation". Regarding ancillary /shadow) data, though the plan is to minimise its gathering as much as possible, in case ancillary is obtained during the course of the research, it must be immediately cancelled.
EPR.4	**Avoidance of discrimination ,**	The Consortium has to avoid that AEGIS demonstrators or AEGIS overall system facilitate discrimination (race, gender, age, religion, disabled) or social sorting. Any possible different treatment has	Mid-term and Final Ethics Report	D, Ex	The European Charter of Fundamental Rights prohibits any kind of discrimination (Article 21).

	harm and social sorting	to rely on a rationale and project's solutions have to avoid to cause undue or unjustified harm to anyone, including wrongfully stigmatisation.[SEP]			
EPR.5	**Assignment of responsibilities**	The data controller has to be appointed, as well as the data processors and, in case, the data sub-processors. Also the data protection officer has to be designated by the controller and the processor in the circumstances set forth by Article 37 of the Regulation (Data Protection Reform). In relation to the role covered, each entity involved in the processing (data controller and data processor or sub-processor) is bound by obligations to be met and principles to be followed. Such obligations ensure that AEGIS data processing conforms to privacy laws and that the data subjects maintain the right to control what information is collected about them, how it is used, who has used it, who maintains it, and what purpose it is used for. Given that the main responsibility for data processing is in charge of the data controller, most duties and obligations are assigned to this figure, whilst the data processor has fewer and limited legal responsibility.	D5.6, Final Ethics Report	D, Ex	The reform maintains the set of provisions regulating the entities involved in data handling and adds the figure of the Data Protection Officer in the cases outlined in Article 37. Such cases include "b) the core activities of the controller or the processor consist of processing operations which, by virtue of their nature, their scope and/or their purposes, require regular and systematic monitoring of data subjects on a large scale" and the sensible data (e.g. health data)
EPR.6	**Informed Consent**	The data subject's informed, explicit and free given consent to the transmission and processing of their data is one of the criteria for rendering the data processing legitimate. Consent is principally explicit under the legal framerwork in force and is an important legal basis of lawful processing in AEGIS (particularly as regards sensitive data). Also when not required as legal ground, seeking consent in AEGIS has to be regarded as best practice. Article 2 (h) of Data Protection Directive provides the definition of consent: "the data subject's consent' shall mean any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed". According to this article, the following specific conditions make the consent valid: ➢ Unambiguity: unambiguous expression of data subject's wishes (no doubt should exist);	AEGIS Consent Form in D9.2, D5.6, Final Ethics Report	D, Ex	This requirement is based on the transparency principle. Article 4, 11) of the Regulation defines the "consent of the data subject" as "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her". Recitals 32 specifies that it can consist of a written statement, including by electronic means, or of an oral statement, provided that the data subject's behaviour clearly indicates his/her acceptance of the data processing. It is relevant to AEGIS also Recital 33 which states that, being often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection, data subjects should be allowed to give their consent to certain areas of scientific research (or parts of research projects) when in keeping with recognised ethical standards for

		➢ Specificity: expression must be intelligible and distinctive, referring "clearly, precisely to the scope and consequences of the data processing". This condition is closely related with the next requirement ("informed"). As an example of invalid consents we can refer to blanket consent; ➢ Information: consent has to be based on accurate, full and understandable information of all relevant issues, as indicated in Articles 10 and 11 of the Data Protection Directive (the nature of the data processed, purposes of the processing, the recipients of possible transfers, and data subject's rights); ➢ Free exercise of choice: the consent has to be freely given, in absence of any sort of intimidation, coercion or risk of negative consequences. This requirement is interlinked with the next requirement; ➢ Possibility of withdrawal: data subject may be able to change is mind and make a different choice at a later time, thus withdrawing the previously given consent and preventing any further processing. Withdrawal may not be retroactive; ➢ Timing: the consent has to be given before the starting of the processing; Consent form has also to be aligned with the applicable national legislation. Each voluntary participant to AEGIS demonstrators has to be provided with the clear information on AEGIS project and on the specific research activity related to the demonstration activity, as well as the information to be collected, how that information will be used and how to exercise his/her rights (e.g. of withdrawal).			scientific research. Recital 42 specifies that "…For consent to be informed, the data subject should be aware at least of the identity of the controller and the purposes of the processing for which the personal data are intended. Consent should not be regarded as freely given if the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment". Recital 54 clarifies that "The processing of special categories of personal data may be necessary for reasons of public interest in the areas of public health without consent of the data subject". However, suitable and specific measures in order to protect the rights and freedoms of natural persons have to be taken. Public health refers to "all elements related to health, namely health status, including morbidity and disability, the determinants having an effect on that health status, health care needs, resources allocated to health care, the provision of, and universal access to, health care as well as health care expenditure and financing, and the causes of mortality".
EPR.7	**Use of private environment/cloud as much as possible**	Being privacy and control more easily retained in a private environment, they should be used when possible for the storage or processing of personal data, in order to retain bigger control of the data being processed.	D3.2-D3.5, D4.1-D4.4	All	-

EPR.8	**Respect for data subject's rights**	The main categories of data subject's rights relevant to AEGIS can be split into two categories: - rights of information - rights of intervention (including rectification and erasure as well as, according to the new Regulation, data portability). This categories relies upon the intervenability protection goal and guiding principles, that encompasses the control exercised by the data subject and the other parties involved in AEGIS processing system. This includes the possibility for them to intervene if necessary. The chance to withdraw the consent can be attributed to this category. The first category comprises transparency or feedback of information, which refers to a set of data subject' rights, first of all his right to access the data stored and processed about him. Several provisions (Article 10, Article 11, Article 12, Article 13) outline this right, according to whether the information is collected directly from the data-subject himself or not. More details and clarifications on modalities and conditions, as well as exemptions, can be retrieved on Articles 12-15 of the Data Protection Directive.	D5.6, Final Ethics report.	D, Ex	The entire chapter III of the Regulation is dedicate to data subject's right and has to be taken into considerations: it describes transparency and its modalities (Section 1), information and access to personal data (Section 2), rectification and erasure (Section 3), the right to object and automated individual decision making (Section 4) and restrictions to the data controller and processor (Section 5). Transparency is considered fundamental by the new Regulation, that includes the same in the key principles.
EPR.9	**Data Quality, including Data Accuracy and Data Security**	The Regulation at Article 5 letter d) expressly refers to data accuracy, stating that "personal data shall be…accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay" (Article 5, letter d of the new Regulation). In AEGIS data accuracy has to be connected to the concept of data quality in data sharing and handling: predefined data handling policies have to be able to ensure data quality and trust. Data Quality, in a privacy-driven perspective, also requires Data Security and Integrity. Personal data shall be "processed in a manner that ensures	D2.1, D2.2, D2.4, D9.1, D5.6, Mid- term and Final Ethics report.	D, Ex	As regards Data Accuracy, this requirement relies upon ethical principles. On the other hand, Data Security and Integrity are two aspects encompassed by the CIA Triad protection goals, which, in addition to privacy protection goals, have been considered as essential for AEGIS methodology for the identification of privacy, data protection and ethical requirements. AEGIS data handling policies have to be able to ensure data quality and trust, besides privacy compliance. The quality level will be described by performing the necessary annotations both at dataset and on dataset element level: this has to be ensured by AEGIS Data Policy framework, which will be

		appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures" (Article 5, letter f of the new Regulation). According to level of security has to be appropriate to the risk taking into account "the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons" (Article 32 new Regulation). AEGIS has to use state-of-the-art technologies for secure storage, delivery, access and handling of personal information, for encryption and anonymisation, as well as for managing the rights of the users. It is necessary to have the complete guarantee that the accessed, delivered, stored and transmitted content will be managed by the right persons, with well-defined rights, at the right time. Where possible (depending on the facilities of each organisation) the data should be stored in a locked server, and all identification data should be stored separately. Tools for monitoring anomalies and activate restraint policy if needed should be used. The Data Policy Framework has to detail the security measures and other tools to be used for ensuring data protection and data quality.			used upon insertion of any kind of data into the platform.
EPR. 10	**Privacy by design and by default**	Security-by-Design, Privacy-by-Design, as well as Security-by-Default and Privacy-by-Default design methodology, has to be adopted in order to minimise the risks of compromising privacy. Efforts should be directed towards compliance with the voluntary standars developed by CEN-CENELEC/JWG 8 'Privacy management in products and services' for implementing data protection by design and by default rules and good practices, and more generally for privacy protection.	D3.2-D3.5, D4.1-D4.4	R	Privacy by Design is at the core of AEGIS approach for the elicitation of privacy and data protection requirements, whilst data protection by default is coherent with data minimisation requirement.

		According to Article 25 of the Regulation, the controller, considering a set of circumstances, shall implement appropriate technical and organisational measures: "such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects". Data protection by default ensures "that, by default, only personal data which are necessary for each specific purpose of the processing are processed".			
EPR. 11	**Record of processing activities**	"Each controller and, where applicable, the controller's representative, shall maintain a record of processing activities under its responsibility" (Article 30 new Regulation)	D5.6, Final Ethics Report	D, Ex	This is a provision of the reform, that specifies also the information that has to be contained in the recording.
EPR. 12	**Data protection impact assessment**	The need for a data protection impact assessment in AEGIS derives by the DoA (WP9), but is also coherent with Article 35 of the new Regulation. Article 35 states that "where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons, the controller shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data".	D9.1, Mid-term and Final Ethics Report	R, D	–
EPR. 13	**Application scrutiny to local/national boards if required by national legislation concerned**	As regards the demonstrator, "authorisation or notification by the National Data Protection Authority must be submitted, where applicable" (WP9). National legislations provide that data controllers and processors have to register at the competent authorities, in order to be allowed to process personal data, and impose differing national requirements for such a registration/authorisation, ranging from none to extensive authorisation processes. In most Member States registration for transfer to another EU Member State is not required,	D9.1, Mid-term and Final Ethics Report	D, Ex	Unlike the Data Protection Directive, the new Regulation doesn't provide "for a general obligation to notify the processing of personal data to the supervisory authorities". Such an obligation has to rely on "effective procedures and mechanisms which focus instead on those types of processing operations which are likely to result in a high risk to the rights and freedoms of natural persons by virtue of their nature, scope, context and purposes. Such types of processing operations may be those which in, particular, involve using

		unlike for cross-border data transfer, where additional or separate requirements may exist (e.g. registration or authorisation or mandatory additions to the standard contractual clauses).			new technologies, or are of a new kind and where no data protection impact assessment has been carried out before by the controller…" (Recital 89)
EPR. 14	**Confidentiality and access restriction**	People in charge of collecting, using or accessing personal data in AEGIS must be subject to an enforceable duty to keep them confidential and secure. Therefore, a confidentiality clause or agreement should be concluded by all research staff that will be having access to personal data in AEGIS. A closed user group has to be established, composed of only authorised persons, contractually obliged to keep confidentiality and meet data security rules. It is recommended an authentication and authorisation infrastructure in AEGIS. In addition to the technical measures that will be taken in view of ensuring confidentiality, publication of AEGIS result will not reveal the data subjects.	D9.1, Mid-term and Final Ethics Report	D, Ex	This requirement is ascribable also to ethical principles and to the chosen privacy protection goal as load-bearing method for elicitating and analysing data protection, privacy and ethical requirements in AEGIS. In particular, it is one of the three well-known security protection goals, named "Classic CIA Triad"
EPR.15	**Involvement of AEGIS Ethics Advisory Board**	This ethical requirement concerns the need to involve this committee to i) monitor ethical and legal issues in the project and report to the Commission; ii) work closely with the consortium in order to address the ethical and legal issues and data privacy concerns, that may arise from accessing user related information	D9.3, Mid-term and Final Ethics Report	D	AEGIS Ethics Advisory Board is expected to act as a sort of Data Protection Officer internal to the project and it has to periodically report to the Commission on the implementation of the ethical concerns (issues) in project and compliance with applicable national and EU regulations.
EPR. 16	**Set of requirements referring to the voluntary participation to AEGIS demonstrators**	The following requirements apply: i) AEGIS Recruitment Procedures for the selection of the voluntary participants for the AEGIS trials have to avoid any sort of discrimination/social sorting and be assessed by the Ethics Advisory Board of the project; ii) informed consent has to be obtained: partners must inform voluntaries and distribute the consent form, to be signed by each voluntary before trials' operations start; iii) Volunteers' dignity has to be safeguard and	D5.2, D5.6, D9.1, Mid-term and Final Ethics Report	All	_

		direct/indirect incentives for participation must not affect it.			
EPR. 17	**Adequate mechanism and tools for safeguarding IPRs on data artefacts and data usage**	This requirement calls for carefully addressing the data ownership aspect and for effectively handling IPRs of each dataset and dataset element.	D2.1, D2.2, D2.4, D9.1, D5.6, Mid- term and Final Ethics report	D	These requirements refers also to the emergent of the Human Data Interaction (HDI) topic, aiming at putting the human beings at the centre of the data driven industry and thus calling attention to address the data ownership aspect more carefully (e.g. who owns this data captured by the sensors? And who should have access to it?)

EPR: Ethical, Privacy and Data Protection Requirement
R: Research phase
D: Demonstration phase
Ex: Exploitation phase of the AEGIS system
All: all the phases, both during the project and after its end.

4.5.4. Guiding principles and recommendations for AEGIS Data Policy Framework

AEGIS cloud based Big Data solution is going to become a distributed database of secure transactions, removing the need for centralised ledgers, trusted parties copies and manual interventions.

AEGIS data sharing, homogenisation and reusability value chain will allow the exchange and documentation of data in a unanimously understandable manner, facilitating knowledge exchange.

It will therefore rely on collaboration activities of diverse sectors and stakeholders. Therefore, AEGIS will offer services supporting data IPR handling, security and privacy. These services will be able to overcome the limits characterising existing solutions, where data confidentiality, privacy protection and IPRs hinder the ability to exchange information in a trustful and transparent manner.

AEGIS Blockchain powered Security, Privacy, Quality and IPR Data Policy Framework (DPF) will power AEGIS Platform with a methodology encompassing aspects related to the exchange of data from business value point of view, focusing on the quality, the IPRs and the privacy of data. Indeed, AEGIS DPF is exactly devoted to set the appropriate security, data privacy, data quality probing and IPR policies to resolve on-the fly how data can be handled by each stakeholder group, based on its content, its value and peer-to-peer agreements that will be reached between the collaborating entities. Hence, AEGIS DPF has to allow the creation of a trustful and rigorous data sharing community, by focusing on data anonymisation and privacy preservation, secure data channels, IPRs on data artefacts and data usage, as well as data quality, going beyond current practises in data sharing and handling.

In this way, organisations, including those previously reluctant, are expected to contribute to AEGIS new data sharing ecosystem around Pubic Safety PSPS related information: stakeholders will be allowed to securely exchange data, on the basis on national, international and business ethics and regulations, and will not incur in the risk of limiting their competitive business advantages.

The cloud based nature of AEGIS infrastructure will made possible the inclusion of new data and knowledge, as well as new sector data and services. Also in this case, the alignment to AEGIS DPF will be requested for enabling producers and consumers to collaborate for the shared value generation and expansion of the overall solution.

AEGIS Data Policy Framework will be driven by the AEGIS Ethical, Privacy, Data Protection and IPR Strategy and resulting requirement list, as provided in this deliverable on the basis of the EU and national legislations, as well as ethical standards. The DPF is going to be elaborated in T2.2 “Data Policy and Business Brokerage Frameworks”. The DPF and the core methods to

be used will be reported in D2.1 “Semantic Representations and Data Policy and Business Mediator Conventions” (M8) and then refined in D2.3 “Update on Semantic Representation and Data handling and Analytics Methods” (M18).

Being the DPF strictly interrelated with the practical implementation of the requirements and recommendations set forth in the AEGIS Ethical, Privacy, Data Protection and IPR Stragegy, in this chapter we intend to briefly underline its key role and providing some guidelines and insights for its development.

The DPF will exploit the new opportunities arisen in the areas of security and privacy through the use of Blockchain technology. In fact, it will represent a novel method of using this technology and micro-services for checking data quality, security, trust and IPRs. As stated in the DoA, “a blockchain can be defined as a digital, chronologically updated, distributed and cryptographically sealed record, of all data transfer activity”, which “enables the transfer of digital assets, representing various manifestations of value or possessing inherent value within themselves. It is a secure way to enable such transactions and various other digital activities as everyone can participate, there is no identity disclosure and as already mentioned, manipulation is difficult due to the distributed nature of blockchain”.

In relation to AEGIS DPF, two of Blockchain’s aspects are particularly relevant: the employment of self-governance of transfer of ownership and the use of cryptography for preserving purposes.

In line with the Blockchain model, specific categories and predefined lists will be selected in the DPF to describe data IPR, security, trust and quality features, identifying how these tags can be applied on both the whole dataset as meta-information but also in their ingredients (e.g. data tuples). In case of classification of personal and sensitive data, extra tags will be used to allow the proper execution of anonymisation techniques by the other AEGIS components, not making it able to link back data to individuals.

The DPF will be used upon insertion of any kind of data into the platform, performing the necessary annotations both at dataset and on dataset element level, ensuring that the accumulated data are fully describe with respect to their IPRs, quality and privacy levels.

The semi-automatic negotiation of micro-contract regarding data exchange based on existing IPR schemes will be made possible by the Business Brokerage framework, by exploiting the DPF core methods, notably the IPR annotations, as well as by utilising Blockchain technology. In particular, following the annotations, all data exchanges are going to be supervised by the Business Brokerage service, which will generate on-the fly micro contracts for data sharing between the different data collaborating parties managing IPRs, data quality and data privacy issues (and thus ensuring that there is no infringement in privacy, no misuse of IPRs and that data quality is guaranteed).

As regards security, privacy, data quality and trust, AEGIS DPF is expected to introduce and research a real-case security usage of Blockchains for ensuring the creation of dynamic secure, isolated privacy-respecting communication channels supporting the communication and data exchange between different stakeholders and users of the platform in a distributed environment without the need for intermediaries. Security incidents and risks that may occur to AEGIS applications and databases will be avoided by matching the security level of any counterpart relational databases without compromising on its operational features. A set of current state-of-the-art security and privacy technologies are currently under evaluation in alignment with the use of Blockchains (e.g. i) anonymisation technologies, such as such CloudTeams Anonymisation Tool, ii) identity and access management applications, standards and methodologies like Directory services, Digital Cards, Service Providers, Identity Providers, Web Services, Access control, Digital Identities, Password Managers, Single Sign-on, Security Tokens, Security Token Services, Workflows, OpenID, WS-Security, WS-Trust, SAML 2.0, OAuth and RBAC; iii) encryption and key management technologies and algorithms, like Triple DES, RSA, Twofish, and AES; iv) Network security and threat management mechanisms, like Firewall; v) vulnerability scanning and Penetration testing, like Metasploit and Samurai Web Testing framework tools).

The AEGIS platform will manage and process closed data (proprietary data) as well as open data and will allow the exchange of data with different IPR. The former will be stored encrypted in the Security Linked Open Data (SLOD) space, whilst the latter will be will stored unencrypted and published under an open data license and will be publicly accessible through the platform. IPR and data sharing agreements through semi-automatic negotiation of micro-contract utilising Blockchain technology will be based on the predefined data handling policies, schemes and annotations defined in the DPF. They will ensure IPRs on data artefacts and data usage in relation to the data to be contributed to the platform.

In this perspective, AEGIS is expected to use a Blockchain-based IP Model, which makes possible unprecedented forms of transparency in copyright information and management. In fact, Blockchain Technology is expected to allow all users to have clear insights and access to all copyright information on the dataset and on any dataset element and, at the same time, to be able to bring an easier payout system to IP owners and licensors of data.

A good example to consider is Ascribe “Ownership Layer”, which provides a powerful tool allowing proof-of existence on Blockchain for IP and innovation (in AEGIS, for daset or data element) and makes easier the whole process of licensing and copyright transfer. Ascribe tackles the compelling need for a workable solution to the ownership and attribution issues by ensuring “ownership processing”, that makes ownership actions of digital property universally accessible. Ascribe’s approach is twofold, being based both on a registry with easy and secure legals and on visibility of data on usage / provenance of the content. It has two components, respectively ensuring IPR transparency and management, based on an ownership registry for easy secure

disposition of rights. There is an ownership registry with easy and secure legals, which formalise (via a creator and consumer-friendly Terms Of Service) existing copyright rights on digital objects traditionally difficult to be leveraged, whilst the bitcoin-inspired blockchain serves for securely recording ownership transactions. In the registry it is possible to register a work, transfer ownership, grant licenses, loans and rentals. The registry also provides the time-stamping evidence of ownership actions through bitcoin-inspired blockchain. Ascribe enables to record intellectual properties on the Bitcoin-inspired blockchain, which is used as a distributed database to store the registry records (that track the history of ownership, the so-called "provenance"). It, thanks to the combination with cryptography, is able to make the registry global, robust, and impairment-resistant, whilst shielding the parties' personal identity (thanks to cryptography again). Ascribe has been proven in several domains and is being used both by individual creators and by institutions (e.g. marketplaces, libraries, archives, museums, galleries) and organisations, including new startups.

The following part of this paragraph will outline the main findings and insights relevant for the definition of AEGIS Data Policy Framework in relation to each of the demonstrators.

Demonstrator 1: Road Safety Indicator

In the automotive demonstrator VIF will engage a number of volunteers to generate **vehicle driving data** and/or **vehicle simulation data** during experiments after having signed an informed consent. Both sources of data are anonymised and do not include any personal information. VIF will provide the collected & anonymised automotive as well as the collected and anonymised simulation data to be published at the AEGIS platform for further data analysis and service generation. Both are sources of non-confidential data and can be shared with the AEGIS consortium to be processed in the AEGIS platform to establish data-driven services. However, other sources of data relevant for the automotive demonstrator and the three scenarios (e.g. weather data or map data) will not be generated by the people volunteering in experiments. They will have to be accessed through other data repositories (e.g. OpenWeatherMap for weather data of OpenStreetMap for map data) which will have their own licenses and disclaimers for using data and services, which have to be studied in detail.

Both vehicle simulation and vehicle usage data is **time series data**. Vehicle simulation data is generated through the software running on the simulator, e.g. through CarMaker[8] - a professional software which has been developed for testing passenger cars and light-duty vehicles in real world test scenarios. Vehicle usage data is generated by the vehicle data logger, a device developed at VIF and used for research purposes. While the **data quality** for simulation data is expected to be very high, the data quality for vehicle usage data measured by sensors from the car as well as by sensors installed on the vehicle data logger may vary. There will certainly by missing or wrong sensor values included in the data which have to be eliminated before exploiting it to establish data-driven services and reports. For instance, GPS data may include

wrong values, if a vehicle is driving under a bridge or through a tunnel as a result of reflections. OBD2 data such as vehicle speed or vehicle rpm may include false values from time to time, too, which have to be corrected, e.g. by applying interpolation mechanisms.

Vehicles are a valuable source of data. The collection of vehicle data from the field – and especially the smart combination of this data with data from other sources – will facilitate the generation of many innovative digital products, third party services and business models. However, such digital services based on vehicle data can only be successful if a critical mass of vehicles shares driving data. Raising awareness in the society on what kind of data a vehicle generates, processes, stores, and potentially transmits to a third party is a challenging yet crucial task. The '***My Car My Data'***[9] campaign launched by Federation Internationale de l'Automobile (FIA) educates car drivers about the potentials and pitfalls of connectivity. The My Car My Data campaign believes that the driver should be the one deciding if vehicle data should be shared and with whom. Europeans should be entirely free to choose with what party they share their vehicle data in the future (eventually on a market that allows services providers to compete in offering the drivers the most added-value for shared data), unless mandated by law.

Regarding direct access to vehicle generated data, **car manufacturers** are in a comparably lucky position. However, they were not very successful in exploiting this market yet to establish a digital ecosystem. The potential to exploit car lifecycle data for purposes other than driving currently remains almost untapped by automotive OEMs. According to the EU research project AutoMat (AutoMat, 2016), the automotive industry has not yet been able to successfully establish an ecosystem for apps and services equivalent to that of smartphone manufacturers. The project mentions three reasons why OEMs are currently struggling: Brand-specific business approaches dominate, and as a consequence there is a lack of brand-independent car lifecycle data. Current proprietary car services focus on the individual customer, what leads to privacy concerns, and few ideas exist how anonymised car data can be used to establish other services. The implied or required collaboration between OEMs on car data and services is considered risky in terms of competition.

Two recent position papers from **VDA - the German association of the automotive industry** discuss the role of German-speaking car manufacturers towards establishing digital ecosystems based on vehicle data. The first position paper '***Data protection principles for connected vehicles***'[10] (VDA 2014) refers to the continuous transformation of vehicles towards 'connected vehicles' with a permanent uplink to the internet and the feasibility to connect various data sources for establishing new services. The position paper suggests three principles for VDA members to handle the advancements in connectivity and the new services associated with respect to responsible data handling as well as with data protection:

- Transparency: The members of the VDA strive for adequate information about the data in connected vehicles and the use of these data.

- Self-determination: The members of the VDA are striving to enable customers to determine themselves the processing and use of personal data through various options.
- Data-security: The members of the VDA strive to implement the strong safety culture in the automotive industry also in the connected vehicle.

The short paper closes with a chart of data categories in connected vehicles and their relevance for protection.

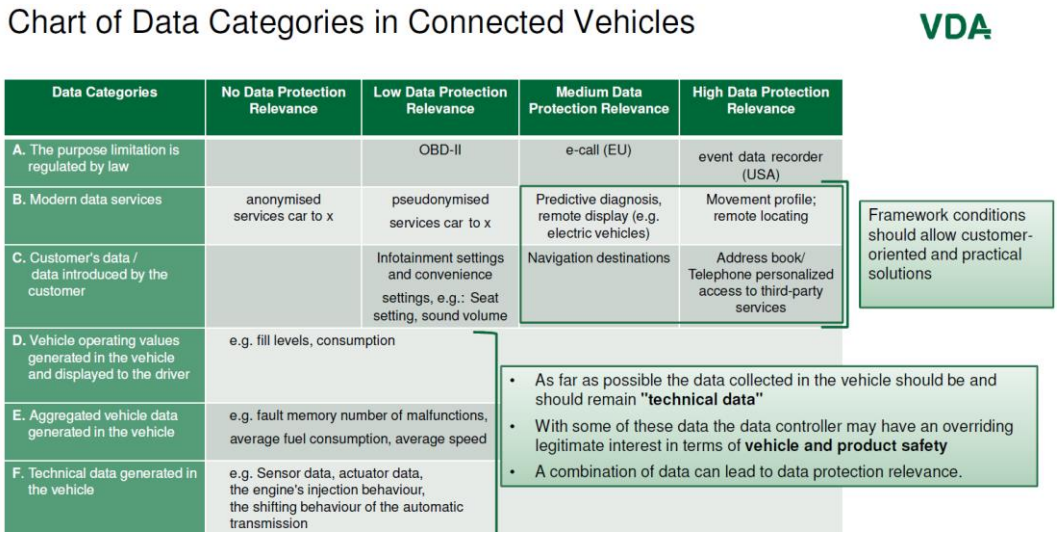

Figure 15: Data categories in connected vehicles (Source: VDA)

The second position paper titled '***Access to the vehicle and vehicle generated data***'[11] (VDA 2016) which has been developed in accordance with the '*EU Commission* ***C-ITS platform project final report***' [12] discusses data-centric requirements for security, privacy, and discrimination free innovation. The C-ITS report cited in the position paper lists five guiding principles to apply when granting access to in-vehicle data and resources:

1. 'Data provision conditions: consent': The data subject (vehicle owner) decides if data can be provided and to whom, including the concrete purpose of the data including an opt-out option.
2. 'Fair and undistorted competition': All service providers should be in an equal position to offer services to the data subject.
3. 'Data privacy & data protection': There is a need for the data subject to have vehicle and movement data protected for privacy reasons.
4. 'Tamper-proof access and liability': Services making use of in vehicle data and resources should not endanger the proper safe and secure functioning of the vehicle.
5. 'Data economy': Standardised access favours interoperability between different applications and facilitates the common use of same vehicle data.

According to the VDA report, each OEM holds the role of a system administrator and is hence responsible for the safe and secure transfer of car data to a business to business (B2B)

OEM interface. Third parties can access this car data directly over the OEM B2B interface or via neutral servers, which gather the data from the cars.

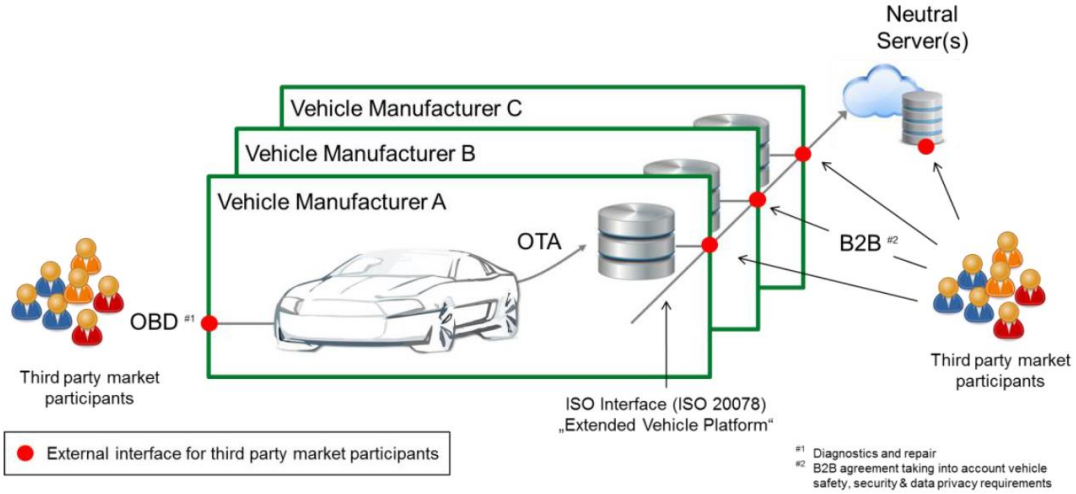

Figure 16: Access to the vehicle (Source: VDA)

The VDA report summarises that the vast majority of vehicle generated data is raw technical data, which is used locally within the vehicle and never stored. The VDA has defined four usage categories for vehicle generated data: Data for the improvement of road traffic safety, data for cross brand services, data for brand specific services, data for component analysis and product improvement, and personal data.

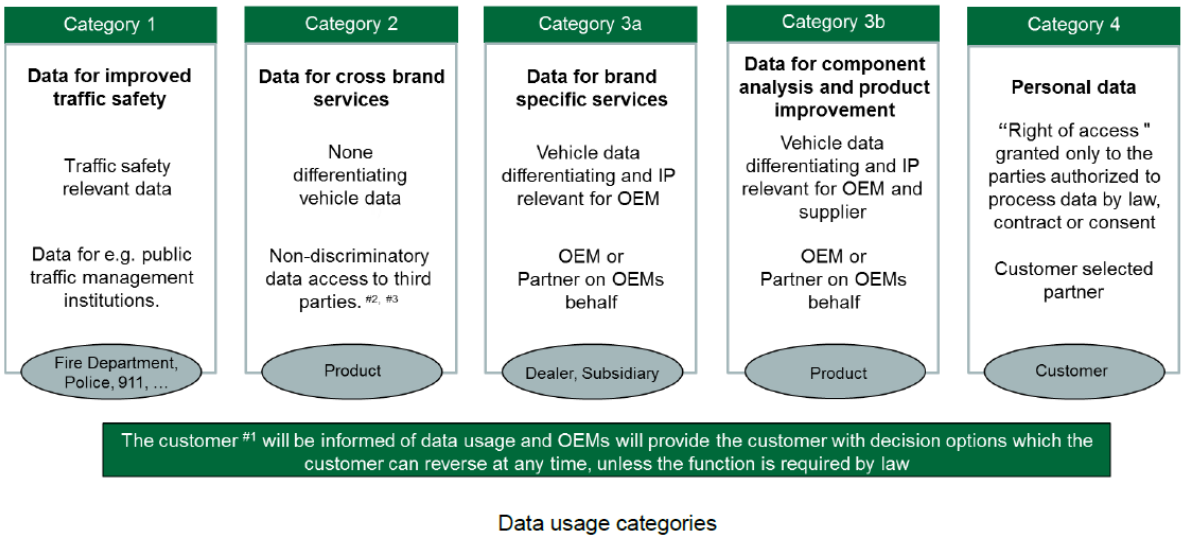

Figure 17: Data usage categories (Source: VDA)

Demonstrator 2: Smart Home and Assisted Living

The main objective of Smart Home and AAL demonstrator is to correlate information coming from ubiquitous IoT devices (building sensors, smartphone data and wearable devices) with open datasets towards the extraction of meaningful services. As Internet of Things (IoT) becomes a

growing reality, more ubiquitous devices are embedded in our daily lives, serving us in a broad range of purposes in everyday life from: personal healthcare to home automation to location based services.

IPR considerations

These devices primarily collect data that is about or produced by people, be it the energy footprint of an individual's home or her location and other situational context. As this unprecedented amount of data is collected, we are challenged with one fundamental research question: **who owns this data and who should have access to it?**

Specifically, the emergent of the Human Data Interaction (HDI) topic which aims to put the human at the centre of the data driven industry, calls attention to the IoT community to address the data ownership aspect more carefully. In this note, it is fundamental within the project to clearly clarify IPR issues on **data artefacts** and **data usage** for Smart Home and AAL demonstrator. The analysis should not only address the demonstrator specific requirements, rather to exploit the potential of commercial exploitation of the AEGIS platform towards providing home automation and AAL services.

Nowadays, the main IP rights in relation to data are copyright, database right and confidentiality. Patents and rights to inventions can apply to software and business processes that manipulate and process data, but generally not in relation to data itself. Trademarks can apply to data products, but again, generally not in relation to the actual data. IPR in relation to data is of uncertain scope at the moment, and the law in this area is likely to continue to develop in the coming years: historically, IPR development has followed the commercialising of innovation and as the value of Big Data rises, so likely will the IP rights underpinning it.

In essence, the owner of machine-generated data (MGD), which covers virtually all of the IoT, is the entity who holds title to the device that recorded the data. In other words, the entity that owns the IoT device also owns the data produced by that device. However, it's not always clear that whomever has possession of the device and/or its output data actually "owns" it. Data may be owned by one party and controlled by another. Possession of data does not necessarily equate to title. Possession is control. Title is ownership. Referred to as usage rights, each time data sets are copied, recopied and transmitted, control of the data follows it. Conversely, transfer of ownership requires a legal mechanism to convey title.

Contract rights in relation to data are technically entirely separate from IPR and their value was confirmed in a UK High Court case in 2006 where the judge said that an owner of data: *is entitled in principle to impose a charge for use of its data by users whether or not it has IP rights in respect of that data.*

By taking into account the high level principles towards managing IPR issues for data coming from IoT devices, we are defining as data controllers: the data scientists of the company providing the equipment to the end users. This is the common case in industry as the IoT solution provider is also the controller of the data streams (IPR on data usage).

Data streamed off devices in home has so far been handled by companies which treat the users as clients **only with no say on how their data should be used**. However, we believe that given a transparent framework and regulations, many users would be willing to share their data. As such, we propose the notion of Data Market as an instrument to enable users to share their personal data locally and globally with monetary benefits, i.e., an individual can trade data produced at her personal space with interested business entities. Such model is currently being considered by a number of data exchange companies, where monetary incentives are offered to end users for correcting erroneous sensor data. The challenge in this case is how to design future infrastructure so to make users aware of the commodity of their data along with the risks of sharing it.

Data Quality Considerations

Extracting high-quality and real data from the massive, variable, and complicated data sets becomes an urgent issue. Data quality is not necessarily data that is devoid of errors. Incorrect data is only one part of the data quality equation. Amongst others, there are several conditions that contributed to the data quality problem such as lack of validation routines, data valid but not correct, mismatched syntax, formats, and structures, unexpected changes in source system, lack of referential integrity checks, poor system design and data conversion errors.

High-quality data are the precondition for analysing and using big data and for guaranteeing the value of the data. Figure 2 shown the first five attributes (i.e. Accuracy, Integrity, Consistency, Completeness and Validity) generally pertain to the content and structure of data, and cover a multitude of sins that we most commonly associate with poor quality data: data entry errors, misapplied business rules, duplicate records, and missing or incorrect data values. But defect-free data is worthless if knowledge workers cannot understand or access the data in a timely manner. The last two attributes (Timeliness and Accessibility) above address usability and usefulness, and they are best evaluated by interviewing and surveying business users of the data.

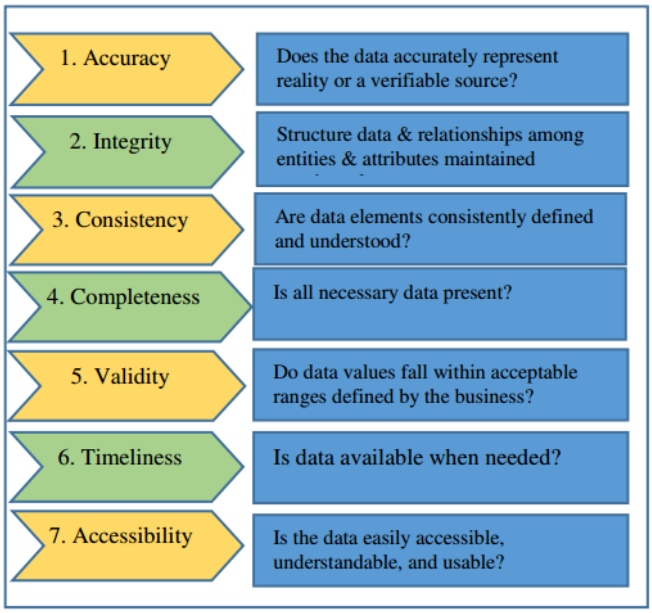

Figure 18: Data Quality Attributes

Data quality is an essential characteristic that determines the reliability of data for making decisions in Smart Home and AAL demonstrator. High data quality is:

- Complete: All relevant data such as in home environment data, wearable devices datasets and open datasets are available.
- Accurate: Common data problems like misspellings, typos, and random abbreviations have been cleaned up.
- Available: Required data are accessible on demand; users do not need to search manually for the information.
- Timely: Up-to-date information is readily available to support decisions.

It is very important to define the associated microservices in AEGIS platform that will meet the demonstrator requirements towards accessing high quality data. This is actually a process highlighted in the definition of the high-level usage scenarios for the associated demonstrator.

Considerations on Data Privacy, security and trust

Considering the nature and the type of the datasets available in Smart Home and AAL demonstrator it is mandatory to adopt Data Privacy, Security and trust policies towards handling the data streams required for PSPS services. Data privacy is suitably defined as the **appropriate use of data**. Privacy assures that personal information are collected, processed (used), protected and destroyed legally and fairly. On the other hand, data security provides protection for all types' information, in any form, so that the information's confidentiality, integrity, and availability are maintained.

We have highlighted the importance of data privacy and security in the description of the high-level usage scenarios for the demonstrator. Visiting the AEGIS platform, the data scientist (of the IoT equipment provider) creates an organisation profile for the company and then creates a project marking it as a "private" project. In the project, the data scientist may invite other users providing classified access to the same project, giving them partial access to the data. At the same time, invitations may be delivered to external collaborators with specific rights to upload data to the project's repository, indicating also the data structures expected and the schemas that need to be uploaded. In line with classified access to the datasets, we are highlighting the importance of anonymisation over the streams of personal data, to meet privacy and security objectives. A list of tools should be provided by the AEGIS platform, to ensure that a prompt anonymisation will be performed over the streams of data by the data scientist.

Finally, some further considerations regarding the aspects that the "Data Policy Framework" will have to take into consideration as regards the Smart Home and Assisted Living Demonstrator.

Purpose of the processing of personal data

The purpose of processing the streams of personal data is defined through the specification of the high-level usage scenario reported in this deliverable. More specifically, the scope of Smart Home and AAL demonstrator is threefold.

- To provide **monitoring and alert Services for Social Care Service Providers**. AEGIS will develop data-driven HMIs to social care services providers (public and private) to enable accurate **monitoring of elderly activities** and **identification of critical incidents** that may require on-the-spot physical interventions and assistance. Towards this direction, data from **smart home** and **wearable devices/ smart phone sensors** (accelerometer, GPS) will be correlated with information coming from open data sources towards the identification of physical wellbeing deterioration and frailty status, signify cognitive deterioration etc..
- Add on feature is the establishment of a **home automation framework** to support the provision of AAL services. Analytics over the streams of in-building datasets will further enable the definition of accurate ambience preference profiles of elderly people in the indoor environment. The extraction of comfort/ discomfort profiles will further facilitate the deployment of human-centric, personalised smart automation strategies over their heating/ cooling and lighting devices, to ensure optimal comfort levels and compliance with special, ambience-related, health requirements.
- Finally, a version of the mobile App for the end users (elderly) will be evaluated as an outcome of the demonstrator. Once again, personal data (location based data as retrieved from smartphones and wearables) will be correlated with data available from open data sources towards providing **human-centric, personalised recommendations and (close**

to) real-time notification to elderly people how they can avoid **risks** imposed by weather conditions, accidents etc..

A more detailed list of processing mechanisms is defined by UBITECH as part of user requirements

Origin of personal data and its collection method

Smart Home and assisted living demonstrator evaluation requires the installation of equipment and usage of wearable devices. The list of datasets to be examined in this demonstrator are presented in the following table.

Occupancy (PIR)
Luminance
Indoor Air Quality
Indoor temperature and humidity
Control actions over lighting and HVAC
In-home Energy Footprint
Wearable Sensor Data
Smartphone Sensors (Accelerometer/ Gyro/GPS)
Personal Health Data (Dummy data)

Towards gathering the required information, a limited number of installations will be performed as part of the demonstrator. Namely we are considering the installation of:

- PIR sensors for tracking occupancy information
- Luminance, temperature humidity, VOC and CO2 sensors for acquiring information about illuminance, indoor temperature and humidity and IAQ levels.
- HVAC controllers and actuators (smart thermostats, actuator interfaces) towards acquiring information about HVAC operation
- Lighting devices controllers and actuators (dimmers, 0-10V actuators, smart lamps) towards acquiring information about lighting devices operation
- Smart metering equipment, clamps and plugs towards gathering information about energy consumption

In addition to in-building equipment installation, datasets will made available from:

- Smartphone devices with build-in sensors as a daemon is running to track accelerometer, gyro and GPS data
- Wearable devices (activity trackers) with build in sensors to track activity and health related parameters

Finally, self-reporting data about personal health conditions is an option considered in the project. As these are sensitive data, a data specific ethics handling methodology has been presented above.

Technological component of the overall AEGIS system processing personal data relevant in this demonstrator

Having defined above 1) the list of datasets (personal data) available in Smart Home and AAL demonstrator and 2) the purpose of processing these datasets, we are further highlighting the list of technological component of the AEGIS system to support the analytics process. The overall analysis takes into account the data value chain schema towards the definition of the associated technological component

- The end users should be able to **upload data to the project's repository**, indicating also the data structures and the schemas that needs to be uploaded. A service should support end users to easily upload datasets examined in Smart Home and Assisted Living demonstrator. Different ways of uploading datasets should be considered in the project as some data are initially uploaded only for experimentation purposes (e.g. some samples of datasets), while also there is the option to connect data to the platform through a project specific API endpoint.
- The platform should support **tools** to **anonymise, clean and transform data** in order to meet the expectations of the data scientist. **Anonymisation** and data privacy preservation methods are required so that no sensitive data is transferred to the platform. Furthermore, **cleansing and normalisation services** should be supported to ensure high quality datasets availability. Finally, **semantic curation** of the datasets should be supported considering the nature of the applications developed in the project (PSPS services)
- Having all the data in one place, the data scientist is now able to invoke several analyses, choosing which data to combine as well as the algorithms to utilise. Those come out of a predefined **algorithms library,** while it is also possible for the end users to **upload their own algorithm** and conduct an analysis. The overall results (datasets and analytics results) are then presented in a dashboard that visualised the outputs of the analyses,

where access can be provided to any member of the project, while the results can be also exported in various formats.

- What is especially interesting is the option to **export the data through an API** that also allows the analyses to be executed remotely by providing an external signal. This can come either from an external stimulus, such as a weekly call from an external system, or from triggers specified and enabled in the AEGIS platform, such as the updating of a dataset or the occurrence of an event.

The definition of the AEGIS technological components towards handling personal datasets is in line with the overall data value chain definition in previous section.

Risks identified when dealing with personal data

Having presented above the overall framework for Smart Home and AAL demonstrator and the usage of personal datasets, an indicative list of risks when dealing with personal data is presented in the following table:

No.	Description
1	Loss of Privacy Control. Participants will be monitored and their personal data will be collected.
2	Difficulty in ensuring the security of shared Personal Data
3	Storage and Process of personal Data - Confidentiality
4	Lack of transparency
5	Delegation of Control Privacy - Incidental Findings
6	Improper use of IT equipment

Demonstrator 3: Insurance Sector. Personalised Early Warning System for Asset Protection

In the Insurance sector demonstrator, through AEGIS technologies and the collection, knowledge of customer data and related real-time risk analysis, will lead to a more personalised mode of calculation of the risk associated with each customer, the provision of an alert system and new insurance models.

In this demonstrator, as in the others, volunteers will be involved and their personal data will be collected and handled.

A unique interface will enable to access and analyse information coming from diverse and heterogeneous data sources (geospatial information, photos, videos, social media, broadcasted news, etc.) and will be combined with the in-house big data platform of the insurance company. The information collected includes the passive collection of the data from the car itself (speed, driving style, fuel consumption, preferred roads, times and places of use, etc.).

In this way, we still deliver personalised intelligence for preventing specific catastrophic incidents related to people lives/assets through early warning notifications. These notifications take the cue from the incidents and threat situations (e.g. severe weather conditions) and will be diffused by the insurance company, together with different recommendations, based on the anticipated impact that the given incidents and threat situations might have, on the basis of an expert's model that store and evaluate their criticality and severity.

The proposed Risk impact analysis e Personalised Early Warning System for Asset Protection Solution will apply semantic analysis to the gathered data and will realise the intelligence extraction from multiple data sources. It will investigate multiple types of threats and refer to the location/asset type, besides capitalising on the already available open data knowledge.

The warning notification will target a restricted group of customers by filtering out the ones belonging to groups that are really in danger (due to location, etc.).

This personalised proactive alert system will allow to risk reduction (of insurance companies' clients) by assisting people to avoid threats and proactively take the necessary precautions, thus resulting in savings for both the insurance companies and the individuals.

The last proposed solution is called Personalised commercial offering. Its goal is to collect data about customers' habits and customs from publicly available information and to use them to create personalised offers. The data will be retrieved from smart home, wearable devices, smart phones and social media, as well as personal health data records reported by the citizens, social media and web trends. All this information will be processed and matched with the company products, in order to create a unique profile that really fits the customer needs. With this profile, the company can choose the most suitable products for this specific customer and propose a personalised offer, preventing the customer from paying for things that are useless to him. This approach can increase the customer's satisfaction also by proposing rewards when reaching established safety goals.

Privacy considerations

In addition to the specific laws on protection and treatment of personal data, HDI's Ethic Code establishes that recipients who, in the exercise of their activities, acquire documents, studies, work plans (including business plans), technological processes, data and Information of any kind related, directly or indirectly, to the activities of HDI, have the obligation to guard and protect them in an appropriate and continuous manner in compliance with the security measures adopted by the Company pursuant to Italian D. Lgs. 196/2003 ("Codice Privacy"). In particular, personal information collected must be processed in accordance with the principles established in the "Codice Privacy" in a consistent and appropriate way to the purpose of their collection.

It is in any case compulsory to refrain from seeking confidential information, which is not functional in the exercise of their functions.

Data Protection considerations

Appropriate security measures are taken against unauthorised access to, or alteration, disclosure or destruction of the data and against their accidental loss or destruction:

- Particular focus is placed on the security of personal data held on portable devices, with appropriate security measures such as encryption applied.
- Robust procedures for limiting access to personal data are in place and that staff are aware of these limits.
- An appropriate external access policy is in place to ensure that only the data subject or their clearly chosen representative has access to their personal data during the course of a policy or claim.
- A confidentiality policy is in place pertaining to the collection, processing, keeping and use of medical and sensitive data.

Access to sensitive data is restricted to authorised staff. In particular it is expected that access to sensitive medical information should be restricted to relevant underwriters, claims assessors and persons needing to access a particular file as part of their role.

Origin of personal data and its collection method

The Insurance demonstrator gathers data both from internal enterprise and both from external sources.

The list of dataset to be examined in the demonstrator are presented in the following table:

Inhouse dataset from Company CRM system
Inhouse dataset from Company Portfolio system
Inhouse dataset from Company Claims system
Third parties customer data (e.g. social activity data, risk indicator)
Institutional Italian databases
Private databases from third parties

National and local press releases
Reports and statistics on Insurance facts
Wearable sensor data
GPS data
Car black-boxes stream data
Weather statistics
Social media (e.g. Twitter, Facebook)
Natural Disaster datasets
Trending topics
Floods
Earthquakes
Health data
National and local press releases

Concerning internal enterprise datasources, data will be adequately anonymised before being processed.

The Insurance demonstrator may require the installation of wearable devices, GPS devices and / or car black boxes. To deal with this data HDI will involve a number of volunteers to generate data during the experimentation; Informed Consent Procedure for gathering the volunteers' consent to the transmission and processing of their data will be followed. Moreover, the simulation of this kind of data is an option considered in the project.

Concerning external datasources, they have their own licences with will be taken into consideration during the experimentation phase.

Purpose of the processing of personal data

The purpose of processing the personal data is defined through the specification of the high-level usage scenario reported in this deliverable. More specifically, the scope of Insurance demonstrator is **to provide customers with efficient added value services**.

- By correlating risk information with internal enterprise datasets, the demonstrator is able to identify assets potentially involved in the risk. This information can be used in order to directly contact customers and to assist them in the event of damages to customer's assets insured with the Company.
- By correlating risk information with real time IoT devices data, the demonstrator is able to identify if a customer may be involved in the risk. This information can be used to provide the customer with a fast and efficient assistance. For example, in the case the customer is found to be potentially involved in an hailstorm, our operators can contact him and suggest him the nearest bodyshop in the area that can assist him.

An advanced customer segmentation enables the Company to provide its customers with personalised offerings that better fit his behaviour and needs.

5. CONCLUSIONS

The deliverable builds on top of the work performed in other tasks and deliverables in order to refine the concept of the project. More specifically, in order to fulfil the needs of all stakeholders (reported in D1.1) and user roles (outlined here and to be formally defined in WP3 to capture the different knowledge background and primary interest in using AEGIS), AEGIS will develop a central platform and a set of additional tools to facilitate (a) big data interlinking under a common PSPS context through commonly agreed upon semantics, (b) big data discoverability and consumption through intuitive and configurable services, (c) big data and intelligence sharing through clearly defined interactions and data privacy and security preserving mechanisms, (d) advanced sensory data manipulation, (e) easy application of big data analysis and visualisation and, ultimately, (f) quicker and easier roll-out of data-enabled applications related to the PSPS domains.

In order to both outline and reveal the reasons why users would consider adopting the project's offerings, the needs they would like to cover and the steps they would expect to go through in order to achieve this, five high level usage scenarios of AEGIS were described in detail. Further analysis of the high-level scenarios was performed and implied features and functionalities were extracted and grouped in order to provide an initial definition of the AEGIS MVP. This work concluded with five main processes, corresponding to different core components that AEGIS should support:

- A flexible data import engine to support all commonly used data formats in PSPS applications (e.g. sensory data).
- A service marketplace for data, analytics and visualisation sharing and acquisition. AEGIS should also implement and provide several configurable data-as-a-service services to cover common stakeholder needs (e.g. for crime, news and weather data) enabling their easy consumption and ensuring they are compatible to be integrated in all analysis processes running on the platform.
- A powerful big data analytics engine, leveraging semantics and linked data in the background to deliver the required intelligence capabilities customised to the PSPS domain.
- A set of intuitive visualisation options, configurable to an extent and easy to combine in user created dashboards.
- Export mechanisms for the visualisations and analysis results for easier consumption and sharing with others. In order to achieve this, AEGIS should provide various export options, indicatively including creation of links to share/embed reports/visualisations and publishing analysis results through RESTful APIs.

Furthermore, the current deliverable defines and presents in detail the first version of the AEGIS integrated methodology towards data-enabled PSPS innovation. Towards this goal, user interactions are modelled and a rich set of workflows to be supported by the AEGIS system is described, in order to enable AEGIS to design, implement, refine and deliver an ecosystem for big data analysis and sharing, aiming to roll-out improved intelligence in the PSPS domains.

Finally, this deliverable defines AEGIS Ethical, Privacy, Data Protection and IPR Strategy, where, besides an in-depth analysis of the provisions of the current European and national

regulatory instruments relevant to AEGIS implementation and overall architecture, key aspects are described for both the project implementation phase and the AEGIS solutions, notably:

- Interrelations between T1.4 and D1.2 with other tasks and deliverables;
- role and activities of the Ethics Advisory Board;
- initial ethics and data protection insights for each of the demonstrators;
- Ethics Procedures, Roadmap and first findings in view of the elaboration of the Data Protection Impact Assessment Methodology;
- Key principles, legal evaluation and assessment of technologies in AEGIS
- Methodology for the elicitation and analysis of Ethical, Privacy, Data Protection and IPR Requirements and list of them;
- Guiding principles and recommendations for AEGIS Data Policy Framework, both at project-level and at demonstrator-level.

This Strategy only reflects an initial standpoint, which may be updated as the AEGIS architectural design develops. Consequently, as AEGIS solutions and demonstrators are not yet shaped in their final fashion, the requirements list has been conceived at a higher level of abstraction to cover various possible future technological choices

The Integrated AEGIS methodology and the MVP definition will be further refined and their updated versions will be presented in D1.3 entitled “Final AEGIS Methodology”.

APPENDIX A: LITERATURE

[1] Warren Samuel and Louis Brandeis, “The Right to Privacy”, Harvard Law Review, Vol. 4, No. 5, 1980.

[2] Westin Alan F., Privacy and Freedom, 1967.

[3] Flaherty David, Protecting Privacy in Surveillance Societies. The Federal Republic of Germany, Sweden, France, Canada & the United States, 1989.

[4] European Commission (2009): Data protection and privacy ethical guidelines, http://ec.europa.eu/research/participants/data/ref/fp7/89827/privacy_en.pdf

[5] Datenschutzbehörde der Republik Österreich, https://www.dsb.gv.at/

[6] Cavoukian, Ann: “Privacy by Design ... Take the Challenge”, 2009

[7] Opinion n.° 28, “Ethics of Security and Surveillance Technologies”, 2014.

[8] CarMaker: ipg-automotive.com/products-services/simulation-software/carmaker/

[9] MyCar MyData: www.mycarmydata.eu

[10] VDA - Data Protection Principles for Connected Vehicles: www.vda.de/en/topics/innovation-and-technology/network/data-protection-principles-for-connected-vehicles.html

[11] VDA - Access to the vehicle (and vehicle generated data: www.vda.de/en/topics/innovation- and-technology/network/access-to-the-vehicle.html

[12] Cooperative Intelligent Transport Systems (C-ITS)-Platform. Final Report: ec.europa.eu/transport/sites/transport/files/themes/its/doc/c-its-platform-final-report-january-2016.pdf

APPENDIX B: NON-DISCLOSURE AGREEMENT TEMPLATE

Non-Disclosure Agreement

between

Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e. V., Hansastrasse 27c, 80686 Munich, Federal Republic of Germany

- hereinafter referred to as »Fraunhofer«-

as legal entity for its

Fraunhofer Institute for Open Communication Systems FOKUS,

Kaiserin-Augusta-Allee 31, 10589 Berlin, Federal Republic of Germany

- hereinafter referred to as »Fraunhofer FOKUS« -

and

GFT Italia SRL, Via Campanini Alfredo, 20124 Milano, Italy

Kungliga Tekniska Hoegskolan, Brinellvagen 8, 100 44 Stockholm, Sweden

UBITECH Ltd., 26 Nikou & Despinas Pattchi, 3071 Limassol, Cyprus

Kompetenzzentrum – Das virtuelle Fahrzeug, Forschungsgesellschaft mbH, Inffeldgasse 21 A, 8010 Graz, Austria

National Technical University of Athens – NTUA, Heroon Polytechniou 9, Zographou Campus, 15780 Athina, Greece

Ecole Polytechnique Federale de Lausanne, Batiment CE 3316 Station 1, 1015 Lausanne, Switzerland, c/o College of Management of Technology (hereinafter referred as to "CDM"), represented by Prof. Andreas Mortensen, Vice-Provost for Research, and Prof. Christopher Tucci, Head of CDM

Suite5 Ltd., Wenlock Road 20-22, N1 7GU London, United Kingdom

Hypertech (Chaipertek) Anonymos Viomichaniki Emporiki Etaireia Pliroforkis Kai Neon Technologion, 32 Perikleous Street, 15232, Chalandri Athina, Greece

HDI Assicurazioni S.p.A., Via Abruzzi 10, 00187 Rome, Italy

- hereinafter together referred to as »Project Partners« -

and

…….

……..
[name, address]

- hereinafter referred to as »Board Member « -

Preamble

The Project Partners co-operate in the project "Advanced Big Data Value Chain for Public Safety and Personal Security" (hereinafter: Project) which is partly funded by the European Commission in the H2020 Research Framework Programme under Grant No. 1290/2013, and which is coordinated by Fraunhofer. The Project Partners have agreed that certain aspects of their work under the Project shall be assessed by experts of an Ethics Advisory Board so to benefit from the Board Members' expertise in order to optimize the Project results. To protect the results and information exchanged between the Project Partners and the Board Member, the following agreement is concluded:

§ 1 Purpose of the Board

The Project Partners will establish the Ethics Advisory Board that will include relevant external, independent expertise to evaluate the Project's progress and the results generated thereunder and to advise the Project Partners how to proceed with the Project ethically correct. Additionally, the Ethics Advisory Board shall supervise the conduct of the Project to ensure that European regulations regarding data protection are fully observed.

The Board Member will be invited to the Project meetings in order to learn about the Project's objectives and approach but shall not have any voting rights.

The Board Member shall contribute to the Ethics Advisory Board's Report that summarizes the evaluation activities of the Ethics Advisory Board and contains the Ethics Advisory Board's recommendations. The report shall be submitted as AEGIS Deliverable 9.3 as attachment to the AEGIS Periodical Reporting in Project Month 18 At the end of the Project, the Ethics Advisory Board shall update its report. The updated report as attachment to the AEGIS Periodical Reporting shall be submitted in Project Month 30.

The Board Member commits itself to actively contribute to the Ethics Advisory Board activities.

The Board Member will get access to deliverables and Results generated by the Project Partners as well as to intended publications from the Project whose drafts need to be treated in confidence.

§ 2 Confidentiality

For the purposes of this Agreement »Confidential lnformation« shall mean

- any technical and/or commercial Information, including – but not limited to – any documents. drawings, sketches or designs, materials or samples disclosed by any of the Project Partners or their subcontractors to the Board Member;
- information obtained from another member of the Ethics Advisory Board;
- deliverables/results generated by the Project Partners or their subcontractors.

The Board Member agrees to treat as confidential all and any Confidential Information – whether obtained directly or indirectly – and to use the same only for the purpose of the execution of its duties as a Board member and not to use or exploit such Confidential Information for its own or any third party purpose, disclose it to any third party or allow any third party access to such Confidential Information, except with the prior written consent of the disclosing Project Partner.

The Board Member will take all necessary precautions to ensure the confidentiality of the Confidential Information.

The restrictions on the use and disclosure of Confidential Information shall not apply to information which is:

(a) proven to have been known to the Board Member prior to the time of its disclosure pursuant to this Agreement; or

(b) in the public domain at the time of disclosure to the of Board Member or thereafter enters the public domain without breach of the terms of this Agreement; or

(c) lawfully acquired by the Board Member from an independent source having a bona fide right to disclose the same; or

(d) independently developed by the Board Member provided that it has not had access to any of the Confidential Information of the disclosing Project Partner.

§ 3 Liability

The Board Member shall be held liable for any damage caused to the Project Partners by breach of its duties under this Agreement.

The Parties agree that any Confidential Information is made available »as is« and that no warranties

are given or liabilities of any kind are assumed with respect to the quality of such Confidential Information, including, but not limited, to its fitness for the purpose, non-infringement of third party rights, accuracy, completeness or its correctness.

§ 4 Non-assignment

This Agreement is personally binding the Board Member and shall not be assigned by the latter without the Project Partners' prior written consent.

§ 5 Intellectual Property

The Board Member agrees not to exploit Confidential Information, in particular not to apply for the registration of intellectual property rights.

All Confidential Information supplied pursuant to this Agreement shall remain the property of the Project Partner disclosing or supplying the same and nothing contained in this Agreement shall be construed as granting to or conferring upon the Board Member any rights by license or otherwise, express or implied, to the Confidential Information, accompanying know how or any underlying intellectual property rights of the Project Partners.

Should any results generated by the Board Member – in particular from evaluating the Project results or deliverables – be eligible for protection under intellectual or industrial property laws or be protected under copyright law, the rights to use such results shall be assigned by the Board Member to the Project Partners. This shall especially apply to the Board Member's contributions to the documents and reports mentioned under § 2. Additionally, the Project Partners shall be entitled to use the contributions and recommendations generated by the Board Member unrestrictedly in time, place and content. For the avoidance of doubt, this right of use contains also the right to implement and develop such contributions and recommendations.

§ 6 Entry into force and Term

This Agreement shall come into force on the date of the last signature and shall thereafter be valid until September 2020, the current planned end date of the AEGIS reporting.The obligation of confidentiality hereunder shall continue to be valid for a period of 10 years after the end of the term of this Agreement.

Upon request of the disclosing Project Partner, any document, sample or material shall be returned by the Board Member to the disclosing Project Partner without delay and at the end of this Agreement at the latest.

§ 7 Miscellaneous

Amendments and additions to this Agreement must be made in writing to have legal effect.

This Agreement is subject to and governed by the laws of the Federal Republic of Germany.

If any provision of this Agreement is determined to be illegal or in conflict with the applicable law, the validity of the remaining provisions shall not be affected. The ineffective provision shall be replaced by an effective provision which is economically equivalent. The same shall apply in case of a gap.

Signed on behalf of the Project Partners, acting through the Coordinator

Place, date

Place, date

_____

Signature of the Board Member

APPENDIX C: EXPERT AGREEMENT – TEMPLATE

Agreement ("Ethics Advisory Board" Expert)

between

NAME of the expert,
Address

hereinafter referred to as "Party or Expert or Board Member"

and

GFT ITALIA SRL (GFT) SRL, 1549351, established in VIA SILE 18, MILANO 20139, Italy, VAT number IT00819200478

hereinafter referred to as "GFT"

WHEREAS,

I. In cooperation with other Beneficiaries, the Coordinator (Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e. V.) has been awarded a Grant Agreement by the European Commission (EC) number 732189 entitled "Advanced Big Data Value Chain for Public Safety and Personal Security, - AEGIS", hereinafter referred to as the Grant Agreement or GA. From this Grant Agreement including its Annexes certain rights and obligations result between the EC and the Coordinator. The Grant Agreement provides the participation of the Expert for certain part of the work;
II. As AEGIS may rise some concerns regarding ethical and privacy issues due to the use of users' personal data after their written consent, the AEGIS Consortium (Annex 1) has decided to put together an advisory board named Ethics Advisory Board (EAB), comprising of known domain experts and practitioners who will work closely with the overall Consortium during the course of the project on tackling ethical and data privacy issues that will have to do with the retrieval, the processing, and the retaining of these data. The EAB will provide independent opinions and thoughts and will advise both the technical and the research partners on issues regarding the AEGIS methodology, the development of the platform and its components and the piloting

operation. The Ethics Advisory Board will be coordinated by the EAB Coordinator, who will be responsible for interfacing with it;

III. GFT, as EAB Coordinator, is in charge for setting up and coordinating the EAB and of subcontracting for engaging the Experts;

IV. In performing the work as member of the Ethics Advisory Board it is anticipated that GFT, the Coordinator and the other partners of the "AEGIS" Consortium disclose to the Expert technical and/or commercial information of a confidential nature presently in their possession and wish to ensure that the same remain confidential. The Expert, the Coordinator and the other AEGIS Project Partners have previously signed a Non-Disclosure Agreement on 00.00.2017, which is still valid and binding.

Now, therefore, it is hereby agreed as follows:

2. The Expert will collaborate with the Coordinator, GFT and the other partners of the "AEGIS" Consortium in order to tackle any ethical issue raised by the project and monitor the legal issues to continuously assess and ensure that the framework being proposed adheres to a minimum set of ethical and legal requirements. The Board Member commits itself to actively contribute to the Ethics Advisory Board activities. In particular, the Expert will perform the work as follows upon demand of GFT:

    a. Provide his/her expertise in specific ethics and privacy areas (as instructed by the Consortium and the EC) during the whole duration of the project. The Expert will contribute to provide independent opinions and thoughts and to advise both the technical and the research partners on issues regarding the AEGIS methodology, the development of the platform and its components and the piloting operation, by providing expertise in specific ethics and privacy areas during the whole duration of the project. The Expert will contribute to propose the Assessment Methodology to be described in D9.1 and followed in WP1 and WP5, including, if opportune, the provision of Templates at an early stage and the coherence with the Ethical Risk Table already named in the AEGIS Annex I;
    b. Participate and/or contribute to AEGIS workshops or meetings, which will be conducted during the project;
    c. Co-create and/or review selected parts of the ethics and privacy related deliverables (e.g. Deliverable D1.2 - Aegis Methodology and High Level Usage Scenarios Aegis Methodology and High Level Usage Scenarios, Deliverable D6.3 - Data Management Handling Plan);
    d. Periodically report to the commission on the implementation of the ethical issues in project and compliance with applicable national and EU regulations. The Board Member shall contribute (in writing) to the Ethics Advisory Board's Report that summarizes the evaluation activities of the Ethics Advisory Board and contains the Ethics Advisory Board's recommendations. The reports will be

based on a common assessment methodology as introduced in D9.1of the AEGIS GA. The report shall be submitted as AEGIS Deliverable 9.3, as attachment to the AEGIS Periodical Reporting in Project Month 18. At the end of the Project, the Ethics Advisory Board shall update its report. The updated report as attachment to the AEGIS Periodical Reporting shall be submitted in Project Month 30.
The Experts are expected to collaborate collectively in the generation of these two reports. The length of each of thesse reports should be adequate and not shorter than 4 pages.

GFT will instruct the Expert in due time as to the dates of operation.

The Expert is responsible for ensuring that the research work meets scientific care, complies with accepted technical, scientific and professional standards, is undertaken by appropriate personnel and carried out in accordance with the the financial provisions laid down in Article 3.

3. The duration of the Grant Agreement is 30 months commencing on 01-01-2017 and terminating on 30-06-2019.The Expert shall commence to perform its part of the work on **00-00-0000** and shall have completed it on 31.05.2019 at the latest to enable GFT to consider the Expert's contributions in the final project report. At the latest by that date all results and reports shall have been delivered to GFT.

   The Expert shall notify GFT in writing without undue delay if it becomes apparent that it might be unable to keep the schedule.

3. The remuneration to be paid to the Expert under this agreement amounts to a lump sum of **5.000,00 Euro,** including VAT, if applicable, and costs for office supplies, communication, insurance, visa, travels, taxes, accommodation, subsistence, telecommunications and any other project related costs if not agreed otherwise herein, and is payable as follows:

   50% (contribution to WP1, WP5 and WP9) on 01.03.2018;
   50% upon completion of this Agreement and acceptance of all deliverables, reports and results.

   The Expert will be liable himself for paying taxes and making his own contributions to social security out of this sum.

Unless otherwise agreed in written form, the Expert shall personally bear travel and subsistence costs incurred by the Expert in connection with providing the services under this Agreement.

GFT shall make the payments to the bank account stated in Appendix 2 upon delivery and acceptance of the performed work and receipt of invoices from the Expert. Such invoices shall quote a reference to the Grant Agreement Number of the European Commission and shall provide a detailed description of the work/deliverables concerned.

4. The Board Member will get access to deliverables and Results generated by the Project Partners as well as to intended publications from the Project whose drafts need to be treated in confidence. The Expert is bound by the Confidentiality Obligation as well as the other obligations set forth in §§ 2, 3, 4 and 5 of the NDA.

5. Should any results generated by the Expert – in particular from evaluating the Project results or deliverables – be eligible for protection under intellectual or industrial property laws or be protected under copyright law, the rights to use such results shall be assigned by the Expert to GFT and the other partners in the AEGIS project. This shall especially apply to the Expert's contributions to the documents and reports mentioned under Section 1. Additionally, GFT and the other AEGIS partners shall be entitled to use the contributions and recommendations generated by the Expert unrestrictedly in time, place and content. For the avoidance of doubt, this right of use contains also the right to implement and develop such contributions and recommendations.

5. This Agreement shall come into force upon the date of its signature and shall thereafter be valid until complete fulfilment of all obligations undertaken by the Expert under this Agreement.

6. Any and all disputes that will arise in connection to this Agreement will be governed by the laws of Italy. Any disputes arising out of the present Agreement which cannot be solved amicably, shall be finally settled under the Rules of Arbitration of the International Chamber of Commerce by one or more arbitrators appointed in accordance with the said Rules. The place of arbitration shall be Milan, Italy if not otherwise agreed by the conflicting Parties. The award of the arbitration will be final and binding upon the Parties. Nothing in this Agreement shall limit the Parties' right to seek injunctive relief or to enforce an arbitration award in any applicable competent court of law.

7. If any provision of this agreement is determined to be illegal or in conflict with the applicable law, the validity of the remaining provisions shall not be affected. The

ineffective provision shall be replaced by an effective provision, which is economically equivalent. The same shall apply in case of a gap.

Signed for the Expert

Place, Date

Signature

Signed for and on behalf of GFT

Place, Date

Signature