Requirements Document: Water LLM – Leak Detection in Water Networks

# 1. Document Control

|  |  |
| --- | --- |
| Project Name | Water LLM – Water Networks & Distribution |
| Use Case Name | Leak Detection |
| Prepared By | Business Analyst, TCS |
| Date | July 21, 2025 |
| Version | 1.0 |
| Reviewed By | Solution Architect, GenAI Engineering Team |

# 2. Purpose

This document captures the functional and non-functional requirements for implementing Leak Detection using GenAI within the Water LLM Engine. The objective is to proactively detect potential leaks using real-time and historical pressure and flow data, reducing non-revenue water and infrastructure damage.

# 3. Scope

- Applicable for urban and rural water distribution networks  
- Monitors real-time sensor inputs from pressure and flow meters  
- Supports predictive and anomaly-based leak identification  
- Integrates with SCADA, GIS, and Work Order systems  
- Delivers GenAI summaries and recommendations

# 4. Actors & Stakeholders

|  |  |
| --- | --- |
| Actor | Role & Responsibility |
| Field Engineer | Acts on leak alerts, verifies leaks on-ground |
| SCADA Operator | Monitors alerts in real-time, validates sensor health |
| Water LLM Engine | Processes sensor data, generates GenAI insights |
| Leak Detection Model | AI/ML model predicting leak probabilities |
| Supervisor / Manager | Approves field actions, reviews summaries |
| GIS System | Maps pipe segments and sensor zones |
| Work Order System | Issues and tracks maintenance tickets |

# 5. Use Cases

UC1: Leak Detection via Flow-Pressure Anomalies  
Trigger: Periodic or real-time data from sensors  
Input: Flow rate, pressure, expected vs. actual consumption  
Output: Leak likelihood, segment ID, location

UC2: GenAI Summary for Leak Insight  
Trigger: Leak identified or anomaly detected  
Input: AI model output, zone history, weather, past incidents  
Output: GenAI-generated narrative explaining root cause, urgency, impact

UC3: Auto-Triage & Work Order Generation  
Trigger: Confirmed leak > threshold severity  
Input: Leak location, urgency, available technicians  
Output: Work Order issued to field ops with details

UC4: Leak Zone Risk Scoring  
Trigger: Daily  
Input: Historical leak data, pipe age, material, terrain  
Output: Zone risk scores (High / Medium / Low)

UC5: Technician Advisory  
Trigger: Leak ticket creation  
Input: Leak type, location, weather, skillset DB  
Output: Recommended technician, safety precautions

# 6. User Stories

|  |  |  |  |
| --- | --- | --- | --- |
| ID | As a... | I want to... | So that... |
| US01 | SCADA Operator | receive real-time alerts on pressure drops | I can notify the field team promptly |
| US02 | Field Technician | view leak location and suggested action via mobile app | I can prepare the right equipment |
| US03 | Supervisor | get daily GenAI summaries of leak-prone zones | I can plan inspections more effectively |
| US04 | Data Scientist | access labeled historical leak data | to improve the leak prediction model |
| US05 | GenAI Engine | summarize incident causes and actions in natural language | to support field learning and documentation |
| US06 | GIS Analyst | overlay leak alerts on the network map | to visually validate correlation with terrain |
| US07 | Asset Manager | track frequently leaking segments | to plan capital investment |

# 7. Functional Requirements

* System must collect real-time pressure and flow data from all connected sensors
* AI model must calculate leak probability per segment every 15 minutes
* GenAI must generate summaries for confirmed leaks including cause, impact, and action
* Alerts must be routed to SCADA and field dashboard with severity tagging
* GIS must show leak segments in red for critical, orange for moderate, green for healthy
* Technician recommendations must consider distance, experience, and certifications
* Historical leak incidents must be exportable in CSV for model retraining
* Users should receive proactive notifications for high-risk zones

# 8. Non-Functional Requirements

* System uptime ≥ 99.9%
* Real-time data latency ≤ 10 seconds
* Alerts delivery time ≤ 1 minute
* Support for at least 10,000 sensors in parallel
* Compliance with GDPR, ISO 27001
* Interface must support desktop and mobile views
* Data archival policy: 2 years minimum for pressure/flow records

# 9. Data Requirements

|  |  |  |
| --- | --- | --- |
| Source | Type | Description |
| Pressure Sensor | Real-time | Per pipe segment, minute-wise |
| Flow Meter | Real-time | Inflow/outflow readings by junction |
| GIS | Static + live | Pipe layout, elevation, terrain info |
| Leak History DB | Historical | Verified past leaks |
| Work Orders DB | Operational | Completed and pending leak repairs |
| Weather API | Real-time | Temperature, rainfall, flood conditions |

# 10. Integration Requirements

- SCADA System → for sensor input and alert broadcasting  
- GIS Platform → to visualize pipe segments and sensor locations  
- Work Order System (Maximo/SAP/IFS) → for auto ticket generation  
- Cloud Storage (Azure/GCP/AWS) → for GenAI model and data lake  
- Notification Services (Email, SMS, App push)

# 11. Open Questions

|  |  |  |
| --- | --- | --- |
| Question | Owner | Status |
| What SCADA protocol is used (Modbus, OPC UA)? | IT Ops | Pending |
| Will leak confirmations be manually verified or sensor-only? | Field Ops | Under review |
| What is the expected frequency of GenAI summaries? | Management | TBD |

# 12. Risks & Mitigations

|  |  |
| --- | --- |
| Risk | Mitigation Strategy |
| False positives from sensor data | Combine AI + rule-based + manual verification |
| Incomplete historical leak records | Use data imputation, seek SME validation |
| Delays in field response | Prioritize critical segments, auto-assign techs |
| GenAI hallucinations in summary generation | Include structured output + review by supervisor |