

Definition of the interface between RA tool and S4RIS

Deliverable 3.3

## Lead Author: FRAUNHOFER

## Contributors: STAM

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.

D3.3 DEFINITION OF THE INTERFACE BETWEEN RA TOOL AND S4RIS		
Deliverable number:	3.3	
Version:	V1.1	
Delivery date:	16/02/2022	
Dissemination	PU – Public	
level:		
Nature:	Report	
Main author(s)	Natalie Miller and Yupak Satsrisakul	Fraunhofer
Contributor(s)	Davide Ottonello and Deborah Hugon	STAM
	Kushal Srivastava	Fraunhofer
Internal reviewer(s)	Antonio De Santiago Laporte	MdM
	Atta Badii	UREAD
	Corinna Köpke	Fraunhofer
	Alexander Stolz	Fraunhofer
	Stephen Crabbe	Fraunhofer
External reviewer(s)	Olivier Waeber	

Document	control		
Version	Date	Author(s)	Change(s)
0.1	28/06/2021	Fraunhofer	Contributions added to sections 1 – 3
0.2	21/07/2021	Fraunhofer	Contributions from STAM added to sections 1 – 3
0.3	30/07/2021	Fraunhofer	Contributions reviewed internally
0.4	08/09/2021	STAM	Combination of STAM and Fraunhofer contributions
0.5	20/09/2021	Fraunhofer	Final edits made to entire document
1.0	12/10/2021	Fraunhofer/STAM	Version 0.5 with incorporation of updates following internal and external review feedback converted into 1.0
1.01	15/02/2022	Fraunhofer/STAM	<ul> <li>Version 1.0 with reviewer comments addressed:</li> <li>Section 1.1 (last paragraph) and 1.2 have been modified to better link the title of the document with the content.</li> <li>In section 3.1 (last paragraph), Q-RA is mentioned and an additional reference has been included to compare with the work presented here.</li> <li>Section 3.2 has been reworked to more clearly highlight the motivation to use ERD for database development.</li> <li>Details in tables in the original section 4.4 have been summarised in the new Table 4-4 and moved to Annexes II and III to concentrate the core content in the main body of the report and to reduce the potential risk of confusion.</li> <li>Section 5.2 has been updated to provide more behavioral views using UML-based sequence diagrams. Two activity diagrams have been added (Figure 5-4 and 5-6) along with additional descriptions.</li> </ul>
1.02	15/02/2022	Fraunhofer	Formatted version
1.03	16/02/2022	Fraunhofer/STAM	Further explanatory text in section 5.2 in multiple places and minor corrections.
1.1	16/02/2022	Fraunhofer	Creation of V1.1 from V1.03.

#### **DISCLAIMER AND COPYRIGHT**

The information appearing in this document has been prepared in good faith and represents the views of the authors. Every effort has been made to ensure that all statements and information contained herein are accurate; however, the authors accept no statutory, contractual or other legal liability for any error or omission to the fullest extent that liability can be limited in law.

This document reflects only the view of its authors. Neither the authors nor the Research Executive Agency nor European Commission are responsible for any use that may be made of the information it contains. The use of the content provided is at the sole risk of the user. The reader is encouraged to investigate whether professional advice is necessary in all situations.

No part of this document may be copied, reproduced, disclosed, or distributed by any means whatsoever, including electronic without the express permission of the SAFETY4RAILS project partners. The same applies for translation, adaptation or transformation, arrangement or reproduction by any method or procedure whatsoever.

© Copyright 2021-22 SAFETY4RAILS Project (project co-funded by the European Union) in this document remains vested in the project partners.

# 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 intracity 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 2004) and combined cyber-physical attacks, which are important emerging scenarios 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 to travellers and other communicated 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 will be validated by two rail transport operators and the results will support the redesign of the final prototype.

# TABLE OF CONTENT

AE	BOUT	SAFI	ETY4RAILS	3
Ex	ecutiv	ve sur	nmary	7
1.	Intr	oduct	ion	8
	1.1	Ove	rview	8
	1.2	Stru	cture of the deliverable	8
2.	Dat	a inte	gration for risk assessment	9
	2.1	Data	a model overview	9
	2.1.	.1	Conceptual level	9
	2.1.	.2	Logical level	9
	2.1.	.3	Physical level	10
	2.2	Data	a modelling types and components	10
	2.2.	.1	Elements	10
	2.2.	.2	Relations	10
	2.3	Terr	ninology	10
	2.3	.1	Data modelling terminology and notations	10
	2.3.	.2	Main terms used in the risk assessment tool	11
3.	RA	data	model integration of cyber and physical threats	13
	3.1	Req	uirements	13
	3.2	RA	conceptual data modeling	14
	3.2	.1	Risk assessment element scheme and relationships	15
	3.2.	.2	Element description and main properties	20
	3.2	.3	Data Dictionary	22
4.	Intr	oduct	ion of Graph Database	24
	4.1	The	Property Graph Model	24
	4.2	Labe	eling and naming rules	25
	4.3	List	of Nodes and Edges	26
	4.4	An e	example of a graph database	27
5.	Inte	gratio	on of the RA tool	30
	5.1	Sec	uRail architecture	30
	5.1.	.1	Databases	30
	5.1.	.2	DB Management System	30
	5.1.	.3	Frontend	30
	5.1.	.4	Authentication provider	31
	5.1.	.5	Rail data service	31
	5.1	.6	Station search engine	31
	5.1.	.7	Risk data service	31
	5.1.	.8	Risk engine	31
	5.1	.9	CBA engine	32

5.1.10	Deployment	32
5.2 See	cuRail workflows	32
5.2.1	Add a new Station to the user network	32
5.2.2	Perform risk assessment	33
5.2.3	Real-time Risk Assessment	
5.3 See	cuRail interfaces	
6. Conclus	sion	39
Bibliography	·	40
ANNEXES		43
ANNEX I.	GLOSSARY AND ACRONYMS	43
ANNEX II	. Edge specifications	44
ANNEX II	I. Code for visualization	47

#### List of tables Table 2-1: Level definitions and processes for the different model types. 9 Table 3-1: The requirements that are fulfilled by the data model and data dictionary 14 Table 3-2: ERD basic symbols used in ra data model. 15 20 Table 3-3: Entities with description and main attributes – Part 1 Table 3-4: Entities with description and main attributes – Part 2 21 22 Table 3-5: An excerpt of the data dictionary created and used as the database for the SecuRail tool Table 4-1: The different naming guidelines with examples. 25 Table 4-2: Table of nodes 26 Table 4-3: Table of relations 27 Table 4-4: Nodes and their relative attributes 28 Table A.0-1: Glossary and Acronyms 43 Table A.0-2: Edge specifications - Part 1 44 Table A.0-3: Edge specifications - Part 2 45 Table A.0-4: Edge specifications – Part 3 46 Table A.0-5: Edge specifications - Part 4 47 List of figures

## 

Figure 3.2: Entity relationship diagram of threat monitoring and risk assessment
Figure 3.3: Entity relationship diagram of network rail operation18
Figure 3.4: Entity Relationship Diagram with the connections between the network Rail operation and Threat and Response
Figure 4.1: An example of nodes and edges with the property graph model
Figure 4.2: Example of a graph database
Figure 5.1: Securail architecture scheme
Figure 5.2: Securail sequence diagram for adding a new station to the network
Figure 5.3: securail sequence diagram for launching a risk analysis
Figure 5.4: SECURAIL activity diagram for offline risk assessment
Figure 5.5: securail sequence diagram for real-time risk assessment
Figure 5.6: SECURAIL activity diagram for real-time risk assessment
Figure 5.7: Flow diagram of the connections of SecuRail to the other tools in S4RIS through the Kafka broker.

# Executive summary

Data models are an initial step when completing a risk assessment of a railway network. This deliverable discusses the steps to create a data model, including a conceptual overview of the different data model levels and types. The components that make up the data model and any supporting terminology are defined and discussed. The data model specifically created for the risk assessment tool, SecuRail 2.0 is explained in this deliverable, as well as the element schemes and relationships. Elements that are included in the model, such as threats, consequences, or stakeholders, are described including their main properties and attributes.

A data dictionary was also created for the risk assessment tool. The dictionary is introduced and described in this deliverable. This data dictionary contains all the attributes for the entities in the data model.

The graph database created by STAM is also introduced and discussed in this deliverable. This graph database is an extension of the conceptual data model and highlights to the connectivity of the data. The labeling is introduced as well as the different nodes and edges.

The last section of this deliverable discusses the integration of the SecuRail tool. This includes the architecture of the tool, with a breakdown of each of the different components that make up the architecture such as the data management system and the risk engine. A few of the workflows are also introduced such as adding a new station or performing a risk assessment. Lastly the interfaces of SecuRail and how it will be integrated into the S4RIS platform is discussed.

# 1. Introduction

## 1.1 Overview

The Risk Assessment (RA) and the threat modelling are significant for S4RIS to represent the causality and effect of each threat scenario. The cyber and physical threats identified in T3.1 and T3.2 will be modelled in the risk assessment tool as one of the outputs of WP3. Combined cyber physical threats need to be explained in terms of a technical description of an attack, detailed features and expected impacts, including their relations between the occurring threat and security measures. SecuRail, i.e. the risk assessment tool built in SAFETY4RAILS, enables the illustration of these pieces of information, computes the likelihood of an incident and evaluates impacts on relevant assets of the railway system. These outputs are shown through the front-end module, in which data can be visualized in the Graphical User Interface (GUI). However, in the back-end module, the computation engine is driven by data management and risk calculation algorithms. Indeed, it is important to aggregate relevant data (of threats, OT/IT monitoring, railway assets, predefined locations, etc.) and identify the relations to provide the computation engine with a consistent and exhaustive data model.

Furthermore, the risk assessment tool aims not only to perform offline analysis of the likelihood and impact of threats on the railway network, but also to use data coming in real-time from sensors to run targeted and fast impact forecasts. Monitoring tools included in the S4RIS system should communicate with risk assessment to transmit alerts triggering the online risk computation engine.

The main objective of this document is to present the interface of the RA tool SecuRail 2.0 and the S4RIS. To provide insights in data management in the RA tool, the data structure based on the data collections of threat features is elaborated with detection and monitoring having been gathered in Task 3.1 and Task 3.2. The document provides a conceptual data modeling, which delivers a data ontology to support the RA database. This database is used for the SecuRail tool development, which is integrated to S4RIS. Furthermore, communication between SecuRail and the rest of the platform (monitoring tools) will be depicted, considering data exchanged and interfaces used for the scope.

## 1.2 Structure of the deliverable

The remainder of this document is structured in the following way. Section 2 provides the data model architecture used for the conceptual data design to develop the RA tool and lists of important technical terms used in this document are also included. The main elements of collected data are categorized and their details are represented. Section 3 explains the data integration of the risk assessment. The data from different sources are combined and a data dictionary is presented as the ontology of the RA tool. In section 4, the properties and elements of the graph data model, which are used in the RA tool development are presented. These three introductory sections build the basis for section 5, in which the actual interface between SecuRail 2.0 and S4RIS is presented along with details on the architecture of the RA tool. Finally, section 6 concludes this document.

# 2. Data integration for risk assessment

Data integration is the process of combining data from different sources. A data model "is a model that describes in an abstract way how data is represented in an information system or a database management system" (Schallehn, 2021). This section introduces the different modeling types and levels, as well as the components that make up models. The terminology and terms are also introduced.

## 2.1 Data model overview

Data models organizing the data are categorized into different stages (Merson, 2009), (Anh, 2009). For example, in the beginning, data models are very conceptual and not many details are included. As more effort is added, the models can evolve to a more logical or physical level. Table 2-1 defines the processes for different model types, which include three different levels/stages: conceptual, logical and physical.

TABLE 2-1: LEVEL DEFINITIONS AND PROCESSES FOR THE DIFFERENT MODEL TYPES.

Model Type	Level	Process
High-level Conceptual Data Model (including Entity-Relationship Diagram (ERD))	Conceptual	Build conceptual data modeling for communication among stakeholders.
Record-based Logical Data Model, Hierarchical models	Logical	Design logical data represented in relational model, network, hierarchical structure.
Physical Data Models	Physical	Physical implementation, create database, data optimization.

#### 2.1.1 Conceptual level

The conceptual data model, explained in (Merson, 2009), includes details on implementation and focuses on the entities and their relationships. High-level conceptual data models provide the concepts for presenting data in a way that is similar to how people discern data (Anh, 2009). The relationships in the model will depend on the problem statement or domain. For communication with stakeholders, this model is the best as it is the most broad and high level. An example of a high-level conceptual data model is an Entity-Relationship Diagram (ERD), that incorporates entities, their attributes and the relationships between them (Anh, 2009). Entities, attributes and relationships are also defined by (Anh, 2009) and further expanded on in the following sections. An entity is a real-world item (such as a rolling stock, or train, or natural disaster threat, etc.). The entities have attributes that represent their properties (threat types, asset number or location, etc.). A relationship is an association among entities. For example, the threat targets a specific asset. The relationship in this example is the word "target". Further details on elements and relations can be found in the following sections 2.2.1 and 2.2.2. An ERD was selected as one of the main models utilized in T3.3. section 3 describes the Entity-Relationship Diagram that was created for the risk assessment tool, SecuRail 2.0.

#### 2.1.2 Logical level

Other types of data models are record-based logical data models which provide the concepts to the user in a similar method to the method computers use to store data. There are well known data models, namely the relational model, the hierarchical model and the network model (Anh, 2009). The relational model represent data in a table format. The network model depicts data as records, while the hierarchical model utilizes hierarchical tree structures for data representation.

The logical data model, as described in (Merson, 2009), adapts the conceptual data model and expands it to data management. Additionally, (Anh, 2009) states that the logical level is where the main descriptions are made regarding what data is stored.

#### 2.1.3 Physical level

At the physical data level, data entities are implemented (Merson, 2009). Optimizations may occur at this level as well, including merging of entities or creating indexes. This level is the lowest level in terms of conception and describes the data storage in actuality (Anh, 2009).

## 2.2 Data modelling types and components

Data models represent the logical concept of organizing data in a database format. The definition of the data model has been described in (Tsichritzis, 1977), as "a set of guidelines for the representation of the logical organization of the data and the relationships between them".

Data modeling becomes a common activity in the software development process of information systems. Data models contain information mentioned by (Merson, 2009), to fulfill the following purposes:

- Describe the object in the system's domain and any relationships it may have.
- Create a plan for the database structure.
- Help with the execution of code units that will access the database.
- Aid in enhancing the performance of data access operations.
- Be utilized to generate access codes and database schema.
- Help smooth the communication with stakeholders in topics of domain analysis and defining requirements.

#### 2.2.1 Elements

The elements in a data model are called data entities. Each entity represents any distinct object in the database. It can be any real-world object, however, many times the scope is limited to those that are necessary from a software perspective. In general, the properties of entities include the following features (as defined in (Merson, 2009)): the name, a description of significant details, a list of attributes representing properties of each entity, the unique identifier or primary key, constraint on the attribute values (as needed), and expected number of instances of the entity.

#### 2.2.2 Relations

There are three kinds of relations as defined by (Merson, 2009) to be:

- 1. **Relationships** are used to designate entities' associations. These can be one-to-one, one-to-many, or many-to-many.
- 2. **Generalization** relations are an 'is-a' relation between entities. The generalization relation is easier to find in conceptual data models than relational databases.
- 3. Aggregation allows for entities to be combined into one. However, this is rarely used.

# 2.3 Terminology

This section introduces terms and notations used in the conceptual data model and risk assessment tool. These terms and notations come from a variety of sources including external and internal (from within the SAFETY4RAILS project).

#### 2.3.1 Data modelling terminology and notations

The following list contains terms/notations commonly used in data modelling, database design, and information engineering.

- Data modelling "is the first step of the process in database design" (Watt, 2014) .
- **Data model** "is a collection of concepts or notations for describing data, data relationships, data semantics and data constraints" (Watt, 2014).
- Entity "an individual object has an identity and does not depend on another object" (Glinz, 2014).
- Node is often referred to as an entity in a graph database (neo4j).
- Attribute is "a characteristic or property of an entity" (Glinz, 2014).
- **Table** is "a named relational database data set that is organized by rows and columns. The relational table is a fundamental relational database concept because tables are the primary form of data storage (Techopedia)."
- **Data type** "defines the way in which the value of an attribute should be physically stored" (Carter, 2003).
- Relationship is "a link between two entities" (Carter, 2003), or between two nodes.
- **Cardinality** is "the minimum (ordinality) and maximum (cardinality) number of objects in a relationship" (Glinz, 2014).
- Value is "a numerical quantity measured, assigned or computed" (National Cancer Institute).
- **Key** "is a field, or combination of fields, in a database table used to retrieve and sort rows in the table based on certain requirements. Keys are defined to speed up access to data and, in many cases, to create links between different tables" (Techopedia).
- Constraint is a rule used to limit a type or length of the data (Project-management.com, 2018).
- **Element** is an abstract collective unit containing essential or characteristic parts, e.g. other elements, entity references, comments, descriptions, etc. (Council of Europe, 2020).

#### 2.3.2 Main terms used in the risk assessment tool

The following terms are the main terms used in the risk assessment tool. The location of the definitions for each of these terms is provided, or a definition is given. See section 3.2.2 for more details.

- **Information Technology (IT) system** is "the use of any computers, storage, networking and other physical devices, infrastructure and processes to create, process, store, secure and exchange all forms of electronic data" (Castagna, et al., 2021).
- **Operational Technology (OT) system** is "the hardware and software component that detects or causes a change, through the direct monitoring and/or control of industrial equipment, assets, processes and events" (contributors, 2021).
- **Operator of Essential Services (OES)** "are private businesses or public entities with an important role to provide security in healthcare, transport, energy, banking and financial market infrastructure, digital infrastructure and water supply" (European Commission, 2018).
- **Monitoring** is the systematic process of collecting, analyzing and using information to track the behavior of a system. In the risk assessment framework, monitoring is mainly concerned with threat monitoring, i.e. the usage of tools and methods to understand the current or upcoming occurrence of an incident or attack (Chartered Accountants ).
- **Asset** is "anything that has value to the organization, its business operations and their continuity, including Information resources that support the organization's mission" (ENISA; International Organization for Standardization (ISO), 2004).
- Target is "a person, object, or place selected as the aim of an attack" (Oxford Languages).
- **Risk** "is the potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm to the organization. (ISO/IEC PDTR 13335-1)" (ENISA).
- **Threat** is "any circumstance or event with the potential to adversely impact an asset through unauthorized access, destruction, disclosure, modification of data, and/or denial of service" (ENISA).
- **Cyber threat/attack** refers to "An attack, via cyberspace, targeting an enterprise's use of cyberspace for the purpose of disrupting, disabling, destroying, or maliciously controlling a

computing environment/infrastructure; or destroying the integrity of the data or stealing controlled information" (National Institute of Standards and Technology).

- **Cyber-physical threat/attack** refers to "a security breach in cyber space that impacts on the physical environment. A malicious user can take control of the computing or communication components of water pumps, transportation, pipeline valves, etc., and cause damage to property and put lives at risk" (International Risk Management Institute, Inc.). In SAFERY4RAILS this also covers threats that originate in the physical environment that impact the cyber space.
- Incident refers to the initial definition in D2.5.
- **Natural Hazard** is a natural extreme event that has the potential to cause an impact to the system under consideration, such as flood, earthquake, wildfires etc. (Organization of American States).
- **Incident** is an unintentional event "that has been assessed as having an actual or potentially adverse effect on the security or performance of a system" (ENISA).
- **Attack** is an intentional event perpetrated against a certain system with the goal of causing a devastating impact (Techopedia, 2016).
- Scenario is a specific event (or a series of events) that describes the occurrence of a threat against a certain target. Scenario can be described by several parameters, according to the modelling purposes (Haugen, et al.).
- **Countermeasure** is a set of measures that target the exposure of the attack, risks in/on the railway operation, infrastructure or relevant systems. The measures can be in the forms of software, hardware, or modes of operative actions, depend on OT, IT or persons who oversee any specific areas. The aims of the measures are to prevent threats (physical/cyber/combined), or mitigate potential impacts, and reduce cascading effects to the whole system.
- Effectiveness of Countermeasure is the capability of a certain countermeasure to prevent or detect the occurrence of a threat, or in the worst case mitigate the severity of the impact of the threat occurred (Piper).
- Vulnerability refers to the initial definition in D2.2 (ENISA).

# 3. RA data model integration of cyber and physical threats

# 3.1 Requirements

Risk Assessment is a complex procedure which has been performed manually for years. To date, due to the growing complexity of systems to be analyzed and the increasing amount of data, performing this process in an adequate manner requires ad-hoc tools or specific application of tool approaches to the system being assessed.

In this context, these risk assessment tools need to be based on solid data models capable to represent the system (infrastructure, assets, people etc.) under consideration as well as the different threats that could affect it. Indeed, a relevant part of the design of these tools is focused on data modelling and database design.

Within SAFETY4RAILS project, SecuRail tool will be further developed: this will be the main risk assessment tool of S4RIS platform and it will be a dedicated web-application for conducting risk assessment and costbenefit analysis of railway networks and infrastructures. The data model of SecuRail is being defined in T3.3 and it will be implemented in T3.4 as the main database of the web-application itself. The data model is available also for other partners of the consortium, to acknowledge the data structure of SecuRail in case of an exchange of information, as well as to promote a common representation of the domain within S4RIS platform.

The data model elaborated in this document will be then exploited for three main goals:

- Allow railway infrastructure managers to represent their own network in a structured manner through the User Interface of the tool
- Allow the railway infrastructure managers to carry out an offline risk assessment of their own network
- Allow the railway infrastructure managers to be informed about a threat occurring and have a rapid estimation of potential impact (real-time risk assessment)

Regarding the real-time risk assessment, the data about an occurring threat will be gathered by monitoring tools developed in WP4 and sent to SecuRail through interfaces (see Figure 3.1) determined as part of the overall S4RIS architecture based on a Distributed Messaging System (DMS). The connections of the data model and the SecuRail tool can be seen in Figure 3.1.



FIGURE 3.1: THE FLOW DIAGRAM OF INFORMATION IN S4R FOR THE DATA MODEL AND SECURAL TOOL.

Indeed, before the definition of the data model, an initial set of requirements has been drafted in order to make the model suitable for SAFTY4RAILS. The data model should:

- Represent the railway domain with a high degree of fidelity.
- Be suitable for the development of risk analysis algorithms based on it.
- Represent both the railway network and its components, as well as the threats that can affect them.
- Contain all the attributes needed to define and analyses a certain scenario within the risk assessment tool.
- Be used to describe the railway network under analysis and the threat landscape in a static way, but also to represent the occurrence of a threat in real-time.
- Model the propagation of threats through the network/infrastructure and be compatible for simulating the possible cascading effects.

A variety of requirements stipulated from the end users can be found in WP2. Specifically, requirements regarding crisis management, crisis communication and smart city resilience are found in D2.5, requirements related to the case studies and threats within the case studies are in D2.2 and standardization and interoperability requirements are in D2.4. Table 3-1 lists the main relevant requirements in connection with the data model and how the data model and data dictionary described in this deliverable will contribute to fulfilling them.

Requirement ID (Deliverable number)	Short name	Fulfilment
UR-CM_R08	Cascade effect simulation	The data model and data dictionary will be utilized by SecuRail, which is a cascading effects simulator and can assess the resilience of the infrastructure.
UR-CM-R11	Detection of abnormal situation/anomalies	The data model includes detectors/sensors.
STD-R46	Infrastructure monitoring	The data model includes detectors/sensors.
STD-R48	Overall security event / incident / vulnerability database	The data dictionary, based on the threat taxonomy completed in T3.3, contains a catalogue of threats as well as their attributes and characteristics.

TABLE 3-1: THE REQUIREMENTS THAT ARE FULFILLED BY THE DATA MODEL AND DATA DICTIONARY

This approach is similar to input collection as highlighted in (Flammini, 2008). The protection mechanism is a list of counter measures. The threats, counter-measures, assets and their inter-relations are rigorously studied to generate a relational database which is explained in section 4. This is similar to the design of Q-RA as outlined in work (Flammini, 2008). However, in this work, emphasis is more on modeling and simulation of cascading effects of the threat events.

# 3.2 RA conceptual data modeling

Risk Assessment (RA) has been elaborated in D2.1 as the process to identify the security gaps and support the operational planning, including risk management. The main processes of RA, as defined in D2.1, includes risk context establishment, risk identification, risk analysis, risk evaluation, and risk treatment.

An entity relationship model (ER model) is formed to represent data/information structure necessary to perform business process, which are then implemented in a database. It represents a business data scheme (Wikipedia, 2016). In order to identify unique entities and define attributes and establish their inter-relations, ERD for the threat modelling and risk assessment, and, network rail operations have been defined in later section.

To establish the risk context in the data modeling, every single potential element has to be defined and the relationships among them clarified. Therefore, this section describes the data ontology providing a relational element structure and description of RA used in SecuRail 2.0.

#### 3.2.1 Risk assessment element scheme and relationships

The scheme contains the risk assessment (RA) elements consisting of key entities and their main attributes. The relationships represent the connected actions between entities. In an Entity Relationship Diagram (ERD), there are common symbols used to represent the relation of data in the risk assessment application. The two basic symbols used in the model represent entities and attributes, and cardinality and ordinality. Descriptions of the symbols can be found in Table 3-2.

oyzoi	
The block represents the entity where the entity name is on top, and the bottom part contains attributes (properties) of the entity. See the following figure.	
The end point symbols represent the different relationship constraints that can exist between entities. Both ends of the connections can have these symbols. (Figure source: (Lucidchart))	
+ One	
Many	

#### TABLE 3-2: ERD BASIC SYMBOLS USED IN RA DATA MODEL.

The preliminary ERD created for the SecuRail 2.0 risk assessment tool as a conceptual design is in the following figures.

In this document, two different ERDs are explained, that is ERD of threat modeling and risk assessment, and ERD of network rail operation.

#### 3.2.1.1 ERD of threat modeling and risk assessment

Figure 3.2 shows the relationships between threat monitoring and response. It represents the data model in relation to risk assessment algorithms, which aims to estimate the impact on the rail network due to the occurrence of threats. The entities modelled in this section include, Threat, Impact, Countermeasure, etc.

A threat can be perpetrated against a certain target; the latter can be any asset, station, or a specific area or section in any location in the railway network modelled. The target, in turn, is protected by one (or more) countermeasures. Due to cascading effects, the threat can trigger further consequences which can affect the same target hit by the initial threat but also connected targets. These consequences cause an impact on the railway network, which is mainly due to physical damages to infrastructure, fatalities, injuries and service interruption/disruption. All these factors cause economic losses which should be quantified. For what concerns the response against a certain threat, the detection phase is modelled through a detector installed in the railway network, which can identify a suspicious event occurring in the modelled environment. The suspicious event can be connected to the occurrence of a threat; indeed, the stakeholder can act to prevent and/or contain the threat (or the triggered consequences) and, in the worst case, recover from impacts affecting the infrastructure.

#### 3.2.1.2 ERD of network rail operation

Figure 3.3 shows the relationships of the network railway operation. A stakeholder can own a network (if it is an infrastructure manager/owner) or it can operate a railway line (if it is a service provider). In a network owned by a single infrastructure manager/owner, it is possible to have several lines operated by different stakeholders (like in Italy, where there is RFI as the railway network owner and Trenitalia, Trenord and Trello as main service provider). But sometimes the owner/manager of the railway network is also the unique service provider, like in Metro de Madrid and in other metro networks.

Trains operate railway lines following a pre-defined timetable made by stops. Stops are the logical layer of Station, which represent the nodes of the railway network. Two adjacent stations are connected by a railway section, which can be part of several lines. Indeed, the path of a line starts from a station and ends in another one crossing several sections and stations (which can be stops or not). Each station is composed by different areas, i.e. a delimited space with a certain functionality. Each area can contain one or more assets; assets can be located also in sections. Assets can be connected to each other if an interdependency exists, e.g. the functioning of the first is essential for the functioning of the second one.





#### 3.2.1.3 Interconnections between the two data models

Figure 3.4 represents the existing connections between the two data models described before. Stakeholders are in charge of managing detectors and carrying out actions to respond against threat occurrence. Area, asset, station and section can be targets of threats, as well as of consequences triggered by threats. Impact caused by consequences will change according to the type of threat and also target. Thanks to the relationships described by these data models, risk analysis algorithms will evaluate propagation of threats, consequences and impacts across the railway network.

The definitions of the different components in the ERD can be seen in Section 3.2.2.



FIGURE 3.3: ENTITY RELATIONSHIP DIAGRAM OF NETWORK RAIL OPERATION.

#### LINK BETWEEN NETWORK AND THREAT



FIGURE 3.4: ENTITY RELATIONSHIP DIAGRAM WITH THE CONNECTIONS BETWEEN THE NETWORK RAIL OPERATION AND THREAT AND RESPONSE.

# 3.2.2 Element description and main properties

TABLE 3-3: ENTITIES WITH DESCRIPTION AND MAIN ATTRIBUTES – PART 1

Entity Name	Description	Main attributes
Action	An incident response defines an "organized approach to addressing and managing" the attack or the aftermath of suspicious event in order to reduce the damage and recovery costs and time (Chai, et al., 2020).	Identity number, incident, crisis management issue, person in charge
Area	A boundary of any location is identified by datum (longitude, latitude), geometric shape or zoning. A spatial area is related to the position and existence of the asset.	Identity number, type, name, size, geometric space
Asset	"Anything that has value to the organization, its business operations and their continuity, including information resources that support the organization's mission" (International Organization for Standardization (ISO), 2004). Any asset belongs in a specific area.	Identity number, type, name, economic value, quantity
Consequence	"Outcome of an event affecting objectives" (International Organization for Standardization (ISO), 2009). There can be more than one consequence from one event. Consequences can range from positive to negative. "Consequences can be expressed qualitatively or quantitively" (International Organization for Standardization (ISO), 2009)	Identity number, name, type, description
Countermeasure	A set of measures targets the exposure of the attack, risks in/on the railway operation, infrastructure or relevant systems. The measures can be in the forms of software, hardware, or modes of operative actions, depend on OT, IT or persons who oversee any specific areas. The aims of the measures are to prevent threats (physical/cyber/combined), or mitigate potential impacts, and reduce cascading effects to the whole system.	Identity number, name of countermeasure, type of countermeasure, cost of capital expectation, cost of operation expectation, resilience scores (prevention, detection, response, mitigation, recovery)
Detector, sensor	A component/module that is monitoring rail capacities to detect defects, anomalies or unusual incidents.	Identification number, device name, type, location
Impact	"The result of an unwanted incident" (International Organization for Standardization (ISO), 2004). In this case, the unwanted incident is expressed in terms of the consequence of a given threat.	Physical damage value, service disruption, economic loses

<b>TABLE 3-4: ENTITIES WITH DESCRIPTION</b>	AND MAIN ATTRIBUTES – PART 2
---	------------------------------

Entity Name	Description	Main attributes
Line	"One or more adjacent running tracks forms a route between two points" (Eurostat Statistics Explained, 2012).	Identity number, type, start station, end station, operator
Network	A collection of one or more lines in a specific area (OECD, 2006). In S4R, this includes stations located along each line. The network also includes communication connections, electrical supply hub and distribution, as well as transportation modals.	Identity number, name, owner, type of network, country
Section	"The line between the departure end yard limit of one location and the arrival end yard limit of another location. A section consists of one or more blocks" (Rail Industry Safety and Standards Board (RISSB)).	Identity number, departure station, arrival station, length, maximum speed, number of tracks
Stakeholder	"Those people and organizations who may affect, be affected by, a decision or activity" (Rail Industry Safety and Standards Board (RISSB))	Identity number, type, country
Station	A place or a building provides service activities related to transportation. It can be a point of connection, starting and ending of a rail route, and a place in which the passenger can get on or off trains.	Identity number, station name, station type, region, city, number of platforms, average occupancy
Suspicious event	An event set off by a particular set of circumstances. The event can be certain or uncertain. The event can be a single occurrence or a series of occurrences (International Organization for Standardization (ISO), 2009).	Identification number, type of event, timestamp, severity level, description
Target	"a person, object, or place selected as the aim of an attack" (Oxford Languages). It could be a vulnerable point of the system or in a particular situation. The target can be protected by a set of countermeasures.	Identification number, name of target, type of target, location of target
Threat	"Any circumstance or event with the potential to adversely impact an asset through unauthorized access, destruction, disclosure, modification of data, and/or denial of service" (ENISA)	Identification number, name of threat, type of threat, severity of threat, likelihood of threat, threat description
Timetable	A timetable defines the schedule for the trains in a railway network. The timetable can include the direction, departure and arrival stations, the time of departure and arrival and track information.	Identification number, day, departure time, arrival time

Train	"A locomotive or self-propelled vehicle, alone or coupled to one or more vehicles" (Rail Industry Safety and Standards Board (RISSB)).	Identification number, train type, train status, maximum number of passengers

#### 3.2.3 Data Dictionary

As defined in (UC Merced Library), a data dictionary provides metadata of the different elements used in the project. The data dictionary collects a list of entities which are captured in a database. The use of a data dictionary is to be a guide on element clarification, depiction and meaning, as well as the rules for data management (UC Merced Library).

The data dictionary is useful for the following reasons (UC Merced Library):

- Helps reduce data inconsistencies,
- Determines the conventions used,
- Allows for more consistency with data collection and data usage in the project,
- Assists in data analysis,
- Applies Data Standards usage.

The data dictionary created in this task will be mainly used in the SecuRail tool and will be filled with data from T3.1 and T3.2, as described in Section 3.1. An excerpt of the data dictionary can be seen in Table 3-5. The excerpt focuses on the threat entity, however, all entities in the ERD diagrams discussed earlier are included (threat, asset, countermeasure, detector, suspicious event, target, consequence, impact, action, station, section, line, network, stakeholder, train, timetable, and area). For each entity, the attributes are defined by name and description. Information including data type, data format, field size, an example and the requirements are also defined. A distinction is made between a primary or foreign key as well, with foreign keys coming from other entities in the dictionary and primary keys being unique to the table. For each entity, the foreign keys are the entity IDs that are input connections for the main entity.

Threat	ID	Text	UUID	10	Acronym of threat type with number	be3b9f70-2aad- 489f-8591- 4f67fbd80803
	Туре	ENUM		10	Threat typology corresponding to acronym in ID	Sabotage
	Likelihood	Integer	Number 1 to 5, or range	10	An estimation of the likelihood of the threat occurrence against systems in railways domain expressed through 5 levels: negligible (1), rare (2), possible (3), frequent (4), certain (5)	2: Rare

TABLE 3-5: AN EXCERPT OF THE DATA DICTIONARY CREATED AND USED AS THE DATABASE FOR THE SECURAIL TOOL

Severity	Integer	Number 1 to 5, or range	10	An estimation of the potential impact of the threat on the target system within the railways domain expressed through 5 levels: negligible (1), minor (2), moderate (3), severe (4), catastrophic (5)	2 to 4
ActionID	Text	UUID	10	ID from Action Entity	f0b4414a-7d67- 484b-bbae- 36c7da578353
SuspiciousEventID	Text	UUID	10	ID from Suspicious Event Entity	9e8635d7-6dbd- 49c5-85f6- 2b2e4dce50c0
Description	Text	UUID	20	General description of the threat	36af9c69-2dbc- 4041-868a- aa89bdf2be91

# 4. Introduction of Graph Database

In section 3, the conceptual data model has been created to express the relationships between potential entities used in the risk assessment. However, these relationships have dependent and complicated connections, which is difficult to handle with relational database management. To cope with this problem, a graph database technology is created as a solution to control the data connections within SecuRail tool and aims to properly implement the relations between components of the railway infrastructure.

The main purpose of the graph database is to deal with highly connected data, and the increase of volume and connectedness of a growing tremendous amount of data. Three major advantages of the graph database are the following aspects (Neo4j):

- Performance: The graph database improves the intensive data relationship by several orders of magnitude. It can avoid the number and depth of relations increase during queries with traditional databases.
- Flexibility: The structure and schema of a graph model flexes as applications. The data team can add it to the existing graph structure without endangering current functionality.
- Agility: The graph database aligns with an agile development process, test-driven development practices, allowing the graph database to evolve in step with the rest of application and any changing requirements. The modern graph database is also consistent in development and convenient in maintenance.

Due to the positive performance of graph database management, SecuRail 2.0 is developed by employing graph database technology to manage the back-end database and reduce data dependency. To understand the concept of the graph database used in RA, Section 4.1 explains a basic element of graph model, how its relationship is represented. Section 4.2 gives a recommendation for labelling and naming rules used for the graph database design to provide consistency in data management. In the last section, the list of potential graph elements used in SecuRail 2.0 is represented and described.

# 4.1 The Property Graph Model

The graph database design used in RA tool is developed on a graph database platform Neo4j. This section provides the concept of the property graph model, in which the data is organized as nodes, relationships and properties. These elements made up the property graph model can be briefly defined by following the Neo4j terms and definition (Neo4j Developer).

"Nodes are the entities in the graph. They can hold any number of attributes (key-value pairs) called properties. Nodes can be tagged with labels, representing their different roles in your domain. Node labels may also serve to attach metadata (such as index or constraint information) to certain nodes."

"Relationships provide directed, named, semantically relevant connections between two nodes entities. A relationship always has a direction, a type, a start node, and an end node. Like nodes, relationships can also have properties. In most cases, relationships have quantitative properties. Due to the efficient way relationships are stored, two nodes can share any number or type of relationships without sacrificing performance. Although they are stored in a specific direction, relationships can always be navigated efficiently in either direction."

Figure 4.1 displays an example relationship connecting between two nodes. In the model figure, it states that a movie is directed, written, produced, acted in and reviewed by a person. Moreover, a circular relationship exists between a person which follows another person. The relationships represent an action described with a verb (such as 'directed', 'reviewed'), while the node represents an object with a noun ('Person', and 'Movie'). The direction of the relationship is always expressed and useful enough to explain the action between the nodes. A bidirectional relationship is possible but not required unless it is needed to duplicate or properly describe the case.



FIGURE 4.1: AN EXAMPLE OF NODES AND EDGES WITH THE PROPERTY GRAPH MODEL

Both relationship and node can have their own properties described in the curly brackets. The properties are a list of paired values used for adding qualities to the node. The values of each property can hold different data types (number, string and Boolean). The naming system of the property graph model is explained in the following section.

The code to be run in Neo4j browser to obtain the graph is reported in the following line:

(:Person {name: string})-[:ACTED\_IN {roles: [string]}]->(:Movie {title: string, released: number})

## 4.2 Labeling and naming rules

This section describes how to name/label the terms used in T3.3 and T3.4.

Naming rules (Corresponding to Neo4j Standardization) (Neo4j, 2020):

- Names should be used consistency
- Names should refer to terms commonly used in the project
- Names always begin with alphabetic characters (a number is not allowed).
- Numbers are allowed to use a suffix or combined with alphabetic characters to represent the order.
- Names are case-sensitive

Examples of other naming guidelines for the different types of parts in the model can be found in Table 4-1.

Name	Recommendation	Example
Node label	Mixed-case names begin with the uppercase alphabetic character	:CounterMeasure
Relationship Type	All upper-case names, separating each word with an underscore.	:IS_PERPETRATED_BY

TABLE 4-1: THE DIFFERENT NAMING GUIDELINES WITH EXAMPLES.

Table	All uppercase names, all lower-case names without space and special character, except underscores ('_'), or mixed-case names without space and special characters (the first letter of each word uses the uppercase)	RISK_ASSESSMENT, risk_assessment, RiskAssessment
Property/Column	Lower-camel-case names begin with a lower-case character	:name, :type, :id

# 4.3 List of Nodes and Edges

The aim of this section is to collect all the nodes and edges defined in Neo4j Database, which should be corresponding to the element scheme presented in section 3. Table 4-2 contains the list of nodes and relations, with their label, key and attributes. A node's label is a list "containing the string representations for all the labels of a node" (Neo4j, 2020). A node's key is a list "containing the string representations for all the property names of a node, relationship, or map" (Neo4j, 2020).

Label	Кеу	Attributes
Threat	Type (enum)	Likelihood (double), Severity (integer), Description (string)
Consequence	Type (enum)	Description (string)
Section	ld (integer)	Length (double), maxSpeed (integer), numberOfTracks (integer), percentangeBridge (integer), percentageTunnel (integer), numberOfLevelCrossing (integer)
Impact	Type (enum)	physicalDamage (double), lethality (double), serviceDisruption (integer), economicLosses (integer)
Station	Name (string)	Coordinates (double, double), Country (enum), city (enum) numberOfPlatform (integer), crowding (array of integer)
Area	Name (string)	Type (enum), economicValue (integer), size (integer) geometry (array of double)
Countermeasure	Name (string)	Id (integer), Type id (integer), Type (enum), CAPEX (integer), OPEX (integer), Quantity (integer), preventionScore (integer), detectionScore (integer), responseScore (integer), mitigationScore (integer), recoveryScore (integer)

Asset	Name (enum)	Id (integer), Type id (integer)		
10001	Name (cham)	Type (enum), economicValue (integer), quantity (integer)		

#### TABLE 4-3: TABLE OF RELATIONS

Connected nodes	Name	Attributes (if applicable)
Section-Station	Starts from, Arrives to	N.A.
Station-Area	Includes	N.A.
Area-Asset	Includes	N.A.
Countermeasure-Area	Protect	efficiencyRate (integer)
Countermeasure-Asset	Protect	efficiencyRate (integer)
Threat-Area	Is perpetrated in	N.A.
Threat-Impact	Generates	N.A.
Impact-Area	Affects	N.A.
Impact-Assets	Affects	N.A.
Countermeasure- Threat	Is effective against	efficiencyRate (integer)
Threat-consequence	Causes	N.A.
Consequence-Impact	Generates	N.A.

## 4.4 An example of a graph database

An example of a graph database has been introduced with the aim of providing an idea on the construction of the complete database which will be used by the risk assessment algorithm. The example (Figure 4.2) considers a terrorist attack characterized by a bomb placed in the central hall of a station (in particular, in Milano Centrale station). The analyzed station is connected with another station, Milano Porta Nuova, with a bilateral section of rail. The bombing attack generates the explosion on the site where the bomb is located and it affects a high number of assets, such as the ticket machine. Moreover, the database presents the countermeasures equipping the main hall: a group of two patrols monitors the site and their actions are effective against the bombing attack. Furthermore, the attack causes the start of a fire where the bomb explodes. This consequence requires the intervention of the main sprinkler to protect the site as much as possible. When the fire reaches a dangerous level, the infrastructure collapses provoking irreversible damages. The minimization of the likelihood

related to the infrastructure collapse is possible through a fireproof structure. This improvement measure can be characterized in different ways, such as structures and materials, resistant to fire, or incombustible, or material for use in making anything fire-proof. This last characteristic aims to protect the central escalator besides all the assets which are present in the main hall.



FIGURE 4.2: EXAMPLE OF A GRAPH DATABASE

These nodes and the relative attributes are inTable 4-4. The attributes associated to each node in the database are reported. As it can be observed, each node has different attributes depending on the type of element it represents (for example, countermeasure, asset, area, threat).

TABLE 4-4: NODES AND THEIR RELATIVE ATTRIBUTE	S
---	---

Name	Classification	Attributes
Bombing	Threat	Description, id, likelihood, severity, type
Main Hall	Area	Economic value, geometry, id, name, size, type, type_id
Main Sprinkler	Countermeasure	CAPEX, detection score, id, mitigation score, name, OPEX, prevention score, quantity, recovery score, response score, type, type_id
Patrol 2	Countermeasure	CAPEX, detection score, id, mitigation score, name, OPEX, prevention score, quantity, recovery score, response score, type, type_id

Milano Central	Station	City, co-ordinates, country, crowding, id, name, number of platform
Milano Porta Garibaldi	Station	City, co-ordinates, country, crowding, id, name, number of platform
Milano Centrale Porta Garibaldi	Section	Id, length, max speed, number of levelC, number of tracks, percentage Tunnel, percentage Bridge
Ticket machine	Asset	Economic value, id, name, quantity, type, type_id
Explosion	Impact	Economic losses, id, lethality, physical damage, service disruption, type
Fire Ignition	Consequences	Description, id, type
Collapse	Impact	Economic losses, id, lethality, physical damage, service disruption, type
Central Escalator	Asset	Economic value, id, name, quantity, type, type_id
Fireproof Element Design	Countermeasure	CAPEX, detection score, id, mitigation score, name, OPEX, prevention score, quantity, recovery score, response score, type, type_id

Furthermore, the edges that connect the nodes with the relative attributes are reported in Annex II. Their purpose is to indicate the correlation that exists between the starting node and the arrival one. Also, sometimes the edges have an associated parameter that indicates its effectiveness. Finally, in Annex III, the code which allows to visualize the graph (Figure 4.2) has been written with the purpose of providing the possibility of easily creating the same graph.

# 5. Integration of the RA tool

## 5.1 SecuRail architecture

SecuRail is a web desktop application which will be deployed using Google Cloud Platform services. The application includes several microservices, which are in charge of undertaking specific tasks useful for the risk analysis.

In Figure 5.1 a scheme reporting the architecture of SecuRail at high level is shown. Here it is possible to identify the main component of the application, which are explained in detail in the following sections.



FIGURE 5.1: SECURAIL ARCHITECTURE SCHEME

#### 5.1.1 Databases

Graph DB is a graph database that contains information about the railway environment components and relationships. It has a static part representing standard items and features of the railway infrastructure, while it has another part which is customized by storing data entered by the user (e.g. the generic element station exists, but the user can create its own station element with different values for each attribute. It is implemented through Neo4j. Other databases include DB Rail and DB Risk. DB Rail is a SQL database which contains all the information related to the railway infrastructure. DB Risk is a SQL database which contains all the information needed for the risk analysis.

#### 5.1.2 DB Management System

This microservice oversees the exchange of data between the three databases described before. While the graph database is suitable for modelling the railway infrastructure in a very detailed and understandable manner, due to computational and development limitations it cannot be directly used as input for the risk analysis algorithms. Indeed, this service takes data from the neo4j DB and transforms them into relational data structure. The microservice exploits the neo4j ETL tool functionalities, which is specifically designed for this purpose.

#### 5.1.3 Frontend

This is the entry point of the whole application. It is an Angular application, served by a *nginx* server that is used also as a proxy to the internal backends. The User Interface is based on the *Angular Material* framework version 9.12 (AngularJS: A modern MVC framework in JavaScript, 2015).

#### 5.1.4 Authentication provider

SecuRail is relying on *Keycloak* as the external authentication provider. It is used by each component of the system all along the application workflow. In fact, the user authenticates himself to the provider, then it uses the provided token to secure each call to the microservices. The communication between the microservices is also secured with service accounts provided by *Keycloak*.

#### 5.1.5 Rail data service

This microservice keeps track of the model of the railway network built by the user in the application. It's used to save stations, lines and every railway asset of the network. Furthermore, it provides *REST API* to interact with the data, used by the frontend or by authenticated rest clients.

The Rail Data Service is written in *Java* upon *Spring* framework. It has a dedicated database (*MySQL*) where it persists the data related to the railway infrastructure of the User.

#### 5.1.6 Station search engine

The User is asked to add the stations of its own railway network to the model built in the application. Of course, each station added to the network should be a real station. To ensure the quality of the data we developed a station search engine to help the user during the network creation process.

It is a *NodeJS* application based upon *ExpressJS*. It provides API interfaces to search stations using three different types of queries:

- Free text query: the user can type the whole name of the station or a part of it.
- Stations in a specific part of the map: the user can visualize a portion of the map and find all the stations within that area.
- Stations in a specific radius around a selected point: the user can define a point on a map and find stations within a certain radius.

The three search functions are based on the following information providers:

- OpenStreetMap A GIS database with information published by users.
- Photon A text search engine based on OpenStreetMap data.
- Google places Used to provide a photo of the station.

#### 5.1.7 Risk data service

This microservice is used to keep the data structure compliant with that needed to run the risk analysis.

Indeed, it transforms the data coming from the Rail data service in a new structure more oriented to elements of the risk analysis. In this way, it is possible to define the risk analysis algorithms out of the railway context, making them multipurpose. But, at the same time, the application contains a layer to translate railway items in the needed input for risk analysis. As an example, an "asset" of the railway station is converted into a "target" which can be hit by a threat.

The Risk data service feeds the Risk Engine with all the inputs to start the risk analysis. As the Rail data service, it's written in *Java* upon *Spring* framework. Also, in this case, the data service has a dedicated database (*MySQL*) where it persists the data related to the risk analysis.

#### 5.1.8 Risk engine

This microservice is a computation engine which is in charge to run several risk analysis algorithms to calculate the output of the risk assessment. It uses as input, the data from the Risk data service and then it feeds the CBA engine with the risk analysis outputs, as well as the frontend to populate the dashboard and provide reports to the user.

The Risk Engine is written in Python and provide some REST API to interact with other SecuRail components.

#### 5.1.9 CBA engine

This microservice is fully devoted to the comparison of protection cost and risks of two different configurations of the railway infrastructure, i.e. the cost-benefit analysis (CBA). The algorithms need both data from the Rail data service (economic figures related to the railway infrastructure, e.g. CAPEX and OPEX of countermeasures) and from the Risk data service (e.g. overall risk level of a certain configuration).

The CBA Engine is written in Python and provide some REST API to interact with other SecuRail components.

#### 5.1.10 Deployment

The entire solution is deployed through Google Cloud Platform. Specifically, the solution uses:

- Cloud SQL 1 instance to host both databases (risk data service and rail data service)
- *Cloud Run* 6 instances, one for each microservice. The services are published as a docker image. Each cloud run instance will run a specific configured image stateless. This allows for easier scale up.
- Search Engine The providers used by the search engine are not hosted by us. They are public services
  with some rules for data access (rate limit or API key access)

## 5.2 SecuRail workflows

In this section, some of the workflows carried out by the SecuRail application to succeed in completing requests of the user are explained. The objective is to give to the reader an overview of how the application is intended to execute the foreseen functionalities according to the architecture explained in the previous section. The main tools used to describe the workflows within SecuRail application are:

- Sequence diagram: a type of interaction diagram because it describes how—and in what order—a group of objects works together. These diagrams are used by software developers and business professionals to understand requirements for a new system or to document an existing process.
- Activity diagram: it depicts the behavior of a system. An activity diagram portrays the control flow from a start point to a finish point showing the various decision paths that exist while the activity is being executed.

Both diagrams are following Unified Modelling Language (UML) conventions, which allows the user to specify, visualize, and document models of software systems, including their structure and design, thanks a common language shared by developers and other stakeholders<sup>1</sup>.

## 5.2.1 Add a new Station to the user network

When a user wants to add a station, he/she should first search the desired station with the search engine through the commands on the UI. See Figure 5.2 for the workflow described as a UML sequence diagram. Three types of search are foreseen: textual search, search by location through bounding box and search by location through pin point on the map.

Taking the textual search as an example, the user should type some letters of the station name on a search box and then the following steps happen:

- After 250ms from the last typed letter the word is sent to the search engine
- The search engine passes the query to a photon that searches for stations with a name similar to the one in the query

<sup>&</sup>lt;sup>1</sup> https://www.uml.org/what-is-uml.htm

- The results are sent back to the frontend, that shows a list of founded stations
- The user then selects a station
- A request of detail is sent to the search engine
- The search engine passes the query to Google Places, that find more details about this place
- The results are sent back to the frontend that shows the result.
- After the user clicks the "Add station" button, all the data about the station are sent to the rail backend
- The rail backend saves the station and asks the risk service to create a relative Target



FIGURE 5.2: SECURAIL SEQUENCE DIAGRAM FOR ADDING A NEW STATION TO THE NETWORK

#### 5.2.2 Perform risk assessment

In this scenario, the user wants to conduct a Risk analysis for a specific configuration of their own railway network. Figure 5.3 represents the work flow as a UML sequence diagram. The steps needed to achieve the goal are as follows:

- The user selects a configuration from the frontend and ask for an analysis
- The request is passed to the Risk Backend
- The service evaluates what entities are needed for the analysis and passes all the data to the risk engine
- The engine then starts the calculation procedure.
- Meanwhile the risk backend polls to get the progress of the procedure
- The progress information is sent back to the frontend until the whole result is ready and visualized



FIGURE 5.3: SECURAIL SEQUENCE DIAGRAM FOR LAUNCHING A RISK ANALYSIS

The workflow is similar also for conducting a Cost-benefit analysis, but in that case, the request is forwarded to the CBA engine. Figure 5.4 represents this request processing in detail using a UML activity diagram, with categorized blocks based on front-end, rail back-end and risk back-end components. This diagram describes all the actions to be undertaken by the user and by the SecuRail application components, from the login to the production and visualization of results. The diagram refers to three main software components:

- **Graphical User Interface (GUI)**: the means through which the user can interact with the software. Through it, the user can perform the required data entry (i.e. the modelling of the infrastructure) and send commands such as set up and launch the risk analysis.
- **Rail back-end**: this component is in charge of managing all the information related to the railway network and infrastructure. It stores data entered by the user through the GUI and it creates the corresponding data model. Furthermore, it sends data to the Risk Back-end to perform the risk analysis.
- **Risk back-end**: this component is responsible for managing the inputs and the outputs needed by the SecuRail computation engine. Data about the railway network coming from the Rail back-end are translated into a new data model suitable for risk estimation. Outputs from computation engine are than stored and elaborated to generate data for the dashboard and the report.



FIGURE 5.4: SECURAIL ACTIVITY DIAGRAM FOR OFFLINE RISK ASSESSMENT

#### 5.2.3 Real-time Risk Assessment

In this scenario, an external event detected by other monitoring tools triggers a risk analysis request. Figure 5.5 has the workflow depicted through the corresponding UML sequence diagram. Indeed, the workflow performed for executing an automatic real-time risk assessment is the following:

- An external event is submitted by the monitoring tool to the broker
- The risk service is subscribed to monitor events and trigger a risk analysis
- The analysis is done as per user request and the results are published back to the broker

When designing the real-time risk assessment tool, an added layer of monitoring is added. This consists of sensors/detectors that use the Kafka broker to communicate with the risk engine. Figure 5.6 represents this request processing in detail using a UML activity diagram, with categorized blocks based on monitoring, front-end, rail back-end and risk back-end components.



FIGURE 5.5: SECURAIL SEQUENCE DIAGRAM FOR REAL-TIME RISK ASSESSMENT

The UML activity diagram reported in Figure 5.6: SECURAIL activity diagram for real-time risk assessmentFigure 5.6 is very similar to the previous one (see Figure 5.4) but it reports also the interaction with the external module of the monitoring tool. The main workflow differences underly in the fact that the risk analysis is not launched manually by the user but it is automatically triggered by a message coming from the monitoring tool. This message is first displayed into the GUI to warn the user about the hazard and then it is used by the Risk Back-end to generate the data needed by the computation engine to perform scenario generation and estimation.





## 5.3 SecuRail interfaces

SecuRail, as one tool part of the S4RIS platform, needs dedicated interfaces to communicate with other system components.

First of all, SecuRail will be integrated into the overall S4RIS platform. As stated in D2.4, the platform will act as a single entry-point for all the applications included in the platform. Indeed, once the user enters the platform it can choose to enter the SecuRail tool and it will be automatically redirected to the URL of the stand-alone

SecuRail web application. Furthermore, it is desirable that the user should make the login just one time, i.e. before accessing to the S4RIS landing page. In this regard, there is the need to define a single sign on mechanism to be implemented in the different tools. SecuRail authentication functionalities are provided by *Keycloak*, but any authentication provider can be used to enable single sign on among S4RIS platform and all the other tools.

Besides the communication between S4RIS platform and SecuRail, the most important interface to be developed is that which allows data exchange with the monitoring tools. See Figure 5.7 for a simple flow diagram. In fact, the real-time risk assessment functionality of SecuRail should be triggered by alerts coming from the monitoring tools, which in turn analyses signals from sensors placed in the railway infrastructure to discover anomalies.

Indeed, there is a need to implement a channel for the data exchange and, together with the monitoring tools providers, a Kafka broker is selected as a suitable solution. The main steps of communication are the following:

- 1. The monitoring tool detects an anomaly
- 2. The monitoring tool publishes a message on the Kafka using the established message format
- 3. SecuRail is subscribed to the Kafka topic and it receives the message
- 4. Using the information of the message, SecuRail launches an immediate targeted risk analysis



FIGURE 5.7: FLOW DIAGRAM OF THE CONNECTIONS OF SECURAL TO THE OTHER TOOLS IN SARIS THROUGH THE KAFKA BROKER.

It is foreseen to use JSON format for the messages published on the broker. The message should contain the following information:

- Type of anomaly detected (agreed between SecuRail developers and monitoring tools providers)
- Location of the anomaly or target
- Timestamp

This information will be then used as input for risk analysis algorithms which will analyze a specific set of scenarios which could correspond to the detected situation.

# 6. Conclusion

This deliverable will lay the groundwork for the implementation of the future use cases defined in the project into the SecuRail tool as well as the implementation of the tool into the S4RIS platform. The general ideas of data modeling were introduced such as the different models and levels of complexity. From this literature review and research, a data model for SecuRail was developed.

This entity-relationship diagram, split into two different sections, incorporates the different entities that will be included in the SecuRail analysis of the railway network such as the trains, lines and timetables. In the threat and response section of this diagram, entities and their relationships are defined as they apply to threat monitoring and risk assessment. The connection between these two sections is also analyzed in the model.

The graph database that was created is based on the entity-relationship diagram and allows for further visualization of the complex and highly connected data related to the railway networks. This model was created in Neo4j and incorporates the same nodes as the entity-relationship diagram, such as the threats, consequences and impacts. The graph database is specific to each use case that is developed with the SecuRail tool. In this deliverable, the Milano use case database is visualized and future work will be to develop the other use cases. Each node in this graph database has defined attributes and examples are in Section 4, as well as examples of the edges or connections between the different nodes. The code for visualization of the graph database is ncluded in this deliverable as well.

Future work in the SAFETY4RAILS project includes extending the data model to include external critical infrastructures that are interconnected to the railway networks. This can include the power grid, the telecommunications network or other forms of transportations such as busses. This extension will help with the investigation of cascading effects of interconnected infrastructures.

The integration of the SecuRail tool began development in T3.3 as well. Section 5 describes the architecture of the tool that includes different components such as databases and database management, and a variety of different engines for risk, CBA and station searching. Specific workflows have also been developed that allow the tool to complete the user requests. These user requests can include the addition of a station to the network incorporated into the tool, as well as a risk assessment. The interface of the tool with the S4RIS platform was developed incorporated with WP6. This interface will be centered around a Kafka broker with JSON format. Next steps include further development of the connections needed for the S4RIS system.

# Bibliography

AngularJS: A modern MVC framework in JavaScript. Jain, Nimisha, Mangal, P. and Mehta, D. 2015. 2015, J. Global Research in Computer Science, Vol. 5, pp. 17-23.

**Anh, Nguyen Kim. 2009.** Database System Concepts. [Online] OpenStax CNX, July 8, 2009. https://cnx.org/contents/tXuHYGiY@1/Database-System-Concepts.

**Carter, Andy. 2003.** *EBU Technical Review: Data-modelling terminology and P\_META.* s.l. : EBU Project Group P/Meta, 2003.

**Castagna, Rich and Bigelow, Stephen. 2021.** Information Technology (IT). *SearchDataCenter.* [Online] SearchDataCenter, August 2021. https://searchdatacenter.techtarget.com/definition/IT.

**Chai, Wesley, Beaver, Kevin and Rosencrance, Linda. 2020.** Incident Response. *SearchSecurity.* [Online] TechTarget, October 2020. https://searchsecurity.techtarget.com/definition/incident-response.

**Chartered Accountants**. Risk Management. *Monitor & Review*. [Online] Chartered Accountants. https://survey.charteredaccountantsanz.com/risk\_management/small-firms/monitor.aspx.

**Collins Dictionary.** Countermeasure. [Online] HarperCollins Publishers. https://www.collinsdictionary.com/dictionary/english/countermeasure.

**contributors, Wikipedia. 2021.** Operational technology. [Online] Wikipedia, The Free Encyclopedia, July 26, 2021. https://en.wikipedia.org/wiki/Operational\_technology.

**Council of Europe. 2020.** COMMON EUROPEAN FRAMEWORK OF REFERENCE FOR LANGUAGES: LEARNING, TEACHING, ASSESSMENT. s.l. : Council of Europe, 2020.

**ENISA.** Glossary. *Risk.* [Online] ENISA. https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/risk-management-inventory/glossary.

-. Glossary. [Online] ENISA. https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/inventory/glossary.

-. Threat. *Glossary - Risk Management.* [Online] ENISA. https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/risk-management/risk-management-inventory/glossary.

**European Commission. 2018.** Questions and Answers: Directive on Security of Network and Information systems, the first EU-wide legislation on cybersecurity [Updated on 28/10/2019]. *Directive on Security of Network and Information systems.* [Online] European Commission, May 4, 2018. https://ec.europa.eu/commission/presscorner/detail/mt/MEMO\_18\_3651.

**Eurostat Statistics Explained. 2012.** Glossary: Railway line. [Online] Eurostat Statistics Explained, July 9, 2012. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Railway\_line.

Flammini, F., Gaglione, A., Mazzocca, N., & Pragliola, C. 2008. Quantitative security risk assessment and management for railway transportation infrastructures. *International Workshop on Critical Information Infrastructures Security*. 2008.

**Glinz, Martin. 2014.** A Glossary of Requirements Engineering Terminology. Zurich : International Requirements Engineering Board (IREB e.V), 2014.

Haugen, Stein and Rausand, Marvin. Risk Assessment 2. The Words of Risk Analysis. Accident scenario. [Online] Norwegian University of Science and Technology (NTNU). https://www.ntnu.edu/documents/624876/1277591044/chapt02-1.pdf/c3c0ec79-aac4-4423-bc03-10129fa5e738. **Hiermaier, Stefan, Hasenstein, Sandra and Faist, Katja. 2017.** Resilience Engineering - How to Handle the Unexpected. [book auth.] Anne-Sphie Nyssen, Mathieu Jaspar and David Woods. *Poised to Adapt: Enacting resilience potential through design, governance and organization.* Liège, Belgium : The Resilience Engineering Association and the University of Liège, Belgium, 2017, pp. 92-97.

International Organization for Standardization (ISO). 2009. ISO Guide 73: Risk Management - Vocabulary. Switzerland : ISO, 2009. ISO Guide 73:2009.

-. 2004. ISO/IEC 13335-1:2004. s.l. : International Organization for Standardization, 2004. ISO/IEC PDTR 13335-1:2004.

International Risk Management Institute, Inc. Cyber-Physical Attack Definition. *IRMI Glossary.* [Online] International Risk Management Institute, Inc. https://www.irmi.com/term/insurance-definitions/cyber-physical-attack.

**Lucidchart.** Entity-Relationship Diagram Symbols and Notation: ER diagram notation. *Lucidchart.* [Online] Lucidchart. https://www.lucidchart.com/pages/ER-diagram-symbols-and-meaning.

Merson, Paulo. 2009. *Data Model as an Architectural View.* Software Engineering Institute, Research, Technology, and System Solutions. Carnegie Mellon : s.n., 2009.

National Cancer Institute. Value: NCI Thesaurus. NCI Thesaurus. [Online] National Cancer Institute. [Cited: July 05, 2021.] https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI\_Thesaurus&ns=ncit&code=C 25712.

**National Institute of Standards and Technology.** Cyber Attack: Computer Security Resource Center. *Computer Security Resource Center.* [Online] National Institute of Standards and Technology. [Cited: July 7, 2021.] https://csrc.nist.gov/glossary/term/cyber\_attack.

**Navathe, Shamkant B. 1992.** *Evolution of Data Modeling for Databases.* Database management, College of Computing, Georgia Institute of Technology. Atlanta, GA : Communications of the ACM, 1992. p. 112.

**Neo4j Developer.** What is a Graph Database? *The Property Graph Model.* [Online] Neo4j Developer. https://neo4j.com/developer/graph-database/.

Neo4j. 2020. Neo4j Cypher Manual v4.3. [online document] San Mateo, CA : Neo4j, 2020.

neo4j. What is a Graph Database? [Online] neo4j. https://neo4j.com/developer/graph-database/.

**Neo4j.** Why Graph Databases? *What Are the Advantages of Using a Graph Database*? [Online] Neo4j. https://neo4j.com/why-graph-databases/.

**OECD. 2006.** Railway Network. *OECD Glossary of Statistical Terms.* [Online] OECD, January 4, 2006. [Cited: 27, 2021.] https://stats.oecd.org/glossary/detail.asp?ID=3913#:~:text=All%20railways%20in%20a%20given,wagon%2Dc arrying%20trailers%20or%20ferries..

**Organization of American States.** What are natural hazards? [Online] Organization of American States. https://www.oas.org/dsd/publications/unit/oea54e/ch05.htm.

**Oxford Languages**. Target . *Google target definition*. [Online] Oxford University Press. https://www.google.com/search?q=target+definition&rlz=1C1CHBF\_deDE888DE888&ei=LAYAYbjCCoGFjLs Pz9244AU&oq=target+definition&gs\_lcp=Cgdnd3Mtd2l6EAMyBwgAEEYQ-QEyAggAMgIIADICCAAyAggAMgIIADICCAAyAggAMgIIADICCAA6BwgAEEcQsAM6BQgAEJECOgQIABBD OggILhDHARCjAjoKCC4Qx.

**Oxford Languages.** Dictionary. *Target.* [Online] Oxford Languages. https://www.google.com/search?q=definition+of+target&rlz=1C1CHBF\_deDE888DE888&ei=hDk7YdfVDZmH 9u8P9Mel-

AM&oq=definition+of+target&gs\_lcp=Cgdnd3Mtd2l6EAMyBQgAEIAEMgUIABCABDIFCAAQgAQyBQgAEIAE MgUIABCABDIFCAAQgAQyBQgAEIAEMgUIABCABDIFCAAQgAQyBQgAEIAEOgUIABCRAjo.

**Piper, John.** *Risk Management Framework: Qualitative Risk Assessment through.* s.l. : NATO. STO-MP-IST-166.

**Project-management.com. 2018.** Risk or Constraint - Project Management Processes. *Risk Management.* [Online] April 15, 2018. https://project-management.com/risk-or-constraint-project-management-processes/.

Quantitative Security Risk Assessment and Management for Railway Transporation Infrastructure. F. Flammini, A. Gaglione, N. Mazzocca, C. Pragliola. 2016. 2016. 978-3-642-03552-4\_16.

**Rail Industry Safety and Standards Board (RISSB).** Section. *Glossary of Terms.* [Online] Rail Industry Safety and Standards Board (RISSB). https://www.rissb.com.au/glossary/?index=S.

-. Stakeholders. *Glossary of Terms.* [Online] Rail Industry Safety and Standards Board (RISSB). https://www.rissb.com.au/glossary/?index=S.

-. Train. *Glossary of Terms.* [Online] Rail Industry Safety and Standards Board (RISSB). https://www.rissb.com.au/glossary/?index=T.

Robinson, Ian, Webber, James and Eifrem, Emil. 2015. Graph databases. Sebastopol, Cal. : O'Reilly, 2015.

Schallehn, Dr. Eike. 2021. Data Models for Engineering Data. *Data Management for Engineering Applications.* [Online] February 16, 2021. https://www.dbse.ovgu.de/en/-p-594-EGOTEC-5g769frdai5m5a7lu9sbp4f151/\_/dmea-04-data%20models.pdf.

Simsion, Graeme and Witt, Graham. 2004. Data Modeling Essentials. Saint Louis : Elsevier Science, 2004.

**Techopedia. 2016.** Dictionary. *Attack.* [Online] Techopedia, December 15, 2016. https://www.techopedia.com/definition/6060/attack.

—. Key: Techopedia Dictionary. *Techopedia Dictionary.* [Online] Techopedia. [Cited: July 05, 2021.] https://www.techopedia.com/definition/1780/key.

-. Table: Techopedia Dictionary. *Techopedia.* [Online] Techopedia. [Cited: July 05, 2021.] https://www.techopedia.com/definition/1247/table.

**Tsichritzis, T. C. and Lochovsky, F. H. 1977.** *Database management systems: Academic Press.* New York : s.n., 1977. p. 388.

**UC Merced Library.** What Is a Data Dictionary. [Online] http://library.ucmerced.edu/data-dictionaries.

Watt, Adrienne. 2014. Data Modelling - 2nd Edition. [book auth.] Adrienne and Eng, Nelson Watt. *Database Design.* Victoria, B.C. : BCcampus. Retrieved from https://opentextbc.ca/dbdesign01, 2014.

**Wikipedia.** 2016. Entity-relationship model. *Wikipedia.* [Online] November 2016. https://en.wikipedia.org/wiki/Entity%E2%80%93relationship\_model.

# ANNEXES ANNEX I. GLOSSARY AND ACRONYMS

TABLE A.0-1: GLOSSARY AND ACRONYMS

Term	Definition/description
CAPEX	Capital expenditure
СВА	Cost-Benefit Analysis
DB	Data Base
DFD	Data Flow Diagram
DMS	Distributed Messaging System
ENISA	European Union Agency for Cybersecurity
ERD	Entity Relationship Diagram
GUI	Graphical User Interface
ID	Identifier
ISO	International Organization for Standardization
IT	Information Technology
N.A.	Not Applicable
OECD	Organisation for Economic Co-operation and Development
OES	Operator of Essential Services
OPEX	Operating expense
от	Operational Technology
RA	Risk Assessment
REST API	Representational State Transfer Application Programming Interface
RISSB	Rail Industry Safety and Standards Board
S4RIS	SAFETY4RAILS Information System
SQL	Structured Query Language
UUID	Universally Unique Identifier

# ANNEX II. Edge specifications

TABLE A.0-2: EDGE SPECIFICATIONS – PART 1

	Bombir	ng – Fire ignition		Bombi	ing – Explosion
An attack perpetrated with explosiv Threat	severity likelih descri	3 0,00256 An attack perpetrated with explosive weapons, e.g. bomb	An attack perpetrated with explosiv Threat	severity likelih descri	3 0,00256 An attack perpetrated with explosive weapons, e.g. bomb
		CAUSES		ļ	GENERATES
begin of a fire inside the facility Consequen	descri id type	begin of a fire inside the facility 1 Fire ignition	3420678	physic servic id type	0,5 72 1 Explosion
	Bomb	ing – Main hall		Patro	l 2 – Bombing
An attack perpetrated with explosiv Threat	severity likelih descri	3 0,00256 An attack perpetrated with explosive weapons, e.g. bomb	Patrol 2	OPEX respon quantity recove	140000 70 1 20
	Ļ	IS_PERPETRATED_IN		↓ is	_EFFECTIVE_AGAINST
Main hall	size type_id name geome	450 3 Main hall 45.486767048011195,	An attack perpetrated with explosiv Threat	severity likelih descri	3 0,00256 An attack perpetrated with explosive weapons, e.g. bomb
	Main spr	inkler – Main hall		Explos	sion – Main hall
efficiencyRate	a 90		3420678	physic servic	0,5 72
Main Sprinkle	respon OPEX quantity	40 2000 1	Impact	type	T Explosion
	Antart	PROTECTS		size	AFFECTS
Main hall	size type_id name geome	450 3 Main hall 45.486767048011195,	Main hall	type_id name geome	450 3 Main hall 45,486767048011195,

21. 10. 21. 21.	Dilana Daisse Cardinaldi		10.000.0

230;700;1410;2120;2250;1870;1...

Milano

Centrale

Station

Main hall

1

Section

Milano

Centrale

1

ctior

Milano Porta

Garibaldi

Station

Minalo Centrale - Main hall

Lombardia

INCLUDES

250;1300;2410;3720;3230;2570;...

25

450

Main hall

Milano Centrale (Starts From)

STARTS\_FROM

250;1300;2410;3720;3230;2570;...

230;700;1410;2120;2250;1870;1...

Lombardia

25

5.6

3

5

1

Milano

Section 1 - Milano Porta Garibaldi (Starts From)

STARTS\_FROM

Lombardia

13

Milano

5.6

3

5

1

45.486767048011195.

3

Milano

country

numbe...

crowdi...

city

size

type\_id

geome...

length

percen...

percen...

country

numbe...

crowdi...

city

length

percen...

percen...

country

numbe...

crowdi...

city

id

id

name

Main hall - Ticket machine

Main hall

45.486767048011195,

Ticket Machines Group A

INCLUDES

450

3

6

5

5.6

3

5

1

Section 1 - Milano Centrale (Arrives To)

ARRIVES TO

250;1300;2410;3720;3230;2570;...

Lombardia

25

5.6

3

5

1

Milano

Section 1 - Milano Porta Garibaldi (Arrives To)

ARRIVES\_TO

Lombardia

13

Milano

size

type\_id

geome...

quantity

type\_id

name

length

percen...

percen...

country

numbe...

crowdi...

length

percen...

percen...

country

numbe...

crowdi...

city

id

city

id

id

name

Main hall

Ticket

Machines

Group A

Asset

1

Section

Milano

Centrale

Station

1

Section

Milano Porta

Garibaldi

Station

	Patro	ol 2 – Main hall		Explos	sion – Main hall
efficiencyRate	99 OPEX respon guantity	140000 70 1	3420678	physic servic id type	0,5 72 1 Explosion
Counterme	recove	PROTECTS			AFFECTS
Main hall	size type_id name geome	450 3 Main hall 45 486767048011195	Main hall	size type_id name geome	450 3 Main hall 45.486767048011195,
E	xplosio	n – Ticket machine	N	lain spriı	nkler – Fire ignition
3420678	physic servic	0,5 72	effectivenessRa	ate 90	
Impact	type	Explosion	Main Sprinkler	respon OPEX quantity	40 2000 1
		AFFECTS	Counterme	recove	0
Ticket Machines Group A	quantity type_id	6 5 Ticket Machines Group A		IS_	_EFFECTIVE_AGAINST
Asset	id	1 +	begin of a fire inside the facility Consequen	descri id type	begin of a fire inside the facility 1 Fire ignition
_	Fire igr	nition – Collapse	Fi	reproof s	structure – Collapse
begin of a fire inside the facility Consequen	descri id type	begin of a fire inside the facility 1 Fire ignition	Fireproof	respon OPEX	10
	ļ	GENERATES	Counterme	quantity detect	1
234578	physic servic	0,8 12		IS	_EFFECTIVE_AGAINST
Impact	type	2 Collapse	234578	physic servic	0,8 12
			Impact	type	2 Collapse

Firep	roof structure – Central escalator	Collapse – Central escalator
efficiencyRate	95	234578 physic 0,8 servic 12
Fireproof structure	respon 10 OPEX 1300	Impact Id 2 type Collapse
Counterme	quantity 1 detect 0	AFFECTS
		Central <b>quantity</b> 2 Escalator <b>type_id</b> 10
Central Escalator	quantity 2 type_id 10	Asset id 1
Asset	name Central Escalator id 1	

# ANNEX III. Code for visualization

Lines of code for visualizing the graph (as mentioned in section 4):

#### create

//ASSET

(as1:Asset {id: "1", type\_id: "5", type: "Ticket Machine", name: "Ticket Machines Group A", economicValue: "7800", quantity: "6"}),

(as2:Asset {id: "1", type\_id: "10", type: "Escalator", name: "Central Escalator", economicValue: "15600", quantity: "2"}),

#### //COUNTERMEASURE

(c1:Countermeasure {id: "1", type\_id: "1", type: "Security Patrol", name: "Patrol 2", CAPEX: "32000", OPEX: "140000", quantity: "1", preventionScore: "60", detectionScore: "80", responseScore: "70", mitigationScore: "40", recoveryScore: "20"}), (c2:Countermeasure {id: "1", type\_id: "1", type: "sprinkler", name: "Main Sprinkler", CAPEX: "58200", OPEX: "2000", quantity: "1", preventionScore: "20", detectionScore: "90", responseScore: "40", mitigationScore: "50", recoveryScore: "0"}), (c3:Countermeasure {id: "1", type\_id: "24", type: "Fireproof element design", name: "Fireproof structure", CAPEX: "5200", OPEX: "1300", quantity: "1", preventionScore: "30", detectionScore: "0", responseScore: "10", mitigationScore: "60", recoveryScore: "0"}),

#### //AREA

(ar1:Area {id: "1", type\_id: "3", type: "Hall", name: "Main hall", economicValue: "1200", size: "450", geometry: "45.486767048011195, 9.203709144885954;45.48768713683475, 9.204772306602763;45.48674061045924, 9.206592767694103;45.48581124642534, 9.205792478091945"}),

#### //STATION

(st1:Station {id: "1", name: "Milano Centrale", coordinates: "45.48719277939978, 9.205436090960848", country: "Lombardia", city: "Milano", numberOfPlatform: "25", crowding:

"250;1300;2410;3720;3230;2570;2180;1600;1540;1320;1890;2130;2910;4200;3290;1780;1060;680;290" }), (st2:Station {id: "2", name: "Milano Porta Garibaldi", coordinates: "45.48460091094036, 9.187366646434361", country: "Lombardia", city: "Milano", numberOfPlatform: "13", crowding: "230;700;1410;2120;2250;1870;1280;670;590;720;1190;1430;2110;2230;1560;780;460;380;180" }),

#### //SECTION

(se1:Section {id: "1", length: "5.6", maxSpeed: "120", numberOfTracks: "9", percentangeBridge: "3", percentageTunnel: "5", numberOfLevelCrossing: "0"}),

#### //THREAT

(t1:Threat {id: "1", type: "Bombing", likelihood: "0,00256", severity: "3", description:"An attack perpetrated with explosive weapons, e.g. bomb"}),

//CONSEQUENCE (co1: Consequence {id: "1", type: "Fire ignition", description: "begin of a fire inside the facility"}),

//IMPACT (i1: Impact {id: "1", type: "Explosion", physicalDamage:"0,5", lethality: "0,6", serviceDisruption: "72", economicLosses: "3420678"}), (i2: Impact {id: "2", type: "Collapse", physicalDamage:"0,8", lethality: "0,2", serviceDisruption: "12", economicLosses: "234578"}),

//Section-Station (se1)-[:STARTS\_FROM]->(st1), (se1)-[:ARRIVES\_TO]->(st2), (se1)-[:STARTS\_FROM]->(st2), (se1)-[:ARRIVES\_TO]->(st1),

//Station-Area (st1)-[:INCLUDES]->(ar1),

//Area-Asset (ar1)-[:INCLUDES]->(as1),

//Countermeasure-Area (c1)-[:PROTECTS {efficiencyRate: "99"}]->(ar1), (c2)-[:PROTECTS {efficiencyRate: "90"}]->(ar1),

//Countermeasure-Asset (c3)-[:PROTECTS {efficiencyRate: "95"}]->(as2),

//Threat-Area (t1)-[:IS\_PERPETRATED\_IN]->(ar1),

//Threat-Impact (t1)-[:GENERATES]->(i1),

//Impact-Area (i1)-[:AFFECTS]->(ar1),

//Impact-Assets (i1)-[:AFFECTS]->(as1), (i2)-[:AFFECTS]->(as2),

//Countermeasure-Threat
(c1)-[:IS\_EFFECTIVE\_AGAINST {effectivenessRate: "90"}]->(t1),
(c2)-[:IS\_EFFECTIVE\_AGAINST {effectivenessRate: "90"}]->(c01),
(c3)-[:IS\_EFFECTIVE\_AGAINST {effectivenessRate: "70"}]->(i2),

//Threat-consequence (t1)-[:CAUSES]->(co1),

//Consequence-Impact
(co1)-[:GENERATES]->(i2)



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