ASSESSMENT OF THE TESTING RESULTS IN THE CASE STUDY OF LISBON

D7.4

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 955273

Deliverable administrative information

Deliverable number	7.4
Deliverable title	Assessment of the testing results in the Case Study of Lisbon
Dissemination level	Public
Submission deadline	31/10/2024
Version number	1.0
Authors	Joana Cunha (CARRIS) João Vieira (CARRIS) Ana V. Silva (A-to-Be) Marco Grassi (CEFRIEL) Marco Comerio (CEFRIEL) Mario Scocca (CEFRIEL) Charis Chalkiadakis (NTUA) Marios Giouroukelis (NTUA) Athina Tympakianaki (Aimsun) Ynte Vanderhoydonc (imec) Mohammadmahdi Rahimiasl (imec) Arka Ghosh (Deusto) Antonio David Masegosa Arredondo (Deusto)
Internal reviewers	Lucie Tristant (ID4MOBILITY)
Document approval	Not applicable

Legal Disclaimer

https://twitter.com/TANGENT_H2020

https://www.linkedin.com/company/tangent-project/

https://www.youtube.com/channel/UCjhz4kwEm_sTHj7fE4zXToA

For further information please visit http://www.tangent-h2020.eu/

Executive summary

The Lisbon Metropolitan Area (LMA) is a grouping of 19 municipalities centered around the city of Lisbon, the capital of Portugal. It has over 2.8 million residents, with Lisbon alone being home to over half-a-million inhabitants. Ever since the COVID-19 pandemic, there has been a progressive worsening of congestion observed in the city, both due to an increase in the population around the LMA, and to a migratory effect of residents shifting away from the city center and towards other municipalities. As a high percentage of jobs are still located in the capital, this has led to a pronounced intensification of commuting traffic entering and exiting the city.

Having taken the pledge to become one of the 100 Climate-Neutral and Smart Cities, the city of Lisbon has the ambition to achieve carbon-neutrality by 2030, and has designed a set of strategies to promote the use of Public Transport (PT) and active modes. Nevertheless, bus and tram operators are subject to the same conditions as cars, and thus likewise negatively affected by the city's congestion.

As a pilot site in the TANGENT project, Lisbon focuses on the experimentation of novel tools for real-time monitoring of traffic conditions, and in the design of strategies to mitigate the negative effects of congestion and disruptions, to improve the public transport services. The case study focuses on the three Westernmost parishes of Lisbon, which are the entryway to the city from the neighboring municipalities, both for car-users and PT passengers.

In Lisbon, three TANGENT services were tested: Service 1, an enhanced multimodal information integration service that enabled the consolidation of real-time data from diverse internal and external sources into a single dashboard that offers users an overview of mobility conditions in the city; Service 2, focused on incident management and the engagement and collaboration of stakeholders to solve incidents; and Service 3, which facilitates the planning of response actions and aids decision-making through simulation.

The main lessons learned during the testing process relate the need for the inclusion of diverse stakeholders and multimodal data sources in the network analysis: it is important to engage and motivate transport entities to collaborate in the development of data sharing protocols that enable the integration of multimodal information, which benefits other operators and (through platforms such as the one developed by TANGENT) themselves. Moreover, an effort should be made to upgrade the existing data sources and data collection infrastructure, to enable the gathering of more detailed and better-quality information, which is essential for creating a more accurate picture of the city's mobility conditions.

The present deliverable describes the testing and validation methodologies conducted during the testing activities, the obtained results, the lessons learned, and next steps for the Lisbon case study.

Key words

Traffic Management, Public Transport Operations, Data integration, Multimodality, Real-time conditions

Table of contents

List of figures

List of tables

List of abbreviations and acronyms

Acronym	Meaning
AMQP	Advanced Message Queuing Protocol
CIM	Cooperative Incident Management
CS	Case Study
CSV	Comma-Separated Values
DRT	Demand Responsive Transport
DTA	Dynamic Traffic Assignment
EC	European Commission
GA	Grant Agreement
HTML	HyperText Markup Language
KCONG	Knowledge Catalog and Governance
KPI	Key Performance Indicator
LMA	Lisbon Metropolitan Area
LoS	Level of Service
NAP	National Access Point
OP	Other Partners

Acronym	Meaning
PT	Public Transport
R&D	Research & Development
RDF	Resource Description Framework
SCS	Stakeholder Cooperation Survey
SNLB	Smart Network Load Balance
TBS	Travel Behavior Survey
TC	Traffic Control
TMB	TANGENT Data Catalogue Management Board
TMS	Traffic Management System
WP	Work Package

1 Introduction

1.1 Attainment of the objectives and explanation of deviations

The main objective of this deliverable is to present the results of TANGENT for the Lisbon case study. In line with the *WP7 "FUTURE TRANSPORT CASE STUDIES"*, the present deliverable summarizes the outputs of *T7.2 "Operation of the tests: testing of subsystems"* and *T7.3 "Upscaling and full operation: testing of the overall system"*, comparing them to the initial expectations defined by the Lisbon case study during *T7.1 "Definition of Case Studies, stakeholders' engagement and tests preparation"*, and specified in deliverables *D7.1 "Definition of the Case Studies, testing methodology and stakeholders' & users' engagement campaign"* and *D7.2 "System deployment and testing specifications"*. These documents present the Lisbon Case Study scope and requirements, proposed testing environment and the scenarios under which the tests were to be conducted.

The present deliverable has the core goal of presenting the results of the testing activities conducted in Lisbon for the three TANGENT services, assessed for the different scenarios: baseline, planned events and unplanned events. The planned event scenarios included a large planned event in Passeio Marítimo de Algés, as well as sports events in the Belém riverside area.

As further explained in this deliverable, due to the higher-than-expected technical complexity of the modules, some testing activities are still pending completion. These testing are ongoing and shall be continued until the end of the project to ensure the full evaluation of all systems.

1.2 Intended audience

This deliverable is public. First, it can be retained, reused and disseminated by CARRIS and A-to-Be, to exploit the results of TANGENT. It may also be consulted by the local stakeholders involved in the activities (workshops, demonstrations, tests) of the project who are not part of the consortium. More broadly, this deliverable can be consulted by any interested party: stakeholders involved in traffic and transport management, whether they are techno providers, other metropolises or transport operators, researchers, or other R&D projects.

1.3 Structure of the deliverable and links with other work packages/deliverables

The present document is structured into four main sections

- "Presentation of the Case Study" briefly describes the scope of the Lisbon demonstration, allowing the reader to understand the context of the TANGENT project and the specific tests conducted in Lisbon
- "Testing of the Subsystems" intend to present the testing activities preformed for each subsystem component, while they were under development by the technical partners and prior to their integration in the consolidated platform. This section outlines the assessment methodology and the results obtained in Lisbon for each tested subsystem
- "Testing of the overall system" reports the results of the tests performed on the final consolidated system, which integrates the developed subsystems as functionalities of the TANGENT tool. It includes the adopted testing methodology and the outcomes attained in the Lisbon case study, for a variety of scenarios

- “Lessons learned and next steps” concludes on the lessons learned in the Lisbon case study, including the challenges encountered and overall impressions gathered during the development and testing stages of the several subsystems and tools. It further highlights the next steps for CARRIS and A-to-Be after the end of the project

The deliverable D7.4 is closely linked to other TANGENT work packages and deliverables, namely:

- Within WP7, this document follows up on D7.1 “Definition of the Case Studies, testing methodology and stakeholders' & users' engagement campaign” and D7.2 “System deployment and testing specifications”, where the Lisbon case study’s scope and requirements are defined. It further links to the other case study deliverables (D7.3 “Assessment of the testing results in the Case Study of Rennes”, D7.5 ”Assessment of the testing results in the Case Study of Greated Manchester” and D7.6 ”Assessment of the testing results in the Case Study of Athens”) in which the same testing methodologies were applied to the remaining TANGENT case studies. Moreover, D7.7 “Impact Assessment” provides an evaluation of the impacts of TANGENT project in the Lisbon case study, through the analysis of the selected KPIs
- This document is also linked with the WP1 workshops, notably the Stage 3 stakeholders workshop in Lisbon, during which a demonstration of the tool took place and feedback from the local stakeholders was gathered
- Furthermore, D7.4 is closely linked with WP6, specifically to deliverable D6.5 “TANGENT tool development and integration. Second release”, which describes the overall tool, the dashboard functionalities and the technical testing. This document is referred to in several sections of D7.4.

2 Presentation of the case study

The Lisbon Case study focuses on the Western part of the city, encompassing 3 parishes: Belém, Ajuda and Alcântara. As shown in Figure 1, the chosen area is bordered south by the Tagus River; by the Monsanto green corridor, at its top and left, enclosing the beginning of A5 from Viaduto Duarte Pacheco (close to Av. de Ceuta) to the Alfragide node and the end of CRIL (Cintura Rodoviária Interna de Lisboa) from there (Alfragide node) to Algés and by Avenida de Ceuta from A5 to Alcântara Mar, perpendicularly to Tagus.

With around ~43,000 inhabitants [Censos, 2021], this is a highly touristic area with several major Museums and Monuments of the city. It is also the location of several university campuses (Agronomy, Veterinary, Architecture and Social Sciences). Commuters enter Lisbon from Cascais and Oeiras either by suburban train or through A5 and N9, and from Amadora through IC19 and CRIL. Some urban and suburban buses also use these high-traffic roads.

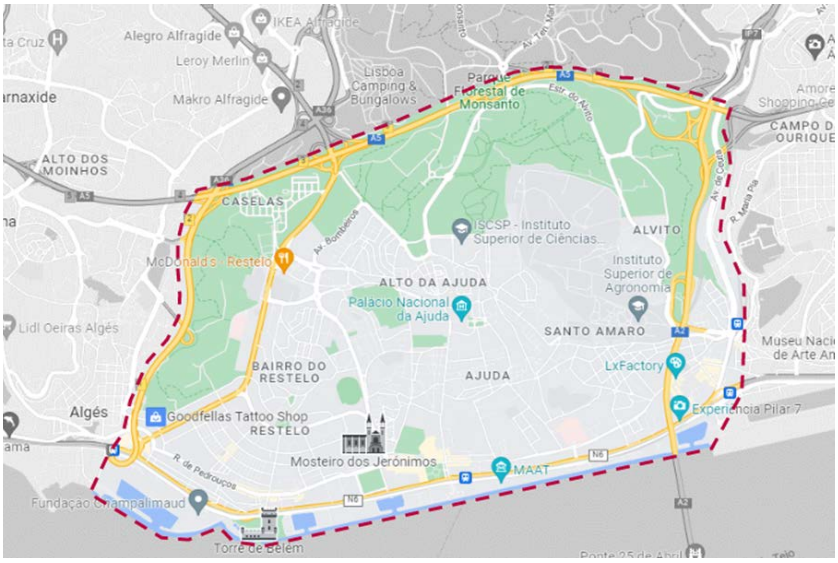


Figure 1: Overview of the selected area for the TANGENT Case Study in Lisbon: Western Lisbon

The chosen case study area is also the stage of several events throughout the year, the most notable being the NOS Alive music festival, which annually draws ~150-200 thousand people to the Passeio Marítimo de Algés venue.

Bearing in mind the scale of the chosen Case Study area, two main corridors (Figure 2) were selected to focus the analysis of the TANGENT tools/functionalities:

- the ‘Express Corridor’ connects the Algés multimodal hub to the A5, offering a fast connection to central Lisbon. This corridor was chosen to test the Service 2 CIM functionality (Traffic Control and Public Transport Synchronization).
- the ‘River Corridor’ connects the Algés multimodal hub to other interfaces alongside the Tagus River. This corridor was selected because it is a high-demand PT corridor that passes near several attraction points and event locations in the case study area.

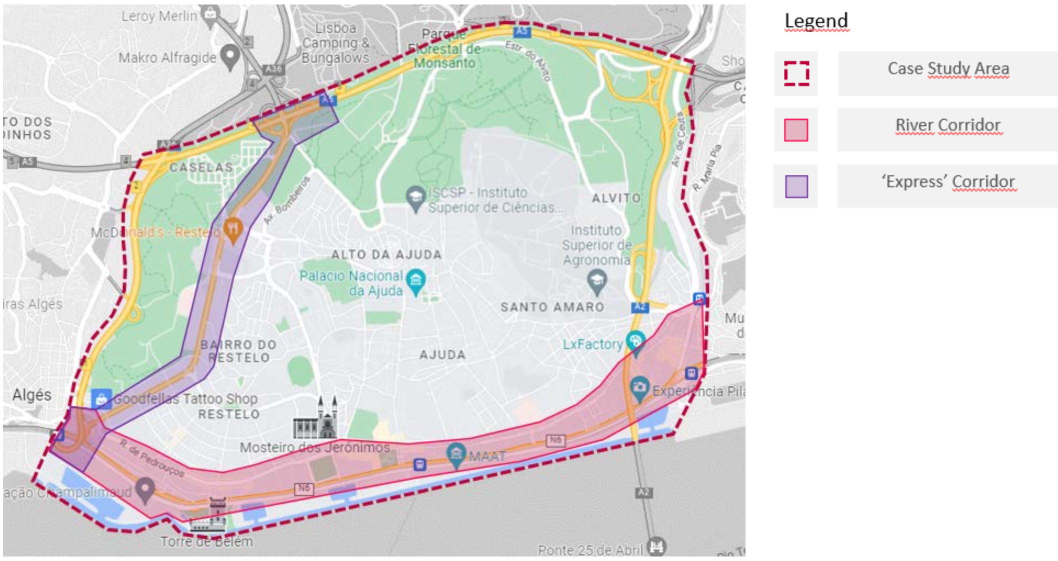


Figure 2: Express and River Corridors

The following scenarios were selected for testing the TANGENT tools:

- The baseline scenario is a typical working day, during peak commuting hours.
    - A typical working day is defined as a work day, excluding school vacation periods (i.e., Summer months, Christmas and Easter). The peak hours considered are 7:00-10:00 and 17:00-20:00
    - Considered PT services: 1 tram line (15E) and 8 bus lines (714, 723, 727, 728, 729, 732, 751, and 79B)
- TANGENT functionalities testing conditions:
    - Services 1 (Current Status), Service 1 (Future Status) and Service 2 (SNLB) to be tested under the baseline scenario conditions
    - Service 2 (CIM) to be simulated only for the 'Express Corridor' and for a subset of 5 bus lines (714, 723, 732, 751 and 79B)
    - What-if scenario 1 to simulate the increase of frequency in a subset of 5 lines (714, 723, 732, 729 and 751)
    - What-if Scenario 2 to simulate the introduction of Dynamic Congestion Pricing on the road access points in the A5 highway

Further details on the Case Study and testing specifications can be consulted in D7.1 and D7.2.

3 Testing of the subsystems

This section describes the testing activities carried out during T7.2 “Operation of the tests: testing of subsystems”, in relation with the technical partners and the first release of their subsystems. The methodology used is based on the “Applied Learning Cycle”: once the technical partners produced the first release of their subsystems, they have organised tests with the case studies to assess the accuracy of their results. The results of these testing activities and the feedback from the case studies were used to improve the second release of their subsystems. Each technical partner produces a different type of subsystem; consequently, they have provided a testing methodology adapted to the elements to assess.

Each following subsection presents the subsystem to be tested, the methodology of the test and the results of the testing activity. A great amount of the following elements is common to the deliverables of the other case studies, mainly the non-virtual ones (D7.3 for Rennes, D7.5 for Greater Manchester). The decision was made to keep and duplicate these sections in all the mentioned deliverables rather than referring to other documents. This way, it maintains consistency in each document, and ensures a complete deliverable for each city, which, after the end of the project, they can use and disseminate as a whole in their local ecosystem.

3.1 Data harmonisation and fusion

The TANGENT Solution for Data Harmonisation and Fusion represents the subsystem developed within WP2 and integrated into the overall TANGENT solution. The first release of the solution is described in D2.3, and the final one, incorporating the feedback received during its testing, is in D2.4. The main components of the TANGENT Solution for Data Harmonisation and Fusion are:

- the TANGENT Data Catalogue: a digital platform enabling data-sharing among different stakeholders;
- the TANGENT Reference Conceptual Model: defining a reference ontology for the transportation-related data handled within TANGENT;
- the Semantic Harmonisation and Fusion Pipelines to fulfil the integration requirements associated with TANGENT data sources;
- the TANGENT Data API: a uniform mechanism to access data sources collected and harmonised by WP2.

The testing of the WP2 subsystem mainly focused on the TANGENT Data Catalogue, which provides a user interface designed to be directly accessible by relevant stakeholders. The testing of the other components has been performed (i) with TANGENT partners from a technical perspective to enable the integration of the different subsystems and (ii) indirectly by stakeholders involved in the testing of other subsystems integrated with the WP2 solution.

The TANGENT Data Catalogue is developed using the Knowledge Catalog and Governance (KCONG) framework by Cefriel. KCONG provides a customizable platform for defining and managing digital asset catalogues, accessible online via a web browser. It supports various digital asset types, each characterized by specific metadata descriptors for user interaction. KCONG allows programmatic access to the list of assets and their metadata, which can be serialized in RDF for advanced querying and automated processing. It also enables the configuration of governance processes for digital assets using the BPMN formalism.

For the TANGENT Data Catalogue, a customized KCONG instance includes:

- organization of digital assets based on TANGENT case studies;
- a Data Source asset type to describe datasets and data services, adhering to the TangentDCAT-AP metadata specification (see D2.2);
- integration of BPMN processes to support TANGENT Data-Sharing Governance rules (see D2.2).

3.1.1 Methodology

The methodology for the TANGENT Data Catalogue considered two distinct types of testing.

The first type involves gathering feedback on the initial version of the TANGENT Data Catalogue. This feedback-driven testing aims to identify areas for improvement to ensure the second and final release of the TANGENT Data Catalogue meets users' requirements. Given that the TANGENT Data Catalogue is a software artefact used uniformly by project partners, the feedback gathered through its testing methodology is not specific to any single use case. Instead, we collected and analyzed anonymous data from all participating partners across use cases to obtain a comprehensive assessment of the catalogue's usability and functionality. By analysing this feedback, we identified common issues and areas for improvement that were used to enhance the overall user experience in the final release of the WP2 solution described in D2.4.

The second type of testing focuses on the catalogue's practical usage within the project. This in-project testing evaluates how effectively the TANGENT Data Catalogue serves its intended purpose in real-world scenarios, ensuring its functionality aligns with the users' needs and expectations.

3.1.1.1 User testing of the TANGENT Data Catalogue

The first version of the TANGENT Data Catalogue, documented in D2.3, was released to partners at M20. To gather opinions and feedback from real users of the catalogue, we designed a set of practical hands-on tasks to be performed by partners on the first version of the Data Catalogue as testing activities within WP7. This testing activity took place from 24th July to 6th September 2023. For the Lisbon case study, CARRIS partners from Lisbon were involved.

Users were introduced to the platform by a short video tutorial explaining the basics of the TANGENT Data Catalogue. Then the users were asked to perform a set of tasks to test the catalogue's intended usage and functionalities. As an immediate follow-up to these tasks, users were asked to respond to a survey to collect explicit and implicit feedback. To improve the quality of the feedback received the survey was made anonymous. Explicit feedback was provided by users in the form of a free text response to an open-ended question. Implicit feedback was collected by asking closed questions related to one of the steps required to complete a task. For instance, if a user provides an incorrect answer to a closed question, it could suggest that the user finds the catalogue’s usage unclear, serving as implicit feedback.

A slightly different set of tasks and accompanying surveys was prepared for case study leaders and other partners taking into account their expected usage of the catalogue. Indeed, case study leaders are expected to use the catalogue to add, edit and browse data sources, while the other users mainly browse available data sources. As such, the tasks and questionnaire were designed to reflect this difference.

As an example, here we discuss a task and the set of associated questions. This task was posed to both case study leaders and technical partners.

Task 1 - Search and analysis of a specific data source

Scenario:

Imagine you need to determine if there was a roadblock in a specific street in Rennes ten days ago. You are invited to check if the "REN_TRAF_11" data source satisfies your needs.

Your assignment is:

Please search the "REN_TRAF_11" data source in the data catalogue and read its metadata description to assess its relevance for addressing the need described in the above scenario.

Answers and follow-up questions:

Did you succeed in finding the REN_TRAF_11 data source?

- *Yes*
  - *Please, describe the steps you did to find the REN_TRAF_11 data source.*
  - *Based on the metadata description, do you think that REN_TRAF_11 data source can be used for addressing the need described in the above scenario? Please, justify your answer.*
- *No*
  - *Please describe any problems you encountered while searching for the data source.*

Through this question, several pieces of information can be gathered. By responding "Yes", we first establish that the user was able to perform the task correctly. The follow-up question then gathers details on how the user performed the task, specifically through which series of steps. This answer is crucial as it reveals whether the user was correctly guided through the intended user experience via the catalogue's user interface. If the user followed an unintended path to complete the task, we can use this information to redesign and improve the user experience. The next follow-up question aims to determine if the adopted metadata profile is clear to the user and if it can be used to solve a task that mirrors a potential real-world use case. The most important information is the one gathered when the user is not able to perform the task. This is crucial, as this information is the main driver for significant changes in the catalogue.

Case study (CS) users and other partners (OP) were assigned various tasks. CS users were given more tasks compared to OP users because they have a more active role in utilizing the catalogue, including adding and editing data sources. In contrast, OP users primarily engage in searching and browsing the catalogue. Feedback from both user groups was collected on the following tasks:

- Searching for a specific data source and analyze its metadata description (CS, OP)
- Searching for a data source that satisfies a specific data requirement (CS, OP)
- Free exploration of the data catalogue and usage of its functionalities (CS)
- Adding and publishing a new data source into the data catalogue (CS)
- Usage of the dashboard functionality to check the status of data sources (CS, OP)

The questions asked to users after these tasks were designed in the same spirit as the previously described task. Similarly, the answers to these questions were used in a similar vein to drive improvements to the catalogue.

3.1.1.2 In-project testing of the TANGENT Data Catalogue

The TANGENT Data Catalogue is designed to manage the diverse data sources collected from the project's four case studies. This catalogue is not merely a static repository of data and metadata, it is intended to serve as a dynamic, evolving archive that adapts to the project's needs. To facilitate this adaptability and support the project's potential developments, a life cycle has been established for the data sources. Each data source can exist in one of several states, transitioning between them based on specific user actions and conditions set by the TANGENT Data Catalogue Management Board (TMB). The TMB oversees the data quality and governance, ensuring the catalogue remains a robust and responsive tool throughout the project's duration. The list of admissible statuses is the following:

- *Submitted*: When a data source is inserted by a user into the catalogue.
- *Published*: When a data source is reviewed by the TMB and approved for publication.
- *Rejected*: When a data source is rejected after a review by the TMB because it is found to not be useful or applicable to the project.
- *Incomplete*: When a data source lacks certain required metadata fields that have not been filled out by the user. The data source is sent back to the user with a request for further information.
- *Not accessible*: A specific case of incomplete metadata where the data source description lacks the access URL to the actual data it describes.

Regarding the in-project testing activity, we report on the number of data sources collected through the TANGENT Data Catalogue and approved for usage within the TANGENT project for the Lisbon Case Study.

3.1.2 Actions achieved and results

In the following, we report the aggregated results obtained for the user testing of the TANGENT Data Catalogue and the results of the in-project testing for the Lisbon case study.

3.1.2.1 User testing of the TANGENT Data Catalogue

The user testing activity involved 14 partners acting as case study leaders and 9 other partners. From these groups, the survey aimed at gathering constructive feedback for the TANGENT data catalogue received responses from 9 out of 14 demo leaders and 3 out of 9 other partners. The testing activity was conducted using the first version of the TANGENT Data Catalogue, described in detail in D2.3 and shown in Figure 3

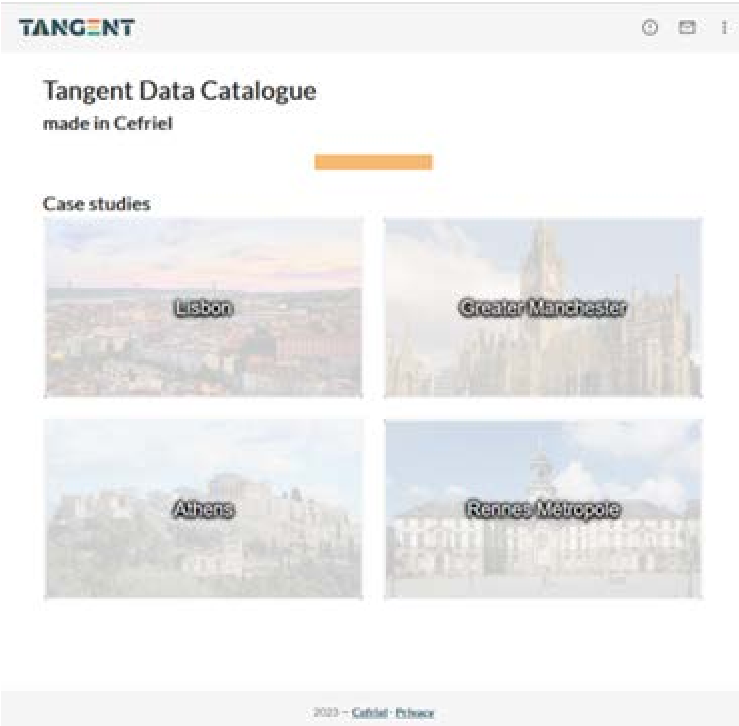


Figure 3: Homepage of the first version of the TANGENT Data Catalogue

After analyzing the current survey responses on the testing of the TANGENT data catalogue, the main areas for improvement have been identified as follows:

1. **Search Capabilities and Interface**: Numerous feedbacks indicated the need for enhanced search functionality, specifically the ability to search by metadata attributes that are not currently available as filter options. Additionally, the filtering capabilities offered through the user interface could be made less confusing and more user-friendly.
2. **Dashboard Information and Interactivity**: The information displayed on the dashboard needs to be expanded and made more interactive. Users have expressed a desire for a more detailed and engaging dashboard experience.
3. **Action Feedback**: Certain actions, such as submitting a data source for publication, currently lack feedback mechanisms that inform the user of the action's result. Providing clear and immediate feedback for these actions would improve the overall user experience.

Based on the feedback we gathered, it became evident that a small but notable portion of users experienced difficulty in successfully completing one of the tasks associated with locating a specific data source. This feedback was collected through both quantitative and qualitative methods. Quantitatively, we measured how many users were able to find the data source. Qualitatively, we invited users to provide open-ended feedback detailing the challenges they faced during this task. This dual approach allowed us to capture not only the extent of the issue but also the specific problems encountered, offering a comprehensive understanding of the user experience and the obstacles they encountered. The answers received for task 2 serve as an excellent example of this quantitative aspect, as is shown in Figure 4.

Scenario: A project partner tells you that a data source satisfies the "WEATHER_DATA" data requirement for your case study.

Your assignment is: Try to find the data source in the data catalogue. Read its metadata description carefully.

Figure 4: On the left the percentage of case study leaders who completed task 2; on the right, the percentage for other partners

By addressing these issues, the final version of the TANGENT Data Catalogue improves upon the first release by introducing new functionalities and improving the user experience. For example, the homepage has been changed to show the latest uploaded data sources, which helps users know the latest changes in the catalogue. Additionally, because users often access the catalogue to search for a specific asset, a search box and corresponding search functionality have been added to the homepage.

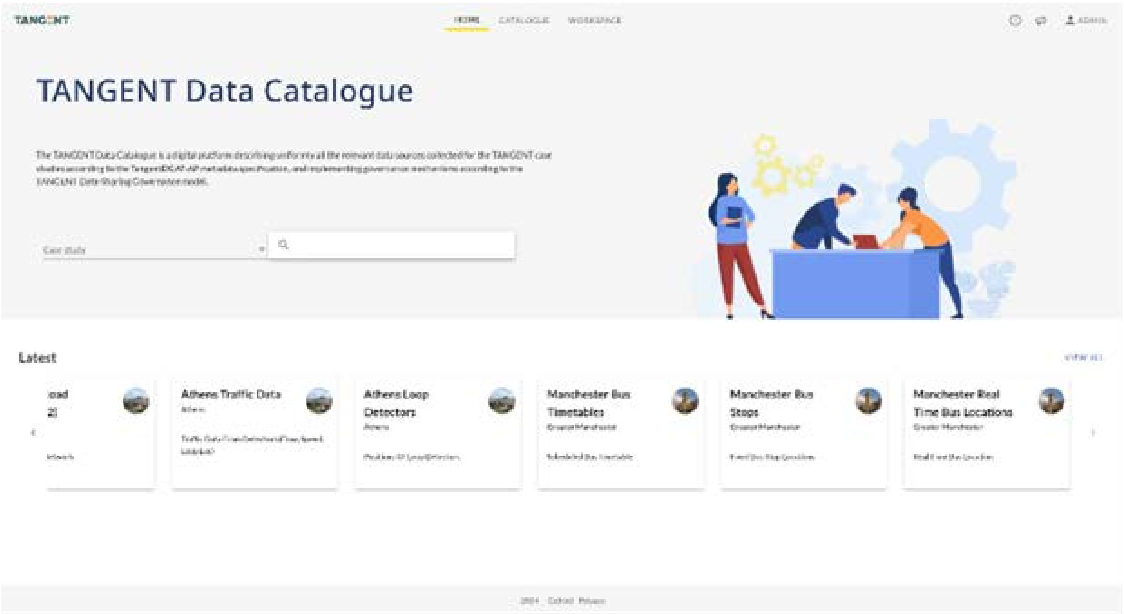


Figure 5: Homepage of the final version of the TANGENT Data Catalogue

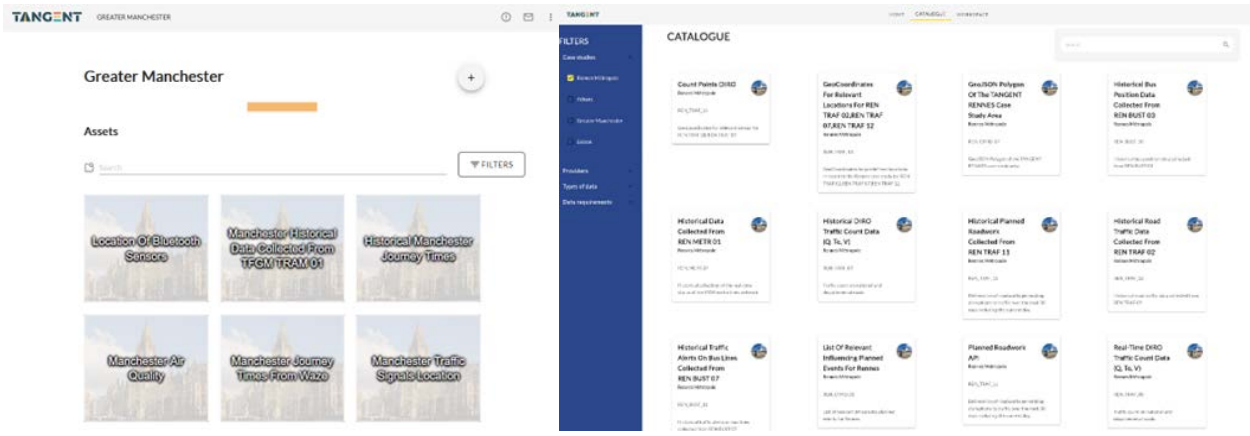


Figure 6: Search functionalities in the different versions of the Tangent Data Catalogue

A site-wide change has been made to separate different functionalities offered by the TANGENT Data Catalogue. As can be seen in the top part of Figure 5, the final release of the catalogue is divided into three main areas: the home, the catalogue and the workspace. As previously described, the homepage is designed to bring the user up to date with the latest changes and offer the shortest possible path to one of the most common interactions with the catalogue, i.e., searching. The catalogue area, shown in Figure 6, allows the user to browse all the data sources available in the catalogue and perform more fine-grained searches. Indeed, as searching and filtering were identified as the main areas of interest during the testing of the first release of the data catalogue, the UI and UX for this functionality were particular points of focus. While the previous version of this page was found to be confusing to users, especially because it was not readily apparent what filters were currently active, the new and improved version clearly shows the active filters on the left side of the page and the used keyword in the search box. Additionally, with this layout, more results can be displayed, making it easier for users to manage searches that retrieve an extensive set of results. Lastly, the workspace area is entered when acting on a single data source or to access an overview of the state of the whole catalogue or the state of the data sources owned by the current user.

3.1.2.2 In-project testing of the TANGENT Data Catalogue

For the Lisbon use case, a total of 34 data sources were collected and submitted through the TANGENT Data Catalogue. After careful review and discussion with relevant project partners, the TANGENT Data Catalogue Management Board (TMB) determined that 11 of these data sources were not relevant and subsequently rejected them. Of the remaining data sources, 22 met the project's data quality requirements and were approved for use. One data source remained inaccessible and was not utilized by any partners.

3.2 Choice model for Synchronization of Public transport and traffic control & Choice model for Dynamic Congestion Pricing (WP3)

The principal output of WP3 is the travel choice module, an analytical tool for modelling travel decision-making under various mobility scenarios. The module, the development of which is thoroughly described in D3.2 and schematically described in Figure 7, builds upon input data elicited using the Travel Behavior Survey (TBS), a structured stated-preferences questionnaire that captures commuters' sociodemographic traits, travel patterns, perceptions of system-level attributes, and preferences for different travel options. At its output, the machine learning and econometric models at its core predict mode choice decisions and identify significant factors influencing commuter decisions.

Figure 7: Travel Choice Module inputs & outputs

Given the design of the modelling framework, the testing of the module aims to ensure that all interface points between the module and the population whose behaviour is modelled fulfil expectations.

- Upstream, the survey’s ability to effectively capture commuter behaviours must be assessed, best achieved through recipient feedback on its clarity, relevance, and effort needed.
- Downstream, it is necessary to evaluate whether the model’s predictions align with common intuitions about travel behaviour in specific cities, with local commuters being again the most qualified evaluators.

As a result, a four-step testing methodology for the WP3 travel behaviour module is developed based on commuter feedback and presented in the sections that follow.

The collected user feedback is critically analysed, offering valuable insights for the development of similar mode choice modules. This methodology is further discussed in the section that follows. Given that the consolidated testing process for all case studies, the sections that follow are virtually identical for D7.3/4/5/6.

3.2.1 Methodology

Step A: Testing sample collection

During the initial phase of testing, the Travel Behavior Survey (TBS) was distributed to commuters from each case study city, including Lisbon. The objectives of this phase are to:

- Familiarise testers with the questionnaire's content and design by having them take the survey themselves.
- Provide the necessary data for executing the travel behaviour module (Step B).
- Obtain the contact points of testers (email addresses) for the distribution of the feedback survey of Step C.

The TBS questionnaire is included in the Annex of D3.2.

To collect the necessary sample, each case study was asked to identify 10 commuters from various age, gender, profession, and usual transport mode groups who had not taken the survey before. With their contact details, the testers were contacted by NTUA with instructions on how to complete the TBS.

Step B: Execution of the travel behaviour module

Following the collection of testers' responses, the mode choice models in the travel behavior module were launched, and their predictions about the factors affecting mode choices were converted into simple, non-technical language. The text communicated to the testers is provided below.

One of the key advantages of the models that we have created is that they are able to reveal the factors that affect commuter's travel choices. Specifically, we have found that:

- *The probability to use a private car over public transport is higher among younger commuters and among commuters with higher incomes.*
- *Increased satisfaction with the public transit system correlates with a higher likelihood of choosing public transportation over a private car.*
- *Rising costs and extended travel times, whether for cars or public transport, diminish the appeal of these modes, leading to a shift towards alternative travel options.*

Step C: Feedback Survey

As an immediate follow-up to these tasks, users were asked to complete a second feedback survey to assess the clarity, interpretability, and intuitiveness of the questions, the time and effort necessary to complete the TBS, and opinions on whether the modeling insights are deemed logical and in line with the testers' expectations.

The survey contained 10 questions, with an estimated completion time of 5 minutes. To ensure honesty in the responses, the survey was anonymous. With the contact details from Phase A, NTUA directly disseminated the feedback survey from 14/02/2024 to 19/02/2024.

Tester feedback was collected using closed questions with predetermined, multiple-choice or scale-based answers. The first question of the questionnaire is provided as an example below.

On a scale of 1 to 5, how easy was it to understand the purpose of the questionnaire?

For specific questions requiring in-depth, qualitative insights that closed-ended questions could not capture, respondents were requested to provide their answers in a free text format to an open-ended question, as illustrated below.

Can you point out some of the questions that you found confusing or ambiguous?

The complete questionnaire is included in the Annex.

Step D: Reporting on the results

The feedback collected during the testing is presented in the "Results" section that follows. Some results have already been presented internally to the consortium during the WP7 bi-weekly meetings dedicated to subsystem testing and the consortium's monthly audio meetings.

3.2.2 Actions achieved and results

A total of 42 commuters from all case studies have been contacted by case study leaders and NTUA, out of which 11 have responded to the feedback survey.

Their responses are pooled together, translated into English and analysed below on a question-by-question basis.

Initially, respondents were asked to evaluate how easy it was to understand the purpose of the questionnaire (Figure 8), rating their experience on a scale from 1 (not clear at all) to 5 (very clear). The majority (81.9%) evaluated their understanding as either 4 or 5, suggesting the effectiveness of the introductory materials and instructions provided with the questionnaire.

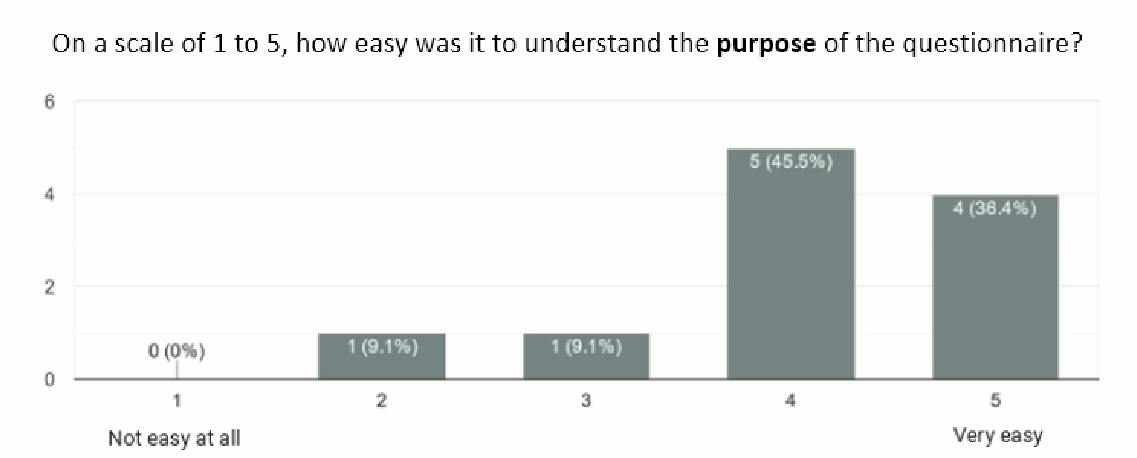


Figure 8: On a scale of 1 to 5, how easy was it to understand the purpose of the questionnaire?

This good understanding of the overall questionnaire can also be attributed to its questions being clear and intuitive. In the relevant question (Figure 9), where an evaluation of 1 suggests that questions were not clear at all and an evaluation of 5 suggests that questions were very clear, the majority of respondents (63.7%) responded with 4 or 5, while respondents (27.3%) found the questions to be moderately clear by selecting "3".

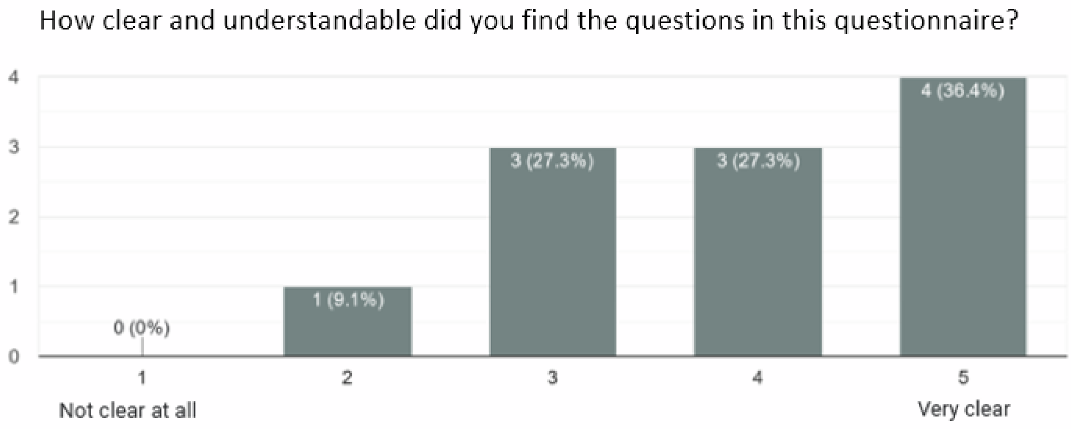


Figure 9: How clear and understandable did you find the questions in this questionnaire?

Two comments were also received in response to the open-ended question on how the overall clarity of the questionnaire can be improved. One respondent suggested the inclusion of a working-from-home option, and a second one suggested that the scale of some questions was somewhat unclear. This feedback can be attributed to Google Forms' practice of providing linguistic interpretations only for the extreme values (1 and 5), leaving the middle scale values as numbers, which can result in some ambiguity.

On a follow-up question of whether the TBS questions were in-context, the vast majority of respondents (81.8%) found all the questions to be relevant, suggesting that the survey was contextually appropriate.

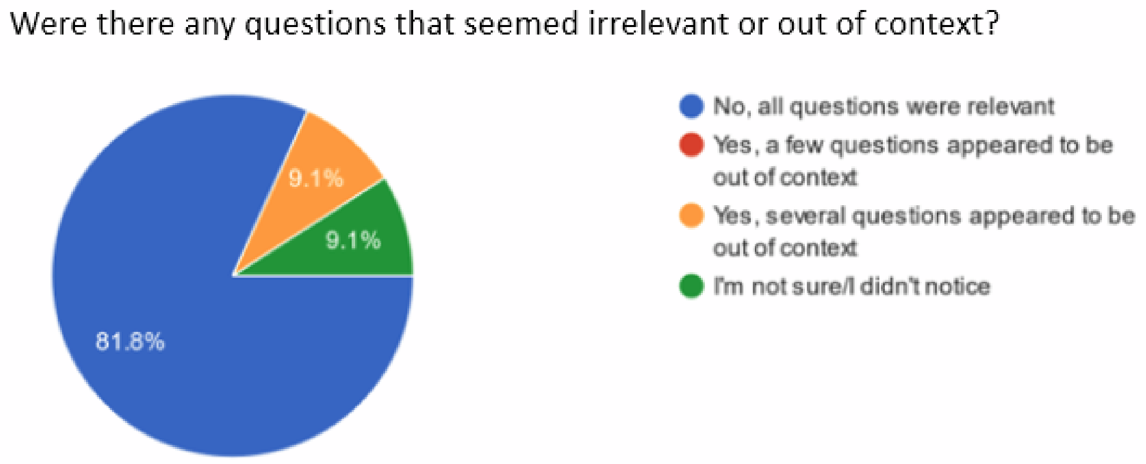


Figure 10: Were there any questions that seemed irrelevant or out of context?

On the overall layout and formatting of the questionnaire (Figure 11), the majority of respondents (54.6%) rated it as either 4 or 5 (good or excellent). Another 36.4% rated it as satisfactory.

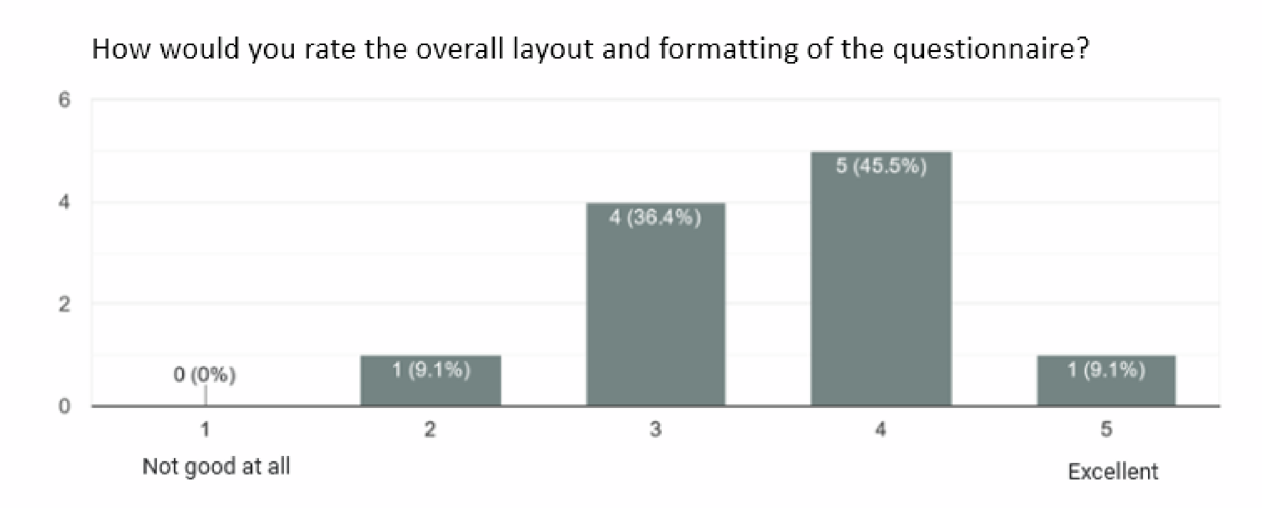


Figure 11: How would you rate the overall layout and formatting of the questionnaire?

Regarding the length and completion time of the questionnaire (Figure 12), most respondents perceived the survey as either appropriate or somewhat lengthy (chart on the left). While this feedback indicates a potential need to shorten the survey to enhance participant satisfaction, the suggestion must be weighed against the benefit that a longer questionnaire allows for the elicitation of more detailed information concerning commuters' mode choice decisions. Considering that 91% of respondents completed the survey within the estimated 15-minute timeframe deemed necessary by the study team (chart on the right), it can be concluded that the completion time was appropriate for the questionnaire's length.

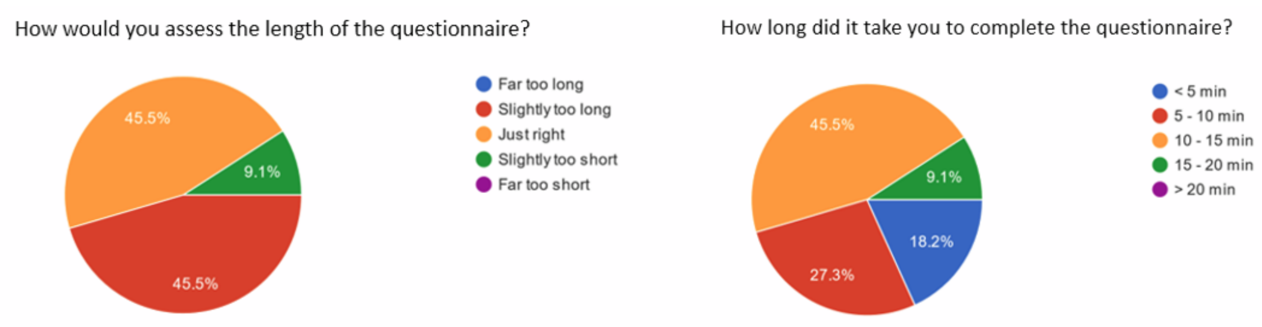


Figure 12: How would you rate the length and completion time of the questionnaire?

Finally, two respondents emphasised the need to revisit and edit previous answers in future questionnaires response to a general open-ended question that elicits overall comments or suggestions on the content and format of the questionnaire. They highlighted that once a page of responses was submitted, it was not possible to return and make corrections.

Modeling insights

In the concluding section of the questionnaire, participants assessed the factors influencing travel choice decisions, as presented in Section B. The overall feedback (Figure 13) indicates that most respondents found the results logical and consistent with their expectations regarding transportation choices, with 63.7% rating their agreement as 4 or 5 (indicating complete agreement with the analysis). Three respondents (27.3%) rated their agreement as 3, suggesting neutrality while one respondent expressed total disagreement but did not provide an explanation in the follow-up question.

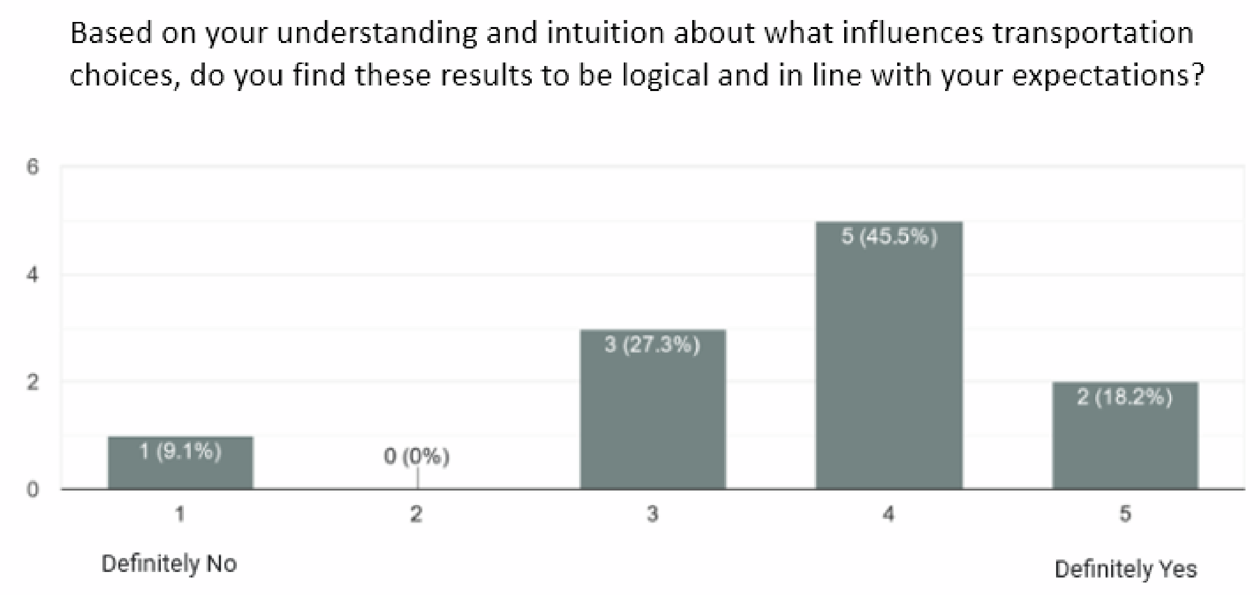


Figure 13: Based on your understanding and intuition about what influences transportation choices, do you find these results to be logical and in line with your expectations?

Concluding, the open-ended question on whether there are additional factors that significantly influence travel choice decisions and could be examined in future models received 5 comments. These comments mentioned the frequency, adherence, reliability, accessibility, and comfort of public transport, as well as the ability to work from home.

3.3 Supply Forecasting & Critical conditions (WP4)

In WP4 deep learning methods were developed to forecast traffic supply. The forecasting methods are applied to the four case studies of the TANGENT project. Results are considered per case study, and are described via different steps, including data preprocessing, training the models, determining forecasting accuracy, etc. in Deliverable 4.2 “Overview of the developed traffic supply forecasting approaches with benchmarks”. The developed anomaly detection algorithm is included in Deliverable 4.4 “Report on the detection and impact analysis of traffic events”, where the results were shown for the case studies of Greater Manchester and Athens. The applied algorithm is similar for Lisbon.

The dataset for the case study of Lisbon that was included in D4.2 contains 8 loop detectors on highways, for which an entire year of data was made available (2023). After submission of D4.2 an extra supply forecasting model for public transport for Lisbon was built, namely a bus travel time prediction algorithm.

For Lisbon bus travel time prediction, the preprocessing steps described in D4.2 are applied. The average speed and travel time for each grid are calculated and are used as input to the first model that makes grid-level travel time predictions. The output of the first model is then fed to a machine-learning pipeline to predict line-level travel times. The description in D4.2 (for Rennes bus travel time prediction) was based on an extraction of the grids that need to be passed from origin to destination, which were aggregated manually. This procedure was extended to a data-driven procedure based on linear regression. A single model is trained for each line of the case study. We couldn't reliably extract the travel times of individual buses in this case study, and therefore used linear regression. It can work with limited data compared to more complex deep learning models. Cyclical time features are concatenated to the grid journey times to provide better context to the model for better accuracy.

3.3.1 Methodology

Both supply forecasting and anomaly detection were tested by the case study leaders. The testing methodology is described in this section.

Predictive model building:

- Build models for the different datasets of Lisbon for supply forecasting, as described in D4.2 and D4.4
- Collection and analysis of the results in maps
- Share results and main conclusions with case study leaders

Feedback:

- The case study leaders provide feedback based on the maps
- In the second stage this feedback can be based on the live dashboard
- Meeting with case study leaders to discuss the feedback

Details regarding the steps above are included below:

Before the supply forecasting and anomaly detection results were added to the live dashboard, the results were shared for testing purposes with the case study leaders. Folders with maps were shared, together with guidelines on how to provide feedback related to the testing.

Supply forecasting
Access to results
After unzipping the files from the testing folder, you'll find an HTML file alongside a "timeseries" folder The HTML file serves as a map, pinpointing the sensors used in the dataset. Clicking on the blue markers will display the sensor's name and prediction error for various horizons. Each prediction horizon represents a 5-minute interval. For instance, horizon 1 corresponds to a 5-minute prediction, while horizon 12 equates to a 1-hour predictions.This data encompasses all sensors and prediction horizons.

The folder contains interactive HTML files, allowing you to zoom in and out and pan across the time series. Missing values are represented by 0 in the ground truth. For the LIS_CAR_01 SPEED dataset, the average Mean Absolute Error (MAE) is 4.065, and the LIS_CAR_01 COUNT dataset has an average MAE of 20.054 cars per 5 minutes. You can also find the individual MAEs for each sensor detailed in the charts and the map file.
Procedure for testing & providing feedback
Then please follow the following procedure: 1. Pick days, times, and locations that interest you. 2. Act as if it's real-time that moment 3. Look at predictions for that time for the upcoming minutes. 4. Check the actual values. Do they correspond? Is this what you expected?

Similarly, results related to anomaly detection were shared.

Anomaly detection
Access to results
After unzipping the files from the testing folder, you'll find a "timeseries" folder. Inside this folder, you can access the time series data for horizon 1 predictions (red charts), their corresponding ground truths (blue charts), normalized anomaly scores (in brown), and detected anomalies in orange vertical lines. Lower anomaly scores represent a higher chance of anomalies. Missing values are represented by 0 in the ground truth. You can use the same map provided for supply forecasting to extract the location of each sensor.
Procedure for testing & providing feedback
Please structure your response in the rows below: 1. Pick days, times, and locations that interest you. 2. Act as if it's real-time that moment 3. Look at detected anomalies for that time for the upcoming minutes.

4. Is this what you expected?

5. Specifically, are there anomalies the model couldn't detect (False Negatives)?

6. Specifically, are there any instances in which the model detected an anomaly, but those are not anomalies (False Positives)?

7. How do you look at these predictions? Do you wait for anomalies to be detected and start traffic management from there, or do you already adjust traffic based on the supply predictions?

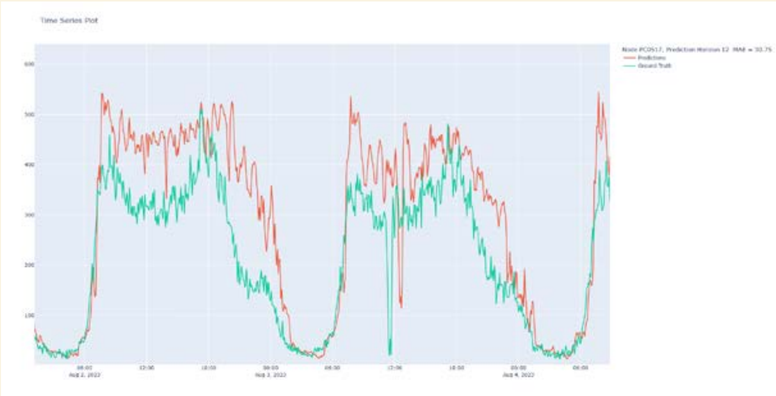
Date and time	Locations (Sensor_id)	Output_of_the_algorithm	Desired_result (Qs 4-5-6)	Notes

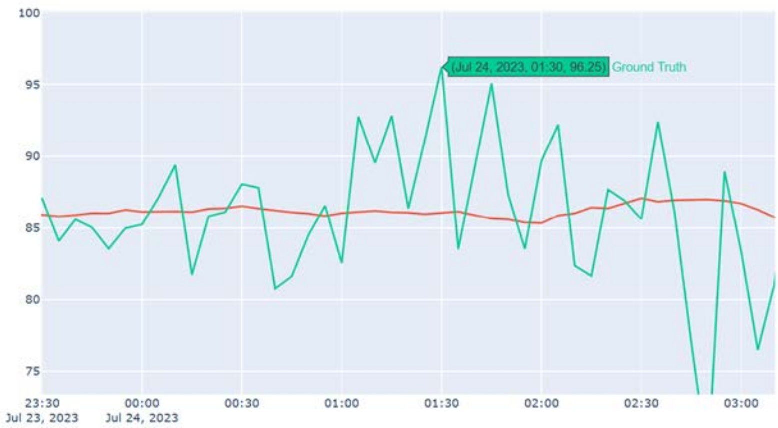
The technical leaders and case study leaders decided to discuss the feedback during a meeting (details described in the following section).

3.3.2 Actions achieved and results

Feedback on the provided results were handled in a meeting between the technical and case study leaders. The global results, resulting from this discussion, are detailed below:

Feedback case study leaders	Reply technical partner IMEC
Traffic count forecasting • Quite accurate for smaller forecasting periods (5-10-20 min), both in typical conditions and during abnormal situations • Looking at longer range (60 min) the prediction seems to “not be able to keep up” with the higher amount of cars	• For training purposes, we used a dataset from 2022-01-01 00:00:00 to 2023-09-30 23:55:00, the first 70% for training and the last 20% for testing. The differences for different months contribute to this, as shown by the mean counts in the train set compared to the ones in the test set. ○ Mean train_data_count: 194.813, Std train_data_count: 139.264 ○ Mean test_data_count: 212.973, Std test_data_count: 144.901 • Research wise this shows the need for retraining and making sure the train set contains the same patterns as the test set, and the real-time application. • Within imec’s research group continual learning is a research topic that is considered to continuously learn on new data, by accumulating knowledge. This track will need more research before it can be applied in real situations such as traffic forecasting.
Comparison traffic count and speed forecasting • The algorithm seems to underestimate the number of vehicles but overestimate speed reductions; in other words, from a congestion evaluation point of view: the	• Similar reasoning as in traffic count forecasting: training period compared to test period • It might be that during the training period, the average speed is higher and the number of cars are lower, also smoothing property of the model ○ Mean train_data_count: 194.813, Std train_data_count: 139.264

Feedback case study leaders	Reply technical partner IMEC
vehicle count prediction algorithm is optimistic, and the speed prediction algorithm is pessimistic.	○ Mean test_data_count: 212.973, Std test_data_count: 144.901 ○ Mean train_data_speed: 83.850, Std train_data_speed: 15.941 ○ Mean test_data_speed: 83.384, Std test_data_speed: 16.341
Traffic count forecasting: event ● Higher amount of car traffic in 2-6 of august - there was effectively a great affluence to the city of Lisbon, due to the World Youth Days	Count ● Exceptional event, Pope visited the country ● No need to adjust the forecasting methodology for such an event ● Research wise it might be interesting to look at the idea of ‘Mixture of experts’, technique where multiple expert networks are used to divide a problem space into regions (normal – event)  Speed 
Traffic speed forecasting	● Most of these errors are unpredictable driver behavior in free flow traffic.

Feedback case study leaders	Reply technical partner IMEC
• Reasonably accurate in predicting travel speeds, but we noticed sometimes it overestimates speed losses and takes longer to correct the prediction to match the ground truth. • As with the vehicle counts, this phenomenon gets pronounced with longer prediction periods; but start displaying this "error" earlier (e.g. from 15 min on)	• Most of the time, there's no trend to capture. Just random fluctuations. 

Feedback case study leaders	Reply technical partner IMEC
Anomaly detection: discussion on methodology used	

- Only the first timestamp is used, several scores are calculated considering the expected error for each sensor.
- Some anomalies are not being detected. The reason is the direction of the anomaly which is more intuitive for traffic speed where we only report anomalies when the predicted speed is lower than the observed one. However, the same method might not be useful for traffic count where both directions might be relevant.
- The discussion resulted in a few ideas:
  - Which metrics are of most interest? Occupancy might be more relevant.
  - For Period P1 between date/time D1 and D2, B(t) as Busy Time for each transaction we calculate Occupancy O as: O(P1) = SUM(t) (B(t)) / (D2 -D1)
  - Is anomaly detection required on the different traffic variables (speed, count, and even journey times for TfGM)? It might be to have one overall measure which reflects the reality in the best way possible.

Table 1: Conclusions from discussion between technical partner (imec) and case study leaders on supply forecasting and anomaly detection

In conclusion, results are good to use as is. Further improvements are listed that can be included as future research directions:

- Change train set to make it more comparable to the test set (retraining)
- Test mixture of experts to separate normal and event days
- Train on occupancy/test on one combined measure used for anomaly detection

In the remainder of the project, we'll check what we can still test/improve within the project. Other elements can be included in future research.

3.4 Demand Estimation & Prediction (WP4)

There was no subsystem to test for Lisbon related to the Demand Estimation and Prediction, due to:

- the delay in acquiring access to the calibrated transportation model (as a consequence of administrative issues regarding signing of NDA with an external entity for the availability of demand data)
- and the macro-level scope of the model, which is inadequate to perform the necessary procedures for Demand Estimation and Predictions

3.5 Synchronization of Public Transport and traffic control (WP5)

There was no subsystem to test for Lisbon related to the Synchronization of Public Transport and Traffic Control, due to:

- the delay in acquiring access to the calibrated transportation model (as a consequence of administrative issues regarding signing of NDA with an external entity for the availability of demand data)
- and the macro-level scope of the model, which is inadequate to perform the necessary optimization experiments

3.6 Consensus Mechanism (WP5)

The goal of Task T5.4, is to develop a comprehensive, automated decision-support tool designed to aid decision-making and foster consensus among stakeholders with divergent opinions. Drawing from actual stakeholder preferences and cooperation patterns obtained through structured surveys, our work, detailed in D5.4 proposes a framework that realistically models the key stages of the consensus-reaching process, until the reach of consensus between stakeholders.

For testing the modelling framework, the same four-step methodology described in section 2.2.1 used for testing the WP3 module, was followed. The adaptation of this methodology to the specificities of the consensus mechanism is outlined below:

3.6.1 Methodology

Step A: Testing sample collection

During the initial phase of testing, the stakeholder cooperation survey (SCS) was distributed to stakeholders from each case study city, including Lisbon. The SCS questionnaire is included in the Annex of D5.4. To collect the necessary sample, each case study was asked to identify 3 stakeholders. Unlike WP3, case study contact persons directly provided the survey material to stakeholders and collected the completed questionnaires, which were then forwarded to NTUA.

Step B: Execution of the consensus mechanism

The consensus mechanism is applied and its output, in terms of consensual opinions regarding stakeholder priorities when a particular traffic management strategy is implemented, is communicated to participants.

Given that not all strategies are relevant to all stakeholders (since the entity they represent might not be involved in the decision-making process for one or more TMS), the opinions communicated were individualized and determined by their responses to the TBS in Step A. For example, the material communicated to the testers who responded that DRT-PT is relevant for their city and organizations, are provided below (Figure 14):

Section 3 of 5

DRT-PT

During the testing phase of the mechanism, we have engaged various stakeholders like you and determined that the collective opinion regarding the different network objectives for the DRT-PT (Synchronization of on-demand and transit modes) traffic management strategy is as follows:

Collective Stakeholder Preferences

Objective A is more important than Objective B					Objectives are Equal	Objective B is more important than Objective A				
Objective A	Strongly		Moderately				Moderately		Strongly	Objective B
Service performance	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1	☐ 2	☑ 3	☐ 4	☐ 5	Network efficiency
Economic performance	☐ 5	☐ 4	☐ 3	☐ 2	☑ 1	☐ 2	☐ 3	☐ 4	☐ 5	Service performance
Service performance	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1	☐ 2	☑ 3	☐ 4	☐ 5	Environment
Network efficiency	☐ 5	☑ 4	☐ 3	☐ 2	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Economic performance
Environment	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1	☐ 2	☐ 3	☑ 4	☐ 5	Network efficiency
Economic performance	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1	☐ 2	☑ 3	☐ 4	☐ 5	Environment

Strategy Description

As a reminder, the objective of this traffic management strategy is the joint optimisation of the running route and the scheduling (departure times) of Responsive Feeder Transit and Responsive Collector Transit systems, to facilitate access and egress connection to public transit services (first/last mile problem).

Service performance: Total passengers served
Network efficiency: Passenger travel time
Economic performance: Operational costs of the DRT fleet
Environment: Emissions of DRT fleet

Figure 14: DRT-PT: Communication of opinion modelling results

Step C: Feedback Survey

As with the testing of WP3 modules, stakeholders were contacted again by NTUA to complete a follow up feedback survey to assess both the questionnaire and the results of the opinion modelling. The distribution of the survey occurred from 19/02/2024 to 28/02/2024. The complete feedback questionnaire is included in Annex II.

Step D: Reporting on the results

The feedback collected during the testing is presented in the "Results" section that follows. Some interim results have already been presented internally to the consortium during the WP7 bi-weekly meetings dedicated to subsystem testing and the consortium's monthly audio meetings.

3.6.2 Actions achieved and results

A total of 12 stakeholders from all case studies, including 3 from the Lisbon case study, were contacted by case studies and NTUA. Six stakeholders responded to the feedback survey, and their responses are pooled together and analysed below.

Stakeholders found the questions of the survey to be rather straightforward and understandable (Figure 15).

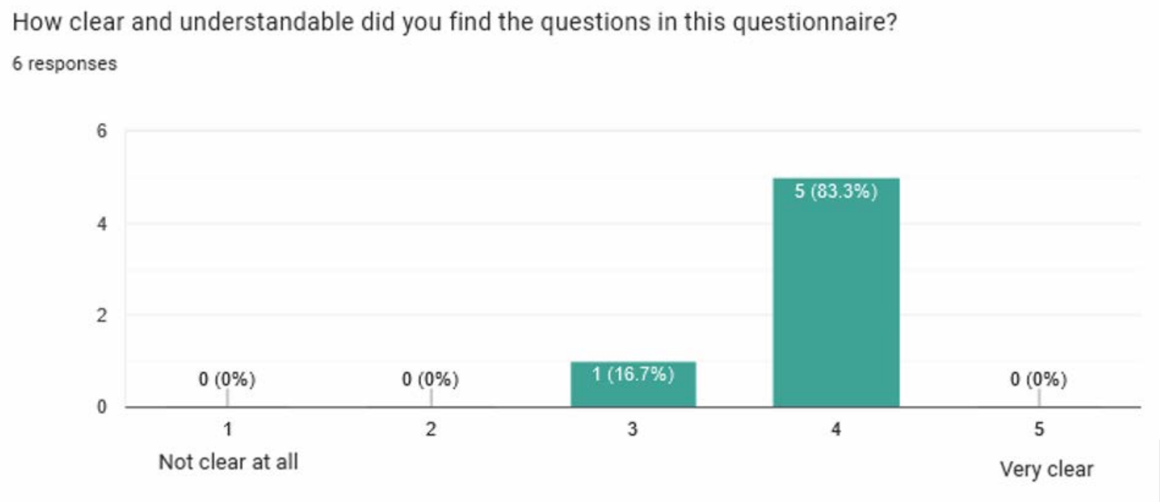


Figure 15: How clear and understandable did you find the questions in this questionnaire?

Perhaps unexpectedly then, 50% of respondents considered that a few questions appeared to be out of context (Figure 16). This issue is attributed to two factors: first, the survey elicited both stakeholder preferences and the institutional relationships between the entities they represent. Although related, these themes are distinct, which is uncommon in similar surveys that are highly focused. Additionally, the indicators used for eliciting relationships, such as the frequency of data exchanges between entities, do not directly reveal their relevance to the overall philosophy of the questionnaire.

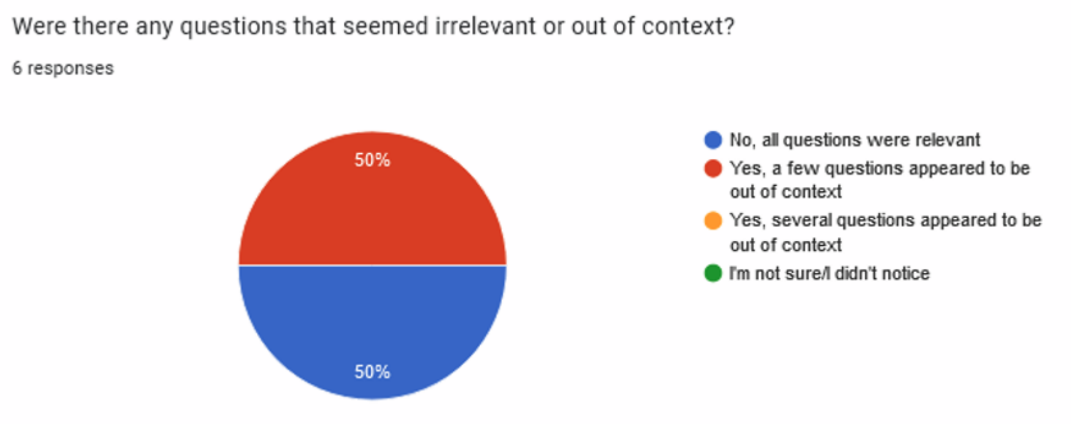


Figure 16: Were there any questions that seemed irrelevant or out of context?

Nevertheless, the instructions provided with the questionnaire were deemed adequate by participants (Figure 17).

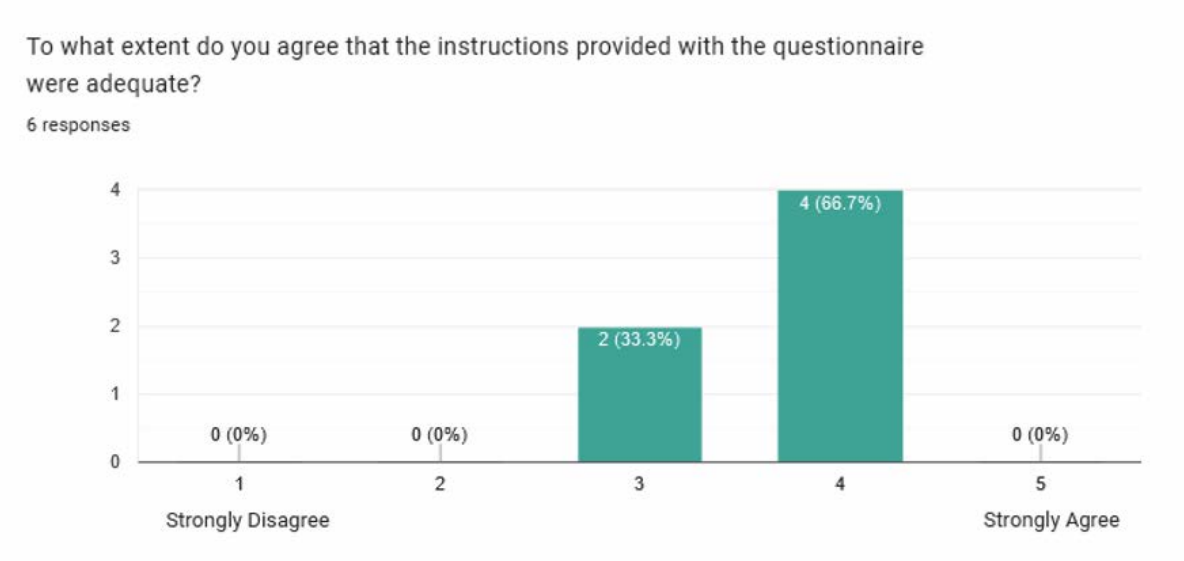


Figure 17: To what extent do you agree that the instructions provided with the questionnaire were adequate?

Opinions were more mixed regarding the overall formatting of the questionnaire. Unlike traditional question design (multiple-choice, open-ended, etc.), the need to provide institutional relationship information by filling a matrix in an Excel format (Figure 18) requires more cognitive effort and may not be immediately intuitive.

Stakeholder Category	Stakeholder Name	**B1**. Communication frequency	**B2**. Integration	**B3**. Data
Government & Authorities				
European Union				
Ministry of transport	Department for Transport	4. 1-2 times/year	2. Partnership: We work together as a formal team with specified responsibilities to achieve common program goals (note: responsibility for each organization is usually outlined in a Memorandum of Understanding or other agreement).	4. Degree 4 – Seamless sharing of information: Universal interpretation of information through data processing based on cooperating applications
Other national ministry				
Regional government				
Local authority				
Neighbouring city	Local councils surrounding Greater Manchester and those across (mainly) North of England	4. 1-2 times/year	3. Collaboration: We work side-by-side and actively pursue opportunities to work together as an informal team (i.e., do not establish a formal agreement; work together "in the spirit of collaboration").	1. Degree 1 – Unstructured data exchange: Exchange of human-interpretable unstructured data such as the text found in operational estimates, analyses and papers
Local transport authority				

Figure 18: Stakeholder relations matrix

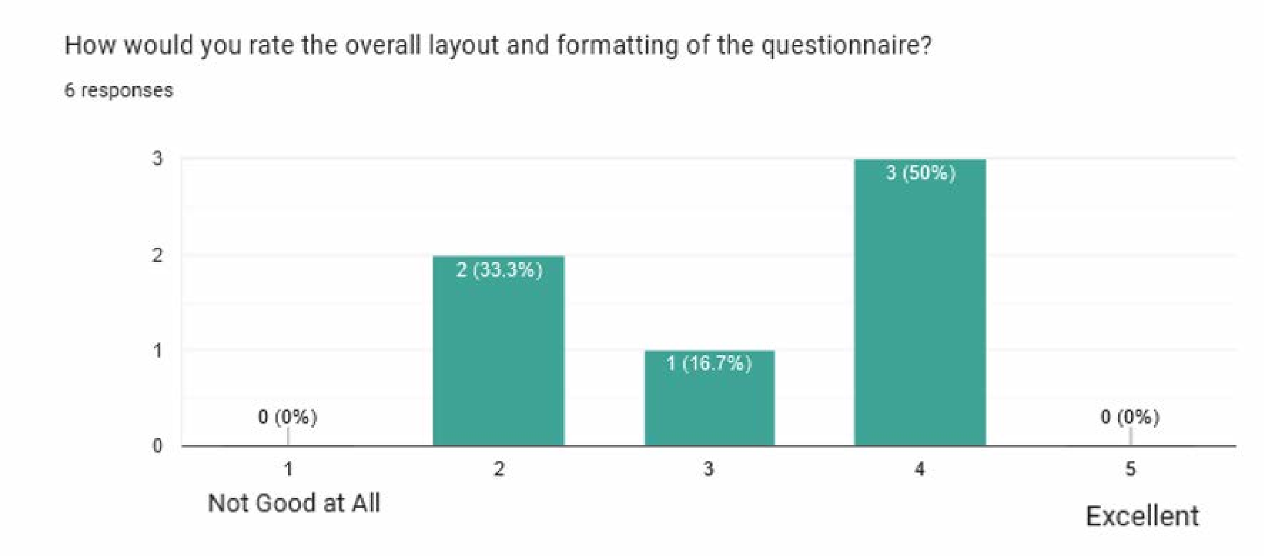


Figure 19: How would you rate the overall layout and formatting of the questionnaire?

The length of the questionnaire was also deemed excessive (Figure 20), with 67% of respondents finding it too long. This result was anticipated during internal test runs of the survey. The decision to proceed was based on the following: this survey targeted the few stakeholders knowledgeable about the decision-making, cooperation, and communication patterns among stakeholders involved in the transport network's management. For instance, the 14 questionnaires collected in the Stakeholder Cooperation Survey contrast with the 1859 responses across the four case studies for the TBS, aimed at the general population. To compensate for the few respondents, we opted to demand more cognitive effort from participating stakeholders, facilitated by case study leaders who primarily distributed the questionnaire within their professional circles. Thus, the 14 questionnaires documented 248 cooperation instances and 46 opinion matrices, numbers suitable for the opinion modelling. This cognitive load was significant for some respondents, as noted by two respondents in open-ended questions on general comments or suggestions on the questionnaire content and format.

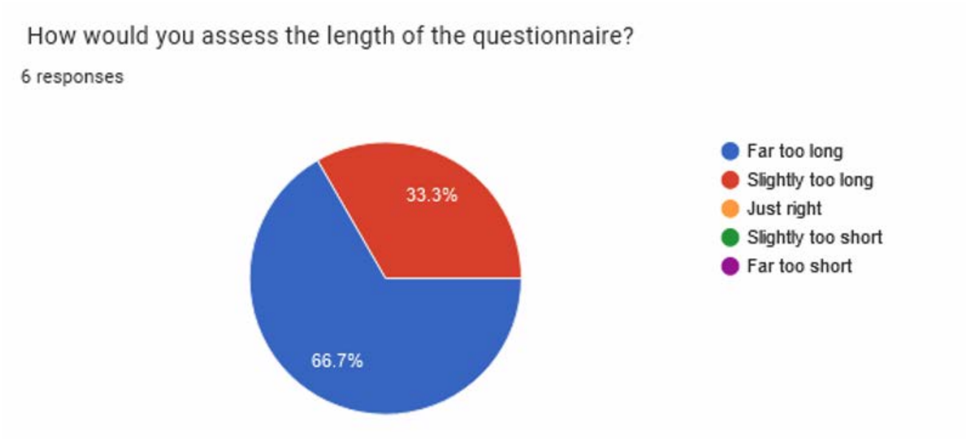


Figure 20: How would you assess the length of the questionnaire?

This cognitive load is also reflected in the fact that most (50%) participants completed the questionnaire in 25 to 30 minutes (Figure 21). For reference, the survey was initially timed at 25 minutes, a duration that most participants did not significantly exceed.

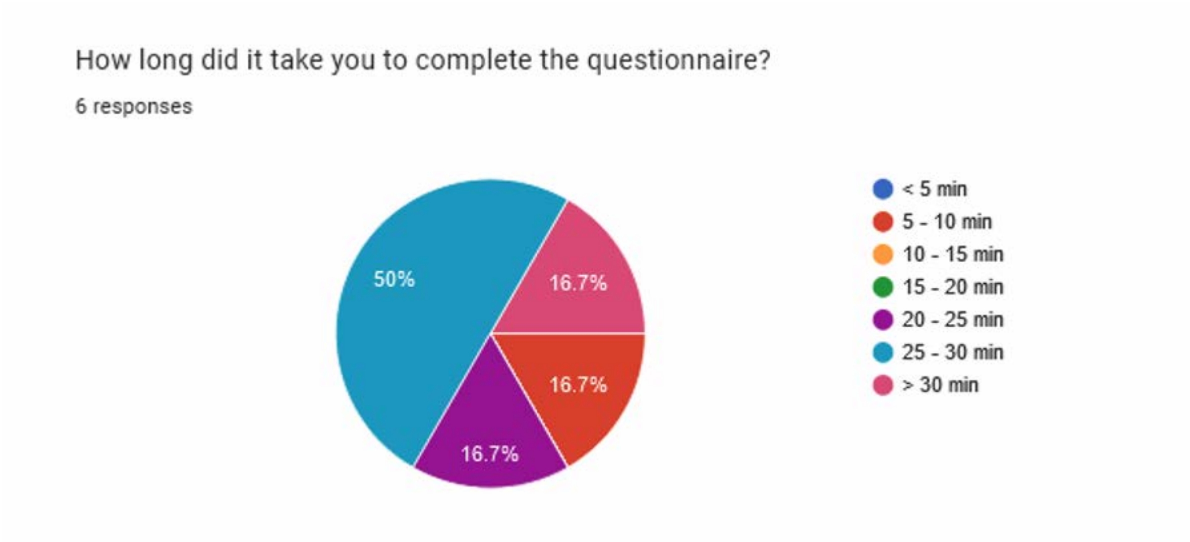


Figure 21: How long did it take you to complete the questionnaire?

Modelling insights

As discussed in D5.4, the opinion modification of the consensual preferences has two main objectives. First, the original opinions of the stakeholders must be preserved to the greatest extent possible, ensuring that each stakeholder's unique perspectives and insights are respected and incorporated into the decision-making process. Second, the consensual collective preference must be accepted by all stakeholders, as this sense of ownership enhances the legitimacy of the decision-making process and ensures that the policies (in this case, traffic management strategies) are more likely to be effectively implemented.

The consensus mechanism successfully achieves both objectives, as stakeholders have explicitly accepted the legitimacy of the consensual opinion Figure 22) and feel that the consensual opinion is not significantly divergent from their own (Figure 23). This result serves as a robust indicator of the consensus mechanism's capacity to determine the true consensual opinion.

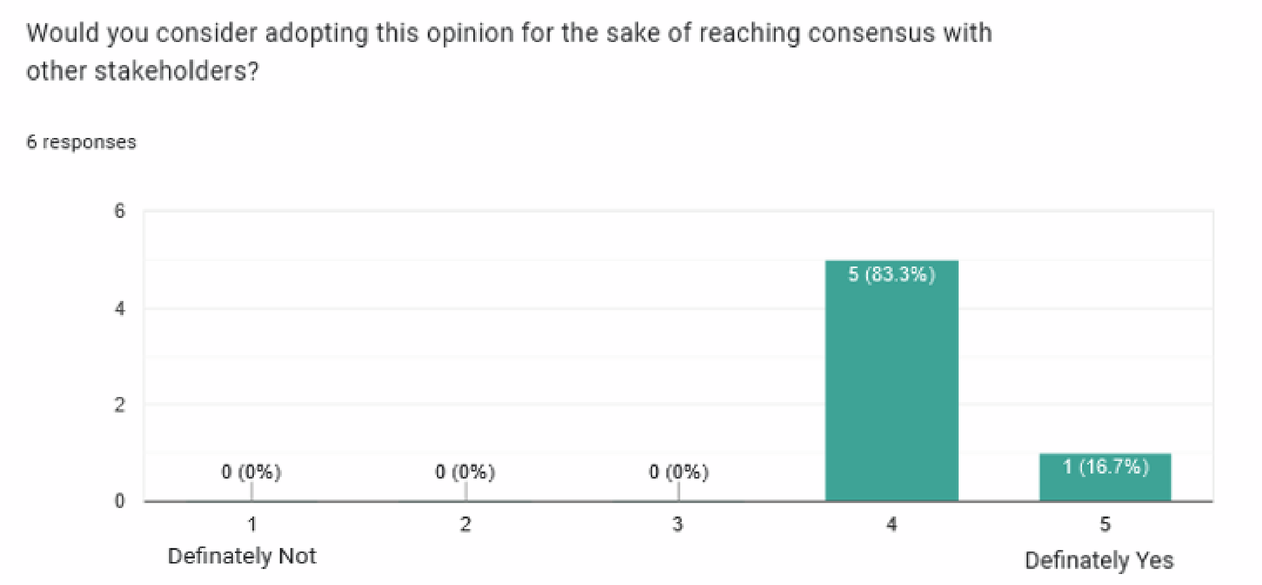


Figure 22: Would you consider adopting this opinion for the sake of reaching consensus with other stakeholders?

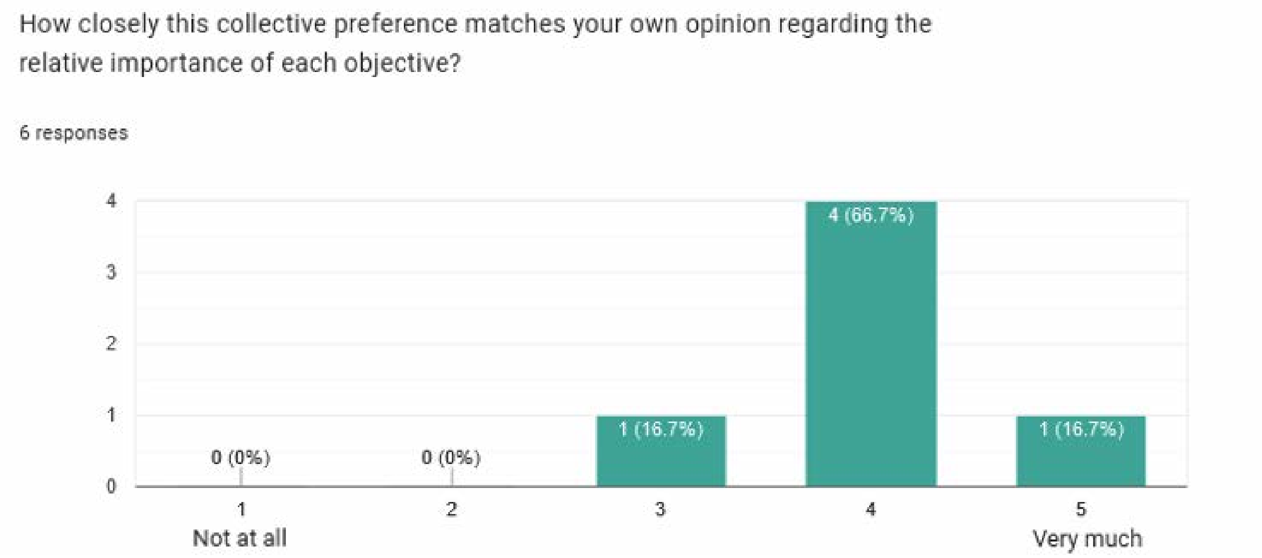


Figure 23: How closely this collective preference matches your own opinion regarding the relative importance of each objective?

Ultimately, stakeholders exhibit general trust in the consensus mechanism's operation, as depicted by the responses to the corresponding question (Figure 24).

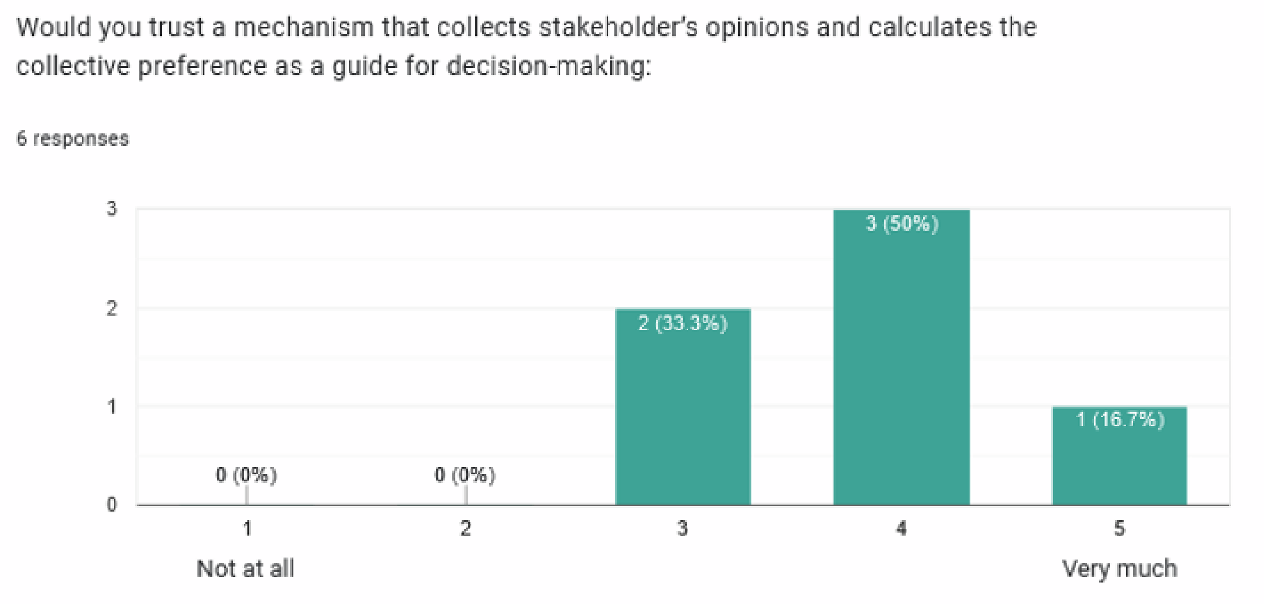


Figure 24: Would you trust a mechanism that collects stakeholder's opinions and calculates the collective preference as a guide for decision-making?

4 Testing of the overall system

The testing activities of the overall system aim to assess the functionalities of the 3 services of the TANGENT tool. They are related to the WP7, more specifically to T7.3 “Upscaling and full operation: testing of the overall system”. The test phase of the overall system is divided in two subtasks:

- ***"Subtask 7.3.1 Technical upscaling and full operation of TANGENT tools"***: Technical tests to assess the proper integration of the subsystems provided by the technical partners and developed under WP2, WP3, WP4 and WP5.
- ***"Subtask 7.3.2 Non-technical testing and business models’ evaluation"***: Performance, usability, and acceptability testing to be performed by the potential end-users of the tools: the case studies.

The test period is also framed by the WP6: Task 6.4 “Integration and development of TANGENT decision-making support tool for coordinated traffic and transport” (M16-M37) and deliverables D6.2 “TANGENT tool development and integration - 1st release” (M32) and D6.5 “TANGENT tool development and integration - 2nd release” (M37).

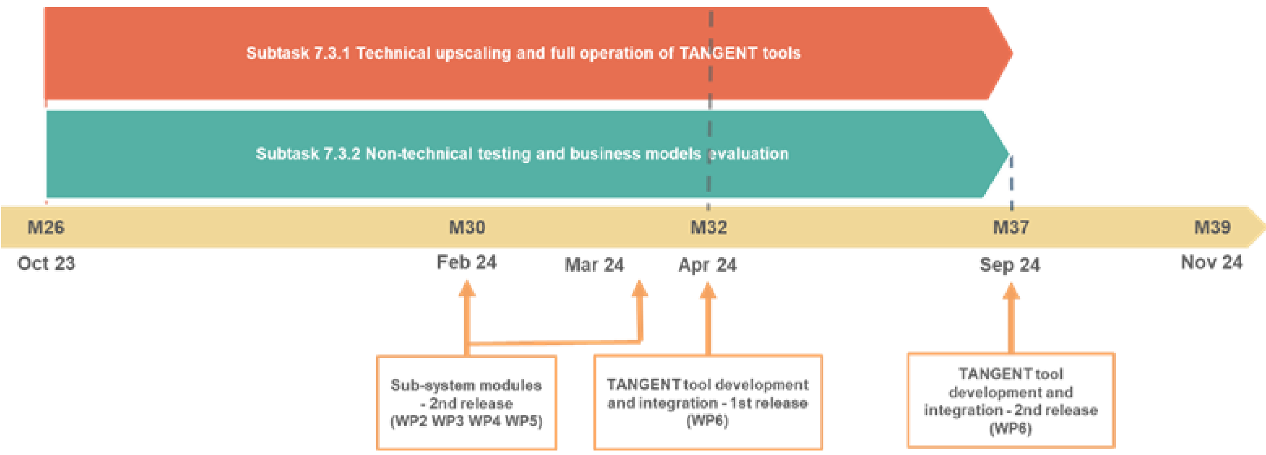


Figure 25: Testing timeline

The above figure represents the global timeline of the overall system testing. The top section represents the execution of Subtask 7.3.1, where the technical testing of the overall tool’s integration has been done progressively, in strict collaboration between the WP6 leader and the technical partner responsible for the subsystem under integration. The bottom section represents the execution of Subtask 7.3.2.

The two T7.3 subtasks are closely linked, and are fully interconnected with WP6 work on subsystem integration (*T6.4 “Integration and development of TANGENT decision-making support tool for coordinated traffic and transport management”*). WP6 itself is dependent on the progress of technical partners in the development of the subsystems. Therefore, the non-technical testing of each functionality by the case studies could only start when its development and integration were completed. The tests were therefore sequential.

During the first months of the project, CARRIS and A-to-Be, in collaboration with WP1 and WP7 partners, defined the Lisbon case study and the services to be implemented and tested under various scenarios, according to the city's challenges and interests. The services and functionalities related to the Lisbon case study are listed in the following table (Table 2).

Service	Functionality	Scenario	Real-time or Simulation/Virtually
Service 1: Enhanced information service for multimodal transport management	Current state of the network	Daily commuting Planned Event Unplanned Event	On-site
	Future state of the network	Daily commuting Planned Event Unplanned Event	On-site
Service 2: Real Time Traffic Management Services	Synchronization of Public Transport and Traffic Control	N/A	N/A
	Informing the Transport Passengers	Daily commuting Planned Event Unplanned Event	On-site
Service 3: Simulation of pre-defined scenarios	Simulation of the following scenarios: 1. Increased Bus and Tram supply 2. Simulation of the following scenario: Implementation of congestion pricing	N/A	Virtually

Table 2: Testing environment for each functionality

4.1 Methodology

4.1.1 Technical testing

The technical testing activities were tackled between the technical partners who developed the different subsystems and A-to-Be, responsible for the integration of the functionalities and the development of the TANGENT global tool.

Therefore, the technical tests have been performed in the scope of T6.4 “Integration and development of TANGENT decision-making support tool for coordinated traffic and transport management”.

4.1.2 Non-technical testing

From April to September 2024, several tests were run to assess the different functionalities which added progressively to the TANGENT integrated tool. The goal was to continuously carry out the testing until the complete tool and all its functionalities were validated before the end of September 2024.

4.1.2.1 Launch of the non-technical testing phase

On March 13th 2024 (M31) during the TMC, the WP6 leader performed a live session launching the TANGENT Dashboard, which has been continuously running since. At this time, the platform featured the first functionality integrated: the current state of the network (Service 1). All partners requesting access were granted credentials to use the platform, according to their assigned roles.

4.1.2.2 Implementation of the non-technical tests

When the technical partners finished developing a functionality adapted for the TANGENT tool specificities, the WP6 integrated it on the Dashboard, and non-technical testing was initiated with the non-virtual case studies, organised in test phases called “Steps”.

For instance, Step 1 was launched on April 3rd until April 29th, in order to test the functionality “Current state of the network” in Daily Commuting scenarios. After April 29th, the functionality remained active within the integrated tool until the end of the project, but the core of its testing was considered finished after that period. This methodology has been iterated for all functionalities until all the services and functionalities were integrated and tested: the whole TANGENT tool was completely tested. In addition, the TANGENT tool has built-in feedback functionality, which has allowed users to provide comments and report issues, extending the testing window of each functionality to the entire usage period.

According to the functionality under assessment, one of the two following types of formats was used to gather feedback:

- A questionnaire was sent to the partners and the local stakeholders that have access to the Dashboard
- A virtual meeting with the partners and the local stakeholders that have access to the Dashboard

In both cases, the testing process required a similar approach, wherein participants were requested to access the TANGENT tool and perform some operations to evaluate various aspects of the platform and the functionality under test.

These tests aimed to assess 3 aspects of the Dashboard:

- **Usability and reliability:** Evaluate the user-friendliness and the clarity of the Dashboard and the functionality
- **Accuracy of the information:** Assess the correctness of the elements displayed, in a qualitative way
- **Compliance with initial requirements:** Review and update the expectations listed in deliverable D7.2 before February 2023

The questions have been iterated for each step of the testing and adapted to the functionality under evaluation, with the support of the relevant technical partner when needed.

In addition to these surveys, the consortium was considering organizing on-site test visits to Rennes, Manchester and Lisbon, between June and September, with the relevant technical partners, particularly for the testing of more complex functionalities for planned and unplanned events, and/or to test the overall platform with all the functionalities integrated. These on-site tests will be decided in the next weeks, once the main functionalities are fully integrated.

While the initial intent was to couple the overall testing with the final workshops led by WP1, in the Lisbon Case Study this did not occur since, at this date, some functionalities (Service 2 - CIM and Service 3) were still pending integration. As the Lisbon stakeholders were closely involved in the testing, it was decided to postpone the overall testing to a later date when new functionalities could be presented to them.

4.2 Tests results

The functionalities have been tested in different scenarios (daily commuting, large planned events and unplanned events). Table 3 summarises the elements tested, and their testing time slots, taking into consideration subsystem development and integration.

STEP	Service	Functionality	Scenario	Non-technical testing dates
1	Service 1	Current state of the network	Daily commuting Planned Event Unplanned Event	April 3rd to April 29th, 2024
2	Service 1	Future state forecasting	Daily commuting Planned Event Unplanned Event	June 14th to September 2nd, 2024
3	Service 2	SNLB: Informing the Transport Passenger[2]	Daily commuting Planned Event Unplanned Event	June 3rd to September 2nd, 2024
4	Service 2/3	Planner features: Synchronisation of Traffic Control & Public Transport[2]	Daily commuting	N/A
5	Service 3	Simulation of predefined scenarios[2]	What-if scenarios	TBD
6	ALL	ALL	ALL	TBD

Table 3: Timeline of Lisbon case study non-technical testing

4.2.1 Step 1: Current state of the network

The first step of the testing aimed at assessing the dashboard of TANGENT and the first functionality of Service 1: **Current state of the network**, during daily commuting.

To this purpose, a questionnaire was sent to the non-virtual case studies, using Google Forms. The questions can be found in Annex I. As this was the first time the partners were asked to review some parts of the TANGENT tool, specific questions about some elements related to the Dashboard (Figure 26) in general, not targeting the functionality assessed, were also asked.

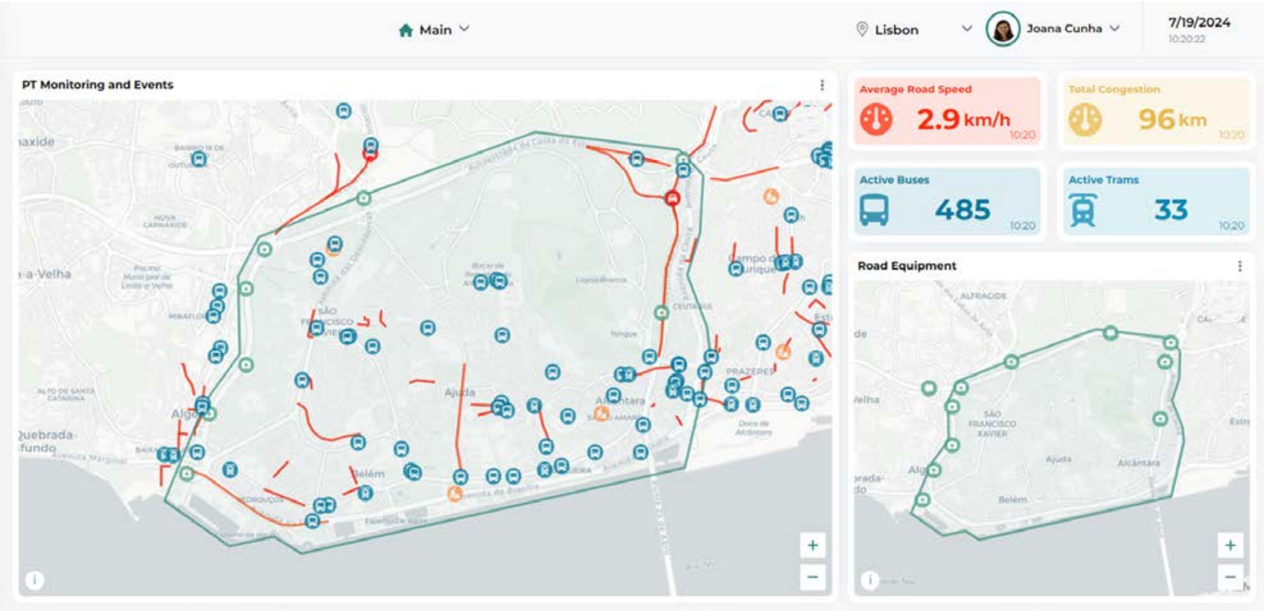


Figure 26: Current State view (without filtering)

Respondents

For this functionality, three respondents sent feedback using the questionnaire:

- 1 Project Manager
- 2 PT Traffic Control Engineers

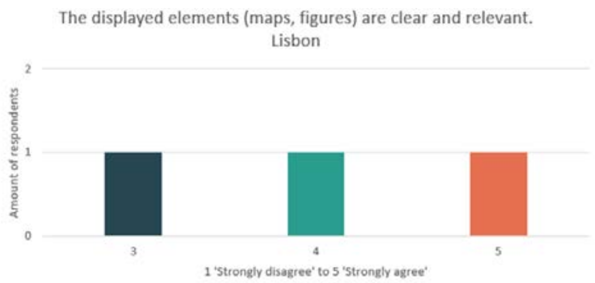
In later contacts, it was possible to garner the feedback from 3 aditional stakeholders. While their responses were not garnered through the questionnaire, their impressions on the functionality were included in the overall analysis presented at the end of this section. The user profile for these respondents is the following:

- 1 City Traffic Manager
- 1 Communication and Dissemination PTO staff
- 1 Customer Care PTO staff

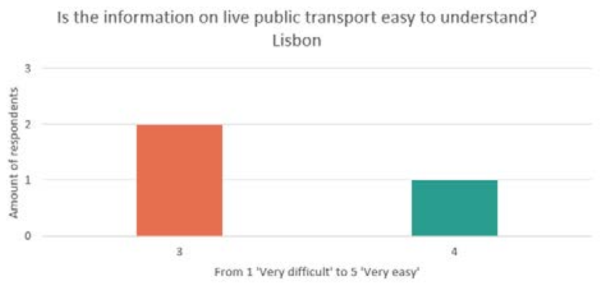
Results

Usability and reliability

Regarding the user-friendliness of the dashboard, 2 respondents assessed that the amount of information on the main page is correct. The 3 respondents disagree on the clarity and relevance of the displayed elements: it is either average, clear or extremely clear.



The information on live traffic and incidents is in majority assessed as easy to understand, but the information on public transport is neutral in terms of difficulty to understand.

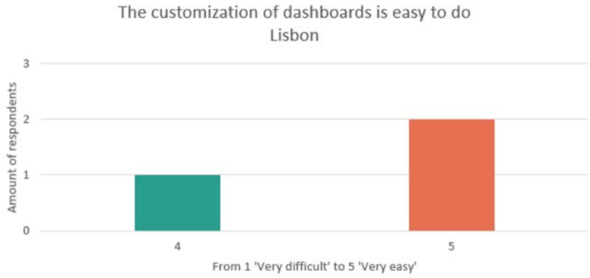


Some comments were left by the respondents:

- Having the Google Maps "base map" display the traffic levels (from Google) would be interesting to help complement the Waze data, but might not be feasible
- It would be interesting to have further details on the PT services. For instance, allowing to filter the displayed lines and displaying the routes.
- The "translation" from the GTFS/GTFS-RT to the map should be reviewed. The lines are being incorrectly identified: the actual line number is concatenated with the vehicle number, and the value displayed in the line is an ID code that is not readable to the final user
- Related to the above comment, the tooltip does not display the destination of the bus service

- The users don't know which way the buses are moving or their route, leaving them unaware of their movement

The customization of the dashboard is assessed as easy or very easy to do.



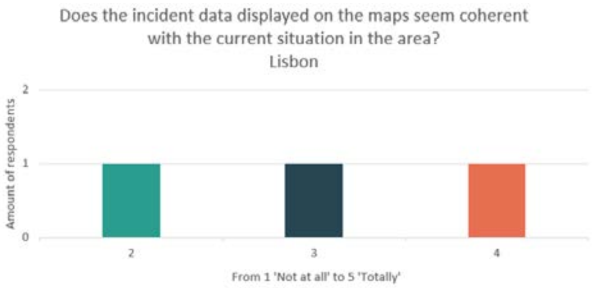
The respondents shared some comments about it:

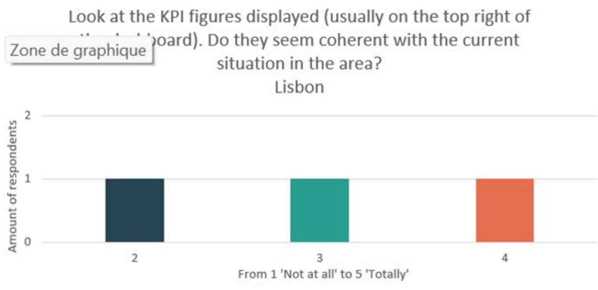
- It would be interesting for the Managers to be able to develop their own indicators (limited to the available datasets, of course);
- Linked to the abovementioned comment, there should be an option to create novel indicators calculated from the existing ones (for instance, having a total number of incidents captured in the original dataset)
- There could be a fixed dashboard, which would be optional to choose: a standard dashboard that is the same for everyone, and impossible to change. In my first contact with the site, I started exploring and then I was never able to post what was there.

For the overall use of the tool, the respondents do not find the tool unnecessarily complex. They thought it was easy-to-use, and imagined that most people would learn to use this tool easily or very easily.

Accuracy of the information

Regarding the correctness of the elements displayed on the dashboard, the responses are very different from one respondent to another.



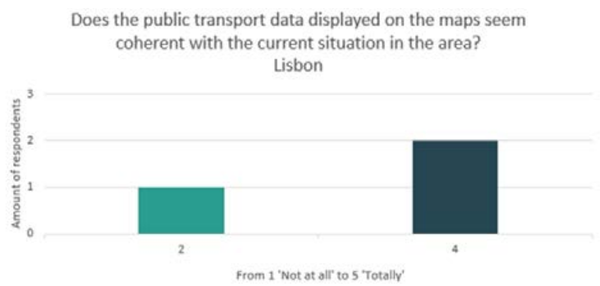


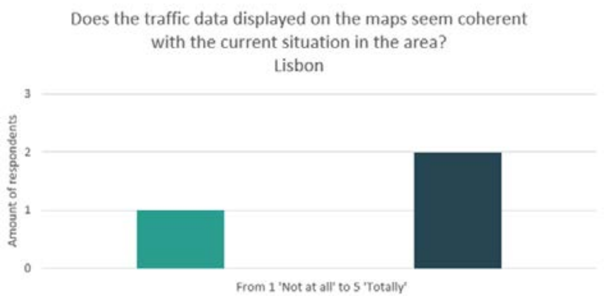
The KPI figures and the incident data displayed on the maps are assessed as either not relevant, neutral or a bit relevant. These different interpretations of the showcased data may stem from:

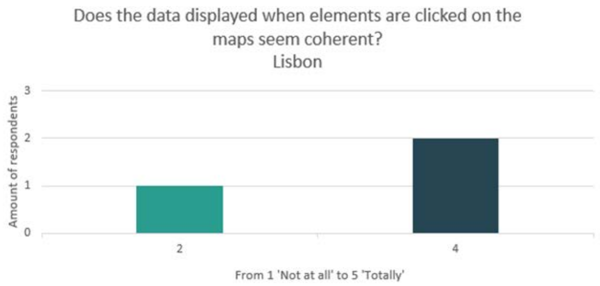
- the distinct profiles of the approached stakeholders, who may look to different information to fit their line of work;

- a different understanding of the scope of the tool, stemming from varying degrees of contact with the project prior to the 'Step 1' testing phase;
- moreover, during this initial testing, the showcased indicators were not yet filtered for the Case Study area, displaying values for the whole city of Lisbon; therefore, stakeholders comparing their perception of the Case Study area with these indicators would necessarily find discrepancies.

Otherwise, the data displayed on the traffic and public transport maps are assessed as coherent by the majority of the respondents, such as the information that appears when elements are clicked on the map.







Some comments highlighted some issue on the average road speed that was abnormally low. These seemingly abnormal values (also observed in Figure 26) stem from a limitation of the Waze Road Congestion dataset which is: only the congested roads are considered in the dataset; therefore, the computation of average speeds disregards the travel speeds in uncongested roads. In the Lisbon Case Study, this indicator is more aptly called "average speed in congested roads".

Compliance with initial requirements

To the question "Do you think the city data sources are well covered?", the respondents answered in majority "Neutral", with a comment explaining that the proper functioning of the platform is heavily penalized by the lack of data.

For the Lisbon case study, CARRIS and A-to-Be defined 5 expectations for this functionality (current state of the network) at the beginning of the project. The following feedback has been shared by the respondents of the survey:

Expectations	Not met	Partially met	Totally met
Expectation 1: Visualise the current congestion levels on the roads		33.3 %	66.7 %
Expectation 2: Visualise the PT lines and find/click the individual vehicles. Clicking on the vehicle should identify the service they are operating and the current delay	33.3 %	66.7 %	
Expectation 3: Visualise the severely delayed services (delay above 5 min), and pinpoint service failures	100 %		
Expectation 4: Pinpoint the location of incidents (such as traffic accidents, congestion or cut roads)	33.3 %	33.3 %	33.3 %
Expectation 5: Compare the current conditions to the weekly average, and to specific days of the week	66.7 %	33.3 %	

Table 4: Lisbon case study expectation for the Service 1: Current Status of the Network functionality

Overall, the respondents from the Lisbon case study are moderately satisfied with the first functionality implemented: the dashboard is easy to use, and the traffic and public transport elements are clear and reliable, but some aspects could be improved, regarding the KPIs and incident reliability. It is important to highlight the fact that the low quality of the data and, critically, the lack of region filtering capabilities at the beginning of the tests impacted the reliability of the elements displayed. In addition, the expectations defined by the Lisbon case study were more detailed and precise than the other case studies, which can explain why they are almost all partially or not met. One such example is the visualization of delayed services: the real-time PT service dataset (GTFS-RT) does not contain this information natively; thus, to be able to display delay information on the TANGENT tool, it would be necessary to run an external computation algorithm that compared the planned services (GTFS dataset) with the vehicle positions (GTFS-RT dataset).

Following this initial feedback from the Lisbon stakeholders, a series of iterations were performed to address the main critiques of the tool, which enhanced the tool and expanded its usability to other areas.

Two main aspects of the tool were revisited that improved the usability and readability of the displayed information, but were deemed as subject for further improvements. These are:

- The identification of vehicle services was reviewed and now correctly displays the operated line and the vehicle ID. Nevertheless, it is still not possible to add the direction of the movement, forcing users to use contextual cues to understand the direction of movement;

- An attempt was also made to display the PT line route in the map visualizations; this endeavour was partially successful but was performed through a manual process that requires intervention when there are significant changes in the planned services (for instance, when changing from Summer vacations to School-period months).

On the other hand, one improvement in particular was found to greatly expand the tool's usability, which was the addition of filters. Two types of filter were added:

- geographical filters (Figure 27, Figure 28 and Figure 29), which allow users to define their own interest areas within the city and analyse events and indicators only for that scope;
- and line-based filters (Figure 30), which allow analyzing the performance of specific lines.

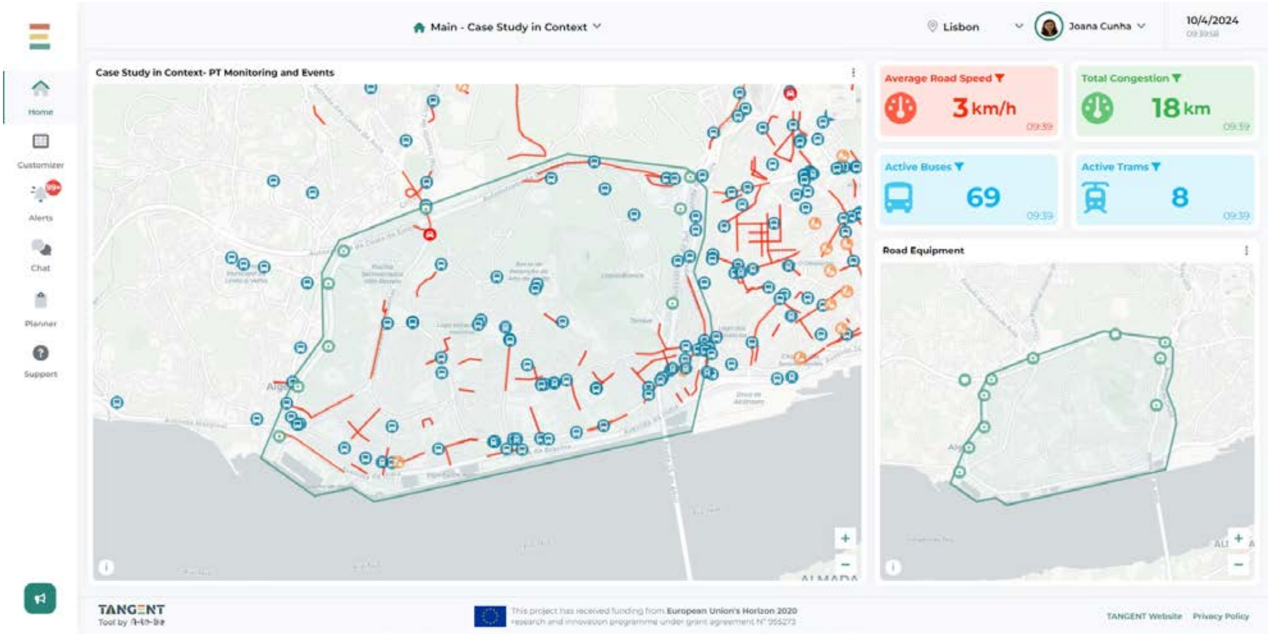


Figure 27: Case Study view with indicator filtering

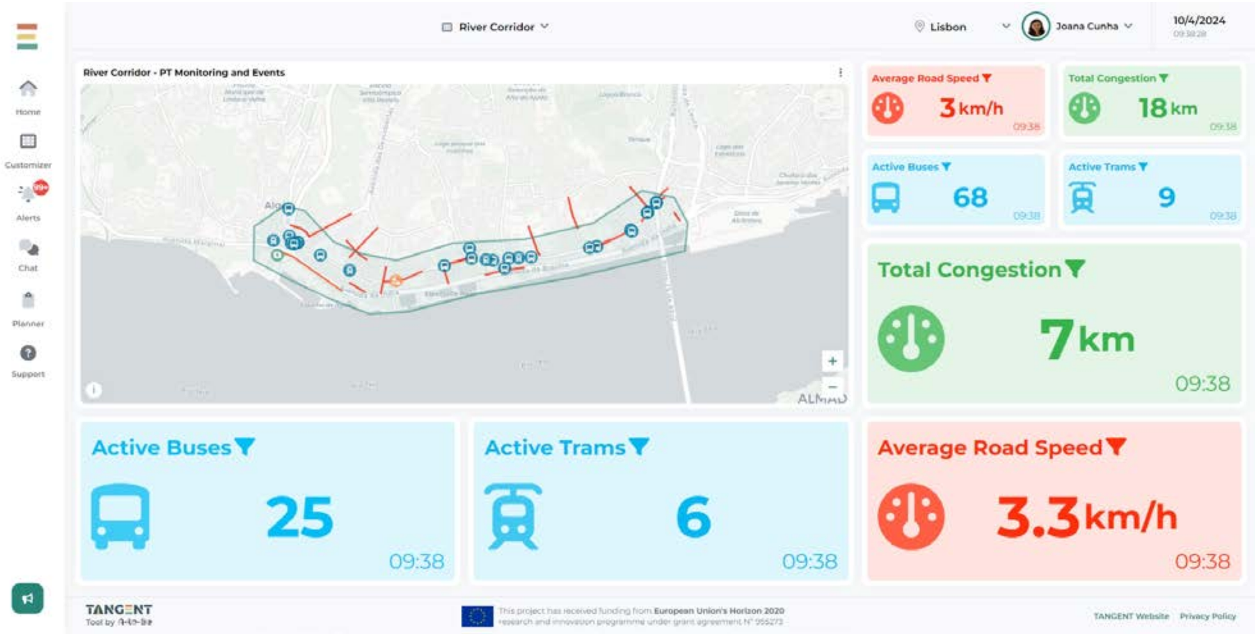


Figure 28: River Corridor View with map elements and indicator filtering

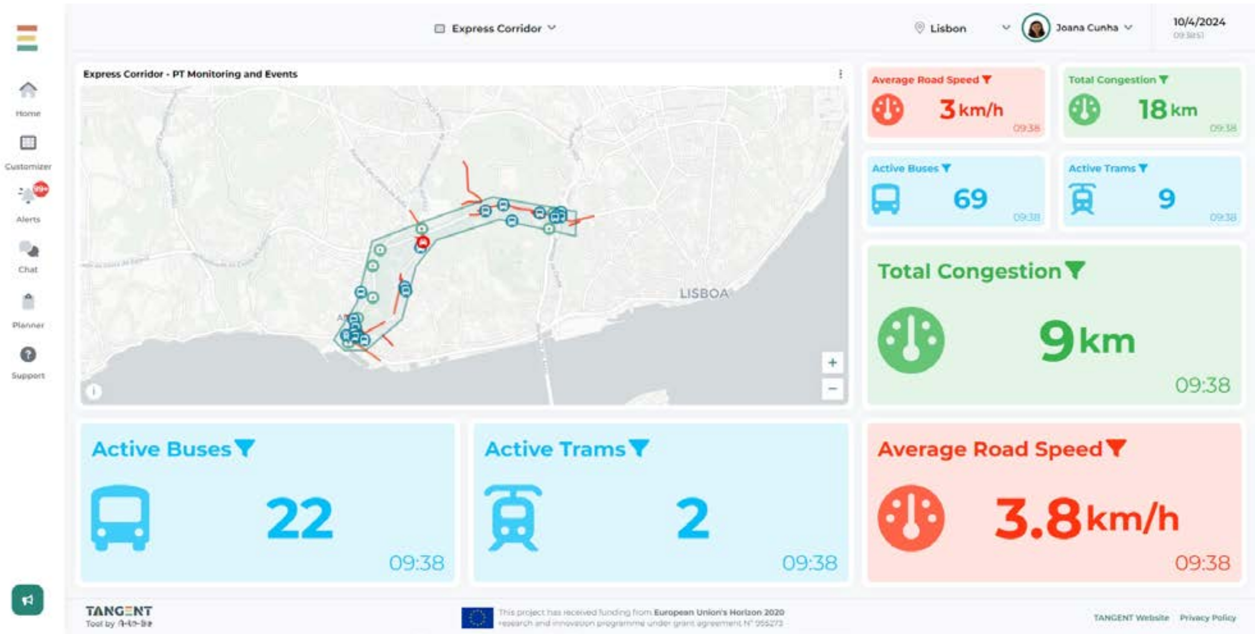


Figure 29: Express Corridor view with map elements and indicator filtering

Figure 30: Line view with map elements and indicator filtering

This filtering capability was successfully implemented and allowed TANGENT to address several comments raised by the stakeholders:

- the gap between the stakeholders' perception of the congestion and the congestion-related indicators is no longer so pronounced, since the user has the option to focus on specific regions of the city;
- PT Traffic Control operatives only monitor specific lines attributed to them, thus the line filtering option was an early request. The ability to filter vehicles by line answers permits a more readable visualization for these operatives. Moreover, geographical filtering can also be employed by them to track events and road conditions along the routes (see Figure 30), allowing them to more easily identify potential disruptions to the service;
- Furthermore, throughout the introduction of these improvements, it was found that the 'Service 1: Current state of the network' functionality could also be of use to other PTO staff, not directly linked to operations management. In the Lisbon Case Study, there was an interest by Customer Support teams to access this tool, since their call centre staff are often contacted by customers complaining/inquiring about delayed/missing buses and trams. At this moment, to address the customers' complaints, they must first ask the Operations teams to explain the cause of the disruption since they are not privy to this information. Through TANGENT's Service 1 functionality, these staff would be able to have an immediate overview of the occurrences in the area and answer in real-time to the customer, providing a quick justification for the inconvenience and (if needed) advising the passengers on alternative PT services they may use.

4.2.2 Step 2: Future state of the network

The step 2 of the testing aimed at assessing the second functionality of Service 1: **Future state of the network**, during daily commuting.

It is important to note that, despite the initial testing of the prediction algorithms for car circulations in the A5 highway (described in Section 3.5), it was not possible to include these predictions in this functionality. This is due to the available dataset not being available through a live feed; rather, it had to be retrieved manually, inhibiting its inclusion in the real-time monitoring and forecasting functionalities. In this context, the predictive functionality was tested for the PT services as it was the only dataset that could feasibly be included.

In contrast to the Step 1 testing phase, the Step 2 testing and validation procedure was not conducted through a survey. Instead, the functionality was addressed through bilateral meetings with relevant stakeholders, where they were asked to provide feedback. This approach was preferred for a few reasons:

- as only one indicator was being tested, it was found that the creation of a dedicated survey to gather the user feedback would lead to undue effort, while resulting in less open responses than those obtained through direct dialogue;
- the interpretation of the forecasted value as the total travel time for new services (rather than for the currently selected vehicle) is not immediate; thus, it was necessary to make sure the meaning of the indicator was correctly understood by all users when asking for their feedback;
- finally, users had no terms of comparison for the forecasted value, since this indicator is not commonly addressed in operation performance analysis, and much less used by traffic/operation controllers; as such, the validation of the forecasting reliability and accuracy had to be assessed based on each individual's sensibility on the subject, which is better elicited through bilateral conversations.

Furthermore, this approach enabled showing respondents the dashboards developed for Rennes and Manchester and, thus, display the potential of the Service 1: Future State of the Network functionality when applied to other datasets.

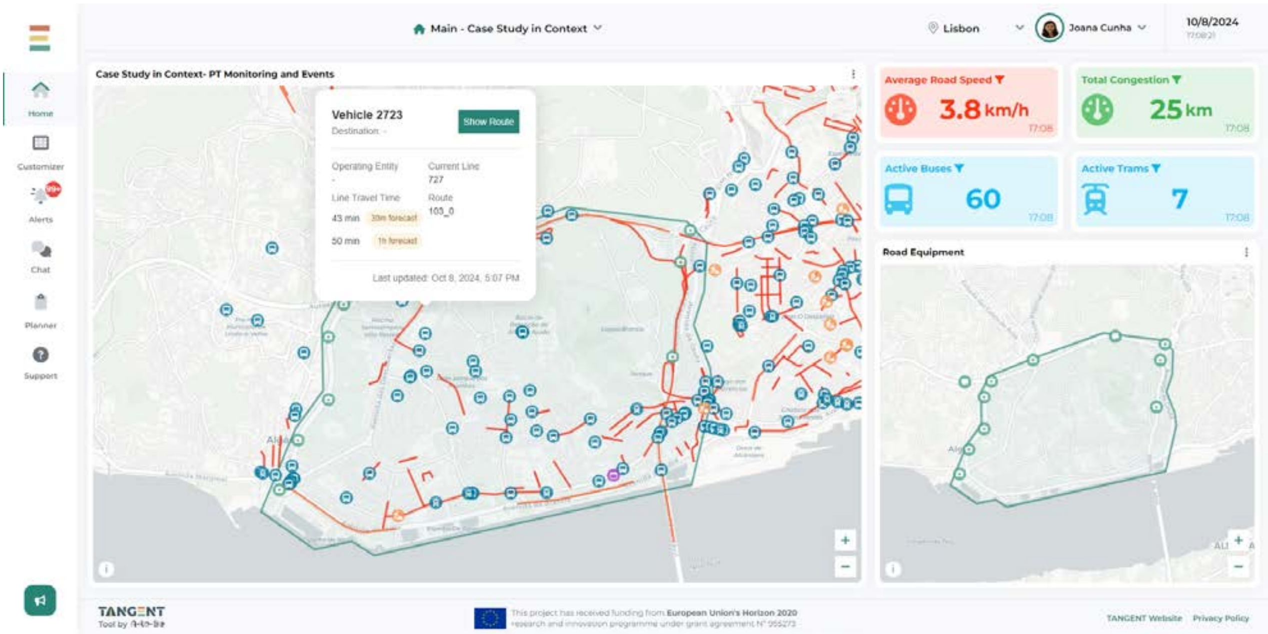


Figure 31: Detail of travel time predictions for vehicles

Respondents

For this functionality, 7 respondents provided their feedback:

- 1 Project Manager
- 2 PT Traffic Control Engineers
- 2 ITS staff
- 1 Communication and Dissemination PTO staff
- 1 Customer Care PTO staff

Results

The overall feedback on the functionality was that it is promising for anticipating possible traffic disruptions and PT service delays. When observing the Manchester and Rennes applications of this service, it displays very interesting indicators that are of use to Traffic Controllers and Operation Managers alike.

In the case of the Lisbon case study, the unavailability of sensoring data, and the lesser richness (in terms of detail) of the BUS real-time data yielded a single indicator which is not straightforward to assess. Nevertheless, tests on the Service 1: Future State of the Network functionality revealed that the forecasted values are sensible in most occasions and, with the right kind of data, may prove a powerful asset in the TANGENT tool ecosystem.

4.2.3 Step 3: SNLB: Informing the transport passengers

Step 3 of the testing aimed at assessing one of the functionalities of Service 2: **SNLB - Informing the Transport Passengers**.

Once integrated in the dashboard, the objective of this functionality is to allow the transport/traffic managers to create dissemination plans that can be activated by a list of designated profiles. Once a dissemination plan is activated, it is triggered when the previously stipulated triggering conditions are met; when this occurs, a parameterized message and real-time data are displayed on the chosen dashboard (see Figure 32) and for the chosen user profiles.

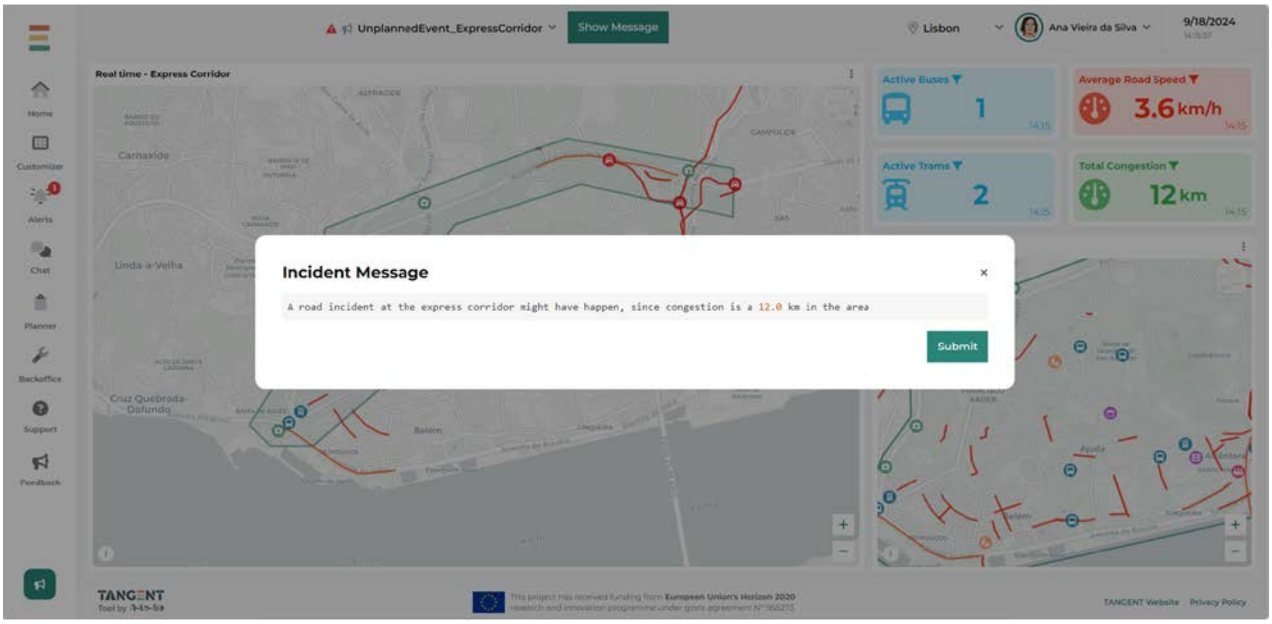


Figure 32: Example of incident message alert

The functionality and the steps of the dissemination plan creation are further described in D6.5.

Lisbon was the second case study to assess this functionality, after Rennes. This assessment was performed through an online meeting with the following partners: CARRIS, A-to-Be and ID4MOBILITY. First, A-to-Be presented the functionality with a demonstration on the platform, creating a dissemination plan. Then, CARRIS tested the functionality by creating their own dissemination plans and finally, the city asked several questions and shared feedback:

- Lisbon users revealed an interest in using this functionality to generate automated warnings regarding events affecting the overall region, or only specific areas/points of interest
- The users also proposed some improvements to this functionality, namely:
    - being able to create calculated indicators from the TANGENT datasets (for instance, the number of detected incidents/accidents in the area), so they can be used as triggering conditions or as inputs for the message to disseminate
    - being able to use forecasted indicators as a triggering condition (e.g., if the forecasted delay is above 5min, generate a warning)
    - creating more complex relationships between triggering conditions: currently, only conjunction or disjunction relationships are allowed when tracking multiple conditions (i.e., sets of conditions much all be either AND or OR); however, with the ability to select different types of relation to compare groups of conditions, it would be possible to refine the search and produce more specific dissemination plans
- One final comment presented an existing challenge for Lisbon users: based on the available datasets, and the ensuing indicators for building conditions, the dissemination plans that can be built are very operational; as such, were are not deemed ready for passenger dissemination. To translate this information to the Passenger Communication channels, human intervention would always be required.

The feedback from CARRIS gives elements to iterate and improve the functionality in order to match the needs and challenges of the use cases. It also gives elements to propose to the other case studies, during their tests. Once the three non-virtual case studies have provided their comments, the evolution proposals that emerged most frequently and that are the most relevant were studied to be integrated in the next iteration.

Common section to D7.3, D7.4, D7.5 (the the Rennes, Lisbon and Greater Manchester case studies deliverables)

The testing meetings with each non-virtual case study to test this functionality were extremely useful to highlight common improvements to implement on the Dashboard, in order to improve SNLB - Informing the Transport Passengers. The following elements were implemented:

- Geographic filtering is now available in the Customizer, in order to create a Dashboard with maps and indicators focused on some specific geographic areas. It can then be used for this functionality: the user can parameter a dashboard on a specific area and choose it when creating a dissemination plan. It is also relevant for the previous functionality tested “Future state of the network”, for which the users can define specific indicators both for real-time and forecast and compare the two types of KPIs.
- The trigger conditions can be defined using any indicator widget. This allows the use of Forecasts and filtered data as triggering conditions too.
- It is possible to add more than one triggering conditions and define the parameters of these several conditions: conjunction (“and”) or disjunction (“or”)
- Data filtering: in the same way as the geographic filtering, when creating a new map or a new indicator in the Customizer, the user can select the data he is interested in visualizing
- The messages of the dissemination plan are published in an AMQP Exchange, allowing integration with other systems for further distribution

End of the common section

4.2.4 Step 4: CIM: Synchronisation of Public Transport and Traffic Control

The Lisbon city model was calibrated for a macro-level scope. As such, it was found to lack the level of detail required to run the optimization algorithms developed in WP5. Therefore, it was not possible to test the “CIM: Synchronisation of Public Transport and Traffic Control” functionality in the Case Study if Lisbon.

4.2.5 Step 5: Simulation of predefined scenarios: ‘Increased Bus and Tram supply’ and ‘Implementation of congestion pricing’

Due to a higher-than-expected technical complexity of the module developed in WP5, required by the functionality *What-if simulations*, this functionality is under testing and can only be reported by mid-November 2024.

4.2.6 Testing the TANGENT functionalities under different scenarios

As previously mentioned, the main goal of the TANGENT tool is to support transport stakeholders manage traffic and operations under different conditions, helping them deal with typical traffic conditions, as well as with events (either foreseen or unexpected) that affect mobility within the city.

In the Lisbon case study, the following functionalities were successfully tested for all scenarios:

- Service 1: **Current state of the network**
- Service 1: **Future state of the network**
- Service 2: **SNLB: Informing the transport passengers**

The following sections present some of the most notable events captured through the TANGENT tools during the non-technical testing stages.

Daily Commuting

The utilization of the Service 1: "Current state of the Network" through the summer months allow the observation of the seasonal nature of Lisbon traffic patterns, and its effects on overall congestion in the city. As can be observed in Figure 33, there was a notable decrease in congestion during the summer, particularly in August (which is the month when most citizens vacation), followed by a steep increase in congestion with the re-start of the school period.

Figure 33: Captured seasonal effects - Left: School Summer Holidays; Middle: August; Right: School period

Planned Events

The selection of the Lisbon case study area was largely influenced by the proximity to the Passeio Marítimo de Algés, which holds the NOS Alive music festival every year in early July. The organization of this event introduces some traffic circulation restrictions around the venue area. Through the usage of Service 1: "Current state of the Network" (Figure 34) it was possible to observe and monitor the effects of this event in the increased congestion in the roads surrounding and leading to the venue.

At the time of the event, Service 1: "Future state of the Network" was also available for testing. As can be seen in Figure 35, the prediction algorithms were able to contemplate the effects of the increased congestion in the travel time predictions: the 30 minute forecast estimated a 55% increase in travel time due to the currently observed congestion levels; while the 60 minute forecast estimated a stabilization of traffic conditions, leading to a travel times in line with the service schedule.

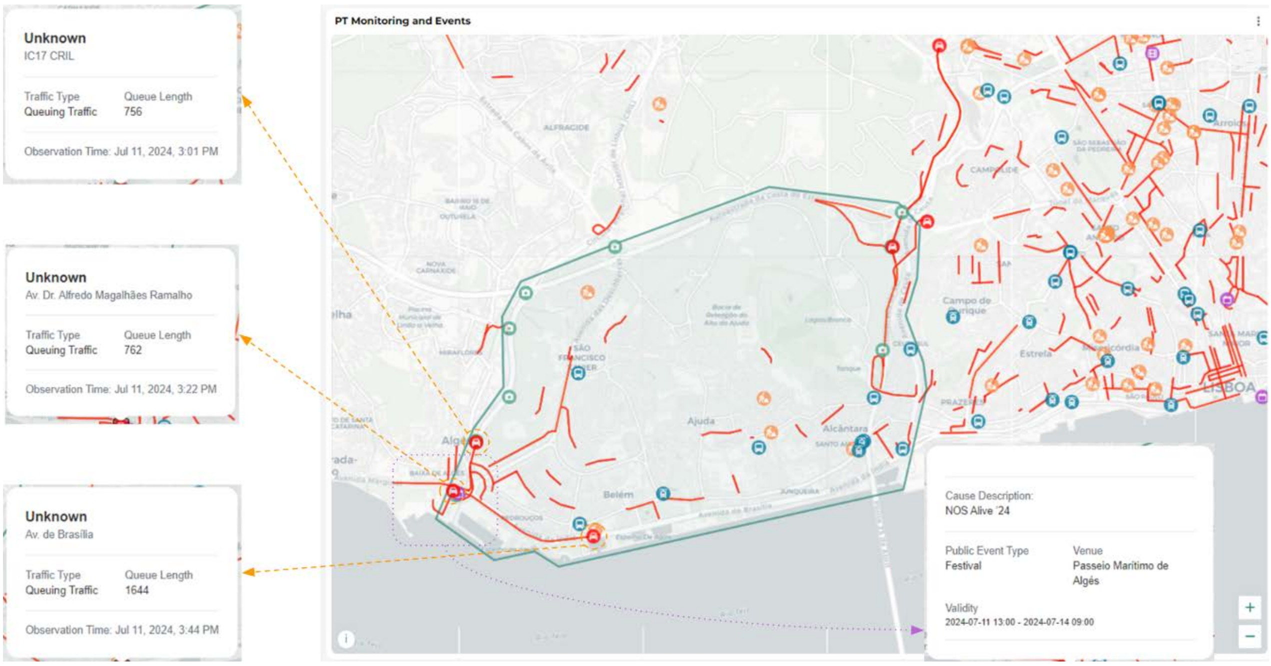


Figure 34: Example of identified incidents/congestion near the NOS Alive venue

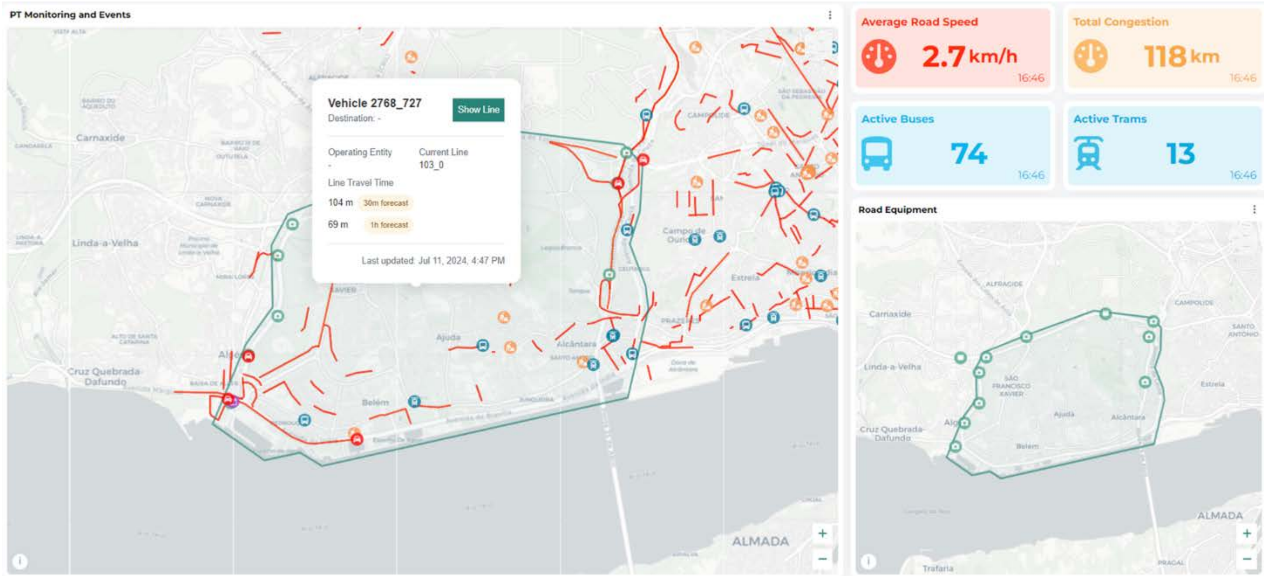


Figure 35: Example of future state predictions during the NOS Alive event

With NOS Alive occurring in early July, at the time of the event it was not possible to test all the TANGENT functionalities, due to them not yet being fully developed and integrated into the platform. Therefore, a new planned event was selected: the October 6th EDP marathon; this is a smaller scale event but prompted to closure of all the roads in the Tagus River front, from Cascais to Lisbon.

With the full toolset of TANGENT at our disposal, and with the ability to filter the areas and indicators of interest, it was possible to produce a highly detailed coverage of the event: a specific dashboard (see

Figure 36) was developed that shows the River Corridor area and collects the congestion indicators for the road where the marathon took place (bottom indicators), as well as for the nearest parallel road (top indicators) that was not closed or affected by the event; it also compares the road-specific congestion indicators with the indicators at the corridor level (on the right).

Figure 36 showcases the evolution of the congestion during and after the event in the River Corridor: it can be observed that during the marathon the speed was 0 km/h, meaning that no cars were able to circulate; upon road re-opening, there was an increase in the queue length, from 1 km to 3 km. Looking at the parallel road, the congestion values remained stable and were seemingly not affected by the event.

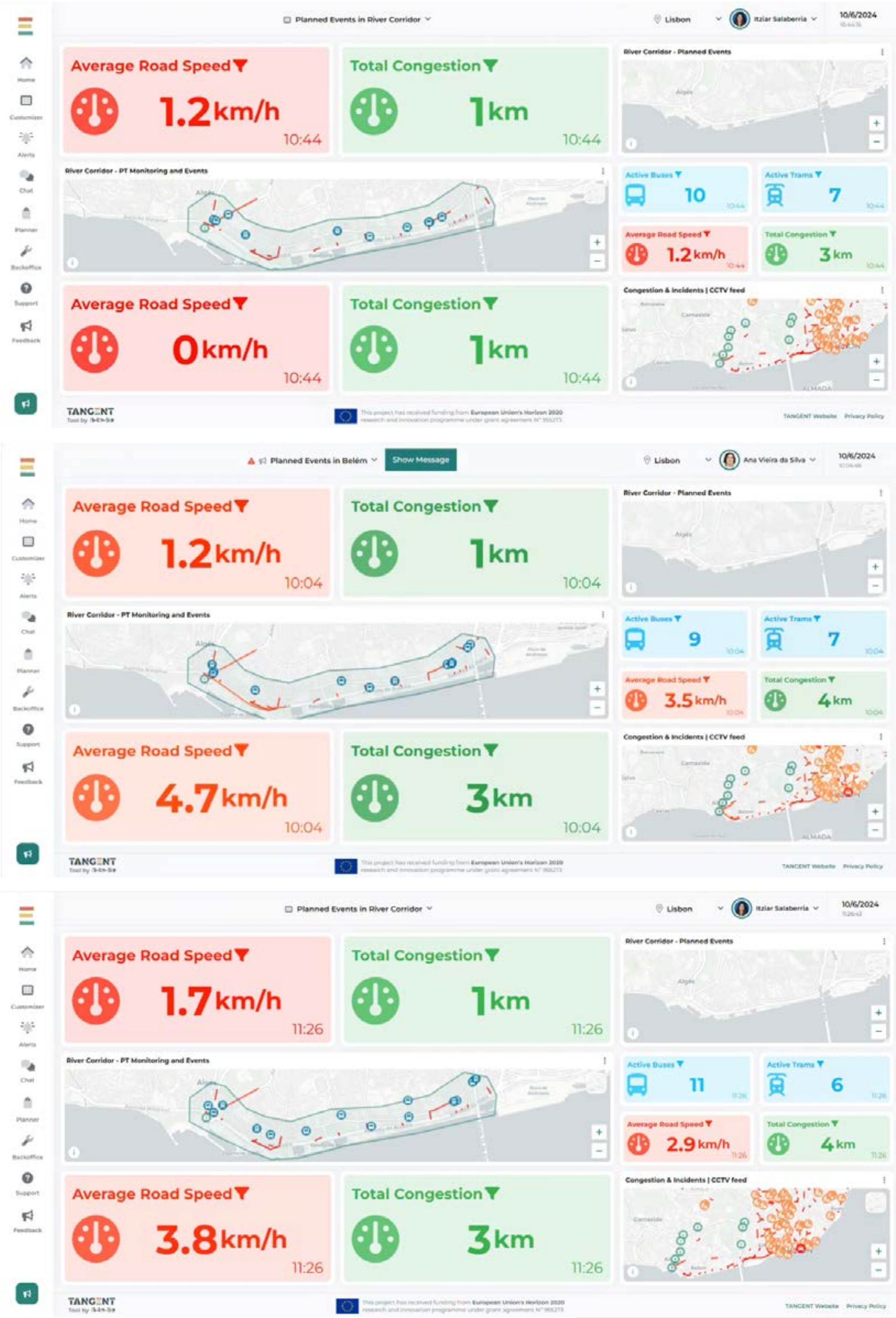


Figure 36: Congestion evolution during the Marathon in Belém - Top: 10:45 (CET); Middle: 11:05 (CET); Bottom: 11:30 (CET)

Comparing the day of the event with a typical Sunday (Figure 37), we can observe the true effects of the event in the region. For both roads evaluated, the speeds in congestion were halved on the day of the event; moreover, the total queues doubled on account of the increased congestion in the road where the marathon took place. Looking at the River Corridor as a whole, these effects were somewhat dissipated, although we can observe the significantly lower speeds during the event.

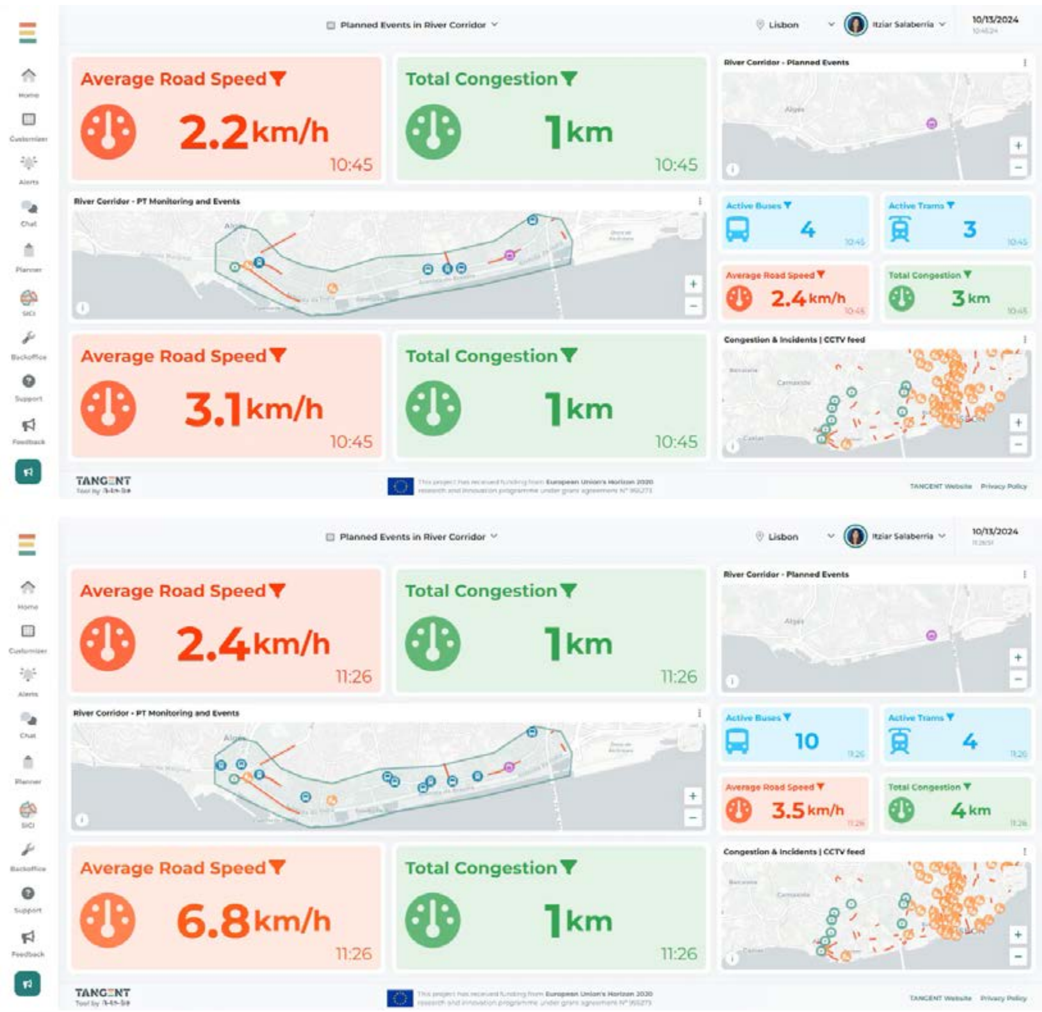


Figure 37: Congestion evolution on a typical Sunday - Top: 10:45 (CET); Bottom: 11:30 (CET)

Finally, the Service 2: “SNLB: Informing the transport passengers” functionality successfully triggered an incident message upon detecting an anomaly in the observed indicators for the roads under analysis.

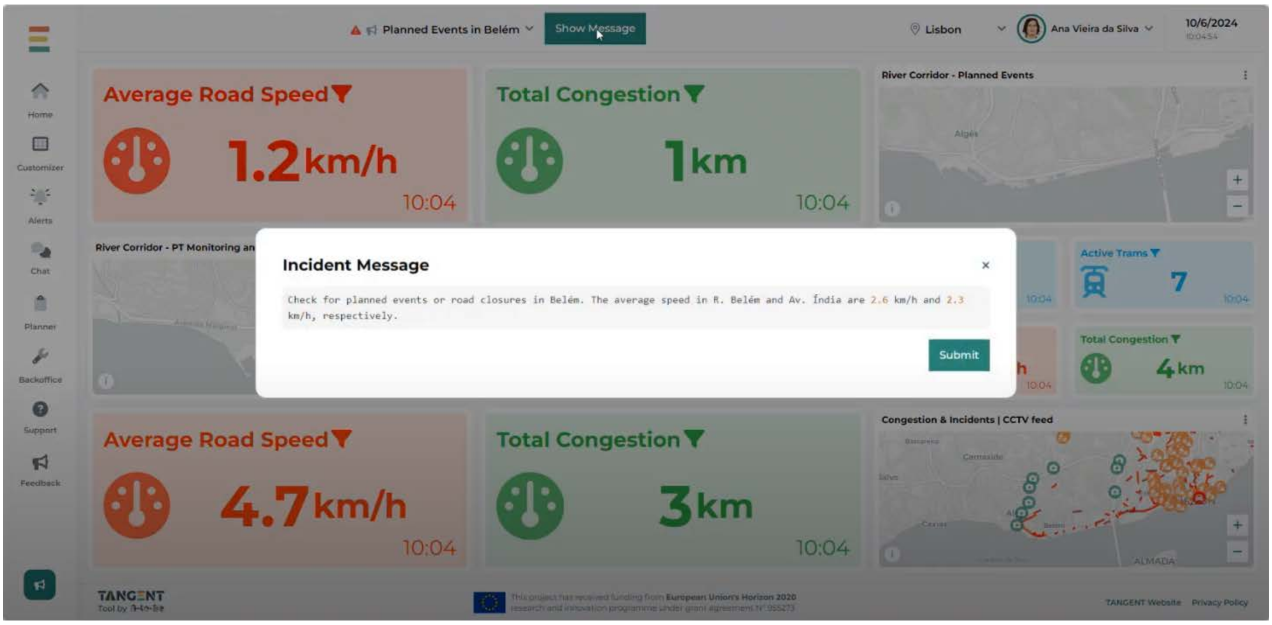


Figure 38: Dissemination Plan triggered during the Marathon in Belém

Unplanned Events

During the non-technical testing, a few unplanned events were detected, such as small accidents and extreme congestion. One notable event captured in the Lisbon Case Study area was the closure of the 25 de Abril bridge due to unplanned police operations.

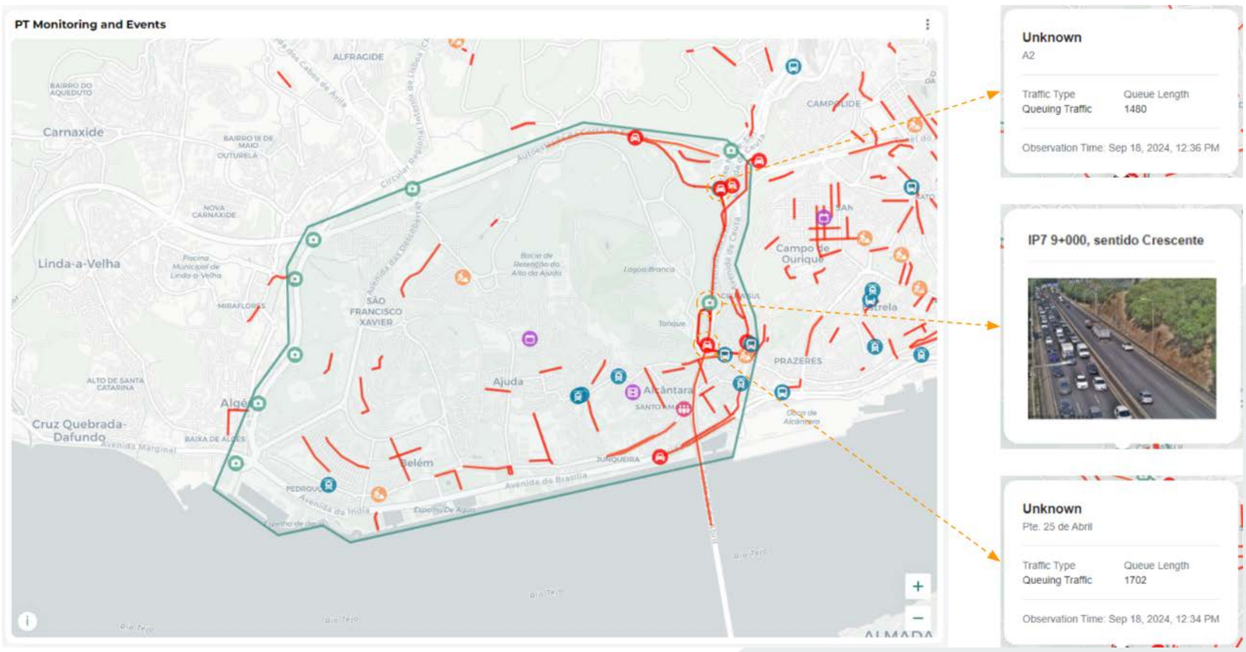


Figure 39: Detection of unplanned event

Figure 39 displays the information observed in the Service 1: “Current state of the Network” functionality. The access to real-time Waze traffic congestion levels and warnings allowed users to have a quick overview of the affected roads. Moreover, the access to road camera footage showcased the intensity of the traffic in these areas, allowing users to quickly validate the Waze information. In this context, the added-value of the TANGENT tools is revealed, whereby it allows users to have a solid grasp on the traffic conditions and adopt some cautionary measures (which could be simply to closely monitor the developments), even if they still weren’t able to assess the causes for the phenomenon (which, in this specific event would be to contact the police).

Moreover, the Service 2: “SNLB: Informing the transport passengers” functionality also proved a valuable assess in this context, since it was able to detect the extremely high congestion levels affecting the Express Corridor. It is important to note that the ability to filter regions of interest played a relevant role in the incident detection for this particular event, since that, if we were evaluating for the full city, the congestion levels would probably not be anomalous enough to trigger the dissemination plan.

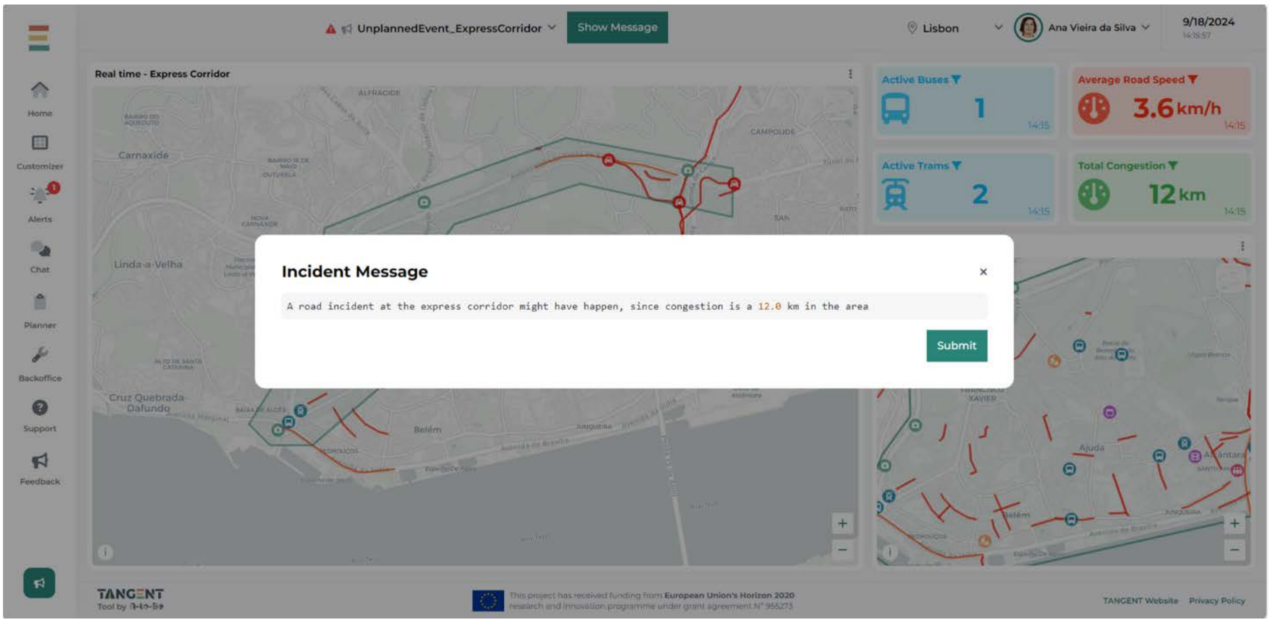


Figure 40: Dissemination Plan triggered during unplanned event

5 Lessons learned and next steps

5.1 The lessons learned from TANGENT for the Lisbon case study

5.1.1 Multi-actor collaboration on traffic management

Using the TANGENT functionalities, and in particular Service 1, we were able to assess the importance of (and need for) multi-actor collaboration for a more efficient traffic and transport management. Throughout the project's development, it became evident that the information owned by different parties is important for other actors in the system to better enhance their own operations and respond to incidents.

Some aspects are already coordinated between entities, albeit in a more "manual" fashion. An example is the pre-existing contact between CARRIS traffic controllers and the Municipality to notify of planned road works and similar instances; likewise, CARRIS is also in constant contact with the Police, to resolve unexpected incidents, such as road accidents.

Through the contact with these entities, it became apparent that all parties were interested in adopting a more automated process for this information exchange. A single platform showcasing all the events in the city, such as the one developed by TANGENT, can be the answer. To achieve this level of integration, however, there is a need for reliable data.

5.1.2 Data sharing / availability / reliability

5.1.2.1 Technical aspects

We realised that many of the preexisting open data sources were unreliable. Namely, some of these sources were not frequently updated (some even became deprecated during the project timeline) or difficult to interpret/translate to the projects' intended indicators.

Moreover, real-time data sources were not available in open data, prompting the need to request external entities to provide the needed information, which presented a challenge in the Lisbon Case Study.

Nevertheless, the efforts undertaken to contact and engage external stakeholders to share their data yielded some positive outcomes: the mutual benefits of data sharing were recognized by the different parties and led to the first steps in establishing data and knowledge sharing protocols between the Lisbon shareholders.

5.1.2.2 Legal aspects

In the Case Study of Lisbon, an obstacle was encountered, being that the critical data sources needed to feed the TANGENT tools were proprietary and/or deemed sensitive by their owners. This circumstance required the negotiation of NDAs between the TANGENT Technical partners and the data owners (the Lisbon Municipality and CARRIS). Likewise, the access to the city model (owned and calibrated by EMEL) was only granted after a bilateral knowledge exchange protocol was established between CARRIS and EMEL. While these obstacles were successful overcome, they implied a lengthy terms agreement process, which delayed the project's access to these elements.

This situation reveals a circumstance that is very common throughout different European cities, whereby: different entities own relevant and complementary data about the transport system, with no established sharing protocols and often adopting a protective attitude towards their data (sometimes, for fear of breaching GDPR). This often hinders attempts to develop new tools that depend on data integration, either due to legal obstacles as those encountered in TANGENT, or due to a lack of knowledge of the existence of some datasets.

5.1.3 Functionalities of TANGENT

5.1.3.1 Current State of the Network

This functionality was found useful by several different stakeholders and to serve different purposes. As abovementioned, CARRIS, the Lisbon Municipality and the Municipal Police have a direct communication channel to warn of planned road closures and road accidents or other incidents; which allows them to expedite the disruption resolution. The "Current State of the Network" functionality allows bypassing this step by automatically showcasing these events in a view that is common to all.

Notwithstanding the need to communicate among themselves to share further details on the incident and coordinate the best response, this common view allows all parties to have a fuller picture of the incident and its surrounding area, as well as the potential ramifications on other services or roads.

For this reason, this feature was found to be of use, not only for managers and operators, but also for customer support operatives. In particular, PTO call center staff receiving queries and complaints from passengers affected by a PT service disruption can make use of the TANGENT dashboard to quickly assess the situation and its cause, and provide an informed response or recommendation to the client.

5.1.3.2 Future State of the Network

This functionality if a very promising feature, as showcased in other case studies (D7.3 and D7.5).

Although the outcomes of the forecasts in the Lisbon Case Study were severely affected by the scarcity of adequate data to feed the predictions, the results yielded in the Rennes and Manchester were able to demonstrate the added-value of this feature when fully implemented.

5.1.3.3 Dissemination Plan

This functionality has the added-value in helping focus practitioners towards emergent incidents and patterns that may impact PT services and overall traffic.

In particular, the ability to tailor each message to specific regions of interest or PT services and to whom the message is disseminated allows users to communicate both internally and with other stakeholders.

Nevertheless, this tool would benefit from further integration with other datasets and from the ability to customize the KPIs that can be monitored (as mentioned in section 4.2.3).

5.1.3.4 Synchronisation of PT and TC

It was not possible to test this functionality in the Lisbon case study.

5.1.3.5 What-if Scenarios

This functionality if currently undergoing testing in the Lisbon case study.

5.1.4 TANGENT Dashboard

The dashboard developed within TANGENT provides a single unified view of the city's traffic conditions integrating real-time data from several sources. If such a solution were in the market, it would be a valuable asset to complement the already existing tools for managing BUS operations.

While the version of the dashboard tested within the project was limited by the available data, the resulting tool has the flexibility to integrate and adapt to new data sources, as they become available (for instance, if new stakeholders decided to share their respective datasets, or if the existing sources were upgraded to contain more detailed information).

The ability to customize views also ensures that the tool can be utilized for different operations, enabling both general overviews of the current mobility conditions (for instance, to inform decision-making), and detailed analysis on specific areas or services (which would fit the need of more operational tasks).

5.1.5 TANGENT project in general

Through CARRIS's participation in the TANGENT project, several learnings were gathered in the Lisbon case study. Namely, it helped have a broader view of the technical specifications and functionalities that will be required when considering system changes/updates.

In particular, in regard to existing data sources: it helped systematize the existing real-time data sources in the city, both open-data and proprietary, and assess the level of detail and reliability of each one. Several gaps were found that hinder the generation of useful data, for instance when aiming to disseminate to the public.

Through the interaction with other cities, it was possible to find common challenges in traffic management and discover new traffic management and data gathering solutions that can be assimilated by Lisbon stakeholders.

5.2 Next steps after the end of the project

Upon the end of the TANGENT projects' timeline, some future developments are envisioned. Firstly, the contacts promoted with other stakeholders revealed a shared interest (and need) on integrating multimodal information, to better inform decision-makers and practitioners in the transportation system. To promote the realization of this goal, further contacts will be encouraged among stakeholders to promote the creation of a data-sharing culture amongst transportation stakeholders.

Moreover, the project revealed gaps in the existing tools that can prompt the need to rethink the system and data requirements, as well as to turn the traffic management system more collaborative and interlinked with the PT operations.

In the future, Lisbon/CARRIS shall continue to participate in R&D and in higher TRL level projects as a pilot testing ground.

6 Conclusions

The deliverable D7.4 aimed to present the activities of the Lisbon case study during the development and evaluation of the TANGENT subsystems and the final integrated tool developed by the technical partners.

In a first stage, the subsystems were tested in collaboration with the technical partners to assess the accuracy of their outputs. Each technical partner provided their own testing methodologies, which differed depending on the subsystem under evaluation (e.g., surveys, meetings, document exchanges), and shared the obtained outcomes.

In the second stage, the Lisbon case study focused on evaluating the overall tool, as the subsystems were gradually integrated into the platform's functionalities. With the support of relevant technical partners, CARRIS and the engaged stakeholders assessed the usability, reliability, and accuracy of the dashboard and its functionalities, as well as their compliance with the expectations and requirements set at the beginning of the project.

The iterative approach to functionality evaluation allowed CARRIS to suggest improvements, enabling the tool to be tailored to meet the different users' needs. Overall, the testing activities revealed that the dashboard is user-friendly, intuitive and provides relevant information in a clear way. The customisation options are powerful and adaptable, however some users were overwhelmed by the freedom this grants. The wide range of features granted by the tool (e.g., planner and customiser functionalities) add to its complexity, and require time and training to be used effectively. Due to higher-than-anticipated technical complexities in the development of the TANGENT functionalities, some elements are still under evaluation; subsequently, testing activities shall continue until the end of the project.

Key lessons were learned by CARRIS, particularly regarding the need to involve local transport stakeholders and to develop and nurture a collaborative culture across the sector. This became apparent during the data gathering stage and further confirmed during the workshops with stakeholders, where it became evident that the access to multimodal real-time information is of interest to operation management and benefits all players.

Another important lesson is the need to upgrade the existing data sources to increase the quality, reliability and level of detail of the collected information. In this, the collaboration with other cities within the TANGENT project gave us the opportunity to access some good examples of relevant and high-quality information sources, with which Lisbon may start to define new technical requirements for data collection.

The next steps for CARRIS include the continued engagement of transport stakeholders to promote the creation of data-sharing protocols within the Lisbon transport ecosystem. Through the TANGENT testing and engagement activities, it was possible to identify specific interlocutors among the most relevant entities.

7 References

1. Franco, D., Juliat, M., Damerau, M. (2022). Deliverable 1.1 “Multi-actor co-creation strategies for each case study”, Public report, TANGENT H2020 Project, Grant Agreement nr. 955273.

2. Franco, D., Juliat, M., Bandyopadhyay, S., Landaluce, H., Ghosh, A., Masegosa, A. D., Serrano, L. (2022). Deliverable 1.2 “NTM needs assessment and system requirements”, Public report, TANGENT H2020 Project, Grant Agreement nr. 955273.

3. Franco, D., Juliat, M. (2022). D1.3 “Multi-actor cooperation models for NTM”, Public report, TANGENT H2020 Project, Grant Agreement nr. 955273.

4. Comerio, M., Scrocca, M., Fiano, A., Grassi, M., Carriero, V., Gizzi, G. (2023). D2.4 “TANGENT API for data harmonisation and fusion. Second release”, Confidential report, TANGENT H2020 Project, Grant Agreement nr. 955273.

5. Konstantinou, C., Giouroukelis, M., Mantouka, E., Vlahogianni, E. (2023). D3.2 “Travel choice modelling (set of models, code). Second release.”, Public report, TANGENT H2020 Project, Grant Agreement nr. 955273.

6. Tympakianaki, A., Dominguez, M., Vanderhoydonc, Y., Rahimiasl, M., Chalkiadakis, C., Fafoutellis, P., Mahdaviabbasabad, S. (2024). D4.6 “Framework for real-time traffic monitoring and forecasting. Second release.”, Confidential report, TANGENT H2020 Project, Grant Agreement nr. 955273.

7. Ghosh, A., A. Laresgoiti, A., S. Larrauri, I., M. Arredondo, A. (2023). D5.6 “Transport Network Optimization Module. Second Release”, Confidential report, TANGENT H2020 Project, Grant Agreement nr. 955273.

8. Dias, T., V. Silva, A. (2022). D6.1 “Technical architecture and resource specification. First release”, Confidential, TANGENT H2020 Project, Grant Agreement nr. 955273.

Annex I: Online survey for the testing of the functionality "Current state of the network"

Introduction questions

What is your full name? *

Votre réponse

What is your role in the city (job title / department / organization) *

Votre réponse

Dashboard Usability and Reliability

This section aims to assess the user-friendliness and the clarity of the dashboard and the functionality.

Use your credentials and access to the TANGENT Dashboard.

Before clicking anything, spend 10-15 seconds looking around the main page.

How do you evaluate the amount of information on the main page? *

	1	2	3	4	5	
Not enough information	○	○	○	○	○	Too much information

The displayed elements (maps, figures) are clear and relevant *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

Let's start playing with the tool!

Click on the top right of a map and select "Expand" to make it bigger. You can move the mouse over the map, zoom in and zoom out, click on elements to get more information. Repeat for other maps as needed.

Is the information on live traffic and incidents (if available) easy to understand? *

	1	2	3	4	5	
Very difficult	○	○	○	○	○	Very easy

Is the information on live public transport easy to understand? *

	1	2	3	4	5	
Very difficult	○	○	○	○	○	Very easy

Any other comment regarding live data?

Votre réponse

Let's customize the dashboard!

If you do not have the access to customize the dashboard, you can skip this question. Otherwise, if you haven't yet, try using the customizer. If you need help, you can take a look at the tutorial video: https://www.youtube.com/watch?v=cybPjJvsNx4

The customization of dashboards is easy to do

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

Any other comment regarding customization?

Votre réponse

Global feedback on the TANGENT Dashboard

Overall review of the usability & user-experience of the TANGENT Dashboard

I found the tool unnecessarily complex *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

I found the tool easy to use *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

I would imagine that most people would learn to use this tool very easily *

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

Any other comment / question on the user experience?

Votre réponse

Accuracy of the elements displayed

This section aims to assess the correctness of the data, maps and other elements available on the dashboard.

Look at the KPI figures displayed (usually on the top right of the dashboard). Do they seem coherent with the current situation in the area? *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Totally

Does the traffic data displayed on the maps seem coherent with the current situation in the area? *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Totally

Does the incident data displayed on the maps seem coherent with the current situation in the area? *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Totally

Does the public tranport data displayed on the maps seem coherent with the current situation in the area? *

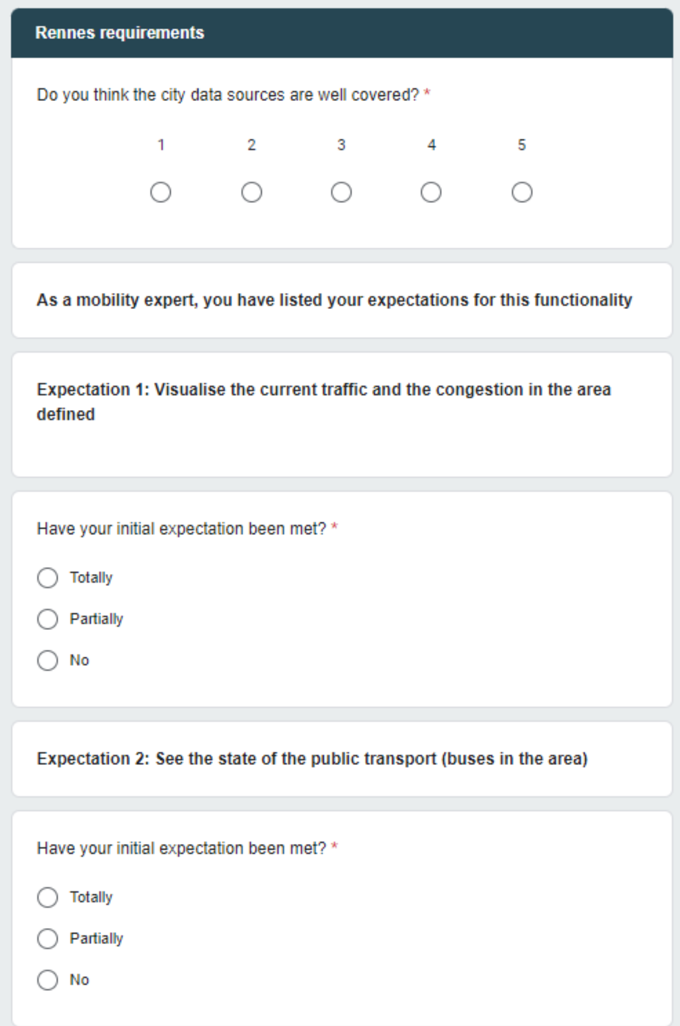
	1	2	3	4	5	
Not at all	○	○	○	○	○	Totally

Does the data displayed when elements are clicked on the maps seem coherent? *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Totally

If you found difficulties in any of the previous elements, please describe them.

Votre réponse



Rennes requirements

Do you think the city data sources are well covered? *

1 2 3 4 5

As a mobility expert, you have listed your expectations for this functionality

Expectation 1: Visualise the current traffic and the congestion in the area defined

Have your initial expectation been met? *

- Totally
- Partially
- No

Expectation 2: See the state of the public transport (buses in the area)

Have your initial expectation been met? *

- Totally
- Partially
- No

Expectation 3: Notice the delay of the various services (buses, shuttles)

Have your initial expectation been met? *

- ◯ Totally
- ◯ Partially
- ◯ No

Expectation 4: Visualise the planned elements that disrupt the traffic (road works, blocked road, etc.)

Have your initial expectation been met? *

- ◯ Totally
- ◯ Partially
- ◯ No

Expectation 5: Display traffic and public transport statistics and choose the granularity

Have your initial expectation been met? *

- ◯ Totally
- ◯ Partially
- ◯ No

Have your initial expectations changed? *

- ◯ Yes
- ◯ No

If the previous answer is yes, what has changed? *

Votre réponse

Any other comment?

Votre réponse

Annex II: Online survey for the testing of the WP5 Consensus Mechanism

How clear and understandable did you find the questions in this questionnaire? *

	1	2	3	4	5	
Not clear at all	○	○	○	○	○	Very clear

Were there any questions that seemed irrelevant or out of context? *

- ○ No, all questions were relevant
- ○ Yes, a few questions appeared to be out of context
- ○ Yes, several questions appeared to be out of context
- ○ I'm not sure/I didn't notice

To what extent do you agree that the instructions provided with the questionnaire were adequate? *

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

How would you rate the overall layout and formatting of the questionnaire? *

	1	2	3	4	5	
Not Good at All	○	○	○	○	○	Excellent

How would you assess the length of the questionnaire? *

- ○ Far too long
- ○ Slightly too long
- ○ Just right
- ○ Slightly too short
- ○ Far too short

How long did it take you to complete the questionnaire? *

- < 5 min
- 5 - 10 min
- 10 - 15 min
- 15 - 20 min
- 20 - 25 min
- 25 - 30 min
- > 30 min

Do you have any other comments or suggestions on the questionnaire content and format?

Short answer text

Modeling insights

The stakeholder's relations and preferences collected through the questionnaire were used as input to the Consensus Mechanism developed in TANGENT, which aims at finding the consensus-based, collective opinion of stakeholders.

Please select one of the four TANGENT problems for which you've indicated your preferences on the network objectives in the initial survey.

- ATTCAV
- DRTPT
- PTTC
- DCP

Depending on the response to the previous question, the survey respondent was directed to a personalized section featuring the modeling results for ATTCAV, DRT-PT, PTTC, and DCP. For illustration, the section for DRT-PT is included below.

DRT-PT

During the testing phase of the mechanism, we have engaged various stakeholders like you and determined that the collective opinion regarding the different network objectives for the DRT-PT (Synchronization of on-demand and transit modes) traffic management strategy is as follows:

Collective Stakeholder Preferences

Objective A is more important than Objective B					Objectives are Equal	Objective B is more important than Objective A				
Objective A	Strongly		Moderately				Moderately		Strongly	Objective B
Service performance	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1	☐ 2	☑ 3	☐ 4	☐ 5	Network efficiency
Economic performance	☐ 5	☐ 4	☐ 3	☐ 2	☑ 1	☐ 2	☐ 3	☐ 4	☐ 5	Service performance
Service performance	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1	☐ 2	☑ 3	☐ 4	☐ 5	Environment
Network efficiency	☐ 5	☑ 4	☐ 3	☐ 2	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	Economic performance
Environment	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1	☐ 2	☐ 3	☑ 4	☐ 5	Network efficiency
Economic performance	☐ 5	☐ 4	☐ 3	☐ 2	☐ 1	☐ 2	☑ 3	☐ 4	☐ 5	Environment

Strategy Description

As a reminder, the objective of this traffic management strategy is the joint optimisation of the running route and the scheduling (departure times) of Responsive Feeder Transit and Responsive Collector Transit systems, to facilitate access and egress connection to public transit services (first/last mile problem).

Service performance: Total passengers served
Network efficiency: Passenger travel time
Economic performance: Operational costs of the DRT fleet
Environment: Emissions of DRT fleet

How closely this collective preference matches your own opinion regarding the relative importance of each objective?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Very much

Would you consider adopting this opinion for the sake of reaching consensus with other stakeholders?

	1	2	3	4	5	
Definately Not	○	○	○	○	○	Definately Yes

Would you trust a mechanism that collects stakeholder's opinions and calculates the collective preference as a guide for decision-making:

	1	2	3	4	5	
Not at all	○	○	○	○	○	Very much

Besides the collective preference, what other output would you like the consensus mechanism to provide?

Short answer text