HORIZON 2020

PROJECT DELIVERABLE REPORT

NAIADES

Greening the economy in line with the sustainable development goals

D1.8 RECOMMENDATIONS FOR DATA POLICY

A holistic water ecosystem for digitisation of urban water sector

SC5-11-2018

Digital solutions for water: linking the physical and digital world for water solutions

"This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820985"

Document Information

Grant Agreement Number	820985	Acronym	NAIADES	
Full Title	A holistic water ecosystem for digitization of urban water sector			
Topic	SC5-11-2018: Digital solutions for water: linking the physical and digital world for water solutions			
Funding scheme	IA - Innovation Action			
Start Date	1st JUNE 2019	**Duration**	36 months	
Project URL	www.naiades-project.eu			
EU Project Officer	Alexandre VACHER			
Project Coordinator	CENTER FOR RESEARCH AND TECHNOLOGY HELLAS - CERTH			
Deliverable	D1.8 – Recommendations for Data Policy			
Work Package	WP1 – Project Management, Quality Assurance and Reporting			
Date of Delivery	**Contractual**	M9	**Actual**	
Nature	ORDP: Open Research Data Pilot	**Dissemination Level**	PU-PUBLIC	
Lead Beneficiary	CERTH			
Responsible Author	Dionysis Bochtis	Email	d.bochtis@certh.gr	
		Phone		
Reviewer(s):				
Keywords				

Revision History

Version	Date	Responsible	Description/Remarks/Reason for changes
0.1	16/12/2019	CERTH	Table of Contents (G. Dolias)
0.14	13/02/2020	CERTH	Report write-up (G. Dolias)
0.15	18/02/2020	DISY	Comments received by DISY
0.15	25/02/2020	Konnekt-Able	Comments addressed by KT
0.16	25/02/2020	CERTH	Comments addressed by CERTH

Disclaimer: Any dissemination of results reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains.

Contents

Abbreviations

EU	European Union
IPR	Intellectual Property Rights
GDPR	General Data Protection Regulation
FAIR	Findable, Accessible, Interoperable, Reusable
IP	Intellectual Property
EC	European Commission
PID	Persistent Identifier
DOI	Digital Object Identifier
URL	Uniform Resource Locator
CC	Creative Commons
SA	Share-alike
ODC	Open Data Commons
NC	Non-commercial
ND	No derivative work
CDLA	Community Data License Agreement
DANS	Data Archiving and Networked services
MIT	Massachusetts Institute of Technology

1 Summary

This report is meant to provide all partners of the NAIADES project with the recommendations and guidelines on aspects, such as intellectual property rights, policies and other legal issues related to NAIADES data. In particular, this document defines the framework upon which data will be collected, stored, processed, shared and reused not only during the development process, but also after the project's completion.

2 Introduction

During the development process of the NAIADES project, large sets of data are expected to be produced. A large part of these data is highly valuable for the industry, government entities and other researchers, thus the creation of a framework for easily accessible data is a central principle of NAIADES project. However, although nowadays there is an increasing number of methods and techniques applied to data processing, the control of data and its access might not be conducted in a secure way. This might lead to data misuse, putting into risk the privacy of individuals.

This report aims to provide policy recommendations regarding the creation of a trusted, open environment for storing, sharing and reusing data produced during the lifetime of NAIADES project. It presents an analysis of the European Union's (EU) privacy and protection requirements and principles, as they apply to NAIDES project. Furthermore, it includes the standards of data processing in a safe and secure manner, an overview of the principles that should be followed towards interoperable and reusable data as well as the conditions concerning fair access to datasets generated within the lifecycle of the project.

More specifically, in Chapter 3, a description of how data creators of the NAIADES project can benefit from their gathered data – through Intellectual Property Rights (IPR) – is presented. Chapter 4 presents the various types of NAIADES data as well as the main aspects of General Data Protection Regulation (GDPR), regarding the protection of data used and produced during the project. Chapter 5 highlights how data will be processed and the people responsible and eligible for such operations as well. The FAIR Data Principles, which facilitate exchange of generated data, are thoroughly described in Chapter 6. Moving on in Chapter 7, the purpose of sharing NAIADES data is specified. The framework upon which data will be openly accessible is determined in Chapter 8 and the licences that can provide access to NAIADES data are presented in Chapter 9.

3 Intellectual Property Rights and legal issues

3.1 Ownership and IPR

Ownership determines the rights related to the data produced during the lifecycle of the NAIADES project. Intellectual Property Rights, such as copyrights, trademarks, patents, etc. allow the creator of the data to use and benefit from his own work. They are attached to the data and are defined in such a way, so that the data is protected. The person who creates the IP (Intellectual Property), owns the rights and provides access rights to everyone who is interested in making use of the data.

3.2 Copyright owner

The copyright owner is considered to be the person who generates the data in activities related to the NAIADES project. Generally, copyrights are granted to the person who creates the data, but they can also be vested in institutes or organisations. Moreover, copyrights expire after a fixed period of time.

3.3 Joint ownership

Joint ownership occurs when several partners of the NAIADES project participate in activities where data is produced and their respective contribution cannot be verified or separated. In order to define joint ownership, specific arrangements shall be made between the partners of the NAIADES project.

3.4 Transfer rights

IPR, related to data of the NAIADES project, can be transferred in two ways. One way is the transfer of user rights to IP under a license, which is valid for a specified period of time. That means that the right to use and exploit the IP is transferred, however the ownership of the IP is still at the partner's side. Users have access rights to the IP under specific conditions. The other option is the permanent transfer of the IP in a way that, even the ownership, is also assigned to the acquirer. Nevertheless, the European Commission (EC) may object to transfers to third parties established in non-associated third countries for ethical, competitiveness or security reasons.

3.5 Data modification

Data modification during the NAIADES project is determined by the IPR in order to protect original work. Any modification is only feasible if either this action is done by the owner, or if certain rights are given to people who are eligible to use data generated throughout the project. As an innovation action close to the market, NAIADES project utilises technologies that aim at developing marketable solutions. Six partners from the private sector comprise the project's consortium, in particular Konnekt-able, ADSYS, SIVECO, DISY, IBATECH, GUARDTIME. Economic sustainability of those partners depends on IPR related to their technologies and data generated by them. Therefore, the NAIADES consortium will protect that data and get approval of concerned partners prior to any data modification and utilization.

The use of original data must be in line with the NAIADES ethical guidelines and integrity must always be respected as well. This means that the right to modify data excludes the right to falsify or change original data. Any work that may derive from the NAIADES project and includes data modifications on copyrighted work is possible, if respective permission is provided.

4 NAIADES data profile

NAIADES data refers to all NAIADES variables produced from operations that fully comply with measurements recommendations, standard operating procedures and quality guidelines established within NAIADES. A list of types of NAIADES data is presented below:

- Quantitative numerical datasets
- Categorical data
- Qualitative data
- Variables from water distribution simulation (EPANET)
- Geographical data (GIS)
- Personal data
- Written and published material.

All partners of NAIADES project have to control all operations related to project's data, e.g. formats and procedures related to quality assurance, quality control, access, data policy, interoperability, licenses etc. In addition, consortium partners must upload and store data from project's operations in a data repository provided by Freedcamp management platform. This platform is chosen due to the fact that it provides security mechanisms that ensure safe storage of data for a long period, even after the completion of the project. Zenodo repository might also be used for sharing datasets outside the project, but this option will be decided later on by the consortium.

4.1 General Data Protection Regulation

During the lifespan of the NAIADES project, personal data derived from human participation (if collected) will be cured under the GDPR framework. Therefore, this type of data from volunteer participants will be protected, as the right to privacy and protection of personal data are regarded as fundamental rights, protected by instruments of the Council of Europe and the EU. European Parliament and the Council of the EU has enacted the General Data Protection Regulation[1] (GDPR) 2016/679, which is an EU law regarding the data protection and privacy for all its individual citizens.

GDPR determines the protection of individuals with regard to the processing of personal data and the free movement of such data. As it is a Regulation and not a Directive, it does not need to be implemented into national law, it is rather applicable in all EU Member States instead of 27 different laws. This regulation applies to data processing that is partially or completely automated, as well as to non-automated processing of data. The specified requirements shall be in compliance with this regulation not only during the NAIADES project, but also after its completion.

Rec 1 As defined by the EU law, there are two key players who process personal data: data controllers and data processors.

They are both legally responsible for complying with the respective obligations under data protection law. Therefore, only people who can be considered responsible under the applicable law, can take up these positions. The distinction between controllers and processors fulfils mainly the need to differentiate their responsibilities.

A data controller is defined as a "natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data"[2]. The role of the controller is to allocate responsibility, determine who shall be complied with data protection rules and how data subjects can call upon their rights in practice. In the private sector, a controller is considered a natural or legal person, whereas in the public sector, it is usually an authority. Groups or institutions without legal personality can be controllers as well, on condition that special legal provisions are provided. The appointment of the controller(s) that take part in the creation of applications and platforms in NAIADES project is important, on the grounds that many partners are involved in the development of the aforementioned entities.

Rec 2 Every time a partner, which participates in the development of such entities, is appointed as a controller, such partner should make sure that all requirements are complied with the European law.

Due to the fact that partners collaborate on several tasks, there is distinct possibility that multiple controllers (called joint controllers) may exist upon the same dataset. Joint controllers should clearly determine the purposes and means of processing according to the law. Under these complex data processing environments, joint controllers should clarify their responsibilities in order to reduce the risks of breaching protection rules. The controllers can participate under different ways in the determination of their joint responsibilities, which do not necessarily have to be equally divided. Thus, controllers play an essential role in the determination of data protection rules throughout the development process of the NAIADES project.

On the other hand, a data processor is defined as a "natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller"[3]. For instance, organisations can be processors, which role is to act as service providers and process data only, on behalf of another organisation (a controller).

[1] *https://gdpr-info.eu*

[2] *Article 4 (7) GDPR; Article 2 (d), Directive 95/46/EC.*

[3] *Article 4 (8) GDPR; Article 2 (e) Directive 95/46/EC*

Rec 3 A controller decides either the process of data by the processor, for instance through staff under his direct authority, or to delegate, all or part of the processing activities, to an external organisation.

GDPR specifies the territorial scope concerning the personal data processing "in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not"[4]. Consequently, the regulation is also applied when the processing takes place in a country which does not belong to the EU, on condition that the processing is in the context of the creator's activities in the Union. If such processing takes place in a third country, data will be transferred to this country by the creator or the processor established in the EU. In addition, the regulation is also applicable even if the creator or the processor is not established in the EU territory, but still offers services to data subjects in the EU.

The processor who is not established in the EU, but to whom EU data protection law applies, should appoint a representative in the EU. This representative should be established in one of the Member States where the data subjects are. The representative should be addressed by supervisory authorities and data subjects. The role of the representative should be under the legal responsibility of the processor who appointed him.

According to Article 2 of Directive 2009/136/EC, which amends the e-Privacy Directive 2002/22/EC, "location data" is defined as "any data processed in an electronic communications' network or by an electronic communications' service, indicating the geographic position of the terminal equipment of a user of a publicly available electronic communications service"[5]. Applications of NAIADES project are able to collect large quantities of data from a device (e.g. personal data stored on the device by the user and data from different sensors, including location). Since such services are utilised within NAIADES, location data will be gathered and processed during the development process of the project.

"Traffic data" is "any data processed for the purpose of the conveyance of a communication on an electronic communications' network or for the billing thereof"[6]. Traffic data might be related to data concerning the routing, duration, time or volume of a communication, as well as the protocol used. These types of data can also be connected with the location of the terminal equipment, the network on which the communication starts or ends, as well as the beginning, end or duration of a connection. Their format can be equivalent to the format of the communication that is transferred by the network.

Rec 4 Location and traffic data may only be processed when they are made anonymous, or with the consent of the volunteer participants, to the extent and for the duration, necessary for the provision of a value added service.

Rec 5 NAIADES users must be informed, prior to the acceptance of their consent, about the type of location and/or traffic data that will be processed, the purposes and the duration of the process, as well as whether the data will be transferred to a third party.

5 Personal data processing

Article 4 of GDPR defines "processing" as "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

[4] *Article 3 (1) GDPR*

[5] *Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC*

[6] *Article 2 Directive 2009/136/EC*

or destruction"[7]. In NAIADES project, the following activities are an example of processing: the collection of data, its storage and analysis, sharing with partners and others, dissemination, etc.

Under GDPR, the process of personal data from special categories may raise a higher risk to the data subjects; their misuse may have long-term consequences on the individual's fundamental rights, the right to privacy and non-discrimination to name a few. Therefore, the protection of personal data requires substantial attention.

GDPR acknowledges the concept of anonymous data. Anonymous data enable the change of a set of personal data in such a way, so that all identifying parts disappear and data subject is no longer identifiable. When personal data is transformed into anonymous data, data protection principles might not be applied to the processing of anonymous information. The decision about whether vital information is considered anonymous or not, and whether individuals are identified by their data, relies mostly on the circumstances. For this reason, each case should be analysed carefully; for instance, if there is statistical information presented as aggregated data and the original sample is not large enough, there is distinct possibility that a user can be identified.

The controllers determine the responsibilities of those who take part in the processing of personal data, thus they account for the confidentiality and security of such operations. Moreover, in order to ensure that the data protection law is applied effectively, liabilities for breaches of data protection law must be given to certain persons.

Rec 6 Controllers are always responsible for data protection breaches, according to the GDPR.

In NAIADES project, every data operation should have a dedicated controller, after an agreement amongst the partners. The processors must process personal data only under the specific guidelines of the controllers.

Rec 7 Controllers should appoint data processing only to processors that can provide evidence regarding their ability to carry out technical and organisational measures.

Rec 8 Processors cannot appoint data processing to another controller without getting the controllers' consent.

An agreement between the controllers and the processors should be made concerning the nature, the purpose and duration of the processing, the type of personal data and the categories of data subjects, the obligations and rights of the controllers.

Rec 9 The processors have to commit that they will act strictly under the controllers' instruction and they will take the appropriate security measures.

Rec 10 Each controller should maintain a record with all processing activities that are carried out under his/hers authorisation, and this should also be done by the processor. These records have to be made available to the supervisory authority upon request.

Data processing can be carried out also by a third party. A third party is considered to be a "natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data"[8]. The legal rights of a third party are generally different from that of the controller. Thus, disclosing data to a third party should be carried out under specific circumstances.

Apart from the third party, another term that is used is the recipient. It refers to a "natural or legal person, public authority, agency or another body, to whom/which the personal data are disclosed, whether a third

[7] *Article 4 (7) GDPR*

[8] *Article 4 (10) GDPR*

party or not"[9]. A recipient can be a person who works either in different company or authority, or in another division, within the same company or authority. The main difference between third parties and recipients is the authorisation status under which they access personal data, held by the controllers.

In general, processing of personal data must comply with the general data protection principles revised by the GDPR, such as fairness, lawfulness and transparency; purpose limitation; data minimisation; data accuracy; storage limitation; security, integrity and confidentiality. In the following paragraphs, an analysis of these principles is provided, in the context of NAIADES project.

5.1 Fair, lawful and transparent processing

In NAIADES, personal data shall be processed lawfully, fairly and in a transparent manner, with respect to the data subject. All partners must have legitimate rights for collecting and using personal data. They should make clear how they plan to use the data and guarantee that they will not misuse them in ways that may cause negative effects on individuals.

These principles are highly related with the term of consent. Data controllers must inform on time all people, whose data will be used during the NAIADES project, about whether the data will be combined with other data stored on a device, or collected from other sources. In addition, they should be informed about the results of such combinations of data, the purpose of further processing as well as to what third parties the data may be transferred to.

5.2 Purpose limitation

Purpose limitation is an important factor, as it is based on the purpose specification and the compatible use of personal data. It defines that personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a way that is incompatible with those purposes. Extended data processing for statistical purposes can take place, providing that it is compatible with the initial purposes and that data protection law is applied.

Rec 11 The acquired personal data will be transformed into anonymous data when it is expected to be used for commercial purposes. Moreover, personal data will be aggregated with other relevant data before any processing, making difficult in this way to identify any data subjects.

Data controllers must take into account the purposes that the personal data will be used for and collect personal data, for those purposes alone. Prior to the start of data processing, they must submit a documentation to the supervisory authority, concerning the purposes of such operations. This documentation has to be available by the supervisory authority upon request, as well as accessible by the person who provides the particular data.

5.3 Data minimisation

As described in the Data Management Plan of NAIADES project, data required for the purpose of the project have been clearly divided into certain categories. Throughout NAIADES project, personal data shall be adequate, relevant and limited to what is necessary, in relation to the purposes for which they are processed. Data controllers should narrow down every personal information that is collected, in order to achieve a clearly defined purpose. They should also maintain that information for as long as it is necessary to accomplish their purpose. The data minimisation process is essential for conducting research and designing platforms in a way that they will collect, store and communicate data at a minimum level.

5.4 Data accuracy

All personal data shall be accurate and, where necessary, kept up to date. Every reasonable step must be taken to ensure that inaccurate personal data are erased or rectified without any delay. Accuracy plays an

[9] *Article 4 (9) GDPR*

important role in data processing. Regular checking and updating of data accuracy are required in order to diminish the risks of incorrect calculations.

5.5 Storage limitation

Personal data shall be kept in a form that allows identification of data subjects for no longer than it is necessary. As a result, the risks of identifying certain people, who participate as volunteers, are minimised. No personal data will be centrally stored. Personal data will be scrambled where possible and abstracted, providing that it will not affect the final project's outcome.

As mentioned in a previous paragraph about data minimisation, during NAIADES, the duration of personal data storage should be no longer than the period for which they are required for conducting a specific purpose. However, the process of personal data may not be conducted according to the specified timetables. Thus, personal data may be retained for longer periods, under the data protection law in any case.

5.6 Data security, integrity and confidentiality

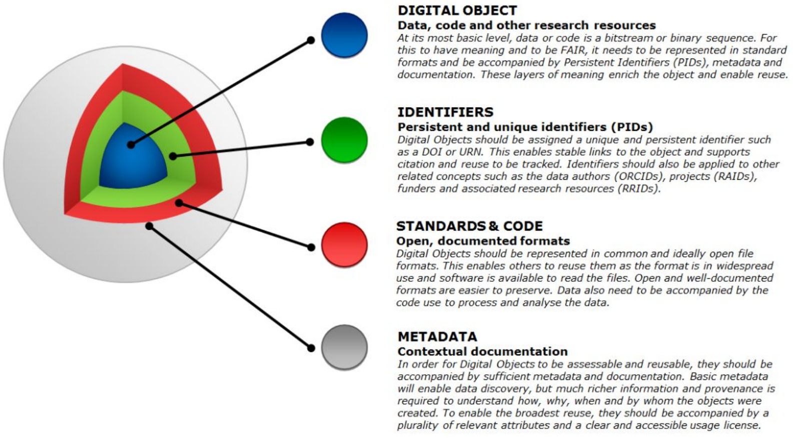
Any personal data collection and storage involving humans will be strictly confidential at any time of the project and the research. This means that all test subjects will be informed and given the opportunity to provide their consent to any data acquisition process. Personal data shall be processed in a manner that ensures appropriate security of personal data. Such actions include protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, by using appropriate technical or organisational measures.

In NAIADES, controllers and processors must implement such measures with respect to the rights of individuals. Some of these measures are listed here: the anonymisation of personal data; the ability to maintain confidentiality, integrity and availability of processing systems and services; the ability to restore and access personal data; the ability to conduct occasional testing, assessing and evaluating, regarding the security of the processing. Furthermore, both controller and processor have to make sure that any natural person with access to personal data, shall process them under their instructions. In any case, individual data on subjects will be used in strictly confidential terms and will only be published as statistics anonymously.

6 FAIR policies

6.1 The FAIR Digital Object

The EC Expert Group on FAIR data launched the concept of the FAIR Digital Object.



DIGITAL OBJECT
Data, code and other research resources
At its most basic level, data or code is a bitstream or binary sequence. For this to have meaning and to be FAIR, it needs to be represented in standard formats and be accompanied by Persistent Identifiers (PIDs), metadata and documentation. These layers of meaning enrich the object and enable reuse.

IDENTIFIERS
Persistent and unique identifiers (PIDs)
Digital Objects should be assigned a unique and persistent identifier such as a DOI or URN. This enables stable links to the object and supports citation and reuse to be tracked. Identifiers should also be applied to other related concepts such as the data authors (ORCIDs), projects (RAIDs), funders and associated research resources (RRIDs).

STANDARDS & CODE
Open, documented formats
Digital Objects should be represented in common and ideally open file formats. This enables others to reuse them as the format is in widespread use and software is available to read the files. Open and well-documented formats are easier to preserve. Data also need to be accompanied by the code use to process and analyse the data.

METADATA
Contextual documentation
In order for Digital Objects to be assessable and reusable, they should be accompanied by sufficient metadata and documentation. Basic metadata will enable data discovery, but much richer information and provenance is required to understand how, why, when and by whom the objects were created. To enable the broadest reuse, they should be accompanied by a plurality of relevant attributes and a clear and accessible usage license.

Figure 1: *The FAIR digital object*[10]

The goal of FAIR Data Principles (Findable, Accessible, Interoperable, Reusable) is to enable the creation, sharing and re-usability of quality, valuable and responsible data. By applying these principles, data can live up to its long-term potential.

Rec 12 Shareable data objects within the NAIADES ecosystem should be considered as FAIR Digital Objects.

6.2 Recommendations for Findability

Findability denotes that the existence of a data object as well as the content related to it (e.g. metadata), must be discoverable on the public Web. Discoverability on the Web may be achieved in two ways:

1. Data objects may be registered in public Web databases;
2. Data objects may just be published on the Web, without being registered in databases.

The first method is mostly applied to data that are ready to be published, whereas the second method arises for data shared within the development process. Data objects can be found by either searching within databases or by using Web search engines. In both cases, data objects have to be well-defined, in order to be reachable from people searching for them. Data objects and repositories of data objects in the NAIADES ecosystem should support both methods. Deposition databases and data resources are required, so that all NAIADES partners can deposit and access data in a data repository.

6.3 Recommendations for Accessibility

In the FAIR Digital Object model, access to a data object is granted through a persistent identifier (PID). There are two main models of PIDs that are primarily utilised:

1. Compact Identifiers (e.g. n2t.net and identifiers.org) mainly used in life sciences, and
2. Digital Object Identifiers (DOIs) used everywhere else.

Regardless the PID form, when data object is being requested by a user, it is presented as a proper Uniform Resource Locator (URL) on the public Web. PIDs should correspond via returning an HTML landing page to the requester, including metadata about the data object.

Rec 13 A data object must have a unique persistent Identifier (PID).

Rec 14 A data object's PID should form part of its metadata record.

6.4 Recommendations for Interoperability

Three key features of interoperability are: good metadata, compatible licensing and open data formats. All partners of NAIADES project are encouraged to develop and implement rich metadata descriptions for all shareable data objects, under the appropriate licenses. Moreover, partners should support and foster sharing and publication of data objects on the Web, in non-proprietary formats.

Rec 15 Data objects should be published on the Web in non-proprietary formats.

6.5 Recommendations for Reusability

EU sui generis database has adopted version 4.0 of the Creative Commons license suite, resulting in an increasing usage of schemes for licensing data. In this way, each license turns into a machine-readable format. A combination of a machine-readable statement, with a data object's metadata record, enhances automated reusability.

[10] *S Hodson et al, Turning FAIR into reality, European Commission Expert Group on FAIR Data, November 2018*

By using licences from this suite, partners clarify which data are available for reuse and the restrictions that govern such actions. In addition, they permit the widest possible reuse of data, they specify the length of time for which data will remain reusable as well as if data is usable by third parties after the end of the project.

Rec 16 Data objects in the NAIADES ecosystem should adopt licenses from the Creative Commons license suite. Where data are publicly available for reuse, these should be one of:

- CC BY 4.0 (attribution),
- CC BY SA 4.0 (distribution under license), and
- CC0 (public domain or rights waiver).

However, as described in Data Management Plan, some of the generated data will be held as confidential for an amount of time, due to the fact that, during the lifecycle of the project, several patents are expected to be developed. Thus, benefits from these patents must be granted to their creators.

7 Data Sharing

Data sharing is a technical issue which needs to be carried out according to the FAIR principles. Data models developed during the NAIADES project should facilitate sharing of data. When data might be shareable for research purposes under specific restrictions, issues regarding data security occur.

A copy of digital data can be reproduced relatively easily and then distributed to a collaborator. Undoubtedly, personal data can be encrypted before sharing it with others. However, despite the security measures taken before the transmission, risks of data leakage significantly increase. The loss of personal data can cause serious damage to a data subject, even if the controller is able to regain data that has been leaked.

During the lifespan of NAIADES, all partners should use computational environments carefully controlled against data leakage. In this way, they can handle personal data, with respect to the code of conduct. Partners are encouraged to use such approaches, especially when they have to link multiple datasets together. However, data linkage in cross-disciplinary teams poses a special case for handling personal data. A combination of publicly available data along with multiple identified data sets, increases risk of accidental (or malicious) disclosure and data leakage. Even de-identification of personal data retains a residual risk, as linking datasets together increases the risk of re-identification.

Rec 17 Shareable data should be used in an appropriate manner.

Rec 18 Data should be shared to trusted users, in order to use them properly.

Rec 19 All data access facilities should take measures against unauthorised use.

Rec 20 The data itself should not include a disclosure risk.

8 Access to NAIADES data

Access concerns every person, team, institution or organisation from private or public sector, that intends to use NAIADES data or any other services. Specifically, NAIADES users may be associated with the following organisations:

- Public research, higher education, international or non-profit private research organisations as well as universities

- Public services which are affiliated with the government, apart from the academia and public research organisations,
- Private companies and businesses, and
- Any other organisation which does not belong to the abovementioned categories, e.g. persons from non-governmental organisations (NGOs).

Open access enhances findability and the use of results produced during the development process. In addition, it determines the access to available data and the acknowledgement of contributors' work. As part of the evaluation process of the NAIADES ecosystem, the opportunities offered by the underlying framework of the project will be tested in demonstrators. The main purpose is to demonstrate to various target groups (Water service providers, Utilities providers, municipalities as well as sensor and technology organisations) the benefits of the project's solutions.

Rec 21 Access to a carefully defined selection of NAIADES data, data products and digital tools, generated in the demonstrators side, will be available through an open access repository. This is done in accordance with the data protection recommendations.

Feedback gained from demonstrators will be important for critical adjustments that will render the project's solutions ready to be introduced in the market.

At a preliminary stage, there is an agreement upon open access publishing. However, in the future, partners may opt for gold or green access to peer-reviewed scientific and non-scientific publications, which might result from the project, depending on the type of information to be published.

Open access principles, combined with certain restrictions, may be implemented for specific data sets, as the access to them could cause a potential misuse and violation of personal data protection laws. Restrictions refer to data which are not openly accessible, without using password. These restrictions are decided case by case and negotiated with the data creators.

9 Licenses

Access rights to data created by the partners of the NAIADES project are defined under specific licenses. Licenses are granted by the copyright owner of the IPR and vary depending on the data type (data, databases, metadata etc.). Their purpose is to prevent the IP from any unauthorised activity, such as malicious reproduction, distribution, reuse etc. Open Data Commons[11] (ODC) and Creative Commons[12] (CC) are public open licenses that propose legal and flexible solutions, in order to provide any user the right to access, share and use content and datasets as well as data repositories.

- ODC have developed licenses that protect rights regarding databases, granted by copyrights. These licenses allow users to freely access, share, modify and use databases' content, subject to specific requirements.
- CC pose several restrictions to user rights, based on four conditions[13]:
  - i. Attribution (BY): It allows copying, distributing, making derivative work and changing of content, in case credits are given to the licensor.
  - ii. Share-alike (SA): It permits distributing derivate work under a similar license.
  - iii. NonCommercial (NC): It allows copying, distributing and making derivative work only for non-commercial purposes.

[11] *https://www.opendatacommons.org*

[12] *https://creativecommons.org*

[13] *https://creativecommons.org/use-remix/cc-licenses/*

iv. NoDerivatives (ND): It permits copying, distributing and displaying, but does not allow making derivative work or changing of its content.

Furthermore, there are two more licenses that can be used in order to grant open access to data generated by NAIADES partners; Community Data License Agreement – Permissive, Version 1.0[14] (CDLA) and Data Archiving and Networked services[15] (DANS).

- CDLA enables users to use, modify and adapt datasets and data within them. Whoever uses datasets shared under this license, he must give credits to the creator of the data. By using this license, there are no obligations or restrictions on results produced from computational use of the data.
- DANS license renders data openly accessible in four different ways:
    1. Fully open access,
    2. Unlimited access only to specific users,
    3. Access with depositor's authorisation only and
    4. Access from another source – data are permanently archived at DANS but they are available through another source.

Another known licenses are the Massachusetts Institute of Technology (MIT) License and Apache License. These licenses are utilised mainly in order to ensure free reuse and modifiability of software, addressing in that way patent rights. Thus, these licenses are principally chosen by developers, allowing secure distribution of source code from their projects, but they are not selected for access rights to datasets and data repositories.

[14] *https://cdla.io/permissive-1-0/*

[15] *https://dans.knaw.nl*

10 Conclusions

This report describes how IPR enable each partner of the NAIADES project to own the rights of the generated data as well as to provide access rights for data which is not reachable in open access repositories. Moreover, it presents clear guidelines regarding the compliance of partners with EU regulations for data protection. This report reflects the FAIR principles that need to be taken into account, in order to render NAIADES data easily findable, accessible, interoperable and reusable within the state members across Europe. In addition, it specifies the requirements needed for effective data sharing, the framework for data retrieval in open access repositories and the licences for access rights, in cases where the NAIADES data is accessed under specific restrictions.

As a result from this report, twenty-one recommendations were arisen, which are listed below:

Table 1: *Recommendations for Data Policy*

Rec 1	As defined by the EU law, there are two key players who process personal data: data controllers and data processors.
Rec 2	Every time a partner, which participates in the development of applications and platforms, is appointed as a controller, such partner should make sure that all requirements in compliance with the European law.
Rec 3	A controller decides either the process of data by the processor, for instance through staff under his direct authority, or to delegate, all or part of the processing activities, to an external organisation.
Rec 4	Location and traffic data may only be processed when they are made anonymous, or with the consent of the volunteer participants, to the extent and for the duration, necessary for the provision of a value added service.
Rec 5	NAIADES users must be informed, prior to the acceptance of their consent, about the type of location and/or traffic data which will be processed, the purposes and the duration of the process, as well as whether the data will be transferred to a third party.
Rec 6	Controllers are always responsible for data protection breaches, according to the GDPR.
Rec 7	Controllers should appoint data processing only to processors that can provide evidence regarding their ability to carry out technical and organisational measures.
Rec 8	Processors cannot appoint data processing to another controller without getting the controllers' consent.
Rec 9	The processors have to commit that they will act strictly under the controllers' instruction and they will take the appropriate security measures.
Rec 10	Each controller should maintain a record with all processing activities that are carried out under his/hers/its authorisation, and this should also be done by the processor. These records have to be made available to the supervisory authority upon request.

Rec 11	The acquired personal data will be transformed into anonymous data when they are expected to be used for commercial purposes. Moreover, personal data will be aggregated with other relevant data before any processing, making difficult in this way to identify any data subjects.
Rec 12	Shareable data objects within the NAIADES ecosystem should be considered as FAIR Digital Objects.
Rec 13	A data object must have a unique persistent Identifier (PID).
Rec 14	A data object's PID should form part of its metadata record.
Rec 15	Data objects should be published on the Web in non-proprietary formats.
Rec 16	Data objects in the NAIADES ecosystem should adopt licenses from the Creative Commons license suite, where data are publicly available for reuse.
Rec 17	Shareable data should be used in an appropriate manner.
Rec 18	Data should be shared to trusted users, in order to use them properly.
Rec 19	All data access facilities should take measures against unauthorised use.
Rec 20	The data itself should not include a disclosure risk.
Rec 21	Access to a carefully defined selection of NAIADES data, data products and digital tools, generated in the demonstrators side, will be available through an open access repository. This is done in accordance with the data protection recommendations.

These policy recommendations are expected to provide a practical guidance towards the effective implementation of data policies within the NAIADES ecosystem. They create a framework under which all consortium members will plan the use of data that they will produce during their work. Furthermore, they present the standards on how data will be preserved and shared in the future for various activities related to the industry, institutions and organisations.