Contract No. H2020 – 826172

SEMANTICS FOR PERFORMANT AND SCALABLE INTEROPERABILITY OF MULTIMODAL TRANSPORT

D3.2 – Performance and Scalability Requirements for the IF

Due date of deliverable: 31/07/2019

Actual submission date: 24/04/2020

Leader/Responsible of this Deliverable: POLIMI

Reviewed: Y

Document status		
Revision	**Date**	**Description**
1	31/07/2019	First issue
2	09/08/2019	Complete version, ready for TMC approval
3	21/08/2019	Final modifications, awaiting TMC approval
4	14/04/2020	Updated version after reviewer comments
5	24/04/2020	Final version after TMC approval and quality check

Project funded from the European Union's Horizon 2020 research and innovation programme		
Dissemination Level		
PU	Public	X
CO	Confidential, restricted under conditions set out in Model Grant Agreement	
CI	Classified, information as referred to in Commission Decision 2001/844/EC	

Start date of project: 01/12/2018 Duration: 25 months

EXECUTIVE SUMMARY

This deliverable describes the analysis of the requirements of the Interoperability Framework (IF). In particular, the deliverable focuses on the performance and scalability requirements of the IF, but it also deals with other types of non-functional requirements, such and privacy and security requirements, usability requirements, and so on.

The analysis starts with the identification of the stakeholders of the IF, an continues with the description of the key scenarios of use of the IF. From the scenarios, a set of requirements and associated Key Performance Indicators are derived.

ABBREVIATIONS AND ACRONYMS

Abbreviation	Description
CP	Carpooling provider
EU	European Union
GA	Grant Agreement
H2020	Horizon 2020 framework programme
KPI	Key Performance Indicator
IF	Interoperability Framework
IM	Infrastructure manager
IT	Information Tehcnology
JU	Shift2Rail Joint Undertaking
MaaS	Mobility as a Service
NAP	National Access Point
SSP	Sharing schemes providers
TEP	Travel Event Provider
TOA	Transport organizing authority
TSP	Transport service provider
TrSP	Travel service provider

TABLE OF CONTENTS

LIST OF TABLES

1. INTRODUCTION

This deliverable describes the analysis of the requirements of the Interoperability Framework (IF) that has been carried out within the SPRINT project. In particular, the analysis focuses on so-called “non-functional requirements”[1] [1]. To drive the analysis, the following, classic approach [2] has been followed:

- First, the relevant stakeholders of the system (the IF in our case) have been identified (Section 2).
- Then, a set of scenarios capturing the main ways in which the IF will be used by the relevant stakeholders has been identified (Section 3).
- From the scenarios, a set of requirements has been identified, (Section 4); in addition, a set of Key Performance Indicators (KPI) has been defined.

The analysis of Section 4 focuses on requirements concerning the performance and scalability of the IF. However, other types of non-functional requirements (concerning security and privacy aspects, usability of features by the IF users, and so on) are also relevant for the IF. An analysis of these requirements is carried out in Section 5.

The requirements and KPIs identified in this deliverable will drive the definition of the testing infrastructure and of the test cases that is the subject of Task 3.3 “Design of the performance and scalability testing infrastructure”, and which will be first described in deliverable D3.3 “Design of Architecture, Testing Infrastructure, Test Cases and Benchmarks of the IF (CREL)”.

In addition, the scenarios identified in this deliverable will be the basis for the demonstration scenarios that will be defined in Task 5.1 “Design of proof-of-concept scenarios”, and that will be first detailed in Deliverable 5.1 “Requirements, scenarios and use cases for the proof-of-concept (CREL)”

[1] It is customary to refer to “functional requirements” as to those concerning the outcomes of the computations of the system; “non-functional requirements”, instead, are all other constraints that instead concerns aspects such as timing, performance, scalability, security, user-friendliness, etc.

2. STAKEHOLDER IDENTIFICATION

This chapter is devoted to the description of the broad set of IF stakeholders. The chapter is based on the outcomes of the GOF4R project, and in particular of Deliverable D2.2 "Analysis of the demand of market actors for the Interoperability Framework". For the sake of completeness, we aimed to provide a comprehensive category of actors in this section. However, the user-stories in Section 3 include a subset of the identified actors in Table 1 and Table 2 since are mainly centered around the assessment of *primary* stakeholders who are most probable to engage with the IF.

Service/infrastructure Providers

Type of the stakeholders	Description and key characteristics
Transport organizing authorities (TOA)	A government or public agency created to perform a single function or restricted group of related activities. The authority pertains to the government entity that is responsible for the organization of the public transport market. It is responsible for transport fare level, route designations and other public transport operating system policies, supervision, regulation and award of operating contracts and franchises. In some cases, the transport operating company and the authority are within the same government unit and perform policy, regulatory, planning, and operating functions. In other cases, the authority is a separate public agency that does not have any operating responsibilities but establishes public transport system policies and acts as the system's regulator[2]. TOA can set up the rules at various territorial levels: local (e.g., city level), national (e.g., country level), European or international for a sustainable mobility and interested in developing an overall market framework and especially an IF facilitating an increase in the use of multimodal seamless mobility solutions combining several travel services.
Transport service provider (TSP) (Transport Operation Carrier)	An individual or an entity, such as a corporation or a partnership, in the business of providing public transport services against payment by the passengers and/or the authority[3]. A TSP is responsible for delivering the transport service and managing the drivers and the operating staff.
Infrastructure manager (IM)	An IM is a body responsible for establishing and maintaining infrastructure. It may also include the management of infrastructure control and safety systems. At the

[2] UITP Glossary of public transport terms – Last update: February 2014

[3] UITP Glossary of public transport terms – Last update: February 2014

	urban scale, it may be often a case that TSP and IM are accomplished by the same company.
Sharing schemes providers (SSP)	One-way-capable rental offer in public spaces, for several target groups, with network characteristics[4]. SSP is a complementary mode of transport.
Carpooling providers (CP)	An arrangement where two or more people share the use and cost of privately-owned vehicles when travelling together to and from pre-arranged destinations[5]. CP is a complementary mode of transport.
Demand responsive (or dial-a-ride) service	Such a service is run with passenger cars, vans or small buses operating in response to calls from passengers to the public transport operator, who then dispatches a vehicle to pick up the passengers and transport them to their destinations. A demand responsive operation is characterized by the following: (a) The vehicles do not operate over a fixed route or on a fixed schedule except, perhaps, on a temporary basis to satisfy a special need; and (b) typically, the vehicle may be dispatched to pick up several passengers at different pick-up points before taking them to their respective destinations and may even be interrupted on route to these destinations to pick up other passengers. These services could be required by law for persons with disabilities and others not able to use fixed-route services[6].
Retailer, Travel Agency, Distributor	A Distributor owns and/or operates a computerized network system to enable transactions between Retailers, Travel Agencies and TSPs. In the context of mobility services that may be distinguished between Technical and Commercial Distributors. A Technical Distributor is essentially a provider of ICT capabilities facilitating the transmission of information between Retailers and TSPs and offering technical services to support Retailers and Travel Agencies in their everyday operations. It does not own or select Retailers or Travel Agencies and does not have responsibility for financial settling of sales. It can be a Global Distribution System. A Commercial Distributor sells services on behalf of TSPs through its own network of Retailers and has full responsibility for financial settlements with relevant TSPs for any sales made through its Retailer network. A Retailer sells the services of TSPs through the ICT capabilities of Distributors. A Retailer may have a direct commercial relationship with a TSPs, supported by a Technical Distributor, whereby it acts as the TSP's appointed agent, and/or it may have an indirect commercial relationship with TSPs mediated by a Commercial Distributor. Retailers handle the relationship with the Customer, including for payment by the customer, ticket delivery and initiating after-Sales

[4] UITP Glossary of public transport terms – Last update: February 2014

[5] UITP Glossary of public transport terms – Last update: February 2014

[6] UITP Glossary of public transport terms – Last update: February 2014

	activities. In the legal context, the Retailer concludes the Contract of Carriage between the Passenger and the Contractual Carrier(s) in the name of and for the account of the latter as a function of the sales agreement which the Retailer has with the TSP, possibly using a Commercial/Technical Distributor as intermediary. A Travel Agent is a Retailer of TSP services which, in addition, it integrates with other travel-related services such as insurance, medical assistance, emergency services, etc., for individual or business customers.
Mobility as a Service (MaaS) providers	A company that operates a mobility distribution model that deliver users' transport needs through a single interface, combining different transport modes to offer a tailored mobility package, or "bundle", through a variety of contractual schemes, such as pay-per-use, subscriptions, etc. MaaS providers may offer customers the ability to customize bundle contents and contractual terms.

Table 1 Stakeholders in Service/infrastructure Providers

Applications/Consumers

Type of the stakeholders	Description and key characteristics
Customers/Travelers	The Customer buys services supplied by TSPs in the name of and for the benefit of Passengers. The Customer, which can be the Passenger, commits to terms and conditions, such as payment, associated with the supply of the Products. A Passenger is the individual actually consuming the services acquired by Customers. In a Mobility-as-a-Service scenario, Customer is a party to the contract for bundled services supplied by a MaaS provider. Customers and Passengers can benefit from door-2-door multimodal services (shopping, tracking, ticketing, on-line payment, etc.). However, the Customer is not interested directly in the IF, but in the services he/she can have access to through a Travel Companion or through a "one-stop-shop" for a trip between a geographical origin and a geographical destination.
Travel service provider (TrSP)	Entity offering travel services (very often online), especially – but not exclusively – on transportation, e.g., providing travel data or other services as a third party. A TrSP can provide the following services to the customers: • Journey planning • Offer building • Trip tricking • Travel companion • Travel Event Provider (TEP)
IT supplier and Software Application (ISA)	An individual or entity that uses the possibilities of the IF (e.g., asset manager) to create new applications, web-services, etc.

Community groups or Social networks	Community groups and social networks (e.g., Twitter, Facebook) play a social role. They contribute to the improvement of the services (e.g., alert functions) and to the choice of the service providers (e.g., transport services and travel services related to transport).

Table 2 Stakeholders in Applications/Consumers

3. IF USER STORIES

In this section, we describe various use-case stories that help us to identify the main design objectives for the development of the IF and to analyze the corresponding major challenges and requirements. In Section 3.1, we list the so-called Basic Scenarios, which indicate the primary functions and accordingly the high-priority requirements that must be addressed in early versions of the IF. Section 3.2, instead, is focused on more sophisticated requirements. It pictures more complex situations to pave the ground for more advanced functions and features of the IF that might be gradually added to IF as it grows.

All scenarios in this section fall in three categories of interactions with the IF:

- **Joining Phase**: It is the starting point and the first compulsory course of actions to be taken by a would-be client of the IF (user/organization) when it joins for the first time the IF ecosystem with its desired roles (Provider/ Consumer).
- **Adaptation Phase**: This type of scenarios includes all the build-time interactions among the IF and a client.
- **Engaging Phase:** This type of scenarios includes all the run-time interactions among the IF and a client.

3.1 BASIC SCENARIOS

Join and search	
Type	Engaging Phase
Actors	**HT-train:** An Italian-based *TSP*, in specific a train service provider
Story	**HT-train** is a train operator based in Italy with a strong offer in historic train travel experiences. They decide to expand their market reach by offering the possibility to reach the departing stations via multiple means of transportation. They decide to join the Shift2Rail ecosystem to be able to build multimodal travel offers and attract more people on their historic trains, and therefore registers with a IF node. They then decide to publish their services on the IF to let other companies discover their services.
Challenge	HT-train needs to register itself as a company with the IF node and is requested to provide relevant information. The joining phase to the S2R ecosystem is regulated by a specific process which checks the quality of the data provided by the new applicant. Upon approval, HT-train must be able to quickly identify relevant assets to start expanding their offerings. The IF node must offer a fine-grained search feature to let HT-train identify other operators serving the departing stations of their touristic historical trains.

Goal	**HT-train** desires to obtain fine-grained information about which services they need to contact to create multi-modal travel solutions.
Involved IF Components	Asset Manager

Table 3 Basic Scenario 1

Distributed service/asset discovery	
Type	Engaging Phase
Actors	**B-Com:** A Belgian-based TrSP which it is hosted by the (National Access Point) NAP of Belgium. **S-com:** A Spanish-based TSP named "S-com" already hosted by the NAP of Spain. S-com offers travel services within and beyond the Spain boundaries. **MyMobility:** An Italian company providing transport services in many regions of Europe.
Story	Summer is approaching and MyMobility anticipates many travelers may target Spain and Belgium's towns. Accordingly, it would like to retrieve more details about catalogues describing data of public transport providers and information of the different public means of transport with Spanish and Belgian travel operators in order to tailor offers and facilities. Accordingly, services offered by S-com and B-com might be interesting for MyMobility. MyMobility searches to discover the desired service/data by querying the IF-node with which it has registered (which in turn federates the query across all IF-Nodes).
Challenge	The services provided by S-com/B-com are stored in one instance of the IF, while MyMobility is initiating the discovery query from a geographically, computationally and administratively separate IF-Node. Hence, the challenge here is to perform a distributed service discovery and grant the maximum service visibility and discovery coverage. In general, a TSP registers itself and its services on one instance IF-Node (e.g., the IF-Node hosted by the NAP of the TSP's country of origin) but its motivation to offer a service is mainly business-oriented and it does not limit to the boundary of a certain country. Accordingly, the potential users of such services might be distributed over various IF-Nodes across Europe.

Goal	It is not desirable and feasible for S-com/B-com to register each of its services to the IF-Node that is more probable to be searched by a client. A generic service provider desires to register its services on the IF-Node in which it is known, and then such a service should be discoverable in any other instance of the IF. It is not desirable and feasible for MyMobility to query the IF-Node that is more probable to contain the desired service. A generic service consumer desires to initiate its search request in the IF-Node that it knows, but the result should include all the services matched with its request across all the IF-Nodes.
Involved IF Components	Asset Manager

Table 4 Basic Scenario 2

Batch Data Conversion	
Type	Adaptation Phase
Actors	**S-com:** Spain-based TSP registered as a service provider on an instance of the IF which is hosted by the NAP of Spain.
Story	S-com's travel services cover a wide variety of services in Spain. Recently, the manager of the company has decided to expand the company by offering services for the whole Europe. Accordingly, for S-com to obtain maximum coverage of potential clients from any region (in Europe), it is required to publish their data including code lists and timetables in a representation/data model compatible with the target consumers' systems and standards. However, such data in the S-com company are not totally static, even though changes are not highly frequent. For instance, timetables change in a monthly manner and code lists on a yearly basis.
Challenge	Most of S-com's IT and logistic infrastructure has been established over many years, and it is mature and fully functioning. Hence, swapping the standards and data models which are the fundamental elements of many higher-level services/functions would lead to a major change in the whole system that would be costly and time-consuming. In addition, even if it were economically reasonable for S-com to re-build everything to move to a different standard, given the heterogeneity of available standards it would still be unfeasible to make its infrastructure compatible and interoperable with all standards!
Goal	S-com desires to convert its codes lists, timetables and other data to other (multiple) data models without revamping its entire software infrastructure.

	S-com prefers to continue to publish/update its data in its current regular basis, and after each publication/alteration, convert the whole dataset into the target model and make it available for interested parties.
Involved IF Components	Converter

Table 5 Basic Scenario 3

Runtime Data/Message Conversion	
Type	Engaging Phase
Actors	**T-A-n:** ISA (Travel Applications for smartphone) **T-O-n:** TSP **BE-Service:** *TrSP* (Offer building service) for land (rail, bus, etc.) travels within central Europe. Its front-end API is used by mobile and web applications (say T-A-1 to T-A-10) and its back-end has access to, and engaged with many train/bus operators (say T-O-1 to T-O-20) in the covered zones.
Story	Alice has T-A-1 installed in her smartphone and initiates a booking request through the dedicated graphical user interface. Upon such request, T-A-1's back-end calls BE-Service to find a list of best offers. BE-Service then returns some offers provided by various T-Os for the requested path. Ultimately Alice (through her T-A-1) opts to buy a ticket from T-O-2. However, the booking standard (data-model, format, terminologies) that is hard-coded in the back-end logic of T-A-1 is different and incompatible with the API of T-O-2. BE-Service hence required to convert the source booking request/confirmation format to the target model – and vice versa – instantly at runtime.
	While the formats/standards that must be converted to each other are predictable for BE-Service, the frequency and types of conversion (i.e., booking, cancellation, tracking, event notification, etc.) varies based on the different parameters such as time of the day, session, etc. In addition, the number and heterogeneity of both BE-Service's clients (i.e., Travel Applications) and back-end services (i.e., Travel Operators) are growing. Consequently, it is not feasible for BE-Service to develop and hard-code the mediators to facilitate the interaction among both sides of the transaction for each and every data model, format and standard.

	The conversion between the source and target format of the booking request/confirmation message is part of a live and runtime transaction. Hence, such a process must be completed fast.
Goal	BE-Service desires to engage with the converter in a loose-coupling manner and utilize it as an external service to its main business logic. BE-Service may discover the right converters in the IF offline, but it desires to utilize as many of the selected ones as necessary at runtime. Accordingly, it desires to follow a pay-as-you-go scheme (in the case of use of commercial assets, if any).
Involved IF Components	Converter

Table 6 Basic Scenario 4

Fast adaptation to peaks	
Type	Engaging Phase
Actors	**BE-Service:** *TrSP* (Offer building service) for land (rail, bus, etc.) travels within central parts of Europe. Its front-end API is used by mobile and web applications (say, T-A-1 to T-A-10) and its back-end has access to, and is engaged with many train/bus operators (say, T- O-1 to T-O-20) in the covered zones.
Story	One of the cities covered by BE-Service is hosting a huge music event, and BE-Service expects a surge of booking requests. BE-Service, therefore, needs to cope with two different scenarios: prepare for the first wave of requests to reach the city, and then to cope with mass requests to reach the music event before its start and to reach the homes and hotels after its end.
Challenge	The infrastructure managing the converters deployed by BE-Service to interact with its partner operators need to dynamically adapt to the load. BE-Service needs to quickly replicate Converters, possibly in a cloud environment, to adapt the infrastructure and avoid denial of service.
Goal	Users reaching the music event (or going back from it) must be able to interact with BE-Services even during an exceptional peak of requests.
Involved IF Components	Converter

Table 7 Basic Scenario 5

Special purpose discovery package	
Type	Adaptation Phase
Actors	**A-comp:** *TSP* **TS-S:** A *TrSP* (Journey Planning) to orchestrate the process of travel shopping initiated by the frontend application such as a travel companion.
Story	A-comp is a company offering various travel services. Recently they have decided to launch a new travel shopping service, TS-S, that manages mobility and travel rights delivery, booking and ticketing. It works as a mediator between the front-end application such as the travel companion and the actual expert operators of a certain region. TS-S must retrieve the list of stop places available in a certain radius of a given point of interest initially queried by the user via the front-end application and ultimately extract Travel Experts covering the available transport infrastructure such as buses, trains and airports of that region and offering various Offer Items and other services for a given route/region.
Challenge	The access, representation, relocation and replication of desired data is hugely fragmented across distributed databases, triple stores and graphs.
Goal	The main goal in this scenario is to mask interacting applications from the physical distribution, access protocols and formats of meta-data and data resources available in the Data Layer. To achieve this goal, it is required to provide access, representation, relocation and replication transparency features. Another goal is to identify networked Travel Experts participating in a coordinated distributed shopping and booking one-stop-shop instance. To achieve this goal, we envision "packaged" data and service discovery.
Involved IF Components	Location resolver and Travel Expert Resolver

Table 8 Basic Scenario 6

Automated Mapping Process for the Conversion	
Type	Adaptation Phase
Actors	**I-organization:** TOA covering the Milan area
Story	Mary is a software developer at I-organization. I-organization has recently applied some new rules and services for the TSPs of Milan as well as a new Data Model (DM-T) and ontology (O-T). However, they are aware of TRANSMODEL and a popular data

	models that are in-use by many target audiences. Accordingly, Mary would like to create a converter for DM-T to TRANSMODEL in order to encourage the consumers (in this case the TSPs) to engage with the new services.
Challenge	Currently, to create a Converter, one should be an expert on both the source and target data model / ontology, so s/he can extract the shared concepts among them and map the terms in the source model to the equivalent concepts in the target model. Clearly, having deep knowledge of both source and target model is time- and cost-consuming.
Goal	Mary would like to benefit from a more automated approach for the extraction of equivalent concepts in the source and target models.
Involved IF Components	Converter/ Mapping IDE

Table 9 Basic Scenario 7

3.2 Advanced Scenarios

Federated identification and access control	
Type	Joining Phase
Actors	**X-com:** *TSP* **BE-Service:** *TrSP* (Offer building service) for rail and road travel within central parts of Europe. Its front-end API is used by mobile and web applications (say, T-A-1 to T-A-10) and its back-end has access to, and, engaged with many train/bus operators (say, T- O-1 to T-O-20) in the covered zones.
Story	Alex is the head of IT department of X-com, which is one of the main rail operators of Country X. He recently has heard about the IF and he is interested to join the ecosystem to offer its travel services as well as various software and IT products/assets to other operators and travel applications, and to learn about their services and assets. However, X-com's services are accessible only under specific use-conditions. Similarly, the company supplies various types of data including ontologies, meta data, service descriptions, and so on. Though the access to some data is fully open, valued and sensitive data are available only to authorized users. The manager of BE-Service is also interested in being involved in the IF ecosystem in order to use its benefits, especially for easier/interoperable engagement with rail/road operators (see basic scenario 5) as well as to enjoy benefits of interoperability services such as converters and resolvers. However, BE-Service is already a client of X-com

	and some other operators, which means it already has gone through the registration process of such organizations.
Challenge	From the service provider viewpoint: In general, not only every single organization/ service provider such as X-com, but even each data source has its own well-established registration process and it applies its own policy and access control strategy, which is not necessarily the same as others.
	From the service consumer viewpoint: The heterogeneity of access control mechanisms of service providers, plus the existence of intermediate organizations running their own user databases (which impose their own user registration mechanisms in addition to the destination organization) increases the complexity of the situation. As a result, a service consumer is required to repeatedly register to each and every service/data provider and follow a separate authentication process for each of the service/data requests it has.
Goal	X-com desires to define (in the IF) the access conditions to its resources according to its organization policy, which might be different from other organizations.
	A typical service consumer desires to minimize the registration and authentication process and to be identified globally and in a cross-organization manner.
Involved IF Components	Asset Manager

Table 10 Advanced Scenario 1

Ad-hoc Converter creation	
Type	Engaging Phase
Actors	**N-rail:** a rail *TSP* which just joined the Shift2Rail ecosystem. **Y-bus and X-bus**: bus *TSPs* already part of the Shift2Rail ecosystem.
Story	Y-bus services joined the S2R ecosystem and contributed a Converter to let its clients interact with X-bus, an allied bus operator. It does so by providing a mapping which “lifts” its own data model to the Shift2Rail ontology, and also a mapping which “lowers” instances of the S2R ontology to the X-bus data model.

	N-rail has just completed mapping its data model onto the Shift2Rail ontology, and now wants to interact with X-bus services. N-rail then provides the Asset manager its mappings asking if they are sufficient to interact with X-bus. The Asset manager assesses such possibility and creates a Converter to let N-rail contact X-bus.
Challenge	The Asset Manager needs to identify fragments of Converters that can be reused to automatically build a new Converter. In details, the Asset Manager needs to check whether the source (or destination) message that needs to be converted is "covered" by a mapping. Then, the Asset Manager must compose a process to implement the message-to-message conversion, gather all the required assets and resources, and build a new deployable artifact.
Goal	N-rail wants to obtain a new Converter reusing parts of already existing (and tested) Converters to reduce the implementation and integration effort.
Involved IF Components	Asset Manager

Table 11 Advanced Scenario 2

4. PERFORMANCE AND SCALABILITY REQUIREMENTS

In section 2 we have gathered the most probable scenarios for communicating with the IF that covered possible interactions in three phases with each IF component. This section analyzes *performance* and *scalability* requirements grouped by component.

Performance and scalability are often categorized as non-functional requirements in many requirement analysis classifications [3]. The term "non-functional requirement" in general is used to refer to a requirement that specifies system properties and physical constraints on a functional requirement [4]. Table 12 describes performance and scalability requirements, along with their most common quality attributes.

Requirement	descriptions	Quality attributes
Performance	How well the system utilizes a resource	Efficiency
	How secure is each operation? How the required task/procedure to grant the security of an operation may affect its efficiency?	Integrity
	What is the confidence level of performing a process?	Reliability
	How easy is to work	Usability
Scalability	How well does the system perform as the number of users/resources/transactions increases?	survivability
	How well does the system performs under adverse conditions?	

Table 12 Performance and Scalability Requirement

Although scalability and performance are categorized as two different properties of a software system, they are highly correlated. For a given environment that consists of properly sized hardware, properly configured operating system, and dependent middleware, if the performance of a software system deteriorates rapidly with increasing load (number of users or volume of transactions) prior to reaching the intended load level, then it is not scalable and it will eventually underperform [5].

In other words, **performance** measures how fast and efficiently a software system can complete certain computing tasks, while **scalability** measures the trend of performance with increasing load. Subsequently, in the rest of the document we first stipulate the required performance properties of the system and then we analyze how such property may be influenced by scalability requirements.

4.1 Extracted Key Performance Indicators and corresponding Scalability requirements for Asset Manager

4.1.1 Performance

Referenced Scenario: Join and search, Distributed service/asset discovery	
KPI	Asset/Service Discovery Response Time
Definition	Time required to perform a SPARQL query on the RDF metadata repository.
Requirement	The activity of looking for assets in the Asset Manager (both as a stand-alone server and in a distributed environment) is likely to be the most used feature related to the catalogue. A discovery task implies multiple retries of a simple search operation, each time adding filters to the previous try. The execution time for a single query on metadata managed by the Asset Manager should therefore not exceed 10 seconds to avoid frustrating users during the discovery task.
Target Value	The execution time for a single query on metadata managed by the Asset Manager depends on the number of assets and the query complexity. Since a query for AM corresponds to a simple search operation and filters, the response time will be low. For example, for 25.000 assets, average response time must be less than 10 seconds. Worst case: 10 seconds
Referenced Scenario: Ad-hoc converter creation	
KPI	Converter artifact generation time
Definition	Time required to generate the Converter downloadable artifact.
Requirement	This scenario requires assembling a downloadable artifact gathering both local and remote resources. Moreover, the request to build a Converter can be governed by a dedicated process requiring the explicit consent of the owners of the referenced assets. In case of immediate consent of the referenced assets owners, we estimate a limit of one hour to complete this operation.
Target Value	Worst case: 1 hour

Table 13 Performance requirements for Asset Manager

4.1.2 Scalability

Referenced Scenario: Join and search	
KPI	Query/search time
Definition	Time required to perform a SPARQL query on the RDF metadata repository.
Requirement	As the discovery/search capabilities are implemented by means of SPARQL queries, the execution time for such queries must stay constant, or at worst grow linearly as the metadata repository size grows.
Target Value	The time depends on the number of results gathered for each query on a metadata repository. Given a metadata repository X and a query Q whose execution time is T1, then the execution time of Q for a metadata repository X' with scale N (X scaled N times) following the same data distribution as X, must be T*N where T is the assumable and desirable time for each increment in the scale of the metadata repository. To estimate the assumable time, the time T1 must be considered. For example, if the assumable time is between 1 and 3 sec and the average time for a query is less than 10 seconds, then for the worst case, the average response time over a metadata repository scaled 5 times must be less than 15 secs (3 sec*5 times) and the average response time over a metadata repository scaled 10 times must be less than 30 secs (3 sec*10 times)

Referenced Scenario: Distributed service/asset discovery	
KPI	Distributed Query/search time
Definition	Time required to perform a SPARQL query on multiple RDF metadata repositories.
Requirement	As the discovery/search capabilities can implemented by means of distributed SPARQL queries, the execution time for such queries must stay constant or at worst grow linearly as the metadata repository size grows.
Target Value	The response time on a distributed metadata repository comprises many factors that can affect it. Mainly, the response time will be affected by the volume of metadata information, the number of metadata repositories, and the number of results returned per query over a specific number of repositories. For a distributed metadata repository, the response time will be longer than a centralized one since it must first select which metadata repositories are of interest for the query and then execute a query for each of the selected repositories in order to subsequently integrate the results obtained by each query. Since the integration of distributed metadata repositories is a more complex process, it is expected that

	the response time will worsen by at least an order of magnitude when the volume of metadata information is large. For example, if the average time for a query is less than 10 seconds, then in the worst case, the average response time on a distributed query could be less than 20 seconds.

Table 14 Scalability requirements for Asset Manager

4.2 Extracted Key Performance Indicators and corresponding Scalability requirements for Converter and related utilities

4.2.1 Performance

Referenced Scenario: Batch Data Conversion	
KPI	Response Time
Definition	The time required to convert the whole data set.
Requirement	Based on the scenario description and given that the process applies during the adoption phase (i.e., Build-time), the target execution time is estimated in terms of few hours to few days, but less than one week per batch conversion. However, it is important to take into account that the *size of data*, and the density *of the transportation graph* are influencing variables to this metric.
Target Value	In the Shift2Rail IP4 domain, conversion of GTFS data for Journey Planning is an example of the Batch Data Conversion scenario. Static GTFS data in the railway domain is generally produced twice per year. Datasets for a local/regional operator (like Trenord in Lombardy region[7] or AMT Genova[8]) may range from below 10MB up to 100MB. For such datasets we expect a worst case scenario of one hour. In is important to consider that conversion time will be largely dependent on the target data model. Our worst case scenario takes into account a complex target data model (like IT2Rail ontology or Transmodel). For simple target ontologies, like Linked GTFS, we expect a lower value for the worst case scenario, like 45 minutes to perform a conversion.
Referenced Scenario: Distributed service/asset discovery	

[7] Trenord GTFS feed data: https://www.dati.lombardia.it/Mobilit-e-trasporti/Orario-Ferroviario-Regionale-Gtfs/3z4k-mxz9

[8] AMT Genova GTFS feed data: https://www.amt.genova.it/amt/GTFS/GTFS_AMT_GENOVA.zip

KPI	Response Time
Definition	The time required to convert one message.
Requirement	Based on the scenario description and given that the process applies during the engaging phase (i.e., run-time) and it is only a part of the transaction process, the target execution time must be such as to avoid degrading the overall transaction time. Hence, its order of magnitude must be in terms of few seconds. However, it is important to take into account that the *type of message*, the size of the message, and the *network latency* are influencing variables to this metric.
Target Value	Since the SPRINT project is not developing new mappings, our initial use case KPI for the Runtime Message Conversion will be based on the outcomes of the ST4RT project. In ST4RT D5.4[9], the total time for converting a TAP/TSI 918-XML PreBooking message to the corresponding message in FSM was 8 seconds. Our aim is to improve such conversion time, so we consider that value as our reference/worst case value.

Referenced Scenario: Automated Mapping Process for the Conversion

KPI	Execution Time
Definition	The time required to learn equivalent concepts of a single term of the source ontology in the target ontology.
Requirement	Given that average number of terms in an ontology (e.g., Transmodel, NeTEX, S2R ontology) is usually in the order of hundreds of terms. the required time to learn one single term should be in the order few seconds.
Target Value	Average time to learn mapping of two data sets with small size, i.e., less than 100 terms must be < 10 minutes

KPI	Usability
Definition	The number of steps for the end user to view and confirm a single mapping.
Requirement	The main factor that makes the automated learning process really useable for the human user is the level of comfort and clarity of the procedure for approving the single

9 http://www.st4rt.eu/D5.4

	mapped concept. We estimate such process must be accomplish within less than 10 steps.
Target Value	<10 steps

Table 15 Performance Requirements of the Converter and relevant utilities

4.2.2 Scalability of IF in respect to interact/deploy Converter/automated Mapping tool

Scalability requirements for the Converter component are largely depending on the deployment type. That is because the entity that is hosting the component and providing the computation resource to run it differs based on the type of deployment as explained as follows.

Direct download of deployable Converter artefact (jar, docker)

In the case of Direct Download, the service consumer downloads a deployable converter artefact (JAR, Docker image) to utilize it locally. Hence the responsible entity to ensure the scalability of the Converter is the Service Consumer.

This deployment approach is most suitable for operations applied at the Adaptation Phase such as Batch Data Conversion and Automated Mapping Scenarios. Given that such processes are accomplished off-line and not in very frequent cycles, the response time requirement is more relaxed. Hence, a **single instance of Converter** with stated performance requirements is enough **for each service consumer**. Clearly, in the case of higher demand, the consumer could **horizontally scale up** its system by **running multiple independent instances** of Converter, which in turn is an external activity with respect to the IF-Node.

Referenced Scenario: Batch Data Conversion /Automated Mapping	
KPI	Scalable Converter Discovery Process
Definition	Following the direct download deployment, the conversion demand decreases to the demand of a single consumer, while the discovery requests to the IF to locate the right converter are coming potentially from all clients similar to any other assets.
Requirement	The requirement is to scale up the system as the number of discovery requests increases in such a way that the system sustains its regular functionality and does not suffer a slowdown in its overall performance.
Target Value	Since the Converter Discovery Process is heavily based on accessing RDF metadata, we expect that the scalability will be mostly influenced by the maximum number of concurrent SPARQL queries which will be allowed by the RDF repository. The KPI will therefore be the same as the defined in Section 4.1.2.

KPI	Scalable Converter download request
Definition	Similar to any other asset, download requests to the IF for downloading the discovered converter are coming potentially from all clients.
Requirement	The requirement is to scale up the system as the number of download request increases in such a way that the system sustains its regular functionality and does not suffer a slowdown in its overall performance.
Target Value	Downloading a Converter is an operation which will just require accessing a URL. If the Converter artifact is stored remotely by the owner, no KPI can be defined. If the Converter artifact will be stored on the Asset Manager, we can expect the scalability to be influenced primarily by the available bandwidth. We expect to be able to sustain up to 512 requests per second without any performance penalty, and to scale linearly from that point on. Anyway, such a workload will be highly unlikely, given the fact that downloading a Converter will not be a frequent operation (it will be required when defining and setting up the interoperability solution)

Table 16 Scalability requirements for Direct Download of a Converter

In summary, in case of direct download the corresponding scalability requirements for deploying and managing a converter are beyond and irrelevant to the IF. In other words, the IF is responsible up to the discovery and download of a single converter, hence the relevant scalability KPIs are about these interactions, as explained in Table 16.

Runtime environment deployment

In theory, consumers might utilize Converters as service if the IF Node is coupled with an IF Runtime environment, for Adaptation Phase and Engaging Phase operations, if they do not want to invest their local computation resources for the conversion process.

Referenced Scenario: Batch Data Conversion / Automated Mapping / Runtime Message Conversion / Fast adaptation to peaks	
KPI	Concurrent Converter Deployment
Definition	The IF should scale up the computation resources for running multiple instances of converters per request of clients, to preserve the declared performance requirement.
Requirement	In this sense, the worst-case scenario is when all potential operators of a single IF instance ask for at least one converter to run on the runtime environment.

	Accordingly, the scalability requirement for the IF is to support the availability of the required computational resources in such a way that all of them meet the most restrictive performance requirement, which is the one declared for the runtime message conversion – i.e., few seconds. In other words, an IF-Node should scale in such a way that the more resource-consuming operation – i.e., Data Conversion (with the performance requirement of few hours to few days) – does not degrade the more demanding operation – i.e., message conversion – at runtime.
Target Value	The IF will not implement a new cloud-based system and will support relying on existing cloud orchestration systems. We expect that the total time to obtain a new deployment of a Converter will be mostly influenced by the size of the Converter artifact, since such artifact will need to be uploaded to the specific cloud node. Given that a Converter will be developed as a microservice with minimal dependencies, we expect that the total time to fulfill a deployment request will be 10 minutes at worst.

Table 17 Scalability requirements for Runtime environment deployment of Converter

However, our suggestion is to follow the direct download approach for off-line operations and the runtime environment approach for run-time operations such as Runtime Message Conversion.

In any case, in the "runtime environment deployment" approach, one instance of the IF Runtime environment is the endpoint for multiple operators willing to convert various types of data/messages. Hence, the relevant scalability requirement is related to the runtime management of multiple instances of Converters, as stated in Table 17.

Direct call to Service endpoint on Service Provider premises

Finally, a service provider might advertise the description of a Converter in the IF and direct the interested consumer to a service endpoint running on the provider premises. Consequently, in such cases, managing of the scalability is up to the service provider and it is an external issue to the IF. Hence, the relevant scalability factors are similar to "*Direct Download Deployment"*, as stated in Table 16.

4.3 Extracted Key Performance Indicators and corresponding Scalability requirements for Resolver

4.3.1 Performance

Referenced Scenario: Special Purpose discovery package	
KPI	Response Time for location Resolver

Definition	The time required to Resolve a Location name to its the geocoordinates data.
Requirement	Based on the scenario description and given that the process applies during the engaging phase (i.e., run-time) and it is only a part of the transaction process, the target execution time must be such as to avoid degrading the overall transaction time. Hence, maximum response time must be less than two second
Target Value	Two different interfaces, using the same ontology to generate NeTEX or IT2Rail data structures, where developed in the IT2Rail project, yielding response times of between 300 and 1600 milliseconds. The SPRINT project targets a maximum of 600 milliseconds for the response time
KPI	Response Time for Travel Expert Resolver
Definition	The time required to discover appropriate Travel Experts based on the requested travel route and location.
Requirement	Based on the scenario description and given that the process applies during the engaging phase (i.e., run-time) and it is only a part of the transaction process, the target execution time must be such as to avoid degrading the overall transaction time. Hence, maximum response time must be less than two second
Target Value	The IT2Rail project achieved response times of between 200 and 600 milliseconds, but with a limited number of entries in the Travel Expert registry. The SPRINT project targets a maximum of 1000 milliseconds response time over a registry of up to 100 entries

Table 18 Performance Requirements of the Resolver

4.3.2 Scalability

Similar to the Converters (See Converter Direct Download), the scalability requirement for the Resolver component depends on the deployment type, and the responsibility to scale the resolver to ensure certain performance must be taken by the entity (i.e., IF, Resolver consumer or Resolver provider) which runs the component.

Direct download of deployable Resolver artefact (jar, docker)

In the case of Direct Download, the service consumer downloads a deployable resolver artefact (JAR, Docker image) to utilize it locally. Hence, the responsible entity to ensure the scalability of the Resolver is the Service Consumer. Accordingly, the interaction of the client with the IF is limited to the discovery and download of the desired resolver, so as the relevant scalability issues as described in Table 19.

Referenced Scenario: Special Purpose discovery package	
KPI	Scalable Resolver Discovery Process
Definition	In direct download deployment, the Resolving demand decreases to the demand of a single consumer, while the discovery requests to the IF to locate the right resolver are coming potentially from all clients.
Requirement	The requirement is to scale up the system as the number of discovery request increases in such a way that the system sustains its regular functionality and does not suffer a slowdown in its overall performance.
Target Value	Since the Resolver Discovery Process is heavily based on accessing RDF metadata, we expect that the scalability will be mostly influenced by the maximum number of concurrent SPARQL queries which will be allowed by the RDF repository. The KPI will therefore be the same as the defined in Section 4.1.2.
KPI	Scalable Resolver download request
Definition	Similar to any other asset, download requests to the IF for downloading the discovered resolver are coming potentially from all clients.
Requirement	The requirement is to scale up the system as the number of download request increases in such a way that the system sustains its regular functionality and does not suffer a slowdown in its overall performance.
Target Value	Downloading a Resolver is an operation which will just require accessing a URL. If the Resolver artifact is stored remotely by the owner, no KPI can be defined. If the Resolver artifact will be stored on the Asset Manager, we can expect the scalability to be influenced primarily by the available bandwidth. We expect to be able to sustain up to 512 requests per second without any performance penalty, and to scale linearly from that point on. Anyway, such a workload will be highly unlikely, given the fact that downloading a Resolver will not be a frequent operation (it will be required when defining and setting up the interoperability solution)

Table 19 Scalability requirements for Direct Download of a Resolver

Runtime environment deployment

Similar to the Converter case, if a client is not interested in employing its local computation resource for the Resolver component, s/he can use a Resolver as a service hosted by the provided IF Runtime environment. So, one IF Runtime environment is the endpoint for multiple operators willing to interact with various types of Resolver. Hence, as stated in Table 20, the IF is responsible to scale up its resources to be able to run multiple instances of the Resolver to satisfy all the requests.

Referenced Scenario: Batch Data Conversion / Automated Mapping / Runtime Message Conversion / Fast adaptation to peaks	
KPI	Concurrent Resolver Deployment
Definition	IF should scale up the computation resources for running multiple instances of the resolver given the requests of clients, to preserve the declared performance requirement.
Requirement	In this sense, the worst-case scenario is when all potential operators of a single IF instance ask for at least one resolver to run on the runtime environment. Accordingly, the scalability requirement for the IF is to support the availability of the required computational resources in such a way that all of them meet the stated performance requirement.
Target Value	The IF will not implement a new cloud-based system and will support relying on existing cloud orchestration systems. We expect that the total time to obtain a new replica of a Resolver will be mostly influenced by the size of the Resolver artifact, since such artifact will need to be uploaded to the specific cloud node. A Resolver is different from a Converter because its features can vary from time to time. Also, its architecture can require the deployment of multiple components. As such, the deployment time of a new Resolver will largely depend on its complexity. If a Resolver will be developed as a microservice with minimal dependencies, we expect that the total time to fulfill a new deployment request will be 10 minutes at worst. The number of parallel deployments will be solely related to the performances and SLA offered by the cloud provider.

Table 20 Scalability requirements for Runtime environment deployment of Resolver

Direct call to Service endpoint on Service Provider premises

As for Converters and any other interoperability services as well as auxiliary services, a service provider might prefer to utilize the IF only for the advertisement of its Resolver. In such cases, the service endpoint would be returned to the client upon his/her discovery request, and the client should then call and interact with the Resolver running on the resources of service provider. Consequently, the managing of scalability issues is up to service provider and it is an external issue to the IF, and the relevant scalability factor for “Direct call to Resolver service” is similar to the one for “Direct Download Deployment”, which is stated in Table 19.

5. OTHER REQUIREMENTS

In addition to performance and scalability requirements discussed in Section 4, we have identified some key non-functional requirements for the IF, as explained in the following.

5.1 SECURITY AND PRIVACY

Security and privacy are among the most imperative aspects of any software system that must be contemplated. In general, security may refer to a vast range of concerns depending on the type and domain of the software. The most widely used terminologies for categorizing the security hazards are the so-called CIA and CIA+, where the former is an abbreviation for Confidentiality, Integrity, and Availability [6], and the latter extends CIA with Authentication (of people, organizations and applications), Access control and Non-repudiation [7]. Such security threats, then, must be identified in various technological aspects such as hardware, databases, networking, etc., and per-component of a system.

Among the various IF components, security and privacy concerns arise mainly in the communication with the Asset Manager, which is the root endpoint for involved parties to interact with the IF by registered users, uploading and discovering various types of data and assets (see Table 3 and Table 4 for more detailed user cases). Furthermore, as discussed in the advanced user-scenario in Section 3.2 (Table 10), authorization and authentication are other relevant security aspects in our case. Finally, it is worth to highlight the fact that the main focus of the IF is to foster semantic interoperability by providing essential tools and building blocks to implement an interoperability solution, rather than a shared centralized data store. Accordingly, its main principle is that data should reside as much as possible in the originating systems while Asset Manager deals mainly with metadata. In this direction, Table 21 shows the identified security requirements for the IF and their priority in the sense of implementation order. *High* priority implies that the requirement must be satisfied in the base version of IF, while *medium* and *low* priority mean such requirements might be addressed in subsequent enhancements of the IF.

Priority	Security and Privacy Requirement
High	IF shall provide access to data according to data protection law and regulations.
Medium	IF shall prevent any unauthorized access to any data stored in AM.
Low	IF shall provide a controlled sharing of data among various transport and mobility operators and applications, enabling them to tune the granularity of access rights to their data from fully open to customized (fine-grained) authorization.
Low	IF shall be compatible with different policy descriptions to be adopted depending on the different business needs of the actors in the transportation domain.

Medium	The IF may provide an interoperable, global and distributed authentication process designed for coping with federated data storage and management.
High	IF governance certifies the safety of assets it generates but ensuring the safety and liability of data provided by third parties is outside the responsibilities of the IF. Similarly, ensuring the security of a mobility transaction is outside the responsibilities out of the IF.

Table 21 Security and Privacy Requirements for IF

5.2 Productivity

The productivity requirement is particularly defined for the Mapping utility of a Converter and it refers to the accumulated time required to produce a complete mapping between the whole terms of the source and target data model.

Productivity time is a combination of Execution Time for learning all terms, plus the time required to finish the approval process and re-learning of the rejected terms. Given that such an approach is low-frequent and must be done off-line, the estimated productivity requirement is less than one week.

Priority	Productivity Requirement
High	The Mapping utility shall provide a complete mapping for an average-size standard in the order of several hours to several days.

Table 22 Productivity Requirement for Mapping Utility

5.3 Precision

Similar to productivity, precision is also defined for the Mapping utility of Converter component. It is stated as the ratio between the suggested mapping and confirmed mapping by a human user, which affects the overall performance of the system. In this context, the precision of more than 50% is acceptable.

Priority	Precision Requirement
High	The suggested equivalent terms in source and target standardization shall be equal to or greater than 50%.

Table 23 Precision Requirement for Mapping Utility

5.4 User centricity

User-friendliness and centricity is another key requirement for any software framework. It refers to a suite of characteristics such as intuitive design and simplicity of functions, unambiguous user interface and navigation through the system, sufficient supports and effective error handling process, etc., which facilitates the interaction of end-users with the system.

Priority	User Centricity Requirement
Low	Asset Manager may provide the possibility of acquiring feedbacks (e.g., in the form of ratings of the assets) so that it could remove faulty assets and promote the most popular ones, which in turn, enhances the productivity of the IF, as well the level of trust that users have in the system.
Medium	The Asset Manager may guide the user through the process of assembling the various pieces of a Converter and configuring it to integrate with the user's IT environment. This process includes finding and evaluating the completeness of lifting and lowering transformations to implement the end-to-end conversion, and the choice of how to expose such feature both in terms of choosing the technological stack (like SOAP, REST, message queues, …) and of deploying the final artifact.

Table 24 User-friendliness Requirement for Mapping Utility

6. REFERENCES

[1] M. Glinz, "On non-functional requirements," in *n 15th IEEE International Requirements Engineering Conference (RE 2007)*, 2007.

[2] K. Pohl, Requirements engineering: fundamentals, principles, and techniques, Springer Publishing Company, 2010.

[3] M. Glinz, "On non-functional requirements," in *In 15th IEEE International Requirements Engineering Conference*, 2007, October.

[4] I. Jacobson, The unified software development process, Pearson Education India, 1999.

[5] H. H. Liu, Software performance and scalability: a quantitative approach, John Wiley & Sons, 2011.

[6] C. P. Pfleeger and P. S. Lawrence, Security in Computing, Prentice Hall Professional Technical Reference, 2002.

[7] A. Simmonds, P. Sandilands and L. Van Ekert, "An ontology for network security attack," in *In Asian Applied Computing Conference*, Berlin, Springer.