i-nergy
Artificial Intelligence for Energy

Deliverable 2.2: AI-related Ethical Guidelines and Recommendations

15th June 2022

Lead: ICCS

Version: 2.00

Classification: Public

I-NERGY project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 101016508

www.i-nergy.eu

Disclaimer

The sole responsibility for the content of this publication lies with the authors. It does not necessarily reflect the opinion of the European Union. The European Commission is responsible for any use that may be made of the information contained therein.

I-NERGY project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 101016508

I-NERGY Project Profile

Grant Agreement No.:	101016508
Acronym:	I-NERGY
Title:	Artificial Intelligence for Next Generation Energy
Type:	Innovation Action (IA)
URL:	https://i-nergy.eu/
Start Date:	01/01/2021
Duration:	36 months

Work Package:	WP2 – Requirements, Specifications, Synergies and Alignment with AI4EU
Deliverable Leader:	ICCS
Authors (Organisation)	George Lampropoulos (ICCS), Sotiris Pelekis (ICCS), Spiros Mouzakitis (ICCS), John Psarras (ICCS), Francesco Saverio Nucci (ENG)
Internal Reviewers:	David Carro Santomé (COMS), Muhammad Zeshan Afzal (DFKI)
Status:	Final
Date of Delivery:	15th June 2022
Version:	2.00
Classification:	Public

Document History

Version	Date	Author (Partner)	Remarks
1.10	10/03/2022	Sotiris Pelekis (ICCS)	ToC Update
1.20	10/05/2022	Sotiris Pelekis (ICCS)	Updated Section 2 Content with new initiatives
1.30	20/05/2022	Ourania Markaki (ICCS)	Section 2.4.2 added
1.40	12/05/2022	Sotiris Pelekis (ICCS)	Restructured deliverable
1.45	21/05/2022	Sotiris Pelekis (ICCS)	Added I-NERGY's activities regarding Trustworthy AI (Section 3)
1.50	26/05/2022	Sotiris Pelekis (ICCS)	Added guidelines for Trustworthy AI in I-NERGY (Section 3)
1.60	31/05/2022	Vangelis Karakolis (ICCS)	Added methodology for Trustworthy AI in I-NERGY (Section 3)
1.80	10/6/2022	Andrej Čampa (COMS)	Internal Review
1.90	13/6/2022	M. Zeshan Afzal (DFKI)	Internal Review
1.91	14/6/2022	Vangelis Karakolis (ICCS)	Addressing reviewers' comments
2.00	15/06/2022	Spiros Mouzakitis, John Psarras (ICCS)	Final Version

Partners

	Participant Name	Short Name	Country Code	Logo
1	INSTITUTE OF COMMUNICATION AND COMPUTER SYSTEMS	ICCS	GR	ICCS ΕΠΙΣΕΥ
2	ENGINEERING – INGEGNERIA INFORMATICA SPA	ENG	IT	ENGINEERING
3	RHEINISCH-WESTFAELISCHE TECHNISCHE HOCHSCHULE AACHEN	RWTH	DE	E.ON Energy Research Center | RWTH AACHEN UNIVERSITY
4	COMSENSUS, KOMUNIKACIJE IN SENZORIKA, DOO	COMS	SI	COMSENSUS
5	FUNDACION CARTIF	CARTIF	ES	[TECHNOLOGY CENTRE] CARTIF
6	DEUTSCHES FORSCHUNGSZENTRUM FUR KUNSTLICHE INTELLIGENZ GMBH	DFKI	DE	DFKI German Research Center for Artificial Intelligence
7	PARITY PLATFORM IDIOTIKI KEFALAIOUXIKI ETAIREIA	PARITY	DE	PARITY
8	FUNDINGBOX ACCELERATOR SP ZOO	FBA	PL	FundingBox
9	CENTRO DE INVESTIGACAO EM ENERGIA REN - STATE GRID SA	R&D NESTER	PT	R&D NESTER
10	ASM TERNI SPA	ASM	IT	ASM ASM Terni S.p.A.
11	SONCE ENERGIJA D. O. O.	SONCE	SI	SONCE | Partner of SUNCONTRACT
12	IRON THERMOILEKTRIKI ANONYMI ETAIREIA	HERON	GR	HERON
13	ZELENA ENERGETSKA ZADRUGA ZA USLUGE	ZEZ	HR	ZEZ Zelena Energetska Zadruga
14	STUDIO TECNICO BFP SOCIETA A RESPONSABILITA LIMITATA	BFP	IT	BFP
15	VEOLIA SERVICIOS LECAM SOCIEDAD ANONIMA UNIPERSONAL	VEOLIA	ES	VEOLIA
16	RIGA MUNICIPAL AGENCY "RIGA ENERGY AGENCY"	REA	LV	REA Riga Energy Agency
17	FUNDACION ASTURIANA DE LA ENERGIA	FAEN	ES	FAEN Fundación Asturiana de la Energía

Executive Summary

This document constitutes a report of the work performed within Task 2.2 "*AI HLEG/AI4EU Ethics Guidelines Compliance Assurance, Data Protection and IPR*". Therefore, in correspondence with the objectives of this task, this Deliverable provides an analysis of the ethical and legal principles that AI systems and solutions within I-NERGY should follow as well as guidance for the handling of possible IPR issues.

To assess these legal principles and provide guidance for compliance with the applicable legal frameworks, I-NERGY follows an approach which considers a) the applicable data protection frameworks b) the network and information security directive and c) some key energy domain regulations and standards. Some of these legal frameworks considered impose direct requirements for the solutions being developed within the project, while the others provide guidance for compliance for the future products and services.

The approach to the ethical dimension of AI systems within I-NERGY aims to identify possible ethical issues that may arise focusing on energy domain, and to provide a set of recommendations to tackle these issues and strengthen the trustworthiness of the developed solutions, focusing mainly on the respective Ethics Guidelines developed by the High-Level Expert Group on AI.

Concerning the IPR implications that may arise in terms of background and foreground, this document aims to provide guidance on how such issues are managed within the I-NERGY consortium.

Table of Contents

Figures

Tables

Glossary

Term	Definition	Term	Definition	Term	Definition
ACER	EU Agency for the Cooperation of Energy Regulators	**ENISA**	European Union Agency for Cybersecurity	**IT**	Information Technology
AI	Artificial Intelligence	**EPES**	Electrical Power and Energy Systems	**LCA**	Life-cycle assessment
HLEG	High-level Expert Group on Artificial Intelligence	**ESCO**	Energy Service Companies	**LE**	Large Enterprise
API	Application Programming Interface	**EU**	European Union	**ML**	Machine Learning
BDVA	Big Data Value Association	**EV**	Electric Vehicle	**NECP**	National Energy and Climate Plan
CA	Consortium Agreement	**FAIR**	Factor Analysis of Information Risk	**NIS**	Network and Information Security
CAHAI	Ad hoc Committee on Artificial Intelligence	**GA**	Grant Agreement	**NISD**	Network and Information Security Directive
CCTA	Central Communication and Telecommunication Agency	**GDPR**	General Data Protection Regulation	**OCTAVE**	Operationally Critical Threat, Asset, and Vulnerability Evaluation
CEER	Council of European Energy Regulators	**GHG**	Greenhouse Gas	**OES**	Operators of Essential Services
CEP	Clean energy for all Europeans package	**ICT**	Information and Communication Technology	**P2P**	Peer-to-peer
CoE	Council of Europe	**IEA**	International Energy Agency	**RES**	Renewable Energy Sources
CRAMM	CCTA Risk Analysis and Management Method	**IEC**	International Electrotechnical Commission	**SAREF**	Smart Appliances REFerence ontology
CSIRT	Computer Security Incident Response Team	**IoT**	Internet of Things	**SIS**	Smart Information Systems
DAIRO	Data, AI and Robotics	**IP**	Intellectual Property	**SME**	Small to Medium Enterprise
DL	Deep Learning	**IPR**	Intellectual Property Rights	**TFEU**	Treaty on the Functioning of the European Union
DLT	Distributed Ledger Technology	**IRAM**	Information Risk Assessment Methodology	**TSO**	Transmission System Operator

DPIA	Data Protection Impact Assessment	**IRENA**	International Renewable Energy Agency	**UC**	Use Case
DSO	Distribution System Operator	**IS**	Information Security	**UDHR**	Universal Declaration of Human Rights
DSP	Digital Service Provider	**ISMS**	Information Security Management System	**WP**	Work Package
EC	European Commission	**ISO**	International Organization for Standardization		
ECHR	European Convention on Human Rights	**ISSP**	Information Security Policy		

1 Introduction

1.1 Purpose

This Deliverable in correspondence with Task 2.2 aims to:

- provide an overview of the applicable legal frameworks in the context of I-NERGY project, concerning the European and International law.
- identify possible implications and ethical issues of AI in scope of I-NERGY.
- provide a set of recommendations alongside a clear methodological framework that will guide I-NERGY partners to incorporate the principles and requirements of trustworthy AI within their developed AI systems.
- raise awareness regarding IPR and their handling within the project.

In this manner, it will provide guidance for the instantiation of AI systems in the energy domain and facilitate the compliance with the ethical and legislative requirements.

1.2 Structure of the document

The deliverable is organised in four sections, as described below:

Section 1 provides an introductory description of the document concerning its scope, its structure and its relations to other tasks of the project.

In Section 2, the regulatory framework along with its relation to the project is presented concerning the dimensions of a) data protection, focusing mainly on the GDPR, b) network and information security, based on the NISD and c) energy sector, considering international and EU legislation, initiatives, and guidelines.

Initially, section 3 indicatively presents some common ethical issues and risks for AI systems in the energy domain and I-NERGY project in particular, alongside tools and methods to be adopted by partners to mitigate those risks. Subsequently, a methodological procedure for I-NERGY technical and pilot partners is set, to ensure trustworthiness within their AI systems.

Section 4 addresses IPR handling and protection within the project.

1.3 Relations to I-NERGY environment

Analysis carried out in Task 2.2 provides a regulatory-oriented benchmark against which I-NERGY concepts, activities, technology and services can be assessed. Activities in this task are closely connected with the following internal activities and achievements of:

- Task 1.4. Data Management Plan, since this task will provide a detailed plan about which of the data collected and generated will be shared publicly, as open data increase the interest towards the project. The results of this task are complementary to the current deliverable. Therefore, partners are strongly advised to refer themselves to the respective deliverable D1.2 especially regarding data protection and privacy issues.

- Task 7.3. Business and Exploitation Planning, since activities of this task include the handling of legal and IPR issues as well as the identification of third party's rights and their licensing schemes. Moreover, it concerns the ownership and access rights for the IP generated within the project.
- Task 3.6 - Security Framework, since this task will technically consider the data protection requirements alongside the security and safety related requirements of AI systems developed within I-NERGY project.

2 Regulatory framework

The overall vision and main objective of I-NERGY is to deliver an energy-specific open modular framework for supporting AI-on-Demand in the energy sector (AI4 Energy), by capitalising on state-of-the-art AI, as well as IoT, semantics and data analytics technologies. Key elements to achieve this goal include among others the sharing of a variety of energy and non-energy-related data, the development of ML/DL models and the utilisation of these models to design and develop energy analytics applications. It is therefore vital for the project to ensure that the whole management of these data, including collection, storage, processing, as well as the development of the models and applications comply with the applicable legislative frameworks at the EU and international level. Compliance with the regulations contributes to the protection of individuals rights and private property and simultaneously increases the trustworthiness of the developed solutions and the projects' impact.

This chapter aims to provide an overview of the legal frameworks pertinent to the context of I-NERGY project. To identify these frameworks, I-NERGY will consider the perspectives of: a) privacy and data protection, b) network and information security and c) energy policy and frameworks.

2.1 Privacy and data protection

It is vital for I-NERGY project to ensure that any personal or private information and data collected or generated will be adequately safeguarded, protecting the privacy of their owners. This includes a variety of proprietary data for the pilots' needs, such as energy and non-energy-related data owned or managed by DSOs, TSOs, ESCOs etc., any personal information that might be collected through questionnaires, forms or submitted in the platform, as for example the information needed for the participation to the Open Calls[1] and any proprietary assets brought by partners into the project. Therefore, it is meaningful to examine the concepts of data protection and privacy and their recognition as fundamental rights.

- **Privacy**. The concept of privacy is closely related to the notion of autonomy and human dignity, to the right to control information about yourself and to be let alone [1]. Right to privacy, also expressed as right to private life, is enshrined in Article 12 of the Universal Declaration of Human Rights, in Article 8 of European Convention of Human Rights and in Article 7 of European Charter of Fundamental Rights.
- **Data protection**. The notion of data protection is connected with the action of keeping safe any personal or proprietary information [1]. It is relevant to the whole lifecycle of data, including, among others, collection, storage, processing, and dissemination. Protection of personal data is a recognised and enshrined right according to Article 8 of European Charter of Fundamental Rights. In addition, Article 16 of the Treaty on the Functioning of the European Union (TFEU) enshrines the right to protection of personal data and obliges European Parliament and Council to lay down the protection rules for the personal data processing [2].

[1] Access link for Open Calls available in project's website: https://www.i-nergy.eu/

I-NERGY will be in line with the principles of privacy and data protection and the conventions and charters where these rights are enshrined as mentioned above. Therefore, it is useful to provide a few details about them.

Universal Declaration of Human Rights (UDHR)

The UDHR was proclaimed on 10 December 1948 in Paris by the United Nations General Assembly "as a common standard of achievements for all people and all nations" [3]. The UDHR "is a milestone document in the history of human rights" and "it sets out, for the first time, fundamental human rights to be universally protected" [3]. Right to privacy is expressed in Article 12 as following:

"No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks."

European Convention on Human Rights (ECHR)

ECHR or Convention for the Protection of Human Rights and Fundamental Freedoms as formally entitled, was adopted by the Council of Europe in Rome on 4 November 1950 and came into force in 1953 [4]. Since then, there were several amendments and supplements in the Convention. The ECHR sets forth a number of fundamental rights and freedoms and all Contracting Parties are obligated to comply with it. To ensure the observance of these obligations, the European Court of Human Rights was set up in Strasbourg in 1959 [4]. The right to privacy is enshrined in Article 8, as following:

"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."

EU Charter of Fundamental Rights

The Charter of Fundamental Rights of the European Union was proclaimed in 2000 by European Parliament, the Council of the European Union and the European Commission and after several amendments were proclaimed again in 2007 [5] The same legal value with Treaties was given to the Charter on 1 December 2009, through the adoption of the Treaty of Lisbon [6] as stated in its Article 6. The Charter enshrines both the right to privacy and right to data protection in Article 7 and Article 8, respectively:

"Everyone has the right to respect for his or her private and family life, home and communications." (Article 7)

"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." (Article 8)

2.1.1 Regulatory frameworks for data protection

This section aims to provide an analysis of the regulatory frameworks that I-NERGY will comply with, in order to ensure data and privacy protection within the project. In this direction, the main legal standards that will be considered include a) the Council of Europe Convention 108 b) the General Data Protection Regulation and c) National legislation for data protection in the pilots' regions.

2.1.1.1 Council of Europe Convention 108

Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108) was the first legally binding international instrument regarding protection of personal data [7]. The Convention was opened for signature by member States and for accession by non-member States in 1981 and entered into force on 1 October 1985. It protects the individual against abuses which may come with the collection and processing of personal data and seeks to regulate the transborder flows of personal data [7].

Convention 108 lays down the basic principles for data protection. There are principles, described under Article 5, concerning the lawfulness of the processing of personal data as well as their quality, namely, they should be accurate, adequate, and relevant to the purposes that they are collected and should not be kept longer than is required for the purposes collected [7]. Emphasis is also given on the security measures and safeguards that should be taken to protect the data and the subject[2]. In addition, the Convention reinforces the right of data subject to know whether his/her personal data are stored and to be able to correct or erase these data. Regarding special categories of personal data (e.g., data revealing racial origin, political opinions), automated processing should not be applied if adequate safeguards are not provided according to Article 6. Furthermore, the Convention describes the exceptional cases when restrictions could be applied to the rights of data subjects for specific reasons such as public safety[3].

Convention 108 underwent a process of modernisation, leading to a major update of the Treaty [8] in 2018 by CoE, when also the Protocol amending the Convention was opened for signature on October 10 [9]. The modernisation of the Convention aims a) to address the challenges for privacy and data protection that arise from the use of new information and communication technologies, and b) to strengthen the Convention's follow-up mechanism. The modernised Convention 108, also referred to as "Convention 108+", provides more detailed and concise guidance, recommendations, and additional safeguards to tackle these new challenges, as well as to ensure the consistency and compatibility with other normative frameworks and in particular EU's law, while maintaining and, at the same time, strengthening the already enshrined data protection principles [9].

Some of the amendments included in the Protocol relate to the principles of proportionality, data minimisation, lawfulness, fairness and transparency of data processing, and accountability of data controllers. In addition, rights of persons in an algorithmic decision-making context are also recognised. Furthermore, the "privacy by design" principle is set as requirement [8].

[2] Convention 108, Article 7 & Article 8

[3] Convention 108, Article 9

2.1.1.2 General Data Protection Regulation

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [10] entered into force on 24 May 2016 and applies since 25 May 2018. The predecessor of GDPR was the Directive 95/46/EC, whose validity ended since the GDPR was fully applied on 25 May 2018. GDPR is directly applicable as a regulation under EU law, in comparison with its predecessor, which as a Directive, had to be incorporated into EU countries' national legislation [11].

GDPR aims to strengthen individuals' rights in the digital age and help them to control better their personal data as well as to provide a set of unified and unambiguous rules for companies, organisations, and public bodies, facilitating administrative processes and increasing their trustworthiness [12]. In addition, GDPR concerns the free movement of personal data within the Union and the rules that should accompany this movement.

Concerning the material scope of GDPR, it applies to partial or complete processing of personal data by automated means and to processing of personal data by means other than automated when these data form or are expected to form part of a filling system.[4] The territorial scope of the Regulation is described under Article 3. Furthermore, exceptional cases where GDPR is not applicable are described in the context of Regulation.

This section does not aim to provide an extensive analysis of GDPR but to briefly describe some key points of the regulation. In addition, many of the following terms and principles are also stated in other regulations, such as CoE Convention 108, which during the modernisation process is more closely aligned to GDPR. Nevertheless, the following paragraphs put focus on the respective sections of GDPR.

2.1.2 Data protection terms

Article 4 of GDPR provides a set of definitions, some of them are listed below and will be used widely in this section:

- ***personal data*** refers to any information that relates to an identified or identifiable natural person. Term 'identifiable' means that a person can be identified either directly or indirectly given an identifier. Examples of identifier could be location data, tax number or an IP address in digital life (Article 4 - para 1).

  Special categories of personal data exist that are particularly sensitive, as for example data revealing racial or ethnic origin. Processing of such data should have occurred only under specific conditions as mentioned in Article 9, otherwise, these data should not be processed[5].

- ***processing*** relates to any operation on personal data, such as collection, dissemination, storage either by automated or other than automated means. (Article 4 - para 2)
- ***data subject*** refers to the person whose data are processed.
- ***data controller*** "means the 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

[4] GDPR, Article 2

[5] See also Recital 51 of GDPR.

data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law" (Article 4 - para 7)

- ***processor*** "means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller". (Article 4 - para 8)
- ***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." (Article 4 - para 5)

  Personal data undergone pseudonymisation should be considered as information on an identifiable person and therefore be appropriately protected according to Regulation. (Recital 26)
- ***third party*** "means 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;" (Article 4 - para 10)
- ***consent*** "of the data subject means 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;" (Article 4 - para 11)

  Conditions for obtaining valid consents are described in Article 7. Consent is one of the legitimate cases for processing personal data as mentioned in Article 6.
- ***filing system*** "means any structured set of personal data which are accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis;" (Article 4 - para 6)

2.1.3 Key principles on processing personal data

Article 5 of GDPR describes the principles that processing of personal data should respect.

Principles (GDPR, Article 5)
〉 **Lawfulness, fairness, transparency** - Processing of personal data should be lawful, fair, and transparent.
〉 **Purpose limitation** - Personal data should be collected and processed for specific and legitimate purposes and any further processing should be compatible with these purposes.
〉 **Data minimisation** - personal data should be adequate, relevant, and limited to what is necessary for the specified purposes.
〉 **Accuracy** - Personal data should be accurate and if necessary, kept up to date. In case of inaccuracy of data, they should be erased or rectified without delay.
〉 **Storage limitation** - Personal data that permit the identification of data subjects should only be kept in this form for the time necessary for the specified purposes that they were collected and processed.
〉 **Integrity and confidentiality** - Processing must be accompanied by appropriate technical and organisational measures to ensure the security of personal data, including protection against unauthorised or unlawful processing, accidental loss, destruction, or damage.
〉 **Accountability** - The controller shall be responsible and able to demonstrate compliance with the principles mentioned above.

Table 1: Key principles on processing personal data

At this point, it is meaningful to also mention the principle of "*data protection by design and by default*" that is addressed under Article 25. The "by design" principle means that the controller shall determine and implement the appropriate technical and organisational measures starting from the early stage of design of processing activities and not only at the time of processing itself[6]. The "by default principle" requires controller to implement appropriate measures to ensure that by default only the necessary, regarding the purposes of processing, personal data are processed[7].

Lawful processing

As described under Article 6 of GDPR the principle of lawfulness requires personal data to be processed based on the consent given by the data subject or on another legal basis according to EU or Member State law. Points (b)-(f) of paragraph 1 of Article 6 describe other conditions that form legal basis:

- "processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract."

[6] GDPR, Article 25, para 1

[7] GDPR, Article 25, para 2

- "processing is necessary for compliance with a legal obligation to which the controller is subject."
- "processing is necessary in order to protect the vital interests of the data subject or of another natural person."
- "processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller."
- "processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child."

Article 6 provides more details regarding conditions for lawful processing and further processing of collected data.

Rights of the data subject

Chapter 3 of GDPR describes the rights of the data subject and the obligations of the data controller to respect these rights. In the following points, these rights are briefly mentioned:

- Right to be informed (Article 13, Article 14)
- Right of access by data subject (Article 15)
- Right to rectification (Article 16)
- Right to erasure (Article 17)
- Right to restriction of processing (Article 18)
- Right to data portability (Article 20)
- Right to object (Article 21)
- Right not to be subject to a decision based solely on automated processing (Article 22)

2.1.4 Responsibilities of controller and processor

A controller is responsible to implement appropriate technical and organisational measures that ensure processing is compliant with GDPR and enable the demonstration of this compliance[8]. These measures may include data protection policies, codes of conduct[9] and certification mechanisms[10] [11]. A set of measures and precautions safeguarding the security of processing are also described under Article 32. Such possible measures include: "a) the pseudonymisation and encryption of personal data; (b) the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services; (c) the ability to restore the availability and access to personal data in a timely manner in the event of a physical or technical incident; (d) a

[8] GDPR, Article 24, para 1

[9] See Article 40 of GDPR

[10] See Article 42 of GDPR

[11] GDPR, Article 24, para 2 & 3

process for regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing”.[12] In addition, assessing the level of security requires taking into consideration the risks deriving from data processing and in particular from “accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to personal data transmitted, stored or otherwise processed”[13] . Moreover, the controller and processor should ensure that persons having access to personal data process them only according to their instructions or due to requirement of EU or national law.[14]

In cases that two or more controllers define the purposes and means of processing they are joint controllers and should determine their respective responsibilities.[15]

Where controller delegates a processor to perform the processing of personal data, must use only processors providing sufficient guarantees to implement adequate technical and organisational measures ensuring compliance with GDPR and protection of the data subject’s rights.[16] In such cases, processing by processor should be performed according to a contract or other legal act under EU and the Member States. Article 28 describes in detail the subject of this contract/legal act between processor and controller as well as the obligations and responsibilities of the processor.

The responsibilities of the processor and controller concerning the maintenance of records of processing activities are described in Article 30. The obligation for keeping such records “shall not apply to an enterprise or an organisation employing fewer than 250 persons unless the processing it carries out is likely to result in a risk to the rights and freedoms of data subjects, the processing is not occasional, or the processing includes special categories of data as referred to in Article 9(1) or personal data relating to criminal convictions and offences referred to in Article 10.”[17]

As mentioned in Article 31, both controller and processor shall cooperate with the supervisory authority upon request. The obligations of controller and processor and the actions they should take in cases of personal data breach are described in articles 33 and 34, including the notification of supervisory authority and data subject if the specified conditions are met.

When processing is likely to result in a high risk to the rights and freedoms of individuals, controller shall carry out a Data Protection Impact Assessment before the processing.[18] A DPIA is also required under the specified conditions of article 35, where also more details are provided. When DPIA indicates that processing would result in high risk if no measures are taken, controller should consult the supervisory authority.[19]

[12] GDPR, Article 32, para 1

[13] GDPR, Article 32, para 2

[14] GDPR, Article 32, para 4

[15] GDPR, Article 36

[16] GDPR, Article 38

[17] GDPR, Article 30

[18] GDPR, Article 35, para 1

[19] GDPR, Article 36

2.1.5 National data protection frameworks

All countries where partners are located, are obliged to comply with Regulation (EU) 2016/679 (GDPR) as Member States of EU and CoE Convention 108[20] mentioned above. GDPR, as Regulation is directly applicable across the EU. Besides the already mentioned regulations an overview of the national data protection frameworks regarding the processing of personal data in pilots' regions is provided below.

2.1.5.1 Croatia

In Croatia the Act on the Implementation of the General Data Protection Regulation (Official Gazette, No. 44/2018) was adopted on 25th May 2018 providing a national implementation of the GDPR. [13] In addition, the Act on confirmation of the Convention for the protection of individuals with regard to automatic processing of personal data (Convention 108) and of the Additional Protocol to the Convention 108 related to data protection authorities and international data exchange (Official Gazette, No. 04/05, International Agreements) implement these legal acts in Croatia. [14] The data protection authority in Croatia is the Croatian Personal Data Protection Agency[21].

2.1.5.2 Greece

National legislation regarding personal data protection in Greece contains Law 4624/2019 and Law 2472/1997. [15] Law 4624/2019 concerns the establishment and operation of the Data Protection Authority (i.e. Hellenic Data Protection Authority[22]), the measures that should be adopted for the implementation of GDPR and the transposition of Directive (EU) 2016/680. Law 2472/1997 regarding protection of individuals with regard to the processing of personal data has been repealed, except for the provisions mentioned in Article 84 of Law 4624/2019. [15]

2.1.5.3 Italy

In Italy, the Personal Data Protection Code (Legislative Decree No. 196 of 30 June 2003) was amended by Legislative Decree No. 101 of 10 August 2018 in order to adapt national legislation to the provisions of GDPR. The Italian Data protection Authority is "Garante per la protezione dei dati personali"[23].

2.1.5.4 Latvia

The Data State Inspectorate "DVI"[24] is the national Data Protection Authority for Latvia providing supervision and guidance on the protection of fundamental rights and freedoms in the field of data protection. The legal preconditions for the implementation of GDPR in Latvia are provided by the "Personal Data Processing Law".[25]

[20] See also Chart of signatures and ratifications of Treaty 108, available at: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108/signatures

[21] https://azop.hr/about-the-agency/

[22] https://www.dpa.gr/en/hdpa/profile

[23] https://www.garanteprivacy.it/

[24] https://www.dvi.gov.lv/lv

[25] Available in Latvian at: https://likumi.lv/ta/id/300099-fizisko-personu-datu-apstrades-likums

2.1.5.5 Portugal

The Portuguese data protection authority is the "Comissão Nacional de Proteção de Dados"-CNPD[26], monitoring and inspecting the compliance with GDPR and other national laws on the protection of personal data. The Law no. 58/2019[27], of August 8th ensures the implementation of GDPR in the Portuguese legal system.

2.1.5.6 Slovenia

Slovenia has not yet implemented the GDPR in its national legal system [16], since the new Personal Data Protection Act "ZVOP-2" is not yet adopted. However, GDPR as a Regulation is applicable in Slovenia. In addition, the Personal Data Protection Act "ZVOP-1" adopted in 2004 is still applicable. The data protection authority for Slovenia is "Information Commissioner of the Republic of Slovenia"[28].

2.1.5.7 Spain

The GDPR is implemented in the context of Organic Law 3/2018, of 5 December 2018, on the Protection of Personal Data and Guarantee of Digital Rights[29] , which also regulates other aspects of data protection and digital rights. The Spanish Data Protection Authority is "Agencia Española de Protección de Datos (AEPD)[30]".

2.1.6 Data protection and anonymisation within I-NERGY

Throughout the project, the handling and processing of data originating from the I-NERGY pilots will be necessary. Therefore, the I-NERGY consortium will continuously monitor all procedures related with these tasks, in order to ensure compliance with the regulatory frameworks that were presented in the previous sections including the GDPR.

2.1.6.1 Data protection

As a general rule, only anonymised or aggregated data (completely disjoined from people identification and profiles) related to the project pilots will be processed and made available.

However, if for specific reasons personal data will need to be processed, the interested partner will appoint a Data Protection Officer and will remain responsible for the data provided during its own research. The partner will also be required to provide evidence of the authorisation to process personal data before access to or use of such data be granted. If defined by the European and national legislative framework, such authorisation will need to be requested by the specific partner to the appropriate competent authority in the partners' countries.

In the case of historical data including personal data, the interested I-NERGY partner will appoint a Data Protection Officer and will be in charge of ensuring that the data sources used will be already complying with the data projection legislation applicable to each of the countries of origin. All original data files will be particularly compliant with the right to access, modify, cancel and object

[26] https://www.cnpd.pt/

[27] Available in Portuguese at: https://dre.pt/pesquisa/-/search/123815982/details/maximized

[28] https://www.ip-rs.si/

[29] Available in Spanish at: https://www.boe.es/eli/es/lo/2018/12/05/3

[30] https://www.aepd.es/es

to further treatment of personal data. If defined by the European and national legislative framework, also for historical data, the interested partner will remain responsible for the data provided during its own research and will also be required to provide evidence of the authorisation to process personal data before access to or use of such data be granted.

All the participants in the project are aware of their obligations as potential data processors (where appropriate) as well as issues beyond data and information protection and privacy described in the previous sections. Actions will be taken to ensure that those handling identifiable information subjects are made fully aware of their responsibilities and obligations to respect best practices and legal requirements under the GDPR. All partners processing personal data will appoint their own Data Protection Officers and will communicate their contact to the Project Coordinator.

2.1.6.2 Data Anonymisation and Pseudonymisation

Regarding I-NERGY anonymisation, it will generally comply with the following statement: The data to be utilised within the context of the project will be already completely anonymised and will not include any personal or sensitive data. The de-identification of datasets has to occur before the beginning of the ingestion: I-NERGY datasets have to be stripped of any direct identifiers (suppression / data masking) or use synthetic data in a way that eliminates the risk of re-identifying the sensitive data.

In order to provide the ability, if desired and applicable, to share data sets (e.g., to the AI4EU community or hackathons) without private information (e.g., household and EV data, asset operational information and identifiers) anonymisation tools can be used during the data ingestion process to protect such information by complete data removal-suppression, generalisation or pseudonymity. In any case, no personal data is foreseen to be used or processed in the I-NERGY services. The main technical tool to be utilised for this purpose is the Data Interoperability and Homogenisation module that will be developed in Task 3.2.

Partners are highly encouraged to refer themselves to the I-NERGY Deliverable 1.2 – Data Management Plan, where a complete description of the anoymisation and pseudonimisation framework and approaches can be found.

2.2 Network and information security

The present project would like to follow the NIS directive in preparing their requirements. This approach has been adopted in order to take into the strong consideration the Network and Information Security (NIS) Directive for all future development and technology transfer after the project completion. The idea of this approach is to have a common understanding of the NIS directive and to spread its adoption to the technical partners during the full project life, specifically during the design and development steps. In this way, the future project results obtained after the the funding phase can be easily compliant with the NIS directives and the guidelines or laws, according to the different countries regulatory guidelines. With this approach, all future products and services carried out by the different project partners under the production and commercialisation steps can be compliant with the network and security rules in a better and easier way, even if the NIS directive cannot be directly applied to the internal mockups and piloting intermediate results under a R&D project.

To this scope in the following a detailed analysis of NIS is presented with a short description of the main reason to design and launch of this directive by the European Commission, including the short history of the directive itself and with the relevant status of the directive in the European ecosystem.

2.2.1 A short history of NIS Directive

On 29 January 2020, the European Commission's new work programme was published. Under the second priority - 'A Europe fit for the digital age', the Commission announced its intention to launch a review of the Directive on security of network and information systems (NIS Directive), in order to 'further strengthen overall cybersecurity in the Union'. According to the adjusted work programme, the review should be adopted in the last quarter of 2020.

The current Directive on security of network and information systems entered into force in August 2016. Member States had to transpose it into their national laws by 9 May 2018. The directive lays down requirements regarding the national cybersecurity capabilities of Member States; rules for their cross-border cooperation; and requirements regarding national supervision of operators of essential services and key digital service providers.

The Commission launched on 7 July 2020 a public consultation on the revision of the NIS Directive that aims to collect views on its implementation and on the impact of potential future changes. The consultation closed on 2 October 2020.

On 16 December 2020, the European Commission and the High Representative of the Union for Foreign Affairs and Security Policy presented a new EU Cybersecurity Strategy that aims to bolster Europe's collective resilience against cyber threats and ensure that all citizens and businesses can fully benefit from trustworthy and reliable services and digital tools. Accordingly, The Commission made two new proposals: a Directive on measures for high common level of cybersecurity across the Union (revised NIS Directive or 'NIS 2'), and a new Directive on the resilience of critical entities.

The NIS Directive has increased the EU national cybersecurity capabilities, requiring Member States to elaborate a National Cybersecurity strategy, to establish Computer Security Incident Response Teams (CSIRTs) and to appoint NIS national competent authorities, improving the cyber resilience of public and private entities in specific sectors and across digital services. However, its implementation proved difficult, resulting in fragmentation at different levels across the internal market. In order to respond to the growing threats due to digitalisation and increase in cyberattacks, the proposed revised NIS Directive NIS 2 repeals the existing NIS Directive. The new proposal broadens its scope, aiming to strengthen the security requirements imposed, addressing security of supply chains, streamlining reporting obligations, introducing more stringent supervisory measures and stricter enforcement requirements including harmonised sanctions regimes across the Member States. It also includes proposals for information sharing and cooperation on cyber crisis management at the national and EU level.

2.2.2 NIS Directive main objectives and scopes

The Directive has three main objectives:

- Improving national cybersecurity capabilities;
- Building cooperation at the EU level; and

- Promoting a culture of risk management and incident reporting among key economic actors, notably operators providing essential services (OES) for the maintenance of economic and societal activities and Digital Service Providers (DSPs).

The NIS Directive is a cornerstone of the EU's response to the growing cyber threats and challenges which are accompanying the digitalisation of our economic and societal life, and its implementation is therefore an essential part of the cybersecurity package presented on 13 September 2017. The effectiveness of the EU's response is inhibited as long as the NIS Directive is not fully transposed in all EU Member States. This was also recognised as a critical point in the Commission's 2016 Communication on Strengthening Europe's Cyber Resilience System [17].

The NIS Directive is the first piece of EU-wide legislation on cybersecurity. It provides legal measures to boost the overall level of cybersecurity in the EU. The Directive on security of network and information systems [18] (the NIS Directive) provides legal measures to boost the overall level of cybersecurity in the EU by ensuring:

- Member States' preparedness, by requiring them to be appropriately equipped. For example, with a Computer Security Incident Response Team (CSIRT) and a competent national NIS authority, cooperation among all the Member States, by setting up a Cooperation Group to support and facilitate strategic cooperation and the exchange of information among Member States.
- A culture of security across sectors that are vital for our economy and society and moreover rely heavily on ICTs, such as energy, transport, water, banking, financial market infrastructures, healthcare and digital infrastructure.

Businesses identified by the Member States as operators of essential services in the above sectors will have to take appropriate security measures and to notify relevant national authorities of serious incidents. Key digital service providers, such as search engines, cloud computing services and online marketplaces, will have to comply with the security and notification requirements under the new Directive.

2.2.3 NISD Guidelines

In the present subsection, a synthesis of the NIS guidelines will be provided to be guidelines to the design of the services and, overall to be baselines for the future products and services implementation after the end of the project in the production phase. These guidelines have been summarised and based on the report provided for DSP and NIS and by ENISA (Guidelines on assessing DSP and OES compliance to the NISD security requirements). The ENISA is the European Union Agency for Network and Information Security, centre of network and information security expertise for the EU, its member states, the private sector and EU citizens. ENISA works with all these groups to develop advice and recommendations on good practices in information security.

First of all, it should be noted that NIS Directive (see Articles 14, 15 and 1), are dedicated to introducing appropriate security measures for operators of essential services (OES) as well as for the digital service providers (DSP). The definition of these operators and their inclusion in a specific list is under the member state control and monitoring.

This approach has been chosen in order to achieve a baseline, common level of information security within the European Union (EU) network and information systems. Information security

(IS) audits and self–assessment/ management exercises are the two major enablers to achieve this objective.

For purposes of completeness and for better understanding, these main three important NISD articles are in the following reported:

Article 14: *"Member States shall ensure that operators of essential services take appropriate and proportionate technical and organisational measures to manage the risks posed to the security of network and information systems which they use in their operations. Having regard to the state of the art, those measures shall ensure a level of security of network and information systems appropriate to the risk posed."*

Article 15: *"Member States shall ensure that the competent authorities have the powers and means to require operators of essential services to provide (b) evidence of the effective implementation of security policies, such as the results of a security audit carried out by the competent authority or a qualified auditor and, in the latter case, to make the results thereof, including the underlying evidence, available to the competent authority."*

Article 16: *"Member States shall ensure that digital service providers identify and take appropriate and proportionate technical and organisational measures to manage the risks posed to the security of network and information systems which they use in the context of offering services referred to in Annex III within the Union. Having regard to the state of the art, those measures shall ensure a level of security of network and information systems appropriate to the risk posed, and shall take into account the following elements: a) the security of systems and facilities, b) incident handling, c) business continuity management, d) monitoring, auditing and testing, and e) compliance with international standards."*

2.2.3.1 NIS specific requirements

After the general introduction, hereinafter a detailed description of the NISD main requirements is presented in the following table.

1.	**Information System Security Risk Analysis**
1.1.	The key personnel should be aware of the main information security risks and the relevant mitigations
1.2.	There should be in place a mechanism for ensuring that all security personnel use the risk management methodology and tools
1.3.	The risk management methodology and/or tools, should be periodically reviewed, taking into account changes and past incidents
2.	**Information System Security Policy**
2.1.	It should be put in place an information security policy (ISSP) and an information security management system (ISMS).
2.2.	Some certifications for specific security risk management standards could be put in place.

2.3. Some information security processes should be reviewed at regular intervals, also taking into account violations, exceptions and incidents which affected other essential operators/ DSP.
3. Information System Security Accreditation
3.1. The systems supporting essential services should be regularly subjected to security scans and they should be integrated within the risk management framework of the organization
3.2. There should be policy/procedures in place for the performance of security assessments and security testing
3.3. The effectiveness of policy/procedures for security testing should be carefully evaluated
4. Information System Security Indicators
4.1. The KPIs implemented in systems supporting essential services should be able to be assessed versus their effectiveness at all times
4.2. Policy/procedures should be put in place for the implementation of security indicators for testing the systems supporting essential services
4.3. The aforementioned policy/procedures should be periodically reviewed and updated
5. Information System Security Audit
5.1. An updated policy and/ or procedure for performing information system security assessments should be put in place, including audits of systems and assets supporting essential services
6. Human Resource Security
6.1. The professional references of key personnel (system administrators, security officers, guards, et cetera) should be validated
6.2. Training material on security issues should be provided to key personnel
6.3. Key personnel should be formally appointed in necessary security roles
6.4. Policies/procedures for the Human Resource security should be regularly reviewed and updated, taking into account any possible changes
7. Asset Management
7.1. Detailed lists of critical assets and configurations of systems supporting essential services should be regularly maintained
7.2. Policy/procedures should be put in place for asset management configuration control
7.3. The asset management policy should be regularly updated, on the basis of changes and past incidents
8. Systems Configuration

8.1. Networks and systems supporting essential services should be configured with information security in mind

8.2. The effectiveness of the security configurations to protect the integrity of systems should be regularly evaluated and reviewed

9. System Segregation

9.1. The information systems should be properly segregated in order minimize the potential consequences when risks occur

10. Traffic Filtering

10.1. A monitoring mechanism of the systems supporting essential services should be put in place

10.2. A traffic monitoring policy of the systems supporting essential services should be defined and put in place

10.3. Specific tools should be defined for supporting the traffic monitoring of the systems supporting essential services

11. Cryptography

11.1. Cryptographic mechanisms should be put in place to protect the confidentiality and integrity of information stored in or out of the company boundaries (digital facilities)

11.2. Implemented cryptographic mechanisms such as digital signatures and hashes to detect unauthorized changes to critical data at rest should be considered and defined

12. Administration Information Systems

12.1. Administration information systems should be solely used for administration purposes and not mixed up with other operations

12.2. The aforementioned resources should be managed and configured by an authorised operator

13. Authentication and Identification

13.1. Some access control mechanisms should be defined and put in place, for network and information systems, in order to allow only authorized use

13.2. All unused or no longer needed accounts should be deactivated

13.3. A mechanism should be defined and put in place for monitoring access to network and information systems and for approving exceptions and registering access violations

14. Access Rights

14.1. Access rights granted in a structured and monitored manner should be defined

14.2. The operator should define access rights to the multiple functionalities of the resource

15. IT Security Maintenance Procedure
15.1. Some procedures should be established for security maintenance in accordance with the security policy
15.2. The conditions for enabling the minimum security level for systems supporting essential services resources should be defined
15.3. Software and hardware resources should be regularly maintained and updated
16. Industrial Control Systems
16.1. Considering that the proper operation of many essential services depends on functioning and secure industrial control systems (ICS), all operators, if applicable, should be taken into account the particular security requirements for ICS

Table 2: NISD main requirements

2.2.3.2 NIS selected standards

In addition to the analysis and the guidelines already presented, a short summary of the main well-known selected international self-risk assessment/management standards and frameworks is presented in these paragraphs. This analysis can be another reference for the I-NERGY team in order to better understand all possible guidelines for the network and security requirements to follow in the project results implementation. The specific analysis describing the main features of these standards has been extracted from the ENISA document and is presented hereinafter. The standards analysed are the following:

- **ISO/IEC 27001** framework for an ISMS;
- **NIST** Special Publication 800-30 Rev. 1, Risk Management Guide for Information Technology Systems;
- **CRAMM** risk management methodology;
- **OCTAVE**, suite of tools, techniques and methods;
- **FAIR**, international standards quantitative mode;
- **IRAM2**, end-to-end approach for performing business-focused information risk assessments;

A short description of all these methodologies is presented in the following subsections.

2.2.3.2.1 ISO/IEC 27001 – ISMS

ISO/IEC 27001 is the international standard for information security management systems (**ISMS**). The ISO/IEC 27001 Standard provides a methodology which can assist OES and DSP to achieve all of their regulatory compliance objectives concerning the NIS Directive by implementing specific controls. Controls recommended by ISO/IEC 27001 are not only technological solutions but also cover people and organisational processes. There are 114 controls in Annex A covering the breadth of information security management, including areas such as physical access control, security staff awareness programmes, procedures for monitoring threats and incident management processes.

The risk assessment process established by ISO/IEC 27001 follows the below procedure:

- establish and maintain certain information security risk criteria;
- ensure that repeated risk assessments "produce consistent, valid and comparable results;
- identify risks associated with the loss of confidentiality, integrity and availability for information within the scope of the information security management system;
- identify the owners of those risks; and
- analyse and evaluate information security risks according to certain criteria.

An ISMS is based on the outcomes of a risk assessment based on the ISO/IEC 27001. OES and DSP will need to produce a set of controls so as to minimize the identified risks resulting from the aforementioned procedure.

2.2.3.2.2 NIST Special Publication 800-30

NIST Special Publication 800-30 is a foundation pillar for developing an effective and adequate risk management program. NIST 800-30 provides both the definitions and the practical guidance required for assessing and mitigating risks identified within IT systems. Additionally, it provides information on the selection of practical and profitable security controls that can be utilised to mitigate risk for the better protection of vital information and the IT systems that process this information. It is composed by well-defined and sequential steps in order to achieve the aforementioned goals, as depicted below:

- system characterisation followed by thread and vulnerability identification; control analysis and likelihood determination;
- impact analysis and risk determination; and
- control recommendations and documentation of the results.

2.2.3.2.3 CRAMM

CRAMM (CCTA Risk Analysis and Management Method) was developed in 1987 by a British government organization, the Central Communication and Telecommunication Agency (CCTA), now renamed into Cabinet Office. CRAMM can be used for all kinds of organisations, but it is especially intended for large organisations, like government bodies and industry [19]. It is in use by NATO and corporations working actively on information security. CRAMM helps in justification of security investments by demonstrating the need for action at management level, based on quantifiable results and countermeasures from the organisation.

CRAMM attempts a qualitative approach that focuses on assets. It provides 10 specific and predefined asset tables which classify the assets in categories. Those tables support the identification and valuation of assets [20]. Therefore, each asset can be classified into a specific category, each with a predefined list of known vulnerabilities and threats that can exploit them. After the completion of identification and valuation of the assets, the provided dedicated tool automatically suggests a set of all possible countermeasures. However, the usefulness of the method is largely dependent on the tool that implements it.

2.2.3.2.4 OCTAVE

OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation) was developed by the Computer Emergency Response Team within the Software Engineering Institute. The goal of the OCTAVE suite of tools, techniques and methods is to allow “risk-based information security strategic assessment and planning” [21]. OCTAVE gives the opportunity to small teams across business units and IT work together to address the security needs of the organisation and face the security challenges. It moves an organisation towards an operational risk-based view of security and addresses technology in a business context.

The methodology is divided in three explicit methods. The primary OCTAVE method forms the basis for the OCTAVE foundation of knowledge. OCTAVE-S is intended for small and medium sized organizations. The main difference with the basic method is that the necessary knowledge is assumed to be known in advance by the analysis group, so the first step of collecting knowledge is omitted. Lastly, OCTAVE-Allegro offers a faster but more limited approach that focuses on information assets. This approach covers only four simplified steps: development of risk measurement criteria, creation of profiles for each critical information asset, identification of threats to these assets and finally, analysis of resulting risks, in order to develop mitigation approaches.

2.2.3.2.5 FAIR

FAIR (Factor Analysis of Information Risk) is an international standard quantitative model for information security and operational risk and provides (a) a model for understanding, analysing and quantifying information risk in financial terms; and (b) a foundation for developing a robust approach to information risk management.

The FAIR framework defines the necessary building blocks for implementing effective risk management programs. FAIR is an ontology of the factors that contribute to risk and how they affect each other. It is primarily concerned with establishing accurate probabilities for the frequency and magnitude of data loss events.

2.2.3.2.6 IRAM2

IRAM2 (Information Risk Assessment Methodology 2) is a complete end-to-end approach for performing business-focused information risk assessments.

- simple, practical, yet rigorous risk assessment approach; focus on the business perspective;
- extended coverage of risks; and
- engagement with key stakeholders.

IRAM2 is supported by four IRAM2 Assistants, each accompanied by a practitioner guide, that help automate one or more phases of the methodology.

These methodologies are the most notable in the field of information security for risk assessment and management. A summary table is presented in the following, comparing the standard with a set of key criteria, specifically:

1. **scope/ domain**: defines the scope and the domain of applicability of the methodology;
2. **focus (RA/ RM)**: defines the focus of the methodology, i.e. risk assessment, risk management or both;
3. **flexibility**: refers to the flexibility of the methodology;
4. **complexity**: refers to the complexity of the methodology;
5. **approach**: refers to the approach of the methodology;
6. **tool Support**: defines whether there is a tool which implements the methodology;
7. **year released/ last update**: refers to the release year and the last update of the methodology; and
8. **target**: refers to the sector and/ or the types of entities that are in the scope of the methodology.

S/N	CRITERIA	ISO 27001	OCTAVE	CRAMM	FAIR	IRAM2	NIST 800-30
1	Scope/ Domain	SME / LE	LE	SME / LE	SME / LE	LE	SME / LE
2	Focus	RA / RM	RA / RM	RA	RA	RA / RM	RM
3	Flexibility	Relatively flexible	Flexible	No Flexible	Relatively flexible	Flexible	Relatively flexible
4	Complexity	Medium	Low	High	Low	Low	Low
5	Approach	Assets and control based	Risk based information security strategy	Qualitative, asset-centric approach	Quantitative approach by filling a questionnaire table	Assessment of risk from a business perspective	Risk based IT related risk management
6	Tool support	no	yes	yes	yes	yes	N/A
7	Year released /Last update	2005/2013	1999/ 2005	1985/ 2011	2001/ 2009	2014/ 2014	2000/ 2012
8	Target	All NISD sector	All NISD sector	All NISD sector	All NISD sector	All NISD sector	All NISD sector

Table 3 Criteria and methodologies

2.3 Energy domain regulatory frameworks

I-NERGY project approaches the energy sector from the point of view of digitalisation and more specifically the adoption of AI for the development of energy services concerning activities such as predictive maintenance of transmission network assets, energy demand prediction, network load forecasting, energy saving verification, DSOs asset management, consumption and flexibility prediction etc. Due to the nature of the project and the services that will be developed, the main effort for compliance with the legislative requirements falls within the fields of data protection and cybersecurity, analysed in previous sections. However, in the following paragraphs, a brief analysis of several initiatives, bodies and organisations, at International and EU level, that may provide guidance for the development of the services and raise awareness on the context of their application, is made.

2.3.1 International legislation and standards

International Energy Agency (IEA)[31] and International Renewable Energy Agency (IRENA)[32] are two intergovernmental organisations with significant contribution to the energy domain. IEA was established in 1974 initially providing response to the disruptions of oil market. Since then, IEA has expanded its role covering the entire global energy domain providing policy recommendations, data and analysis for a wide range of aspects of energy sector including among others renewable energy, energy efficiency, electricity market, energy security and clean energy technologies [22]. IRENA was founded in 2009 to promote the adoption and sustainable use of renewable energy [23]. IRENA provides a repository of policies, financial knowledge, data, statistics, studies and the latest information on renewable energy [23]. In addition, IRENA provides through International Standards and Patents in Renewable Energy (INSPIRE)[33] guidance on the use of standards and patents for the development of renewable energy technologies. Both organisations provide resources and references concerning policies, standards, codes, studies, data, statistics and analysis that may provide useful insights for the development of I-NERGY applications. For instance, financial data and analysis may contribute to I-NERGY activities concerning energy efficiency investments de-risking and energy saving verification. I-NERGY services and modules will comply with relevant standards to maximise their reliability and their compatibility. I-NERGY will also promote new work items in relevant SDOs from the very beginning and throughout all the project lifetime and thereby actively participate in work to standardise the interfaces, APIs, business models and architecture relating to I-NERGY framework. Through the memberships of many I-NERGY partners, the project will contribute to relevant standardisation activities in the respective working groups dealing with AI/big data, prosumer flexibility (SAREF4Energy, AI, blockchain, smart grid at the IT and energy domain.

2.3.2 EU legislation

The energy policy framework in the EU underwent a significant update through the Clean Energy for all Europeans Package (CEP), facilitating a clean energy transition and contributing to the Energy Union Strategy [24]. It is meaningful to examine very briefly the legislative acts that are part

[31] https://www.iea.org/

[32] https://www.irena.org/

[33] http://inspire.irena.org/Pages/home.aspx

of CEP since they regulate areas of energy domain that are also related to the context of services developed within the project.

The initial proposal for CEP was published by EC in 2016 and the publication of final texts was completed in 2019 following the agreement of the Council and European Parliament [25]. CEP contains eight legislative acts that concern the following areas of the energy domain:

- **Energy performance in buildings**

Improving the energy performance of buildings is of outmost importance for EU's energy and environmental goals since the energy consumption and the CO2 emissions coming from buildings are estimated at 40% and 36%, respectively of total energy consumption and CO2 emissions in the EU [25]. The relevant legislative framework includes the Energy Performance of Buildings Directive 2010/31/EU (EPBD) [26], which was amended in 2018 by the Directive (2018/844/EU) [27]. The EPBD provides a wide range of measures and policies that will help EU countries to improve the energy efficiency of their buildings, concerning also among others the setting of minimum energy performance requirements for buildings, energy performance certificates, buildings requirements for EVs and smart technologies [28].

- **Renewable energy**

Energy from renewable resources is vital to reduce greenhouse gas emissions and tackle climate change, protect the environment and to contribute to the vision for a climate neutral Europe [29]. CEP includes the recast Renewable Energy Directive (2018/2001/EU) [30], which entered in December 2018 establishing the framework for the promotion of renewable forms of energy, setting the EU's binding target regarding renewable energy in 2030 and describing the actions towards this direction [31]. The target for energy coming from renewable sources is set at 32% in 2030. The Directive includes among others provisions regarding support schemes for promoting renewable energy, financial support for electricity coming from RES, cooperation mechanisms between EU and non-EU countries, renewable energy in heating, cooling and transport sector [31]. In addition, recasting of Directive brings provisions about self-consumers and their rights to generation of renewable energy, storage and sale of surplus energy as well as renewable energy communities, their rights and a framework facilitating their development [31].

- **Energy efficiency**

Energy efficiency is central to the energy union strategy. Prior to CEP, the energy efficient Directive (2012/27/EU) [32] has set the target of improving energy efficiency by 20% by 2020 in comparison with 1990 levels [33]. This Directive provides a set of measures contributing to energy efficiency covering all the stages of energy chain from generation to consumption. In 2018 the Directive was amended by Directive (2018/2002/EU) [34] updating the existing policy framework and setting the goal of a 32.5% improvement in energy efficiency by 2030 [35]. In addition, the new Directive brought several amendments including stricter rules for energy metering and billing, requirements for utility companies to assist consumers in energy saving, requirements for EU countries to define their contributions to energy efficiency for the decade 2020-2030 and to have clear national rules for allocation of costs relating to heating, cooling and hot water when these services are shared. [35]

- **Governance regulation**

Regulation (EU) 2018/1999 [36] on the governance of the Energy Union and Climate Action entered into force on 24 December 2018 as part of CEP aiming to ensure the implementation of the Energy Union Strategy and the achievement of the targets set in the EU policy framework for climate and energy (2020 to 2030) and Paris Agreement on climate change [37]. It describes how EC and EU countries should work together and on their own to contribute to the Energy Union's goals. According to the Regulation, EU countries should establish a 10-year integrated National Energy and Climate Plan (NECP) starting from 2021-2030. For this period, the NECPs should have been submitted till the end of 2019. The regulation also requires that Member Statesreport to EC regarding the progress of implementation of their NECPs from 2021 and provides specific requirements for this progress reporting. In addition, the Regulation repealed the existing monitoring and reporting mechanism for GHG emissions (Regulation (EU) No 525/2013).

- **Electricity market design**

CEP brings the necessary updates to the legislative framework in the Electricity market to meet the new requirements, including the adoption of new technologies, the increase of energy coming from RES and the need for a more flexible market [25].

CEP includes the Directive (EU) 2019/944 [38] on common rules for the internal market for electricity which revises and replaces Directive 2009/72/EC as of 1 January 2021. As described in Article 1, this Directive contains the rules that are pertinent to a wide range of electricity market activities including the "generation, transmission, distribution, energy storage and supply of electricity, together with consumer protection provisions, with a view to creating truly integrated competitive, consumer-centred, flexible, fair and transparent electricity markets in the Union" [39]. Therefore, these rules also describe the obligations of DSOs, TSOs and aggregators. In addition, this Directive recognises certain categories of citizen energy initiatives at the Union level as 'citizen energy communities'[34] including provisions about their rights and obligations as well as the requirements for the Member States to provide an appropriate enabling framework[35].

Another legal act included in CEP is the Regulation (EU) 2019/943 on the internal electricity market [40], which is applicable since 1 January 2020 and describes the rules and the principles that the operation of the electricity market should respect [41]. The regulation covers, among others, the network access and the congestion management describing also the conditions and requirements for capacity mechanisms [41].

Regulation (EU) 2019/941 [42] on risk-preparedness in the electricity sector is another piece of legislation included in CEP and is applicable since 4 July 2019. It contributes to the identification of possible electricity crises and the assessment of the related risks, to the establishment of risk-preparedness plans and to the management and overall handling of such situations based on common methodologies and cooperation between EU countries [43].

The Agency for the Cooperation of Energy Regulators (ACER)[36] was established in March 2011 by the Third Energy Package legislation [44]. CEP updated the role of ACER through Regulation (EU) 2019/942 [45] introducing additional tasks and considering also the new energy market design introduced [46]. ACER mission is "to achieve a transition of the European energy system in line with political objectives set, reaping benefits of increased energy market integration across Europe, and

[34] See Article 2 point 11 for definition

[35] See Article 16

[36] Available at: https://www.acer.europa.eu/

securing low-carbon supply at least possible cost for European businesses and citizens" [47]. ACER responsibilities include coordinating and assisting regulatory authorities with their tasks, developing common network and market rules, monitoring the markets of electricity and natural gas, coordinating and overseeing the regional coordination centres, issuing opinions and recommendations to TSOs and DSOs [46], [48]. ACER works also closely with the Council of European Energy Regulators (CEER)[37], which aims to facilitate the development of a single and efficient Internal Energy Market [49]. CEER fosters cooperation as well as information and best practice exchange between European national energy regulators and strengthens their presence at EU and international level [49].

2.4 Regulation and Guidelines for Ethical AI in Europe

The advancement of AI is accompanied by great opportunities for economic development and addressing societal challenges. The potential of AI is expected to radically transform the energy sector, revolutionise the way that Electric Power and Energy Systems (EPES) community is undertaking the business processes and have a significant impact on society and environment. However, AI can put pressure on ethical values and fundamental rights that drive our lives and our societies. Ethical considerations often underlie the law and represent the rationale and thus overlap with law to a certain degree. Ethical considerations could also serve as guidance where the law is not entirely adapted to new phenomena, e.g., where technology enables practices, which the legislator had not anticipated. More specifically, in newly emerging fields such as AI, legislation may not cover sufficiently all ethical implications or may have no clear rules on them, hampering compliance. Hence, I-NERGY partners should ensure that, in addition to respecting legal obligations are guided by ethical considerations and the values and principles on which the EU is founded. In this context, it is necessary for the development of relevant AI systems to be in line with ethical principles and requirements, preventing any harmful implications. In the same direction, it is crucial to identify all possible ethical issues and implications, within the I-NERGY project, thus mitigating the associated risks and maximising project's trustworthiness, impact and sustainability. To achieve this, I-NERGY examines and continuously monitors key initiatives and ongoing efforts around ethics in AI, leveraging their results to guide its internal ethical development and assessment process that will be described in the following section. Those initiatives are briefly described in the following paragraphs.

2.4.1 Ethics Guidelines for Trustworthy AI

With trust being a prerequisite for human-centered AI, EC set up the High-Level Expert group on AI (AI HLEG) in June 2018 to provide advice on its Strategy [50]. The AI HLEG prepared and published in 2019 the Ethics Guidelines for Trustworthy AI, where key concepts and requirements of Trustworthy AI are prescribed, and EC highlighted these requirements through its Communication on "Building Trust in Human Centric Artificial Intelligence" [51]. In parallel to HLEG, the European AI Alliance[38] was established, bringing together multiple stakeholders for an open discussion on AI, including its impacts.

[37] Available at: https://www.ceer.eu/

[38] Available at: https://digital-strategy.ec.europa.eu/en/policies/european-ai-alliance

At first, it is meaningful to examine the term “Trustworthy AI”. According to AI HLEG, AI systems should meet the following conditions to be deemed Trustworthy [52, p. 5]:

- Be in line with all applicable laws and regulations. ***Lawful***.
- Comply with ethical principles and values. ***Ethical***.
- Be robust concerning both a technical and social perspective. ***Robust***.

AI HLEG puts fundamental rights, enshrined in EU Treaties, EU Charter for fundamental Rights, ECHR and other international human rights law, at the center of a trustworthy approach of AI. These rights contribute to the lawful dimension of AI systems as they are legally binding, but they also form the basis for the ethical principles and guidelines that AI systems should follow [52, pp. 9-10]. AI HLEG describes the categories of fundamental rights that are suitable for AI systems. The reflection of these rights will raise awareness within I-NERGY about the aspects that should be considered when assessing an AI system from an ethical perspective. Therefore, these rights, described in detail in AI HLEG Guidelines [52, pp. 10-11], are listed below:

- **Respect for human dignity**
- **Freedom of the individual**
- **Respect for democracy, justice and the rule of law**
- **Equality, non-discrimination and solidarity**
- **Citizens’ rights**

Grounded on fundamental rights, Ethics Guidelines for Trustworthy AI lists the four principles, that AI systems, and therefore AI solutions within I-NERGY must respect in order for their development and operation to be deemed trustworthy. These principles, as described in the Guidelines [52, pp. 11-13], are briefly listed as follows:

- **The principle of respect for human autonomy**. AI systems must not affect the freedom and autonomy of human beings and must ensure that humans are “able to keep full and effective self-determination over themselves and be able to partake in the democratic process” [52, p. 12].
- **The principle of prevention of harm**. AI systems must not cause any harm or pose any negative impact to humans. Instead, they must ensure the protection of physical and mental integrity and human dignity.
- **The principle of fairness**. AI systems must be developed and operated in a fair manner. According to authors, fairness has a substantive dimension which “implies a commitment to: ensuring equal and just distribution of both benefits and costs, and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatisation” [52, p. 12] and a procedural dimension which “entails the ability to contest and seek effective redress against decisions made by AI systems and by the humans operating them” [52, p. 13].
- **The principle of explicability**. The purposes of development of AI systems, the way these systems operate, how they make decisions and what they are able to do should be transparent, explainable and well communicated to those affected. In this manner, these systems can gain and maintain users’ trust.

2.4.1.1 Requirements for Trustworthy AI

The requirements for trustworthy AI, as provided by the HLEG Ethics Guidelines for Trustworthy AI, constitute a non-exhaustive list of seven key equally weighted factors that should be considered through the implementation stage of AI systems. These requirements derive from the previously mentioned 4 ethical principles and come to put them in practice within the implementation stage of an AI system. The 7 requirements are described in detail in the Guidelines [52, pp. 14-18], and are briefly summarised as follows.

- ***Human agency and oversight***

"AI systems should support human autonomy and decision-making" [52, p. 15] as well as foster fundamental rights and social values such as democracy and equality. When AI systems may pose threats to fundamental rights it is important to assess this impact prior to their development. Concerning the aspect of human agency, AI systems should provide information and tools that enable users to understand their mechanisms and interact with them, in order to be able to take informed decisions. Human oversight is essential to ensure human autonomy in AI systems. Degree of human oversight may differ depending on the application area and the potential risks. To achieve human oversight, appropriate government mechanisms should be defined. Examples of these mechanisms are the human-in-the-loop (HITL), human-on-the-loop (HOTL), or human-in-command (HIC) approaches.

- ***Technical robustness and safety***

AI systems must be technically robust to prevent any unintentional or unexpected harm and ensure physical and mental integrity of humans. One aspect of that robustness is "resilience to attack and security". AI systems should be safeguarded against vulnerabilities as they pose threats for the security and operation of the systems. Moreover, situations that AI systems may be potentially misused by malicious actors should be considered to prevent and limit the negative consequences of such cases. Another aspect of robust systems is the "fallback plan and general safety". It is crucial for AI systems to prescribe fallback plans that will be used in case of problems. In addition, processes towards identifying the potential risks of the application of AI systems should be set up and depending on their findings appropriate measures may be necessary to be developed and tested. Accuracy of AI systems should be also taken into account during the development and evaluation of AI systems, as inaccurate results can have adverse effects. Moreover, it is important for AI systems to operate properly under different circumstances and to be reliable and reproducible.

- ***Privacy and data governance***

Adequate data governance is crucial for ensuring privacy. AI systems must ensure data protection and privacy throughout their lifecycle. Information either collected by users or generated by system must be protected. Systems must also guarantee that the data collected will be used lawfully and fairly. Data integrity and quality must also be ensured as collected datasets may contain biases or inaccuracies that must be addressed before the training phase of models to prevent their further reproduction. Access policy to data must also be clearly defined and implemented. In this manner, only the defined personnel and under specific conditions should have access to data.

- ***Transparency***

It is important for AI systems to be characterised by traceability. This means that datasets, algorithms, processes used for decision making by AI systems should be documented to the greatest possible degree. This enables the identification of reasons for erroneous decisions and therefore possibly their correction. AI systems should also concern the aspect of explainability. This means that processes of AI systems and decision making can be understood by humans. In addition, users should be able to know when they are interacting with AI systems and what capabilities and limitations these systems have.

- ***Diversity, non-discrimination and fairness***

In order for AI systems to be deemed trustworthy, they should foster inclusion and diversity. Datasets used in AI systems may include unfair biases that may lead to unintended prejudice and discrimination. These biases should be addressed in the collection phase. In addition, biases can be introduced through the development of AI algorithms. To mitigate this risk appropriate oversight processes and diversity of opinions should be considered. AI solutions should be designed in a manner allowing all people use their services or products and providing accessibility for person with disabilities concerning the relevant standards and design principles. These approaches foster equal access and participation of all people to the benefits of technology. Furthermore, the engagement and participation of relevant stakeholders helps the development of trustworthy systems.

- ***Societal and environmental well-being***

AI systems should be designed, developed and operate in the most environmentally friendly manner and measures towards this direction are encouraged. In addition, the social impact and the effects that AI systems may have on people's mental and physical health should be considered and monitored. Attention should also be paid to the possible impacts of AI systems on society and democracy.

- ***Accountability***

Appropriate mechanisms should be set up to ensure responsibility and accountability of AI systems. Auditability is related to the assessment of AI systems and in combination with evaluation reports can contribute to their trustworthiness. It is also crucial to identify, report and minimise the negative impacts, and impact assessments can be useful for this effort. In case that tensions arise between the requirements appropriate trade-offs should be considered. These trade-offs should be documented, and their impacts should be evaluated, especially regarding ethical principles. In case these trade-offs violate any ethical principles, the development of such systems must be prohibited. In addition, appropriate mechanisms that would allow affected parties to obtain redress should be prescribed.

2.4.1.2 Technical and non-technical methods

To further help the realisation of Trustworthy AI, the Guidelines suggest a set of technical and non-technical methods for implementing the defined requirements. The technical methods as described in Guidelines [52, pp. 21-22] include the following:

- *Architectures for Trustworthy AI.* Architecture should prescribe appropriate processes and rules that define the acceptable and restricted behaviours as well as processes that monitor compliance with these rules and restrictions.

- *Ethics and rule of law by design.* AI systems should follow values-by-design approach and therefore compliance with the ethics requirements should be implemented starting from the design phase. This includes for example appropriate measures to safeguard the data, the outcomes and to prevent potential risks and attacks.
- *Explanation methods.* It is very important for the AI systems to be explainable. Therefore, methods towards this direction are meaningful for users to understand the systems behaviour and for systems to increase their reliability.
- *Testing and validating.* The nature of AI systems requires extensive testing and validation throughout their whole lifecycle, including training, deployment and operation. While testing and validating, it must be ensured that the outcomes of the system are consistent with the given input and the defined requirements. In addition, all components of AI systems should take part in testing and validation. It is also beneficial, to develop multiple metrics and test the system from different perspectives and by diverse teams.
- *Quality of Service Indicators.* Setting quality of service indicators for AI systems can contribute to evaluation of systems development and testing from different perspectives such as security, functionality, usability, etc.

Besides the technical methods, the Guidelines suggest a set of non-technical methods that contribute to the trustworthiness of AI and will be considered to the extent that they can apply. These, as described in the Guidelines [52, pp. 22-23] include:

- *Regulation*. AI system should comply with existing regulations and legislative frameworks.
- *Code of conducts*. Organisations can consult the Guidelines to update accordingly their policies and their codes of conduct in an effort to realise Trustworthy AI.
- *Standardisation*. Different standards provide valuable information, rules and guidance and can act as quality management criteria.
- *Certification*. Organisations that would be able to certify that AI systems are compatible with requirements such as transparency might be beneficial for better information for the public.
- *Accountability via governance frameworks*. AI systems should be accompanied by appropriate government mechanisms that help to ensure accountability regarding ethics.
- *Education and awareness to foster an ethical mind-set*. It is beneficial for all stakeholders to be well informed, educated and trained around capabilities and impact of AI and how they can participate to the evolvement of society.
- *Stakeholder participation and social dialogue*. It is important to involve stakeholders and general public to discussions around AI in order to ensure equal access to its benefits.

- *Diversity and inclusive design teams*. Involvement of diverse teams in the development cycle of AI systems is significant and can contribute to the development of realistic and objective systems.

2.4.1.3 Assessment list for Trustworthy AI

AI HLEG presented on 17 July 2020 the final Assessment list for Trustworthy AI (ALTAI) [53]. This list makes ethics central to the development of AI systems. It acts as a self-evaluation tool for assessing AI systems under the key requirements defined in the Guidelines. [53] The list contains a set of questions relevant to the requirements that provide guidance for their practical implementation. In addition, this list raises awareness around the potential impact and risks of the proposed AI systems and the kind of measures that can be taken to mitigate these risks. The Trustworthy AI assessment list comes to seal the process of Trustworthy AI by enabling the inspection and validation of the final AI system, hence completing the lifecycle of Trustworth AI development as shown in Figure 1.

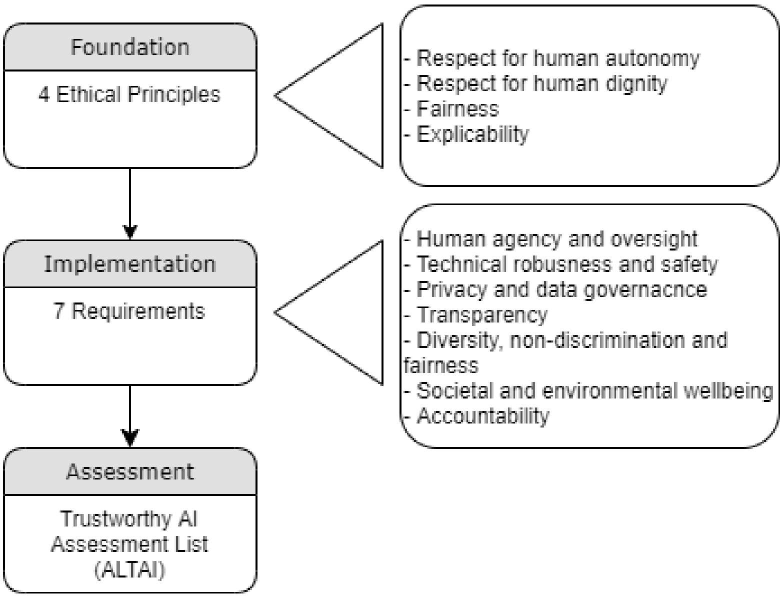
Foundation
4 Ethical Principles
- Respect for human autonomy
- Respect for human dignity
- Fairness
- Explicability
Implementation
7 Requirements
- Human agency and oversight
- Technical robusness and safety
- Privacy and data governacnce
- Transparency
- Diversity, non-discrimination and fairness
- Societal and environmental wellbeing
- Accountability
Assessment
Trustworthy AI Assessment List (ALTAI)

Figure 1 - The Trustworthy AI Framework as established by HLEG

2.4.2 Artificial Intelligence ACT

On 21 April 2021, the European Commission unveiled the draft AI Act, a new proposal for an EU regulatory framework on artificial intelligence. The draft act signals the European Commission's shift from a **soft-law** approach, as indicated by the publication of its non-binding Ethics Guidelines for Trustworthy AI [54] and Policy investment recommendations [55] towards a **legislative** one [56]: The draft AI act is the first ever attempt to enact horizontal regulation of AI, applicable to all AI systems, developed, placed on the market or used in the Union, establishes a technology-neutral definition of the latter in EU law, and lays down for them a classification with different requirements and obligations tailored on a **'risk-based approach'**, whereby legal intervention is dependent upon the concrete level of risk [57]. In particular, the draft AI act distinguishes between AI systems posing

(i) **unacceptable risk**, (ii) **high risk**, (iii) **limited risk**, and (iv) **low or minimal risk**. Under this approach, AI applications would be regulated only as strictly necessary to address specific levels of risk[39].

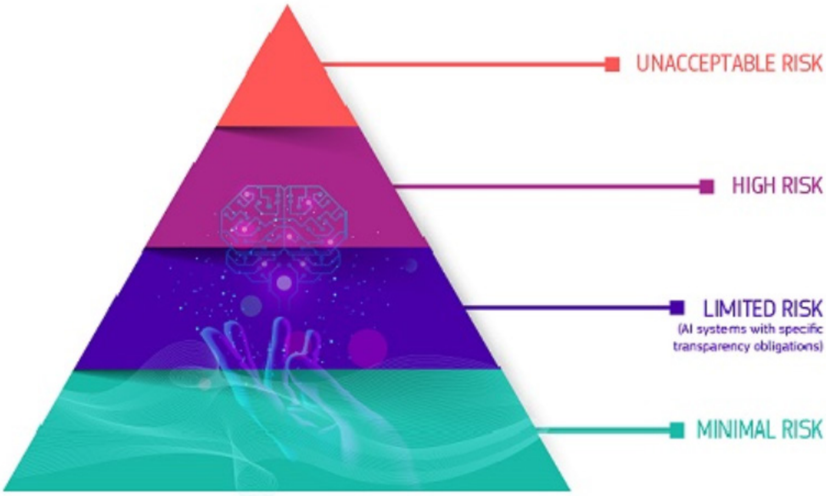

Figure 2 - AI Act risk-based approach [58]

Along the above lines, and due to the *'unacceptable risk'* they create, the act (Title II, Article 5) explicitly prohibits the placing on the market, putting into service or use of AI systems that:

- use 'subliminal techniques' to manipulate a person's behavior in a manner that may cause psychological or physical harm;
- exploit specific vulnerable (in terms of physical or mental disability) groups;
- serve social scoring purposes by public authorities;
- enable 'real-time' remote biometric identification in publicly accessible spaces for law enforcement purposes with the exception of certain time-limited public safety scenarios.

The AI act (Title III, Article 6) further regulates *high-risk* AI systems that create an adverse impact on people's health and safety or their fundamental rights. These are AI systems:

- used as a safety component of a product or as a product falling under Union health and safety harmonisation legislation (e.g. toys, aviation, cars, medical devices, lifts) and hence subject to third party ex-ante conformity assessment;
- deployed in the following eight specific areas (Annex III), which the European Commission may update as necessary via a delegated act (Article 7):
    - biometric identification and categorisation of natural persons;
    - critical infrastructure management and operation, that could put the life and health of citizens at risk;
    - education and vocational training, that may determine the access to education and professional course of someone's life;
    - employment, worker management and access to self-employment;
    - access to and enjoyment of essential public and private services;
    - law enforcement that may interfere with people's fundamental rights;
    - migration, asylum and border control management;

[39] See impact assessment at pp. 48-49. A risk approach is also adopted in the United States Algorithmic Accountability Act of 2019 and in the 2019 Canadian Directive on Automated Decision-Making.

- administration of justice and democratic processes.

As mentioned above, the act mandates an ex-ante conformity assessment for high-risk AI systems. AI products and services governed by existing product safety legislation, will fall under the existing third-party conformity assessment structures and regulatory frameworks that already apply. Providers of AI systems that are not currently governed by explicit regulatory frameworks are obliged to conduct their own conformity assessment (self-assessment) and register their systems in an EU database managed by the Commission before placing them on the market or putting them into services (Title VII). Proportionate obligations are also placed on users and other participants across the AI value chain (e.g., importers, distributors, authorised representatives).

Additional technical and auditing requirements for high-risk AI systems (Articles 8 to 15) concern:

- **risk management:** creating and maintaining a risk management system for the entire lifecycle of the system.
- **data and data governance**: establishing appropriate data governance controls, including the requirement that all training, validation, and testing datasets be complete, error-free, and representative.
- **technical documentation**: production of detailed technical documentation, including around system architecture, algorithmic design, and model specification.
- **record-keeping**: automatic logging of events while the system is running, with the recording conforming to recognised standards.
- **transparency and provision of information to users**: system designed with sufficient transparency to allow users to interpret its output.
- **human oversight** system designed to maintain human oversight at all times and prevent or minimise risks to health and safety or fundamental rights, including an override or off-switch capability.
- **accuracy, robustness and cybersecurity**: system designed and developed in such a way that it performs consistently in these respects throughout its lifecycle.

Certain *limited-risk* AI systems are covered by the Act (Title IV) under transparency requirements. Systems that interact with humans, emotion recognition or biometric categorisation systems, and systems that generate or manipulate image, audio or video content that resembles authentic content ('deep fakes') should respectively notify users that they are interacting with an AI system, that their emotions or characteristics are recognised or that the content is generated through automated means. This would not apply if it is "obvious from the circumstances and the context of use" or in case of legitimate purposes (law enforcement, freedom of expression).

All other AI systems presenting only *low or minimal risk* (*non-high-risk* AI systems) are not subject to legal obligations. However, the AI Act proposal envisages the creation of codes of conduct to encourage providers of non-high-risk AI systems to apply voluntarily the mandatory requirements for high-risk AI systems. Those codes may include as well voluntary commitments indicatively related to environmental sustainability, accessibility for persons with disability, stakeholder participation in the design and development of AI systems, and diversity of development teams (Title IV).

In the direction of creating a legal framework that is innovation-friendly, future-proof and resilient to disruption, the draft act encourages national competent authorities to set up regulatory sandboxes, i.e. controlled environments to facilitate the development, testing and validation of innovative AI systems for a limited time before they are put on the market (Title V).

In terms of governance and implementation mechanisms (Title VI), at EU level, the act designates the establishment of a *European Artificial Intelligence Board*, composed of representatives from the Member States and the Commission, to facilitate effective and harmonised implementation of the regulation and to ensure cooperation between the national supervisory authorities and the Commission. At national level, it foresees that Member States designate one or more competent authorities, including a national supervisory authority, which will be tasked with supervising the application and implementation of the regulation.

Market surveillance authorities should further be responsible for assessing compliance with the obligations and requirements for all high-risk AI systems already placed on the market (Title VIII, Chapter 3), taking corrective measures to prohibit, restrict, withdraw or recall AI systems that do not comply with the AI act requirements or that, although compliant, present a risk to health or safety of persons or to fundamental rights or other public interest protection. Non-compliance with the AI act is anticipated to draw administrative fines of up to €30M or 6 percent of the total worldwide annual turnover (whichever is higher), depending on the severity of the infringement (Title X, Article 71).

2.4.3 AI4EU

Significant contribution to the field of ethics on AI is also provided by the AI4EU project. One of the strategic objectives of the AI4EU project is to promote European values for Ethical, Legal, Socio-Economic and Cultural (ELSEC) issues in AI (Figure 1). Societal concerns over the use and misuse of AI are addressed with the organisation of an Ethical, Legal, Socio-Economical and Gender-Aware observatory, providing the AI community as well as European and national authorities with detailed, accurate and up to date information regarding the consequences of use and misuse of AI. While AI technologies can provide great benefit for European Society, misuse can pose grave risks. To protect European Society from abuse of AI, AI4EU has created the AI4EU Ethical Observatory working to assure respect of European values and to assure that respect for these values provides an important competitive advantage both within the EU and in larger international markets.

The AI4EU project also developed an abbreviated assessment framework, based on the Assessment List for Trustworthy AI (ALTAI), developed by the European High-Level Expert Group on AI. This abbreviated assessment list is mainly meant to assess the AI applications shared through the AI4EU catalogue, but it also can support organisations perform a 'quick scan' of their AI systems. This list can be used as a self-assessment tool to quickly identify the relevant elements of trustworthy AI and the level of compliance to these elements. It will help determine the level of impact of the AI applications and provide options to balance different tensions and interests.

2.4.4 International Outreach for Human-Centric Artificial Intelligence Project (InTouchAI.eu)

The EC's Service for Foreign Policy Instruments (FPI) and the Directorate General for Communications Networks, Content and Technology (DG CONNECT), in collaboration with the European External Action Services (EEAS), launched a large foreign policy instrument project namely International Outreach for Human-Centric Artificial Intelligence project (InTouchAI.eu) to engage with international partners on regulatory and ethical matters and promote the responsible development of trustworthy AI at global level. The project aims to support the European Commission in setting up a framework for ethics and trust to enable the growth of AI. This is in accordance with EU values, and should prepare the ground for global coalition building in this field.

To achieve this goal, the specific objectives of the project are to support the Commission to:

- Develop responsible leadership in global discussions around AI;
- Create the conditions for the uptake of policies and good practices and standards that ensure an appropriate ethical and legal framework on AI;
- Enhance public awareness on the challenges and opportunities linked to AI.

The initiative organises activities relating to:

- Dialogue and joint initiatives with like-minded partners:
- Public outreach and technology diplomacy
- Research, intelligence gathering, and monitoring of AI developments

2.4.5 Project SHERPA – An Initiative Focused on the Smart Grid and the Energy Sector

At this point, it is important to define the famous concept of the Energy Trilemma. According to the World Energy Council's[40] definition, energy sustainability is based on three core dimensions: Energy Security, Energy Equity, and Environmental Sustainability of Energy Systems. Balancing these three goals constitutes a 'Trilemma' and well-balanced systems enable prosperity and competitiveness of individual countries. The efforts towards the "resolution" of the Energy Trilemma comprise—to a greater or lesser extent—the main cause of today's technological growth along with a series of ethical issues and implications that arise for an optimal equilibrium amongst its three core factors.

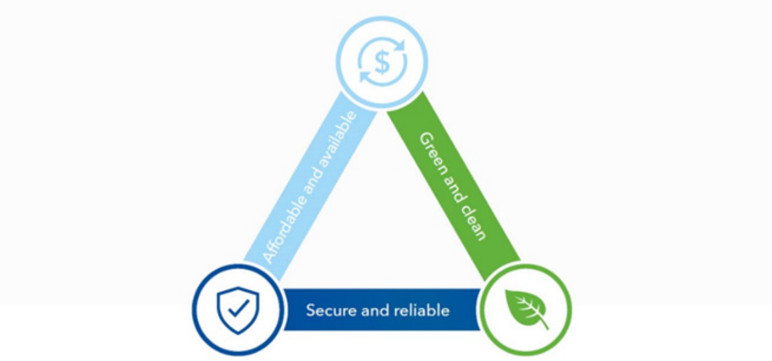

Figure 3 The Energy Trilemma[41]

Although AI systems can help solve the Energy Trilemma, it is of uttermost importance that the related issues that emerge from an ethical perspective are directly and responsibly addressed. However, such issues have been significantly underreported and they represent a very little fraction of today's research interest and publications. In the same context, consumer research that has

[40] https://www.worldenergy.org/

[41] Source: https://safety4sea.com/gas-to-account-29-of-world-energy-mix-by-2050/

taken place relies mostly on pilots with interested parties and looks at the response and use of such technologies, from a functional rather than an ethical perspective.

Relatively to this trilemma, the H2020 project SHERPA[42] has investigated and analysed the ways in which smart information systems (SIS; the combination of AI and big data analytics) impact ethics and human rights issues. Amongst other results, a case study [59] has been published, within the context of this project, analysing the principal ethical issues that occur in the use of SIS in electricity grids. These ethical issues are described as follows.

2.4.5.1 Privacy and informed consent

AI services rely on the collection and processing of granular data on household energy usage via smart meters. In this context, newcomers in the energy market, such as aggregators, propose customer engagement programs involving thorough household data monitoring with the purpose of dynamic consumption advice in order to maximise energy savings and efficiency. Three serious issues arise here:

- Such data can reveal information about people's private lives within their homes. In addition, AI algorithms can often reveal patterns and habits that even a "malicious" human being would be unable to trace. Therefore, how should these issues regarding in-home surveillance and consumers' privacy interests can be effectively addressed both at technical and policy level issues?
- Will poor families truly have the option to opt out from their consent in such programs if this is the only way to gain access to affordable energy or will they indirectly be forced to do it?
- Will those that live in shared privates have the chance to individually opt out from giving their consent irrespective given that the landlord is the controller of the energy supply?

2.4.5.2 Energy Security

To date, cyber-attacks on the energy grid have been sparse, albeit raising significant concerns as the use of sensors and IoT networks renders the grid vulnerable to them, leading to the danger of disruptions to the distribution of energy and even to the infrastructure itself. As electrical energy is fundamental for modern living, such disruptions can directly affect people's wellbeing. Cyberattacks on the energy grid have serious economic implications for citizens at a national level, while according to ENISA[43] , global losses have been estimated to reach the amount of 1.69 billion euros in 2018 [60].

2.4.5.3 Energy Equity and Affordability

While smart grids are seen as one of the solutions to effecting energy justice or equity, their main focus is energy abundance so that there is enough supply to cover the disproportionate increases in energy demand [61]. Nevertheless, AI-based smart grids raise significant ethical issues around energy justice:

[42] Available at: https://www.project-sherpa.eu/

[43] Available at: https://www.enisa.europa.eu/

- Affluent consumers, who can afford cutting edge technology equipment (e.g., EVs) will be able to benefit from AI and smart grids much sooner and at a larger scale, while the cost of the energy grid is funded via taxation and hence shared between citizens.
- In terms of electricity distribution, AI algorithms deployed in smart energy grids can become biased against small, albeit essential, loads (e.g., washing machine) of poor households in favor of affluent consumers whose large loads (e.g., EV charging) are more likely to influence the model training in a significant manner.
- Dynamic and incentive-based pricing supported by AI flexibility forecasting algorithms is being designed to minimise energy spendings. However, it is doubted that poor citizens will be able to profit from it, given that their energy consumption numbers are already very low leading to very small profit margins, while the cost of energy saving services can be unaffordable for them.

Therefore, it becomes obvious that instead of establishing energy equity and affordability, AI-based energy applications can nurture inequalities leading to further reinforcement of the blight of energy poverty[44] contrary to the missions of the EU.

2.4.5.4 Sustainability

Regarding sustainability, AI technologies and smart devices are the core components of smart grids forming a major part of the EU's decarbonisation strategy[45] as real-time grid modelling and forecasts allow for more precise management of electricity flows, leaving room for higher penetration of renewables. Nonetheless, smart grid equipment (smart meters, servers and networks) makes intense use of electricity to function, while AI models specifically require huge volumes of storage along with high amounts of processing power leading to considerable electricity consumption and thus GHG emissions [62]. Additionally, several studies such as [63] have demonstrated that, from an LCA [64] perspective, the introduction of smart devices in the grid is not always beneficial for the reduction of GHG emissions.

2.4.6 Other Initiatives

EC through its "White Paper on AI – A European approach to excellence and trust" expressed the need for a European ecosystem of excellence and trust [65]. The Big Data Value Association/Data, AI and Robotics (BDVA/DAIRO)[46] expressed through its position paper, its alignment with the vision of EC for this ecosystem and provided its feedback on the identified issues and the policy and regulation options discussed [66]. This paper was part of a wider effort of BDVA/DAIRO Task Force 5 which aims to "structure the debate and members' opinions on a number of relevant issues in the domain Policy & Societal implications of Big Data–driven innovations" [67].

Another important initiative is the establishment of the Ad Hoc Committee on Artificial Intelligence (CAHAI) in September 2019 by the Committee of Ministers of the Council of Europe. CAHAI will examine the feasibility of a legal framework for the adoption of AI, based on "Council of Europe's standards on human rights, democracy and the rule of law" [68]. Useful references for guidelines,

[44] Available at: https://www.energypoverty.eu/

[45] Available at: https://ec.europa.eu/clima/policies/strategies/2030_en

[46] Available at: https://www.bdva.eu/

recommendations, policy, initiatives and other legal instruments on AI issued by CoE bodies and committees can be found on CoE's webpage[47].

[47] Available at: https://www.coe.int/en/web/artificial-intelligence/work-in-progress

3 Trustworthy AI Framework in I-NERGY

In addition to respecting legal obligations, the project's results need to be guided by the ethical considerations, the values, and the principles on which the EU is founded. In this context, it is necessary for the development of relevant AI systems to be in line with ethical principles and requirements, preventing any harmful implications. In the same direction, it is crucial to address all possible ethical issues and implications, within the I-NERGY project, thus mitigating the associated risks and maximising project's trustworthiness, impact and sustainability. To achieve that, throughout the progress of the project, the I-NERGY consortium participates in various activities and initiatives relating to Trustworthy AI. Moreover, specific measures for Trustworthy AI have already been adopted by the consortium aiming at a coordinated and effective effort amongst partners. Last but not least, a methodological procedure has been adopted for trustworthy AI assessment —mainly based on the HLEG AI Guidelines for Trustworthy AI— that will help partners and Open Call (OC) projects to carefully examine the AI solutions that are being developed and to identify and face potential ethical issues that may arise. This procedure has been shared with the ICT-49 projects with the goal to form the basis of the trustworthy AI assessment for solutions uploaded to AIOD platform.

3.1 Action for Trustworthy AI in I-NERGY

The I-NERGY consortium has actively participated in numerous events and initiatives related to the ethical and trustworthy AI concepts. In this context, some of these activities along with their main takeaways are presented as follows:

- **Participation in ICT-49 Trustworthy AI Working Group:** The ICT-49 projects have established an ICT-49 Trustworthy AI WG (TAI WG) to enhance collaboration regarding the approach and the activities of AIOD in the realm of Trustworthy AI. The TAI WG has a short-term objective to define a common methodology for the trustworthy assessment of proposals from the OC and for the AIoD's platform assets and provide feedback to the EU Trustworthy AI requirements (HLEG). The long-term objective is to improve the "L-service" taxonomy and the strategy to integrate them with the AIOD platform. I-NERGY actively participated in all meetings of the WG and shared its approach for the trustworthy AI assessment in the context of the project and the Open Calls. As of today, the TAI WG has drafted a initial action plan in which the definition of the first version of the common methodology for assessing the trustworthiness of AI assets for the Ocs has been assigned to the I-NERGY project. The draft plan is included in Annex I of this document.

- **Participation in "European AI Excellence and Trust in the World":** I-NERGY was represented by its project manager Dr. Spiros Mouzakitis at the event "European AI Excellence and Trust in the World" [69]. The event hosted workshops on AI Ethics and specifically AI for Sustainability. The workshop was organised by the InTouchAI.eu project that has been described in section 2.4.4. The key takeaways from this event are the following [70]:
  - Considering the potential of AI in fighting the climate crisis (AI for sustainability) together with its environmental impacts (sustainability of AI) is crucial to ensure a positive net effect.

- Prioritisation of non-technical solutions over AI-based ones should be key for social and sustainability challenges to avoid technology push that is not strictly required.
- The human-centric approach to AI should account for vulnerable and marginalised groups that are affected most by the climate crisis, as well as it should be widened to a planet-centric approach.
- Tools and methods to measure the environmental (and social) impact of AI should be developed to increase transparency, and Public-Private Partnerships should be further harnessed in this field.
- Europe needs more investments to reduce market and research fragmentation and achieve bigger goals for global-level climate challenges, leveraging the excellence of EU academia and industry.
- The concept of sustainable AI should be embedded into the AI Act to help creating the European market and foster the global leadership role Europe can play in AI for Sustainability.
- Global solutions to enhance data sharing and effective AI solutions to tackle climate change should be developed, with the possibility of data sources to be accessed remotely to foster a holistic approach to SDGs that engage also developing countries.

Feedback from the event is leveraged to reinforce the sustainability of the project's AI systems in the following sections also taking into account the background research that framed the AI for sustainability workshop.

- **Participation to the European AI Alliance:** Following the outcome of the "European AI Excellence and Trust in the World" workshop I-NERGY is committed to actively contribute to the European AI Alliance by presenting its results and setting a point of reference for the relevant discussions and debates within this community[48].

- **Trustworthy AI provisions in I-NERGY:**
  - I-NERGY is aware and active regarding Trustworthy AI throughout its technology transfer programs in WP6. The trustworthy AI experts of the project developed a self-assessment procedure to be followed by participating FSTP beneficiaries in order to ensure Trustworthy AI within the developed AI systems.
  - The Trustworthy AI experts of I-NERGY have been continuously looking for new emerging methodologies and technologies that can help ensure trustworthiness in the project's AI systems. Such tools can be found in the following section.
  - I-NERGY Trustworthy AI experts have been continuously monitoring the legislative framework and especially on the AI ACT and its potential future application.

3.2 Framework for Trustworthy of AI within I-NERGY

This section initially proposes a non-exhaustive set of guidelines for I-NERGY partners on how to identify the ethical risks during the development lifecycle of I-NERGY AI services. Subsequently, the methodological framework is established so that pilot and technical partners can monitor the developed AI solutions in terms of trustworthiness following a specific procedure defined by the

[48] https://futurium.ec.europa.eu/en/european-ai-alliance/blog/sustainable-artificial-intelligence-energy-sector

project. The Ethics Guidelines for Trustworthy AI are in the heart of the I-NERGY Framework for Trustworthy AI, which however incorporates recommendations and guidelines from multiple of the aforementioned initiatives such as the aforementioned AI4EU (methodology) and SHERPA (energy domain) projects and the European AI Excellence and Trust in the World (sustainability). Regarding the AI ACT, it is of utmost importance for I-NERGY given that electrical grid related AI systems would be categorised as high risk. Nonetheless, it has been used in a complementary manner, given its still ongoing / draft status. Finally, the GDPR also plays a central role within the proposed guidelines.

3.2.1 Guidelines for Identification and Management of Ethical Risks within I-NERGY AI systems

This section is structured based on the 7 requirements proposed by Ethics Guidelines for Trustworthy AI. Each requirement is accompanied by a table that identifies potential risks along with proposed technical or non-technical methods (non-exhaustive) for their mitigation. Specifically, the objective of this section is to identify potential ethical risks and implications of AI relating to the energy domain, and specifically I-NERGY AI services during the implementation phase of the project. The approaches and tools proposed can be leveraged by I-NERGY pilot and technical partners as an additional set of recommendations and tools, while applying I-NERGY's methodological framework for Trustworthy AI. Regarding the description of the 7 requirements, they have already been presented in Section 2.4.1. Additionally, we strongly encourage the reader to revisit the AI Guidelines for Trustworthy AI [52] along with the ALTAI [71] and only use the guidelines of the following sections as a search engine for specific tools and methods for their AI system lifecycle assessment.

3.2.1.1 Human agency and oversight

AI systems should be aimed at supporting human agency and decision-making and they are not meant to replace them. Table 4 indicatively lists the technical and non-technical methods that can be leveraged for the compliant of I-NERGY AI systems with this requirement.

Table 4: Indicative Guidelines for the Identification and Mitigation of Ethical Risks with respect to Human Agency and Oversight

Risks	Methods & Tools	Method Type
User Misperception, Deception, Addiction, Manipulation	Fundamental rights impact assessment. Available tools: - AI & Equality: Human Rights Toolbox[49] - An approach for Fundamental Rights Impact Assessment to Automated Decision-Making [72]. - EU tool for fundamental rights & human rights[50]	Non-technical

[49] https://womenatthetable.net/project/ai-equality-human-rights-toolbox/

[50] https://ec.europa.eu/info/sites/default/files/file_import/better-regulation-toolbox-28_en_0.pdf

	Consider the Article 22 of GDPR (Automated individual decision making)	Non-technical
	Keep the end-user informed about the level of automations and the reasoning behind decisions. Ensure user-friendly and intuitive front-end of the AI system.	Technical
	Establish governance mechanisms (human-in-the-loop, human-on-the-loop, human-in-command [52]) for the AI system. (human oversight)	Non-technical

3.2.1.2 Technical robustness and safety

Technical robustness and safety is mostly a technical requirement that is related with the resilience of an AI system to cyberattacks. All I-NERGY pilots and technical partners should continuously monitor ethical issues and risks related to energy security, given that all AI solutions strongly rely on the extensive usage of SIS which will always be at a certain degree vulnerable to cyber-crime. This requirement is critical for the power grid and related services as vital infrastructures can be affected in case of malfunctioning or downtime of its components. In the same direction, I-NERGY partners should effectively ensure all three attributes—Confidentiality, Integrity, Availability—of information security regarding the data stored and served within its AI solutions (note here that this statement is strongly linked to requirement of privacy and data governance as well). In this context, Table 5 lists the risks relating to this requirement and mainly proposes some technical methods for their mitigation.

Table 5: Indicative Guidelines for the Identification and Mitigation of Ethical Risks with respect to Technical Robustness and Safety

Risks	Methods	Method Type
Cyber incidents and cyber attacks / Technical Faults	Red Teaming / Penetration Testing	Technical
	Strong documentation	Technical
	Consider ISO standards and certifications (EU cybersecurity ACT [73])	Technical
	Fill the Machine Learning Canvas[51]. This can help validate that all the stages of the Machine Learning Lifecycle have been addressed ensuring technical robustness of the AI system.	Non-technical
	Be sure to comply with the already established measures adopted by the I-NERGY consortium (Security framework – T3.6)	-

[51] https://www.ownml.co/about

Established measures within I-NERGY Consortium

Note here that the requirement of technical robustness and safety is special for the project as the consortium as a whole will explicitly establish a security framework in the context of the dedicated Task 3.6. More details regarding the security policies of I-NERGY can be located on the respective deliverables. Therefore, what is of utmost importance is that all AI system owners comply with this framework and align / integrate their AI services with it.

3.2.1.3 Privacy and data governance

The prevention of harm to privacy requires data governance procedures that ensure the confidentiality and integrity of the data to be used and processed by AI systems. In this context, Table 6 indicatively lists several potential risks and methods, associated with the requirement of privacy and data governance within I-NERGY.

Table 6: Indicative Guidelines for the Identification and Mitigation of Ethical Risks with respect to Privacy and Data Governance

Risks	Methods	Method Type
Privacy breaches, Reidentification, Bad data quality / Integrity loss	Conduct a Data Protection Impact Assessment (DPIA) – Proposed in the GDPR [10]	Non-Technical
	Fill a Data Ethics Canvas[52] to gain awareness of the data processes relating to the use cases	Non-technical
	Ensure that no personal data are included in the datasets. (e.g. names, addresses, email addresses, location data, IPs, cookie IDs etc.). (GDPR)	Non-Technical
	Privacy-by-design (anonymisation, data minimisation, encryption, etc.). This should be aligned with the I-NERGY security framework. (Task 3.6).	Technical
	Apply quality controls on datasets.	Technical
	Consider low granularity / resolution of datasets. Consider timesteps larger than 15-30 minutes (instead of storing real time measurements).	Technical
	Be sure to comply with the already established measures adopted by the I-NERGY consortium and consider them in the development of the AI solution. - Informed consent procedures (Deliverable 1.2 – Data Management Plan)	-

[52] https://theodi.org/article/the-data-ethics-canvas-2021/

	- Anonymisation and pseudonimisation framework (Deliverable 1.2 – Data Management Plan) - Security framework (Task 3.6) - Data integrity (Task 3.2)	

Established measures within I-NERGY Consortium

In an effort to maintain privacy, personal data processing is highly discouraged and is not foreseen within the I-NERGY project. Nonetheless, I-NERGY partners should pay specific attention to cases that potentially include data coming from smart meters (e.g. prosumer and household data: UC3, UC6, UC7, UC8, UC11, UC12, EV charging data: UC10), ensuring that their utilisation will be transparent and according to the purposes collected preventing any unintended use. In case there is a need for that, the involved partner should follow the data protection procedures as indicated in section 2.1.6 and especially the Deliverable 1.2 – Data Management Plan. The I-NERGY security framework (Task 3.6) is also crucial here, securing data privacy via authentication, authorization, anonymisation and encryption mechanisms and compliance is required. Finally, the Data Interoperability and Homogenisation module that is developed in Task 3.2 is essential for ensuring data integrity, data quality through curation and imputation techniques.

3.2.1.4 Transparency

Transparency refers to 3 main concepts: i) traceability ii) explainability and iii) communication regarding the limitations of the system. Table 7 gives an overview of indicative risks and mitigation actions that can be indicatively adopted by the pilot and technical partners while developing AI services.

Table 7: Indicative Guidelines for the Identification and Mitigation of Ethical Risks with respect to Privacy and Data Governance

Risks	Methods	Method Type
Inability to contest a decision, fake expectations, untransparent decisions, excessive trust in the AI system	Communication of the abilities and limitations of the AI system across its various users. The user should know that she / he is interacting with AI.	Technical
	Strong documentation of the development process (models and datasets) and decision-making mechanisms to ensure traceability and ability of the user to contest decisions.	Technical
	Pilots can consider tools to ensure transparency of data management processes and their usefulness for users. Such a tool is: - Datasheets for Datasets [74]	Non-technical

	Consider interpretability / explainability related software components within the AI system, such as: - Lime[53] - SHAP[54]	Technical
	In the case of Deep Learning methods, consider using interpretable architectures such as: - NBEATS [75] - Temporal Fusion Transformer (TFT) [76]	Technical

3.2.1.5 Diversity, non-discrimination and fairness

It is of utmost importance that AI systems avoid discriminatory bias. Discriminatory bias refers to systematic errors in AI algorithms that lead to decisions against specific groups of people. The bias can be caused by (i) inappropriately trained AI algorithms, (ii) datasets that are not representative of reality (due to bad data collection or pre-processing). Table 8 lists indicative risks and methods directed to pilot and technical partners for the development of their AI systems.

Table 8: Indicative Guidelines for the Identification and Mitigation of Ethical Risks with respect to diversity, non-discrimination and fairness

Risks	Methods	Method Type
Discrimination, Deteriorations of social inequalities, Marginalisation, Unfair competition	Monitoring of data collection process to ensure the representativeness and quality (inclusion of different social groups)	Technical
	Consider producing a bias report through fairness metrics and open-source tools, such as: - AIF360[55] - Aequitas[56] - Audit-AI[57] - FairML[58] - Fairness Measures[59] - Fairtest[60]	Technical

[53] Avaiable at: https://github.com/marcotcr/lime

[54] Available at: https://github.com/slundberg/shap

[55] Available at: https://github.com/Trusted-AI/AIF360

[56] Available at: https://github.com/dssg/aequitas

[57] Available at: https://github.com/pymetrics/audit-ai

[58] Available at: https://github.com/adebayoj/fairml

[59] Available at: https://github.com/megantosh/fairness_measures_code

[60] Available at: https://github.com/columbia/fairtest

	- Themis-ML[61] - Fairness Comparison[62] In I-NERGY, emphasis should be given to energy equity and affordability issues as described in section 2.4.5.	
	Consider removing variables that introduce bias in the models	Technical
	Comply with accessibility standards, hence allowing all groups of people to utilise the AI system.	Technical
	Adopt participatory approaches to development. (Working groups, Questionnaires etc.)	Non-technical

Note that all pilot cases should examine measures to ensure and promote energy equity and affordability. For example, UC4, and UC13 involve energy efficiency investments evaluation based on smart meter measurements and thus they should consider criteria to ensure the fair treatment of citizens by investors. Similarly, UC7, UC8 and UC11 involve prosumer segmentation according to their capacity to provide flexibility to the grid. Hence, efficient and fair segmentation and evaluation approaches will be considered by the consortium leading to promotion equal opportunities for participation.

3.2.1.6 Societal and environmental well-being

Based on a study [77] from McKinsey, AI systems can accelerate most of the UN SDGs (Sustainable Development Goals). Nonetheless, R. Vinuesa et al. show in [78] that AI may act as an enabler on 134 targets (79%) across all SDGs, while 59 targets (35%, also across all SDGs) may experience a negative impact from the development of AI. In this context, I-NERGY partners should ensure the alignment of their AI systems with the UN's sustainability goals. Amongst them energy poverty, energy efficiency and climate change are some of the most crucial and are potentially strongly linked with power grid related AI systems.

Table 9: Indicative Guidelines for the Identification and Mitigation of Ethical Risks with respect to societal and environmental well-being

Risks	Methods	Method Type
Hinder the realisation of UN SDG's Deterioration of human skills	Monitoring and alignment with UN's SDG's[63] and European legislation and directives that have been presented in section 2.	Non-technical
	Establish frameworks for assessing the greenhouse gas (GHG) emissions of Machine Learning. The framework should encompass:	Technical

[61] Available at: https://github.com/cosmicBboy/themis-ml
[62] Available at: https://github.com/algofairness/fairness-comparison
[63] https://www.un.org/sustainabledevelopment/sustainable-development-goals/

(social skills, job loss), Reinforcement of authoritarian behaviour, Social scoring systems Acceleration of climate change	- **Compute-related impacts**, caused by both the electricity used for AI computations (bottom-up approach) and the embodied emissions associated with computing infrastructure and hardware (top-down approach) - **Immediate impacts**, tied to the short-term outcomes of applications of AI. For instance, some AI applications might decrease the cost of emissions-intensive activities (e.g., accelerating oil and gas exploration and extraction by decreasing production costs and boosting reserves) while at the same time increasing their consumption (e.g., greater use of fossil fuels). - **Structural or "system-level" GHG effects** induced by AI applications that can have broader societal implications beyond their immediate impact. We can witness to rebound (i.e., when improved efficiency may yield lower GHG emissions per use and a decrease in cost, resulting in increased consumption of the same or another good) and to lock-in effects (i.e., when AI may prevent other, e.g., low-carbon technologies, from entering the market). For example, autonomous vehicles can improve fuel efficiency, but also lead to higher rates of individualised vehicle travel (rebound effect) and ingrain the role of trucks and private cars as the dominant means of transportation, instead of enabling infrastructure for less emissions-intensive rail, public transit, and micro-mobility options (lock-in effect). Another possible system-level impact occurs when broader lifestyle changes across society, for example by changing the demand for goods and services, are induced by AI applications (e.g., advertising). **Some tools have already been proposed in** [79] from L. Kaack et al. regarding: - **Reports** for measuring AI/ML model energy use and carbon emissions - **Metrics** for reporting model accuracy as a function of computational budget - **Benchmarks** measuring training and inference efficiency)	

	Consider KPIs that will provide a measurable approach on reporting the project's compliance to SDGs for reduction of energy bills, increased RES integration and reduction of environmental footprint.	Technical
	Consider the AI Project Canvas (more general) or Human-Centered AI Canvas [80] (focused on human / social wellbeing)	Non-technical
	Revisit and comply with international and European legislation and directives regarding RES, energy efficiency and climate change as presented in section 2.3.	Non-technical

3.2.1.7 Accountability

Accountability refers to procedures relating to the responsibility during the development, deployment and use of AI systems. It involves risk management, internal and external auditing capabilities of the AI system and the management of trade-offs. This requirement can be considered as a superset of the rest of requirements and therefore can be easily managed if put into practice together with the previous ones. Table 10 indicatively lists risks and proposed methods relating to accountability.

Table 10: Indicative Guidelines for the Identification and Mitigation of Ethical Risks with respect to accountability

Risks	Methods	Method Type
Allegation, Prosecution, Opaque development processes (also look at transparency), Distrust	Conduct algorithm impact assessments (e.g. DPIA).	Non-technical
	Prepare for and conduct internal and external audits.	Non-technical
	Consider the establishment of an AI ethics review board.	Non-technical
	Strong documentation regarding trade-offs (e.g. accuracy vs explainability, privacy vs safety).	Technical

Established measures within I-NERGY Consortium

The I-NERGY project has already defined a risk management plan with Deliverable D1.1 – Project Management Handbook. This plan can be used by partners as a starting point; however, it refers to the project as a whole and not to the specific AI systems to be developed.

3.2.2 Application Methodology

In order to draw a common approach regarding trustworthy AI services, a framework that is based on the Assessment List on Trustworthy Artificial Intelligence (ALTAI), as well as on the proposed guidelines for the identification and management of ethical risks, as described in the previous section, has been designed. The sections below describe the application of this common procedure for the Open Calls and the Pilot partners:

Open Calls

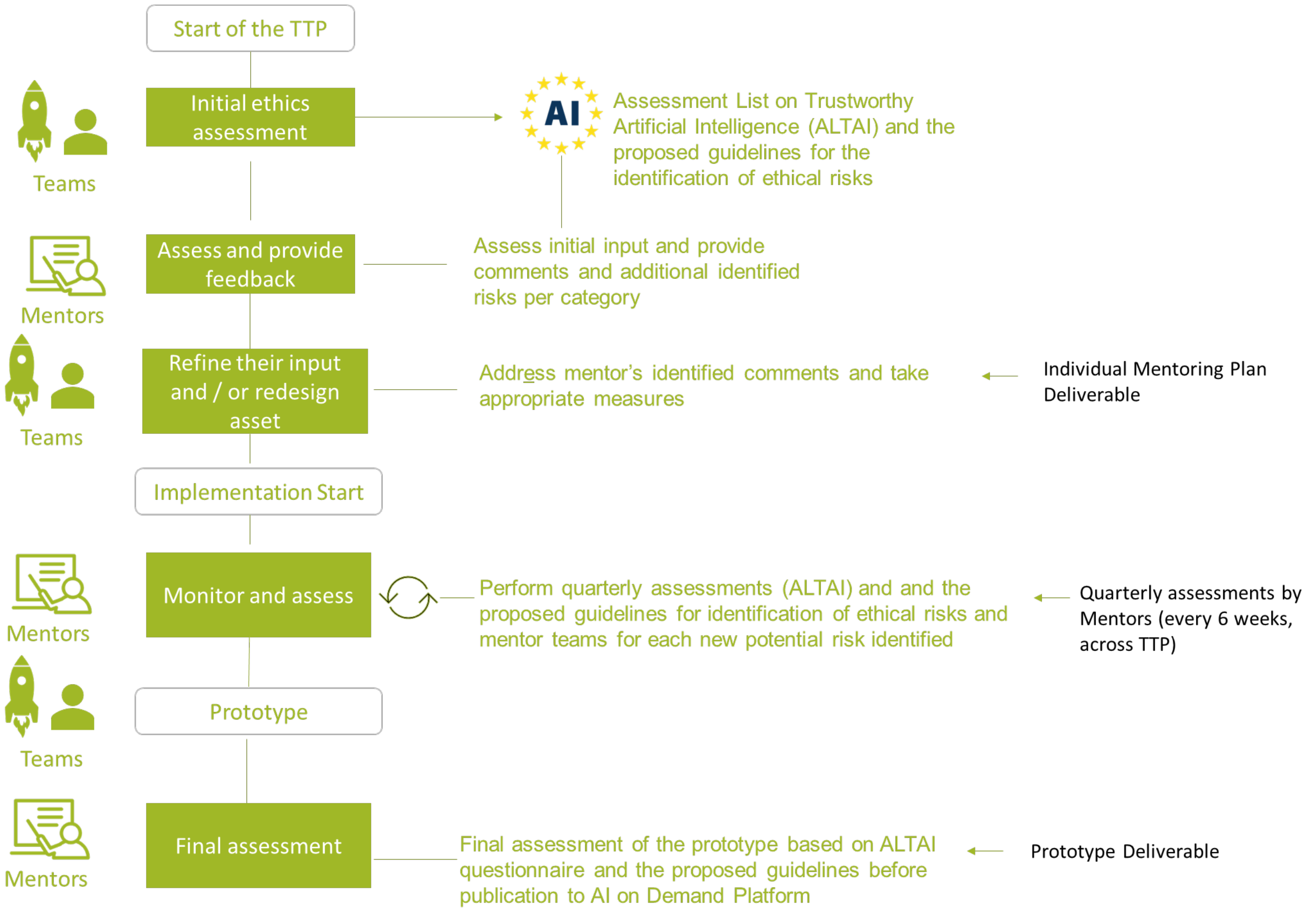
Start of the TTP
Initial ethics assessment
Teams
AI
Assessment List on Trustworthy Artificial Intelligence (ALTAI) and the proposed guidelines for the identification of ethical risks
Assess and provide feedback
Mentors
Assess initial input and provide comments and additional identified risks per category
Refine their input and / or redesign asset
Teams
Address mentor's identified comments and take appropriate measures
Individual Mentoring Plan Deliverable
Implementation Start
Monitor and assess
Mentors
Perform quarterly assessments (ALTAI) and and the proposed guidelines for identification of ethical risks and mentor teams for each new potential risk identified
Quarterly assessments by Mentors (every 6 weeks, across TTP)
Prototype
Teams
Final assessment
Mentors
Final assessment of the prototype based on ALTAI questionnaire and the proposed guidelines before publication to AI on Demand Platform
Prototype Deliverable

Figure 4 - I-NERGY Procedure for trustworthy AI services for the Open Calls

From the Open Call procedure, the teams are initially asked to perform an initial asssement on their AI systems to be reported in the Individual Mentoring and Proof of Concept deliverable. The assessment is based on providing input of identified AI ethical risks and proposed mitigation actions per ALTAI category (section 3.2 categories). The mentors evaluate the initial input and provide further suggestions and comments per category. For this procedure the mentors can collaborate and communicate on common identified issues across teams through the established private mentors channels. Thereafter during the implementation phase the mentors perform periodic assessments, whereas among others, they monitor and evaluate the ethics risks of the prototype at its current phase. This assessment is based on the following questionnaire:

Table 11 - Mentors Ethics quarterly assessment questionnaire

Requirements based on the Ethics Guidelines for Trustworthy AI and the ALTAI	Score (0-10)	Identified Issues / Risks by Mentors	Proposed actions/Comments
Indicator			
Fundamental Rights			
REQUIREMENT #1 Human Agency and Oversight			
REQUIREMENT #2 Technical Robustness and Safety			
REQUIREMENT #3 Privacy and Data Governance			
REQUIREMENT #4 Transparency			
REQUIREMENT #5 Diversity, Non-discrimination and Fairness			
REQUIREMENT #6 Societal and Environmental Well-being			
REQUIREMENT #7 Accountability			

A final assessment will take place in the context of the final deliverable (Prototype) for the FTSP beneficiaries and before the final update on the AIOD platform.

Pilot applications

For the pilot applications, the same procedure is applied with respect to the pilot application time plan and evaluation results. For each AI service that is being developed under the context of the I-NERGY project, continuous assessment of the risks for all categories that have been described by ALTAI takes place by both the pilot that intend to use each AI service and the technical partners that are developing the latter, as well as by experts in Trustworthy AI. After each assessment cycle, the implementation teams address the experts' comments and the identified risks and take the appropriate measures to mitigate them during the next iterations of the services' development. The experts group for the assessment mainly consists of the mentors from the Open Calls TTP.

The aforementioned procedure is illustrated in the following figure:

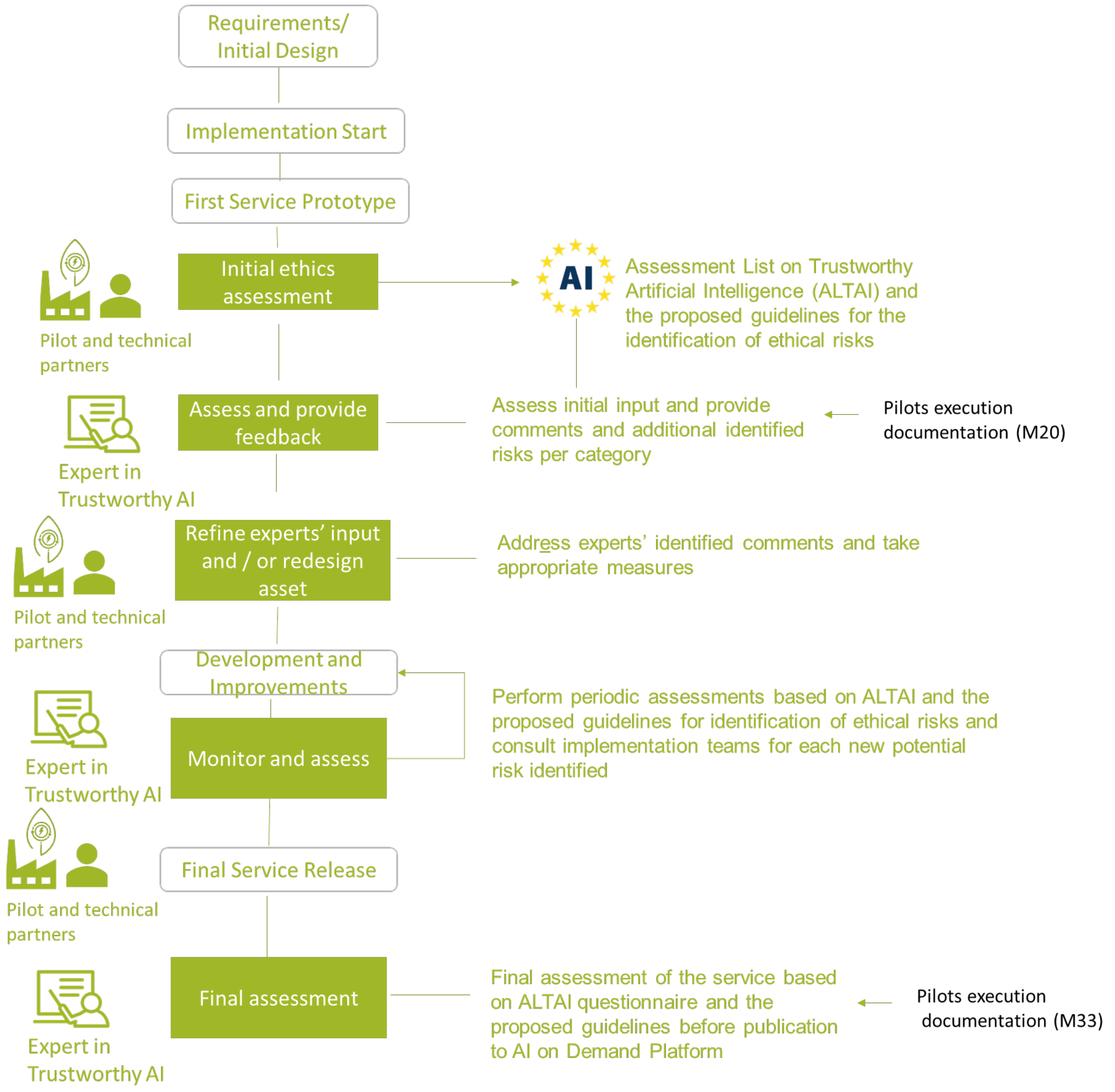
Requirements/ Initial Design
Implementation Start
First Service Prototype
Initial ethics assessment
Pilot and technical partners
AI
Assessment List on Trustworthy Artificial Intelligence (ALTAI) and the proposed guidelines for the identification of ethical risks
Assess and provide feedback
Expert in Trustworthy AI
Assess initial input and provide comments and additional identified risks per category
Pilots execution documentation (M20)
Refine experts' input and / or redesign asset
Pilot and technical partners
Address experts' identified comments and take appropriate measures
Development and Improvements
Expert in Trustworthy AI
Monitor and assess
Perform periodic assessments based on ALTAI and the proposed guidelines for identification of ethical risks and consult implementation teams for each new potential risk identified
Pilot and technical partners
Final Service Release
Expert in Trustworthy AI
Final assessment
Final assessment of the service based on ALTAI questionnaire and the proposed guidelines before publication to AI on Demand Platform
Pilots execution documentation (M33)

Figure 5 - I-NERGY Procedure for trustworthy AI services

After each assessment cycle, the most important risks for each one of the requirements described in the previous section (Section 3.2.1) are identified and reported. Moreover, several technical and non-technical methods and tools that can help addressing the identified risks are proposed, along with specific actions that are proposed to mitigate the latter. To this end, after an assessment cycle, a table with the aforementioned information should be filled in for each service and requirement. The tables of Annex II illustrate the list of specific identified risks regarding the ethics conformance and considerations, according to ALTAI and the guidelines of the previous section.

The results of the proposed approach will be reported by each pilot partner for all services they are using in the Deliverables 5.3 and 5.4 that are planned to be submitted on months 20 and 33 respectively.

4 IPR guidelines within I-NERGY

In the context of I-NERGY project, appropriate IPR handling is very important to maximise exploitation and dissemination results and overall project's outcomes while at the same time ensure protection of partners intellectual property. Therefore, it is crucial to identify the IPR of the datasets, the software, the tools and the knowledge that will be used or produced in the project, including licensing schemes, terms of usage and access rights of these assets. Prior to examine how IPR are managed within project it is useful to see how IPR are defined and can be protected.

"Intellectual property rights (IPR) are legal rights aimed at protecting the creations of the intellect, such as inventions, the appearance of products, literary, artistic and scientific works and signs, among others." [81]

The most common types of IPR [82] include:

- Copyrights, which refer to the rights that creators have over their literary and artistic works. Copyrights are applicable to a wide range of work including, for example, books, music and also computer programs and software.
- Patents, which aim to protect inventions, namely new solutions and ways to solve technical problems. Patents allow their owners to define how their inventions can be used.
- Trademarks, which refer to the signs, such as logos, that are used to recognise the products or services from different enterprises.
- Industrial designs, which protect the appearance of an article, which is related to attributes such as its shape, colours or materials.
- Trade secrets, which protect confidential information that has commercial value, are known by a limited number of people, and appropriate measures are taken for its protection.
- Geographical indications, that are signs used for goods that have characteristics or reputation which is linked with a specific geographical origin.

This section aims to provide guidelines on how and which through means IPR are managed within the project.

4.1 IPR management

IPR management is addressed privately within the project and is governed in terms of the Grant Agreement and Consortium Agreement signed by all partners covering both the rights and obligations related to background and foreground/results. In particular, aspects covered in the documents include:

- ***Access rights***, which are the rights to use the background or the results/foreground. Partners have identified and agreed on the background for the project and have also specified a background that is subject to legal restrictions or limits within CA. In addition, GA and CA provide details about the process of requesting and granting an access right. Access rights granted for different purposes, including implementation, exploitation and dissemination activities and to different entities, including beneficiaries and affiliated parties, are also covered within these documents. Moreover, access rights for parties leaving or entering the Consortium are described in CA. Specific provisions also exist in CA concerning the access rights to software.

- ***Ownership and protection of results***. Results are owned by the party that generates them. However, there may be cases where two or more beneficiaries own the results. The ownership and joint ownership regimes are described in detail in CA and GA where also the obligations to protect the results are prescribed. In addition, both CA and GA provide provisions regarding the transfer and licensing of results.
- ***Exploitation and dissemination of results***. From the perspective of IP protection, specific provisions of GA and CA describe the obligations of partners regarding the exploitation and dissemination activities.

Regarding the protection of the results/foreground, the Consortium has already identified and reported in GA some initial requirements and guidelines concerning copyright protection and licensing schemes within the project. These guidelines are expressed below:

- I-NERGY consortium embraces the vision that open-source policies facilitate the growth and spread of knowledge and innovation as well as foster the scientific progress. Therefore, the Consortium commits to open-source policies whenever possible, since there will also be results based on proprietary/legacy components that cannot fit with open-source licenses.
- the I-NERGY solution will be copyright protected using a licensing scheme that will not violate the terms and conditions of the discrete components comprising it.
- the components that can be provided open source will be delivered under such a license while components and modules that cannot be delivered open source, will be copyright protected but freely available to Consortium members to use to produce foreground.
- the Consortium will examine the software licenses of all algorithms, components and modules to be used, and decide under which license the I-NERGY framework will be released, considering the aspects of each license and the possible limitations that could consequently arise.

In order to make sure that the agreed terms are followed, to avoid disputes and to facilitate business planning, the Steering Committee will maintain an IPR Directory throughout the lifetime of the project. This document will list all items of knowledge concerning both background know-how and results/foreground, and make explicit for each item its owner, nature, status and dissemination and protection measures. The directory will be regularly updated and distributed to all partners. It will form a key tool to enable knowledge management. The Consortium has already considered and agreed on an initial approach for the IPR issues of the main project’s results, as shown below.

Initial Agreement on IP And Use Rights	Contributing Partners	Consortium Partners
DLT/blockchain tool for P2P Data/Information Sharing and Management	IPR	Use rights
Big data energy generation/demand forecasting and loads segmentation	IPR	Use rights
Asset- and System-level Digital Twin services	IPR	Use rights
Federated Learning and analytics/	IPR	Use rights
AI Energy Analytics Suite	IPR	Use rights
Best practices – Applications	Public (open access)	

Table 12: IPR strategy related to result type

In addition, task 7.3, *Business and Exploitation Planning*, is closely related with IPR management mostly from the perspective of foreground/results since it concerns the planning of the exploitation and sustainability of the project's result and will identify all the exploitable assets of the project, along with the results to be sustained following the end of the project. These activities also include the handing of IPR issues, the identification of third party's rights, comprehensive patent and trademark searches, licensing agreements with third parties, if necessary, to avoid any infringement.

The project needs to consider whether its outcomes can be exploited without infringing on existing patents claims and if its innovations can be patented. A preliminary analysis of the patents indicated that no patents/applications address AI energy-related environments, which is the main innovation of I-NERGY. Nevertheless, exploitation of project's outcomes should examine the patents landscape to avoid any infringement and consider possible applications for patents. Useful guidance for the searching and applying for patents can be found on the websites of World Intellectual Property Organization (WIPO)[64] and European Patent Office (EPO)[65].

[64] Available at: https://www.wipo.int/portal/en/index.html

[65] Available at: https://www.epo.org/

References

[1] European Data Protection Supervisor, “Data Protection,” [Online]. Available: https://edps.europa.eu/data-protection/data-protection_en. [Accessed 31 March 2021].

[2] The Member States, “Consolidated version of the Treaty on the Functioning of the European Union,” *Official Journal of the European Union (OJ),* vol. 55, no. C 326, pp. 47-390, 26 October 2012.

[3] UN General Assembly, “Universal Declaration of Human Rights,” Paris, 1948.

[4] Council of Europe, “Convention for the Protection of Human Rights and Fundamental Freedoms,” 4 November 1950. [Online]. Available: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/005.

[5] The Member States, “Charter of Fundamental Rights of the European Union,” *Official Journal of the European Union (OJ),* vol. 55, no. C 326, pp. 391-407, 26 October 2012.

[6] The Member States, “Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007,” *Official Journal of the European Union (OJ),* vol. 50, no. C 306, pp. 1-271, 17 December 2007.

[7] Council of Europe, “Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data,” 28 January 1981. [Online]. Available: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108.

[8] Council of Europe, “Protocol amending the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data,” 10 October 2018. [Online]. Available: https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/223.

[9] “Modernisation of the Data Protection ‘Convention 108’,” [Online]. Available: https://www.coe.int/en/web/portal/28-january-data-protection-day-factsheet. [Accessed 31 March 2021].

[10] “Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (GDPR),” *Official Journal of the European Union (OJ),* vol. 59, no. L 119, pp. 1-88, 4 May 2016.

[11] “Regulations, Directives and other acts,” 29 September 2020. [Online]. Available: https://europa.eu/european-union/law/legal-acts_en.

[12] Publications Office of the EU, “Protection of personal data (from 2018) : Summary of Regulation (EU) 2016/679,” EUR-Lex, 21 December 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=CELEX:02016R0679-20160504. [Accessed 10 March 2021].

[13] “National legislation,” Croatian Personal Data Protection Agency, [Online]. Available: https://azop.hr/national-legislation/. [Accessed 5 April 2021].

[14] “Legal framework,” Croatian Personal Data Protection Agency, [Online]. Available: https://azop.hr/legal-framework/. [Accessed 5 April 2021].

[15] “Personal data legislation,” Hellenic Data Protection Authority, [Online]. Available: https://www.dpa.gr/index.php/en/enimerwtiko/legal_framework/personal_data/personal_legislation. [Accessed 02 March 2021].

[16] “Reforma evropskega zakonodajnega okvira za varstvo osebnih podatkov [Reform of the European legislative framework for the protection of personal data],” Informacijski pooblaščenec [Information Commissioner], [Online]. Available: https://www.ip-rs.si/zakonodaja/reforma-evropskega-zakonodajnega-okvira-za-varstvo-osebnih-podatkov/. [Accessed 5 April 2021].

[17] European Commission, “Strengthening Europe's Cyber Resilience System and Fostering a Competitive and Innovative Cybersecurity Industry,” EU, 2016.

[18] European Parliament, Council of the European Union, “Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union,” *Official Journal of the European Union (OJ),* vol. 59, no. L 194, 19 July 2016.

[19] “Inventory of Risk Management / Risk Assessment Methods,” ENISA, [Online]. Available: http://rm-inv.enisa.europa.eu/methods .

[20] S. H. Houmb, “Decision Support for Choice of Security Solution: The Aspect-Oriented Risk Driven Development (AORDD)Framework,” Norwegian University of Science and Technology, Trondheim, 2007.

[21] “CERT (Computer Emergency Response Team). OCTAVE® (Operationally Critical Threat, Asset, and Vulnerability Evaluation),” [Online]. Available: http://www.cert.org/resilience/products-services/octave/index.cfm.

[22] “About IEA: Mission,” IEA, [Online]. Available: https://www.iea.org/about/mission. [Accessed 7 April 2021].

[23] “International Renewable Energy Agency (IRENA),” European Commission, 17 March 2020. [Online]. Available: https://ec.europa.eu/energy/topics/international-

cooperation/international-organisations-and-initiatives/international-renewable-energy-agency_en. [Accessed 7 April 2021].

[24] European Commission, “Energy union strategy,” EU, 2015.

[25] “Clean energy for all Europeans package,” European Commission, 18 December 2020. [Online]. Available: https://ec.europa.eu/energy/topics/energy-strategy/clean-energy-all-europeans_en. [Accessed 20 April 2021].

[26] European Commission, “Energy Performance of Buildings Directive,” EU, 2010.

[27] European Parliament, “DIRECTIVE (EU) 2018/844 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 30 May 2018 amending Directive 2010/31/EU on the energy performance of buildings and on energy efficiency,” EU, 2018.

[28] “Energy performance of buildings directive,” European Commission, 12 April 2021. [Online]. Available: https://ec.europa.eu/energy/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en. [Accessed 20 April 2021].

[29] “Renewable energy directive,” European Commission, 31 March 2021. [Online]. Available: https://ec.europa.eu/energy/topics/renewable-energy/renewable-energy-directive/overview_en. [Accessed 20 April 2021].

[30] European Parliament, “Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources,” EU, 2018.

[31] Publications Office of the EU, “Renewable energy: Summary of Directive (EU) 2018/2001,” EUR-Lex, 28 March 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=uriserv:OJ.L_.2018.328.01.0082.01.ENG. [Accessed 20 April 2021].

[32] European Parliament, “The Energy Efficiency Directive,” 2012.

[33] “Energy efficiency directive,” European Commission, 2 December 2020. [Online]. Available: https://ec.europa.eu/energy/topics/energy-efficiency/targets-directive-and-rules/energy-efficiency-directive_en. [Accessed 20 April 2021].

[34] E. Parliament, “Directive (EU) 2018/2002 of the European Parliament and of the Council of 11 December 2018 amending Directive 2012/27/EU on energy efficiency,” 2018.

[35] Publications Office of the EU, “Energy efficiency: Summary of Directive 2012/27/EU, Directive (EU) 2018/2002,” EUR-Lex, 11 March 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=celex:32018L2002. [Accessed 20 April 2021].

[36] European Parliament, “Regulation (EU) 2018/1999 on the Governance of the Energy Union and Climate Action, amending Regulations (EC) No 663/2009 and (EC) No 715/2009,” EU, 2018.

[37] Publications Office of the EU, “Governance of the Energy Union: Summary of Regulation (EU) 2018/1999,” EUR-Lex, 26 March 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=uriserv:OJ.L_.2018.328.01.0001.01.ENG. [Accessed 20 April 2021].

[38] European Parliament, “Directive (EU) 2019/944 on common rules for the internal market for electricity,” EU, 2019.

[39] European Parliament, Council of the European Union, “Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU,” *Official Journal of the European Union (OJ),* vol. 62, no. L 158, p. 125–199, 14 June 2019.

[40] European Parliament, “Regulation (EU) 2019/943 on the internal market for electricity,” EU, 2019.

[41] Publications Office of the EU, “Cross-border exchanges in electricity: Summary of Regulation (EU) 2019/943,” EUR-Lex, 13 September 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=CELEX:32019R0943. [Accessed 20 April 2021].

[42] E. Parliament, “Regulation (EU) 2019/941 on risk-preparedness in the electricity sector,” 2019.

[43] Publications Office of the EU, “Risk-preparedness in the electricity sector,” EUR-Lex, 25 August 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=uriserv:OJ.L_.2019.158.01.0001.01.ENG. [Accessed 20 April 2021].

[44] European Commission, “Third Energy Package legislation,” EU, 2009.

[45] European Parliament, “Regulation (EU) 2019/942 of the European Parliament and of the Council of 5 June 2019 establishing a European Union Agency for the Cooperation of Energy Regulators,” EU, 2019.

[46] Publications Office of the EU, “Agency for the cooperation of national energy regulators: Summary of Regulation (EU) 2019/942,” EUR-Lex, 11 October 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=CELEX:32019R0942. [Accessed 20 April 2021].

[47] “ACER: Mission,” European Union Agency for the Cooperation of Energy Regulators (ACER), [Online]. Available: https://www.acer.europa.eu/en/The_agency/Mission_and_Objectives/Pages/default.aspx. [Accessed 22 April 2021].

[48] “An enhanced role for ACER,” European Union Agency for the Cooperation of Energy Regulators (ACER), [Online]. Available: https://acer.europa.eu/en/Electricity/CLEAN_ENERGY_PACKAGE/Pages/An-enhanced-role-for-ACER.aspx. [Accessed 22 April 2021].

[49] “CEER: About,” Council of European Energy Regulators, [Online]. Available: https://www.ceer.eu/eer_about. [Accessed 22 April 2021].

[50] “High-level expert group on artificial intelligence,” European Commission, 20 April 2021. [Online]. Available: https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai. [Accessed 23 April 2021].

[51] European Commission, “Communication on Building Trust in Human-Centric Artificial Intelligence,” EU, 2019.

[52] AI HLEG, “Ethics Guidelines for Trustworthy AI,” 8 April 2019. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[53] AI HLEG, “Assessment List for Trustworthy AI (ALTAI),” 17 July 2020. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment.

[54] E. Commission, “Ethics guidelines for trustworthy AI | Shaping Europe’s digital future,” EU, 2019.

[55] European Commission, “Policy and investment recommendations for trustworthy Artificial Intelligence | Shaping Europe’s digital future,” EU, 2019.

[56] European Commission, “Legislation in Progress, Artificial Intelligence Act,” EU, 2022.

[57] European Commission, “Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (AI ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS,” EU, 2021.

[58] European Commission, “Regulatory framework proposal on artificial intelligence | Shaping Europe’s digital future,” EU, 2021.

[59] T. Hatzakis, R. Rodrigues and D. Wright, “Smart Grids and Ethics,” *ORBIT Journal,* vol. 2, no. 2, pp. 1-28, 2019.

[60] T. Dan, T. Nikolakopoulos and E. Darra, “The cost of incidents affecting CIIs,” ENISA, 2016.

[61] B. K. Sovacool and M. H. Dworkin,, “Energy justice: Conceptual insights and practical applications,” *Applied Energy,* vol. 142, pp. 435-444, 2015.

[62] A. Lacoste, A. Luccioni, V. Schmidt and T. Dandres, “Quantifying the Carbon Emissions of Machine Learning,” *arXiv,* 2019.

[63] G. G. Sias, “Characterization of the Life Cycle Environmental Impacts and Benefits of Smart Electric Meters and Consequences of their Deployment in California,” UCLA, 2017.

[64] L. Reijnders, “Life cycle assessment of greenhouse gas emissions,” in *Handbook of Climate Change Mitigation*, Springer, New York, NY, 2012, pp. 13-41.

[65] European Commission, Directorate-General for Communications Networks, Content and Technology, “White Paper on Artificial Intelligence - A European approach to excellence and trust,” 19 February 2020. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0065.

[66] BDVA/DAIRO, “BDVA’s response to the European Commission’s Whitepaper on Artificial Intelligence ‘A European approach to excellence and trust’,” May 2020. [Online]. Available: https://bdva.eu/sites/default/files/BDVA%27s%20reponse%20to%20the%20European%20AI%20whitepaper%20-%20May%202020%20-%20ed1.pdf.

[67] “Task Force 5: Legal and Policy,” BDVA/DAIRO, [Online]. Available: https://www.bdva.eu/task-force-5. [Accessed 1 April 2021].

[68] “The Council of Europe established an Ad Hoc Committee on Artificial Intelligence - CAHAI,” Council of Europe, 11 November 2019. [Online]. Available: https://www.coe.int/en/web/artificial-intelligence/-/the-council-of-europe-established-an-ad-hoc-committee-on-artificial-intelligence-cahai. [Accessed 1 April 2021].

[69] European Commission, “European AI Excellence and Trust in the World,” EU, 2022.

[70] European Commission, “European AI Excellence and Trust in the World,” EU, Dubai, 2022.

[71] European Commission, “European AI Excellence and Trust in the World,” EU, 2021.

[72] H. L. Janssen, “An approach for a fundamental rights impact assessment to automated decision-making,” *International Data Privacy Law,* vol. 10, no. 1, pp. 76-106, 2020.

[73] E. Commission, “The EU Cybersecurity Act,” EU, 2022.

[74] T. Gebru , J. Morgenstern, B. Vecchione, J. Wortman Vaughan, H. Wallach, H. Daumé and K. Crawford, “Datasheets for Datasets,” 2018.

[75] B. N. Oreshkin, D. Carpov, N. Chapados and Y. Bengio, “N-BEATS: Neural basis expansion analysis for interpretable time series forecasting,” 2020.

[76] B. Lim, S. Arik, N. Loeff and T. Pfister, “Temporal Fusion Transformers for interpretable multi-horizon time series forecasting,” *International Journal of Forecasting,* vol. 37, no. 4, pp. 1748-1764, 2021.

[77] M. Chui, R. Chung and A. Van Heteren, “Using AI to help achieve Sustainable Development Goals,” 2019.

[78] H. A. I. L. M. B. V. D. S. D. A. F. S. D. L. M. T. F. F. N. Ricardo Vinuesa, “The role of artificial intelligence in achieving the Sustainable Development Goals,” *Nature Communications,* vol. 11, no. 233, The role of artificial intelligence in achieving the Sustainable Development Goals.

[79] P. D. E. S. G. K. F. C. D. R. Lynn Kaack, “Aligning artificial intelligence with climate change mitigation,” *HAL,* 2021.

[80] A. Maillet, “Introducing the Human-Centered AI Canvas,” 2019.

[81] “What are intellectual property rights (IPR)?”,” [Online]. Available: https://intellectual-property-helpdesk.ec.europa.eu/regional-helpdesks/european-ip-helpdesk/europe-frequently-asked-questions_en#Intellectual_Property_Rights. [Accessed 1 April 2020].

[82] “Types of intellectual property,” 1 April 2021. [Online]. Available: https://www.wipo.int/about-ip/en/.

[83] G. Danezis, J. Domingo-Ferrer, M. Hansen, J.-H. Hoepman, D. L. Metayer, R. Tirtea and S. Schiffner, “Privacy and Data Protection by Design - from policy to engineering,” January 2015. [Online]. Available: https://arxiv.org/abs/1501.03726.

[84] European Commission, Directorate-General for Communications Networks, Content and Technology, “Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe,” 25 April 2018. [Online]. Available: https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe.

[85] European Commission, Directorate-General for Communications Networks, Content and Technology, “COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Coordinated Plan on Artificial Intelligence,” 7 December 2018. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:795:FIN.

[86] “Europe fit for the Digital Age: Artificial Intelligence,” 21 April 2021. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682. [Accessed 28 April 2021].

[87] "New rules for Artificial Intelligence – Q&As," 21 April 2021. [Online]. Available: https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683. [Accessed 28 April 2021].

[88] European Commission, Directorate-General for Communications Networks, Content and Technology, "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS," 21 April 2021. [Online]. Available: https://eur-lex.europa.eu/legal-content/DA/TXT/?uri=CELEX:52021PC0206.

[89] European Economic and Social Committee, "Opinion of the European Economic and Social Committee on 'Artificial intelligence — The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society' (own-initiative opinion),," *Official Journal of the European Union (OJ),* vol. 60, no. C 288, pp. 1-9, 31 August 2017.

[90] "Artificial intelligence: threats and opportunities," 29 March 2021. [Online]. Available: https://www.europarl.europa.eu/news/en/headlines/society/20200918STO87404/artificial-intelligence-threats-and-opportunities. [Accessed 31 March 2021].

[91] European Union Agency for Fundamental Rights, Council of Europe, European Court of Human Rights, European Data Protection Supervisor, Handbook on European data protection law 2018 edition, Luxembourg: Publications Office of the European Union, 2018.

[92] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics and Data Analysis,* vol. 38, no. 4, pp. 367-378, February 2002.

[93] L. Breiman, "Random forests," *Machine Learning,* vol. 45, no. 1, pp. 5-32, October 2001.

[94] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye and T.-Y. Liu, "LightGBM: a highly efficient gradient boosting decision tree," in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*, 2017.

[95] "Ethical Observatory description of functions, oversight powers, specific agenda and interactions with other groups," 21 August 2019. [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5c6ce1f0d&appId=PPGMS%5d.

[96] European c, "Energy Performance of Buildings Directive".

Annex I – Draft Trustworthy AI Working Group Action Plan

Draft Trustworthy AI Working Group Action Plan

ID	Stream	Activity	Owner	Co-owner	Status	Start date	Deadline	Key output
1.1	**WG coordination**	ICT49 TAI WG kick off meeting	DIH4AI	All the other projects	finished	12/04/2022	12/04/2022	
1.2	**WG coordination**	Collection of ICT49 projects' TAI activities and objectives	All the other projects	DIH4AI	finished	12/04/2022	3/5/2022	
1.3	**WG coordination**	ICT49 TAI WG May meeting	DIH4AI	All the other projects	finished	3/5/2022	3/5/2022	
1.4	**WG coordination**	Draft the first version of the TAI WG Action Plan	DIH4AI	All the other projects	finished	3/5/2022	31/5/2022	
1.5	**WG coordination**	Provide feedback and integrate the TAI WG Action Plan	All the other projects	DIH4AI	ongoing	31/5/2022	6/6/2022	
1.6	**WG coordination**	Finalisation of the TAI WG Action Plan	DIH4AI	All the other projects	to start	7/6/2022	7/6/2022	TAI WG Action Plan
1.7	**WG coordination**	ICT49 TAI WG June meeting	DIH4AI	All the other projects	to start	7/6/2022	7/6/2022	
1.8	**WG coordination**	ICT49 TAI WG July meeting	DIH4AI	All the other projects	to start	5/7/2022	5/7/2022	
1.9	**WG coordination**	ICT49 TAI WG August meeting	DIH4AI	All the other projects	to start	2/8/2022	2/8/2022	
1.10	**WG coordination**	ICT49 TAI WG September meeting	DIH4AI	All the other projects	to start	06/09/2022	06/09/2022	
2.1	**L-services Taxonomy**	Define template for collecting feedback and mapping available services	DIH4AI	All the other projects	to start	8/6/2022	30/6/2022	

2.2	**L-services Taxonomy**	Fill in the template	All the other projects	DIH4AI	to start	30/6/2022	31/7/2022	
2.3	**L-services Taxonomy**	Definition of the improved L-services taxonomy	DIH4AI	All the other projects	to start	31/7/2022	31/8/2022	Improved L-services taxonomy
2.4	**L-services Taxonomy**	Definition of the strategy to integrate the L taxonomy with OSAI / AIoD Platform	StairwAI	?	to start	31/7/2022	31/8/2022	
2.5	**L-services Taxonomy**	Definition of the business model of the L-services	Bonsapp	?	TBD			
3.1	**TAI assessment methodology - development**	Analisys of the best practices (including ALTAI) and of the previously developed ICT49 projects' assets	DIH4AI	?	ongoing	23/5/2022	30/6/2022	
3.2	**TAI assessment methodology - development**	Definition of the requirements for the common assessment methodology	AI4Copernicus	i-nergy, DIH4AI	to start	13/6/2022	30/6/2022	
3.3	**TAI assessment methodology - development**	Definition of the first version of the common methodology for assessing the trustworthiness of AI assets for the Ocs	i-nergy	AI4Copernicus StairwAI, DIH4AI	to start	4/7/2022	31/8/2022	First version of the TAI common assessment methodology
4.1	**TAI assessment methodology - validation**	Define the testing strategy and plan	StairwAI	AI4Copernicus, i-nergy, DIH4AI, Bonsapp, AIPlan4EU	to start	1/9/2022	16/9/2022	Common template for collecting feedback

4.2	**TAI assessment methodology - validation**	Feedback from each ICT49 project	DIH4AI	AI4Copernicus, i-nergy, Bonsapp, AIPlan4EU, StairwAI	to start	19/9/2022	15/10/2022	
4.3	**TAI assessment methodology - validation**	Engage the EU Digital SMEs Alliance and perform the testing	TBC	TBC	to start	19/9/2022	15/10/2022	
4.4	**TAI assessment methodology - validation**	Engage the pool of experts of the network of excellence and perform the testing	TBC	TBC	to start	19/9/2022	15/10/2022	
4.5	**TAI assessment methodology - validation**	Engage the pool of legal experts and perform testing	TBC	TBC	to start	19/9/2022	15/10/2022	
4.6	**TAI assessment methodology - validation**	Analysis of feedback and definition of lesson learned	DIH4AI	TBC	to start	19/9/2022	15/10/2022	
3.4	**TAI assessment methodology - development**	Refinement of the assessment methodology on the basis of the feedback and lesson learned from the testing	i-nergy/ AI4Copernicus	TBC	to start	15/10/2022	15/11/2022	Final version of the TAI common assessment methodology
3.5	**TAI assessment methodology - development**	Definition of a strategy for integrating the common assessment methodology in AI4EU	StairwAI	TBC	to start	15/10/2022	15/11/2022	
4.7	**TAI assessment methodology - validation**	Provision of feedback on the ALTAI and on the WG activities	DIH4AI	AI4Copernicus, i-nergy, Bonsapp, AIPlan4EU, StairwAI	to start	15/10/2022	30/11/2022	Brief Report

Annex II – Templates for listing specific identified risks regarding the ethics conformance and considerations

Table 13 - List of identified risks related to Human Agency and Oversight

Risks related to Human Agency and oversight	Methods/Tools/Mitigation Actions

Table 14 - List of identified risks related to Technical Robustness and safety

Risks related to technical robustness and safety	Methods/Tools/Mitigation Actions

Table 15 - List of identified risks related to privacy and data governance

Risks related to privacy and data governance	Methods/Tools/Mitigation Actions

Table 16 - List of identified risks related to transparency

Risks related to transparency	Methods/Tools/Mitigation Actions

Table 17 - List of identified risks related to diversity, non-discrimination and fairness

Risks related to diversity, non-discrimination and fairness	Methods/Tools/Mitigation Actions

Table 18 - List of identified risks related to societal and environmental well-being

Risks related to societal and environmental well-being	Methods/Tools/Mitigation Actions

Table 19 - List of identified risks related to accountability

Risks related to accountability	Methods/Tools/Mitigation Actions