# A detailed report on the development process of the Replicable Sustainable Public Toolkit

OPTI-UP

Version 1.0
July 2026

This publication has been developed within the OPTI-UP project, supported by the Interreg CENTRAL EUROPE Programme with co-financing from the European Union. Its content is the sole responsibility of the authors and the project partnership and does not necessarily reflect the views or positions of the Interreg CENTRAL EUROPE Programme, its programme bodies or the European Union. Neither the programme authorities nor the European Union can be held responsible for any use that may be made of the information contained herein.

# Content

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| AF | Application form |
| AHP | Analytic Hierarchy Process |
| BRT | Bus Rapid Transit |
| DRT | Demand-responsive transport |
| EV | Electric Vehicle |
| ICE | Internal Combustion Engine |
| JSON | JavaScript Object Notation |
| KPI | Key Performance Indicator |
| LLM | Large Language Model |
| LUTI | Land-Use and Transport Interaction (modelling) |
| MCDM | Multi-criteria decision making |
| ML | Machine Learning |
| OSM | OpenStreetMap |
| PP | Project partners |
| PT | Public transport |
| RAG | Retrieval-Augmented Generation |
| TOPSIS | Technique for Order of Preference by Similarity to Ideal Solution |

# Disclaimer

This deliverable has been produced within the framework of the OPTI-UP project, co-funded by the Interreg Central Europe Programme (Grant Agreement No. CE0200713). The views expressed herein are those of the authors and do not necessarily represent the official position of the European Union, the Interreg Central Europe Programme, or its managing and national authorities, which bear no responsibility for any use that may be made of the information contained in this document.

The contents reflect the state of work at the time of preparation and remain subject to revision as the project progresses. Any references to the OPTI-UP toolkit and its outputs are subject to the limitations described in Section 6 of this document, which users are strongly encouraged to read before drawing conclusions on the basis of information presented here. The toolkit and accompanying documentation are provided "as is", without warranty of any kind. The authors and project partners accept no liability for any consequences arising from the use of this document or the toolkit outputs it describes.

# 1. Overview

## 1.1. General

The OPTI-UP project aims to improve the planning, optimization and sustainability of public transport (PT) systems in small and medium-sized cities in Central Europe.

The project combines analytical work, pilot actions and data-driven approaches in order to develop practical tools that support public authorities and transport planners in improving public transport systems.

One of the key outcomes of the project is the development of the OPTI-UP toolkit, a decision-support instrument designed to assist stakeholders in analyzing public transport conditions and identifying appropriate optimization measures.

The OPTI-UP toolkit has been designed as a transferable project output that extends beyond the immediate needs of the partner cities. While it is based on the experience gained through OPTI-UP pilot actions, the challenges addressed by the toolkit are common to many small and medium-sized cities across Central Europe and wider European regions: limited technical capacity, constrained financial resources, dispersed travel demand, fragmented service provision and the need to improve public transport in line with sustainability objectives.

By converting project experience and expert knowledge into a structured decision-support tool, the toolkit enables other municipalities to carry out an initial assessment of their public transport context and to explore potentially suitable improvement strategies. This directly supports the Interreg logic of knowledge transfer, capacity building and wider uptake of tested solutions. The tool helps ensure that the results of OPTI-UP are not limited to individual pilot locations, but remain usable for future planning processes in cities facing comparable mobility challenges.

The toolkit should therefore be understood as a practical entry point for municipalities that wish to identify relevant public transport improvement pathways before undertaking more detailed studies, modelling or feasibility assessments. In this sense, it contributes to the durability, transferability and transnational added value of the project results.

The development of the toolkit is based on the results of the previous project activities, including analytical work and pilot actions implemented in several partner cities. The toolkit integrates lessons learned from these activities and transforms them into a structured tool that can support public transport planning and optimization. To complement the information from the previous project activities, the toolkit also uses 200 additional expert case assessments as benchmark data, which are collected as part of the toolkit development process.

This document presents **Deliverable D3.1.1: Report on the Development Process of the Toolkit**, which describes the process of toolkit creation, its structure and functionality, as well as the preparation of training materials supporting its practical use.

## 1.2. Purpose and scope of the OPTI-UP toolkit

The purpose of the OPTI-UP toolkit is to provide a user-friendly toolkit and related documentation and guidelines that supports decision-making processes related to public transport planning and optimisation in small and medium-sized cities.

The toolkit is designed to support stakeholders in small- and medium-sized cities (users) in exploring potential public transport solutions. Based on user inputs, it provides indicative recommendations on which approaches may be worth considering, along with a brief overview of the relevant context. Please note: this toolkit does not replace detailed local planning or transport modelling. Its suggestions should be treated as a starting point for further investigation, not as definitive guidance. Any final decisions on introducing or improving public transport should be informed by thorough local analysis and professional transport expertise.

The toolkit builds on the expert rules collected during the WP3 implementation and on the experiences gained from the pilot actions implemented by project partners. These experiences provide practical insights that contribute to the development and validation of the toolkit.

The toolkit is intended to be used primarily by:

- public authorities
- transport planners
- regional mobility organisations
- experts involved in public transport planning and management

A screenshot of the toolkit's interface is given in Figure 1. The following chapters describe the individual steps involved in the development of the toolkit, including the conceptual design, technical development, testing, implementation and preparation of training materials.

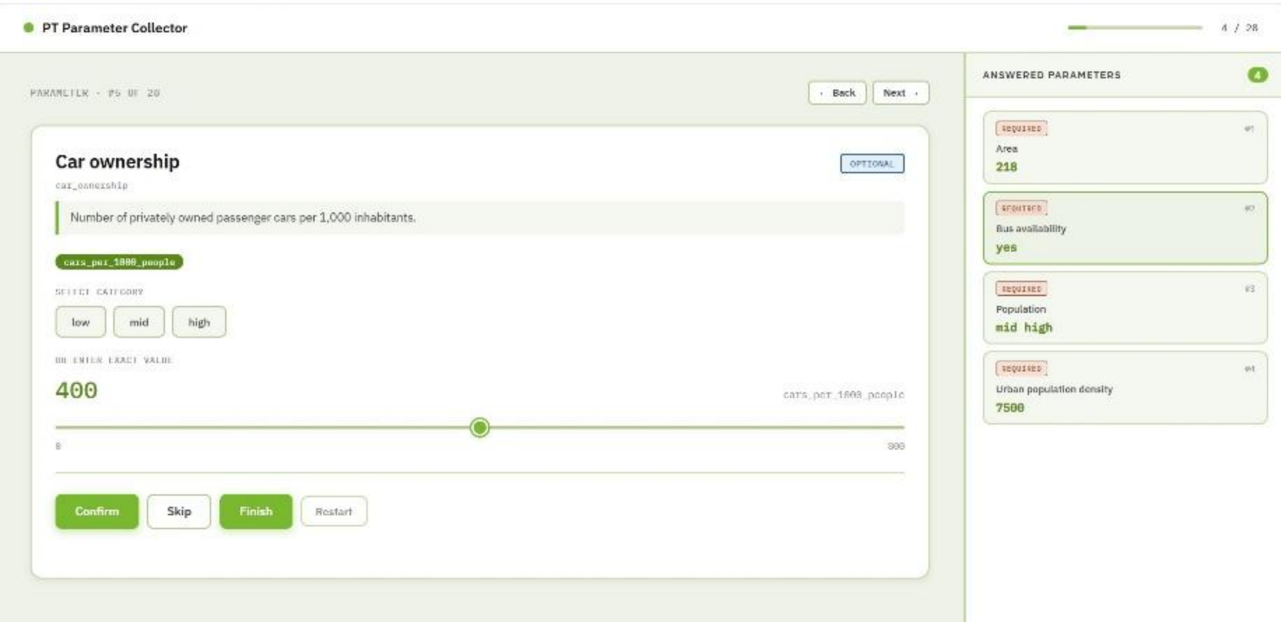


Figure 1: Screenshot of the toolkit interface.

## 1.3. Document structure

This document is structured into seven main chapters and two appendices.

Chapter 1 provides an overview of the OPTI-UP toolkit, including its purpose, scope and intended users. Chapter 2 describes the background and context of public transport optimisation challenges in small and medium-sized cities and presents the improvement strategies considered within the toolkit.

Chapter 3 explains the conceptual foundation and technical development of the toolkit, including the methodology applied, variable definitions, expert knowledge encoding and recommendation logic. Chapter 4 presents the practical implementation of the toolkit and its main functionalities.

Chapter 5 describes the testing activities, stakeholder involvement, feedback collection and corrective actions undertaken during the validation process.

Chapter 6 outlines the limitations of the toolkit and provides the necessary disclaimer regarding the interpretation and use of its outputs. Finally, Chapter 7 summarises the current status of the toolkit and presents recommendations for future development, maintenance and uptake by municipalities.

Appendix A contains the complete list of variables used in the toolkit; Appendix B shows a summary of the expert case coverage.

# 2. Background and Context

## 2.1. Problem statement

Public transport systems in small and medium-sized cities across Central Europe face a range of common challenges that affect their efficiency, attractiveness and long-term sustainability. While the specific local context differs between cities, many municipalities experience similar difficulties related to demographic change, urban development patterns, limited financial resources and increasing expectations regarding sustainable mobility.

Compared to larger metropolitan areas, small and medium-sized cities often have lower population densities and more dispersed travel patterns, making it difficult to provide frequent and financially efficient public transport services. In many cases, public transport networks have developed incrementally over time and may no longer fully reflect current mobility needs. As a result, municipalities frequently face declining ridership, increasing operational costs and difficulties in maintaining service quality.

At the same time, local authorities are expected to **contribute to broader environmental and climate objectives through the promotion of sustainable transport solutions**. This requires informed decision-making regarding potential measures such as network optimisation, demand-responsive transport services, fleet modernisation or the introduction of new public transport services. However, identifying the most appropriate intervention is often a complex task.

One of the main barriers encountered by municipalities is the **limited availability of technical expertise, detailed transport data and analytical resources**. Comprehensive transport modelling, feasibility studies and advanced planning tools are often costly and require specialised knowledge or specific data or resources to support the modelling and monitoring processes, all of which may not be readily available within local administrations. Consequently, without the support of personnel with the necessary technical background, decision-makers may find it difficult to assess which public transport improvement measures are most suitable for their local conditions and development objectives.

The experiences gathered through the OPTI-UP project partners and pilot activities confirmed that there is a need for a practical and user-friendly decision-support tool that can assist municipalities during the initial stages of public transport planning and optimisation. Such a tool should help users evaluate local conditions, explore potential improvement strategies and identify relevant examples of implementation without requiring extensive technical expertise or complex modelling procedures.

The OPTI-UP toolkit was therefore developed to address this gap. By combining **expert knowledge**, **experiences from pilot actions** and **a structured evaluation methodology**, the toolkit provides indicative recommendations that support the identification of potentially suitable public transport strategies for small and medium-sized cities. The toolkit is intended to facilitate informed discussions and preliminary assessments, while recognising that detailed planning and final decision-making must remain based on comprehensive local analyses and professional expertise.

## 2.2. Overview of PT improvement strategies

Public transport systems can be improved through a wide range of operational, organisational and technological measures. The selection of an appropriate strategy depends on local conditions, including population size and density, settlement structure, travel demand patterns, existing transport services, financial resources and policy objectives. As these factors vary considerably between cities, there is no universally applicable solution for public transport development.

Based on the analytical work conducted within the OPTI-UP project and the experiences gathered through pilot activities, the toolkit focuses on a selected set of public transport improvement strategies that are particularly relevant for small and medium-sized cities. These strategies represent practical and scalable approaches that can be considered by municipalities seeking to improve the accessibility, efficiency and sustainability of their transport systems.

The first strategy considered is **Demand-Responsive Transport (DRT)**. DRT services provide flexible transport solutions that adapt to passenger demand rather than operating according to fixed routes and timetables. Such services may be particularly suitable in areas with low population density, dispersed settlements or limited passenger demand, where conventional fixed-route public transport may not be economically viable.

The second strategy is **public transport network optimisation**. This approach focuses on improving the structure and performance of existing transport services through adjustments to routes, timetables, service frequencies and network connectivity. Network optimisation aims to better align transport supply with travel demand while making more efficient use of available resources. It should be noted that DRT can also be considered a form of PT network optimisation in specific areas (e.g., low density or PT demand) to supplement inefficient fixed routes. Within OPTI-UP, they are treated as two strategies to ensure the recommendations remain operationally actionable. However, given the many overlapping considerations, cities receiving a recommendation for either strategy should also explore the other or hybrid approaches in their planning process.

The third strategy addressed by the toolkit is **bus fleet electrification**. The transition from conventional diesel-powered vehicles to electric buses can contribute to the reduction of greenhouse gas emissions, air pollution and noise levels. However, successful implementation requires consideration of factors such as operational requirements, charging infrastructure, vehicle range, investment costs and funding opportunities.

In addition to these improvement measures, the toolkit also considers situations where the **introduction of public transport services** may be relevant. This scenario may apply to municipalities or areas where organised public transport provision is currently limited or absent and where the establishment of a basic transport service could improve accessibility and social inclusion.

Finally, the toolkit includes situations in which **the available information does not allow a sufficiently reliable recommendation** to be generated. In such cases, the result should not be interpreted as a failure of the methodology, but rather as an indication that additional data collection, expert assessment or more detailed analysis may be required before an appropriate strategy can be identified.

It should be noted that the strategies included in the toolkit do not represent the full range of public transport improvement options available to municipalities. Other measures, such as Bus Rapid Transit (BRT), integrated mobility services, rail-based solutions, active mobility interventions or broader urban development policies, may also contribute significantly to sustainable mobility objectives. These approaches were not included within the scope of the current toolkit due to project constraints and the need to maintain a focused and practical decision-support framework.

The strategies presented in this section represent a curated selection based on the analytical work and pilot experiences conducted within the OPTI-UP project, and should not be read as an exhaustive typology of public transport improvement options. The inclusion or exclusion of a strategy from the toolkit reflects project scope and practical constraints rather than a judgement on its general merit or applicability. Users are encouraged to consider the toolkit's recommendations as a reference or starting point for more detailed planning processes, and to seek professional transport planning advice when evaluating strategies not covered by the toolkit or when local conditions suggest that alternative approaches may be more appropriate.

# 3. Toolkit Development

This section documents the development of the analytical engine underpinning the OPTI-UP PT strategy toolkit. Specifically, a fuzzy-interpolation multi-criteria decision-making (MCDM) system is adopted, which translates city-level input characteristics into ranked PT strategy recommendations and matching reference solutions from the six OPTI-UP pilot cities. It corresponds to Tasks 3.1.1-3.1.3 of Activity 3.1 and feeds into deliverables D.3.1.1–D.3.1.3.

## 3.1. Development timeline

The backend of toolkit was developed using Python programming language with specifications and user inputs formulated as JavaScript Object Notation (JSON) key-value pairs. The toolkit is iteratively developed over nine major versions (V1 through V9). The development of the toolkit benefited from LLM Claude Code, and verifications were carried out with test cases (pseudo city profiles). Through each version, the developing team holds consultations with the front-end developer and PT planning experts to progressively adjust the city characteristics variable set, expand the expert case bank, and check the validation logic in response to internal testing and partner feedback. Key figures from the development process are summarized below in Figure 2.

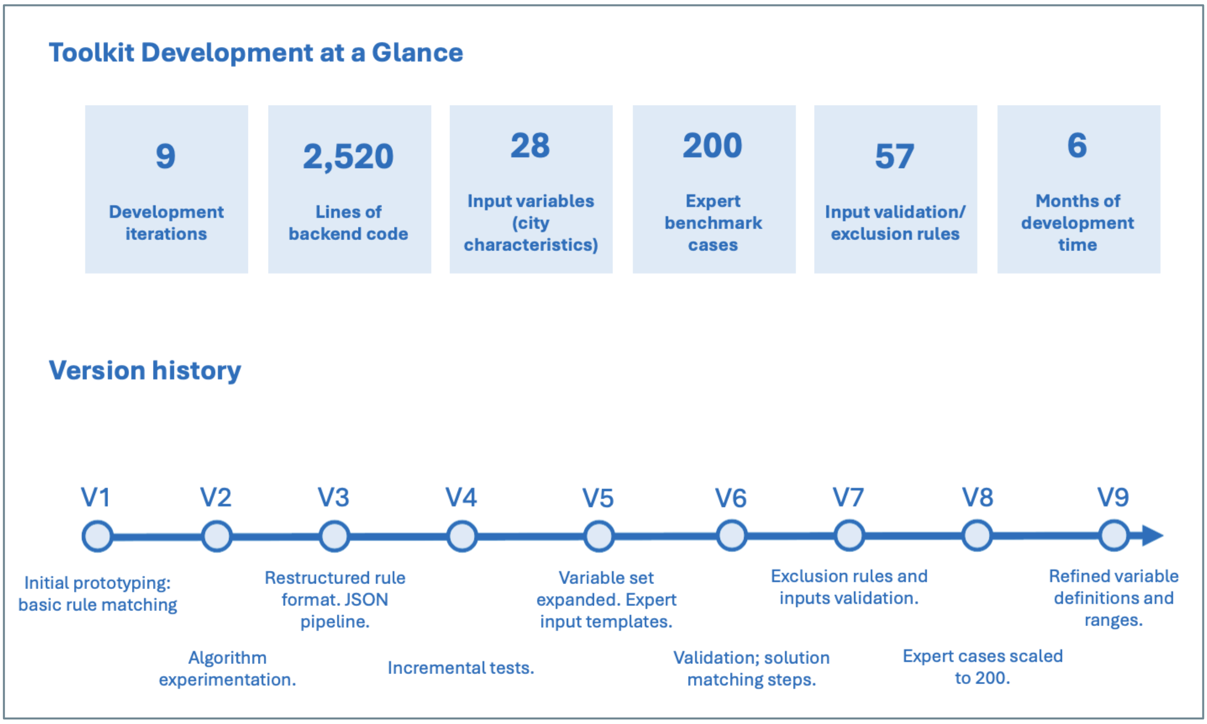


Figure 2. Toolkit overview and development timeline.

## 3.2. Conceptual Fondation

### 3.2.1. Review of alternative methodologies

Several families of multi-criteria decision-making methods were considered for the toolkit's backend recommendation engine before settling on a fuzzy rule-interpolation approach. The list below reflects the core practical limitations that the toolkit has to embrace:

Cities have highly diverse profiles. Following repeated evaluations, the OPTI-UP consortium identified 28 variables as necessary for PT planning (Appendix A). Since each variable may cover a wide range of possible values (quantitative) or categories (qualitative), it is not feasible to define rules that explicitly match all permutations of city profiles (basically rules out crisp rule-based expert system, AHP and TOPSIS methods in Table 1). The number of available expert-validated benchmark cases is necessary (as training data) but is limited (~200 cases, see Section 3.3c), so the chosen method must be capable of producing reliable recommendations even when a city's profile does not exactly match any stored case.

- Potential users of the toolkit (e.g., typically PT planner or decision makers in a small- or medium-sized city) typically have a good qualitative understanding of the local situations but may not have complete quantitative data at hands regarding the city's characteristics before a dedicated study is carried out (e.g., survey or other observations). This limits the suitability of methods that rely strictly on complete, precise numeric inputs and offer no mechanism for partial or qualitative information. Yet users' local insights should still be leveraged, e.g., using linguistic term expressions such as "the city has high population density", "the PT ridership in the city is usually low";

Table 1. Comparison of different decision-making methods.

| Method | Pros | Cons |
|---|---|---|
| Crisp rule-based expert system (IF-THEN) | Directly encodes expert reasoning in an auditable, interpretable form;<br>Each rule can be inspected and justified individually;<br>No statistical training required. | Brittle at category boundaries ("cliff-edge" effect);<br>Requires exact input matches;<br>Combinatorial explosion of rules needed to cover the full diversity of city profiles. |
| Machine learning / regression (data-driven) | Captures complex non-linear relationships with pretrained coefficients;<br>Improves with more data. | Requires large, labelled training datasets that do not exist for this domain (six pilot cities plus around 200 expert cases for 28 variables);<br>Low transparency / hard to justify to public-sector decision-makers;<br>Overfitting risk. |
| Weighted Sum / AHP | Simple and transparent;<br>Widely understood by non-technical stakeholders; | As a structured deliberation tool, it does not produce a reusable recommendation model. |
| TOPSIS / classic MCDM (distance-to-ideal) | Handles multiple criteria simultaneously;<br>Automatically normalises across different units and scales. | Requires complete input vector;<br>Produces a ranking of alternatives without drawing case-based expert knowledge. |
| Fuzzy rule interpolation<br>* selected approach | Interpolates between expert benchmark cases rather than requiring an exact match;<br>Handles both linguistic and numeric inputs natively;<br>Transparent and auditable scoring per rule. | Output quality remains dependent on the completeness and consistency of the underlying expert case bank.<br>Requires careful design of membership functions and aggregation logic; |
| Bayesian logistic regression with expert priors (based on the work of Nussbaumer, 2026) | Handles small datasets by encoding expert knowledge as priors rather than relying on data alone;<br>Captures complex decisions through prior distributions and updated posteriors. | Cannot natively accommodate linguistic or qualitative data without pre-processing;<br>Computationally more demanding than deterministic methods; |

| | | |
|---|---|---|
| | Provides full uncertainty quantification over predictions;<br><br>Supports explainability analysis. | Results and uncertainty estimates can be difficult to communicate to non-technical decision-makers. |

The comparison shows that none of the first four methods adequately address both constraints simultaneously: AHP and TOPSIS require complete numeric inputs; crisp rule-based systems require exact matches and are brittle at boundaries; and machine learning requires a volume of labelled data unavailable in this context. Fuzzy rule interpolation, on the other hand, can natively accommodates both qualitative and incomplete inputs while generalising across a limited expert case bank. It is therefore selected as the basis for the toolkit's recommendation engine, for the following reasons:

- Do not require large training data: unlike data-driven/ML approaches, fuzzy interpolation does not require a large training dataset, which is not available given the project's six pilot cities and not possible to obtain with the help of PT experts in limited time.
- Native support for linguistic and numeric data: city stakeholders can describe their city-specific characteristics context quantitatively (if they have such data) or qualitatively (if they do not have the quantitative data).
- Reuse of expert knowledge without requiring an exact match: interpolating between the nearest expert benchmark cases avoids the "cliff-edge" failure mode of crisp rule systems and allows the toolkit to function reliably even where a city's profile does not correspond exactly to any stored cases.

It's worth mentioning that Bayesain logistics regression, which has been explored in a student project at TU Wien (Nussbaumer, 2026), holds promise as a feasible methodology alternative. Since the development of the relevant technical aspects and evaluation process would take more time, it is highlighted as an alternative approach in D.3.1.1.

Beyond these properties of the method class itself, the OPTI-UP implementation adds further design choices to make the toolkit usable in practice, as described in Section 3.2.2.

### 3.2.2. Specific considerations in the OPTI-UP backend

The toolkit's scoring engine is conceptually related to Kóczy-Hirota (K-H) fuzzy rule interpolation (Kóczy and Hirota, 1993), which estimates a conclusion for an observation that does not exactly match any stored rule by interpolating between the closest surrounding rules in the antecedent (input) space. The OPTI-UP implementation adapts this idea to the project's practical constraints rather than applying the classical K-H formulation directly:

- Sparse, non-grid rule base: classical K-H interpolation assumes rules are arranged so that the two (or more) closest "flanking" rules can be identified along each variable axis. The OPTI-UP expert case bank (~200 collected cases evaluated by PT experts in the consortium) is not arranged on a regular grid across 28 variables (28 dimensions, with 2-5 possible levels per dimension), so the toolkit instead computes a weighted similarity between the user's case and every stored expert rule (~200 in total), rather than identifying a strict pair of flanking rules per axis.
- Per-variable membership similarity instead of geometric distance: each available user input variable contributes a membership-based similarity score (see Section 3.3.4(a)), which is then aggregated per rule (Section 3.3.4(b)) instead of computing a single multi-dimensional geometric distance as in the original K-H ratio method. For variables that are not provided by the user, this step is skipped and the coverage penalty is considered in the "coverage weighting" step.

Coverage weighting: The aggregated similarity is further adjusted by a coverage penalty that discounts rules whose specified conditions exceed what the user has provided, ensuring that recommendation confidence scales with input completeness rather than rewarding unverifiable rule claims (Section 3.3.4(c)). This is a project-specific addition not present in the classical interpolation method.

In short, the toolkit borrows the central idea of K–H interpolation: inferring a conclusion from the surrounding body of expert knowledge rather than requiring an exact rule match, while replacing the strict geometric interpolation procedure with a similarity- and coverage-weighted aggregation. This leads to the benefit that the toolkit can function robustly given the **sparse, irregular, high-dimensional expert rule base** available in this project.

Another related benefit of the modification is that the toolkit can also gracefully handle **use cases with incomplete city characteristics inputs**. For example, a user may ask for recommendations for their cities PT improvements without knowing the city's current fleet electrification rate. In this case, they may choose to skip providing an answer to this question. Such missing variables are handled at the rule-evaluation level (Section 3.3.4(a) returns no similarity contribution for unsupplied variables) and renormalizes each rule's score over only the variables the user actually provided. Only four variables are mandatory (population, area, density, bus availability) due to their high relevance in guiding the general directions of PT planning, strategy selection and solution recommendation, while they are also variables that are relative easy to obtain. The toolkit still produces a recommendation as fewer of the remaining 24 are supplied, though with less confidence.

The same fuzzy-similarity logic is reused first to rank PT strategies against expert rules (Section 3.3.4) and then to match a chosen strategy against specific reference solution variants (Section 3.3.5), rather than K–H interpolation being used for a single inference step.

## 3.3. Technical Framework

The overall technical framework of the toolkit is divided into four phases, where the first three phases are addressed in the backend algorithms. The framework is illustrated in Figure 3.

**Phase 1** is the collection of user inputs. The city's characteristics need to be supplied as 28 user input variables across five categories, including socio-demographics, PT supply, PT demand, fleet characteristics, and policy goals. The specific user input variables will be introduced in Section 3.3.1. These variables are eventually collected through a dynamic choice-based front-end interface, as introduced in Section 4.2.

- **Phase 2** recommends the general PT improvement strategies, including three strategies tested in OPTI-UP ("DRT", "bus electrification", "PT network optimization"), as well as "PT implementation" considering some small cities in Europe do not currently have PT services, and "None of the above" for situations where the best strategies are not one under the OPTI-UP study. The strategy recommendation is based on matching of user inputs and expert rules, with optional scripts provided for checking the validity of user inputs and exclusion rules to flag problematic inputs to the users. These are covered in Section 3.3.2 and Section 3.3.3.
- **Phase 3** recommends the specific solution tested in one of the six OPTI-UP case study cities (e.g., DRT implementation solution in Modena is different from Grosuplje due to the different city sizes and other considerations). If the recommended strategy is "PT Implementation" (Strategy label S4) or "No Recommendation" (Strategy label S5), for which no reference pilot case comparison applies, the toolkit skips Phase 3 and proceeds directly to Phase 4. This is mainly discussed in Section 3.3.5.
- **Phase 4** belongs to the frontend visualization step, which uses LLM to guide users through the technical outputs and helps readers to better understand the strategies to be adopted. The specific OPTI-UP tool visualization development is introduced in Section 4.

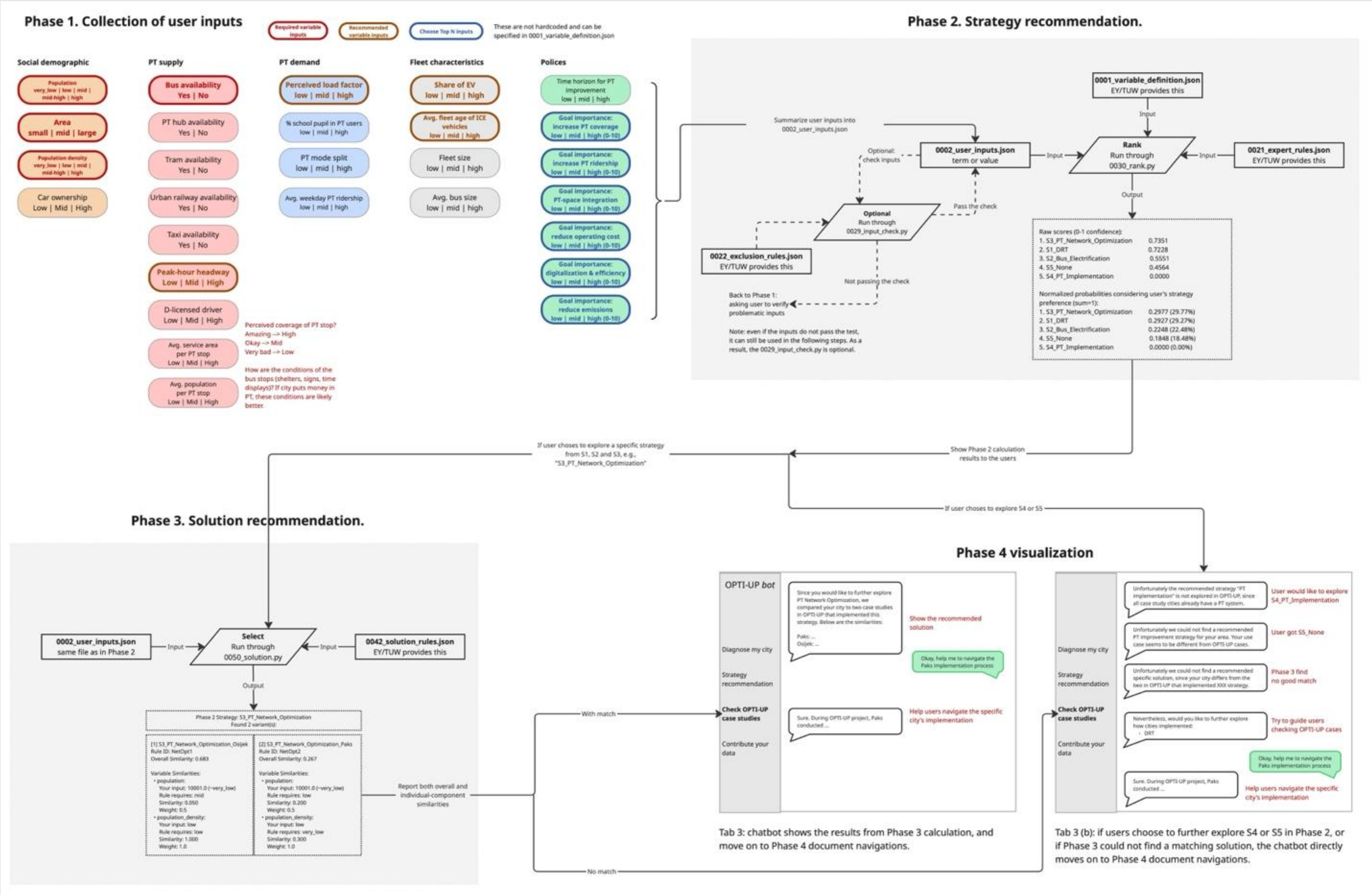


Figure 3. Overall processing pipeline of the toolkit, from variable definition to final recommendation output.

### 3.3.1. Variable definition

The user input variable set was narrowed down from a longer candidate list compiled jointly with the project partners during the early development versions (V1–V5), guided by three criteria: (i) relevance to the PT strategies that the toolkit can recommend, (ii) realistic data availability for small- and medium-sized European cities (population census data, transport-operator records, and open data such as OpenStreetMap stop locations), and (iii) the goal of keeping mandatory inputs minimal so that cities with limited data can still receive a recommendation.

The current configuration defines 28 variables across five categories. each variable tagged as required, recommended, or "rank the top" (e.g. the city's top policy priorities) categories. The full configuration (list of variables, their linguistic terms, numerical ranges, and types) is stored as JSON (0001_variable_definition.json) rather than hardcoded, so the variables and associated contents can be adjusted without changing the underlying code, allowing the variable set to evolve after the pilot phase.

Each variable is defined either as a fuzzy triangular type (with linguistic terms mapped to overlapping numeric ranges) or as a crisp boolean (yes/no), depending on its nature. Only three example variables are shown in Table 2 as an illustration; the complete 28-variable table is provided in Appendix A.

Table 2. Example input variables.

| Category | Variable | Definition | Type | Linguistic term and Numerical range |
|---|---|---|---|---|
| Socio-demographics | Population | Total number of residents within the study area.<br>Unit: people | Linguistic (fuzzy);<br>required | "very_low": [-1; 5,000; 15,000]<br>"low": [10,000; 25,000; 50,000]<br>"mid": [40,000; 100,000; 200,000]<br>"mid_high": [150,000; 300,000; 500,000]<br>"high": [400,000; 1,000,000; 3,000,000] |
| PT supply | Bus availability | Indicates whether there is an organised local public bus or trolleybus service operating within the study area, with scheduled services serving urban public transport stops. | Boolean;<br>required | "no", "yes" |
| Policy goals & time horizon | Goal importance of increase PT coverage | Importance of increasing PT coverage as a policy goal.<br>Unit: 0 (less important) - 10 (more important) | Linguistic (fuzzy) | "low": [-0.1; 2; 4]<br>"mid": [2; 5; 8]<br>"high": [6; 8; 10] |

### 3.3.2. Input validation and exclusion rules

Before a user-input case is scored, the toolkit checks user inputs in different aspects. The validation include:

- Required variables: the four required variables are all present;
- Goal selections: the users indicate three high-priority goals;
- Input validity: all input variables have valid terms or numerical ranges;
- Exclusion rules: the user inputs are compared against 57 exclusion rules (0022_exclusion_rules.json) designed to flag combinations that are statistically or logically unrealistic. For example, a mid-to-high population in a small area combined with very low

population density, or a city reporting no bus service while also reporting a tram, urban railway, or PT hub, would be illogical combinations. Each rule specifies the offending condition, a severity level, a human-readable warning message, and guidance on how to correct the input.

Two examples of the backend output from the validity check are given in Figure 4. Specifically, Figure 4(a) shows that the user input passes all validation checks, i.e., all required variables are provided, the priority goals are selected, all inputs have valid terms or values, and no disallowed combinations are present. Figure 4(b) shows a contrary case, where the users select that there is no bus service in their area, while still provided information such as "peak-hour departures per line" and "perceived load factor".

The input validation step is optional but recommended: it helps catch data-entry errors before they silently distort the recommendation, while still allowing the subsequent ranking and solution steps to run even if validation is skipped or only partially passed.

It should be noted that in the current implementation of the toolkit frontend, the validity check results are implemented as user-interface constraints, rather than directly provided to the users to avoid confusions. Specifically, users must provide answers to all required variables before being able to continue to the next step, and user inputs are constrained by selection buttons and toggle bars to avoid invalid numerical or terms. The exclusion rules are not examined due to the limited time, but it can nevertheless be enabled in the future once this additional check is deemed necessary.

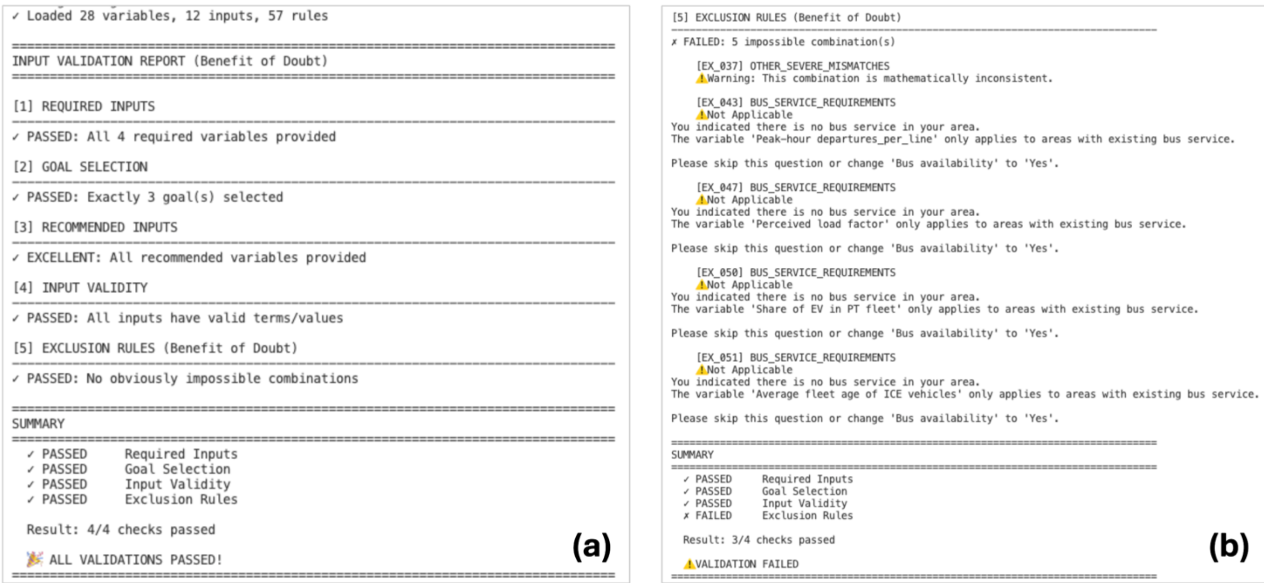


Figure 4. Example of input validity check. (a) All checks passed; (b) Inputs contain disallowed variable combinations.

### 3.3.3. Expert knowledge encoding

The toolkit's recommendations are grounded in a bank of 200 expert-defined benchmark cases, jointly developed and reviewed by members of the project team (mainly by three experts from EY). Each expert is given 50-100 input variable combinations (with a minimum of 7 variables and a maximum of 15 variables specified per case), and they are asked to evaluate the most suitable PT improvement strategies for a city with these characteristics. The 200 collected cases have been converted into the structured fuzzy-rule format used by the ranking engine (0021_expert_rules.json). This expert case bank is the toolkit's equivalent of a training set in a data-driven model, but consistent with the conceptual choice described in Section 3.2, it is used through interpolation and similarity matching rather than statistical model fitting or direct look-up, both of which would require significantly more expert cases to cover a parameter space of 28 dimensions (28 variables).

Figure 5 provides a screenshot of the expert rules, where the column “Recommended strategy” is provided by the experts. A summary of the frequency of the appearance of each variable in all expert cases is given in Appendix B.

| 1 | Case ID | Template | Recommended strategy | Var 1 | Level 1 | Weight 1 | Var 2 | Level 2 | Weight 2 | Var 3 | Level 3 | Weight 3 | Var 4 | Level 4 | Weight 4 | Var 5 | Level 5 | Weight 5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | C1 | social_dem | S3_PT_Network_Optimization | population | low | 1 | area | small | 1 | population_ | mid_high | 1 | bus_availab | yes | 1 | peak_hour_departures_ | low | 1 |
| 3 | C2 | social_dem | S3_PT_Network_Optimization | population | mid_high | 1 | area | large | 1 | population_ | mid_high | 1 | bus_availab | yes | 1 | peak_hour_departures_ | low | 1 |
| 4 | C3 | fleet_chara | S2_Bus_Electrification | population | mid | 1 | area | small | 1 | population_ | mid | 1 | bus_availab | yes | 1 | peak_hour_departures_ | mid | 1 |
| 5 | C4 | fleet_chara | S4_PT_Implementation | population | mid | 1 | area | small | 1 | population_ | mid_high | 1 | bus_availab | no | 1 | | | |
| 6 | C5 | policies | S2_Bus_Electrification | population | high | 1 | area | mid | 1 | population_ | high | 1 | bus_availab | yes | 1 | peak_hour_departures_ | high | 1 |
| 7 | C6 | pt_demand | S2_Bus_Electrification | population | mid | 1 | area | small | 1 | population_ | mid | 1 | bus_availab | yes | 1 | peak_hour_departures_ | high | 1 |
| 8 | C7 | pt_supply | S4_PT_Implementation | population | mid | 1 | area | small | 1 | population_ | high | 1 | bus_availab | no | 1 | | | |
| 9 | C8 | fleet_chara | S3_PT_Network_Optimization | population | mid_high | 1 | area | mid | 1 | population_ | mid_high | 1 | bus_availab | yes | 1 | peak_hour_departures_ | mid | 1 |
| 10 | C9 | policies | S4_PT_Implementation | population | mid_high | 1 | area | large | 1 | population_ | low | 1 | bus_availab | no | 1 | | | |
| 11 | C10 | social_dem | S3_PT_Network_Optimization | population | mid_high | 1 | area | large | 1 | population_ | low | 1 | bus_availab | yes | 1 | peak_hour_departures_ | mid | 1 |
| 12 | C11 | social_dem | S3_PT_Network_Optimization | population | mid | 1 | area | small | 1 | population_ | high | 1 | bus_availab | yes | 1 | peak_hour_departures_ | mid | 1 |
| 13 | C12 | pt_demand | S4_PT_Implementation | population | mid_high | 1 | area | large | 1 | population_ | mid_high | 1 | bus_availab | no | 1 | | | |
| 14 | C13 | pt_supply | S1_DRT | population | very_low | 1 | area | small | 1 | population_ | low | 1 | bus_availab | yes | 1 | peak_hour_departures_ | mid | 1 |
| 15 | C14 | fleet_chara | S2_Bus_Electrification | population | low | 1 | area | small | 1 | population_ | low | 1 | bus_availab | yes | 1 | peak_hour_departures_ | mid | 1 |
| 16 | C15 | policies | S1_DRT | population | low | 1 | area | mid | 1 | population_ | very_low | 1 | bus_availab | yes | 1 | peak_hour_departures_ | low | 1 |
| 17 | C16 | pt_supply | S3_PT_Network_Optimization | population | mid_high | 1 | area | mid | 1 | population_ | mid | 1 | bus_availab | yes | 1 | peak_hour_departures_ | high | 1 |
| 18 | C17 | social_dem | S3_PT_Network_Optimization | population | mid_high | 1 | area | large | 1 | population_ | mid_high | 1 | bus_availab | yes | 1 | peak_hour_departures_ | low | 1 |
| 19 | C18 | pt_supply | S3_PT_Network_Optimization | population | high | 1 | area | mid | 1 | population_ | high | 1 | bus_availab | yes | 1 | peak_hour_departures_ | low | 1 |
| 20 | C19 | social_dem | S3_PT_Network_Optimization | population | mid_high | 1 | area | small | 1 | population_ | high | 1 | bus_availab | yes | 1 | peak_hour_departures_ | high | 1 |
| 21 | C20 | pt_supply | S4_PT_Implementation | population | high | 1 | area | large | 1 | population_ | mid_high | 1 | bus_availab | no | 1 | | | |

Figure 5. Screenshot of expert rule bank (first 20 rules and first 5 variables).

### 3.3.4. Strategy recommendation

Given a (possibly incomplete) set of user inputs, the toolkit narrows down on one of five possible outcomes as recommendations for PT improvement strategies:

- **S1 Demand-Responsive Transport (DRT):** A flexible, on-demand service using smaller vehicles, particularly suited to areas where conventional fixed-route PT is inefficient or absent. Typical use cases include sparsely populated areas with low or unpredictable demand, zones with gaps in existing PT coverage, and routes where fixed schedules cannot be justified economically.
- **S2 Bus/Fleet Electrification:** Transitioning existing ICE vehicles to electric ones in the PT fleet, ideally timed to coincide with scheduled fleet replacement cycles.
- **S3 PT Network Optimization:** Restructuring existing routes and adjusting service frequencies or timetables to improve efficiency. This is one of the most commonly applicable measures for cities with an existing but inefficient PT networks or services.
- **S4 New PT Implementation:** Launching a new fixed-route bus service from scratch, suited to areas with sufficient population density to sustain regular scheduled services but currently does no have existing PT provision.
- **S5 No Recommendation:** the area is not suitable for any of the above intervention, or the available expert rules do not support a confident recommendation.

This strategy identification process is achieved by scoring the user’s case against every active expert rule and ranking the resulting strategy scores, rather than by following a fixed decision tree. The adopted method allows multiple strategies to be shown as viable alternatives when the evidence is mixed.

#### (a) Membership function

Numeric variables are converted into **fuzzy linguistic terms** using triangular membership functions with deliberate overlap between adjacent terms (for example, "small" and "mid" areas share a transition zone rather than meeting at a hard threshold, as shown in Figure 6). This avoids artificial discontinuities at category boundaries: a city with an area of 60 km² receives partial membership in both "small" and "mid", meaning that expert rules specifying either term can both contribute a non-zero similarity score, weighted by the degree of membership, rather than one category being entirely excluded. Each variable can have between two and five linguistic terms depending on how much differentiation is meaningful for that variable.

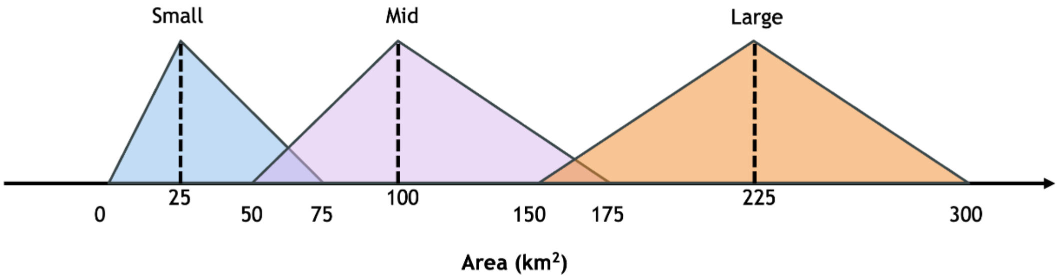


Figure 6. Example of fuzzification through the triangular membership function: for variable Area (km2).

### (b) Condition aggregation methods

Each expert rule typically specifies conditions across several variables simultaneously (e.g. low population AND small area AND mid-high density AND …). The similarity between a user's case and a rule across all of its conditions can be aggregated as a **similarity score** in two configurable ways: (1) a minimum ("fuzzy AND") aggregation, where the weakest-matching condition dominates the rule's overall score (the strict setting for strategy ranking); or (2) an average aggregation, where all condition similarities are weighted and averaged (a more lenient setting used for solution-variant matching). The choice of aggregation method is a configuration parameter (0000_config.json) rather than hard-coded, allowing it to be tuned as the toolkit is validated against pilot-city feedback.

### (c) Coverage-weighted score

Each expert rule specifies between 7 and 15 conditions (median 12, out of 28 variables). When a user provides only the four mandatory variables, a typical rule will have the majority of its conditions unverifiable. Without correction, it could still score well on the four matched conditions, producing an artificially confident recommendation. To address this, each rule's similarity score is multiplied by a coverage penalty.

The penalty is based on the ratio of the total weight of the rule's conditions that the user has actually provided ("*provided_weight*") to the total weight of all conditions the rule specifies ("*total_weight*"). This ratio, raised to a configurable exponent (default 1.5), is multiplied into the rule's similarity score:

A well-covered rule (ratio close to 1.0) receives little penalty; a poorly covered rule is discounted substantially. The result is an absolute confidence score per strategy (0–1), which can be normalised across all five strategies to a probability distribution summing to 1 if so desired.

## 3.3.5. Solution recommendation

Once the user is given a recommended strategy, the toolkit moves to Phase 3 (0050_solution.py), which reuses the same fuzzy-similarity machinery to narrow down on a specific reference pilot case from the OPTI-UP consortium that is associated with that strategy. This step compares the user's full input profile against the characteristics of each candidate reference solution and returns an overall similarity score together with a variable-by-variable breakdown, so that users can see precisely where their city resembles (or diverges from) the closest matching real-world example. For example, in Figure 7, the user indicates that they are looking for a particular for electrification solution for a mid-sized, mid-density city. This profile matches better with the profile of Pécs in the OPTI-UP pilot cities (similarity score of 0.653). In the subsequent steps of the interactive toolkit, the users will be recommended to refer to the Pécs local plan.

```
================================================================================
SOLUTION VARIANT COMPARISON
================================================================================

Phase 2 Strategy: S2_Bus_Electrification

Found 2 variant(s):

────────────────────────────────────────────────────────────────────────────────
[1] S2_Bus_Electrification_Pecs
────────────────────────────────────────────────────────────────────────────────
Rule ID: EV1
Overall Similarity: 0.653

Variable Similarities:
  • population:
      Your input: mid
      Rule requires: mid
      Similarity: 1.000
      Weight: 0.5
  • population_density:
      Your input: mid_high
      Rule requires: mid
      Similarity: 0.480
      Weight: 1.0

────────────────────────────────────────────────────────────────────────────────
[2] S2_Bus_Electrification_Cesky_Krumlov
────────────────────────────────────────────────────────────────────────────────
Rule ID: EV2
Overall Similarity: 0.150

Variable Similarities:
  • population:
      Your input: mid
      Rule requires: very_low
      Similarity: 0.050
      Weight: 0.5
  • population_density:
      Your input: mid_high
      Rule requires: low
      Similarity: 0.200
      Weight: 1.0
```

Figure 7. Example output from the solution recommendation script.

This final step is intended to ground the toolkit's strategic guidance in concrete, replicable precedents rather than abstract category labels alone. In the subsequent visualization frontend, the users will be given the opportunities to ask questions regarding the experiences from the pilot case, assisted by the LLM to answer specific questions from the user based on the pilot study reports from WP2 of the OPTI-UP project (OPTI-UP, 2026).

# 4. Toolkit development

## 4.1. Development Process

The development of the toolkit was conceived as a comprehensive, iterative process aimed at transforming expert knowledge in the field of urban public transport into a digital decision-support tool. The entire process was divided into several key phases, each addressing a specific aspect of the user experience and technical implementation.

The first step involved an extensive requirements analysis. During this phase, it was necessary to identify the parameters that are critical for selecting an appropriate transport strategy. Collaboration with domain experts resulted in the creation of a structured set of variables, which are now stored in the 0001_variable_definition.json file. For more details, readers can refer to Section 3.3.1. It was also decided that the system should be web-based to ensure easy accessibility and sharing among public authorities and consultants. As part of this strategy, the domain name mivoma.cz was registered and serves as the central hub for hosting the application and providing stable access to the user interface.

The next phase focused on designing the system architecture. The system was divided into four logical stages (Pipeline), ensuring modularity and maintainability. The first stage is dedicated to data collection (Parameter Collector), the second to strategy ranking using fuzzy logic, the third to detailed comparison of alternative solutions, and the fourth to the integration of generative AI for the interpretation of the results. The details of each step are introduced in Section 3.3.

The implementation phase followed an agile development approach. First, the backend was developed in Python, followed by the implementation of a web-based graphical user interface using the Flask framework. Particular attention was devoted to integrating the **Retrieval-Augmented Generation (RAG)** pipeline, which required the creation of a vector database from the existing project documentation. The final implementation step involved optimizing prompts for the **large language model (LLM)** to ensure that the generated responses were both understandable and factually accurate.

In the final stage of development, the toolkit was deployed to the production environment, and its web infrastructure was reorganized. The application itself was moved to the toolkit.opti-up.net subdomain, which is dedicated exclusively to hosting the Toolkit. The original opti-up.net domain was subsequently transformed into an informational landing page serving as the primary entry point for users. The landing page provides essential information about the project, its objectives, and its potential applications while also offering direct access to the Toolkit through the dedicated subdomain. This separation improved the organization of the project's web presence, enhanced usability, and simplified the future maintenance and further development of the overall system.

## 4.2. Main Toolkit Functionalities

The toolkit represents a synergy between deterministic rules and probabilistic intelligence. Its main functionalities include:

1. Dynamic Parameter Collection: The system does not operate as a static questionnaire but as an adaptive parameter collection tool. The ParameterCollector class enables the definition of both mandatory and optional parameters. Support for fuzzy terms (e.g., low, medium, high) allows users to provide input intuitively without requiring precise numerical values in every case.An example of the parameter collection interface is shown in Figure 8.

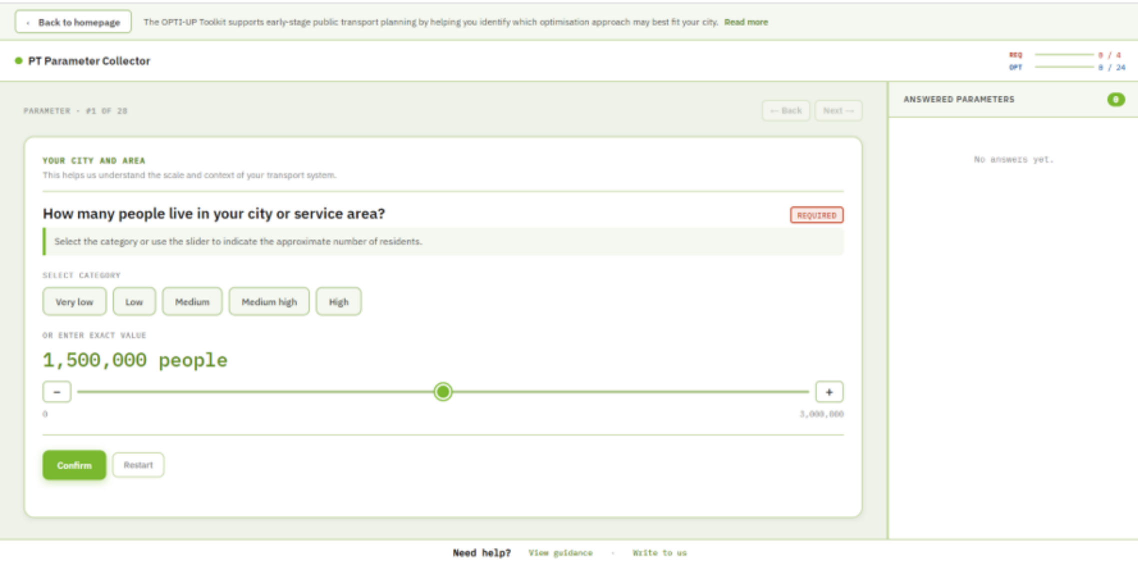


Figure 8. Parameter collection interface of the OPTI-UP Toolkit. Users can provide input using predefined categories or exact numerical values for each transport-related parameter.

2. Fuzzy Strategy Ranking: The core of the system is the inference engine implemented in the opti_rank2 module. It employs triangular membership functions to determine the degree of similarity between the user's profile and the expert-defined rules. The output is a ranked list of strategies with probabilistic similarity scores, providing a transparent basis for decision-making. An example of the fuzzy strategy ranking together with its AI-generated interpretation is presented in Figure 9.

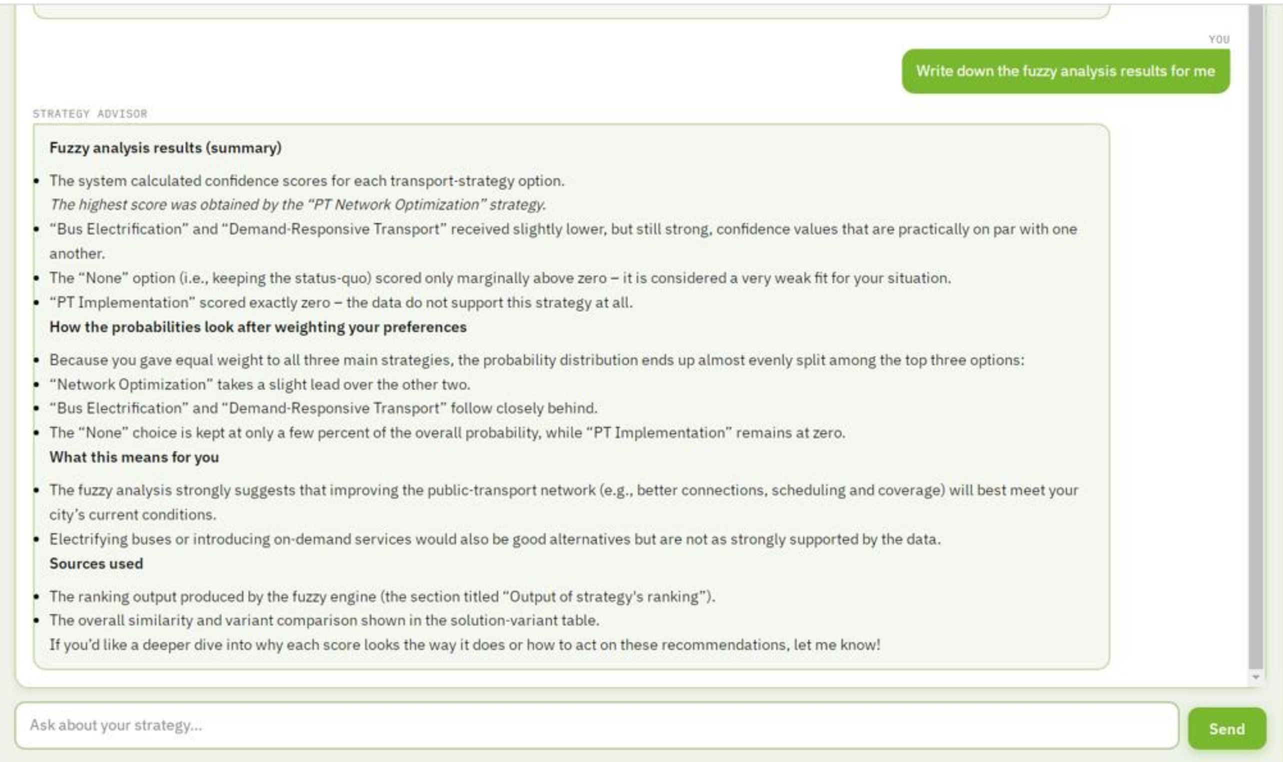


Figure 9. Fuzzy strategy ranking results and AI-generated interpretation provided by the OPTI-UP Toolkit

3. Analysis of Solution Alternatives: Rather than stopping at a general strategy recommendation, the system performs detailed matching against specific solution alternatives within the opti_solution2 module. This enables differentiation between, for example, various electrification approaches or specific public transport prioritization models. Figure 10 illustrates the detailed recommendation generated by the toolkit, including the selected implementation alternative, justification, and matching pilot example.

STRATEGY ADVISOR

STRATEGY ADVISOR

**Chosen strategy**

- The system ranked the **Bus Electrification** option as the most suitable overall.
- This choice came after weighing your city's characteristics—medium-sized population, medium density, and existing bus services—and comparing them against all available rules.
- Because Bus Electrification scored highest confidence among the three user-chosen strategies (S1 = Dynamic Routing Tool, S2 = Bus Electrification, S3 = Public-Transport Network Optimisation), it was selected as the primary recommendation.

**Why this strategy fits your context**

- Your city has a fairly large number of people but does not yet have a very dense network.
- Bus services are already present and would benefit from cleaner fuel technology.
- The best-practice analysis highlighted that cities with similar conditions see notable gains when switching to electric buses—this is also reflected in the related Best Practice survey section that focuses on alternative fuel technologies (see Section 7.2).

**Pilot experiment that matches your data**

- Two pilot variants were evaluated for Bus Electrification:
  *Variant 1* – "Pecs"
  *Variant 2* – "Cesky Krumlov"
- Variant 1 shows the closest match to your city's profile:
- Population size is treated as **high** in the rule, which matches your figure.
- Population density requirement is **mid**, identical to your input.
- The overall similarity score for this variant is significantly higher than that of Variant 2.
- Therefore, the **Pecs pilot experiment** is recommended as the best fit for your city's data.

**Where to find more detail**

- The ranking and probabilities are listed in the "Output of strategy's ranking" section.
- Pilot variants and their similarity details appear under the "Recommended solution – Solution Variant Comparison".
- Additional context about bus electrification can be found in the Best Practice survey's discussion on alternative fuel technologies (Section 7.2).

Feel free to review these sections for deeper insight or let me know if you'd like a walk-through of how each rule contributed to the final decision.

Figure 10. Example of a recommended implementation alternative with explanation and matching pilot case generated by the OPTI-UP Toolkit.

4. Intelligent AI Assistant (RAG): The integration of the Qwen/Llama model via the Ollama server enables users to interact with the toolkit using natural language queries. Instead of relying solely on the model's internal knowledge, the AI primarily retrieves relevant passages from the connected documentation (Retrieval), significantly reducing the risk of hallucinations while improving the factual accuracy and domain relevance of the generated responses. An example of an interactive conversation with the AI assistant, providing context-aware recommendations based on the toolkit results, is shown in Figure 11.

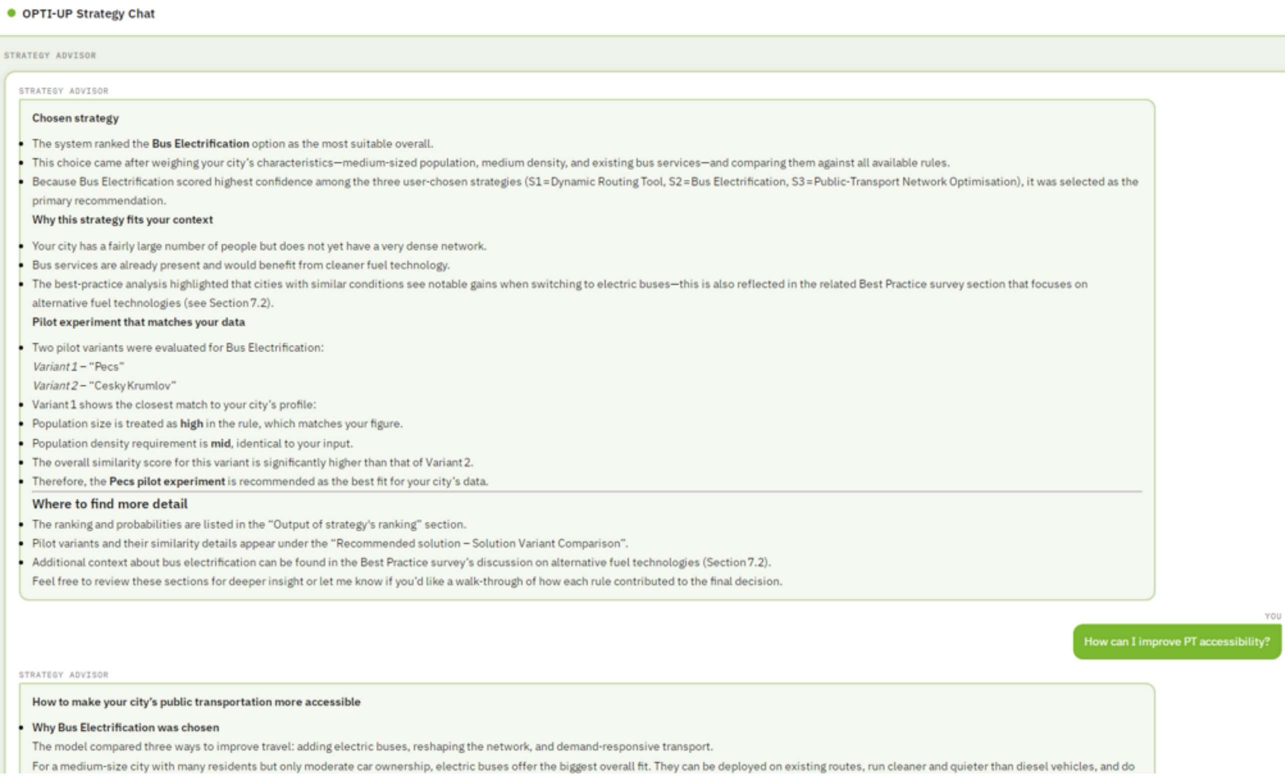


Figure 11. AI assistant providing context-aware recommendations and implementation guidance based on the toolkit results.

## 4.3. Technologies and Methods Used

To achieve high stability and accuracy, the following technology stack was selected:

- **Programming Language and Frameworks:** Python 3.11 was chosen due to its extensive ecosystem of libraries, including Pandas, NumPy, and Flask. Flask serves as a lightweight web framework that connects the frontend with the backend processing logic.
- **Fuzzy Logic:** To process uncertain linguistic variables, the system implements fuzzy set theory. The weighted_average method and triangular membership functions are used to model gradual transitions between categories (e.g., from a small city to a large city), which more accurately reflects real-world urban planning scenarios than rigid threshold-based classification.
- **RAG Pipeline:** The RAG implementation includes document chunking, embedding generation, and storage within a vector database. Information retrieval is based on cosine similarity, ensuring that the AI model receives only the most relevant sections of the documentation when generating responses. As an input knowledge base the reports about the pilot plans of all 6 pilot cities (OPTI-UP, 2026) were used together with D 1.1.1 - Comprehensive project report (OPTI-UP, 2024).
- **LLM and Infrastructure:** The system utilizes a locally hosted Ollama server for running large language models. This approach was chosen to ensure data security and eliminate dependence on external APIs. The model was configured with a temperature setting of 0 to produce consistent, deterministic, and factually reliable responses. The reason for local model are also the costs and their predictability. Using the LLM server, we will be able to change the model in the future. Nowadays the gpt-oss:20b model is used (https://openai.com/open-models/)

- **Frontend:** The user interface was developed using HTML5, CSS3, and JavaScript, with a strong emphasis on responsiveness and an intuitive user flow.

## 4.4. Overall Solution Architecture

The toolkit architecture is designed as a sequential pipeline. The data flow begins in the ParameterCollector module, where the user profile (output.json) is created. This profile is then passed to the ranking engine (opti_rank2), which generates a ranked list of transport strategies. The highest-ranked strategy is subsequently processed by the opti_solution2 module to identify the most appropriate implementation alternative. Finally, all generated results, together with the original user inputs, are forwarded to the RAG pipeline (opti_rag3), which provides contextual information for the LLM-based chatbot. The overall implementation architecture and workflow of the toolkit are illustrated in Figure 12.

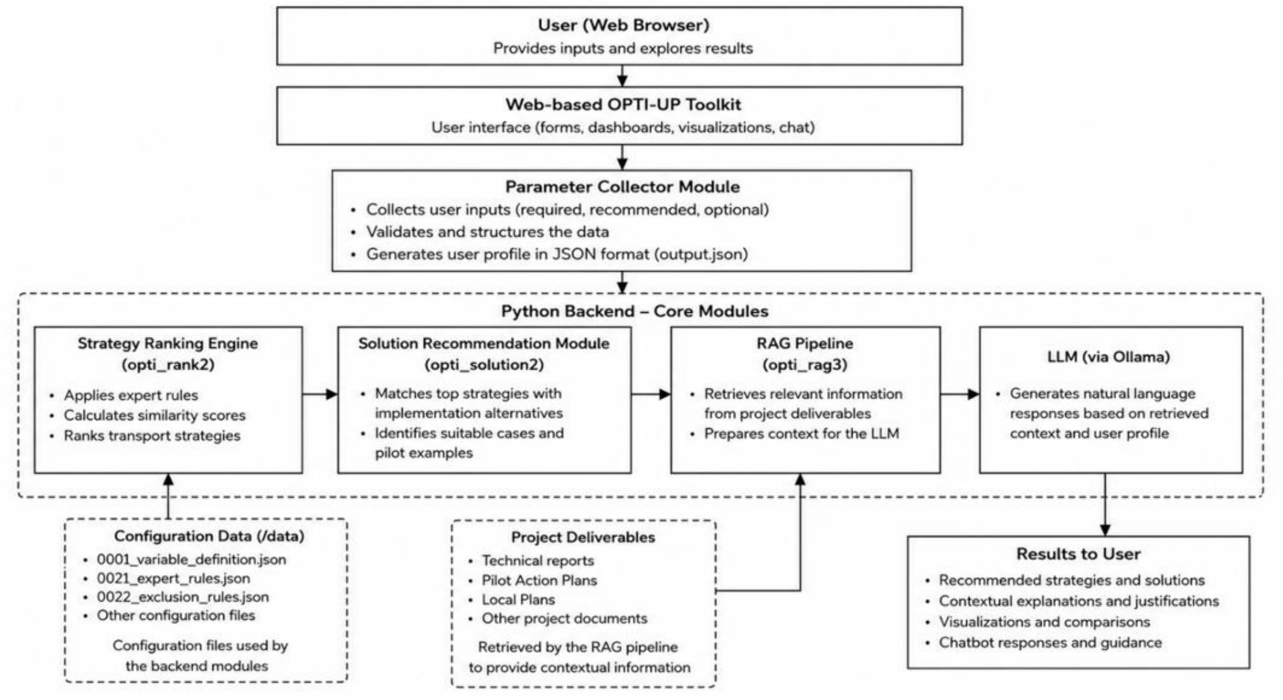


Figure 12. Overall implementation architecture and workflow of the OPTI-UP Toolkit.

The separation of configuration data from the application logic through JSON files stored in the /data directory enables expert rules and parameters to be updated without modifying the program code. This design significantly simplifies system maintenance and accelerates future updates and extensions.

## 4.5. Disclaimer Regarding the Use of AI

Section 4 of the system incorporates generative AI. Users are informed that the AI generates responses based on probabilistic language models. Although the implementation of the RAG pipeline grounds the generated responses in the underlying project documents, there remains a possibility of syntactically correct but factually inaccurate outputs. The AI assistant is intended to serve as a guide to the available documentation rather than as a final decision-making or approval authority. All recommendations generated by the AI module should therefore be verified against the original source documents, which are explicitly cited in the "Sources" section of every response.

# 5. oolkit testing

## 5.1. Testing and feedback procedure

Toolkit testing was carried out continuously throughout the entire development process using an iterative approach to verify the functionality of individual system components. The primary objective was to ensure the correct operation of the implemented algorithms, the stability of the application, and the overall usability of the toolkit for its intended users.

The toolkit was tested between April and June 2026. The testing phase lasted approximately three months and included several iterative evaluation cycles involving project partners. The collected feedback was systematically incorporated into the refinement of the toolkit before its final release.

The first phase of testing focused on verifying the correctness of individual modules. This included validating the functionality of parameter collection, data processing based on fuzzy logic, the calculation of strategy rankings, and the subsequent selection of the most appropriate solution alternatives. In addition, the correctness of the generated output files and their use in subsequent stages of the processing pipeline were verified.

The second phase focused on the integration of the system components. Particular attention was paid to the communication between the frontend and backend, the correct transfer of data between individual modules, and the functionality of the RAG pipeline in retrieving relevant information for the AI assistant. The stability of the web application was also evaluated under repeated use and with various combinations of input parameters.

Feedback was collected continuously from both project team members and project partners who used the toolkit to evaluate its individual functionalities. The collected comments and recommendations were regularly reviewed and incorporated into subsequent versions of the system, enabling continuous improvements in reliability, recommendation accuracy, and overall user experience.

## 5.2. Stakeholders involved and their feedback

Three main stakeholder groups were involved in the toolkit testing process: transport experts and urban planners, IT specialists, and potential end users. Each group evaluated the toolkit from a different perspective and provided feedback corresponding to their expertise and expected use of the system.

Table 3. summary of stakeholder feedback process.

| Stakeholder group | Number of stakeholders involved | Number of feedback items | Example feedback |
|---|---|---|---|
| Transport experts and urban planners | 5 | 28 | Clarify the scope of the analysis for municipalities served by regional transport; improve the wording of questions (e.g. railway, stop coverage, passenger capacity); refine parameter definitions and recommendation methodology. |
| IT specialists | 2 | 56 | Reduce the loading time of the first response; simplify the initial interface and remove unnecessary scroll bars; improve navigation and fix UI-related issues. |
| Potential end users | 5 | 24 | Add a descriptive landing page and Q&A page with toolkit objectives and instructions; provide clearer guidance for completing the questionnaire (e.g. population input and qualitative values); improve the step-by-step workflow and overall usability. |

Transport experts and urban planners primarily focused on the technical validity of the implemented rules and the recommended transport strategies. They assessed whether the system's decision-making logic reflected established practices in urban mobility and public transport planning. Their feedback resulted in the refinement of several expert rules, adjustments to the weighting of selected parameters, and the addition of further explanations for certain recommended strategies.

IT specialists evaluated the technical aspects of the application. They verified the functionality of individual modules, communication between the frontend and backend, the integration of the RAG pipeline, the operation of the vector database, and the overall stability of the application. Their recommendations also included performance optimizations, reducing the chatbot's response time, and resolving technical issues identified during testing.

Potential end users evaluated the toolkit primarily in terms of usability and intuitiveness. Most of their comments focused on the visual appearance of the application and the overall user experience. They recommended improving the clarity of the user interface, reducing the amount of displayed information, removing unnecessary interface elements, and providing clearer descriptions of individual input parameters. They also emphasized the need to clearly explain the expected input values, the meaning of the different categories (e.g., Low, Medium, High), and the interpretation of the generated results.

A significant portion of the feedback also concerned the AI chatbot. Test participants highlighted issues such as overly long introductory messages, extended loading times for the initial response, ambiguously phrased questions, language inconsistencies, and certain aspects of the conversation flow that were difficult to follow. They also recommended including direct links to relevant reference documents, improving the clarity of user communication, and providing a clearer separation between informational content and the parameter collection process.

The collected feedback was continuously evaluated and incorporated into subsequent stages of the Toolkit's development. The recommendations formed the basis for improvements to the user interface, optimization of the AI chatbot, refinement of the expert rules, and overall enhancements to the clarity, usability, and effectiveness of the system.

## 5.3. Corrective actions

Based on the testing results and the feedback provided by the different stakeholder groups, a series of improvements were implemented throughout the Toolkit's development to enhance its functionality, clarity, and overall user experience.

The most significant changes were made to the user interface. The layout of individual screens was redesigned, unnecessary interface elements were removed, and the arrangement of controls was improved to provide a clearer and more intuitive user experience. In addition, descriptions of individual parameters and explanatory texts were added to assist users in correctly completing the input forms. Several question formulations and parameter names were also revised to ensure they were easily understandable, even for users without a technical background. For a specific example:

- Old question formulation: Are there any major public transport hubs in the area?
- New question formulation: Is there a transport hub in the city offering regional, interregional, national, or international transfer?
- Substantial improvements were also introduced to the AI assistant module. The prompts used by the LLM were optimized, the structure of the introductory messages was revised, and the presentation of generated responses was improved to make them more concise, easier to follow, and better aligned with users' queries. At the same time, the RAG pipeline

was optimized to improve the relevance of retrieved information and enhance the overall quality of AI-generated responses. The examples of improvements based on testing feedback can be the modification of system prompt of the LLM to not use links to internal rules, be more user friendly, be more explainable, and so on. Based on feedback from domain experts, several expert parameters were refined and the rules used for the fuzzy evaluation of transport strategies were updated. These modifications improved the interpretation of input data and increased the accuracy of the recommendations generated by the system. An example of this refinement is given below:

- During iterative refinement, domain experts identified a semantic inconsistency in how service frequency was represented. The original variable “peak-hour headway” is an inverse measure of service quality (a higher value means less frequent service) which caused a mismatch when encoding expert rules.
- This variable was subsequently changed to “peak-hour departures per line”, then “average number of minutes between departures during peak hours” to make it more understandable by the users.

The corrective actions also included a number of technical improvements aimed at increasing application stability, resolving issues identified during testing, and optimizing communication between the frontend and backend. In RAG system, the most modifications were made in LLM system prompt to make the answers to the user mode friendly, well explained and understandable without technical details. All feedback was continuously incorporated into subsequent versions of the toolkit as part of the iterative development process.

# 6. Toolkit limitations and disclaimer

The OPTI-UP toolkit is intended to support, rather than replace, professional PT planning judgement. This section outlines the toolkit's known limitations, clarifies the interpretive status of its outputs, and describes the due diligence expected of users. It applies to the version of the toolkit described in this deliverable and is expected to be revisited as the toolkit develops further in the remaining time of the project duration.

(A) **Technical limitations**: The toolkit requires a minimum of four input variables (population, area, population density, and bus availability) and performs best when more variables are provided, as a richer input profile enables closer matching with the expert benchmark cases. Numeric variables are represented through overlapping triangular membership functions. Recommendations are generated by comparing user inputs against a finite bank of 200 expert-defined benchmark cases, compiled within the project's time and resource constraints and reflecting the perspectives of a limited expert panel; systematic gaps or inconsistencies in coverage cannot be excluded. Absolute confidence scores are further subject to a coverage penalty: users supplying only the mandatory variables should therefore expect lower scores, reflecting input incompleteness rather than an unfavourable assessment of the city's suitability.

(B) **Influencing factors:** The current variable set focuses primarily on observable, supply-side characteristics of the public transport system, such as service frequency, network coverage, fleet composition, and demographic context, together with some demand indicators such as mode split and ridership. Other aspects, including fare affordability, perceived safety, information accessibility, digital exclusion, and cultural or behavioural attitudes towards PT, are not represented as input variables. These factors can be significant determinants of ridership growth in small and medium-sized cities, where social norms and car-dependency may constrain demand independently of service quality. As the expert rule bank was elicited primarily from operational and planning perspectives, strategies addressing social equity, communication, or behaviour change may be underweighted in the current outputs. Future iterations of the toolkit should consider incorporating demand-side and socio-cultural variables, provided sufficient expert knowledge can be encoded to support reliable rule-based differentiation.

(C) **Nature of recommendations**: Toolkit outputs are similarity-based indications rather than prescriptions. Recommended strategies and reference solutions reflect analytical patterns drawn from the expert knowledge encoded in the system and scored using fuzzy-logic methods. They are not exhaustive and do not represent a complete ranking of all possible interventions; alternative approaches may be equally or more suitable depending on local factors the toolkit does not capture. A high similarity score reflects resemblance to expert-validated cases, not a guarantee of outcomes if a strategy is implemented.

(D) **Required due diligence**: The toolkit is best understood as a screening and orientation aid, intended to narrow the field of options for further investigation rather than to conclude the planning process. Before implementing any strategy or solution it indicates, users are expected to undertake the analyses appropriate to their context, typically including the four-step models, the transport and land-use interaction (LUTI) modelling, both introduced in deliverables D.1.2.1 and D.1.3.1 of the OPTI-UP project (OPTI-UP, 2026); financial feasibility assessment; regulatory and legal review; and stakeholder consultation and environmental impact assessment.

(E) **Scope and professional advice:** The toolkit offers high-level strategic guidance calibrated to the public transport context of the OPTI-UP pilot cities and expert benchmark cases. It is not a substitute for comprehensive mobility planning, urban design, or engineering expertise, and its outputs should not be read as legal, financial, technical, or operational advice. Engagement with qualified mobility planners, transport engineers, and urban planning consultants is recommended as part of any full project evaluation.

**(F) Data accuracy:** The accuracy of toolkit outputs depends on the data supplied by the user. OPTI-UP project partners do not independently verify externally imported or user-supplied data, and outputs should be interpreted with this in mind. Responsibility for decisions made on the basis of toolkit outputs rests with the user, in consultation with the relevant planning and technical authorities.

**(G) Status of this version:** The toolkit is under active development within the OPTI-UP project (Interreg Central Europe, Grant Agreement No. CE0200713) and is expected to evolve as further pilot testing, stakeholder feedback, and expert review are incorporated. This section reflects the toolkit's status at the time of this deliverable and may be updated in subsequent releases as appropriate.

**(H) Liability notice**: OPTI-UP partners do not independently validate the accuracy of user-supplied data or inputs. The toolkit, its underlying expert knowledge base, and accompanying documentation are provided "as is" and "as available," without warranty of any kind, whether express or implied, including but not limited to warranties of accuracy, completeness, fitness for a particular purpose, or non-infringement. All final decisions, and their consequences, remain the sole responsibility of the user. The OPTI-UP project partners, and the organisations they represent, assume no liability for any direct, indirect, incidental, or consequential damages arising from actions taken or omitted on the basis of toolkit outputs.

# 7. Conclusion and next steps

## 7.1. Summary of the status of the toolkit

The current version of the OPTI-UP toolkit represents a functional and operational decision-support instrument that integrates the analytical outcomes of previous project activities into a practical digital application. The toolkit combines expert knowledge, experiences from the six OPTI-UP pilot cities, and a fuzzy rule-interpolation methodology to support municipalities in identifying suitable public transport improvement strategies.

At the current stage of development, the analytical engine, recommendation logic and web-based user interface have been successfully implemented and integrated into a coherent workflow. Users are able to describe the characteristics of their city through a structured set of variables, receive recommendations for appropriate public transport strategies, compare their local context with relevant pilot solutions, and access additional project knowledge through the integrated AI-supported documentation assistant.

The toolkit has undergone several rounds of technical verification and stakeholder testing throughout its iterative development. Feedback received from transport experts, project partners, IT specialists and potential end users has been incorporated into successive development versions, improving both the quality of the recommendations and the overall usability of the application.

Although the toolkit has reached a level of maturity suitable for demonstration, stakeholder engagement and practical pilot use, it should be regarded as a decision-support instrument rather than a fully automated planning solution. Its recommendations provide an informed starting point for further investigation and discussion, while detailed transport planning and investment decisions continue to require comprehensive local analyses and professional expertise.

Overall, the current version demonstrates that the objectives of this deliverable have been achieved by providing a structured methodology, a functional software implementation and the accompanying documentation describing the development process of the OPTI-UP toolkit.

## 7.2. Recommendations for further development and maintenance

While the current version of the OPTI-UP toolkit provides a robust basis for supporting public transport planning in small and medium-sized cities, several opportunities exist for its further development beyond the present deliverable.

The first priority should be the continuous expansion and refinement of the expert knowledge base. Increasing the number and diversity of expert benchmark cases would improve the robustness of the recommendation engine and enable the toolkit to better represent the variety of public transport conditions encountered across European municipalities. Additional validation activities involving transport professionals and local authorities would further strengthen the reliability of the recommendations.

Future development could also broaden the range of public transport strategies supported by the toolkit. While the current version focuses on DRT, public transport network optimisation, bus fleet electrification and the introduction of new public transport services, future releases could incorporate additional strategies such as multimodal integration, mobility hubs, active mobility measures, Bus Rapid Transit (BRT), shared mobility services or rail-based solutions where appropriate.

From a technical perspective, further improvements may include extending the flexibility of the recommendation engine, enhancing explainability of the generated recommendations and expanding multilingual support to facilitate wider international adoption. Continuous maintenance of the knowledge base, software components and supporting documentation will also be essential to ensure that the toolkit remains aligned with evolving transport planning practices, technological developments and policy priorities.

As additional pilot experiences and implementation examples become available, these should be incorporated into the toolkit in order to continuously improve its practical relevance and provide users with an increasingly comprehensive collection of reference solutions. In this way, the toolkit can evolve from a project output into a living knowledge platform supporting evidence-based public transport planning.

## 7.3. Pathways for uptake by municipalities

The OPTI-UP toolkit has been designed to support municipalities during the early stages of public transport planning by providing accessible, transparent and evidence-based guidance for the selection of suitable improvement strategies. Its intended users include local authorities, transport planners, regional mobility organisations and consultants involved in the planning and management of public transport systems.

The adoption of the toolkit by municipalities should be considered as part of a broader planning process rather than as a stand-alone decision-making mechanism. The toolkit can assist users in identifying promising strategic directions, highlighting comparable implementation examples from the OPTI-UP pilot cities and facilitating discussions among decision-makers and stakeholders before more detailed technical analyses are undertaken.

Because the toolkit requires relatively limited input data and accepts both quantitative and qualitative information, it is particularly well suited to municipalities with limited analytical resources or without access to advanced transport modelling capabilities. By lowering the initial barrier to strategic assessment, the toolkit can help cities identify appropriate development pathways and prioritise further planning activities.

To maximise long-term uptake, municipalities should be supported through appropriate guidance materials, training activities and dissemination events demonstrating the practical use of the toolkit. Collaboration with regional authorities, transport operators and professional networks can further encourage knowledge exchange and facilitate the wider transfer of good practices developed within the OPTI-UP project.

Beyond the lifetime of the project, the toolkit has the potential to serve as a reusable decision-support resource for municipalities across Central Europe and beyond. Continued maintenance of the software, periodic updates of the expert knowledge base and the incorporation of additional case studies will help ensure that the toolkit remains relevant and valuable for future public transport planning initiatives.

# References

Kóczy, L., & Hirota, K. (1993). Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases. Information Sciences, 71(1–2), 169–201. https://doi.org/10.1016/0020-0255(93)90070-3

Nussbaumer, N. (2026). Enhancing public transport strategy selection in low-data environments: A Bayesian approach with expert priors [Unpublished student report for the TU Wien course: Interdisciplinary project report]

OPTI-UP (2024). Comprehensive data report on existing public transport networks and best practices. https://www.interreg-central.eu/wp-content/uploads/2025/01/OPTI-UP_-D.1.1.1._Comprehensive_report.pdf

OPTI-UP (2026). OPTI-UP deliverables published: From transport modelling to 6 local public transport plans. https://www.interreg-central.eu/news/opti-up-deliverables-published-from-transport-modelling-to-6-local-public-transport-plans/

# Appendix A. List of variables

Complete list of the 28 input variables used by the toolkit (Version 9). Numerical ranges shown are the support of each triangular fuzzy term (i.e. the full interval over which membership is non-zero); adjacent terms deliberately overlap. Boolean variables list their two possible terms.

| Variable | Definition | Type | Linguistic term and Numerical range |
|---|---|---|---|
| Socio-demographics | | | |
| Population | Total number of residents within the study area.<br>Unit: people | Linguistic (fuzzy); required | "very_low": [-1; 5,000; 15,000]<br>"low": [10,000; 25,000; 50,000]<br>"mid": [40,000; 100,000; 200,000]<br>"mid_high": [150,000; 300,000; 500,000]<br>"high": [400,000; 1,000,000; 3,000,000] |
| Area | Total land area of the study region.<br>Unit: km$^2$ | Linguistic (fuzzy); required | "small": [-1; 25; 75]<br>"mid": [50; 100; 175]<br>"large": [150; 225; 300] |
| Urban population density | Average number of people per square kilometre in the urban area.<br>Unit: people/km$^2$ | Linguistic (fuzzy); required | "very_low": [-1; 150; 400]<br>"low": [250; 500; 900]<br>"mid": [700; 1,200; 2,000]<br>"mid_high": [1,600; 2,500; 4,000]<br>"high": [3,500; 8,000; 15,000] |
| Car ownership | Number of privately owned passenger cars per 1,000 inhabitants.<br>Unit: cars/1,000 residents. | Linguistic (fuzzy) | "low": [-1; 150; 250]<br>"mid": [200; 350; 500]<br>"high": [450; 600; 800] |
| PT supply | | | |
| Bus availability | Indicates whether there is an organised local public bus or trolleybus service operating within the study area, with scheduled services serving urban public transport stops. | Boolean; required | "no", "yes" |
| PT hub availability | Presence of major public transport hubs such as bus terminals, railway stations, or ports. | Boolean | "no", "yes" |
| Tram availability | Existence of a tram or light rail system in the city. | Boolean | "no", "yes" |
| Urban railway availability | Presence of urban or suburban rail services (e.g. metro, S-Bahn). | Boolean | "no", "yes" |
| Taxi availability | Existence of a conventional taxi system operating in the service area. | Boolean | "no", "yes" |

| Variable | Definition | Type | Linguistic term and Numerical range |
|---|---|---|---|
| Peak-hour headay | Interval (in minutes) between two departures during peak hours?<br>Unit: minutes | Linguistic (fuzzy) | "low": [-1; 5; 15]<br>"mid": [10; 20; 35]<br>"high": [30; 60; 90] |
| Number of D-licensed driver compared to fleet in operation at peak demand | Low/Insufficient: peak service cannot be fully operated without cancellations or overtime; Mid/Just sufficient: peak service can be operated, but with little flexibility; High/Sufficient: enough drivers to operate peak service with reserve<br>Unit: ratio | Linguistic (fuzzy) | "low": [-0.1; 0.5; 0.9]<br>"mid": [0.8; 1.0; 1.2]<br>"high": [1.0; 1.2; 2.0] |
| Average distance between two PT stops | Estimated average distance between two consecutive PT stops to measure the density of PT coverage.<br>Unit: metre | Linguistic (fuzzy) | "low": [-0.1; 400; 700]<br>"mid": [500; 750; 1,200]<br>"high": [1,000; 2,000; 3,000] |
| Avg. population per stop | Average number of residents served by a single public transport stop. Note data from OSM sometimes count multiple directions of the same stop as two separate stops. Please only count them as one stop in this case.<br>Unit: people | Linguistic (fuzzy) | "low": [-0.1; 200; 400]<br>"mid": [300; 700; 1,000]<br>"high": [800; 3,000; 5,000] |
| PT demand | | | |
| Perceived load factor | Passenger perception of vehicle occupancy during operation. Mostly empty: < 20%; sometimes have to stand: 20-50% hard to find a seat: >50%<br>Unit: percent | Linguistic (fuzzy) | "low": [-0.1; 10; 20]<br>"mid": [15; 35; 50]<br>"high": [40; 75; 100] |
| Share of school pupils among PT users | Share of school-aged pupils in the total population, reflecting seasonal demand variability.<br>Unit: percent | Linguistic (fuzzy) | "low": [-0.1; 10; 20]<br>"mid": [20; 35; 50]<br>"high": [50; 75; 100] |
| PT mode split | Percentage of all trips made using public transport modes. This statistics can be found in the city's transport strategies and SUMPs.<br>Unit: percent | Linguistic (fuzzy) | "low": [-0.1; 5; 15]<br>"mid": [10; 15; 25]<br>"high": [20; 35; 60] |
| Average weekday PT ridership | Approximate number of passenger boardings on a typical working day within the local or urban public transport system.<br>Unit: passengers/day | Linguistic (fuzzy) | "low": [-1; 2,500; 5,000]<br>"mid": [4,000; 15,000; 30,000]<br>"high": [25,000; 50,000; 75,000] |
| Fleet characteristics | | | |

| Variable | Definition | Type | Linguistic term and Numerical range |
|---|---|---|---|
| Share of EV in PT fleet | Proportion of electric vehicles in the total public transport fleet.<br>Unit: percent | Linguistic (fuzzy) | "low": [-0.1; 15; 25]<br>"mid": [20; 40; 70]<br>"high": [60; 80; 100] |
| Average fleet age of ICE vehicles | Mean age of internal combustion engine vehicles in the public transport fleet.<br>Unit: years | Linguistic (fuzzy) | "low": [-1; 3; 6]<br>"mid": [4; 8; 12]<br>"high": [10; 15; 25] |
| Fleet size | Total number of public transport vehicles operated by the authority or operator.<br>Unit: vehicles | Linguistic (fuzzy) | "low": [-1; 10; 23]<br>"mid": [20; 60; 100]<br>"high": [80; 150; 300] |
| Average number of seats in a bus | Represents the average seating capacity of buses used in regular public transport service within the study area, reflecting the typical vehicle size operated on most routes.<br>Unit: number of seats | Linguistic (fuzzy) | "low": [-0.1; 5; 15]<br>"mid": [10; 30; 50]<br>"high": [40; 70; 100] |
| Policy goals & horizon | | | |
| Time horizon for PT improvement | Planned time horizon for implementing public transport improvements or strategies.<br>Unit: years | Linguistic (fuzzy) | "low": [-0.1; 2; 4]<br>"mid": [3; 5; 7]<br>"high": [6; 13; 20] |
| Goal importance of increase PT coverage | Importance of increasing PT coverage as a policy goal.<br>Unit: 0 (less important) to 10 (more important) | Linguistic (fuzzy) | "low": [-0.1; 2; 4]<br>"mid": [2; 5; 8]<br>"high": [6; 8; 10] |
| Goal importance of increase PT ridership | Importance of increasing PT ridership as a policy goal.<br>Unit: 0 (less important) to 10 (more important) | Linguistic (fuzzy) | "low": [-0.1; 2; 4]<br>"mid": [2; 5; 8]<br>"high": [6; 8; 10] |
| Goal importance of integrate PT with public space | Importance of integrating PT with the public space as a policy goal.<br>Unit: 0 (less important) to 10 (more important) | Linguistic (fuzzy) | "low": [-0.1; 2; 4]<br>"mid": [2; 5; 8]<br>"high": [6; 8; 10] |
| Goal importance of reducing operating costs | Importance of reducing the PT operating cost as a policy goal.<br>Unit: 0 (less important) to 10 (more important) | Linguistic (fuzzy) | "low": [-0.1; 2; 4]<br>"mid": [2; 5; 8]<br>"high": [6; 8; 10] |
| Goal importance of improve | Importance of improving PT system digitalization and efficiency as a policy goal. | Linguistic (fuzzy) | "low": [-0.1; 2; 4]<br>"mid": [2; 5; 8]<br>"high": [6; 8; 10] |

| Variable | Definition | Type | Linguistic term and Numerical range |
|---|---|---|---|
| digitalization and efficiency | Unit: 0 (less important) to 10 (more important) | | |
| Goal importance of reducing emissions | Importance of reducing emissions as a policy goal.<br>Unit: 0 (less important) to 10 (more important) | Linguistic (fuzzy) | "low": [-0.1; 2; 4]<br>"mid": [2; 5; 8]<br>"high": [6; 8; 10] |

# Appendix B. Expert case variable coverage

Diagnostic summary of how often each input variable is specified across the 207 expert benchmark cases (0012_filled_expert_cases_full.xlsx), broken down by linguistic term. A variable's total appearance count is the number of expert cases that specify any value for it; variables used in fewer than 207 cases means that they did not appear in all cases for expert to evaluate, consistent with the toolkit's support for partial/optional inputs (see Section 3.3.4).

| Variable | Appearances by linguistic term (count) | Total appearances (out of 207 cases) |
|---|---|---|
| Socio-demographics | | |
| population | high: 24 • low: 50 • mid: 56 • mid_high: 53 • very_low: 18 | 207 |
| area | large: 68 • mid: 75 • small: 58 | 207 |
| population_density | high: 31 • low: 46 • mid: 37 • mid_high: 45 • very_low: 42 | 207 |
| car_ownership | high: 11 • low: 7 • mid: 9 | 27 |
| PT supply | | |
| bus_availability | no: 38 • yes: 163 | 207 |
| pt_hub_availability | no: 12 • yes: 10 | 22 |
| tram_availability | no: 12 • yes: 10 | 24 |
| urban_railway_availability | no: 11 • yes: 14 | 31 |
| taxi_availability | no: 9 • yes: 13 | 22 |
| d_licensed_driver_availability | high: 10 • low: 7 • mid: 5 | 23 |
| avg_service_area_per_stop | high: 5 • low: 7 • mid: 7 | 19 |
| avg_population_per_stop | high: 10 • low: 4 • mid: 8 | 22 |
| PT demand | | |
| perceived_load_factor | high: 62 • low: 49 • mid: 52 | 164 |
| share_of_school_pupils_among_pt_users | high: 5 • low: 6 • mid: 9 | 23 |
| pt_mode_split | high: 6 • low: 13 • mid: 4 | 23 |
| avg_weekday_pt_ridership | high: 5 • low: 7 • mid: 8 | 26 |
| Fleet characteristics | | |
| share_of_ev | high: 56 • low: 46 • mid: 61 | 167 |

| Variable | Appearances by linguistic term (count) | Total appearances (out of 207 cases) |
|---|---|---|
| average_fleet_age_of_ice_vehicles | high: 56 • low: 50 • mid: 57 | 167 |
| fleet_size | high: 8 • low: 6 • mid: 9 | 29 |
| average_bus_size | high: 6 • low: 5 • mid: 6 | 17 |
| Policy goals & horizon | | |
| time_horizon | high: 4 • low: 8 • mid: 12 | 24 |
| goal_importance_increase_pt_coverage | high: 103 | 103 |
| goal_importance_increase_pt_ridership | high: 89 | 89 |
| goal_importance_integrate_pt_with_public_space | high: 106 | 106 |
| goal_importance_reduce_operating_costs | high: 106 | 106 |
| goal_importance_improve_digitalization_and_efficiency | high: 99 | 99 |
| goal_importance_reduce_emissions | high: 100 | 100 |