Key Performance Indicators definition and evaluation data

Deliverable Number	D.2.4
Deliverable Type	R - Document, report
Dissemination Level	PU - Public
Deliverable Responsible	Falk Fernbach (LUX)
Document Version & Status	V1.0 | Final

Project Acronym	CARMONY
Project Title	Coordinated Action for Responsive Mixed Orchestration and Network Yield
Grant Agreement Number	101202858
Project Coordinator	Virtual Vehicle Research GmbH
Project Website	www.carmony-project.eu

The document is submitted to the European Commission for review but has not been accepted so far. As soon as the final version of this document is available, this document should be deleted.

Funded by the European Union

CARMONY has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101202858.

Author(s)

Name	Organisation
Falk Fernbach	LUX
Patrick Van Egmond	LUX
Alexia Bierlaire	LUX

Reviewers

Name	Organisation	Date
Gamal Elghazaly	UNILUX	2026-01-27
Ana Martínez Roselló	ETRA	2026-01-20

Change History

Version	Date	Name/Organisation	Description
V0.1	2026-01-06	Falk Fernbach (LUX)	Version 1
V0.2	2026-01-13	Falk Fernbach (LUX)	Final Draft
V0.3	2026-01-19	Falk Fernbach (LUX)	Final internal version
V0.4	2026-01-23	Falk Fernbach (LUX)	Corrected Version after review 1
V0.5	2026-01-29	Falk Fernbach (LUX)	Corrected Version after review 2
V1.0	2026-01-30	Katrin Walzer (ViF)	Quality check and submission

Table of Contents

List of Figures

1. Executive Summary

Deliverable D2.4 consolidates the final outcomes of Task 2.5 and establishes the evaluation approach for CARMONY, covering both the Interlinked Orchestrator and the impacts of the six use cases (UC1-UC6) across the two pilots in Murcia (urban) and Luxembourg (highway/corridor). CARMONY targets mixed-traffic management under increasingly heterogeneous conditions, where manually driven vehicles, connected vehicles, and automated vehicles coexist, and where multiple stakeholders must coordinate responses to congestion, incidents, policy constraints (e.g., Low Emission Zones), and disruptive events.

The core output of D2.4 is a refined and validated KPI register that translates proposal-level ambitions into precise, measurable indicators ready for implementation in real-world pilots and simulation-based demonstrations. The register ensures full traceability from each “Initial KPI” in the proposal to the final KPI definition used for evaluation and specifies, per KPI, the relevant use case reference, evaluation context (real-world or simulation), indicator and target value, baseline and data source, expected data formats, collection frequency, storage location, validation status, and required clarifications. This common structure ensures that all partners apply a consistent evaluation logic and that results can be compared reliably across pilots, scenarios, and test conditions.

KPI refinement followed a structured, partner-led co-creation process focused on clarity and feasibility under operational constraints.

Phase 1 harmonised the initial KPI set into a consistent structure and gathered partner and stakeholder feedback, identifying insufficient precision as the main barrier to consistent implementation.

Phase 2 used dedicated workshops to resolve open points and align methodology, including baseline definitions (simulation vs. real-world reference), system-level versus vehicle-level measurements, success thresholds (e.g., seconds/minutes/percentages), and harmonised measurement windows and geographic boundaries.

Phase 3 incorporated these clarifications, pilot feasibility constraints, and inputs from other work packages on data availability and measurement capabilities, followed by a final validation round with partners and a pilot-level check of practical measurement details (e.g., time resolution and refresh/update cycles).

Overall, D2.4 delivers a transparent and operational KPI framework that supports credible assessment of orchestration performance and use case impacts in both real-world conditions and scalable simulation environments.

Additionally, D2.4 defines a set of orchestrator-operation KPIs to ensure that the Interlinked Orchestrator can be evaluated as a reliable operational service, consistent with established monitoring practice for distributed systems.

Keywords: Key Performance Indicators, Evaluation methodology, Mixed traffic management, Cooperative, CCAM, Real-world pilots, Simulation-based validation, Murcia (urban pilot), Luxembourg (highway/corridor pilot);

2. Introduction

CARMONY (Coordinated Action for Responsive Mixed Orchestration and Network Yield) is a Horizon Europe project developing and demonstrating an Interlinked Orchestration Framework for traffic management in mixed traffic conditions, where conventional driven vehicles, connected vehicles, and increasingly automated vehicles coexist. The project takes a user-centric approach, aiming to translate travellers' preferences and needs into traffic management and network control decisions. Its goal is to help cities and road operators manage traffic more efficiently and safely, while supporting sustainability objectives through better, network-wide orchestration, and at the same time meeting user preferences. Terminology related to driving automation levels is aligned with SAE J3016[1].

Traffic is increasingly heterogeneous[2]: different vehicle types, different automation levels, different mobility services, and different stakeholder objectives (public authorities, fleet operators, citizens). This makes it difficult to coordinate responses to congestion, incidents, low-emission policies, and disruptive events in a way that stays safe, efficient, and fair for everyone. CARMONY addresses this challenge by developing orchestration methods that can work across actors and tools, supported by real-world data and credible simulation.

In CARMONY, the orchestrator is a decision-support and coordination service that combines (i) traffic management tools, (ii) data from infrastructure, fleets and users, and (iii) simulation-based forecasting to propose actions such as routing recommendations, warnings, mitigation measures, and coordination strategies. A key feature is that decisions and actions can be tracked and evaluated to support long-term trust and fairness, while the orchestrator communicates with both CCAM actors and human users via apps and dashboards.

CARMONY pilots the orchestration approach in two complementary contexts:

- Spain (Murcia): urban scenarios including low-emission zone (LEZ)-type strategies and incident/disruption management.
- Luxembourg: highway/corridor context with cross-border characteristics, focused on recurring congestion and disruptive events.

CARMONY is implemented through two pilots, covering six use cases (UC) in total, with three urban use cases (UC1–UC3) and three highway use cases (UC4–UC6).

In addition to assessing use-case impacts on traffic efficiency, safety and sustainability, CARMONY also requires an explicit evaluation of the operational performance of the Interlinked Orchestrator itself. As the orchestrator acts as a central decision-support and coordination service, its reliability, responsiveness and capacity directly influence the feasibility and credibility of all use cases.

For this reason, Task 2.5 defines not only use-case-specific KPIs, but also a dedicated set of system-level KPIs addressing orchestrator operation, such as service availability, response latency, error rates and recovery behaviour. These system KPIs complement the use-case

[1] SAE International, 2021, https://www.sae.org/standards/j3016_202104-taxonomy-definitions-terms-related-driving-automation-systems-road-motor-vehicles.
European Commission, 2022)
[2] CCAM Partnership (2023) *Strategic Research and Innovation Agenda (SRIA) Update 2023.* Available at: https://www.ccam.eu/wp-content/uploads/2023/11/CCAM-SRIA-Update-2023.pdf.

KPIs by ensuring that observed impacts can be interpreted in light of the orchestrator's operational behaviour and constraints.

Note: This deliverable was prepared by the project partners. AI-based tools (including large language models) were used solely to support language editing, grammar correction, and improvement of textual clarity. All technical content, methodological choices, interpretations, and conclusions remain the sole responsibility of the authors.

2.1 Relation to other WPs and deliverables

D2.4 directly feeds the implementation, integration and validation work in the subsequent work packages. In WP3 and WP4 (architecture, service development and integration), the KPI definitions and data specifications act as functional measurement requirements, ensuring that the orchestrator services, interfaces and logs are implemented with the necessary observability (timestamps, identifiers, geofencing, message traces) to compute each KPI. In WP5 (pilot preparation and execution) and WP6 (testing, simulation and demonstrations), D2.4 provides the reference evaluation framework to design test campaigns, define baseline scenarios (real-world and/or counterfactual simulation), and align data collection plans and measurement windows across Murcia and Luxembourg. Finally, in WP7 (impact assessment, exploitation and dissemination), the KPI register provides the basis for aggregating results, comparing impacts across use cases, and substantiating project-level claims on efficiency, safety, sustainability and user-centric outcomes with traceable evidence.

2.2 Structure of this deliverable

This deliverable is structured to guide the reader from the project context to the operational KPI register used for evaluation.

- Section 2 introduces CARMONY, the pilots and use cases, and clarifies the objectives of Task 2.5 and Deliverable D2.4.
- Section 3 describes the KPI evaluation process and methodology, including the co-creation and validation approach and the definitions of all KPI register fields used consistently throughout the document.
- Section 4 presents the final KPI register structured by use case, including for each KPI the indicator and target value, baseline and data sources, expected formats, collection frequency, storage location, validation status, and clarifications.
- Section 5 synthesises how the KPI framework enables consistent assessment across pilots and scenarios, and Section 6 provides the list of abbreviations.

2.3 The objectives of deliverable D2.4 and Task 2.5

Deliverable D2.4 "Key Performance Indicators definition and evaluation data" provides the final KPI set and the associated evaluation data logic that will be used to assess the success of the project's use cases and the Interlinked Orchestrator.

Task 2.5 "Key Performance Indicators definition and evaluation data" defines the KPIs needed to evaluate:

- whether the use cases achieve their intended impacts (with the consolidated impact assessment reported in Task 7.1), and

- whether the orchestrator performs reliably and effectively under real operational constraints, with validation carried out in simulation prior to, and alongside, real-world operations.

This KPI work is part of WP2 (Requirements & KPIs), which provides the foundation for later technical development, integration, testing and demonstrations.

CARMONY's KPI framework is designed to reflect the project's ambition to improve:

- Traffic efficiency (e.g., reducing travel time and congestion),
- Safety (e.g., safer behaviour and fewer high-risk interactions),
- Sustainability (e.g., reduced emissions and energy use),

while ensuring that orchestration works across mixed actors and stakeholders in both urban and highway environments.

In addition, CARMONY explicitly combines real-world demonstrations with simulation-based demonstrations, so that short-term feasibility and long-term/large-scale impacts can both be assessed using consistent KPI logic.

3. KPI Evaluation Process and Methodology

This section explains how the KPIs were developed and made measurable, including a short overview of the six use cases, the evaluation approach per pilot (real-world, simulation, or hybrid), the partner co-creation and validation process, as well as the definitions of all KPI register fields used consistently in the document.

3.1 The use case overview

This section provides a short overview of the use cases; further details on the use case definitions, scope, stakeholders, and assumptions are provided in Deliverable D2.1.

3.1.1 UC1 - Traffic Management Optimization in an Urban Environment

In a city context, where low-emission policies constrain access, travellers still need reliable routes and clear guidance. The orchestrator supports trip planning by providing route recommendations, LEZ compliance information, and alternative options such as transfer points (e.g., Park & Ride) when a vehicle cannot enter the zone.

LEZ (Low-Emission Zone): a defined area where access for certain (more polluting) vehicles is restricted or only allowed under specific conditions, to reduce harmful emissions and improve air quality.[3]

Figure 1 Upsticks Spain (2025) 'Low Emission Zones ZBE in Spain, coming to a town near you', Upsticks Spain, 21 March (updated 6 April 2025). Available at: https://upsticks.es/low-emission-zones-zbe-in-spain-coming-to-a-town-near-you/ (Accessed: 28 January 2026).

Why it matters: Helps cities meet air-quality and climate objectives while keeping mobility practical and fair. Reduces unnecessary traffic (e.g., detours, searching) and supports smoother flows.

What makes it unique: It is policy-driven orchestration: decisions depend on vehicle eligibility (e.g., emission class) and access constraints. It explicitly considers modal alternatives when direct access is restricted.

Where / how evaluated: Murcia (Infante Juan Manuel area). Hybrid approach: demonstrated in real conditions where possible and complemented with simulation to scale and compare scenarios.

[3] Flanders.be (2025) *Low-emission zones (LEZ)*. Available at: https://www.vlaanderen.be/en/low-emission-zones-lez?utm.

3.1.2 UC2 - Critical Road Blockage

A vehicle breakdown or incident partially blocks a key urban intersection. The orchestrator aims to detect/confirm the disruption, inform relevant actors, and reduce secondary congestion through coordinated information and traffic re-routing strategies.

Why it matters: Faster, better coordinated incident response reduces queue growth, delays, and crash risk. It supports authorities/operators by shortening the time between detection, communication and action.

What makes it unique: It is event-triggered orchestration: the system reacts to an abnormal event, not regular demand. Focus is resilience and traffic stabilization, rather than policy compliance.

Where / how evaluated: Murcia urban context. Real-world + simulation: real disruption cases plus simulated “stronger disruption” variants to test impact without operational risk.

3.1.3 UC3 - Traffic Orchestration During a Disruptive Event

A connected emergency vehicle needs to reach its destination faster and more safely. The orchestrator supports an emergency corridor by combining route/corridor selection, traffic signal priority, warnings to nearby vehicles, and alerts to connected Vulnerable Road Users (VRUs) in the affected area.

Why it matters: Potentially reduces emergency response time (high societal value). Improves safety for everyone by reducing abrupt manoeuvres and providing early warnings to connected VRUs and connected vehicles.

What makes it unique: It directly involves priority operations (traffic signals) and safety-critical messaging. Strong focus on “right people get warned” (coverage + avoiding unnecessary alerts).

Where / how evaluated: Murcia emergency corridors. Real-world demonstration is central for this UC.

3.1.4 UC4 - Carpool Lane Optimization, innovative mobility services

On the A3 corridor, congestion is recurrent during peak periods. This use case tests (in simulation) whether a dynamic reservation/slot approach can improve the utilisation and throughput of a carpool (and/or managed) lane, including controlled access strategies under defined conditions.

Figure 2 Paperjam (2025) 'A3 motorway: the carpool lane is opening', Paperjam English News, 11 May. Available at: https://en.paperjam.lu/article/a3-motorway-the-carpool-lane-is-opening (Accessed: 28 January 2026). Photo credit: MMTP.

Why it matters: Improves corridor performance without building new infrastructure. Supports cross-border commuter flows by making capacity management smarter.

What makes it unique: It is a lane access / capacity management use case (not only routing or warning). Works with concepts like time slots/eligibility to optimise throughput.

Where / how evaluated: Luxembourg A3 corridor. Simulation-only, to explore effectiveness and policies safely before real deployment constraints.

3.1.5 UC5 - Cross-Border Highway Traffic Optimization in case of interruption

This use case demonstrates orchestration in real corridor operations on the A13 towards the cross-border area (Schengen/Perl direction). It integrates infrastructure and vehicle data to support operational decisions and user recommendations, covering both commuter and freight perspectives.

Why it matters: Shows the orchestrator works with real data flows and real operational constraints. Cross-border corridors are complex: demand peaks, disruptions, and heterogenous users.

What makes it unique: It is explicitly real-world corridor orchestration, not only a simulation study. Strong focus on data integration and operational feasibility in live conditions.

Where / how evaluated: Luxembourg A13 corridor. Real-world demonstration.

3.1.6 UC6 - Speed Reduction for Jam Avoidance

Stop-and-go traffic is often caused by speed variability and unstable flow. This use case evaluates, in simulation, whether the orchestrator can prevent or mitigate congestion by harmonising speeds, either through in-vehicle / app-based speed advice or by influencing Variable Speed Limits (VSL) where applicable.

Why it matters: Smoother traffic means less delay, fewer harsh braking events, and potentially lower emissions. Supports safety and comfort through more stable driving dynamics.

What makes it unique: It explicitly studies penetration (how many connected users are needed for network-wide benefits). It targets traffic stability rather than just re-routing.

Where / how evaluated: Luxembourg A3 corridor. Simulation-only, ideal for testing thresholds and edge cases repeatedly.

3.2 The main differences between the use cases

CARMONY addresses both urban and highway contexts by distinguishing use cases according to environment, operational mode, and evaluation approach.

In the urban pilot in Murcia, UC1 focuses on low-emission zone and policy-related traffic optimisation, UC2 addresses incidents at intersections, and UC3 targets emergency corridor activation with particular attention to VRUs, while in the highway pilot in Luxembourg, UC4 explores managed lane access through simulation, UC5 focuses on the live A13 cross-border corridor, and UC6 examines speed harmonisation using simulation.

Across these use cases, normal traffic operations are optimised in UC1, UC4, UC5, and UC6, whereas disruption and emergency situations are explicitly addressed in UC2 through incident management and in UC3 through emergency response and safety messaging.

From an evaluation perspective, UC3 and UC5 provide real-world demonstrations to test feasibility and operational constraints, UC1 and UC2 combine real-world elements with simulation to enable scaling and deeper analysis, and UC4 and UC6 rely on simulation-only demonstrations to safely support large-scale testing, repeatability, and penetration or threshold analysis.

3.3 Design process and co-creation workshop

To ensure that the KPI framework is measurable, operationally realistic, and aligned with the pilot sites' data availability, the KPI work in Task 2.5 was developed through a structured co-creation process with project partners and pilot stakeholders (operators, authorities, and data owners). The work followed three main phases.

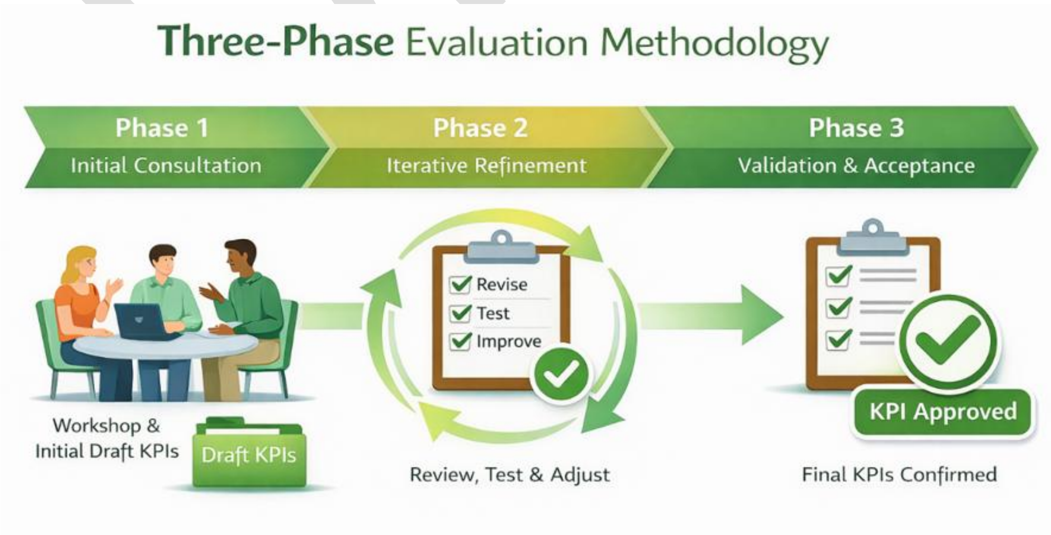

Figure 3 Three-phase KPI refinement and validation workflow. LuxMobility, based on Task 2.5 process documentation, CARMONY D2.4 (2026).

Phase 1 focused on KPI structuring and first refinement. The initial KPI list was consolidated into a consistent KPI structure, followed by a first round of feedback collected from project partners and pilot stakeholders focusing on clarity, measurability, and data feasibility. A key challenge in this process was the lack of sufficient precision and detail in the initial KPI definitions, which made it difficult to assess whether KPIs could be measured consistently across use cases, pilot sites, and data sources. Many of the resulting comments were therefore linked to use-case-specific conditions, pilot site constraints, and data availability, such as the existence of real-time data, accessibility of system logs, and the definition of credible baselines for evaluation.

Phase 2 focused on the workshop-based clarification of open points. After an additional round of revisions by LuxMobility, a dedicated workshop was organised to resolve remaining ambiguities and open questions, including baseline definitions (simulation-based versus real-world references), system-level versus vehicle-level measurements, precise success thresholds (e.g. seconds, minutes, percentages), and the harmonisation of measurement windows and geographic boundaries.

In several cases, KPI performance is measured in real-world pilots (e.g. via system logs), while the baseline is derived from simulation to represent the counterfactual “no-orchestrator” scenario, which cannot be reproduced safely, ethically, or consistently in live operations. This approach preserves comparability across scenarios while reflecting practical constraints at the pilot sites.

Following the workshop, LuxMobility consolidated the feedback from all participants and produced an updated version of the KPIs with a refined structure, which was then circulated for another round of validation and comments.

Phase 3 focused on the updated KPI register and final validation round. The KPI register was revised to integrate the clarifications agreed in earlier phases, feasibility constraints identified at pilot sites, and additional information gathered from other work packages on data availability and measurement capabilities. A final validation round with project partners confirmed the KPI set and its measurement approach, ensuring readiness for implementation in both real-world pilots and simulations. In addition, the pilot sites carried out a last, more detailed validation of specific KPIs, refining parameters such as achievable time resolution (e.g. seconds or minutes) and thresholds, to ensure the KPIs were as precise and operationally realistic as possible at this stage of the project.

3.4 KPI Register Field Definitions

The KPI register includes the following fields: Use case reference, Category, Initial KPI (target), Use case type, Number, Final KPI description and development, Baseline, Baseline data source, Data format, Collection frequency, Data storage location, Validated, and Clarifications.

In the following sections, each of these KPI register fields is explained in detail.

3.4.1 Use case reference

Links each KPI to the relevant use case(s) (UC1–UC6).

3.4.2 Category

Tags each KPI with the impact area it supports: Traffic efficiency, Safety, Sustainability.

3.4.3 Initial KPIs

The Initial KPI expresses the original ambition defined in the CARMONY proposal for a given use case, representing the intended performance improvement that the project aims to achieve through orchestration.

3.4.4 Use Case type

The use case type indicates whether a KPI is evaluated in a real-world trial or a simulation.

- Simulation KPIs are used to test the orchestrator in a controlled and repeatable environment, enabling large-scale scenarios, sensitivity analyses, and "what-if" comparisons under identical conditions.
- Real-world KPIs are used to validate the simulation findings, demonstrate technical feasibility, and confirm that the system performs reliably under real operational constraints (e.g., data availability, latency, stakeholder processes, and user behaviour).
- Hybrid evaluation refers to KPIs where performance is measured using real-world pilot data, while the baseline and/or complementary comparison scenarios are generated in simulation (e.g., counterfactual "no-orchestrator" runs or scaled scenarios). This approach is used when a fully comparable real-world "no-orchestrator" reference period cannot be obtained for operational, safety, ethical, or repeatability reasons.

Because the real-world trials cannot always provide a fully comparable "no-orchestrator" reference period for operational, safety, and repeatability reasons, the baseline for some real-world KPIs is generated in simulation as a counterfactual scenario without orchestration. KPI performance is still measured in the real-world pilot using operational logs and field data, and results are compared against the simulated baseline to quantify the orchestrator's impact. Where a real-world baseline is feasible, it is used instead; for example, in UC3 the baseline is collected in the real-world trial.

3.4.5 Number

Number refers to the unique KPI identifier, enabling stakeholders to quickly find, reference, and consistently track each KPI across documents and reports.

3.4.6 Final KPI description and development

This field describes the final KPI, providing a more precise and detailed definition and illustrating how the KPI has evolved compared to the Initial KPI (Section 3.2.2) through refinement during the project.

It captures the outcomes of the co-creation process and stakeholder input that influenced the KPI scope, wording, and success criteria, showing what was clarified, reformulated, split, or adjusted to ensure that the KPI is measurable and unambiguous.

3.4.7 Indicator

The indicator specifies, in percentages or absolute values (e.g. seconds), the exact performance target the KPI aims to achieve, ensuring a high level of precision about the intended outcome.

3.4.8 Baseline

The baseline defines the reference scenario against which KPIs are assessed. “A baseline is useful because it provides the ‘before’ (or ‘no-intervention’) reference condition against which post-deployment performance can be compared to quantify the effect of the intervention (FHWA, 2019).[4]”

3.4.9 Baseline - Data source

Baseline - Data source specifies where the baseline values originate from, such as simulation outputs or real-world measurements, depending on the use case.

3.4.10 Data format

Expected format (CSV/JSON/logs) and the minimum required fields (e.g., timestamps, vehicle ID, segment ID). Further specifications on data structures and required attributes are provided in Deliverable D2.2.

3.4.11 Collection frequency

How often data is recorded and by what mechanism (e.g., periodic sampling at a defined frequency, per simulation timestep/run, or event-triggered logging when a specified condition occurs). This affects what can be concluded (e.g., high-frequency sampling supports latency analysis, while event-triggered logs capture discrete system responses without a fixed sampling rate).

3.4.12 Data storage

Where data is stored/shared in the project (e.g., project data space / repository), ensuring traceability. Further specifications on data storage are provided in Deliverable D2.2. The CARMONY Data Space concept and interoperability approach are consistent with FIWARE guidance and the IDS Reference Architecture Model[5].

3.4.13 Validated

This element provides clear traceability, showing which KPIs required the most modifications before validation. It also records the validation status, confirming whether a KPI was accepted directly during the workshop or refined over several iterations following the first consultation round.

3.4.14 Clarifications

Some KPIs require further clarification, as the KPI statement alone may not be sufficiently self-explanatory. For this reason, an additional “Clarifications” field is included to support validators and pilot sites in specifying any relevant assumptions, constraints, or methodological details related to the KPI definition and its measurement.

---

[4] FHWA (2019) *Evaluation Methods and Techniques: Advanced Transportation and Congestion Management Technologies Deployment Program.* Available at: https://ops.fhwa.dot.gov/publications/fhwahop19053/fhwahop19053.pdf.

[5] FIWARE Foundation, 2021, https://www.fiware.org/wp-content/uploads/FF_PositionPaper_FIWARE4DataSpaces.pdf; International Data Spaces Association, 2022, https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Tech-Talk-IDS-RAM.pdf

3.5 Orchestrator System KPIs

In line with the objectives, the KPI framework distinguishes between use-case impact KPIs and orchestrator system operation KPIs. While use-case KPIs assess the effects of orchestration on traffic efficiency, safety and sustainability, system KPIs focus on the operational performance of the Interlinked Orchestrator as a digital service.

According to Google[6], Orchestrator operation KPIs follow established monitoring practice for distributed digital services, focusing on a small set of “service health” signals that are widely used to evaluate whether a system is responsive and reliable under real conditions.

Orchestrator system KPIs are designed to capture whether the system operates reliably under real operational conditions and pilot constraints. They cover key aspects such as service availability, end-to-end response latency, processing reliability (error rate), successful delivery of orchestration actions to downstream channels, auditability/log completeness, and recovery time after failure. These KPIs are measured using system logs, platform monitoring tools and simulation outputs, depending on the evaluation context. In simulation-based evaluations, they support repeatable stress and scalability testing, while in real-world pilots they provide evidence of stable and dependable operation. Together, system KPIs and use-case KPIs ensure that CARMONY’s evaluation framework addresses both what the orchestrator achieves and how reliably it operates.

Additional engineering metrics (e.g., throughput or saturation) may be monitored during WP6 integration and testing but are not part of the core KPI set reported in this deliverable.

To ensure consistency across the deliverable, orchestrator system operation KPIs follow the same KPI register logic defined in Section 3.4 (Category, Use case type/evaluation context, Indicator/target, Baseline and baseline data source, Data format, Collection frequency, Data storage, Validation status, Clarifications). However, because system KPIs evaluate the orchestrator as a digital service rather than a use-case impact, they do not always require certain fields that are central for impact KPIs, most notably an “Initial KPI” derived from the proposal’s use-case ambitions. Instead, system KPIs are structured using a service-monitoring KPI template, in which the core definition is expressed through (i) a precise service-quality metric, (ii) an explicit measurement source, and (iii) an operational baseline.

In addition, the selection of system KPIs is consistent with widely used software quality models that emphasise characteristics such as reliability and performance efficiency (ISO, 2023)[7].

This harmonised approach preserves a common evaluation logic across the whole KPI register while adapting the presentation where necessary to reflect the different nature of system-level performance assessment and the specific measurability requirements.

[6] Google (n.d.) Monitoring Distributed Systems: The Four Golden Signals, Google Site Reliability Engineering. Available at: https://sre.google/sre-book/monitoring-distributed-systems/.

[7] Dynatrace (n.d.) What are golden signals?, Dynatrace Knowledge Base. Available at: https://www.dynatrace.com/knowledge-base/golden-signals/.

4. The Key Performance indicators of CARMONY

This section presents the final set of Key Performance Indicators (KPIs) defined in CARMONY, structured by use case and aligned with the project's two pilots and six use cases. For each use case, the baseline conditions are first described, followed by the detailed specification of the associated KPIs. Each KPI is documented in a consistent format, including its original ambition as defined in the proposal, the refined KPI description, the evaluation context (real-world or simulation), the indicator and target value, data sources, formats, collection frequency, storage location, validation status, and any necessary clarifications.

4.1 Use Case 1

The baseline for this use case is established through simulation, with CVs operating without orchestrator support (lacking route recommendations, LEZ restriction info, and parking spot notifications). The KPIs are always compared to not using the orchestrator.

4.1.1 KPI 1

- Category: Traffic efficiency.
- Initial KPI: The individual travel time is reduced by 15% and the travel time of all vehicles is reduced on average by 10%.
- Use case type: Simulation
- KPI description and development:

  The individual travel time (simulated CVs) is reduced by 15%, over one-hour simulation period
- Indicator: 15%
- Baseline - Data source: Simulation
- Data Format: CSV/JSON (vehicle ID, trip start/end time, travel duration)
- Collection frequency: 1 sec per simulation timestep (over one-hour simulation period)
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.1.2 KPI 2

- Category: Traffic efficiency.
- Initial KPI: The individual travel time is reduced by 15% and the travel time of all vehicles is reduced on average by 10%.
- Use case type: Simulation
- KPI description and development:

  The travel time of all vehicles is reduced on average by 10%. ->Average travel time within the LEZ does not increase by more than 10% after orchestration activation.
- Indicator: 10%
- Baseline - Data source: Simulation
- Data Format: CSV summary (vehicle category, avg speed, delay)
- Collection frequency: Per run
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)

- Validated: Yes
- Clarifications: Not applicable

4.1.3 KPI 3

- Category: Traffic efficiency.
- Initial KPI: At least 98% of trips' origins and destinations are processed by the orchestrator without delays or malfunctioning.
- Use case type: Real-world
- KPI description and development:
- **At least 98% of all trip origin-destination requests are processed by the orchestrator successfully, meaning a valid routing or orchestration output is generated within the maximum response time of 5 seconds, and without system errors, data loss, or incomplete results.**
- Indicator: 98%
- Baseline - Data source: Simulation
- Data Format: Log file (timestamped orchestrator events)
- Collection frequency: 1 log per routing
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Baseline is derived from simulation to represent the counterfactual "no-orchestrator" scenario, as an equivalent no-orchestrator reference period cannot be reproduced consistently in live operations. KPI performance is measured using real-world orchestrator system logs.

4.1.4 KPI 4

- Category: Sustainability.
- Initial KPI: Reduction of CO2 emissions within the LEZ by at least 8%.
- Use case type: Simulation
- KPI description and development:

  Average CO_2 emissions within the Low-Emission Zone reduced by at least 8% compared to the baseline (no orchestration).
- Indicator: 8%
- Baseline - Data source: Simulation
- Data Format: CSV (vehicle emission values from simulation output)
- Collection frequency: 1 sec per simulation timestep
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.1.5 KPI 5

- Category: Sustainability.
- Initial KPI: Reduction of fuel consumption by 8% for connected vehicles.
- Use case type: Simulation

- KPI description and development:

  Average normalized fuel (or energy) consumption per 100 km for CV is reduced by 8% compared to the baseline scenario without orchestration.
- Indicator: 8%
- Baseline - Data source: Simulation
- Data Format: CSV (fuel [g/s] or [l/100 km])
- Collection frequency: 1 sec per simulation timestep
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.1.6 KPI 6

- Category: Traffic efficiency and Sustainability.
- Initial KPI: The parking space utilization is improved by 25%.
- Use case type: Simulation
- KPI description and development:

  Average utilization of public and managed off-street parking spaces within the LEZ increased by 25% compared to the baseline (no orchestration)
- Indicator: 25%
- Baseline - Data source: Simulation
- Data Format: JSON/CSV (slot ID, occupied state, timestamp)
- Collection frequency: 1-5 min, occupancy status
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Collection frequency may need to be adjusted to minute-level updates, subject to system refresh cycles.

4.1.7 KPI 7

- Category: Traffic efficiency and Sustainability.
- Initial KPI: Increase the number of public transports users in people entering LEZ by 10%.
- Use case type: Simulation
- KPI description and development:

  Increase the number of public transports users in people entering LEZ by 10%.
- Indicator: 10%
- Baseline - Data source: Simulation
- Data Format: CSV (PT boardings, lightings per stop)
- Collection frequency: Per simulated hour
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.1.8 KPI 8

- Category: Traffic efficiency and Sustainability.
- Initial KPI: Reduce the number of routes avoiding the LEZ by 50%
- Use case type: Simulation
- KPI description and development:

  The increase in traffic volumes on the perimeter roads surrounding the Low Emission Zone (LEZ) shall not exceed 5% compared to the baseline scenario.
- Indicator: 5%
- Baseline - Data source: Simulation
- Data Format: CSV (route ID, avoidance flag)
- Collection frequency: Per simulation run
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: The KPI specification required extensive clarification to ensure precision and measurability.

4.2 Use Case 2

To assess baseline KPIs, a traffic simulation will feature a damaged vehicle blocking a Murcia intersection. CVs, CAVs, and the replacement vehicle will receive no orchestrator support. The KPIs are always compared to not using the orchestrator.

4.2.1 KPI 9

- Category: Traffic efficiency and Safety.
- Initial KPI: Time of informing authorities about the accident must be less than 10 seconds.
- Use case type: Real-world
- KPI description and development:

  Time to inform authorities about an accident must be less than 10 seconds; this metric is measured at the system level, not at the vehicle level, with the start time defined as the moment the system detects or confirms the accident due to traffic arising.
- Indicator: 10 sec
- Baseline - Data source: Simulation
- Data Format: Log file (event ID, time sent/received)
- Collection frequency: event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: KPI performance is measured in the pilot at system level (e.g., detection and notification timestamps from control-centre/orchestrator logs), while the baseline is derived from simulation to represent the counterfactual “no-orchestrator” scenario, which cannot be safely or consistently recreated in a real incident.

4.2.2 KPI 10

- Category: Traffic efficiency and Safety.

- Initial KPI: Time of technical assistance to arrive at accident must be less than 10 minutes.
- Use case type: Real-world
- KPI description and development:

  Time for technical assistance to arrive at the accident must be less than 10 minutes; this metric is measured at the control center level, not at the vehicle level. The start time is defined as the moment the system detects or confirms the accident due to traffic arising. The 10-minute target is relative to the location of the nearest technical assistance unit, representing the average response time monitored.
- Indicator: 10 min
- Baseline - Data source: Simulation
- Data Format: CSV (incident ID, dispatch/arrival timestamps)
- Collection frequency: event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: KPI performance is measured in the pilot at system level (e.g., dispatch and arrival timestamps where available), while the baseline is derived from simulation to provide a consistent counterfactual “no-orchestrator” reference, since real incidents cannot be repeated under identical conditions for baseline measurement.

4.2.3 KPI 11

- Category: Traffic efficiency.
- Initial KPI: The total length of the traffic jam, measured in distance, is reduced by 40%.
- Use case type: Simulation
- KPI description and development:

  Number and total duration of congestion events (average speed < 30 km/h sustained for > 2 minutes over ≥ 500 m segment) reduced by 40% compared to baseline
- Indicator: 40%
- Baseline - Data source: Simulation
- Data Format: CSV (queue length [m], time stamp)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.2.4 KPI 12

- Category: Sustainability.
- Initial KPI: Reduced amount of fuel from vehicles influenced by the accident by 25%.
- Use case type: Simulation
- KPI description and development:

  Average fuel consumption per affected vehicle and per kilometer in the incident area is reduced by 25% compared to the baseline (no orchestration)

- Indicator: 25%
- Baseline - Data source: Simulation
- Data Format: CSV (fuel [g/s] per vehicle)
- Collection frequency: 1s
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.2.5 KPI 13

- Category: Traffic efficiency and Safety.
- Initial KPI: Reduction of the time needed for the incident to be communicated to CVs & CAVs by 20%.
- Use case type: Simulation
- KPI description and development:

  Average communication latency between confirmed incident detection and message reception by CVs/CAVs must be below 2 seconds; this metric is measured at the system level, not at the vehicle level.
- Indicator: 2 sec
- Baseline - Data source: Simulation
- Data Format: JSON log (message ID, time sent/received)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Incident detection is performed through traffic monitoring and analysis of sensor and control center data. The KPI reflects system communication efficiency once an incident is validated and serves as a time-based indicator in the absence of a defined baseline.

4.2.6 KPI 14

- Category: Safety.
- Initial KPI: Reduction of vehicles speed approaching the damaged vehicle by 10%.
- Use case type: Simulation
- KPI description and development:

  Average speed of CVs/CAVs within 200 m of a detected damaged vehicle must decrease within 5 seconds after system warning messages are received.
- Indicator: 5 sec
- Baseline - Data source: Simulation
- Data Format: JSON log (message ID, time sent/received)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes

- Clarifications: This metric is measured at the system level to assess responsive driving behavior, serving as a performance indicator in the absence of a baseline.

4.3 Use Case 3

The baseline for this use case will be evaluated in a real-world trial. The emergency vehicle will receive a notification, but no orchestrator support, and nearby CVs, CAVs and connected VRUs will not receive warnings.

KPIs will be compared to scenarios without orchestrator support.

4.3.1 KPI 15

- Category: Traffic efficiency and Safety.
- Initial KPI: Emergency vehicle response time to be reduced by 15%.
- Use case type: Real-world
- KPI description and development:

  The average time between emergency incident notification and the start of the emergency vehicle's response route is reduced by 15% compared to the baseline, due to faster detection, communication, and route activation by the orchestrator.
- Indicator: 15%
- Baseline - Data source: real-world trial
- Data Format: JSON (event ID, notification, arrival timestamps)
- Collection frequency: Per Run
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Reformulated for clarification

4.3.2 KPI 16

- Category: Safety.
- Initial KPI: Travel time of the emergency vehicle is 25% less.
- Use case type: Real-world
- KPI description and development:

  Average travel time of the emergency vehicle from dispatch to arrival reduced by 25% compared to the baseline (no orchestration).
- Indicator: 25%
- Baseline - Data source: real-world trial
- Data Format: CSV (GPS trace, speed profile)
- Collection frequency: 1 Hz
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.3.3 KPI 17

- Category: Traffic efficiency.

- Initial KPI: Additional travel time of all other vehicles due to emergency corridor is 50% less.
- Use case type: Real-world
- KPI description and development:

  The activation of the emergency corridor shall not increase the average travel time or delay for non-emergency vehicles on cross-streets in more than 10% compared to normal operating conditions (no orchestration and no signal priority).
- Indicator: 10%
- Baseline - Data source: real-world trial
- Data Format: CSV (vehicle ID, delay seconds)
- Collection frequency: 1 Hz
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Has been changed to "real world", since we do not have simulation for this UC. Has been reformulated to study the effect of the emergency corridor in vehicles out of it, since the vehicles that are already in the corridor will benefit from the signal priority. The measurement will be either based on the extra red signal time for cross streets at the intersections with signal priority and/or the detected queue times at the affected segments.

4.3.4 KPI 18

- Category: Safety.
- Initial KPI: All required vehicles are informed within 10 seconds after the event is communicated to the orchestrator.
- Use case type: Real-world
- KPI description and development:

  Vehicles within the affected road segment are informed within 10 seconds after the event is communicated to the orchestrator.
- Indicator: 10 sec
- Baseline - Data source: real-world trial
- Data Format: System log (message timestamp per receiver)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.3.5 KPI 19

- Category: Sustainability.
- Initial KPI: The information of the approaching emergency vehicle is communicated to all required vehicles and to no vehicles who are not influenced.
- Use case type: Real-world
- KPI description and development:

The information of the approaching emergency vehicle is communicated to all required vehicles and to no vehicles who are not influenced.

- Indicator: 99%
- Baseline - Data source: real-world trial
- Data Format: JSON (recipient list IDs, event ID)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.3.6 KPI 20

- Category: Safety.
- Initial KPI: 50% less critical driving manoeuvres of dodge vehicles.
- Use case type: Real-world
- KPI description and development:

  50% less critical driving manoeuvres of dodge vehicles (Sudden braking or lane changes, Sharp acceleration or deceleration, Unsafe merges…)
- Indicator: 50%
- Baseline - Data source: real-world trial
- Data Format: CSV (acceleration, jerk, lane change flags)
- Collection frequency: 10 Hz
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Has been changed to "real world", since we do not have simulation for this UC.

4.3.7 KPI 21

- Category: Safety.
- Initial KPI: All connected VRUs in the hazardous area are informed at least 5 minutes before the arrival of the emergency vehicle.
- Use case type: Real-world
- KPI description and development:

  At least 95% of connected VRUs located within 200 m of the emergency vehicle's (CEV's) route receive a warning message at least 20 seconds before the CEV reaches their position.
- Indicator: 95%
- Baseline - Data source: real-world trial
- Data Format: JSON (VRU ID, alert timestamp)
- Collection frequency: Event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.3.8 KPI 22

- Category: Safety.
- Initial KPI: The emergency vehicle information is sent to all of connected VRUs within the delimited area and to none outside it.
- Use case type: Real-world
- KPI description and development:

  The emergency vehicle information is sent to all of connected VRUs within the delimited area and to none outside it.
- Indicator: 99%
- Baseline - Data source: real-world trial
- Data Format: JSON (VRU ID, position, alert flag)
- Collection frequency: Event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.3.9 KPI 23

- Category: Traffic efficiency and Safety.
- Initial KPI: Emergency vehicle finds a green light in the emergency corridor 25% more times.
- Use case type: Real-world
- KPI description and development:

  The emergency vehicle encounters green traffic signals and enable the CEV to pass through their route 25% faster within the designated emergency corridor compared to the baseline (no orchestration).
- Indicator: 25%
- Baseline - Data source: real-world trial
- Data Format: CSV (intersection ID, signal phase state)
- Collection frequency: 1 Hz
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.3.10 KPI 24

- Category: Traffic efficiency and Safety.
- Initial KPI: Emergency vehicle finds a green light at least 75% of the times when approaching a signalised intersection
- Use case type: Real-world

  KPI description and development: Emergency vehicle finds a green light at least 75% of the times when approaching a signalised intersection.
- Indicator: 75%
- Baseline - Data source: real-world trial

- Data Format: CSV (signal state, EV arrival timestamp)
- Collection frequency: 1 Hz
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.4 Use case 4

Baseline results for KPIs will be generated using real-world data simulations. KPIs are relative to scenarios without the orchestrator.

4.4.1 KPI 25

- Category: Traffic efficiency.
- Initial KPI: Reduced average travel time by 10% for all users on the highway.
- Use case type: Simulation
- KPI description and development:

  Average travel time per vehicle reduced by 10% compared to the baseline (no orchestration) over the same highway segment and time period.
- Indicator: 10%
- Baseline - Data source: Simulations
- Data Format: CSV (vehicle ID, start/end time)
- Collection frequency: Per simulation run
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.4.2 KPI 26

- Category: Traffic efficiency.
- Initial KPI: 25% less traffic jams on the highway.
- Use case type: Simulation
- KPI description and development:

  Number and total duration of congestion events (average speed < 30 km/h sustained for > 2 minutes over ≥ 500 m segment) reduced by 25% compared to baseline
- Indicator: 25%
- Baseline - Data source: Simulations
- Data Format: CSV (congestion segments, duration)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.4.3 KPI 27

- Category: Traffic efficiency.
- Initial KPI: Improved capacity management of the carpool lane by at least 25%.
- Use case type: Simulation
- KPI description and development:

  Average occupancy rate and traffic throughput of the carpool lane improved by at least 25% compared to baseline (no orchestration), measured as vehicles per hour.
- Indicator: 25%
- Baseline - Data source: Simulations
- Data Format: CSV (lane occupancy, flow)
- Collection frequency: 45 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Initially, a 5-second update interval was targeted; however, this was identified as too short for reliable estimation. While values could technically be reported every 5 seconds, the system would likely output unchanged values until a stable update is available. Therefore, an update interval of 30 seconds, or even 60 seconds, is considered more appropriate to ensure meaningful and reliable reporting.

4.4.4 KPI 28

- Category: Safety.
- Initial KPI: 10% less accidents and injuries during the operation of the orchestration.
- Use case type: Simulation
- KPI description and development:

  The average Time-to-Collision (TTC) across all vehicle interactions increases by 10%, indicating safer inter-vehicle distances.
- Indicator: 10%
- Baseline - Data source: Simulations
- Data Format: JSON log (incident ID, time, severity)
- Collection frequency: Event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: The original KPI (No.28) has been split into two separate KPIs. KPI No. 29 represents an additional KPI introduced as part of this refinement.

4.4.5 KPI 29

- Category: Safety.
- Initial KPI: 10% less accidents and injuries during the operation of the orchestration.
- Use case type: Simulation
- KPI description and development:

  The maximum deceleration rate observed in the network decreases by 10%, indicating smoother traffic flow and fewer harsh braking events.

- Indicator: 10%
- Baseline - Data source: Simulations
- Data Format: JSON log (incident ID, time, severity)
- Collection frequency: Event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: This KPI (Nr. 29) represents an additional indicator introduced following the split of the original KPI No. 28.

4.4.6 KPI 30

- Category: Sustainability.
- Initial KPI: 5% less fuel consumption for all vehicles on the highway.
- Use case type: Simulation
- KPI description and development:

  Average fuel consumption per vehicle-kilometre reduced by 5% compared to the baseline (no orchestration)
- Indicator: 5%
- Baseline - Data source: Simulations
- Data Format: CSV (fuel per vehicle [g/s])
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.4.7 KPI 31

- Category: Sustainability.
- Initial KPI: 10% less fuel consumption for vehicles using the carpool lane.
- Use case type: Simulation
- KPI description and development:

  For eligible vehicles on corridor, the average energy/fuel consumption per vehicle-km is more than 10% lower when the carpool lane is active and orchestrated than in a matched baseline period when the lane is inactive/not used.
- Indicator: 10%
- Baseline - Data source: Simulations
- Data Format: CSV (lane ID, fuel avg)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.5 Use Case 5

Baselines will be estimated based on real world measurement provided by Luxembourg. National Road Administration. KPIs are relative to scenarios without the orchestrator.

4.5.1 KPI 32

- Category: Traffic efficiency.
- Initial KPI: Reduced average travel time by 10% for all users on the highway.
- Use case type: Real-world
- KPI description and development:

  Average travel time per vehicle reduced by 10% compared to the baseline (no orchestration) over the same highway segment and time period.
- Indicator: 10%
- Baseline - Data source: real world measurement
- Data Format: CSV (loop detector speed/time)
- Collection frequency: 1 min
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.5.2 KPI 33

- Category: Traffic efficiency.
- Initial KPI: 25% less traffic jams on the highway.
- Use case type: Real-world
- KPI description and development:

  Number and total duration of congestion events (average speed < 30 km/h sustained for > 2 minutes over ≥ 500 m segment) reduced by 25% compared to baseline
- Indicator: 25%
- Baseline - Data source: real world measurement
- Data Format: JSON/CSV (segment status, congestion index)
- Collection frequency: 1 min
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.5.3 KPI 34

- Category: Safety.
- Initial KPI: 10% less accidents and injuries during the operation of the orchestration.
- Use case type: Hybrid
- KPI description and development:

  The average Time-to-Collision (TTC) across all vehicle interactions increases by 10%, indicating safer inter-vehicle distances.

- Indicator: 10%
- Baseline - Data source: real world measurement
- Data Format: CSV/JSON trajectories
- Collection frequency: Event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: The original KPI (No. 34) has been split into two separate KPIs. KPI No. 35 represents an additional KPI introduced as part of this refinement.

4.5.4 KPI 35

- Category: Safety.
- Initial KPI: 10% less accidents and injuries during the operation of the orchestration.
- Use case type: Hybrid
- KPI description and development:

  The maximum deceleration rate observed in the network decreases by 10%, indicating smoother traffic flow and fewer harsh braking events.
- Indicator: 10%
- Baseline - Data source: real world measurement
- Data Format: CSV/JSON trajectories
- Collection frequency: Event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes (Validation refers to KPI definition)
- Clarifications: This KPI (No. 35) represents an additional indicator introduced following the split of the original KPI No. 34. We are not sure if we have means of available data to measure maximum deceleration rates (Feasibility of real-world data extraction will be confirmed during WP6).

4.5.5 KPI 36

- Category: Sustainability.
- Initial KPI: 10% less fuel consumption for all vehicles on the highway.
- Use case type: Real-world
- KPI description and development:

  Average fuel consumption per vehicle-kilometre reduced by 10% compared to the baseline (no orchestration)
- Indicator: 10%
- Baseline - Data source: real world measurement
- Data Format: CSV (fleet sensor data [l/100 km])
- Collection frequency: 1 min
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.6 Use Case 6

The baseline KPIs are generated from simulation results using real-world highway data. KPIs are relative to scenarios without the orchestrator.

4.6.1 KPI 37

- Category: Traffic efficiency.
- Initial KPI: Reduce average travel time by 10% for every highway user.
- Use case type: Simulation
- KPI description and development:

  Average travel time per vehicle reduced by 10% compared to the baseline (no orchestration) over the same highway segment and time period.
- Indicator: 10%
- Baseline - Data source: Simulations
- Data Format: CSV (vehicle ID, trip time)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.6.2 KPI 38

- Category: Sustainability.
- Initial KPI: Reduce the average fuel consumption per driven km by 8% for every highway user.
- Use case type: Simulation
- KPI description and development:

  Average fuel consumption per vehicle-kilometre reduced by 8% compared to the baseline (no orchestration)
- Indicator: 8%
- Baseline - Data source: Simulations
- Data Format: CSV (fuel [g/s], distance [m])
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.6.3 KPI 39

- Category: Safety.
- Initial KPI: Reduce the number of traffic accidents by 10%.
- Use case type: Simulation
- KPI description and development:

  The average Time-to-Collision (TTC) across all vehicle interactions increases by 10%, indicating safer inter-vehicle distances.

- Indicator: 10%
- Baseline - Data source: Simulations
- Data Format: JSON (event ID, time, vehicles involved)
- Collection frequency: Event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: The original KPI (No. 39) has been split into two separate KPIs. KPI No. 40 represents an additional KPI introduced as part of this refinement.

4.6.4 KPI 40

- Category: Safety.
- Initial KPI: Reduce the number of traffic accidents by 10%.
- Use case type: Simulation
- KPI description and development:

  The maximum deceleration rate observed in the network decreases by 10%, indicating smoother traffic flow and fewer harsh braking events.
- Indicator: 10%
- Baseline - Data source: Simulations
- Data Format: JSON (event ID, time, vehicles involved)
- Collection frequency: Event-based
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: This KPI (No. 40) represents an additional indicator introduced following the split of the original KPI No. 39. We are not sure if we have means of available data to measure maximum deceleration rates.

4.6.5 KPI 41

- Category: Traffic efficiency and Safety.
- Initial KPI: Reduce the percentage of vehicles traveling more than 10% below or above the average speed on a road from 20% to 10%, aiming for a 20% reduction in overall accident rate.
- Use case type: Simulation
- KPI description and development:

  Reduce the percentage of vehicles traveling more than 10% below or above the average speed on a road from 20% to 10% (aiming for a 20% reduction in overall accident rate)
- Indicator: 10%
- Baseline - Data source: Simulations
- Data Format: CSV (vehicle speed records)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes

- Clarifications: Not applicable

4.6.6 KPI 42

- Category: Traffic efficiency.
- Initial KPI: Reduce the number of traffic jams by 25%.
- Use case type: Simulation
- KPI description and development:

  Number and total duration of congestion events (average speed < 30 km/h sustained for > 2 minutes over ≥ 500 m segment) reduced by 25% compared to baseline
- Indicator: 25%
- Baseline - Data source: Simulations
- Data Format: CSV (jam length/duration metrics)
- Collection frequency: 1 sec
- Data storage: CARMONY Data Space (FIWARE/IDSA-compliant cloud repository)
- Validated: Yes
- Clarifications: Not applicable

4.7 Orchestrator System Operation KPIs

4.7.1 Purpose and scope

In addition to use-case impact KPIs (traffic efficiency, safety, sustainability), CARMONY defines a concise set of KPIs that assess the operational performance of the Interlinked Orchestrator as a digital service. These KPIs ensure that the orchestrator can operate reliably under real operational constraints and provide measurable evidence of service quality during pilots and simulation-based demonstrations.

The selection is informed by established practice in monitoring distributed systems, in particular service-health signals such as latency and errors, which capture whether users can successfully obtain timely outputs under operational load.

In parallel, the chosen operational KPIs align with widely used software quality models that treat reliability and performance efficiency as key system qualities for evaluating operational readiness and robustness (ISO, 2023).

Where applicable, these system KPIs complement use-case KPIs by enabling interpretation of impacts in light of the orchestrator's observed operational behaviour[8].

[8] [8] Google (n.d.) Monitoring Distributed Systems: The Four Golden Signals, Google Site Reliability Engineering. Available at: https://sre.google/sre-book/monitoring-distributed-systems/. Dynatrace (n.d.) What are golden signals?, Dynatrace Knowledge Base. Available at: https://www.dynatrace.com/knowledge-base/golden-signals/.

4.7.2 Orchestrator Operation KPIs - KPI 43 Orchestrator service availability

Orchestrator service availability measured as the percentage of time the orchestrator API/services are operational and able to process requests.

Indicator: ≥ 99.0% availability during agreed evaluation windows

Baseline: Staging/pre-pilot window

Baseline data source: Platform monitoring / logs / APM

Collection frequency: 1 min (or finer if available)

4.7.3 Orchestrator Operation KPIs - KPI 44 Orchestrator end-to-end response latency

End-to-end API response time for key orchestrator endpoints (e.g., routing request, incident publish, advisory generation), reported as P95 and P99 latency.

Indicator: P95 ≤ 2 s; P99 ≤ 5 s

Baseline: Staging/pre-pilot window

Baseline data source: Platform monitoring / logs / APM

Collection frequency: Event-based

4.7.4 Orchestrator Operation KPIs - KPI 45 Orchestrator error rate

Ratio of failed orchestration transactions to total transactions.

Indicator: ≤ 1.0% failed transactions (or tighter for critical endpoints)

Baseline: Staging/pre-pilot window

Baseline data source: Platform monitoring / logs / APM

Collection frequency: Per request + aggregated per 1 min

4.7.5 Orchestrator Operation KPIs - KPI 46 Action delivery success rate

Percentage of orchestrator-generated actions successfully delivered to the intended downstream channel (apps, dashboards, infrastructure interfaces), based on delivery acknowledgements or log confirmations. This KPI complements the orchestrator error rate by focusing on downstream delivery confirmation rather than internal transaction success.

Indicator: ≥ 98% successfully delivered

Baseline: Staging/pre-pilot window

Baseline data source: Platform monitoring / logs / APM

Collection frequency: Event-based

4.7.6 Orchestrator Operation KPIs - KPI 47 Auditability and log completeness

Share of orchestration transactions for which the minimum audit trail is complete (event trigger, inputs version/IDs, decision output, dissemination record).

Indicator: ≥ 99% of transactions have complete trace fields

Baseline: Staging/pre-pilot window

Baseline data source: Platform monitoring / logs / APM

Collection frequency: Daily validation + per run

4.7.7 Orchestrator Operation KPIs - KPI 48 Recovery time after failure (service resilience)

Mean time to recover (MTTR) after orchestrator service failure or critical degradation (e.g., restart, redeploy, dependency outage).

Indicator: MTTR ≤ 15 minutes

Baseline: Staging/pre-pilot window

Baseline data source: Platform monitoring / logs / APM

Collection frequency: Event-based

5. Conclusion

This deliverable establishes the evaluation foundation for CARMONY by delivering a refined, validated, and implementation-ready KPI register for all six use cases across the Murcia and Luxembourg pilots. By converting proposal ambitions into measurable KPI definitions with clear indicators, baselines, and data requirements, D2.4 provides the project with a single reference framework to guide evaluation planning, data collection, and later impact assessment.

The KPI register ensures consistent assessment across heterogeneous environments (urban and highway), across operational modes (normal optimisation and disruption/emergency response), and across evaluation approaches (real-world trials, hybrid assessment, and simulation-only testing). Importantly, the KPI refinement process explicitly accounted for real operational constraints, such as data availability, logging access, system latency, and realistic update cycles, thereby reducing ambiguity and limiting implementation risks in subsequent work packages. The final validation rounds, supported by detailed input from pilot sites, ensured that KPI targets and measurement assumptions are as precise and feasible as possible at this stage of the project.

In total, D2.4 defines 42 use-case KPIs across UC1-UC6, complemented by 6 orchestrator system operation KPIs that assess the Interlinked Orchestrator's service reliability and performance.

These figures provide a transparent overview of evaluation coverage and support planning of test campaigns, data collection effort, and analysis workload.

While the KPI set is finalised for implementation, further refinements will be carried out during WP6 integration, testing, and demonstrations, including the detailed specification of KPI calculation methods, measurement windows, data cleaning rules, and aggregation procedures, as well as any final parameter adjustments required by pilot instrumentation and operational conditions. With this KPI register in place, CARMONY is prepared to implement a consistent evaluation methodology and to quantify orchestrator performance and use case impacts on traffic efficiency, safety, and sustainability under mixed traffic conditions, while supporting transparency, comparability, and reproducibility of results across pilots and scenarios.

In addition to use-case impact assessment, D2.4 explicitly introduces a set of orchestrator system operation KPIs, addressing availability, latency, and error rates. These KPIs ensure that the Interlinked Orchestrator is evaluated not only in terms of its effects on traffic and users, but also as an operational service deployed under real-world and simulated conditions.

By combining system-level KPIs with use-case-specific KPIs, the evaluation framework supports a transparent interpretation of results, allowing observed impacts to be linked to the orchestrator's operational performance. This dual perspective strengthens the credibility of the evaluation and provides a solid basis for subsequent testing, demonstrations and impact assessment activities in later work packages.

6. Abbreviations

Please include any abbreviations, terms etc. used within the deliverable in alphabetical order.

Term	Definition
A13	Motorway corridor in the Luxembourg-Germany cross-border area
A3	Motorway corridor connecting the French border to Luxembourg City
CAV	Connected Automated Vehicle
CCAM	Cooperative, Connected and Automated Mobility
CEV	Connected Emergency Vehicle
CO_2	Carbon Dioxide
CV	Connected Vehicle
FIWARE	Open-source framework for data spaces and smart applications
HMI	Human-Machine Interface
IDSA	International Data Spaces Association
KPI	Key Performance Indicator
LEZ	Low Emission Zone
MMTR	Mean time to recover
PT	Public Transport
SOP	Standard Operating Procedure
TTC	Time-to-Collision
UC	Use Case
VSL	Variable Speed Limits
VRU	Vulnerable Road User
WP	Work Package

7. References

CCAM Partnership (2023) *Strategic Research and Innovation Agenda (SRIA) Update 2023.* Available at: https://www.ccam.eu/wp-content/uploads/2023/11/CCAM-SRIA-Update-2023.pdf (Accessed: 28 January 2026).

Dynatrace (n.d.) What are golden signals?, Dynatrace Knowledge Base. Available at: https://www.dynatrace.com/knowledge-base/golden-signals/ (Accessed: 28 January 2026).

European Commission (2023) *Better regulation toolbox*. Available at: https://commission.europa.eu/law/law-making-process/better-regulation/better-regulation-guidelines-and-toolbox/better-regulation-toolbox_en (Accessed: 28 January 2026).

FHWA (2019) *Evaluation Methods and Techniques: Advanced Transportation and Congestion Management Technologies Deployment Program.* Available at: https://ops.fhwa.dot.gov/publications/fhwahop19053/fhwahop19053.pdf (Accessed: 28 January 2026).

FIWARE Foundation (2021) *FIWARE for Data Spaces: Position Paper (Version 1.0, June 2021)*. Available at: https://www.fiware.org/wp-content/uploads/FF_PositionPaper_FIWARE4DataSpaces.pdf (Accessed: 28 January 2026).

Flanders.be (2025) *Low-emission zones (LEZ)*. Available at: https://www.vlaanderen.be/en/low-emission-zones-lez (Accessed: 28 January 2026).

Google (n.d.) Monitoring Distributed Systems: The Four Golden Signals, Google Site Reliability Engineering. Available at: https://sre.google/sre-book/monitoring-distributed-systems/ (Accessed: 28 January 2026)

International Data Spaces Association (2022) *IDS Reference Architecture Model (IDS-RAM)*. Available at: https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/IDSA-Tech-Talk-IDS-RAM.pdf (Accessed: 28 January 2026).

SAE International (2021) *J3016_202104: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. Available at: https://www.sae.org/standards/j3016_202104-taxonomy-definitions-terms-related-driving-automation-systems-road-motor-vehicles (Accessed: 28 January 2026)

A. Annex

A table with the final list of KPIs (KPI ID, description, use case/pilot site, evaluation context, target value, etc.) is enclosed:

CARMONY Workshop D.2.4 - Task 2.5 - Version 6.11.25.xlsx