HARMONY

Holistic Approach for Providing Spatial & Transport Planning Tools and Evidence to Metropolitan and Regional Authorities to Lead a Sustainable Transition to a New Mobility Era

D3.3 Transport and Spatial Data Warehouse 1st Prototype

@Harmony_H2020

#harmony-h2020

https://www.linkedin.com/company/harmony-h2020/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

Part of: 2020 CIVITAS

THE CIVITAS INITIATIVE IS CO-FINANCED BY THE EUROPEAN UNION

SUMMARY SHEET

PROJECT

Project Acronym:	HARMONY
Project Full Title:	Holistic Approach for Providing Spatial & Transport Planning Tools and Evidence to Metropolitan and Regional Authorities to Lead a Sustainable Transition to a New Mobility Era
Grant Agreement No.	815269 (H2020 – LC-MG-1-2-2018)
Project Coordinator:	University College London (UCL)
Website	www.harmony-h2020.eu
Starting date	June 2019
Duration	42 months

DELIVERABLE

Deliverable No. - Title	D3.3 - Transport and Spatial Data Warehouse 1st Prototype
Dissemination level:	Public
Deliverable type:	Report
Work Package No. & Title:	WP3 - Data collection tools, data fusion and warehousing
Deliverable Leader:	ICCS
Responsible Author(s):	Efthimios Bothos, Babis Magoutas, Nikos Papageorgioy, Grigoris Mentzas (ICCS)
Responsible Co-Author(s):	Panagiotis Georgakis (UoW), Ilias Gerostathopoulos, Shakur Al Islam, Athina Tsirimpa (MOBY X SOFTWARE)
Peer Review:	Panagiotis Georgakis (UoW), Ilias Gerostathopoulos (MOBY X SOFTWARE)
Quality Assurance Committee Review:	Maria Kamargianni, Lampros Yfantis, Emmanouil (Manos) Chaniotakis (UCL)

DOCUMENT HISTORY

Version	Date	Released by	Nature of Change
0.1	24/07/2020	ICCS	First draft, ToC
0.4	15/09/2020	ICCS	Description of Data Management Framework
0.7	26/10/2020	ICCS	Description of API and interfaces
0.9	20/11/2020	ICCS	Version ready for internal review
1.0	30/11/2020	ICCS	Final Version

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Explanation
API	Application Programming Interface
CRUD	Create - Read - Update - Delete
HTTP	Hypertext Transfer Protocol
JSON	JavaScript Object Notation
LUTI	Land Use Transport Interaction
MS	Model Suite
NoSQL	non Structured Query Language or non relational
ODBC	Open Database Connectivity
REST	REpresentational State Transfer
SQL	Structured Query Language
TSDW	Transport and Spatial Data Warehouse
URI	Uniform Resource Identifier
URL	Uniform Resource Locator

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269
Part of: 2020 CIVITAS

EXECUTIVE SUMMARY

This deliverable, the third of WP3, describes the 1st prototype of HARMONY's Transport and Spatial Data Warehouse (TSDW) and is the outcome of the work performed in tasks T3.3 "Data management mechanisms", T3.4 "Data security & Privacy" and T3.5 "TSDW Integration, and population".

The work performed in these tasks has resulted to the specification and implementation of the HARMONY Data Management Framework. The need for such a framework stems from the fact in HARMONY, the integration of simulators is performed at different levels (operational, tactical and strategic), which needs to efficiently communicate, exchange information and data in an automated manner, while the HARMONY platform has to be able to store and manage the required input and output data. Moreover, in HARMONY different users and user groups can have access to the functionalities as well as the stored information. Users access rights are required in order to ensure that sensitive information can be viewed only by particular users (including transport modelers of a city or an area), while parts of the information can be accessible by a wider set of stakeholders. Additionally, part of the stored data refers to analytics and insights that the platform generates based on the output of the simulations. Such insights support decisions of transport stakeholders with respect to the integration of new mobility services.

The HARMONY data management framework allows managing the end to end lifecycle of transport simulation scenarios, including their definition, configuration, execution, tracking, comparison etc. Transport modelers use the HARMONY framework to track transport simulations, organize their work for future reuse, and structure output data for insights and decision support. Teams of modelers use the framework to log and compare results across multiple users working on the same problem. They can try different simulators to tackle the same problem and then run the same simulations again on new data to compare results in the future. The data management framework supports and drives the development of the platform and provides the basis for making sense of transport simulations in the integrated HARMONY platform. Note that the HARMONY TSDW is the instantiation of the proposed framework and consists of data stored in database systems as well as artifacts (files) stored in the filesystem of the server where the HARMONY platform is deployed.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269
Part of: 2020 CIVITAS
THE CIVITAS INITIATIVE IS CO-FINANCED BY THE EUROPEAN UNION

1 Introduction

1.1 Project Summary

HARMONY envisions to assist metropolitan areas with evidence-based decision making, by providing a state-of-the-art model suite that quantifies the multidimensional impact of various policies, investments and mobility concept applications, while simultaneously identifying the most appropriate solutions and recommending ways to exploit disruptive mobility innovations. HARMONY proposes an integrated approach through the development of the HARMONY Model Suite (HARMONY MS) which integrates new and existing sub-models. This integrated approach is necessary to understand if, how and to what extent new policies, investments and mobility concepts can produce results that are in line with the objectives set by authorities.

The HARMONY MS enables end-users such as planners, decision makers, researchers and transport operators/providers to couple/link independent models and analyse a portfolio of regional and urban interventions for both passenger and freight mobility. The main objective behind the model system's architecture is to enable the evaluation of such interventions with regards to their impact on land-use, economic growth, transportation networks, energy, vehicular noise and emissions, while, at the same time, provide recommendations for Sustainable Urban Mobility Plans (SUMPs) of the new mobility era.

HARMONY follows a multi-scale approach (described in detail in D1.3) consisting of the **Strategic Level** (Long-term), the **Tactical Level** (Mid-Term) and the **Operational Level** (Short-term) models.

- **Strategic Level**: It is mainly composed of regional economic, demographic forecasting, land-use, spatial freight interaction and long-term mobility choice models. It operates on a long-term horizon (year-to-year) and is mainly responsible for generating i) disaggregate household and firm population and the locations for different types of activities such as employment, housing, and education, ii) aggregate commodity flows between employment sectors and iii) long-term mobility choices of individuals (agents) including car-ownership or subscriptions to different mobility services.
- **Tactical Level**: it is a fully agent-based passenger and freight demand model and it consists of two sub-models one for passenger and one for freight which model agents' choices on a day-to-day level.
- **Operational Level**: it represents the transport supply and demand interactions at high granularity (e.g. second to second, minute to minute). It can be characterised as a multimodal network assignment model system that is responsible for loading the demand into different types of networks, while simultaneously capturing travellers' route choices and dynamic schedule re-evaluation choices due to supply conditions. It also includes dedicated modules that emulate disruptive new mobility service operations and their interactions with agents (e.g. traveller, vehicles) of the system. This simulator is mainly responsible for generating traffic volumes and impedance measures of the form of skim matrices (e.g. travel time, cost distance) per model and spatial unit of analysis.

1.2 Deliverable Objectives

This deliverable, the third of WP3, describes the 1st prototype of HARMONY's Transport and Spatial Data Warehouse (TSDW) and is the outcome of the work performed in tasks T3.3 "Data management mechanisms", T3.4 "Data security & Privacy" and T3.5 "TSDW Integration, and population".

More specifically, T3.3 deals with the development of methodologies to integrate multidisciplinary data sources used by the HARMONY simulators and provides the necessary input/output and management mechanisms to automatically prepare the data in the formats required. The import interfaces will include

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269
Part of: 2020 CIVITAS
THE CIVITAS INITIATIVE IS CO-FINANCED BY THE EUROPEAN UNION

(i) the mechanisms for accepting, storing, and processing static and near real-time data streams, (ii) employment of mechanisms to convert the aforementioned data, and (iii) implementation of semantic data integration modules for storing external data into the central TSDW reference data schema and ensuring data harmonization. T3.4 provides privacy enhancing principles and technologies for sensitive data. In order to prevent data misuse, this task introduces measures for secure data transfer and proper access control. T3.5 puts together the methods, reference data models, and data management tools developed in T3.3 and T3.4 and focuses on their implementation and deployment in the TSDW.

The aforementioned work has resulted in the specification and implementation of the HARMONY Data Management Framework which provides a holistic approach for managing the end to end lifecycle of transport simulation scenarios, including their definition, configuration, execution, tracking, comparison etc. Note that the HARMONY Data Warehouse is the instantiated Data Management Framework and comprises of data and metadata stored in the HARMONY database as well as artifacts (files) stored in the HARMONY server.

The objective of this deliverable is to describe the data entities of the HARMONY Data Management Framework and provide the technical specification of the data structures and data access API that consist the 1st prototype of HARMONY's TSDW.

1.3 Deliverable Structure

The deliverable is structured as follows. Section 2 provides an overview of the HARMONY Data Management Framework while Sections 3-9 focus on describing the data entities that comprise the framework:

- Section 3 focuses on the description of user groups, access rights and roles.
- Section 4 describes the transport simulators registry, a data entity which is used to keep track of the different simulators that are part of the HARMONY MS.
- Section 5 explains how simulation scenarios are specified in the data warehouse.
- Section 6 provides the metadata used to describe the input data required by the harmony simulators.
- Section 7 lists the metadata used to describe the log data generated by the harmony simulators.
- Section 8 describes how the output data of the simulators are stored in the data warehouse.
- Section 9 shows how key performance indicators inferred from the output data of the simulators are stored.

The deliverable concludes in Section 10 with the TSDW implementation status and plan for next steps.

2 Data Management Framework of the TSDW

The need for the HARMONY Data Management Framework stems from the fact that in HARMONY, the integration of simulators is performed at different levels (operational, tactical and strategic), which have to efficiently communicate, exchange information and data in an automated manner, while the HARMONY platform has to be able to store and manage the required input and output data. Moreover, in HARMONY different users and user groups can have access to the functionalities, as well as the stored information. User access rights are required, in order to ensure that sensitive information can be viewed only by particular users (e.g. transport modelers of a city or an area), while parts of the information can be accessible by a wider set of stakeholders (e.g. transport modelers of a different city). Additionally, part of the stored data refers to analytics and insights that the platform generates based on the output of the simulations. Such insights support decisions of transport stakeholders with respect to the integration of new mobility services.

In this landscape, a number of actions require specific designs in terms of data management mechanisms. Such actions include:

- **Versioning** (i.e. keep track) of transport simulation models, results, users, inputs, configurations, and generated outputs. Transport simulation models are expected to change frequently when transport modelers examine different mobility services and integration scenarios. Furthermore, per simulation there is a wealth of data generated that need to be stored in a structured manner for further processing and analysis.

- **Searching** for simulation scenarios and related outputs. When the output data of simulations are available and as the platform is populated with different scenarios, transport modelers need to be able to search and find relevant scenarios that need to be analyzed to support their decisions
- **Reproducing** results and re-running scenarios. Reproducibility is also an important aspect that needs to be considered as once a scenario has been executed, it should be reproducible by other users so that the results can be cross-checked and verified. When a scenario contains stochastic elements, performing several runs of it can also be used for obtaining better understanding of the effect of such elements on the key metrics of the scenario.
- **Aggregating** the results of multiple simulation runs. As simulation tools are underpinned by probabilistic models, aggregation of results in the form of averages will result in more reliable outputs. Furthermore, aggregated data can be used for sensitivity analysis of various models/algorithms where necessary.
- **Understanding** the outputs and **comparing** different simulation scenarios. The analysis of the simulation outputs needs to be properly managed and users of the platform need to be able to understand the outputs in an intuitive manner as well as be able to compare the outputs of different scenarios.

Considering the above points, HARMONY has developed and implemented a holistic framework that allows managing the end to end lifecycle of transport simulation scenarios, including their definition, configuration, execution, tracking, comparison etc. Transport modelers would use the HARMONY framework to track transport simulations, organize their work for future reuse, and structure output data for insights and decision support. Teams of modelers would use the framework to log and compare results across multiple users working on the same problem. They could try different simulators to tackle the same problem and then run the same simulations again on new data to compare results in the future. The data management framework supports and drives the development of the platform and provides the basis for making sense of transport simulations in the integrated HARMONY platform.

The HARMONY data management framework and its position within the HARMONY platform is shown in Figure 1. The main components of the platform are the following:

- The Scenario Configuration and Simulators Orchestration GUI that provides access to the platform's functionalities. This is a web application (WP2) that encloses all necessary screens and views which practically allows users to configure simulation scenarios, run them and analyse the generated results.
- The Pool of Simulators is a set of executable programs or scripts that implement the strategic, tactical and operational modelling levels of the HARMONY MS.
- Orchestration Engine. This is the heart of the HARMONY MS as it executes the workflow that stands for the technical realization of a scenario in the HARMONY MS and includes the order with which the executables are run and the way they communicate with each other.
- Data Management. This is "the memory" of the HARMONY MS as it stores all data required to run the simulations and the knowledge that is generated by them.

The HARMONY data management framework has been designed such that it captures all needed information and adheres to the needs of the transport modellers developing the simulators as well as the transport stakeholders who need to access, analyse and understand the simulation outputs.

The 1st prototype of the HARMONY TSDW implements the proposed Data Management Framework and comprises of all data and metadata stored in the HARMONY database as well as artifacts (files) stored in the HARMONY server.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

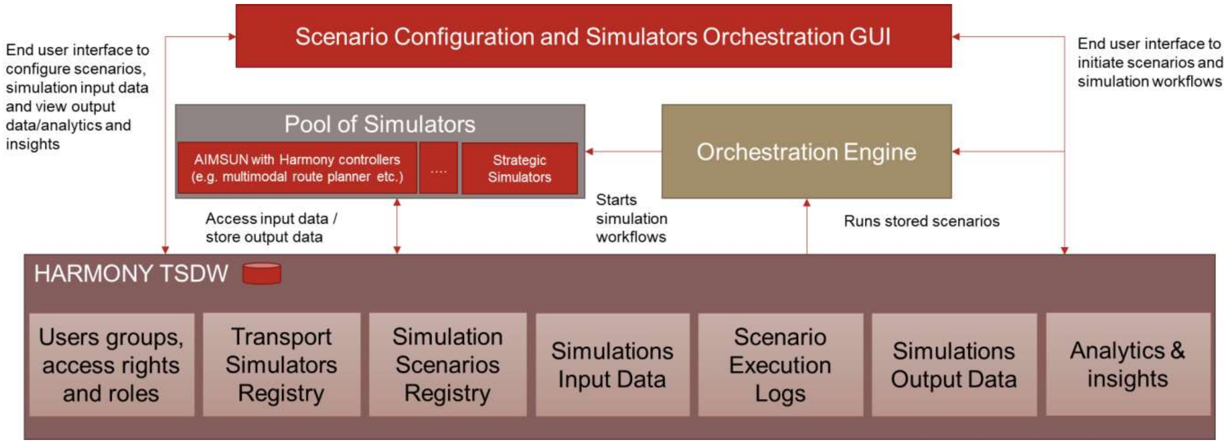

Figure 1: The HARMONY data management framework as part of the TSDW and its position within the HARMONY platform.

It consists of seven main data entities as follows:

- **User groups, access rights and roles** that encompasses data related to access control and security.
- **Transport Simulators registry** that maintains data and metadata related to the available transport simulators and their versions.
- **Simulation Scenarios Registry** that stores data related to configured scenario templates and related scenarios that users execute. A scenario template provides a simulation workflow. A template can be instantiated in different scenarios with a set of input parameters for a specific modelling need. For example, dynamic freight scheduling scenario template can be used to evaluate micro freight logistics with low emission zones.
- **Simulations Input Data** that refers to input information required by the simulation to run. For example, the network models and other external data.
- **Scenario Execution Logs** that hold logs which are produced when the simulation run. These can be useful especially for transport modellers implementing the simulators.
- **Simulations Output Data** that include raw output information coming out of the simulations.
- **Analytics and Insights** that infers KPIs and stores them in a format that can be readily used by the HARMONY front-end applications so that end users can analyse the results and take transport planning decisions.

3 User groups, access rights and roles

The HARMONY platform is accessed by different users, with varying interests regarding the functionalities and information stored in the platform. The platform needs to be aware of the access level of users and allow access to the functionalities and information relevant for the particular user. For this purpose, the ‘user groups, access rights and roles’ data entity stores related information.

A Role-based access control (RBAC) approach is followed, which establishes permissions based on groups (defined sets of users) and roles (defined sets of actions). Individuals can perform any action that is assigned to their role and may be assigned multiple roles as necessary. Users are not permitted to change the level of access control that has been assigned to their role.

The RBAC model allows HARMONY administrators to create roles for various job functions and assign permissions to perform certain operations to specific roles. Members or staff (or other system users) are assigned particular roles, and through those role assignments acquire the permissions needed to perform particular system functions. With RBAC users are not assigned permissions directly, but acquire them through their role (or roles), which means that the management of individual user rights becomes a matter of simply assigning appropriate roles to the user’s account; this simplifies common operations, such as adding a user, or changing a user’s department.

Four main roles are foreseen for HARMONY at this point as follows:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269
Part of: 2020 CIVITAS

- Admin users: This role has access to all functionalities and data of the platform. The admin user can configure user accounts, manage the platform and have global access to the data which are available in the platform.
- Transport planners/modellers: This role has access to all functionalities and data for a specific area i.e. read and write access (change the network, introduce new service types, etc.) and can configure and run simulation scenarios.
- Policy/Decision makers: Policy makers and observers are mainly interested in the results of transport simulation and access related output data and generated analytics and insights. These users cannot configure or run simulations. Their main aim is to check and understand scenarios and respective results for decision making within the area/city.
- External stakeholders (e.g. policy maker / observers from another city/area): These users are interested stakeholders from other areas/cities and in general users who want access to analytics/insights and check/understand scenarios and results for decision making in other areas. These users need access to aggregated results and not detailed simulation outputs.

In terms of implementation, the HARMONY deployment can be multitenant, which means that a single server hosts users and data coming from different areas or single tenant which means that the platform is deployed in the premises of an organization (e.g. a municipality) and is used by the transport modelers of that specific organization. In the case of multi-tenant deployments where more than one organizations make use of the same server, different user groups can be configured having access to the specific data of their organization.

For example, if we consider the Oxfordshire City Council (OCC), a pilot partner of HARMONY, the HARMONY administrator configures a group of Transport modelers and another of Policy makers. When a new user account is configured in this instance of the HARMONY platform, the user is allocated to one of the preconfigured groups depending on her/his role. The respective definitions for this example, in a JSON format used to store them in the HARMONY TSDW can be seen bellow.

```
//HARMONY platform roles / user groups for a typical deployment
{
  "id": "d290f1ee-6c54-4b01-90e6-d701748f0851",
  "name": "Admin",
  "type": "admin",
  "organization": "HARMONY",
  "description": "This is a group for the admin users",
  "access_rights": ["read_all", "write_all", "execute_all"],
}
{
  "id": "d290f1ee-6c54-4b01-90e6-d701748f0852",
  "name": "Modeller_OCC",
  "type": "modeler",
  "organization": "OCC",
  "description": "This is a group for the transport modelers of OCC",
  "access_rights": ["read", "write", "execute"],//users of this group can view
/write data created within the OCC organization and run simulation scenarios
}
{
  "id": "d290f1ee-6c54-4b01-90e6-d701748f0852",
  "name": "Policy_maker_OCC",
  "type": "policy_maker",
  "organization": "OCC",
  "description": "This is a group for the policy / decision makers modelers of
OCC",
  "access_rights": ["read"],//users of this group can view data created within the
OCC organization
}
```

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269
Part of: 2020 CIVITAS

```
{
  "id": "d290f1ee-6c54-4b01-90e6-d701748f0853",
  "name": "External users",
  "type": "external",
  "description": "This is a group for users who can view public HARMONY data",
  "access_rights": ["read_public"],//users of this group can view public data only
}
//HARMONY platform example user for a typical deployment
{
  "id": "d290f1ee-6c54-4b01-90e6-d701748f0854",
  "name": "Patrick Mueller",
  "username": "pmueller@gmail.com",
  "password_hash": "234234sfdefwf23fe",
  "password_salt": "234232efef",
  "date": "2020-10-18T16:00:00Z",
  "group": "Modeler_OCC",
}
```

4 Transport Simulators Registry

The HARMONY platform is composed of a number of simulators and the Transport Simulators Registry holds information which allows the platform to run the simulators in related workflows and also keep track of different simulator versions.

As defined in HARMONY's technical specification (Deliverable D2.1) the following simulators (also referred to as components or models of the HARMONY platform) are being developed:

1. Strategic Level:
   a. S1. Demographic forecasting module: Examining the most current demographic data on case study areas and making short term projections (1 year) of future demographic profile
   b. S2. Regional economy module: This component aims to generate future employment (including services, health and educational activities) which structure the demand for physical travel.
   c. S3. Land Use Transport Interaction (LUTI) model.
   d. S4. Vehicle ownership module.
   e. S5. Spatial Interaction Freight Model.
   f. S6. Aggregate to disaggregate population and employment translators (Synthetic population model).
2. Tactical level:
   a. T1. Tactical freight simulator.
   b. T2. Tactical passenger simulator: This component will embody the tactical level simulator for passengers. It is essentially an activity-based simulation model for the designated areas.
   c. T3. Day-to-day learning module.
3. Operational level:
   a. O1. Meso/Micro Simulator.
   b. O2. Freight service controller: Controls freight distribution at operational level.
   c. O3. Multimodal Passenger service controller system: Embodied into the operational level simulator it is a manager of a mobility service, including its staff or vehicles and traveller trip requests.
   d. O4. Within-day re-evaluation controller.
   e. O5. Energy and emission calculator.
   f. O6. Noise calculator.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

For each simulator, several metadata are configured in the HARMONY warehouse and are regularly updated to reflect the simulators' current status. These include:

- Simulator_ID: System identifier of the simulator which refers to an object id generated when a new simulator entity is created in the database.
- Name: A descriptive name of the simulator.
- Description: Short description of the simulator highlighting its main purpose.
- Type: Can be one of the following: strategic/tactical/operational.
- Current Version Number: As simulators are regularly updated this filed holds its version and is used to identify with which version a scenario has been configured with.
- Date: Datetime object showing when the simulator was last updated.
- Binary file location / script: The location of the script or executable of the simulator in the HARMONY server. This is used by the platform in order to run the simulator within a workflow.
- Arguments required to run the simulator: Array of arguments that are used to run the simulator, including input data.
- Output data: List of output data that are generated by the simulator.
- Version History: Array of past versions of the simulator.

As an example, the entry for the operational freight service controller (O2) is the following:

- Simulator_ID: ObjectId("507f191e810c19729de860ea")
- Name: Operational Freight Service Controller
- Description: Controls freight distribution at operational level.
- Type: operational
- Current Version Number: v0.91
- Date: 2020-10-18T16:00:00Z
- Binary file location / script: /harmony/simulators/o2_freight_controller/v0.9/o2_fc.py
- Arguments required to run the simulator: {"heavy_freight_tt": "refers to a file containing trips and tours for heavy freight", "parcels_delivery_tt": "refers to a file containing trips and tours for parcels delivery", "activity_schedules": "refers to a file containing activity schedules"}
- Output data: {"travel_times_for_tours": "travel time data for tours", "delivery_delays": "delays in delivering goods", "fleet_utilization": "fleet utilization data"}
- Version History: ["v0.7", "v0.8"]

And the corresponding json format is as follows:


```
{
  "id": "d290f1ee-6c54-4b01-90e6-d701748f0851",
  "name": "Operational Freight Simulator",
  "description": "Operational Freight Simulator",
  "date": "2020-10-18T16:00:00Z",
  "arguments": [ {"heavy_freight_tt": "refers to a file containing trips and tours
for heavy freight", "parcels_delivery_tt": "refers to a file containing trips and
tours for parcels delivery", "activity_schedules": "refers to a file containing
activity schedules"}],
  "type": "operational",
  "currentVersion": "v0.91",
  "versionHistory": ["v0.7", "v0.8"],
  "binaryFileLocation": "/simulatorsRegistry/tactical/tacticalfreight_v0.91.py"
  "outputData": [ {
    "dataType": "travel_times_for_tours",
    "format": "csv",
    "location": "",
    "descrpition": "travel time data for tours"
  }, {
    "dataType": "delivery_delays",
    "format": "csv",
```

```
    "location": "",
    "descrpition": "delays in delivering goods"
  }, {
    "dataType": "fleet_utilization",
    "format": "csv",
    "location": "",
    "descrpition": "fleet utilization data"
  } ],
}
```


5 Simulation Scenarios Templates and Scenarios Registry

In HARMONY, simulations are configured through simulation scenario templates. A scenario template addresses a specific modelling need and is composed of a processing workflow that is executed by running different simulators which interact and finally provide the simulation output that in turn supports policy decisions. A scenario template can be instantiated in a scenario with specific input data, calibrated and defined for a city/area of interest. Transport modelers configure scenarios for their modeling needs (e.g. for examining zero emission zones and green districts or for planning consolidation centers). Scenarios can be run with different input configuration data in order to allow transport modelers to compare different “treatments” and understand the effects of different configurations and designs.

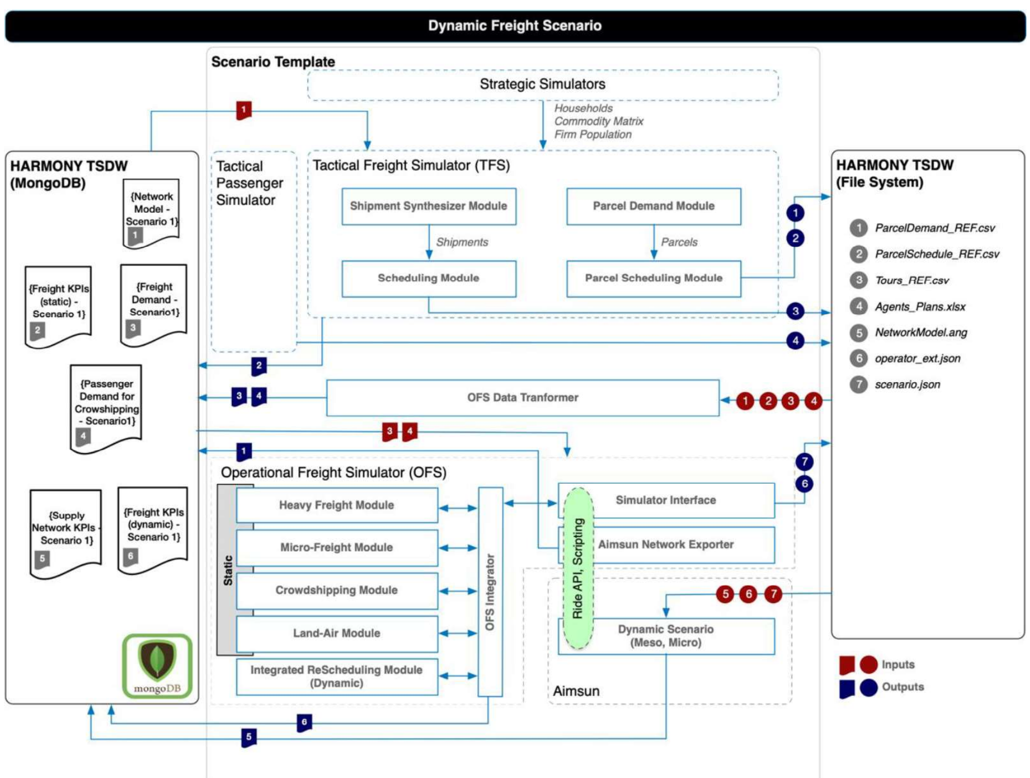

Figure 2: The scenario template for dynamic freight simulation and the data configurations which result in a related scenario.

Figure 2 shows the scenario template that realises the evaluation of freight services/operations using Micro/Meso simulation. The scenario commences with the generation of freight demand (heavy freight and parcels) using the different modules included in the Tactical Freight Simulator (TFS) and the generation of passenger demand, from the Tactical Passenger Simulator (TPS), in the form of activity schedules. The former, involves two synthesizers (one for shipments and one for parcels) which generate the demand in the form of O/D matrices, and two scheduling modules which convert the O/D matrices into tours that different types of vehicles need to undertake. The demand and schedules from the tactical layer are fed into the Operational Freight Simulator (OFS) which process it and generates alternative schedules by integrating heavy- and micro-freight, crowd shipping and air freight services. These schedules are then simulated using dynamic traffic assignment in the supply simulator and rescheduling of vehicles/services is actioned when necessary. The HARMONY TSWD hosts the input and output data required to realize different scenarios based on this scenario template. The following subsections provide an overview of the data used to describe a scenario template and corresponding scenarios.

5.1 Scenario Templates

For each scenario template the following metadata are configured:

- ScenarioTemplate_ID: System identifier of the scenario template which refers to an object id generated when a new scenario template is created in the database.
- Name: Descriptive name of the scenario template.
- Description: Detailed information which shows the aim and scope of the scenario template.
- Simulation levels: Involved strategic/tactical/operational simulators required to run the scenario template.
- Simulation workflow: The list of simulators that are part of this scenario template. It refers to an array of IDs of simulators that the platform needs to run in order to execute the scenario template workflow.
- Category: Category of this scenario template
- KPIs for evaluation: list of KPIs that are calculated with the use of the output data of the simulation workflow
- Mandatory data: Array of mandatory data that are needed to run scenarios based on this template, including the network model.

As an example, the entry for a dynamic freight scheduling scenario template is the following:

- Scenario_ID: ObjectId("607f145e810e154329abe430df")
- Name: Dynamic freight scheduling scenario template.
- Description: This scenario template runs the tactical freight simulator to generate the demand and subsequently the operational freight simulator and meso/macro simulators for dynamic scheduling of different freight services.
- Simulation levels: tactical, operational
- Simulation workflow: ["T2", "O2", "O1"] // where T2 is the tactical freight simulator, O2 is the Operational Freight simulator and O1 is the meso/macro simulator
- Category: freight scheduling
- KPIs for evaluation: freight delivery delay times, number of trips per logistic segment, km per logistic segment and vehicle type, fleet utilization, freight CO2 footprint
- Mandatory data: network_model, tours and trips for heavy freight and parcels delivery, firms population, commodity matrix and households, activity schedules, fleets configuration

And the corresponding json format is as follows:


```
{
  "id": "d290f1ee-6c54-4b01-90e6-d701748f0851",
  "name": "Dynamic freight scheduling scenario template",
  "description": " This scenario template runs the tactical freight simulator to
generate the demand and subsequently the operational freight simulator and
meso/macro simulators for dynamic scheduling of different freight services",
  "date": "2020-10-18T16:00:00Z",
  "simulation_levels": ["tactical", "operational"],
  "simulation_workflow": ["T2", "O2", "O1"],
```


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

```
  "KPIs": "freight delivery delay times", "number of trips per logistic segment",
"km per logistic segment and vehicle type", "fleet utilization", "freight CO2
footprint",
  "mandatory_data": [{"network_model": "refers to a file containing the newtwork
model of the area of interest to be used by the simulators", "heavy_freight_tt":
"refers to a file containing trips and tours for heavy freight",
"parcels_delivery_tt": "refers to a file containing trips and tours for parcels
delivery", "activity_schedules": "refers to a file containing activity schedules",
"fleets_configuration": "configuration of vehicle types, their characteristics and
composition", "firms_population": "", "commodity_matrix_and_households": ""}]
}
```


5.2 Scenarios

For each scenario, the following metadata are configured:

- Scenario_ID: System identifier of the scenario which refers to an object id generated when a new scenario is instantiated from a scenario template and created in the database.
- Scenario Template: Object id of reference scenario template. The simulation workflow of the corresponding scenario template is used to run the simulations of this scenario.
- Name: Descriptive name of the scenario.
- Description: Detailed information which describes the specifics of the scenario.
- InputData: The set of input data that are needed to run the scenario. These are instances of the scenario template mandatory data.
- Output data for KPI calculation: raw data generated from the simulators of the scenario.
- KPIs for evaluation: List of key performance indicators calculated at the end of the scenario for decision support. These are instances of the KPIs defined in the scenario template.
- Category: Scenarios fall under predefined categories for example “planning of drone landing points”, “planning of consolidation centers”, “low emission zones”, etc.

As an example, an entry for a micro freight logistics with low emission zones scenario is the following:

- Scenario_ID: ObjectId(“30fd231e342c128642de098ed”)
- Name: Assessment of low emission zones.
- Description: In this scenario we consider evaluation of micro freight logistics with low emission zones.
- Category: “low emission zones”
- Tags: "freight scheduling", "CO2 emissions", "low emission zones"
- Input data:
    - Network model describing the low emission zones under consideration
    - Fleets configuration / vehicle types, their characteristics and composition (e.g. 10 vans, microbikes)
    - Trips and tours for heavy freight
    - Trips and tours for parcels delivery
    - Activity schedules
    - Firms population
    - Commodity matrix and households
- Output data for KPI calculation:
    - Travel time data for tours
    - Delays in delivering goods
    - Fleet utilization data
- KPIs:
    - Freight CO2 footprint

And the corresponding json format is as follows:


```
{
  "id": "30fd231e342c128642de098ed",
```


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

Part of: 2020 CIVITAS
THE CIVITAS INITIATIVE IS CO-FINANCED BY THE EUROPEAN UNION

```
  "name": " Assessment of low emission zones",
  "description": " In this scenario we consider evaluation of micro freight
logistics with low emission zones",
  "date": "2020-10-18T16:00:00Z",
  "category": "freight scheduling", "CO2 emissions", "low emission zones",
  "tags": "freight scheduling", "delay times", "logistics services",
  "KPIs": "freight CO2 footprint",
  "input_data": [{"network_model":
ObjectId("30fd231e342c128642de0981d"),"heavy_freight_tt":
ObjectId("30fd231e342c128642de0982d"), "parcels_delivery_tt":
ObjectId("30fd231e342c128642de0983d"), "activity_schedules":
ObjectId("30fd231e342c128642de0984d"), "fleets_configuration":
ObjectId("30fd231e342c128642de0985d"), "firms_population":
ObjectId("30fd231e342c128642de0986d"), "commodity_matrix_and_households":
ObjectId("30fd231e342c128642de0987d")}]
  "output_data": [{"travel_times_for_tours":
ObjectId("30fd231e342c128642de0988d"), "delays_in_delivering_goods":
ObjectId("30fd231e342c128642de0989d"), "fleet_utilization":
ObjectId("30fd231e342c128642de098ad")}]
}
```


6 Simulations input data

The simulators' input data include mainly configuration and initialization of transport related variables (for example the network model to be used or travel demand). The related data are stored in the HARMONY TSDW, as part of the database or as data artifacts in the HARMONY server filesystem. Deliverable D3.2 provided an overview of the HARMONY simulators input data while the data management framework provided in this deliverables deals with the metadata required to describe these data and readily use them as input to the simulators.

The following metadata properties are considered per input data instance:

- InputData_ID: System identifier of the input data which refers to an object id generated when created in the database.
- Name: Descriptive name of the input data.
- Format: Can be database object or file (artifact) stored in the filesystem.
- Access details: Can be a query string for fetching the data or a path in the filesystem.
- Type: Denotes the type of data, for example a network model, an OD matrix etc.
- AssociatedScenarios: List of scenarios where the data are used (as ObjectIds).

As an example, a subset of the entries for a micro freight logistics with low emission zones scenario are the following:


```
{
  "id": "30fd231e342c128642de0981d",
  "name": "Network model for Oxfordshire – valid for low emission zones
evaluation",
  "date": "2020-10-18T16:00:00Z",
  "format": "artifact",
  "access_details": "/harmony/network_models/occ_model.ang",
  "type": "network_model",
  "associated_scenarios": [ObjectId("30fd231e342c128642de098ed")],
}
{
  "id": "30fd231e342c128642de0982d",
  "name": "Trips and tours for heavy freight at Oxfordshire",
  "date": "2020-10-18T16:00:00Z",
```


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

```
  "format": "artifact",
  "access_details": "/harmony/network_models/heavy_freight_tt_v1.csv",
  "type": "heavy_freight_tt",
  "associated_scenarios": [ObjectId("30fd231e342c128642de098ed")],
}
{
  "id": "30fd231e342c128642de0983d",
  "name": "Trips and tours for parcels delivery",
  "date": "2020-10-18T16:00:00Z",
  "format": "artifact",
  "access_details": "/harmony/network_models/parcels_delivery_tt_v1.csv",
  "type": "parcels_delivery_tt",
  "associated_scenarios": [ObjectId("30fd231e342c128642de098ed")],
}
{
  "id": "30fd231e342c128642de0984d",
  "name": "Activity schedules at OCC",
  "date": "2020-10-18T16:00:00Z",
  "format": "artifact",
  "access_details": "/harmony/network_models/activity_schedules_v1.csv",
  "type": "activity_schedules ",
  "associated_scenarios": [ObjectId("30fd231e342c128642de098ed")],
}
```


7 Scenario Execution Logs

Each simulator generates logs that provide a detailed view of how the simulator runs. Depending on the implementation approach of each simulator the logs can be stored as part of the TSDW either on the HARMONY database or in the server's filesystem. When creating a log, the simulators provide the following metadata:

- LogID: System identifier of the log data which refers to an object id generated when created in the database.
- AssociatedSimulator: The simulator that creates the log.
- AssociatedScenario: The scenario associated with the log.
- Format: Can be database object or file (artifact) stored in the filesystem.
- Access details: Can be a query string for fetching the data or a path in the filesystem.

As an example, a subset of the entries for a micro freight logistics with low emission zones scenario are the following:


```
{
  "id": "30fd231e342c128642dd0911d",
  "date": "2020-10-18T16:00:00Z",
  "associated_simulator": "O2"
  "format": "artifact",
  "access_details": "/harmony/network_models/logs/30fd231e342c128642de098ed
/O2_log.csv",
  "associated_scenarios": [ObjectId("30fd231e342c128642de098ed")],
}
```


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

8 Simulations Output Data

At the end of a simulation scenario a set of output data are stored by the simulators for further processing. These data are mainly used for the calculation of scenario KPIs. The metadata which are associated with the output data of a simulator include:

- OutputData_ID: System identifier of the output data which refers to an object id generated when a new output data entry is created in the database.
- Name: Descriptive name of the output data.
- Format: Can be database object or file (artifact) stored in the filesystem.
- Access details: Can be a query string for fetching the data or a path in the filesystem.
- Associated Scenario: The scenario that produces the output data.

As an example, a subset of the output data for a micro freight logistics with low emission zones scenario are the following:

```
{
  "id": "30fd231e342c128642de0911d",
  "name": "Travel time data for tours",
  "date": "2020-10-18T16:00:00Z",
  "format": "artifact",
  "access_details": "/harmony/network_models/travel_time_data_for_tours.csv",
  "type": "travel_time_data",
  "associated_scenarios": [ObjectId("30fd231e342c128642de098ed")],
}
{
  "id": "30fd231e342c128642de0922d",
  "name": "Delays in delivering goods",
  "date": "2020-10-18T16:00:00Z",
  "format": "artifact",
  "access_details": "/harmony/network_models/delays_in_delivering_goods.csv",
  "type": "delays_in_delivering_goods",
  "associated_scenarios": [ObjectId("30fd231e342c128642de098ed")],
}
{
  "id": "30fd231e342c128642de0983d",
  "name": "Fleet utilization data",
  "date": "2020-10-18T16:00:00Z",
  "format": "artifact",
  "access_details": "/harmony/network_models/fleet_utilization_data.csv",
  "type": "fleet_utilization_data",
  "associated_scenarios": [ObjectId("30fd231e342c128642de098ed")],
}
```

9 Analytics and Insights

This data entity is used to describe Key Performance Indicators (KPIs) associated with a Scenario. The metadata describing the KPIs include:

- KPIName: Descriptive name of the KPI
- Date: Datetime object showing when the KPI was last updated.
- AssociatedScenario: A scenario for which the KPI is calculated.
- Type: This can be a time series or a single value.
- KPI calculated values: An object providing the calculated values of the KPI depending on its type.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

As an example, a subset of the KPIs for a micro freight logistics with low emission zones scenario are the following:

```
{
  "id": "30fd231e342c128642de0911d",
  "KPIname": "average freight CO2 emissions",
  "date": "2020-10-18T16:00:00Z",
  "AssociatedScenario": "[ObjectId("30fd231e342c128642de098sd")]"
  "type": "single_value",
  "calculated_values": "105gr/km",
}
{
  "id": "30fd231e342c128642de0911d",
  "KPIname": "average hourly freight CO2 emissions",
  "date": "2020-10-18T16:00:00Z",
  "associated_scenario": "[ObjectId("30fd231e342c128642de098ed")]"
  "type": "time_series",
  "calculated_values": {"09:00": "100gr/km", "10:00": "105gr/km", "11:00":
"110gr/km", "12:00": "110gr/km", "13:00": "105gr/km", "14:00": "100gr/km"}
}
```

10 Data Warehouse Implementation Status and Next Steps

The TSDW and the related data management framework have been instantiated in the HARMONY development servers aiming to support the platform implementation within the first period of the project. For this purpose, a dockerised instance of the MongoDB NoSQL database has been configured and a filesystem structure that accommodates the execution of the HARMONY simulators has been setup.The data entities described in sections 3-9 have been defined and codified in the Swagger platform (https://app.swaggerhub.com) which provides the means to generate RESTful interfaces for developing and consuming APIs by effectively mapping all the resources and operations associated with them. All defined interfaces together with their documentation are provided in Appendix I.

In terms of next steps, the HARMONY TSDW follows the implementation of the HARMONY platform and will be adapted in order to accommodate the data required for the simulation scenarios defined by the transport modelers.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815269

Part of: 2020 CIVITAS

THE CIVITAS INITIATIVE IS CO-FINANCED BY THE EUROPEAN UNION