SAFETY4RAILS Ethical Compliance Framework (ECF)

Deliverable 9.1

Lead Author: UREAD, Atta Badii

Contributors: LAU, Tuomas Tammilehto, Janel Coburn

Dissemination level: PU – Public

Security Assessment Control: passed

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 883532.

D9.1 SAFETY4RAILS Ethical Compliance Framework (ECF)		
Deliverable number:	9.1	
Version:	1.1	
Delivery date:	24/01/2022	
Deliverable due date:	31/01/2021	
Dissemination level:	PU - Public	
Nature:	Report	
Main author(s):	Atta Badii (ECF) Tuomas Tammilehto, Janel Cubrun (SIA)	UREAD LAU
Contributor(s) to main deliverable production:	All	
Internal reviewer(s):	Stephen Crabbe Antonio De Santiago Laporte	Fraunhofer MdM
External reviewer(s):	Rebecca Daniels	UREAD – not involved directly in project

Document control			
Version	Date	Author(s)	Change(s)
0.1	24/11/2020	UREAD (ECF)	Initial content structuring
0.2	02/11/2020	UREAD (ECF) Tuomas Tammilehto, Janel Coburn (SIA)	Various sections contributed
0.3	02/02/2021	UREAD (ECF) All questionnaire returns	Recompiling to abridge
0.3	09/02/2021	UREAD (ECF)	Draft Report
0.4	02/03/2021	UREAD (ECF)	Final edited post reviews
0.5	22/03/2021	UREAD (ECF)	Update following review
1.0	25/03/2021	Fraunhofer	Creation of V1_0 from V0_5 with minor updates/formatting.
1.1	24/01/2022	Fraunhofer	Update of links under Table 4 (page 16) and on page 46, the last but one paragraph.

DISCLAIMER AND COPYRIGHT

ABOUT SAFETY4RAILS

SAFETY4RAILS is the acronym for the innovation project: **Data-based analysis for SAFETY and security protection FOR detection, prevention, mitigation and response in trans-modal metro and RAILway networkS**. Railways and Metros are safe, efficient, reliable and environmentally friendly mass carriers, and they are becoming even more important means of transportation given the need to address climate change. However, being such critical infrastructures turns metro and railway operators as well as related intermodal transport operators into attractive targets for cyber and/or physical attacks.**The SAFETY4RAILS project delivers methods and systems to increase the safety and recovery of track-based inter-city railway and intra-city metro transportation.** It addresses both cyber-only attacks (such as impact from WannaCry infections), physical-only attacks (such as the Madrid commuter trains bombing in 2014) and combined cyber-physical attacks, which an important emerging scenarios are given increasing IoT infrastructure integration.

SAFETY4RAILS concentrates on rush hour rail transport scenarios where many passengers are using metros and railways to commute to work or attend mass events (e.g. large multi-venue sporting events such as the Olympics). When an incident occurs during heavy usage, metro and railway operators have to consider many aspects to ensure passenger safety and security, e.g. carry out a threat analysis, maintain situation awareness, establish crisis communication and response, and they have to ensure that mitigation steps are taken and communicated to travellers and other users. **SAFETY4RAILS will improve the handling of such events through a holistic approach.** It will analyse the cyber-physical resilience of metro and railway systems and deliver mitigation strategies for an efficient response, and, in order to remain secure given everchanging novel emerging risks, it will facilitate continuous adaptation of the SAFETY4RAILS solution; this is validated by two rail transport operators and the results supporting the re-design of the final prototype.

TABLE OF CONTENT

List of tables

List of figures

Executive summary

Through the ongoing Data Protection engagement within the SAFETY4RAILS consortium, this document, Deliverable D9.1, has established a methodologically-guided initial analysis of the range of “Purposes and Contexts” (Ps & Cs) underpinning the proposed SAFETY4RAILS data processing pipelines, and, accordingly set out the implicated stakeholder and data types. The above preliminary analysis has led the deliverable to examine the typology of data types as currently planned for processing and respective de-identification strategies where personal data would be involved. This has mapped out the essential data processing in SAFETY4RAILS and thus the requisite compliance measures to be planned, deployed and monitored throughout the project lifecycle. Accordingly the deliverable has developed a systematic evidence-based framework for legally-based actionable safeguarding steps to ensure data protection compliance; comprising: **i)** a Data Controllers’ Reference Compliance table (the “Ethical Compliance Framework Console”, ECFC), as shall be periodically updated by the Ethics Manger to enable a streamlined process whereby the Data Controller can readily determine the requisite compliance steps and legal basis by selecting the relevant recommendations as indicated on the ECFC for the particular data processing pipelines as proposed by partners; **ii)** a Consent Form Master Template, to be used to derive the appropriate Explicit Consent forms for each of the data processing purposes proposed to-date as likely to be needed within the SAFETY4RAILS project. This should help partners create and distribute the correct consent forms together with all the requisite information prior to any data acquisition. A Social Impact Analysis (SIA) is also conducted to inform socio-ethically reflective design for acceptability and accountability to ensure socially responsible innovation and its adoption. The deliverable builds on the above, finally, to set out the local and consortium governance structures to support ethical, data protection and risk-aversive measures to prevent and mitigate against any breaches of information security, and, misuse.

1 Introduction

According to the Description of Action (DoA)[1] this deliverable, D9.1, is the initial deliverable to result from task T9.1 which is to analyse and define the legal, security, societal and ethical frameworks in which SAFETY4RAILS is to operate. The task is essentially to identify and take into consideration legal, ethical and societal dimensions of the planned innovation and resulting solution including in particular ethical and legal compliance at the frontline of project implementation based on specific operational guidelines. Guidelines are required to support the adoption and implementation of carefully considered and actionable approaches to ethical and data protection as well as socio-ethically reflective design to support socially responsible innovation. Accordingly, this deliverable has developed an initial operational framework and a set of guidelines to enable the implementation of the SAFTEY4RAILS Compliance Framework to support the above objectives. The Deliverable sets out the framework of governance structures and safeguarding measures put in place by the consortium to ensure ethical, legal and security compliance monitoring and management throughout the project lifecycle. The deliverable also includes a Societal Impact Assessment, which will be conducted twice during the project.

This document is the initial version of the deliverable which shall be updated to provide the final version (version 2) consistent with the project implementation timeline. The situated context of this deliverable is characterised as a manual to support the ethical and data protection compliance management at the operational frontline of a 24-month intensive innovation project which is to enhance and integrate 17 tools to deliver the Safety4RAILS Information System (S4RIS) Platform to support the railways operations management with robust security protection capabilities for the railway network. The S4RIS platform is foreseen to be comprised of three layers:

- **Multi-lingual modular risk assessment tool** for the analysis of combined cyber-physical threats
- **Monitoring tool** for the analysis of the current situation (detection, forecasting, mitigation) providing tools to process real-data and detect anomalies
- **Simulation tool** for the analysis of possible what-if-scenarios scenarios

The S4RIS platform will provide enhanced real-time situated risk-assessment to support preparedness, predictive and rapidly responsive security protection capabilities; notably including:

- **Identification of highest impact scenarios**
- **Evaluation of mitigation strategies and countermeasures**
- **Real-time Risk Assessment for early identification of hazardous conditions**
- **Cascading effect evaluation for a fast and effective response**

1.1 Scope and Positioning of this Deliverable

This deliverable is to establish initial guidelines incorporating a framework to support:

- Healthy Consent seeking as a routinised process where it is needed to ensure ethical and legal compliance
- Seamless collaboration of the Ethical Board and the Data Controller to achieve a streamlined approach to determining and adopting the requisite steps for personal data protection and proactive monitoring for compliance assurance
- Local and consortium level governance structure for monitoring and reporting to safeguard against potential misuse of results (Information Security)

The above objectives are achieved through a methodologically-guided approach structured as follows:

[1] SAFETY4RAILS Grant Agreement, Annex 1, Description of Action, Part A, page 46-50

Chapter 1: Presents a brief introductory overview to the SAFETY4RAILS Information System (S4RIS) Platform as the ultimate objective motivating the planned innovation in this project. It presents the structure of the deliverables and its implementation logic - the rationale for the starting point of the analysis, how the analysis base is evolved through the stages and what it is to lead to and why.

Chapter 2: Provides the essential clarifications of the fundamental concepts in data protection logic including the criteria that a Data Controller has to bear in mind in determining, indeed prescribing, the steps that need to be taken to ensure ethical and legal data protection depending on the Processor, Purpose, Context (who, why, what) of any proposed processing of data that may involve personal data at some level. This Chapter lays down the cornerstones of our shared understanding to promote commitment to, and thus concerted action for, the shared objectives of ethical and legal compliance and socially responsible innovation.

Chapter 3: Honours the debt of transparency and justification of the proposed data processing and its minimisation, by essentially setting out the roots of WHY in outlining of the innovation objectives of the project, focusing on the distinct "Purposes-and-Contexts" (the Ps & Cs) involved. The Chapter then presents the analysis instruments: **i)** the Questionnaire; **ii)** The "data stakeholder types" Check Table, and, **iii)** the Data Processing Purposes and Contexts Check Table

Chapter 4: Sets out the list of implicated stakeholder types, as the Stakeholder Typology, and, the Data Typology as derived from information elicited from the Partners, including the Data Controller and the Coordinator, through the analysis instruments, plenary meetings, bilateral interviews and discussions to specify the Ps & Cs of the proposed data processing pipelines. This Chapter then discusses each type of Personally Identifiable Information (PII) as declared in the Check Tables as needed for the implementation of specific project Task(s) and for each such type of personal data, prescribes specific safeguarding measures for data protection such as anonymisation, k-pseudonymisation, generalisation through aggregation, shuffling the fields and various approaches to image/video/audio masking such as blurring, pixelization and scrambling.

Chapter 5: Presents the essential data processing as concluded from the analysis performed in the foregoing Chapters and accordingly sets out the GDPR compliance assurance requirements to be met and the legal basis adopted to ensure adherence with the GDPR supported by the Ethical Compliance Framework Console; (ECFC) to ensure the protection of privacy, and thus, dignity, fundamental rights and freedoms of data-subjects.

Chapter 6: This provides a "Consent Seeking Form Constructor Kit" as the Master Template to support the creation of "Healthy Consent" Tables, when appropriate as may be prescribed by the Data Controller in accordance with ECFC, for each of the categories of Purpose-and-Context as identified through the prior analysis of Ps & Cs and stakeholder-data typologies as explicated within Chapters 3 and 4.

Chapter 7: Introduces the analysis base for the Social Impacts Assessment including the presentation of the results of the anonymous online survey performed through a secure online instrument, namely a Questionnaire served by the EUSurvey tool, to explore the views and expectations of the stakeholders on the potential impacts of the type of solution system to result from the planned invitation in SAFETY4RAILS.

Chapter 8: Sets out the Local and Consortium Level Governance Structure including policy, monitoring and compliance assurance structures. This specifies the roles of specific members of the Ethical Board with special responsibility in relation to data protection and information security and risk-aversive measures to avoid misuse.

Chapter 9: Concludes this deliverable with an outline of the key results and insights achieved and looks forward to the ongoing examination of emerging data protection issues within the project to support ethical and legal compliance assurance as the essential component of the project management.

1.2 Deliverable Implementation Logic

For the initial phase of the project: the objective of this deliverable is to clarify the data processing that is envisaged at this stage and accordingly provide an initial analysis of the Safety4Rails Data Protection Requirements considering the data processing planned to-date and its justification responsive to the 7 GDPR principles including, in particular, the Data Minimisation and Purpose Limitation Principles and the key requirements appertaining to protection of the fundamental rights and freedoms of the data subjects as protected under the provisions of GDPR.

This process is to ensure transparency, accountability and lawful rationale for the purpose and context of any data processing. This includes protection of the rights of data subjects e.g. the right to free and informed Consent ("Healthy Consent") specifically regarding the processing of their personal data including the transfer of any data between EU and non-EU countries and the legal basis for the approaches adopted to ensure compliance.

The data protection considerations and respective legal requirements need to be explained to inform legal and ethical compliance assurance. Accordingly, this deliverable has focussed on the explication and formalisation of the nature of the SAFTEY4RAILS data processing pipelines as planned to-date pending the confirmation of the user requirements and use-cases. As data processing requirements evolve, further considerations can then be included to inform the analysis base and the resulting updated Ethical Compliance Framework Console (ECFC) as required to support to the Data Controller in ensuring the ethically and legally compliant conduct of the project.

Thus, as a first step this deliverable has provided the partners with a set of questionnaires (Table 2) and data protection check tables (Tables 3 and 4) which they have used in order to provide information about any data processing requirements of their respective task(s) and work package(s). In this way a methodologically-guided approach has been pursued to set a sharp focus on the innovation objectives of each Work Package and Task to motivate the clarification of **i)** the proposed data-subject groups, **ii)** data types, **iii)** processing pipelines, **iv)** processing context and purpose, and **v)** any data transfers in particular to non-EU countries from which some partners participate in the project.

Therefore, in the sections below, first the outline of the salient objectives of each Work Package is set out followed by the latest versions of the analysis document set deployed which comprises of the questionnaire, the Data Types Check Table (

Table 3) and the Data Processing Purpose and Context Check Table (Table 4). These have been discussed with the partners in various meetings leading to the tabularised listing of the data processing requirements as they currently stand for this phase 1 analysis. As can be seen in the follow-on sections, this analysis has in turn enabled the classification of the data subject classes and the data type classes involved in the envisaged data processing pipelines. This has culminated in the first analysis of the Data Protection compliance requirements, the Data Controller's Ethical Compliance Framework Console, to support the essential data processing within the Safety4Rail Project and its legal basis.

2 Clarification of Data Protection Concepts

2.1 Compliance Framework Conceptual Foundations

To help achieve shared understanding in support of the shared objective of ensuring ethical and legal compliance, it is important to clarify the conceptual and practical significance of the essential terms as referred to within GDPR as these are fundamental to distinguishing the different compliance requirements that have to be met and their respective legal basis in each case.

For this it is necessary to clarify the essence of what constitutes "personal data" and to provide guidelines to help partners assess whether or not and to what extent a proposed data processing pipeline involves "*personal data processing*". This is dealt with in the context of the following definition of the key concepts related to data protection as follows[2]:

Natural person: this refers to any individual who is alive. This term is used in contradistinction to a legal person (a company/organisation/enterprise) or a deceased person, the information about either of whom is not considered to be personal data.

Processing data: This includes any automated or manual operation involving capturing collecting, recording, organising, structuring, storing, adapting, altering, transferring, erasing/destroying/deleting of (personal) data.

Anonymisation: This refers to the process of transforming the data, either reversibly or irreversibly, into a string as an encrypted structure.

De-identification/De-linking: this is to transform a personal data element, almost always reversibly, such that it is not possible to readily link the data to an individual although through authorised/unauthorised access to the code used for the transformation or by combining other elements of data (which are not de-identified) the person could be identified.[3]

(Re)Identifying: This is to identify a person by combining elements of personal data, some of which may have been pseudonymised/masked to some extent to render them de-identified.

Data-subject: This is any identified or identifiable natural person whether their data is already processed or not; e.g. a person is a data-subject already even when although they are merely being contacted to consider signing a consent form to permit the use of specific personal data of theirs for a particular purpose and context of processing over a specific period – because they have been identified and similarly when they are invited to participate in a workshop and/or act as an advisor etc they are a data-subject by virtue of having been identified even if the extent of processing their data may have been simply to emailed them.

Data Controller: This is the entity (person/organisation) who determines the purposes and means of the processing of personal data; a Data Controller could be i) public organisation/NGO, ii) a research institution, iii) a private company/enterprise; The purpose of a Data Controller would be typically related to their role and responsibilities which mainly could be distinguished as falling under the two categories of "social/public-duty, not-for -profit" objectives or "business, marketing" objectives.

Data Processor: This is the entity (person/organisation) who processes data on behalf of a Data Controller.

[2] What is Personal Data? - European Reference Legal Information Site
[3] Article 29 European Data Protection Working Party

Personal Data: this is any data that relates to a directly or indirectly identifiable natural person; such data could include direct or indirect (secondary/linkable) identifier of a person)[4]:

- Direct identifier can include, for example: name; surname, home address, email address (name.surname@company.com) identification number; location data; and online identifier such IP/MAC address, cookie ID, the advertising identifier of a person's mobile phone.
- National ID/passport number
- Data held by a hospital or doctor, which could be a symbol that uniquely identifies a person.
- Physical/behavioural attributes: e.g., shape/colour/markings/distinguishing feature of face/any body part (e.g. hair/eye colour, birthmark, voice, gait etc.).

Other factors can identify an individual (the above is a non-exhaustive list), other elements of data that can directly identify an individual is personal data although processing any (potential) identifier of an individual would not in itself mean that the Data Controller is processing personal data because the purpose and context of processing of any elements of personal data is crucial to determining whether personal data is being processed; this will be further explained later in the context of Relate-ability and Link-ability of personal data.

Special category data: this is Sensitive personal data, including: genetic data relating to the inherited or acquired genetic characteristics; healthcare data (physical, mental health status and healthcare services used); biometric data to uniquely identify a natural person (including facial images, fingerprint, gait); racial or ethnic origin; political, religion, philosophical opinions and/or beliefs; religious or philosophical beliefs; trade union membership; sex life or sexual orientation; information about any criminal history.

2.2 Data Controller's Compliance Requirements Assessment Process

For a Data Controller to perform the requisite situation assessment in order to determine whether the proposed data processing is indeed to be classified as processing personal data, the processing would need to be examined to assess if there are any steps envisaged that may relate-to the data-subject (and concomitantly link to link and place the data-subject) and thereby lead to their identification. Accordingly, the Data Controller would need to examine the extent of Relate-ability, Linkability and Place-ability; by reflecting on the following observables:

Relate-ability Criteria:

An entity is said to be processing personal data if:

i) It is processing any of the above types of data or other data that can potentially (re) identify the person directly or indirectly -
ii) If the data is not irreversibly anonymised (reversibly anonymised/de-identified/de-linked/ pseudonymised data is still personal data)
iii) If the purpose and context of the data processing is such that it ***relates to*** the individual
iv) If it relates to the individual even if it is factually incorrect it is still personal data
v) If there is uncertainty about whether or not a processing of some data can be linked to an individual and this amount to personal data processing, then it is safer to assume that it is personal data processing.

As stated above a crucial qualifier here is whether the processing ***relates to*** the individual and this depends on:

A. The Data Controllers' position and purpose (whether intent on identifying the person or not);
B. The content of the data; the purpose and context of processing;
C. The likely impact or effect of that processing on the individual.

[4] Article 29 Working Party Opinion 4/2007 on the concept of personal data

It is possible that processing the same data of an individual would amount to personal data processing or not depending on any one of the above 3 factors.

A situation can occur whereby the same data does not relate to an identifiable individual for a particular controller, e.g., a public institution carrying out a general classification or survey for which the linking of a linkable personal data to identify a person is not the purpose whereas the same information in the hands of another controller, e.g. a private enterprise conducting a marketing campaign, would amount to personal data processing, an example of this is as follows:

If any researcher wishes to examine the trend in price of houses of a particular size in a particular area they may (at random) select a particular house prototypical of the category of interest (e.g., a typical 4-bedroom house) to examine the trend in its valuations over the period of interest. The address would be linkable to the owners but the processing of information about someone's house at the intended level would not amount to processing personal data. As soon as the information is used for, say, the purpose of finding out the pattern of electricity consumption or any other aspect that begins to involve the owners' identity and lifestyle information e.g., for marketing purposes then processing their address becomes personal data processing. Another example is a tourist taking a photo of a monument or a landmark feature of a city that they are visiting, e.g. in the vicinity of Big Ben in London, near the Parliament Square, with other persons in the background (no personal data processing) versus the same photo if it were to be requested and processed by the police as part of a surveillance operation (personal data processing).

Linkability-Place-ability: In the above example the clock on Big Ben will show personal data as it carries Linkability -Place-ability information for all identifiable persons in the photograph. Even more so would be the case were this to have been a mobile-phone/cam video recording with audio stream presenting a characteristic siren of a distant London ambulance/police car approaching/or moving away from the location (Doppler Effect) indicating where the individuals were at that moment of time as shown by the clock – thus placing and locating the individuals and associating them potentially with ambient events. Thus, if for the purposes of a controller, the identity of the individuals is irrelevant, the data therefore does not relate to individuals (no personal data processing) but if the identity is important then it obviously does relate to them (personal data processing). Even although the information by itself may not directly involve the identity of the person it may still be linkable by the controller or anyone interested and determined to link other information e.g., by those who know the individual. For example, a person videoed wearing a hoodie walking away from a cctv camera at night with poor street lighting would most likely be recognised by his friends and family (e.g., based on his gait or his trainers).

Such Linkability can also occur in the case of data that may be obscured/masked in other data modalities such as textual data e.g., with pseudonymised data which is viewed as personal data. The linkability potential of indirect identifiers in leading to the identification of an individual has a dynamic value that evolves as the specificity-uniqueness of indirect information and its context changes over time. If in the example above the manufacturer of the trainers were to have discontinued the particular model of trainers worn by the person in the hoodie, then he/she would become even more easily and readily linkable by more people than previously.

Another example of linkability potential is e.g., the role/occupation information e.g. compare two persons for whom the occupation is given to indicate that in one case person the person works for Starbucks and in another for a micro-SME. Clearly the role information in the former case can hardly be classed as linkable to the person's identity (Starbucks employs over 350,000 persons worldwide) and so is not personal data whereas it would certainly constitute personal data in the latter case as the information could be easily linkable to a particular person given the micro-SME has only a couple of employees. However, were the SME have a meteoric growth requiring many more employees, then to that extent the same occupation information for the same employee will have less of a linkable potential and in the limit, it may no longer be considered as personal data were the Ex-SME, now a massive enterprise, to be employing a very large number of employees indeed.

3 Requisite Analysis Base to Support the Compliance Framework Development Methodology

The above review of the key concepts underpinning the criteria for data protection led to an analysis of the purposes and contexts of the data processing pipelines as proposed within the SAFETY4RAILS project to-date. This constituted the cornerstone of our methodological approach to:

i) Characterise and classify the proposed data processing to arrive at an inclusive set of distinct data processing categories.
ii) Assess the compliance challenges to be met responsive to each type of proposed data processing and determine the relevant legal basis and respective actionable steps to be taken by the partners involved to ensure ethical data protection compliance.

3.1 Systematic Top-down and Bottom-up Transparency of the Rationale for Data Processing and the Choice of Data Protection Compliance Strategy

Following the 7 GDPR principles, in particular, the need for transparency in setting out the rationale for adherence to data minimisation and purpose limitation it was necessary to set out the objectives of the project with a view to highlighting the linking of WP-level objectives (Table 1) to the downstream purposes and contexts that justifiably ensue from them at the task level and that demonstrate the rationale for the proposed data processing consistent with the above GDPR principles; in particularly transparency and data/purpose minimisation. Accordingly, in the analysis chain that follows, we first set out an outline of WP-level objectives in such a way to contextualise the specific purpose and contexts of data processing as subsequently elicited through the following questionnaire and tabularised in the check tables (

Table 3, Table 4). This led to our situation assessment of the spectrum of proposed data processing as required by various tasks in the project. As a result it was possible to identify the distinct types of data processing that broadly shared the same categories of purpose and context for each of which it has been possible to determine an appropriate compliance strategy and its respective legal basis as set out in Table 7 referred to as the Data Controller's Ethical Compliance Framework Console (ECFC). The ECFC shall be kept updated to assess and classify any data processing requirements that may emerge responsive to stakeholders' needs. In this way the ECFC shall serve as a live reference grid for the Data Controller in order to help assess the nature of a proposed data processing in terms of its *Purpose and Context* and accordingly advise the partners of the appropriate steps to be taken to safeguard the ethical and data protection compliance assurance for current and any emerging data processing requirements.

3.2 Outline of SAFETY4RAILS Work Package Objectives

TABLE 1 OUTLINE OF WORK PACKAGE OBJECTIVES

WP1 Project management
Project coordination and management, including budgets, quality assurance, review and submitting deliverables Maintaining the relationship with EC project officer assigned to the project
WP2 Requirements, specification and architecture of the S4R Framework **Establish the specification of the Safety4Rails architecture** based on past failures and OSINT data analysis, Identify requirements and specification for the SAFETY4RAILS Information System (S4RIS) to ensure tools developed in WP3and WP4 meet the needs of end-users.
WP3 Development of a multilingual risk assessment tool for combined cyber-physical threats in S4RIS Develop Risk Assessment tool for dealing with cyber-physical threats against Railway infrastructures and networks Upgrade and Extend the SAFETY4RAILS, SecuRail tool for cyber-threats protection to be integrated with the simulation and real-time monitoring tool to exchange input and output data.

WP4 Monitoring Methods for S4RIS – Detection, Forecasting, Response and Recovery **Analysis of cyber-physical incidents, monitoring, anomalies detection, threat propagation, and predictive monitoring** Crisis Management methodology for coordination of response adaptation, mitigation and recovery management Analysis of Blockchain technology capabilities applicable to prevention, detection, response and recovery requirements
WP5 Simulation methods for S4RIS- Prevention, Preparedness and Risk Mitigation **Develop a taxonomy of risks and vulnerabilities, resilience strategies, disaster response for decision support** **Develop and integrate into S4RIS a conceptual toolkit and innovative strategies to identify, prevent and mitigate the impact of various threats, and the technical challenges to be overcome in adopting preparedness, prevention and risk mitigation strategies of S4RIS and their technical challenges** **Identify-Assess the impacts of any attacks on e-components used in the railways control systems exposed to internet in EU** **Develop an agent-based simulator to discover vulnerabilities, estimate parameters useful for risk mitigation, and test the effectiveness of surveillance and security policies and estimate the risk of possible criminal events.** **Provide a data ingestion and decision support tool for insights for operators, search and visualization capabilities**
WP6 Implementation of SAFETY4RAILS Information System (S4RIS) **Develop, implement, integrate a**nd usability test the overall S4RIS platform and its various features and tools
WP7 Policy planning and investment measures of prevention-detection-response-mitigation **Develop a comprehensive** approach to resilience, preparedness and prevention by including financial and budgetary elements from the inception of the SAFETY4RAILS framework.
WP8 Simulation Exercises and Evaluations in Operational Environment **Validate the SAFETY4RAILS solution with different organisations and scenarios for responsive security protection** **Prepare, run and evaluate the results of the simulation exercises covering various scenarios to ensure full spectrum applicability**
WP 9 Ethics, Legal, Privacy and Societal Aspects **Ensure the ethical, legal and societal issues are considered at every stage of the project; conduct social impact assessment** **Develop the S4R Compliance Framework taking into account the broader spectrum of all legal and ethical issues** **Develop the S4RIS (privacy-by-design) according to the railway sector EU regulation and security standards** **Provide Guidelines for Ethically Responsible and Sustainable crisis communication and information sharing** **Perform analysis of regulatory standards and compliance requirements and define the legal framework for the deployment certification and standardisation of the SAFETY4RAILS solution** **Recommendations for a consistent ethical and security framework responding to threats and attacks in the EU railways.**
WP10: Communication, Dissemination and Exploitation **Deliver a strategy with milestones to be reached and KPIs to be maintained** **Implement the tools for efficient communication and dissemination of results** **Provide a market analysis and value proposition as well as a Business Plan**

3.3 Questionnaire on the Current Data Processing Plan

TABLE 2 QUESTIONNAIRE ON THE CURRENT DATA PROCESSING PLAN

This to provide information on the data processing planned so as to assess the data protection requirements
Q1. What data type do you need for executing what part of your development task for developing which sub-system **Q2.** For what functional, non-functional part of the development planned in your task do you need the data? **Q3.** What are the minimum number of data elements (fields) that you will need to access/process/store/transfer? **Q4**. State which elements for which processing to develop what module to serve which functionality (use-case, use-scenario) of the system? **Q5**. Which data elements are absolutely necessary for which use-case (implies use-context)? **Q6**. Are any of such data elements part of or related to personally identifiable information (PII) or PII-related? **Q7**. What are the PII elements that are essentially to be processed? - e. g. name, address, age, gender? **Q8**. Is the required data to be captured (as by IOT or by data streams feed/synthesised/received from Railway operation stakeholder) obtained in any other way? -which? **Q9.** Will data be anonymised at source or do you plan to render it de-identified and if so how and at what stage? **Q10**. Will the data be used as pseudonymised data? **Q11**. Will the data be transferred to other third countries, if so which? **Q12**. How long will the data need to be processed and thus kept stored before it could be deleted? **Q13**. Where and how is that data to be kept in storage? **Q14.** What access control safeguards will be in place to protect the data at-rest and in-transit?

TABLE 3 DATA TYPES CHECK TABLE

Please tick √ in the relevant cell to indicate the type of data expected to be used/exchanged/disseminated													
Data Source /Type	Data Format	SAFETY4RAILS Activity Types that may need to Capture/Acquire/Generate/Process/Store/Transfer/Delete Data											
		Consortium Meetings	Questionnaire Surveys Requirements Interviews & Workshops			Website Clicks Cookies, Webinar Attendee's Registration data	Demonstrator Trials User feedback & Usability Evaluations	All Technical R&D WP3, WP4, WP5, WP6, WP7					Open Data Mangmnt
			WP2	WP10	WP9	WP10	WP8	WP3	WP4	WP5	WP6	WP7	WP1-WP9

TABLE 4 DATA PROCESSING PURPOSE AND CONTEXT CHECK TABLE

SAFETY4RAILS Project Data Protection Requirements Analysis Checklist Table								
WP#	WP Objec -tives	What Data Types is Used ? This is as specified in the data type table above which is to be filled first and each data type ticked in that table needs to be entered here in a separate row so that the various processing could be described along the row under each column of this table. This is so that a data protection situation assessment can be done, and any safeguarding measures stated as possibly needed	Already Anonymised -at-Source or not? The answer can be Yes - anonymised at source OR No - Not anonymised-a or de-identified? OR Not Applicable - because this data type would not contain any personally identifiable (or related) information.	The Scope of the Processing -what are the data elements? In this column please state what the information content will be in the data and if it involves persons (data-subjects), will the data include personally identifiable Information (PIIs) or Special Category Data (SCD) The data-subjects demography and how long the data will be kept? When deleted? Also confirm deletion if a consent form were to be revoked and if data-subject still identifiable in the aggregated data- (no need if data already irreversibly anonymised /aggregated).	The Data Processing Sequence -what processing is planned? In this column please state 1) What will happen to the data from injection (when you access it) through various stages of (pre) processing (e.g. feature extraction, modelling etc.) 2) What data exchanges are planned, if any, involving non-EU countries 3) What data protection safeguarding measures will be taken for data transit and data-at-rest in either direction re any data transfers to/from non-EU countries	The Purpose -why is there a need for this processing of this data? In this column please state why each element of data is necessarily required and the why each stage in the sequence of the data processing that you have outlined in the previous column is necessary e.g., any pseudonymisation, random sampling, the feature extraction, modelling, training and testing etc. The point here is to see if the processing will ensure privacy - preserving and unbiased modelling etc.	The Nature of the Processing: - what protocols you plan to perform before data capture, storage, transfer (e.g. consent, secure access) In this column please state the protocol planned to be followed from the point of pre-capture to final deletion e.g. if the data capture is to involve persons (data-subjects) then one needs to mention the necessary Explicit Consent process and if any Personally Identifiable Information or related data (trace-able back to potentially identify the person) are involved, then one needs to mention any (pseudo/ano)nymisation-at-capture (if not already done-at-source, i.e. at pre-capture/pre-transfer), state where the data will be stored (data-at-rest) and how (encrypted?) and how access controlled and eventually when to be deleted unless already deleted before responsive to revocation of consent as applicable.	The Context of the Processing – for what use-cases ? In this column please state the use-scenario context where the data is used during development and during testing and deployment in the operational workflow context of the stakeholders e.g. for situation assessment or, detection or classification, or alarm raising and decision support.

Partners were invited to respond to the above questionnaire (Table 2) and the Check Tables (Table 3, Table 4) and through a process of plenary and bilateral Q & A meetings, the following Purpose & Context Table (Table 5) was concluded to indicate the spectrum of the proposed data processing requirements to support the planned tasks as shown in below.

TABLE 5 OVERVIEW OF PARTNER'S PLANNED DATA PROCESSING

partner Acronym	Work Packages & Tasks	The Purpose & Context (P & C) of Data Processing	Personal Data Processing Planned Yes/No	Transfer of any Personal Data Planned Yes/No
AC	WP4- T4.3	Blockchain records, BCT Feasibility Analysis	**No**	**No**
EOS	WP2, WP8, WP10, T10.1., T10.2, T10.3, T10.4	Involving requirements elicitation and/or usability evaluation, workshops engaging the stakeholders in the railways sector, interviews with prior consent formally sought, data may include: name, email, gender, region, type of organisation, data pseudonymised, de-identified. Involving the use of social network data (Facebook, LinkedIn, Twitter) through respective APIs, and Mailchimp to access the stakeholders for dissemination; data types would include data as held on such social media (name, surname, address, age, sex, localisation, interests, preferences, photographs, network, gender. -Not anonymised at source but anonymised at the point of capture, data is processed for aggregation and modelling; data stored in secure role-based access server within passworded documents; data be kept for no longer than necessary to pursue the research objectives and for no other purpose.	**Yes** Appropriate de-identification steps to be advised by the Data Controller based on available methods for the data type table 6) and the options indicated as per the Ethical Compliance Console (table 7).	**No**
	WP9-T9.2	Information analysis for communication and information sharing platform,	**No**	**No**
Fraun-hofer	WP1 T1.1-1.4	Quality management and coordination activities involving the Advisory Board and end-user groups. Data generation and collection in the course of administrative, financial scientific and technical management, User questionnaire consent forms etc as applicable; data would include contact details, aggregated data on personal cost per beneficiary. The data types to be largely textual data collected/generated -not yet fully defined. The formats of the data are also not yet fully defined but are likely to include: docx, .pdf, xlsx, .txt, .html, .jpeg / png etc .pptx. For data other than those already included in public sources, data protection steps will be taken as advised depending on type of data to be processed when this shall become clear.	**Yes** This data processing is essential to the conduct of the project management. As such this is legitimate use as determined by the Data Controller *-only limited to the processing of the minimum of personal data solely for the purpose of project management & administration.*	**Yes** Any personal data will only be shareable on a strict need to know basis for the purpose of the management of the project subject to the contractual obligations to data protection.
	WP2	Involving end-user requirements elicitation: consent forms and questionnaires processing	**No**	**No**
	WP4, WP5 T4.4, T5.1, T5.3, T5.4	Data to include the essential parameters as required for railways system modelling; e.g. parametric (sub)system functional performance data, threat surfaces, and, historical threat incidence data for cascading effects simulation.	**No**	**No**
	WP10. T10.1 to T10.5.2	Data types include: emails, databases and presentations, use cases, feedback on evaluation and validation results, MS Teams and Outlook type of images, plus other personal data that end-users may already have online. All partners have access to the names and images of all staff working on the project partners UIC, MdM, UREAD, ECIRA and FHG have access to names and contact details of Board Members 3rd parties will have access to the names and contact details of beneficiaries' staff participating in events.	**Yes** This data processing is essential to the conduct of the project management. As such this is legitimate use as determined by the Data Controller *-only limited to the processing of the minimum*	**Yes** Any personal data will only be shareable on a strict need to know basis for the purpose of the

partner Acronym	Work Packages & Tasks	The Purpose & Context (P & C) of Data Processing	Personal Data Processing Planned Yes/No	Transfer of any Personal Data Planned Yes/No
			of personal data solely for the purpose of project management administration.	management of the project subject to the contractual obligations to data protection.
IC	WP2,	Literature survey data, operational parametrics of sensors, legacy cyber-physical incident detection analysis,	**No**	**No**
	WP3, WP4, WP5 T3.1, T3.2, T3.3, T4.1 T4.4	T3.2 Analysis of end-user needs and requirements for risk analysis tool task T3.3, Processing raw data collected by Task 3.1 (text and numbers) T4.1 Sensor data, active system monitoring, AI-based anomaly detection, identifying potentially unknown issues, multiple data sources based situation assessment of the operational and social context of events.	No Data expected to be anonymised at source, if not, then appropriate de-identification steps to be taken as advised by the Data Controller based on ECFC (table 7).	**No,** T3.3 Data transfer to SecuRail Tool by STAM in Italy, data encrypted in-transit & at-rest.
	WP5 T5.1, T5.3, T5.4	T5.3, modelling cascading effects on the railway system using system components attributes functionality and performance data, threats experienced by the system and mitigation measures currently deployed; system characteristics and situated context of attacks and their impact – developing a threat catalogue.	**No**	**No**
		T5.4 to elaborate new mitigation measures based on those currently deployed as known through WP2.	**No**	**No** Aggregated results only maybe shared.
INNO	WP4, T4.2,	Involving IoT/SCADA and related system data; data on system components, installation, configuration, attack vulnerabilities and attack vectors. Some data may be sourced from crawlers to find security related data updates using IoT and SCDA data.	**No**	**No**
INTRACOM	WP4, WP5	Literature, text, open-data re attack and illicit activities vector profiles, explicit network traffic excluding data on origin and content; only target address, port, protocol and commands are processed for intrusion detection to be deployed on railway operators' network with no data transferred outside.	**No**	**No**
LAU	WP9 WP10	Statistical data from web pages and social media pages, personal data of the targeted stakeholders and citizens to be engaged. Literature, references, consent forms, questionnaires, deliverables, contact details. Data elements: first and last name, email addresses, country, type of organisation, region, gender. Data Acquisition: through publications and posts received from railway operator or any other partner Purpose and Context: to be able to reach the stakeholders and citizens targeted in an efficient way to publish and post relevant content, to analyse relevance of content shared on the websites/social-media	**Possibly** If any, then appropriate de-identification steps to be advised by the Data Controller based on available methods for the data type table 6) and the options indicated as per the Ethical Compliance Console (table 7).	**No**
MTRS	WP3	Literature, references. open-source data	**No**	**No**
RINA	WP5 -T5.1	Past incidents, risk scenarios, attack vectors and mitigation as input to domain taxonomy Data de-identified at-source	**No**	**No**

partner Acronym	Work Packages & Tasks	The Purpose & Context (P & C) of Data Processing	Personal Data Processing Planned Yes/No	Transfer of any Personal Data Planned Yes/No
			Data is anonymised at source, else appropriate de-identification steps to be taken as advised by the Data Controller based on ECFC (table 7).	
	WP9 - T9.3	Technical data re standards and regulations and interoperability requirements	**No**	**No**
RMIT	WP7, T7.1	Data on maintenance, repair and rehabilitation records of rolling stock and systems. Assets Inventory maintenance and investment plans, Developing an investment assessment model, for cost-benefit evaluation of risk mitigation and recovery strategies. Data Source: literature on maintenance management, not including personal data but in any case, data can be de-identified at-source.	**No** No personal data expected, but if any then deidentified at source.	**No**
UIC	WP2, WP7, WP8, WP10	Involving interviews/ questionnaires regarding critical assets and for end-user feedback, vulnerabilities, threats and cascaded effects analysis and end-users' decision support needs; Published literature, references, images, video files used for the simulation exercises to assess S4RIS usability; person profiling within the simulation exercises; Data types including video, images, xml, Json files (offline sensors data), documents, software generated data, NDAs, consent forms, railway infrastructure data, sector professionals contact detail, consent forms to be used prior to data acquisition. Personal data could be included either in the input set or output set of the system being tested, depending on the requirements to be specified; NDAs and consent forms will be deployed as required data will be anonymised, Audio and video data images to be de-identified through masking and blurring techniques.	**Yes** Appropriate de-identification steps to be advised by the Data Controller based on available methods for the data type (table 6) and the options indicated as per the Ethical Compliance Console (table 7).	**No**
UNEW	WP2, WP6	Integrating the enhanced tools resulting from WP3,4,5,7 into the SAFETY4RAILS Information System S4RIS Data elements to be specified later – to be anonymised as required. Person profiling in xml and Json not anonymised, and it contains: Account ID, email, password, data used solely to access the S4RIS platform, it is stored internally in the system and not shared outside that application, data used solely for logging-in for S4RIS; online authentication, encryption and secure access.	**Yes** Data is essential for user registration and its use limited to access control, authentication and service provision and secure storage. Appropriate measures to be taken as advised by the Data Controller based on the ECFC (table).	**No**
UREAD	WP7,	Domain ontology and semantic modelling of railway systems, use-context based, threat-driven risk severity calculus and responsive countermeasures optimisation.	**No**	**No**
	WP9	Published information on use-cases, user-scenarios, partners' secure server setup, information on contact detail of partner's DPOs, the Data Controller and the Ethical Advisory Board.	**Yes** Minimum personal data for the legitimate use of the project.	**No,** Minimum personal data shared for the legitimate use only.

partner Acronym	Work Packages & Tasks	The Purpose & Context (P & C) of Data Processing	Personal Data Processing Planned Yes/No	Transfer of any Personal Data Planned Yes/No
WINGS	WP4, T4.1	Active system monitoring, forecasting and detection of anomalies using AI methods, data will be an integral part of the functionality,	**No** Data is de-identified at source; else-de-identified at acquisition as advised by the Data Controller.	**No**

4 SAFETY4RAIL Stakeholder & Data Typology

4.1 SAFETY4RAILS Stakeholder Types

The Stakeholder Pool for this project includes two main types of stakeholders, in direct involvement with the project activities, both of whom are targeted to be end-users or beneficiaries of the proposed framework:

- End-users: These beneficiaries shall be the final end-users of the SAFETY4RAILS targeted system as planned. These are the operational frontline Railways staff and the service delivery managers; in particular, those with safety, security and maintenance planning responsibilities.
- Experts: in the context of SAFETY4RAILS project these are end-users with considerable sectoral experience in this particular application domain of Railway Systems, and broadly, in Critical Infrastructure and Cyber-Physical Security; possibly also acting as consultants within the sector.

Of course, there exist other stakeholder types who are indirectly involved with the project activities, these include the local and central government and security forces and indeed the passengers who constitute the other stakeholders as both the customers and member of the society, who would overall be affected, directly or indirectly, by the impact of the level of efficacy or otherwise of the resulting S4RIS platform and rightly expect a safe, secure and privacy-preserving solution that responsibly addresses their needs. Indeed, this demands the socially responsible and ethically reflective innovation as motivated by the Social Impact Analysis that provides the other important analysis base as incorporated within this deliverable to inform the overall ethically and societally reflective design of the SAFETY4RAILS solution architecture and its innovation pathways.

4.2 SAFETY4RAILS Data Types

4.2.1 Network Security and Intrusion Detection Data

No recruitment strategy is required because the Networking security and intrusion detection data will be collected through synthetic data supplemented by automatic logging of cyber operations with the prior consent of, and under the domain of the end-user partners. For the networking security and intrusion detection system some profile data would be needed for the system to most effectively learn to prevent cyber-attacks by observing various user attributes on the network.

4.2.2 Requirement Engineering

The Requirement engineering data requires a recruitment strategy since the process has to have an interaction with the stakeholders in order to elicit and prioritise the users, needs. The experts working with the organisations within the consortium will identify users' needs by conducting interviews and surveys (electronic or printed). The responses are anonymised and irreversibly de-linked from user's identities.

4.2.3 Click-Through Data of Website

Since the click frequencies on the SAFETY4RAILS project website will be collected transparently but with an initial announcement of the purpose and expecting the visitor's consent, there will be a recruitment strategy. The visitors' responses to projects' online outputs will be collected for the purpose of assessing the level of interest on project results. The website statistics will be reported in an anonymised way. The responses are combined to provide aggregated analysis and clustering of webpage visits.

4.2.4 Usability Evaluation Data

The experts are the potential subjects to provide usability evaluation data by testing the developed tools and filling the questionnaires to assess and evaluate the usability of the project outcomes. The responses are anonymised, pooled and just overall findings are shared. The related data modality will

be in a questionnaire format where each subject will be asked to fill a form by presenting their profile. User profile data here may include level of education, gender ("Male", "Female" or "rather not say"), age, level of expertise in the financial and the business domain all of which are not sensitive data. The Person's identity will be anonymised at source and the data aggregated so as to enable aggregation of user sub-groups for clustering analysis of the perceived usability of the system e.g., as viewed by categories of users with similar/different gender/age/skills etc.

4.2.5 Social Network Data (secondary use)

To support simulation studies as well as for stakeholder and end-user group engagement, efficient outreach and best adapted dissemination of project results, some secondary social network data is obtained from the well-established social networks, namely Facebook, LinkedIn, and Twitter. This does include personal data; the elements include those commonly available on these media such as name (or in some cases pseudonym), surname, gender, native language, networked others and links and favourites (likes, social-network metadata (age, ethnicity etc) and it may have images and audio-visual material within or linked. In order to conduct stakeholder opinion analysis relevant to the expectations regarding the SAFETY4RAILS innovation results it may be necessary to de-identify some fields of such data through pseudonymisation whilst other fields may be anonymised or irreversibly masked/alerted or deleted e.g. here is a range of standardised irreversible transformation techniques available to prevent persons being re-identifiable in images/videos (scrambling, pixelization, blurring , masking) whilst still retaining some non-identifying information such as number of persons in a photo. Such techniques together with generalisation and randomised shuffling of the order of K-anonymised/pseudonymised fields can be used in conjunction with field deletion/irreversible anonymisation where possible to bring to bear a managed mix of multiple of data protection approaches on this secondary-use data type; as shown below in Table 6.

Based on the analysis of the typology of stakeholders and data types that could be processed within the SAFETY4RAILS project as outlined above, Table 6 below sets out the distinct categories of data (data modalities) and respective de-identifications approaches that could be deployed for each type as part of the data protection strategy to be determined by the Data Controller based on consideration of recommendations as per the Ethical Compliance Framework Console (Table 7) .

TABLE 6 DATA TYPE DE-IDENTIFICATION MEASURES

Data Subject's Personal Data	Safeguarding measures	Possible implementation
Name, surname, signature such as for example might appear on a consent form		Character masking of subject's name and surname, total (irreversible) masking of the signature
Person's image,	Selective or total privacy masking of the image	Reversible or irreversible masking, blurring, pixelization, scrambling etc
Profile such as age, gender, education, region, profession, rank Social-networks metadata such as the user's location, language spoken, biographical data, and/or shared links.	Clustering into as many subsets as practical, aiming to keep the number of subsets as high as possible for example, can have a cluster of 7 age subsets to give as high a K-anonymity as possible	K-anonymisation for age, region, property/house size, residence size, plus random shuffling of the order of the records and encrypted storage.
Usability evaluation	The questionnaire design follows the data minimisation and purpose limitation principle and avoid asking privacy sensitive and personal data related questions if at all possible, however, in any event. Considering the perceived value of a particular performance set of a system as a sort of	Generalisation and data aggregation, single or double pseudonymisation, plus random shuffling of the order of the records, encrypted storage

Data Subject's Personal Data	Safeguarding measures	Possible implementation
	subjectively evaluated Key Performance Indicator (KPI) as being influenced by a person's profile background such as education, gender, age, computer savvy, can use the mean opinion scoring principle to assess the satisfaction rates or we can generalise the opinions to arrive at an overall aggregated performance indication rather than a person specific indication.	
Metadata, network traffic metadata bearing the IP or MAC address of the data subject's device, mobile phone identifiers etc.	In order to protect the identity of the data subject's devices which might be involved in online evaluations such as the computer, MAC or IP address steps have to be taken to protect this data through masking.	De-identification through masking computer data in the metadata.

5 SAFETY4RAILS Essential Data Processing and Legal Basis

The consortium, within the framework of its quality planning has adopted a risk-aversive strategy to project management and in particular to ethical and data protection compliance as well as to ensure social responsibility and acceptability of the resulting innovation. This is further supported by its Governance structures as outlined in Chapter 8, whereby the Project Ethical Board oversees the coordination and management of ethical and data protection compliance involving the local DPOs, the project Data Controller, the Ethics manager and the Coordinator with periodic reporting to, and issue-specific consultation with the Ethical Advisory Board as required. This is to ensure that the innovation process remains fully compliant with GDPR and relevant ethical and data protection requirements of all jurisdictions involved.

Accordingly, in adherence to the 7 GDPR principles, in particular data minimisation and purpose limitation, the following specific guidelines shall form the basis of our ethical and data protection compliance assurance.

- The use of any personal data shall be kept to the minimum essential for the conduct of the SAFETY4RAILS innovation objectives to support the enhanced safety and security of Railways as Critical Infrastructure as is in the public interest.
- The rights and freedoms of the data-subjects are to be fully respected throughout the lifecycle stages of any personal data that may be deemed to be essential, starting from the pre-acquisition stage, i.e. the preliminary data-set and data-element, and finally data-subject selection and recruitment and subsequent treatment of data-subjects to ensure a “Healthy Consent” process is conducted and later on, de-identification/k-pseudonymisation at the point of sourcing and thereafter secure processing, storage, transfer and eventually deletion once the research and relevant essential studies are completed.
- Additional safeguards apply as mandatory precautionary measures to be undertaken in the case of data exchanged between EU and non-EU countries as described below.
- As far as possible any essential personal data processing is conducted locally and only the de-identified and/or aggregated results shared between partners
- Consistent with the Article 6 GDPR, processing of personal data is, in summary lawful: **a)** with consent of the data subject; **b)** performance of a contract to which the data-subject is a party or at the request of a data-subject prior to entering a contract; **c)** compliance with legal obligations to which the controller is a subject; **d)** necessary to protect the vital interests of the data-subject or of another natural person; **e)** 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; and/or **f)** 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. The processing of personal data in SAFETY4RAILS needs to be analysed as lawful under 1 or more of the lawful possibilities **a) – f)** before it is permitted to proceed.

5.1 Data Controller’s Ethical Compliance Framework Decision Console (ECFC)

Table 7 below represents a Data Controller’s Console in setting out the essential data processing involved in the SAFETY4RAILS project and the alternative or combined measures open to the Data Controller to ensure full compliance with the ethical and Data Protection Regulations; with all jurisdictions involved. By examining the GDPR closer, we can notice that when it comes to the question which lawful basis should be used when processing personal data in general, the most important parameters to take into consideration are:

i) The identity of the Data Controller,
ii) The Purposes of Processing, and,

iii) The Context of Processing.

It is clear that based on the above parameters, the Data Controller can decide which lawful basis to use for processing and what pre-requisite safeguarding measures have to be taken. This is supported by the approach taken throughout our analysis chain in this Deliverable which has led to the development of the Data Controllers' Ethical Compliance Framework Console as presented in this section.

TABLE 7 THE DATA CONTROLLERS' ETHICAL COMPLIANCE FRAMEWORK CONSOLE (ECFC)

The Purpose & Context of Data Processing that involve various levels of Personal Data Processing As per planned work to date	De-identification Steps	Updated Range of Options as Relevant Legal Basis & Data Protection Measures considered by the Data Controller -Responsive to Ps & Cs		
		Prior Explicit Content Seeking if possible and meaningful	Data Transfer	Legal Basis within GDPR and relevant Nation States Regulations
Personal Data: comprising name, email, gender, region, contact details **Other linked data:** type of organisation, **Acquisition purpose:** requirements elicitation and /or usability evaluation, engaging railway sector or experts in the domain **Context:** selection and consent seeking, recruitment and selection of data subjects, consent-seeking (as advised by Data Controller), interviews, stakeholder group workshops, securely processed online questionnaire, expert advisor consultation **Data types and formats**: documents, software generated data, railways infrastructure data, forms (NDAs, consent forms etc) **Anonymisation Status at Source:** not anonymised	Data is to be rendered de-identified through K-pseudonymisation,- anonymisation, scrambling the order of records, and generalisation through aggregation plus encrypted storage.	Explicit consent seeking per healthy consent process as advised by the Data Controller. (For information to be provided see e.g.: Art. 13. Art.14 Art. 22)	No transfer involved. However if this were to change then can apply Art. 46(3) a	See last bullet point under chapter 5 above, primarily foreseen: Art. 6.1.e OR Art. 9.2.J in accordance with Art 89
Personal Data: comprising log-in and access control data including account ID, email and password data **Other linked data:** possible affiliation data as per user registration **Acquisition purpose:** solely to access the S4RS platform **Context:** prior registration, subsequent user log-ins, authentication **Data types and formats**: text, character string, xml and json **Anonymisation Status at Source:** not anonymised, secure access encryption and online authentication	Data is to be rendered de-identified through K-pseudonymisation,- anonymisation, scrambling the order of records, and generalisation through aggregation plus encrypted storage.	Explicit consent seeking per healthy consent process as advised by the Data Controller. (For information to be provided see e.g.: Art. 13. Art.14 Art. 22)	No transfer involved. However if this were to change then can apply Art. 46(3) a	See last bullet point under chapter 5 above, primarily foreseen: Art. 6.1.e OR Art. 9.2.J in accordance with Art 89
Personal Data: comprising audio, images, video **Other linked data:** offline sensor data **Acquisition purpose:** for simulation exercises and testing of the system and aggregation and modelling, **Context:** personal data could be involved either as part of the input set or the output set of the system being tested depending on the requirements to be specified	Data is to be rendered de-identified through K-pseudonymisation,- anonymisation, scrambling the order of records, and generalisation through aggregation plus encrypted storage.	Explicit consent seeking per healthy consent process as advised by the Data Controller.	No transfer involved However if this were to change then can apply Art. 46(3) a	See last bullet point under chapter 5 above, primarily foreseen: Art. 6.1.e OR Art. 9.2.J in accordance with Art 89

The Purpose & Context of Data Processing that involve various levels of Personal Data Processing As per planned work to date	De-identification Steps	Updated Range of Options as Relevant Legal Basis & Data Protection Measures considered by the Data Controller -Responsive to Ps & Cs		
		Prior Explicit Content Seeking if possible and meaningful	Data Transfer	Legal Basis within GDPR and relevant Nation States Regulations
Data types and formats: media files (audio, video, images), xml, json files **Anonymisation Status at Source:** not anonymised	Audio/video data masking using appropriate techniques such as scrambling, blurring, pixelization etc as appropriate	(For information to be provided see e.g.: Art. 13. Art.14 Art. 22)		
Personal Data: comprising social network data (name, surname, address, age, gender, localisation, interests, preferences, photograph, network) **Other linked data:** not sought but potentially could be other linkable data **Acquisition purpose:** selection of relevant stakeholders for invitation to events and/or dissemination of project results **Context:** secondary usage of social network data (Facebook, LinkedIn, Twitter) **Data types and formats**: text, image, audio **Anonymisation Status at Source:** not anonymised	Data is to be rendered de-identified through K-pseudonymisation,- anonymisation, scrambling the order of records, and generalisation through aggregation plus encrypted storage Audio/video data masking using appropriate techniques such as scrambling, blurring, pixelization etc as appropriate.	Secondary data usage thus consent impractical to obtain,	No transfer affected However if this were to change then can apply Art. 46(3) a	See last bullet point under chapter 5 above, primarily foreseen: Art. 6.1.e OR Art. 9.2.J in accordance with Art 89
Personal Data: unlikely but possibly, a range of unexpected personal data may occur in the output received from a simulation tool should a cyber- attack have led to personal data breach **Other linked data:** none expected **Acquisition purpose:** simulated cyber-attack scenarios impact assessment **Context:** using synthetic, data to simulate cyber-attacks which may have led to personal data breach **Data types and formats**: As yet unknown depending upon what data will become available **Anonymisation Status at Source:** simulations will use synthetic and thus anonymised data but any unexpected personal data may not be anonymised	Data is to be rendered de-identified through K-pseudonymisation,- anonymisation, scrambling the order of records, and generalisation through aggregation plus encrypted storage. If any data were to comprise of audio/video data, irreversible masking will be applied using appropriate techniques such as scrambling, blurring, pixelization etc as appropriate.	Not applicable as any personal data that may arise would be unsought, unexpected secondary data	Data transfer to SecureRail Tool by STAM in Italy. encryption of data in-transit and data at-rest. Art.46 (3) a	See last bullet point under chapter 5 above, primarily foreseen: Art. 6.1.e OR Art. 9.2.J in accordance with Art 89.

5.2 Data Controllers' Compliance Strategy Decision Criteria

There are a number of criteria that would require consideration by the Data Controller (DC) in determining the appropriate compliance steps to be followed to ensure ethical and data protection compliance for a proposed data processing pipeline; as follows:

The status and Objective of the Data Controller or the Data Processor to whom the DC may have conferred *delegated responsibility* in respect of a specific (element of) proposed data processing (in which case the DC still remains responsible for all aspects of Data Processing) is a primary consideration amongst several; as follows:

i) The DC or the responsibility is a Public Entity (Authority/Institution) with specific public duties)
ii) The DC or the entity with delegated responsibility is a Research Organisation
iii) The DC or the entity with delegated responsibility is a Private Entity (enterprise).
iv) If i) is the case, then is it also the true, as deemed by the DC, that that the proposed data processing is consistent with the normal everyday duties of the public institution concerned as established in the Nation State Law?
v) If i) is not true is it the case that the proposed data processing constitutes legitimate interest as deemed by the DC
vi) If ii) is true then is it the case that, as deemed by the DC, the proposed data processing can be pursued subject to certain provisors based on nation State Law, on the grounds of it being for a scientific research purpose
vii) If Data is to be transferred to a non-EU country without an Adequacy Decision what specific relevant exemptions or derogations from requirements under the GDPR and the non-EU state regulations are provided for
viii) If Obtaining Consent is proposed as a legal basis then do the arrangements proposed provide assurance that a Healthy and/or Meaningful and Practicable consent seeking process will result?
ix) If the objectives of research could be viewed as being in the public interest and any personal data is processed as necessary for the performance of a task carried out in the public interest or in the exercise of official authority.
x) Any special category data can be processed only if absolutely essential to support the performance of public tasks in the public interest for scientific research purposes.

A range of public and private organisations are partners within the consortium, some may have public responsibilities including, crucially, responsibility for maintaining public safety. However, partners can verify their public task status in accordance with Art 6.2 and 6.3 GDPR as well as Rec. 45 GDPR where it is stated that Union or Member State law shall define whether the controller performing a task of public interest can be a legal person governed by public law or by private law.

The Project Data Controller and several other partners have a public institution status. For such partners, such data processing can proceed on the grounds that it supports the performance of a public duty. Processing certain personal data is deemed essential to support their responsibilities as charged with under the member state law and thus supporting certain aspects of the day-to-day running of their public services, for example, maintaining public safety.

Thus, for partners it is possible to perform essential personal data processing either under Art 6.1(e) GDPR as processing that is necessary for the performance of a public task, or under Art 9.2.j as processing necessary for archiving purposes in the public interest, scientific or for historical research purposes or statistical purposes in accordance with Article 89(1).

5.3 Obtaining Explicit Consent

5.3.1 The Requirements for Explicit Consent for Data Processing within the Innovation Activities

For the consent seeking process, where possible and meaningful as consistent with the provisions of Art.13. and Art. 14 GDPR, Explicit Consent shall be obtained including compliance with the principles of healthy consent by ensuring compliance with guidelines for recruiting and informing potential data-subjects. This is to avoid recruiting those who for any reason may be, or may feel, vulnerable, and for those selected, providing prior information about the extent and nature of any processing covering the requirements in Article 13/14 including Explicit Consent Article 6.1(a), 9.2(a) and 22(1) (b) for profiling strictly for research purposes that will have no legal effect whatsoever on the participants.

As far as possible and where practicable and meaningful, the Explicit Consent may be sought as a matter of compliance with additional ethical standards and procedural obligation consistent with Art 29 Working Party Guidelines on consent under Regulation 2016/679, p 28. In the context of operating under the provisions of Art 6.1(e) GDPR, although the consent of data subjects may be sought by a public institution, this could not be relied upon as the legal basis but rather a supplementary procedural obligation.

However in some circumstances such as for secondary use of social network data, it can be difficult or impossible to seek informed consent and here consideration needs to be given to the criteria for proportionality of efforts, practicality and meaningfulness of informed consent seeking[5]. Consistent with Article 5 1(b) GDPR subject to safeguards that the Data Controller will observe (Art. 89 GDPR) the partners can process such data where relevant exemptions or derogations from requirements under the GDPR are provided for within member state law. For example where consent cannot be obtained, where full transparency information cannot be provided, or where access rights could not be fulfilled. These exemptions are referenced in Art 89 GDPR; however, these take effect only within member state law. So, for example, for the UK, this is the exemption within the UK Data Protection Act 2018[6]. This exemption can also apply to restrictions on international transfers; however each partner will have to ensure that they have equivalent exemptions, within local regulations, that they can rely on. This exemption can also apply to restrictions on international transfers (and providing safeguards within Article 89 GDPR can be deemed to have been met in such context).

5.3.2 The Requirements for Explicit Consent for Data Processing within Project Management & Administrative Activities

The legitimate processing of the absolute minimum of personal data within the consortium for the purpose of the necessary communication, administration and project management tasks should proceed without the requirement for explicit consent. This is because processing of the minimum of personal data such as emails, address and affiliation could be readily seen as essential for the implementation and management of the project for which all the partners share the signed contract, including data protection obligation, and thereby are committed to cooperation in the joint endeavour which requires the mutual ability to use such minimum of personal data to support communication and management of the project. Such processing of data within the consortium does not extend to the use of any personal data that falls outside the definition of “legitimate use for the conduct of the project and its management”; In particular, it does not extend to the use of sensitive personal data such as special category data as defined by GDPR -irrespective of whether it is limited to project partners or not.

[5] An Analysis of the Consequences of the General Data Protection Regulation on Social Network Research, Dec 2019

[6] Guide to the General Data Protection Regulation (GDPR) Exemptions (UK-DPA-2018)

However, the above minimum use of personal data (email, address, organisational affiliation) for legitimate purpose also extends to:

i) Such minimal personal data of "project advisors"/"potential advisors"/specific external collaborators who by virtue of either having already signed an NDA/MOU/Consent Form or being already known to one of the partners would be legitimately contactable by the members of the consortium specifically for purposes related to the project;
ii) The use of the essential minimum of personal data (name, email, address) merely to invite data-subjects established as candidates to participate in stakeholder meetings etc by virtue of having a specific sectoral expertise and role and having been selected a such as a potential participant for system requirements and/or evaluation for which they can subsequently be invited to consider signing a consent form; given all the required project information.

5.3.3 Processing Profiling Data

The use of any profiling data is to be kept to the absolute minimum that can be justified as essential for the implementation of the necessary innovation tasks as per previously established in the Description of Action (DoA). Any profiling data shall not include any sensitive personal data and shall be strictly access controlled and anonymised, de-identified or K-pseudonymised at-source and as such (irreversibly) de-linked from the rest of the data. Such data is solely to be processed for aggregated analysis and clustering of the needs of user sub-groups and shall be deleted once it is no longer needed to serve the stated innovation objectives. Such data is solely to be processed for aggregated analysis and clustering of the needs of user sub-groups and shall be deleted once it is no longer needed to serve the stated research objectives.

In the SAFETY4RAILS project such data shall be processes within the following exclusive purpose and context categories:

- **Users' system requirements viewpoint clustering** (professional role & responsibilities area, relevant sectoral services expertise, gender, age)
- **Users' system usability viewpoint clustering** (professional role & responsibilities area, financial services sector expertise, relevant Fintech operations skills, gender, age)
- **Disseminations and Communication with Stakeholder Groups** (professional roles, interests and affiliations)
- **Scenario Simulation, in particular modelling cyber-threats** and developing and testing attack scenario simulation and impact assessment tools
- **User registration information** (login data such as name, password) as required to enable user registration

5.3.4 Lawful Data Transfer

The following safeguarding guidelines have to be carefully considered so that an adequately safeguarded procedure is implemented to ensure ethical and data protection compliant with GDPR and the non-EU nation State; as follows:

- As far as possible data are to be locally processed on a role-based secure access basis by specific persons responsible for the relevant tasks. Accordingly, such data are processed within the domain of each relevant Data Controller for clustering analysis and data aggregation; as such, this shall provide the required data protection measures consistent with GDPR.
- As for any data transfers involving partner organisations located within non-EU countries; the following guidelines need to apply at all stages of data transfer, particularly regarding minimisation and de-identification:
    - As an adequacy decision with respect to two of the non-EU partners namely Australia and Turkey are currently pending, any data transfer involving the partners from the above two countries shall be strictly limited to only the exchange of data that is encrypted pre-transfer and whilst stored (i.e., encrypted in-transit and at rest) under the responsibility of the local DPO reporting to the respective Data Controller. This is a mandatory pre-requisite; all parties involved must ensure that the processing remains

compliant with GDPR requirements as well as with the regulatory requirements of any third countries involved in the data transfer.
    - Data transfer shall be minimised and limited to only that which is deemed by the Data Controller to be essential for the objectives of the joint tasks and responsibilities for the collaborative system development and performance evaluation.
- Data shall be irreversibly anonymised.
- The data to be transferred shall be inspected to be verified as synthesised data and/or irreversibly anonymised.
- Formal Notification by all partners involved shall be provided to respective local Data Protection Officers confirming the name of the data file inspected, its anonymisation status verification and date, the name of the actual data file transferred and the date of its transfer which is thus approved for transfer in encrypted form.
- The planned exchanges of data shall be for the limited duration of particular phases of the project on a non-repetitive basis and for the sole purpose of research which is permitted within the provisions of GDPR on a legal basis as deemed appropriate by the Data controller such as on the grounds of performing a public task (Art. 6.1.e GDPR), or for legitimate interest of Data Controller (Art.9.2.j GDPR) or by obtaining the Explicit Consent of the data-subject (Art. 6.1.a GDPR)
- On the condition of adherence to the above compliance assurance process, such a transfer is permitted without further provisions as would be otherwise mandated by GDPR - This shall be ensured through strict monitoring locally and at consortium level within the SAFETY4RAILS Governance Structures Framework as outlined in Chapter 7.
- The categoric safeguard basis for such transfers would seem to be Article 46(3) (c) namely “Standard Contractual Clauses” but as it is planned to use only anonymised data in such data exchanges, GDPR would not apply. However, this still requires an assessment of any risks that may be posed to data-subjects specific to the Third country involved. Indeed, the consortium recognises that in circumstances whereby the activities undertaken in non-EU countries may be deemed to raise ethical issues, it must ensure that the research conducted outside the EU is legal in at least one EU Member State.

5.3.5 Ethical and Social Considerations

In all Data Processing within a collaborative project involving both EU and non-EU countries, it is mandatory that the legal and data protection regulations of all jurisdictions involved and SAFETY4RAILS partners from Non-EU countries shall ensure that any data processing conducted by them remains compliant with the data protection requirements of their country which in almost all cases should be consistent with GDPR requirements.

The scope of responsibility from a GDPR viewpoint is only for the data processing actions that are to be taken for innovation purposes within the project and do not extend to the legal basis of the data processing actions that future users of the solution system may perform. Of course, any such innovation requires both legal compliance assurance as well as social acceptability which demands methodologically-guided and socio-ethically reflective co-design to be followed as an integral part of the SAFETY4RAILS Ethical Compliance Framework (ECF). Within this framework, SAFETY4RAILS also addresses the societal impact which relates to the PESTEL criteria (Political, Economic, Social, Technological, Legal, Ethical and Environmental) and for which the initial SIA analysis is presented in Chapter 7. This will include consideration of how the data, in particular any repurposed data (such as the social media data) will be protected from misuse, any uses outside of the research purposes, any uses that could potentially expose citizens to the risk of any harm and/or hurt, including physical or psychological e.g. inequitable treatment, stigmatisation, isolation, bias and discrimination.

6 Healthy Explicit Consent Process Documentation

The following is to serve as a general template to help the structuring of a Consent Form consistent with GDPR requirements (Art. 6.1.a GDPR, Recital (40)-(43), information to be provided: Art. 13, Art.14) to serve the particular requirements for consent seeking in respect of each of the three Purposes and Contexts of signal processing within SAFETY4RAIL (survey/ interview involving stakeholders, profiling, data transfer) as deemed necessary and accordingly as advised by the Data Controller. As such this "consent form constructor" template has a more extensive structure than the consent forms derived from it. The "consent form" is in fact a "concept form kit" hence it looks denser than any single concept form that could be derived from it.

The sections on the form appearing in green/blue/purple font define the three distinct purposes and contexts of data processing as planned to-date within the SAFETY4RAIL project. This consent form constructor template shows a mapping across from each of the GDPR-specified data-subject rights to be included (as appear on the left-hand column) to the type of corresponding information the form should have, as necessary, in the righthand column to be adequately responsive to each of the rights according to GDPR. The relevant consent form can be instantiated in each case by simply keeping the paragraphs describing the SAFETY4RAIL project objectives and the relevant purpose and context of processing to also include the data transfer and deleting each paragraph covering the purpose and context that is irrelevant to the particular consent being sought on each occasion as required.

This form is subject to updates to accommodate any new purpose and contexts of data processing as may arise or simply to continuously improve the form responsive to feedback from consent seekers, data-subjects and as may be advised by the Data Controller. This form is to be submitted to any potential "data-subject" prior to the intended date of data acquisition; giving adequate time for the request to be properly considered. It is to be translated into their language along with the translation of the project information pack. The potential "data-subject" is thus invited to consider the consent request, the reasons for the requested data and the undertakings of the consortium with respect to data processing and protection, and to freely state simply whether they would wish to participate in the process or not with no explanation being needed for their decision

For each type of data being requested and for each purpose a separate such form has to be submitted to the data-subject ideally at least two weeks prior to their expected date for filling the form to provide Explicit Consent or reject the request. Should further information be required, the Data Controller can be contacted as detailed on the form.

Appropriate procedures shall be followed for correct recruitment of potential data-subjects to ensure an effective consent seeking process fully compliant with the GDPR requirements as monitored and advised by the Data-Controller. This form and the accompanying project information pack have translated versions covering the 6 other partner languages (French, Finish, German, Greek, Italian, Turkish).

TABLE 8 THE ENGLISH VERSION OF THE SAFETY4RAILS CONSENT FORM

SAFETY4RAILS Consent Form & Process Feedback Template for Participation in Questionnaire/Interview/Focus Group

***Age and/or Vulnerability Eligibility Criteria:* To be eligible for this case study you must be over 18 years old and not be in a situation of any vulnerability due to physical /mental conditions and/or have any perceived obligation to agree; if you are under 18 and/or perceive any pressure whatsoever to take part in this study please do not proceed further as we are not permitted to invite you to participate in this and agree to anything under such circumstances- Thank You**

Date:

Our Project Name: SAFETY4RAILS

Project Data Controller:
Mr Antonio De Santiago Laporte

Data Controller's address:
Metro de Madrid, c/ Cavanilles 58, 28007 Madrid,

T +34 913790263,
e-mail
antonio.desantiago@metromadrid.es

Data controller's Phone Number:

T +34 913790263

The Information you need includes this form plus the project information pack all translated in your native language and of course you can seek further clarification from the Data Controller via the above contact details.

The project website can be reached at:
https://safety4rails.eu/

Railways and Metros are safe, efficient, reliable and environmentally friendly mass carriers, and they are becoming an even more important means of transportation given the need to address climate change. However, being such critical infrastructures makes metro and railway operators as well as related intermodal transport operators attractive targets for cyber and/or physical attacks. The SAFETY4RAILS project delivers methods and systems to increase the safety and recovery of track-based inter-city railway and intra-city metro transportation. It addresses both cyber-only attacks (such as WannaCry infections), physical-only attacks (such as the Madrid commuter trains bombing in 2014) and combined cyber-physical attacks, which are amongst the emerging threat scenarios, given the increasing IoT infrastructure integration.

SAFETY4RAILS concentrates on rush hour rail transport scenarios where many passengers are using metros and railways to commute to work or attend mass events (e.g., large multi-venue sporting events such as the Olympics). When an incident occurs during peak times, metro and railway operators have to consider many aspects to ensure passenger safety and security, e.g., carry out threat analysis, maintain situation awareness, establish crisis communication and response, and ensure that mitigation steps are taken and communicated to travellers and other users.

SAFETY4RAILS will improve the handling of such events through a holistic approach. It will analyse the cyber-physical resilience of metro and railway systems and deliver mitigation strategies for an efficient response, and, in order to remain secure given ever-changing emerging risks, it will facilitate continuous adaptation of the SAFETY4RAILS solution; this is validated by two rail transport operators and the insights arising from validations will inform the re-design of the final prototype.

The Specific Purpose of this Research Study in which you are invited to consider participating and/or the research purposes and legal basis of any automatic data capture (if included in the study) and the extent of any profiling and/or justification for and notification of any transfer of data

The Specific Data Requested is your prioritised requirements, the use-cases and the resulting usability features of the proposed system as presented. This is needed for the purpose of user-centred co-design of the solution system to be developed and validated as described above for the Purpose of Enhanced Safety and Security of the Railway Systems and Passengers. Thus, the objectives of the research are in the public interest and compliant with Art. 6.1.e GDPR or Art. 9.2.J in accordance with Art. 89 9(1)

SAFETY4RAILS Consent Form & Process Feedback Template for Participation in Questionnaire/Interview/Focus Group	
to any other entities including in particular to organisations in Third Countries (without data protection provisions deemed equivalent to EC) and your rights to ensure full transparency on the basis of an Informed Consent Process as mandated by the General Data Protection Regulations (GDPR). All the above to be explained, fully, clearly and specifically	to respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the "data-subject" The type of data that we seek through the questionnaire/interview/workshop is needed for usage over the lifecycle of the project and for long enough afterwards to support follow-on research as deemed essential. To support collaborative work, there may be a need for data transfer to and from our EU-partners to non-EU partners in Switzerland, Turkey and Israel. In accordance with Art 46 (3) (a) - Standard contractual clauses shall be established between the Data Controller (EU exporter) and their counterpart Data Controller (importer – third country) to stipulate the legally binding safeguards for the protection of personal data even although data shall be encrypted in-transit and at-rest. The Specific Data Requested is your assessment of the performance and usability of the solution system as deployed for trialling purposes and user-centred evaluation and responsive refinement of the system design features. This type of data is needed for user-centred co-design of the solution system to be developed and validated as described above for the Purpose of Enhanced Safety and Security of the Railway Systems. Thus, the objectives of the research are in the public interest and compliant with Art. 6.1.e GDPR or Art. 9.2.J in accordance with Art. 89 9(1) to respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject. The type of data that we seek through the questionnaire/ interview/workshop is needed for usage over the lifecycle of the project and for long enough afterwards to support follow-on research as deemed essential. To support collaborative work, there may be a need for data transfer to and from our EU-partners to non-EU partners in Switzerland, Turkey and Israel. In accordance with Art 46 (3) (a) - Standard contractual clauses shall be established between the Data Controller (EU exporter) and their counterpart, Data Controller (importer – third country) to stipulate the legally binding safeguard for the protection of personal data even although data shall be encrypted in-transit and at-rest. The Specific Data Requested is profiling data of the potential end-users as needed for secure access to the right tool to provide a range of suitably adapted system functionalities (e.g., simulation of cyber-physical incidents/attack and anomalous flows detection) as part of the solution system to be developed and validated to promote the enhanced safety and security of the Railways. Thus, the objectives of the research are in the public interest and compliant with Art. 6.1.e GDPR or Art. 9.2.J in accordance with Art. 89 9(1) to respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject. To support collaborative work, there may be a need for data transfer to and from our EU-partners to non-EU partners in Switzerland, Turkey and Israel. In accordance with Art 46 (3) (a) - Standard contractual clauses shall be established between Data Controller (EU exporter) and Data Controller (importer – third country) to

TABLE 8 THE ENGLISH VERSION OF THE SAFETY4RAILS CONSENT FORM

SAFETY4RAILS Consent Form & Process Feedback Template for Participation in Questionnaire/Interview/Focus Group	
	stipulate the legally binding safeguards for the protection of personal data even although data shall be encrypted in-transit and at-rest.
You have the right to complain through your national Data Protection Authority (or by contacting the Data Controller responsible at the address provided above) about any aspect of the conduct of this consent seeking process and of course the right to accept or reject our invitation without any explanation whatsoever, However should you decide to participate you would continue to have certain rights under the data protection law which are: • Withdraw your consent, for example if you opted in to be added to a participant register • Access your personal data or ask for a copy • Rectify inaccuracies in personal data that we hold about you • Be forgotten, that is your details to be removed from systems that we use to process your personal data • Restrict uses of your data • Object to uses of your data, for example retention after you have withdrawn from a study	A list of relevant names and addresses of Data Protection Officers to be provided, The name and address of the person in charge of this particular consent process to be provided at the end of this form The rights of data subjects to full information under the provisions of Articles (13) and (14) of GDPR to be set out in full. For this to be a "Healthy Consent" Process, it is mandatory that there exist no factors directly or indirectly compelling your acceptance of this invitation to participate. *You are entirely free to reject this invitation without any reason for your rejection or any consequences arising from it.* Any participant's data will be rendered de-identified or pseudonymised securely stored for only as long as it is essential to enable the research planned during the lifecycle of the project and the relevant follow-on study as deemed necessary for the purpose of this research. Only the minimum data justified as essential for study will be asked for and its use will be strictly limited for the purpose clearly stated above. All data shall be subject to strict Data Protection as planned for and monitored by the Data Controller Antonio De Santiago Laporte, (contact details provided above on this form). Despite our collective commitment to full compliance assurance with GDPR and local data protection regulations, in the case of any unexpected data breach, were this in anyway to have exposed your data to any privacy protection risks, the respective Data Protection Officer and the Project Data Controller shall be notified in accordance with Articles 33-34-GDPR and you will be formally advised of the data affected and the preventative action taken to ensure that any breaches will be fully investigated to establish cause and prevent recurrences consistent with Art 19, 35, 17. Any consent given can be withdrawn at any time without any explanation. Participant's data can be deleted, subject to the provisions of Art. 17 GDPR, at any time upon request. Whilst of course we would endeavour to uphold your rights, however we may face some restrictions that would apply to the above rights where data is collected and used for research purposes. You should take sufficient time to consider the invitation and make your decision when you are satisfied that you have fully understood the nature of the data processing. The project website https://.Safety4Rails/) and the Information pack in your language provides more information. However, should you require further clarification please contact the Project Data Controller through email: antonio.desantiago@metromadrid.es
Consent/Reject Participation:	**Consent/Yes** **Reject/No**

TABLE 8 THE ENGLISH VERSION OF THE SAFETY4RAILS CONSENT FORM

SAFETY4RAILS Consent Form & Process Feedback Template for Participation in Questionnaire/Interview/Focus Group

Your Considered Decision:

I was informed about the purpose of the project and the specific purpose of the research study in which I have been invited to participate.	Yes ☐	NO ☐
I was given the opportunity to ask questions about the project and about this study.	Yes ☐	NO ☐
I was made aware that I could withdraw from the study at any time for any reason including in case of an unanticipated incidental finding arising from my participation.	Yes ☐	NO ☐
I have been reassured that after analysis of the minimum information sought and processed, any personal data of mine shall be fully protected and deleted after the conclusion of the study.	Yes ☐	NO ☐
Yes, I comprehend the information above and consent to participate in the focus study; Yes ☐ Date:	No, I do not consent to participate in the focus group study; NO ☐	

Should you have any questions, please call or write to the following contact persons

Data Controller Antonio De Santiago Laporte

Metro de Madrid, c/ Cavanilles 58, 28007 Madrid, T +34 913790263, e-mail antonio.desantiago@metromadrid.es

The SAFETY4RAILS Coordinator: Stephen Crabbe

stephen.crabbe@emi.fraunhofer.de

Fraunhofer Institute (EMI), Am Klingelberg 1 · 79588 Efringen-Kirchen · Germany, Tel: +49 (0)7628 9050 645

7 Methodologically-Guided Social Impact Analysis

Until perhaps the 1970s, there was no doubt in the minds of policymakers that investment in research and development would have nothing but a positive impact on, for example, work life, housing, clothing, food, health, communication of methods of transportation – even the length and quality of life itself. Some urged that any investment in science is inherently good for the society.[7] However, gradually, when businesses and corporations were faced with consumer demands for corporate social responsibility a paradigm shift was starting to emerge: actions, no matter how good or evil, have consequences, and often they are unintentional or unforeseen.[8]

This realisation of the notion that actions are followed by impacts, now a seemingly obvious thing, spread from business life into government activities, for example, in the form of Regulatory Impact Assessment (RIA)[9], and more generally, including all sorts of R&D activities.

Without going into too many details on the history of different impact analyses, we can conclude that the Societal Impact Assessment (SIA) of SAFETY4RAILS is a manifestation of a tradition of assessing outcomes of actions. SIA is not the only assessment of SAFETY4RAILS impacts, nor is it the first. Actually, there are some desirable impacts that have already been acknowledged in the very beginning of the SAFETY4RAILS R&D process, i.e., in the proposal phase. It is stated in the proposal

"[…] the SAFETY4RAILS project is centred around ensuring positive societal impact through enhancing security for EU citizens by promoting resilience of railway systems against physical, cyber and combined cyber-physical hazards. In line with relevant EU policies, the activities will aim to reduce the loss of human life, health, environmental, economic and material damage from natural and man-made disasters, such as cyber-physical threats. Thus, it is clear that the intention is to have a very positive impact on the society: save lives and preserve the environment." [project number 883532-H2202 https://safety4rails.eu/]

Why then a specific SIA, one might ask. Maybe, because it gives the opportunity to paint with a wide brush the possible impacts of the project other than what was detailed in the proposal Description of Action? Alternatively, maybe because conducting a SIA enables a narrowing down, specifying and concretising both hopes and worries of the project outcomes. Perhaps the best justification for SIA is that it is a process that lives within the project and is accomplished jointly with those participating in the project. Thus, changes can be made when needed, new worries can be considered etc. The aim is to focus on identifying, analysing, monitoring and managing the intended and unintended social consequences, risks and changed processes resulting from the SAFETY4RAILS project, whether they are from the project itself or from the tangible outcomes of the project, not forgetting the future use of the solutions. SIA is, thus, more than just predicting impacts in a regulatory context as in the above-mentioned RIA. In fact, a SIA covers a much wider perspective than traditional impact assessments in that it focuses solely on economic, legal or environmental impacts, or assessments that focus on the measurement of the impacts retrospectively. On the contrary, a SIA is an active process of managing the social aspects of development. The benefit of a SIA is that when impacts are identified in advance, better decisions can be made regarding how the project should proceed and how the outcomes and their use should be. Following this, mitigation measures can be implemented, to both minimise the harm, and

[7] Bush, V. (1945). Science: The endless frontier. [A report to President Truman outlining his proposal for post-war U.S. science and technology policy.] Washington, DC: United States Government Printing Office. Quoted from Bornmann, L. (2013), What is societal impact of research and how can it be assessed? A literature survey. Journal of American Society of Information Science Technology, 64: 217-233.

[8] See for example, Hill, S (2016). Assessing (for) impact: future assessment of the societal impact of research. Palgrave Commun 2, 16073.

[9] See for example, OECD (2020). Regulatory Impact Assessment, OECD Best Practice Principles for Regulatory Policy, OECD Publishing, Paris, https://doi.org/10.1787/7a9638cb-en.

maximise the benefits.[10] Pivotal too, is the respect for human rights; they should underpin all actions.[11] Overall, a SIA possesses some common aspects with risk management of the project and its outcomes.

The process of the SIA of SAFETY4RAILS is to prepare a minimum of three main impact assessment tasks during the project execution. The first is the Initial Societal Impact, which is typically carried out during the first six months of the project. This provides not only initial guidance and information for the developers and the whole consortium, but also stirs the persons participating in the project to think about the consequences and impacts this project can have. The results of the initial SIA are presented in this document. The second phase would be analysis of the requirements and/or scenarios defined by the project from the Societal Impact and acceptability perspective in order to provide further guidance and recommendations for the developers. Finally, the third phase is the Final Societal Impact Review. It summarises the societal impact issues that have been raised in the previous assessments and describes how they have been addressed during the project. It should also mention the potential Societal Impact issues facing the deployment of the solution.

Typically, the contents of the social impacts touch on the following aspects of the society:

1. **Way of life, fears and aspirations** (e.g. how people live and interact with each other, their perceptions about their safety and the safety of their communities, future aspirations…);
2. **Culture and community** (e.g. shared beliefs, customs, values and languages, as well as the cohesion, stability and character of their communities…);
3. **Political systems** (e.g. participation in the decisions and processes that affect peoples' lives, the nature and functioning of democratic processes, and the resources available to support peoples' involvement in these…);
4. **Environment** (e.g. clean air, water, and other natural resources and access to them, as well as the level of exposure to pollutants and harmful substances and the adequacy of sanitation…);
5. **Health & well-being** (both physical and mental well-being…); and
6. **Personal and property rights** (e.g. economic effects, civil rights and liberties including privacy, personal disadvantages…).[12]

For the initial SIA of SAFETY4RAILS, the content was collected from consortium partners via an online tool called EUSurvey, provided by the European Commission. A link to the survey was sent by email to all partners as well as the so-called end-users using the predefined email-list that the project management had created. Altogether 16 respondents replied in January 2021, all anonymously, to three questions (in addition to the two questions related with consent for participation and related ethics and privacy issues).

The first question was: "Please write down any possible ethical challenges that you consider may emerge and/or negative impacts on individuals or societies you foresee arising with increased safety and security to railway systems, especially concerning cyber and/or physical attacks. You have the option to put as many as you want, specify order of importance, or highlight a specific one etc." The second question was: "Please write down how you think the responses you provided above to ethical challenges could be resolved and/or how the problems could be mitigated." The third and final question was: "What in your opinion could be the possible positive outcomes or impacts to individuals or societies when increasing safety and security to railway systems, specifically related to cyber and/or physical attacks?" The replies were then analysed and grouped into like themes during several so-called brainstorming sessions at Laurea. Below is an example from the online question.

Below are the results of the initial SIA in brief. This analysis provides to the consortium and relevant partners realistic information regarding ethical challenges and potentially negative impacts as a result of

[10] Morrison-Saunders, A., Bond, A., Pope, J. & Retief, F. (2015). Demonstrating the benefits of impact assessment for proponents, Impact Assessment and Project Appraisal, 33:2,108-115.
[11] Kemp, D. & Vanclay, F (2013). *Human rights and impact assessment: clarifying the connections in practice*, Impact Assessment and Project Appraisal, 31:2, 86-96.
[12] Vanclay, F., and Esteves, A.M (Eds.) (2011). *New directions in Social Impact Assessment. Conceptual and Methodological Advances*. Cheltenham (UK).

the project work; to be avoided by their early consideration. This analysis can be used as a guide for developers and other project workers to avoid creating solutions for the SAFETY4RAILS project that could create these negative outcomes.

Table 9 below represents a compacted display of the SIA Questionnaire as presented on single page.

TABLE 9 THE SIA QUESTIONNAIRE

Disclaimer

The European Commission is not responsible for the content of questionnaires created using the EUSurvey service - it remains the sole responsibility of the form creator and manager. The use of EUSurvey service does not imply a recommendation or endorsement, by the European Commission, of the views expressed within them.

Dear SAFETY4RAILS Societal Impact Assessment (SIA) participant,

Before starting the SIA, we ask you to carefully read the **Information sheet** (under Background documents, see the column on the right), confirm that you have read the document, and give your consent for the participation of this survey.

Please tick the related boxes before beginning with your answers.

* I have carefully read the information sheet about the project, about SIA, about my participation, and my rights as a participant of the SIA, including the data protection.

at most 1 choice(s)

☑ Yes

* I hereby give my consent for my participation in the SAFETY4RAILS Societal Impact Assessment.

at most 1 choice(s)

☑ Yes

And now, let's begin with the SIA.

And now, let's begin with the SIA.

Thank you very much for taking the time to answers questions on possible ethical challenges and societal impacts when developing solutions to enhance the safety and security of railways.

In order to answer the questions, we would like to ask you first to imagine solutions that enhance passenger safety and security of track-based railway systems related to both cyber and physical attacks (such as the Madrid commuter trains bombing in 2014) and combined cyber-physical attacks, which is an important emerging scenario given the increasing IoT infrastructure integration.

Picture in your mind solutions that enhance the safety and security of passengers, or solutions that prevent or mitigates attacks and their effects.

You do not need to know what the solutions are, how they work, or what their functions are in practice,

It is more critical that you use your imagination regarding the overall effects of such potential solutions or just specific aspects of the solutions.

Now, after some moments of reflection when hopefully several ideas have come to your mind, about potential solutions you could imagine; please reply to our three open ended questions, and once again, thank you for your valuable contribution.

1) Please write down any possible ethical challenges that you consider may emerge and/or negative impacts on individuals or societies you foresee arising with increased safety and security to railway systems, especially concerning cyber and/or physical attacks. You can mention as many aspects as you want, specify these in the order of importance, or highlight a specific one etc. [1000 character(s) maximum]

I could imagine that the solution may require some form of cyber-physical security monitoring including perhaps passenger profiling, crowd monitoring, possibly cctv surveillance and as such may lead to exposure to privacy protection risks if not appropriately safeguarded.

2) Please write down how you think the problems you stated in response to question 1 above could be avoided/resolved and/or mitigated. [1000 character(s) maximum]

At the earliest possible stage in the design and development, a complete list of data types planned to be used for each processing pipeline has to be provided and the justification given as to the reason each data element is needed and how each data type is intended to be processed. In this way a privacy risk analysis could be performed leading to recommendations as to how the required data protection safeguarding measures could be implemented as an integral part of the solution.

3) What in your opinion could be the possible positive outcomes or impacts to individuals or societies with increasing safety and security of railway systems, specifically related to cyber and/or physical attacks?

7.1 The Concerns

The Table 10 below identifies from the survey responses a total of 15 clusters that represent all participants responses. The clusters represent common themes into which the participants' responses were grouped. The column on the left, shows the number of times this theme occurred in the total participants responses. The first survey question asked for the possible ethical challenges that may emerge and/or negative impacts on individuals or societies arising from increased safety and security to railway systems, especially concerning cyber and/or physical attacks. Respondents were free to express as many answers as they wished or leave it unanswered. A total of 14 participants provided responses to the first question. Below they are presented in the order of prevalence, together with the related aspect of societal impact. It must be also highlighted, that the prevalence of a concern does not necessarily mean that any of the other ethical challenges or negative impacts are not as valid or possible.

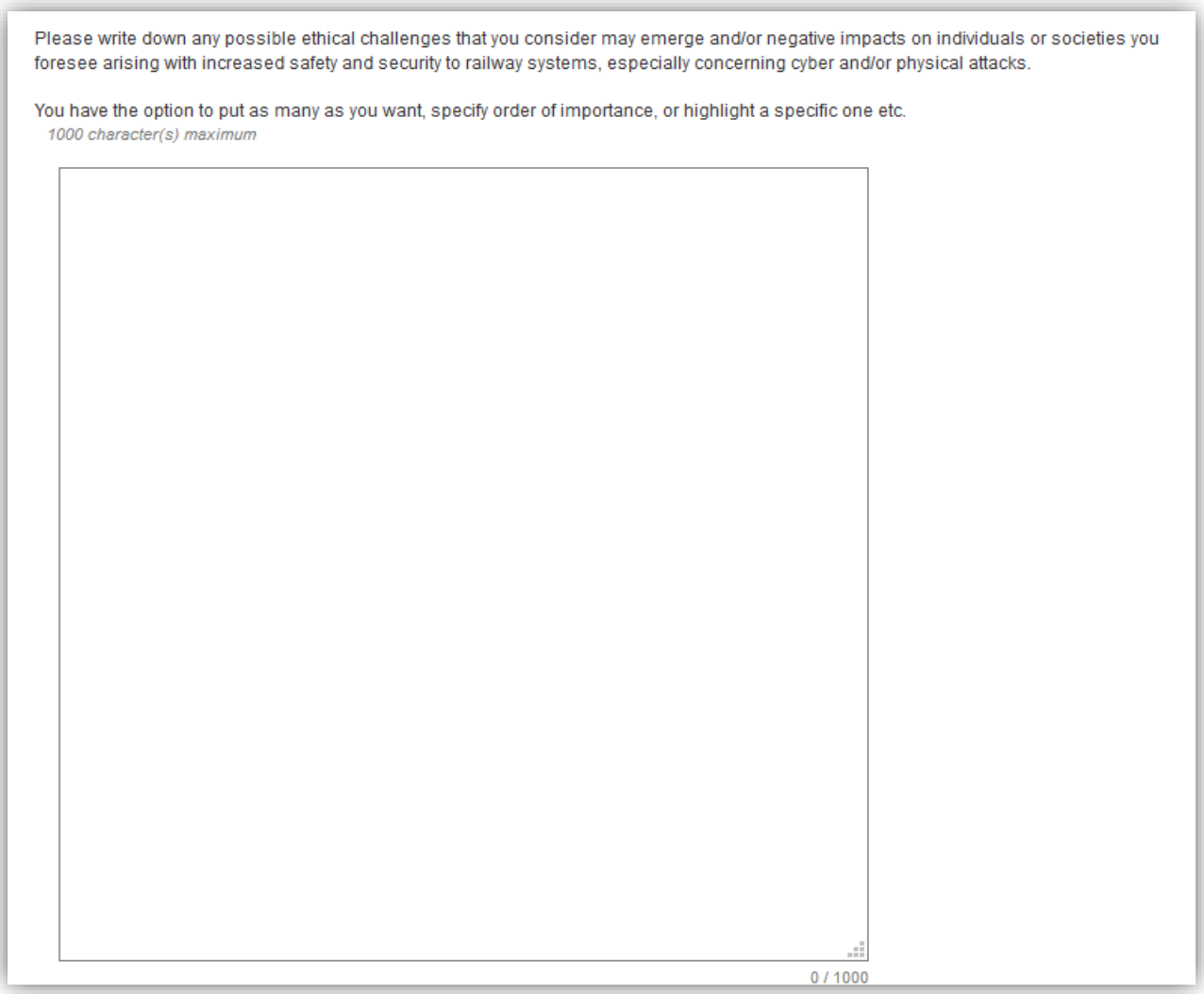

FIGURE 1 SCREEN CAPTURE FROM THE ESURVEY

TABLE 10 ACCEPTABILITY CONCERNS I

Prevalence	Clusters of Concern	Aspects
12x	Loss of passenger privacy	Way of life, fears and aspirations Culture and community Personal and property rights
5x	Heightened passenger fear of wrongdoing	Way of life, fears and aspirations Culture and community and Health & well-being
4x	Loss of personal data via cyber attack	Way of life, fears and aspirations Culture and community Health & well-being Personal and property rights
5x	Distrust in integrity of ICT systems personnel	Way of life, fears and aspirations Culture and community and Personal and property rights
3x	Distrust in ICT systems	Way of life, fears and aspirations Personal and property rights
2x	Loss of company data via cyber attack	Way of life, fears and aspirations Personal and property rights
2x	Inappropriate information management during and after a crisis situation	Way of life, fears and aspirations Culture and community and Personal and property rights
1x	Loss of rail passengers	Way of life, fears and aspirations Environment
1x	Loss of freedom of movement	Way of life, fears and aspirations Culture and community Political systems Personal and property rights
1x	Prioritisation conflicts between rail operator and state-owned investigative agencies	Personal and property rights
1x	Political motivation potentially conflicting with railway operating directives	Personal and property rights
1x	Problem resolution of one security risk creating a risk in a different area	Way of life, fears and aspirations Health & well-being Personal and property rights
1x	Loss of security via cyber attack	Way of life, fears and aspirations Health & well-being Personal and property rights
1x	Risk of too much technological interdependency increasing security risks	Way of life, fears and aspirations Health & well-being Personal and property rights
1x	Distrust in rail operator company/ government	Way of life, fears and aspirations Personal and property rights

In short, the loss of passenger privacy was mentioned as a concern by nearly all respondents. Thus, this aspect must be addressed in some way or another in SAFETY4RAILS. Also, the heightened fear of passenger wrongdoing was mentioned in a substantial number of answers. These answers were interpreted to mean that some less confident passengers might have a persistent fear of being arrested or otherwise prosecuted for even accidentally violating new rules. Some other passengers might feel that they are constantly under surveillance and being suspected of being capable of committing serious crimes. As a result of this persistent fear, it could be inferred that passengers may come to believe that the government and/or transit system operator does not trust the average citizen. Along with concerns of privacy, the fear of loss of personal data via a cyber-attack was expressed several times. Also, distrust in the integrity of ICT systems was mentioned

in roughly a third of the answers. Further, the loss of company data via cyber-attack was a concern, all making the data security and data privacy an area of major concern.

Naturally, this problem also needs to be addressed. Several concerns received few or only a single mention. Some were related to the management of railway safety and security, such as inappropriate information management during and after a crisis situation, problem resolution of one security risk inadvertently creating a risk in a different area, loss of security via cyber-attack, even a risk of too much technological interdependency increasing security risks was mentioned. Likewise, higher level political and/or societal concerns were expressed via the potential for prioritisation conflicts between rail operator and state-owned investigative agencies and political motivation potentially conflicting with railway operating directives, and general distrust in rail operator/government, resulting ultimately in a loss of rail passengers (to another means of commuting) and loss of freedom of movement.

7.2 Mitigation

The second question of the survey question asked how to solve and/or mitigate the issues raised in the first question. A total of 14 participants provided responses to this question, and as with the first question, the prevalence of certain types of solutions or types of mitigation does not mean that they are correct, sufficient, or most pivotal. Below are the presented the answers in their clusters.

TABLE 11 ACCEPTABILITY CONCERNS II

Prevalence	Clusters of Solutions	Aspects
5x	PR Campaigns and methods that build mutual trust between society and governance/rail operators and demonstrate data privacy and protection as a priority	Way of life, fears and aspirations Culture and community Political systems Personal and property rights
4x	Internal protection of sensitive personal data	Way of life, fears and aspirations Culture and community Personal and property rights
3x	External Data security protection	Way of life, fears and aspirations Culture and community Personal and property rights
3x	Clearly announce and/or display all passenger expectations	Way of life, fears and aspirations Culture and community Personal and property rights
3x	Create a plan that addresses/justifies to the public, the potential loss of individual privacy/collection of personal information as it relates to providing security for the masses	Way of life, fears and aspirations Culture and community Personal and property rights
2x	Appropriate information management during and after a crisis situation	Way of life, fears and aspirations Culture and community
2x	Creation of ethical guidelines and monitoring of their use	Way of life, fears and aspirations Culture and community Personal and property rights
2x	Ensuring safe levels of technology interdependency	Way of life, fears and aspirations Personal and property rights
1x	Ensure measures are applied equally with no preferential treatment or profiling.	Way of life, fears and aspirations Culture and community Political systems Personal and property rights
1x	Holistic consideration of cause and effect during problem resolution	Way of life, fears and aspirations Culture and community Personal and property rights
1x	Risk mitigation	Way of life, fears and aspirations Culture and community

Prevalence	Clusters of Solutions	Aspects
		Personal and property rights
1x	Invest in analytical tools for cost benefit analysis/financial planning	Way of life, fears and aspirations Way of life, fears and aspirations Culture and community Personal and property rights
1x	Prohibit physically intrusive measures	Way of life, fears and aspirations Health & well-being Personal and property rights
1x	The use of the most current security measures to mitigate cyberattacks.	Way of life, fears and aspirations Health & well-being Personal and property rights
1x	Validation of technology against user requirements.	Way of life, fears and aspirations Health & well-being Personal and property rights

Based on the answers, the importance of a good PR-campaign should be well noted. With the help of a good campaign, trust between society and government/rail operators could be built. Data privacy and protection should also be demonstrated as a priority. Trust building overall was seen as a way to resolve ethical and other challenges. For example, the responses included implementation of methods that enable trust and demonstrate to the public the importance of privacy, and how their access to personal data is handled, protected, and taken seriously. Also related with public relations are the recommendations for justification/explanation to passengers for the need of their sensitive data, clearly announcing and/or displaying passenger compliance expectations, and creating plans that address and justifies to the public, the potential loss of some individual privacy/collection of personal information, as it relates to providing security to the masses. All of the above were mentioned several times. There were also many practical ICT-design related suggestions, such as internal protection of sensitive personal data, anonymisation and or minimal data gathering techniques used whenever possible and implementing data protection and privacy practices at all levels of operator access were mentioned. This together with external data security protection was mentioned several times. Furthermore, management practices were also addressed, for example, suggestions for appropriate information management before, during, and after a crisis situation and the creation of ethical guidelines and monitoring that these guidelines are followed. All of the above can be seen as recommendations for the developers when developing solutions forSAFETY4RAILS.

7.3 The Positive Outcomes and Impacts

The third and final survey question sought to examine the positive outcomes, in a way that justifies the SAFETY4RAILS project by revealing the value it creates. There was a total of 14 usable responses from question number three. Again, the prevalence should not be taken as a direct indication of importance; all are valid and possible.

TABLE 12 ACCEPTABILITY CONCERNS III

Prevalence	Clusters of Positive outcomes	Aspects
11x	Promotes increased security of rail travel	Way of life, fears and aspirations Culture and community Political systems Environment Health & well-being Personal and property rights
8x	Promotes increased use of rail travel	Way of life, fears and aspirations Culture and community Environment

		Health & well-being
8x	Reduces negative impact to the environment	Way of life, fears and aspirations Culture and community Environment Health & well-being
5x	Improved public opinion of rail travel	Way of life, fears and aspirations Culture and community
3x	Monetary gains/savings	Way of life, fears and aspirations Culture and community
3x	Decreased potential for loss of life and other bodily injury	Way of life, fears and aspirations Culture and community Health & well-being
3x	Improved efficiency of railway operating systems benefiting operator and public	Way of life, fears and aspirations Culture and community Environment
1x	Promotes social interaction	Way of life, fears and aspirations Culture and community Political systems

Related to the third question of the initial SIA, the potential value of SAFETY4RAILS is promoting increased security of rail travel: nearly all responses contained an item within this cluster. Additional information included passengers experiencing an increased sense of personal safety. SAFETY4RAILS could also create an increased use of rail travel and thus reduced negative impacts to the environment, both of these responses were mentioned often. All of the above are naturally linked to an improved public opinion of rail travel, which was also mentioned several times. Besides promoting safety, security, and an improved public opinion, a decreased loss of life and other bodily injury was pointed out as a positive outcome of the project. Actual tangible positive impacts which were mentioned included monetary gains/savings and improved efficiency of railway operating systems, not forgetting promoting social interaction, although the last is hardly very tangible.

7.4 Concluding Remarks

This initial SIA should be considered a demonstration that the project partners are well aware of the potential social impact challenges and benefits connected with the outcomes of the SAFETY4RAILS project. It is as much an exercise in ethical thinking as it is a way of determining the range of issues that need to be considered when both designing the solutions and thinking of the future use of these solutions. Also important, are the changing governance models that the solutions will inevitably cause, not forgetting changes in the division of labour. The category of *Way of life, fears and aspirations* is most prominent in this SIA, which is obvious since SAFETY4RAILS is about safety and security. Issues on *Culture and Community*, *Health & Well-being*, *Personal and Property Rights*, and *Environment* were also well represented, whereas the answers did not touch many aspects of the *Political System*, for obvious reasons. As stated, this is the initial SIA, which gives a good start for the two more rounds which are forthcoming during the project. Thus, the next steps will be developing the EUSurvey tool for the next round, if the preferable method of face-to-face meetings is out of question, due to the COVID-19 –situation.

In closing, the main messages should be highlighted: first, SAFETY4RAILS' value is in making railway transportation more safe and secure. Second, when doing so, privacy issues must be addressed, for example following a privacy-by-design –approach.[13] Finally, good communication is pivotal. These three in mind, most ethical concerns can be overcome and SAFETY4RAILS will have a positive impact to societies.

[13] Following privacy-by-design approach privacy is by default incorporated into technology and systems. It means that products are designed with privacy as a priority, along with whatever other purposes the system serves.

8 Local and Consortium Level Ethical and Data Protection Governance Structure

In terms of governance structures for safeguarding compliance with ethical and data protection regulations, the structure put in place attempts to provide mutually complimentary support at the frontline of everyday data processing within the project to ensure that the ethical and regulatory requirements are met and in particular any data processing remains fully GDPR-compliant.

As shown in Table 13 below, this comprises of a framework of bottom-up monitoring and reporting by local project managers to local DPOs at each data processor's site and reporting up to the project Data Controller who will proceed to involve the project Coordinator and together with the Ethical Manager they consider any operational issues that may emerge appertaining to planning for and implementation of ethical and data protection compliance for proposed data processing pipelines.

The Ethical Advisory Board, comprising of the three ethical and socio-ethical experts, will be periodically updated and consulted to ensure that the partners continue to remain compliant with both the legal and data protection regulations, specifically GDPR and equivalent regulations of relevant non-EU partner countries to which some of our partners belong, as well as the safeguards for ethically reflective and socially-responsible innovation.

TABLE 13 THE SAFTEY4RAIL ETHICAL & DATA PROTECTION GOVERNANCE STRUCTURE

Ethical Management Board (3 Advisors, the Ethical Manager, the Project Coordinator)
Project Data Controller (Data Protection Compliance Operational Monitoring & Reporting)
Individual partners' Local Data Protection Officers (DPOs) Legal & Data Protection Compliance & Reporting
Local Project Managers (Data Protection Compliance Implementation, Monitoring & Reporting)
Work Package Leaders Data Protection Compliance Assurance Monitoring & Reporting
Task Leaders Data Protection Compliance Assurance Implementation Management
Individual Staff Responsible for Data Protection Execution throughout all stages of exposure to compliance risk, namely:
Capture, Injection, Processing, Access, Storage Transfer Deletion
Data Sourcing, Data-Subject Recruitment, Healthy Consent Process
Anonymisation-at-Source, Encryption Pre-Transfer, Role-Based Access, and Incident Reporting

There are in particular three critical data protection related operational contexts and one critical ethical and social impact safeguarding context that are particularly signposted for special monitoring at the local and consortium level monitoring reporting points; these are:

A) **Preparatory planning well in advance of proposed data acquisition and processing:** this is to ensure full compliance with the relevant data protection regulatory requirements by clear and timely specification of the purpose and context of the data processing and seeking advice from the Data Controller. By reference to the relevant guidelines as set out in chapter 5 of this document and by consulting the Ethical Manager and Coordinator, the DC will be able to advise the partner(s) concerned as to the necessary steps to be a taken to ensure compliance with data protection regulations. Examples of this include deciding on any of the following compliance routes to be followed to by the partners depending on the context:

   i) the criteria to be considered to determine whether, when and how to conduct an Explicit Consent process either as the legal basis for a particular context or indeed as an optional extra safeguard where practicable

and meaningful (Art. 6.1.a GDPR, Recital (40)-(43), further reference: Art.13, Art 14. Art 22-GDPR); Chapter 5 has addressed the purpose and Contexts for which Explicit Consent seeking could be considered (e.g. for research as a volunteer per se or for personal data processing); the key criteria outlined should suffice to enable the DC to make the correct determination.

ii) Art 6.1.e 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 (personal data),

iii) Art 9.2.j (where any special category data applies) processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

iv) Of course the compliance safeguards shall not necessarily be limited to the above alternative legal basis but appropriate routes to compliance shall be concluded through consultation with the Ethical Manager and the Coordinator responsive to any further data processing needs that may be needed to support new use-cases requirements that may emerge through our on-going stakeholder-centred co-design process.

v) Mandatory Protocol for Data Transfers: Data Controller and the local DPO shall be informed of any planned data exchanges with partners from EU and non-EU countries with an “Adequacy Decision” pending (Turkey). Any such data transfer will adhere strictly to the data encrypted in-transit and at rest rule and in any case unless the data is irreversibly anonymised or aggregated, the transfer would have to proceed on the basis of Art. 46 (3) a, requiring contract clauses between the two Data Controller/processors committing both parties to ensure full compliance with regulations for data protection specifically GDPR and LPPD (in the case of Turkey -Law on the Protection of Personal Data No. 6698 dated April 7, 2016).

B) **Avoiding design decisions with potentially adverse and irreversible impacts downstream at the user acceptance and social acceptability levels:** Close engagement with the stakeholders and collective awareness of ethically reflective and socially responsible approaches to system design needs to be facilitated through enhanced communication across the consortium and in particular engagement of the partners with the relevant gateway partners and partners with expertise in socio-technical approaches to design as addressed in Chapter 6 of this document.

C) **Risk-Aversive Measures to prevent and mitigate against Misuse:** Some results of SAFETY4RAILS could be at the risk of potential misuse hence these are classified as Confidential or Restricted and stored with advanced encryption software to avoid distribution of sensitive information by conventional means. Recognising the potential risks of Dual Use, and Misuse depends on awareness raising and training about such risks and measure to avoid or minimise them. Accordingly a special seminar will be arranged by the Ethical Manager to provide the essential awareness raising for consortium partners. The partners are contractually obliged to comply with data protection and information security requirements and are thus expected to remain compliant and report to the Chair of the project Security Advisory Board -SAB) any issues of concern arising from the risk of misuse. The Security monitoring as shall be described later entails the assessment of security level of deliverables and risk assessment to determine the appropriate risk-aversive response for security protection.

We believe this governance structure represents a clear and thus workable approach to ethical and data protection compliance especially as it enables streamlined collectively responsive action by partners to support compliance at the implementation frontline of our planned innovation tasks. This process is already embedded through the engagement of partners with the SAFETY4RAILS Ethical Compliance Framework as demonstrated by the high level of interest evident in participation in the plenary and bilateral meetings focused on the detailed forensic analysis of the data processing requirements of the project. Partners have responded effectively thereby enabling the preliminary situation assessment re the SAFETY4RAILS data processing needs to be arrived at as shown in Table 5. This has in turn enabled the typological mapping of the 5 distinct classes of purposes and contexts of data processing, as proposed to-date, to be concluded as a reference table (Table 7) to enable the determination of the appropriate safeguarding measures and legal basis for each of the 5 classes of data processing that could possibly arise.

As such Table 7 is to serve as an Ethical Compliance Framework Console (ECFC) of indicative safeguards and legal basis, to be updated by the Ethical Manager and the Coordinator, in the light of any emerging new class of data processing that might be needed responsive to any new use-case(s) that may be demanded by the stakeholder group.

In this way the Data Controller, by reference to the CFC, and in consultation within the Ethical Board shall be able to conclude the appropriate ethical and data protection measure(s) to be taken prior to each intended data processing or indeed suggest how such proposed data processing may be revised for assuredly safeguarded data processing to proceed. In this way the efforts of the Ethical Board and the Security Advisory Board should mobilise our mutual resolve to ensure the avoidance of security and privacy risks for a successful and socio-ethically responsible SAFETY4RAILS solution.

SAFETY4RAILS has a set of Ethical and Data Protection Requirements as specified by the EU Ethical Committee and a number of supportive tasks and deliverables to ensure that: i) systemic embedding of security and privacy compliance will be achieved and adhered to at each task level throughout the project lifecycle (D9.1, D9.4, D11.1, D11.2, D11.3, D11.4) and ii) agreed operational standards for consortium implementations and procedures for all partners (D1.1, D1.5). These establish an operational management framework to support the above governance structure (Table 13) for the SAFETY4RAILS consortium to ensure a rigorous approach to security and privacy protection and generally ethically and legally safeguarded conduct of the project. This includes taking into consideration the likelihood of, and as far as possible, safeguarding against, mis-use in our approach to the socially responsible design of our innovation for risk-aversive operational deployment to the extent any risks, including any future misuse, could possibly be prevented within the boundaries of design-time control within this project.

Accordingly, the Data Controller shall respond to any situation assessment and resolution of ethical issues at the implementation front line on the basis of the ECFC as updated by the Ethical Manager and the Coordinator and collective determination and resolution of issues to ensure ethical and data protection compliance.

This is supported by the relevant collective experience within the Ethical Board including advice from the Ethical Advisory Board and local monitoring and advice of the respective Data Protection Officers reporting to the project Data Controller. The management framework includes tasks (e.g. T1.1, T1.4, T1.5) that are to provide the requisite coordination support to the overall ethical and quality management including the quality of our monitoring and reporting.

8.1 Compliance Risk Mitigation Measures

Our compliance monitoring will include, primarily, prompt reporting to the Data Processor's DPO who will report to the DC and EB re any likely or actual exposure to the risk of data breach of personal data including failure to adhere at the right time to agreed procedures e.g. re appropriate level of de-identification of any Personal Data Elements [Personal Identifying Information (PIIs)] consistent with the compliance strategy as determined by the DC, including in particular data encryption before any data transfer.

Following the above reporting step, a risk-aversive assessment of any compliance issues-of-concern will be performed by the DPO(s) involved and this will be completed within a period as shall be specified by the DC responsive to the severity of the risk. The DC and EB will make a joint report on the matter to the Coordinator who will initiate a process of situation assessment which may involve conference calls with the DPOs and partner teams concerned in order to take the appropriate action which may include red-flagging any ongoing data process or a temporary halt to the processing whilst remedial steps are taken; the Coordinator and the Ethical Manger shall conclude a situation assessment and seek advice from the Ethical Advisors accordingly.

As part of the commitment made to any data-subject whose personal data may have inadvertently been exposed to any risk, they shall be notified of the exposure to risk (consistent with Art. 19-GDPR) and the remedial steps taken to identify the cause(s) of any data breach, to assess the likely impacts (consistent with Art.17 GDPR) and to resolve the issues to ensure full compliance.

8.2 The Project Ethical Board (EB)

The Project Ethical Board comprises of i) the three independent ethical experts collectively as the Ethical Advisory Board, ii) The Coordinator and iii) the Ethical Manager; as described in Section 3.1.5 of Deliverable D1.1 (Project Management Manual) with key responsibilities, outlined here as follows:

The Project Ethical Board (EB) shall be constituted with the appointment of three Ethical Advisors early in the project to ethical and data protection compliance. The Ethical manager (EM) as the EB Chair is to ensure

i) Compliance monitoring with the ethical, privacy and data protection requirements during the project lifetime;
ii) Maintaining of on-going assessment of the ethical sensitivity of deliverables before any publication,
iii) Support of the objectives of ethical compliance, and safety of project activities for researchers and future users;
iv) Provision of supportive advice responsive to the needs of the local managers and the Data Controller
v) Monitoring of gender equality issues by appointing a Gender Equality Manager to monitor all the issues,
vi) Support of the monitoring and responsiveness posture of the project with regards to any risk of misuse.

8.2.1 The Independent Ethical Advisors

Three Advisors as Ethical Experts operate independently to support the ethical and data protection compliance of the project through their contribution to the Project Ethical Board. They are to be consulted periodically and as required to ensure a risk-aversive approach to planned research and innovation efforts.

The Project Ethical Board represents an effective integration of expertise including:

i) **The Ethical Advisors**

- **Dr Irina, Marsh** (Senior Consultant, CBRNE Ltd)
  Privacy & Data Protection Legal Compliance, Privacy Risk Analysis, DPA
- **Dr. Paul Raphael Stadelhofer** (Applied Ethics, TU Dresden)
  Applied Ethics and Social Impacts Analysis
- **Prof. Benjamin Scharte** (University of Zurich)
  Resilience Engineering as a way of navigating the complexity of sociotechnical systems

Due to the classification of this deliverable as “public” the Board Members CVs, and contact details that are approved for release are available on request and can be provided as a separate addendum to this document with a “private” classification; likewise the name and contact details of the Data Protection Officers (DPOs) at each partner organisation can be provided as a further addendum with a private classification.

The Ethical Board shall agree the timing of the periodic meetings in advance of which the Coordinator shall liaise with and agree an agenda of items to be discussed and provide the relevant information to the Ethical Advisors.

A first meeting of the Ethical Board has already been held within the first three months of the project (17th December 2020) and further periodic meetings are planned (M9, M16, M23); however additional meetings may be held as required throughout the project lifecycle to support the on-going compliance management.

This first meeting at M3 introduced the project objectives and data processing plans defined to-date. The Ethical Advisors have expressed their keen interest in the project objectives which they regard as very much in the interest of public safety and security and have stated their readiness to support awareness raising within the consortium about the data protection issues and the need for a responsible socio-ethically reflective and precautionary approach to ensure data protection compliance. This meeting proved a valuable early opportunity to ensure that the project is set for compliance with ethical and data protection requirements.

It should be noted that, consistent with the H2020 requirements, the costs incurred in attending Ethical Board meetings will be paid from the project funds; the consortium greatly appreciates the generosity of the Ethical Advisors in accepting our invitation to join the Ethical Board without any form of recompense other than travel and accommodation expenses incurred in attending Ethical Board meetings.

8.2.2 Project Management Members of the Ethical Board

The Coordinator and the Ethics manager represent the Project Management Team (PMT) within the Ethical Board; thus supporting efficient communication between the operational management frontline and the project ethical compliance framework; as follows:

i) **Project Coordinator Stephen Crabbe** (Fraunhofer EMI Institute)

(Project management and legal background)

ii) **Project Data Controller Antonio De Santiago Laporte** (Metro De Madrid)

(Industrial and Technical Engineering & Management in the Railways Sector)

iii) **Ethical Manager Prof. Atta Badii (University of Reading)**

(Privacy-by- design, social responsibility and acceptability of research and innovation)

8.3 The Security Advisory Board (SAB)

The Grant Agreement, Annex 1, Part B, Section 6 includes the proposed members of the project Security Advisory Board (SAB). Antonio Santiago de Laporte (Metro de Madrid -MdM) as Project Data Controller (DC) and Security Officer (SO) is the Chair of SAB which includes 2 external experts. The responsibilities of the SAB as described in section 6.2 of the Grant Agreement and also D1.1 section 3.1.6 are outlined here as follows:

i) Security sensitivity assessment of each deliverable and if approved to confirm this on the front sheet of the deliverable before submission to EC; this to be conducted with special reference to various information security risks to be avoided such as:

   a) Release of information on man-made threats,
   b) Critical Infrastructure (CI) security vulnerabilities,
   c) Security protection systems,
   d) Detailed attack and response scenario on CIs -historical or simulated

ii) Overseeing and safeguarding the use of security sensitive information within the project tasks and in any interaction with third parties;
iii) Managing cooperation on security issues among the project partners;
iv) Reporting to the Project Management Board (PMB) and the Project Coordinator (PC), any risk of security sensitive and/or misuse arising as part of the management reporting.

Section 6.2 of the Grant Agreement also provides details on deliverables requiring limited dissemination due to security reasons as identified by the European Commission during the negotiation of the Grant Agreement

The SAB members are to operate in two modes:

A) Routine Mode: this is to review all deliverables from the project to ensure that no sensitive information will be wrongly disseminated, and, in
B) Responsive Mode:
   a) This is to meet responsively to decide any risks likely to arise from any intended dissemination activity relating to any of the confidential deliverables as referred to the SAB.
   b) In addition any partner proposing a scientific publication is to submit this to the SAB, 30 days before due date; this is to enable the SAB to assess the security sensitivity of the contents of the proposed publications and the partners to raise any of security concern with respect to the contents of the proposed publication.

The EB and SAB will work cooperatively but independently to ensure the integrity and efficiency of the security scrutiny of the dissemination processes are maintained to prevent mistaken dissemination of security sensitive information.

8.4 Local Data Protection Policy Implementation

In terms of local security policy and governance structures for Data Protection Compliance Assurance, Monitoring and Verification at specific *points-of-inspection* e.g. at data-acquisition-point (at-source), data-synthesis-point, anonymisation-point, encryption-point and particularly before any data transfer, anonymisation and deletion, the following operational standards represent the basic measures to which all partners already conform:

As a minimum the data should be stored and processed in a manner to ensure that:

- Any consent forms, duly obtained as required per advice from the Data Controller, are securely stored.

- If at all possible, the data is either synthesised or if real data is to be used then any personal data elements are anonymised at-source with any Personal Identifying Information (e.g. name, surname) deleted/anonymised/pseudonymised and any personally linkable images subjected to appropriate masking techniques to ensure de-identification
- In cases where it is practically impossible to receive the data as already anonymised at-source, personal data shall be rendered de-identified at acquisition-point; this can be achieved through various techniques as set out in Table 6, ranging from irreversible anonymisation to strong pseudonymisation of textual and ID data and the masking of personally identifying or co-locating images/videos/audio data (e.g., through blurring/pixelisation/scrambling as appropriate)
- Data is stored in secure server using firewalls, and anti-virus
- Data access is secured ideally through Multi Factor Authentication (MFA)-based protected access (in particular if accessed off-site) but at least through a secure role-based access control including passworded documents etc.
- Personal Data encryption at-rest and in-transit, in particular personal data is encrypted prior to data transfer to non-EU countries.
- Data is protected by secure regular back ups
- Data is not hosted by a 'data processor' (as defined in Art 4 (8) GDPR) with whom the Data Controller does not have written contractual data processor agreements in place for data protection

(For the avoidance of doubt, the following are examples of Data Processors – Dropbox, AWS, GCLOUD/GSUITE, ICLOUD, Survey Monkey etc.) It is true that these Data Processors state that they have taken steps to be GDPR compliant. However legally speaking Private accounts with these suppliers will not cover the necessary contractual obligations under Art 28 of GDPR, as these need to be between the Data Controller (partner holding the data and processing it) and the sub-processor. In a nutshell, it needs to be a location that has been cleared by the Data Controllers' Data Protection Officer (DPO).

Whilst the Data Protection and Cyber Security Policy and Structures of all partner organisations satisfy all the above, some of the partners have a more extensive local data protection policy and governance framework as, for example, is the case with the major Data Processors who also have a very close collaboration with end-users. As such they have a key role in the implementation of the consent seeking processes and data acquisition in the consortium particularly for acquiring and processing User-Specified Requirements, User-Expressed Evaluation and Usability Data as well as synthesising other data as required.

9 Conclusions

This deliverable has set out the SAFETY4RAILS compliance framework of actionable guidelines for the Data Controller and partners to adopt appropriate data protection safeguarding steps to be followed as required to ensure ethical and data protection as well as information security compliance. The deliverable has also provided an analysis of approaches to socially responsible and acceptable innovation and the initial analysis of the data processing requirements, legal data protection and ethical and social impact analysis within the SAFETY4RAILS project. A framework has been set out upon which the further analysis to guide the socio-ethical and legal compliance within the project can be maintained. This deliverable has in particular addressed the key aspects of such compliance as follows:

- The Requirement for Explicit "Healthy Consent" as part of the legal basis for processing essential personal data as necessary; ensuring that any consent seeking process is conducted appropriately, meaningfully and where it is practicable and meaningful to do so.
- Data modality and privacy-sensitivity-specific approaches to de-identification of personal and otherwise linkable data.
- The categoric data protection safeguarding measures and respective legal basis responsive to the spectrum of proposed data processing in SAFETY4RAILS to-date; this is to serve as the continuously updated Data Controller's Ethical Compliance Framework Console to enable the determination of the most appropriate compliance safeguarding steps to be taken to ensure data protection with respect to any of the data processing pipelines proposed to-date.
- Localisation of sub-system development (e.g. model training and testing) to avoid transfer of any mistakenly included and identifiable personal or linkable data of real persons.
- Socio-technical, user-centred and social acceptability analysis for responsible and responsive innovation.
- Risk-aversive approach to prevention and mitigation against the risks of information security breach and misuse.

****

BIBLIOGRAPHY

Bush, V. (1945). Science: The endless frontier. [A report to President Truman outlining his proposal for post-war U.S. science and technology policy.] Washington, DC: United States Government Printing Office. Quoted from Bornmann, L. (2013), What is societal impact of research and how can it be assessed? A literature survey. Journal of American Society of Information Science Technology, 64: 217-233.

Hill, S (2016). Assessing (for) impact: future assessment of the societal impact of research. Palgrave Commun 2, 16073.

OECD (2020). Regulatory Impact Assessment, OECD Best Practice Principles for Regulatory Policy, OECD Publishing, Paris, https://doi.org/10.1787/7a9638cb-en.

Morrison-Saunders, A., Bond, A., Pope, J. & Retief, F. (2015). Demonstrating the benefits of impact assessment for proponents, Impact Assessment and Project Appraisal, 33:2,108-115.

Kemp, D. & Vanclay, F (2013). Human rights and impact assessment: clarifying the connections in practice, Impact Assessment and Project Appraisal, 31:2, 86-96.

Vanclay, F., and Esteves, A.M (Eds.) (2011). New directions in Social Impact Assessment. Conceptual and Methodological Advances. Cheltenham (UK).

SAFETY4RAILS Project, Project Number 883632, https://safety4rails.eu/

General Data Protection Regulation (EU) 2016/679), https://gdpr-info.eu/

SAFETY4RAILS Grant Agreement, Annex 1, Description of Action

What is Personal Data -European Reference legal Information Site

https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en

Article 29 European Data Protection Working Party

https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

An Analysis of the Consequences of the General Data Protection Regulation on Social Network Research, Dec 2019, ACM Transactions on Social Computing ACM Digital Library:

https://dl.acm.org/doi/10.1145/3365524

Guide to the General Data Protection Regulation (GDPR) Exemptions (UK-DPA)

https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/exemptions/

SAFETY4RAILS

Partners:

Fraunhofer EMI
UIC
Metro de Madrid
EGO Genel Müdürlüğü
RFI Rete Ferroviaria Italiana – Gruppo Ferrovie dello Stato Italiane
Leonardo
ceis avisa partners
STAM Mastering Excellence
IC information company
RMIT University
University of Reading
etra I+D
Demokritos National Centre for Scientific Research
RINA
WINGS ICT Solutions
Newcastle University
EOS European Organisation for Security
Innova Integra
Laurea Ammattikorkeakoulu University of Applied Sciences
CybeServices Made in Europe
erGE
MTRS
Intracom Telecom
Universitas Miguel Hernández
TREE Technology
Elbit Systems C4I and Cyber
ProRail
Comune di Milano

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 883532.