This report is a fictional example illustrating how a Citcom Responsible AI Label assessment should be compiled. All entities referenced (the company, the system, the city, and all named individuals) are fictional, with the sole exception of LIST and Technology and its named assessors. The assessors' opinions expressed in this example are fictional and do not necessarily reflect the conclusions that would be reached were a similar system submitted for real assessment.

CITCOM RESPONSIBLE AI LABEL

Assessment Report

UrbanPulse Traffic Intelligence Platform

LuxSmartMobility S.A. | Villange, Luxembourg

Assessment Context

Report reference	LUXSMARTMOBILITY-2027-LU
Assessment date(s)	9 February 2027 – 6 March 2027
TEF site	LIST – Luxembourg Institute of Science and Technology
Lead assessor	Dr. Alessio Buscemi
Contributors	Marouane Lahmidi
Lead reviewer	Dr. German Castignani
Report release date	27 March 2027

Provider Profile

Organisation name	LuxSmartMobility S.A.
Type of organisation	Private company (scale-up)
Country of establishment	Luxembourg
Contact person	Claire Mangen, Chief Product Officer (claire.mangen@luxsmartmobility.lu)

Subject of Assessment

System name	UrbanPulse Traffic Intelligence Platform (UTIP)
Version	v1.0-beta (pilot build)
Brief description	UTIP is an AI-driven adaptive traffic signal control system designed to ingest real-time data from road sensors, CCTV cameras, and connected vehicle feeds to dynamically optimise signal timing across

	an urban arterial network. A secondary module generates predictive congestion alerts and recommends rerouting to city operators. The system is currently operating in a supervised pilot on 14 intersections in Villange's Kourchberg district, with city operators retaining primary control at all times.
Smart city domain(s)	Mobility; Urban Infrastructure
TRL	Pilot (TRL 7)
Deployment location(s)	Villange, Luxembourg – Kourchberg district (14 intersections, supervised pilot since October 2026)
Who operates the system	LuxSmartMobility S.A., supervised jointly with Service de la Gestion de la Mobilité, Villange under pilot agreement
Primary user	LuxSmartMobility S.A. technical team and Villange pilot observers

How this Assessment was Initiated

This assessment was initiated at the request of LuxSmartMobility S.A. The provider approached LIST in November 2026 seeking independent expert assessment of UTIP as part of its preparation for commercial tendering with municipalities. LuxSmartMobility S.A. regards the Citcom Label as a way to demonstrate responsible AI development to prospective public-sector clients and to identify areas for further development ahead of full deployment. The Villange, as pilot host, facilitated access to operational staff. All three badges were placed in scope at the provider's request. LIST notes that assessment findings and recommendations are calibrated to TRL 7 and should be revisited upon any transition to full operational deployment.

Badges covered by this assessment: [X] Technical Testing [X] Governance [X] Impact

Materials Reviewed

The following materials were reviewed during the assessment. All items were provided by LuxSmartMobility S.A. unless otherwise noted.

Document	Version / Date	Notes
UTIP System Design and Architecture Document	v1.0, Jan 2027	
Model Performance Report (internal benchmarking, lab + pilot)	v1.0, Dec 2026	
Pilot Monitoring Report (Kourchberg, Oct 2026–Jan 2027)	Jan 2027	3 months of supervised pilot data
Risk Assessment (pre-pilot)	v1, Sept 2026	Update in progress
Data Protection Impact Assessment (DPIA)	v1, Aug 2026	Scoped to pilot; full deployment version planned
Pilot Agreement: Villange / LuxSmartMobility S.A.	Signed Sept 2026	Provided in full

Documentation is appropriate in scope and quality for a system at TRL 7. The provider was transparent about which documents are in draft or planned for expansion, and engaged constructively with all requests from the assessment team.

Interactions

Person / Role	Organisation	Format	Date
Claire Mangen, Chief Product Officer	LuxSmartMobility S.A.	Interview (2h)	16 Feb 2027
Thomas Braun, Lead ML Engineer	LuxSmartMobility S.A.	Interview + live system demo (3h)	19 Feb 2027
Sophie Kremer, Data Protection Officer	LuxSmartMobility S.A.	Written Q&A	23 Feb 2027
Jean-Paul Schmit, Traffic Control Manager	Villange	Interview (1.5h)	25 Feb 2027
Fatou Diallo, Senior Traffic Operator	Villange	Observation + interview (2h)	25 Feb 2027

The provider made all relevant staff available and responded promptly to follow-up questions. The Villange pilot team engaged openly with the assessment.

TECHNICAL TESTING BADGE

Testing Methodology: What was tested, under what conditions, against what benchmarks or datasets, and were these appropriate for the actual deployment environment?

UTIP's model has been evaluated against two complementary evidence bases: a controlled lab benchmarking exercise using 12 months of historical traffic data from the Kourchberg district (January–December 2025), and three months of live pilot data from the 14 operational intersections (October 2026–January 2027). The combination of retrospective and live evaluation is methodologically sound for a system at this stage and gives a clearer picture than either source alone. The benchmarks used, i.e. average intersection delay, queue length variance, and green phase utilisation, are well-chosen and appropriate for signal optimisation. The live demonstration on 19 February showed credible real-time adjustment to a congestion build-up on Avenue John F. Kennedy, with the system responding within expected latency parameters. The assessors note that the current evidence base reflects a single district under predominantly stable conditions; broadening the evaluation to cover a wider range of traffic patterns and urban contexts is a natural next step as the system matures, and the provider has indicated this is planned for the next pilot phase.

Independence of Testing: Who conducted the testing, the development team, a separate internal team, or an external party, and what does that imply for the reliability of reported results?

Testing has been conducted by LuxSmartMobility S.A.'s internal team, with the development and evaluation functions carried out by different team members. The pilot monitoring data has been reviewed jointly with Villange operators, which provides a degree of external scrutiny of live performance. Full independence of testing, through a separate external evaluation, is not expected at TRL 7 and its absence is not a concern at this stage. The provider has demonstrated awareness of this and indicated that broader external validation would be sought as part of any full deployment preparation. The assessors regard the current approach as proportionate and credible for a pilot-stage system.

Performance: Are the chosen metrics appropriate for the task and deployment context, are results reported on representative data, and is uncertainty or variability acknowledged?

The reported metrics show an 11% average reduction in intersection delay and an 18% reduction in queue spillback relative to the legacy fixed-timing plan at the 14 pilot intersections. These figures are consistent with what independent literature would suggest for a well-functioning adaptive signal control system in a district of this type, which gives the assessors reasonable confidence in their plausibility. The provider appropriately acknowledges that the three-month pilot period covers a single season and that performance under higher-variability conditions — winter disruption, major public events, peak tourist season — remains to be observed. This is an honest and technically sound characterisation of the current evidence. Extending the pilot period and reporting confidence ranges alongside point estimates are natural development steps the provider has already identified.

Fairness: Does the testing evidence show how the system performs across different groups and contexts, by neighbourhood, demographic, or infrastructure type, or was it evaluated only on aggregate metrics?

Performance data is currently reported at the network level across the 14 pilot intersections. Disaggregation by road user type — pedestrians, mobility-impaired persons, cyclists — has not yet been conducted, which is common at this stage of development. The provider's lead ML engineer engaged openly with the fairness question during interview and acknowledged that the current objective function prioritises vehicle throughput, with pedestrian phase lengths treated as constraints. This is a design choice the provider is aware of and is actively considering. The assessors note it as an area for development rather than a failing: broadening the optimisation objectives to explicitly incorporate pedestrian and active travel performance, and evaluating the system's effects on different road user groups, would strengthen the system's positioning for municipalities with active travel commitments. The Kourchberg pilot, as a predominantly commercial district, is a limited context for this analysis; a

future pilot in a more mixed residential environment would provide richer evidence.
Robustness & Failure Modes: Can the provider speak concretely about how their system fails and what mitigations are in place?

The provider demonstrated good awareness of the system's failure behaviour. The lead ML engineer described the primary failure mode encountered in the pilot, i.e. sensor data dropouts triggering reversion to static timing fallback, clearly and with specificity about the conditions that produce it. A model latency threshold of 340ms, beyond which the system defaults to pre-set timing plans, is documented and was verified during the live demonstration. Both failure responses functioned correctly during testing. The provider has not yet encountered or deliberately stress-tested correlated multi-intersection failures or degraded input conditions at scale, which is expected at this pilot stage. The failure mode picture will naturally broaden as the pilot expands. The assessors regard the provider's current level of self-knowledge about system behaviour as a positive indicator: a team that can articulate how and why their system fails is better positioned to manage it responsibly than one that cannot.
Safety & Harmful Outputs: Have outputs been evaluated for safety, bias, or harmful behaviour, and how rigorous was that evaluation relative to the deployment context?

Safety evaluation has appropriately focused on the most critical failure mode for a signal control system: preventing conflicting signal states. The provider has implemented and tested conflict-detection logic and fail-safe fallback behaviour, both of which functioned correctly during assessment. Under the current pilot arrangement, where city operators confirm all timing changes before implementation, the residual safety risk is well-managed. The assessors note that safety evaluation will need to expand as the system moves toward more autonomous operation, particularly with respect to the effects of signal timing decisions on pedestrian and cyclist safety.

Explainability: To what extent can the system's outputs be explained, and has explainability been tested or documented for the relevant use cases?

The system generates textual rationale summaries for its signal timing decisions, displayed in the pilot monitoring interface. During the observed session on 19 February, these summaries described the primary contributing factors at a level that both operators present found useful and intelligible — for example, indicating which approach had the highest detected queue and by how much the green phase had been extended in response. The operators confirmed that the explanations were consulted regularly and had helped them build confidence in the system's recommendations. This is a well-implemented explainability feature for an operator-facing context. The predictive congestion module produces less developed explanations at this stage, which is appropriate given its experimental status. The assessors regard the current explainability provision as a genuine strength of the system.

Gaps & Limitations: Where was evidence unavailable, access restricted, or a confident assessment not possible?

The evidence base reflects the pilot stage: three months, one district, 14 intersections under supervised conditions. Performance under a broader range of environments, road user compositions, and operational conditions has not yet been observed. These are natural limitations of a TRL 7 system rather than shortcomings in the assessment process, and the provider has characterised them accurately in its own documentation.

GOVERNANCE BADGE

Risk Management: Is risk management specific to this system and actively updated as circumstances change?

LuxSmartMobility S.A. has produced a risk assessment specific to the Kourchberg pilot that addresses technical failure modes, data quality risks, and operational contingencies with reasonable specificity. The document was drafted before the pilot began and an update incorporating findings from the first three months of operation is currently in progress, an update cadence the assessors regard as appropriate and indicative of a risk management practice that is treated as a living process rather than a one-time exercise. The current risk assessment is scoped to the pilot and does not yet address the broader risks of full-city deployment, which is both expected and appropriate at this stage. The provider indicated that a comprehensive risk assessment for deployment contexts would be developed as part of any procurement process, which is the right sequencing.

Human Oversight: Can human operators actually override the system in real time with genuine authority?

Under the current pilot arrangement, human oversight is the primary control mode. Villange operators confirm all signal timing changes before implementation; the system operates in an advisory capacity and cannot act autonomously. The override mechanism, which is a single-click rejection of any proposed change, was demonstrated live on 25 February and is immediately accessible from the main operator interface. The two operators observed during the site visit were clearly comfortable with the interface and confident in their authority over the system. One operator noted that working alongside UTIP had made the rationale for timing decisions more visible than under the legacy system, which she found useful. The assessors regard the current oversight model as well-designed and operationally genuine.

Accountability: When something goes wrong, who is concretely responsible?

The pilot agreement between LuxSmartMobility S.A. and the Villange was provided in full and reviewed. It assigns clear operational and technical responsibility to LuxSmartMobility S.A. and places the city in a supervisory and co-evaluation role, with liability provisions appropriately scoped to the pilot context. The accountability structure is clear and unambiguous for the current phase. The accountability framework for a full deployment contract, including the boundary between provider and municipal responsibility in operational scenarios, will need to be developed as part of procurement negotiations. The provider is aware of this and regards it as a standard part of the contracting process. The assessors have no concern about the current pilot arrangements.

Monitoring & Incident Response: Is the system monitored post-deployment and has incident response actually been exercised?

Pilot monitoring is conducted in real time via a shared dashboard accessible to both LuxSmartMobility S.A.'s technical team and Villange operators. Two sensor dropout incidents during the pilot were logged with root cause notes, and both were resolved without operational disruption. The shared monitoring arrangement is a good practice: it gives the city visibility into system behaviour from the outset and builds the operational familiarity that will be important when a full deployment transition is considered. A formal incident response procedure is being developed and is planned for finalisation before any full deployment. The two logged incidents demonstrate that the team responds to anomalies in a structured way even without a formally codified procedure, which is a positive signal.

Governance Over Time: Has the provider considered how governance adapts as the system evolves or ownership shifts?

The pilot is governed by a time-bound agreement with a defined review clause, which provides a natural checkpoint for assessing whether governance arrangements remain appropriate as the system evolves. The provider has thought through the transition from pilot to full deployment in operational terms, even

if the governance documentation for that transition is still to be developed. Questions about longer-term governance, i.e. continuity provisions, data portability, and arrangements in the event of changes to the company's ownership or structure, are not yet fully resolved. The provider acknowledged this openly and noted that these would be addressed in deployment contracting. For a company at this stage actively seeking its first municipal contracts, raising these questions early and being transparent about their status is the right approach.

Documentation: Is the documentation complete, specific, and candid?

Documentation quality is consistently good for a system at TRL 7. The system design document is technically detailed and clearly written. The pilot monitoring report is factual and candid, including the two sensor dropout incidents without minimisation. The algorithm description summary is accessible and gives a clear account of what the system does and on what basis it makes recommendations. Some documents, the human oversight procedures and the full-deployment DPIA, are in draft or planned for expansion, and this is clearly flagged in the materials. The assessors regard the overall documentary picture as honest and well-organised. The provider's transparency throughout the assessment, including on questions where the answers were incomplete, reflects well on its approach to responsible development.

Gaps & Limitations: Where was evidence unavailable, access restricted, or a confident assessment not possible?

No significant access restrictions were encountered during the assessment. The process badge reflects a company at an early but well-managed stage: governance arrangements are appropriately designed for the current pilot, with clear acknowledgment of what needs to be developed before full deployment. The assessors found no evidence of governance being treated as a formality, which is the most important indicator at this stage.

IMPACT BADGE

Affected Populations: Who is directly and indirectly affected by this system, and has the provider mapped populations beyond their primary users?

The provider's impact documentation focuses primarily on drivers and city operators as the intended beneficiaries, which reflects the current pilot context in Kourchberg — a predominantly commercial and institutional district. When the question of broader affected populations was raised during interview, the CPO engaged with it thoughtfully and acknowledged that a full deployment context would require a more comprehensive mapping, including pedestrians, residents of affected streets, and businesses whose accessibility depends on traffic patterns. The assessors regard this as an appropriate level of impact thinking for a pilot stage. The Kourchberg context limits the range of affected communities compared to what a residential deployment would entail, and the provider is aware of this distinction. Developing a more complete affected population mapping will be a natural and necessary step as the system is considered for broader deployment.

Fairness & Differential Impact: Do the system's outputs or effects differ systematically across groups, and is disaggregated data available?

Disaggregated impact analysis by road user type or neighbourhood has not yet been conducted, which is common at pilot stage. The assessors note that the current optimisation objective, vehicle throughput, with pedestrian phases as constraints, is a design choice that has implications for how benefits and burdens are distributed across road users. This is not a problem specific to UTIP; it reflects a common default in traffic signal engineering, and addressing it explicitly is an opportunity for the provider to differentiate its product for municipalities with strong active travel commitments. The provider engaged constructively with this question and indicated that incorporating pedestrian and active travel performance metrics into the objective function is under consideration for a future version. The assessors encourage this direction.

Transparency & Recourse: Are affected people informed when AI is being used, and do they have meaningful recourse?

At pilot scale, formal public transparency measures have not yet been introduced, and this is not unusual for a supervised pilot of this scope. A single community information session was held in Kourchberg in October 2026. The provider has indicated that a public-facing description of the system and a communication plan for deployment contexts are being developed. The assessors note that even at pilot stage, basic transparency, informing the public that a pilot is taking place and providing a point of contact, is good practice and straightforward to implement. This is a low-effort, high-visibility action that would strengthen the system's responsible development credentials ahead of any procurement process.

Broader Societal Effects: Does the deployment create new dependencies, change how public space functions, or affect employment or social systems at scale?

At the current pilot scale, broader systemic effects are limited. Importantly, the pilot has been designed to run in parallel with the existing legacy system, which remains active and could resume primary control immediately. No legacy infrastructure has been decommissioned. This is a well-considered pilot design that preserves reversibility and avoids creating dependencies before the system has been validated at scale. The assessors regard this as a responsible approach. The broader dependency and lock-in questions that arise at full deployment scale, i.e. single-vendor concentration, decommissioning of fallback systems, alignment with urban mobility policy, are appropriate to consider at the contracting stage rather than the pilot stage, and the provider is aware of them.

Participation & Community Input: How willing and transparent has the provider been about community effects, and have communities had any role in design or deployment?

Community engagement has been limited to the Kourchberg information session, which is proportionate

to the pilot scope. The provider acknowledged openly that broader community engagement, particularly with residential communities, would be needed before any deployment in a more diverse urban context. The assessors found the provider willing to discuss community impact questions in depth, including on topics where the current evidence is thin. The assessors note that Kourchberg's predominantly institutional character makes it a relatively simple community context. Engaging with more diverse communities, including residents, elderly populations, and people with mobility limitations, will be important for demonstrating that the system has been developed with the full range of affected people in mind, and would strengthen its position in procurement Governance with municipalities that have active community engagement requirements.

Environmental Impact: What is known about the system's energy consumption and environmental footprint?

An internal estimate from November 2026 puts UTIP's compute energy consumption at approximately 8 MWh per year for the 14-intersection pilot configuration, based on current cloud hosting usage. The provider also notes that adaptive signal control systems can reduce vehicle idling and associated emissions at intersections, an environmental co-benefit that is well-supported in the literature for systems of this type. A more rigorous lifecycle analysis and an extrapolation to full-deployment scale would be useful additions to the provider's evidence base for procurement discussions.

Gaps & Limitations: What could not be assessed at this point in time, and what additional evidence would be needed?

The impact assessment reflects the pilot stage and its limited community and environmental scope. Effects on diverse affected populations, fairness across road user types, and broader societal implications at scale cannot be fully assessed from a single-district supervised pilot. These are acknowledged gaps that the provider is aware of and has appropriately framed as development priorities rather than resolved questions. A richer impact picture will emerge as the pilot expands and as the provider engages with a wider range of communities and deployment contexts.

Overall Assessment Summary

The assessors find that LuxSmartMobility S.A. has approached the development and pilot of UTIP with technical care, an honest awareness of the system's current limitations, and a constructive attitude throughout the assessment process. The combination of controlled lab benchmarking and live pilot evaluation is methodologically sound. The operator-centric explainability features are a genuine strength. The pilot governance model, in which city operators retain primary control and monitoring is shared with the city, reflects responsible practice for a system at this stage. The areas for development identified across the three badges are consistent with a system at TRL 7 preparing for its first full deployment: broadening the evaluation to cover more diverse urban contexts and road user types; developing the governance, accountability, and transparency documentation needed for operational contracting; and extending impact analysis to cover a wider range of affected communities. None of these represent fundamental concerns about the system's design or the provider's intentions; they are the natural development agenda for a promising system approaching the market. The assessors draw particular attention to one design consideration that prospective municipal clients should be aware of: the current optimisation objective prioritises vehicle throughput, with pedestrian and active travel phase lengths treated as constraints. This is a common engineering default and not a disqualifying feature, but municipalities with strong active travel commitments will want to understand how this can be configured or extended before contracting. LuxSmartMobility S.A. has indicated openness to developing this capability, which the assessors regard as an important signal. The following development recommendations are offered to support the provider's preparation for full deployment:

Ref	Recommendation	Badge
R1	Extend the pilot to a mixed residential district and across a full seasonal cycle to broaden the performance evidence base.	Technical Testing
R2	Introduce disaggregated evaluation by road user type (pedestrians, cyclists, mobility-impaired persons) in the next pilot phase.	Technical Testing + Impact
R3	Consider extending the optimisation objective to incorporate pedestrian and active travel performance alongside vehicle throughput, and document the design rationale explicitly for municipal clients.	Technical Testing + Impact
R4	Finalise the human oversight procedures and update the risk assessment to incorporate pilot learnings before any full deployment.	Governance
R5	Develop a full-deployment DPIA and accountability framework as part of procurement preparation.	Governance
R6	Conduct structured engagement with residential and mixed-community stakeholders before any deployment beyond Kourchberg.	Impact
R7	Develop a quantified energy and emissions estimate for full deployment scale, incorporating the potential vehicle idling reduction co-benefit.	Impact

Reviewer Note

This report was reviewed by Dr. German Castignani on 23 March 2027. Dr. Castignani confirmed that the assessment findings are well-grounded, that the tone and calibration are appropriate for a TRL 7 system seeking the Label as a market signal, and that the recommendations are constructive and proportionate. He highlighted recommendation R3 as the finding most likely to be material in procurement conversations, and the assessors have given it appropriate prominence in the summary.