# **CHAPTER 1. FUNCTIONAL REQUIREMENTS**

## **1.1 Solution Intent**

To form an architectural design, first understand the content of the solution itself that is being designed. Since for educational purposes we do not plan to form a commercial solution, we will limit ourselves to its MVP-version. Therefore, in this issue, it is first necessary to determine the high-level requirements for the very idea (Solution Intent) of creating software for processing biometric information. For this, an approach to creating a Business Case was used, which involves formulating the main idea and global characteristics (Epics and Features) of the designed application. It is worth adding that due to certain limitations on the scope of the final qualification work, an abbreviated version of the business case was used.

For this, the following milestones were identified for creating a solution and the corresponding justification of its architectural design: determining business requirements for creating a solution; justifying system requirements; forming an architectural design for the MVP-product model

The subject area of ​​the Solution design was chosen to be the field of work with Cloud‑Native Serverless Architecture for Real‑Time Sound Analytics in IoT. In recent years, it has been gaining increasing popularity. The use cases for software solutions providing the Real‑Time Sound Analytics in IoT are growing very rapidly.

**Business Context & Vision (Solution Intent)**

Real‑time analysis of environmental sound is critical for smart‑city scenarios such as traffic monitoring, public‑safety siren detection and urban noise mapping. However, the massive, highly‑variable sound event streams generated by distributed microphones and edge devices exceed the latency and cost envelopes of server‑centred architectures. This research proposes a cloud‑native serverless architecture that ingests, processes and analyses sound‑only IoT data with millisecond‑level end‑to‑end latency while maintaining pay‑per‑use cost efficiency. Building on recent advances in Function‑as‑a‑Service (FaaS), event‑driven pipelines and intelligent pre‑warming, we will design, implement and experimentally evaluate multi‑cloud prototypes on AWS Lambda and Azure Functions. Controlled workloads derived from UrbanSound8K, CitySoundscapes and ESC‑50 will be replayed through IoT simulators to benchmark cold‑start delay, throughput, accuracy of online audio classification and cost‑performance trade‑offs. Deep‑reinforcement‑learning (DRL) policies will be investigated for proactive scaling. The expected outcome is an empirically‑validated, cost‑aware orchestration blueprint and a set of open research artifacts (datasets, code, dashboards) that advance the state‑of‑the‑art in serverless audio analytics.

**1.1 Stakeholders & Benefits**

|  |  |
| --- | --- |
| Stakeholder | Key Benefits |
| City Operations / DOT | Live traffic flow characterization; congestion and incident detection; heatmaps for planning. |
| Public Safety (911, EMS) | Automatic siren detection, escalation triggers, and response ETA analytics. |
| Urban Planning / Environment | Noise mapping by street/zone; compliance reporting; long‑term trends. |
| Cloud / Platform Engineering | Multi‑cloud IaC, automated scaling, observability, and cost control. |
| Data Science / Research | High‑quality labeled streams; reproducible experiments; model performance dashboards. |

The Stakeholders & Benefits identifies the main groups who would use or rely on the cloud-native serverless sound analytics solution and highlights the value each group receives. City operations and transportation departments benefit from live traffic characterization and congestion detection to improve planning and response. Public safety agencies gain rapid siren detection and automatic alerts that can shorten emergency response times. Urban planners and environmental authorities obtain noise mapping and compliance insights that inform long-term policy decisions. Finally, cloud engineering and research teams benefit from reproducible multi-cloud deployments, improved observability, and high-quality datasets for experimentation and innovation.

**1.2 Core Value Proposition**

The proposed architecture is designed to achieve millisecond-level end-to-end latency, a requirement that ensures timely and reliable operational decision-making in real-time urban environments. Beyond latency, the system emphasizes elastic scalability, maintaining resilience under bursty workloads of up to 10,000 requests per second without resorting to costly over-provisioning. From an economic perspective, the architecture is guided by the principle of cost efficiency, achieved through a pay-per-use model complemented by intelligent pre-warming strategies and adaptive memory allocation. Finally, the solution prioritizes portability, providing a multi-cloud blueprint that leverages Infrastructure-as-Code (Terraform, AWS CDK, Azure Bicep) to ensure reproducibility and seamless deployment across heterogeneous cloud providers.

**1.3 High‑Level Use‑Case Domains**

Application Domains

|  |  |  |
| --- | --- | --- |
| Domain | Example Uses | Stakeholder Value |
| Traffic monitoring | Vehicle density estimation, honk events, congestion alerts | Faster incident response; planning insights |
| Public‑safety siren detection | Siren classification & georouting | Prioritization for emergency services |
| Urban noise mapping | Zonal SPL/Leq estimates, temporal heatmaps | Compliance & policy making |
| Industrial sound monitoring | Anomaly detection on machinery spectra | Reduced downtime; predictive maintenance |
| Environmental & wildlife | Species / event acoustic detection | Biodiversity studies; conservation actions |

The table of high-level use case domains outlines the breadth of applicability for cloud-native serverless sound analytics in contemporary smart-city and industrial contexts. One of the most prominent domains is traffic monitoring, where audio data such as vehicle noise and honking can be leveraged to infer traffic density, detect congestion, and provide actionable insights for transport authorities. In this domain, the proposed architecture contributes to reducing delays in incident detection and enables the generation of real-time traffic heatmaps that support both operational responses and long-term infrastructure planning. A second domain is public-safety siren detection, which involves automatically classifying emergency vehicle sirens and georouting them to dispatch centers. This application is particularly important because it shortens emergency response times, ensures faster prioritization of road clearance, and enhances situational awareness for emergency services.

The urban noise mapping domain addresses the growing concern of environmental sound pollution by producing zonal measurements of sound pressure levels and aggregated temporal heatmaps. Such data are valuable not only for compliance and regulatory purposes but also for guiding urban planning strategies that aim to improve citizens’ quality of life. Beyond city governance, industrial sound monitoring constitutes another relevant domain, where acoustic signatures of machinery can be analyzed to detect anomalies and prevent breakdowns. This proactive form of monitoring can reduce downtime, lower maintenance costs, and contribute to predictive maintenance frameworks already in use across Industry 4.0 settings.

Finally, the environmental and wildlife monitoring domain extends the scope of sound analytics to ecological applications. Acoustic event detection in natural habitats provides a non-intrusive means of tracking species presence, migration, and ecosystem health. Such insights are indispensable for conservation initiatives and biodiversity assessments, particularly in areas where traditional observation methods are impractical. Taken together, these use case domains illustrate the versatility and societal value of real-time sound analytics, demonstrating its potential to deliver both operational efficiency and long-term sustainability outcomes across diverse sectors.

## **2.2 Users specification**

The Users Specification table delineates the principal categories of stakeholders who will engage with the cloud-native serverless architecture for real-time sound analytics in IoT.

|  |  |  |  |
| --- | --- | --- | --- |
| Users Group | Users’ Basic Needs | Features of the Solution Explored | Primary Use Cases |
| Municipal Authorities & City Operators | Real-time traffic insights, noise compliance monitoring, long-term planning data | Event-driven ingestion pipeline; real-time latency monitoring; Grafana dashboards | Traffic flow monitoring, congestion detection, urban noise mapping |
| Public Safety & Emergency Services | Rapid detection of sirens/alarms, situational awareness, reduced response times | CNN-based audio classification (ESC-50 baseline); serverless pre-warming for low latency | Siren detection and georouting, anomaly alerts in urban areas |
| Cloud & Platform Engineers | Cost control, scalability, multi-cloud deployment, reproducibility | IaC templates (Terraform, AWS CDK, Azure Bicep); autoscaling with DRL (StableBaselines3 PPO); monitoring | Deployment automation, cross-cloud testing, observability & cost optimization |
| Data Scientists & AI Researchers | Reproducible datasets, metrics for evaluation, experimentation environment | Pretrained CNN models; feature extraction with librosa; Jupyter notebooks; public GitHub repository | Benchmarking classification accuracy, training/testing new models, publishing artifacts |
| Industrial Operators & Environmental Agencies | Predictive maintenance, regulatory compliance, ecological monitoring | Acoustic anomaly detection; workload replay simulator (UrbanSoundReplayer); metrics collection APIs | Industrial machinery monitoring, compliance noise mapping, wildlife acoustic tracking |

The first group, municipal authorities and city operators, represents entities responsible for transportation management and environmental governance. Their primary need is the acquisition of timely and accurate acoustic data to inform operational decisions, enforce compliance, and plan long-term infrastructural investments. The solution addresses these needs by offering event-driven data ingestion, real-time latency monitoring, and intuitive dashboards that support both immediate interventions and strategic planning.

The second group, public safety and emergency services, relies on rapid acoustic classification for enhanced situational awareness. Their operational efficiency depends on the capacity to detect and interpret signals such as sirens or alarms in real time. The architecture’s integration of pretrained convolutional neural networks (CNNs), combined with serverless pre-warming strategies, ensures low-latency inference. In practice, this allows emergency services to shorten response times and prioritize interventions based on the geographical routing of detected acoustic events.

A third group, cloud and platform engineers, requires robust mechanisms for deployment, scalability, and cost optimization across multiple cloud providers. Their work is facilitated through Infrastructure-as-Code templates and reinforcement learning–driven autoscaling policies, which guarantee both reproducibility and cost efficiency. For these users, the system is not merely a research tool but an operational framework that ensures high availability and optimized resource allocation.

The fourth group, data scientists and AI researchers, depends on reproducible datasets and transparent evaluation pipelines. The solution provides them with open-source resources such as pretrained CNN models, Jupyter notebooks, and metrics repositories. By supporting rigorous benchmarking and iterative experimentation, the system enables researchers to improve classification accuracy and contribute new models back to the community.

Finally, industrial operators and environmental agencies constitute an extended but equally important user group. Industrial operators require predictive maintenance and anomaly detection in acoustic signatures of machinery, while environmental agencies depend on sound-based monitoring for regulatory compliance and ecological conservation. For both categories, the system provides advanced tools for anomaly detection, synthetic workload replay, and high-frequency metrics collection. These capabilities allow stakeholders to anticipate equipment failures, reduce downtime, and monitor biodiversity in non-intrusive ways.

Taken together, the table illustrates the multi-stakeholder relevance of the proposed solution. It demonstrates that real-time sound analytics, when implemented via a serverless multi-cloud architecture, serves not only operational efficiency but also regulatory, scientific, and environmental objectives. By addressing the specific needs of diverse user groups, the architecture maximizes societal, economic, and scientific impact, positioning itself as a versatile and sustainable technological innovation.

## **2.3 Use Journey definition**

Basic Uses Cases for the developed solution are given in Appendix A.

The real-time traffic monitoring use case (UC-1) is designed to support urban mobility management by leveraging continuous acoustic data streams from distributed road-side microphones. The main flow involves the ingestion of audio via the IoT Hub, feature extraction through librosa, and subsequent classification of traffic-related acoustic events (e.g., honking, congestion signatures) using pretrained CNN models. Results are aggregated into dashboards that allow city operators to make evidence-based interventions. Potential exceptions include network disruptions that may cause partial data loss, misclassification due to overlapping sound events, or cloud function cold-start delays that temporarily degrade responsiveness. Nevertheless, the benefits are significant: traffic management authorities gain timely indicators for congestion mitigation, law enforcement agencies can respond more effectively to incidents, and urban planners obtain longitudinal datasets that inform infrastructure design (Picture 1).

A diagram of a company

AI-generated content may be incorrect.

Picture 1. Flow Chart for the UC-1

The public-safety siren detection use case addresses the need for rapid identification and localization of emergency vehicles in dense urban environments. The main flow involves real-time ingestion of acoustic signals, detection of siren patterns through CNN-based classification, and geolocation annotation of events for dispatch centers. This automated pipeline minimizes reliance on manual reporting and reduces delays in emergency response coordination. Possible exceptions may include false positives triggered by acoustically similar events (e.g., construction noise) or degraded accuracy in high-noise urban corridors. However, the benefits are distributed across multiple stakeholders: dispatch operators gain actionable alerts that improve response routing, emergency services experience reduced arrival times at incident sites, and citizens benefit from faster and more reliable public-safety interventions (Picture 2).

A diagram of a diagram

AI-generated content may be incorrect.

Picture 2. Flow Chart for the UC-2

.

Urban noise mapping exemplifies the integration of acoustic analytics into regulatory and planning contexts. The main flow involves continuous ingestion of environmental sound, computation of aggregated acoustic indices (e.g., SPL, Leq), and spatial visualization through temporal heatmaps. These outputs provide regulators and urban planners with empirical insights into noise patterns across city zones. Exceptions may occur in the form of incomplete coverage due to sensor deployment constraints, calibration drift in edge devices, or temporary loss of connectivity between IoT endpoints and the cloud pipeline. The benefits, however, are substantial: environmental agencies gain robust compliance reports for enforcement, municipal governments can design targeted noise mitigation strategies, and citizens indirectly benefit from improved quality of life through better noise control policies (Picture 3).

A screenshot of a computer

AI-generated content may be incorrect.

Picture 3. Flow Chart for the UC-3

.

Industrial acoustic monitoring applies real-time sound analytics to manufacturing and production environments. The main flow entails capturing machinery sound signatures, extracting spectral features, and detecting anomalies indicative of early equipment failure. Alerts are generated and stored for diagnostic use, providing maintenance teams with actionable insights. Potential exceptions include false alarms triggered by temporary operational variations (e.g., tool changes or non-critical vibrations) or failure to detect anomalies when background noise masks the acoustic patterns of interest. Despite these risks, the benefits are clear: plant operators reduce unplanned downtime, maintenance engineers improve efficiency through predictive scheduling, and management achieves cost savings and improved safety. The same framework also supports industrial compliance with health and safety regulations (Picture 4).

A diagram of a flowchart

AI-generated content may be incorrect.

Picture 4. Flow Chart for the UC-4

.

Environmental and Wildlife Monitoring use case extends the solution into ecological research and biodiversity conservation. The main flow consists of ingesting environmental acoustic data from edge devices in natural habitats, identifying species or ecological events, and storing annotated data for long-term analysis. Outputs are used to track species presence, migration, and overall ecosystem health. Exceptions may arise from limited connectivity in remote regions, misclassification of overlapping animal sounds, or environmental conditions (e.g., storms) obscuring audio signals. Nonetheless, the benefits are multifaceted: ecologists gain a scalable and non-intrusive method of monitoring wildlife, conservation agencies access high-quality datasets for biodiversity assessments, and policymakers benefit from evidence-based recommendations for environmental protection (Picture 5).

A diagram of a computer

AI-generated content may be incorrect.

Picture 5. Flow Chart for the UC-5

Across all five use cases—Traffic Monitoring (U1), Siren Detection (U2), Noise Mapping (U3), Industrial Acoustic Monitoring (U4), and Environmental & Wildlife Monitoring (U5)—exception handling is essential to maintain the reliability, validity, and operational resilience of the serverless sound analytics pipeline. Although each domain introduces specific contingencies, several categories of exceptions can be generalized.

1. Data Integrity and Payload Validation

In all scenarios, corrupted or incomplete ingestion payloads represent a fundamental exception. These may arise from device malfunction, network packet loss, or encoding errors. The system applies strict schema validation, rejecting invalid frames while logging events for later inspection. This ensures that downstream models are not exposed to unreliable inputs.

2. Sensor Degradation and Calibration Drift

Environmental sensors, industrial microphones, and traffic detectors are all vulnerable to hardware wear, environmental contamination (dust, moisture), or calibration loss. Exceptions of this type are managed by cross-sensor consistency checks, self-diagnostic routines, and exclusion of suspect data. This preserves analytic quality across the full use-case spectrum.

3. Acoustic Ambiguity and Masking Effects

Whether in urban streets, factories, or wildlife habitats, overlapping sound sources often produce ambiguous signals. Common confounders include background noise (wind, rain), transient bursts (construction, fireworks), and acoustic look-alikes (musical sirens, species calls). Exceptions of this type are mitigated via confidence thresholds, persistence rules (N-of-M verification), and secondary classification passes.

4. Latency Breaches and System Backpressure

Since all use cases are real-time in nature, exceeding the end-to-end latency service-level objective (SLO) is a critical exception. Backpressure may arise from bursty workloads (e.g., >10k req/s in U1) or unexpected event surges (e.g., mass siren activations in U2). Exception handling relies on dynamic autoscaling policies, workload shedding of non-critical tasks, and graceful degradation (e.g., widening aggregation windows).

5. Model Drift and Misclassification

All use cases depend on machine learning classifiers. Over time, acoustic distributions shift due to seasonal patterns, machinery wear, or evolving urban soundscapes. Drift manifests as increased false positives or false negatives. Exceptions are managed through continuous monitoring of classification confidence, scheduled retraining, and human-in-the-loop validation for borderline cases.

6. Connectivity Loss and Sparse Coverage

Field-deployed IoT sensors frequently operate under constrained connectivity. Exceptions occur when network outages delay or prevent data arrival. To mitigate this, the system incorporates tolerance for late-arriving data, interpolates across neighboring sensors, and generates sensor health alerts when coverage gaps persist.

7. Governance, Safety, and Compliance Exceptions

Some exceptions are regulatory or safety-related, rather than purely technical. For example, triggering a false road-closure alert (U2) or a premature industrial shutdown (U4) carries socio-economic risks. To prevent cascading errors, these exceptions are addressed with dual-verification protocols, escalation to human operators, and comprehensive audit logging.

# **CHAPTER 2. NON-FUNCTIONAL REQUIREMENTS AND ARCHITECTURAL DESIGN**

## **2.1 General architecture overview**

Here to provide the general explanation of arch decisions and patterns for the solution.

## **2.2 Architecture definition**

## **2.2.1 Component View diagram**

Here to provide the diagram with its description for 12-15 sentences

## **2.2.2 UML Class diagram**

Here to provide the diagram with its description for 12-15 sentences

## **2.2.3 UML Sequence diagram**

Here to provide the diagram with its description for 12-15 sentences.

!!! provide brief explanation of how different solution components are integrated, what is the communication flow across them

## **2.2.4 Architecture Diagram**

Here to provide the diagram with its description for 15-20 sentences.

!!! this diagram is the generalization for the solution architecture

## **2.3 NFRs specification**

Provide some tables with their description including the following:

* Security requirements (mandatory!!!)
* Performance requirements (mandatory!!!)
* Resilience and Availability (include DR description) (mandatory!!!)
* Scalability (mandatory!!!)
* Maintainability (mandatory!!!)
* Other NFRs (optionally based on your opinion)

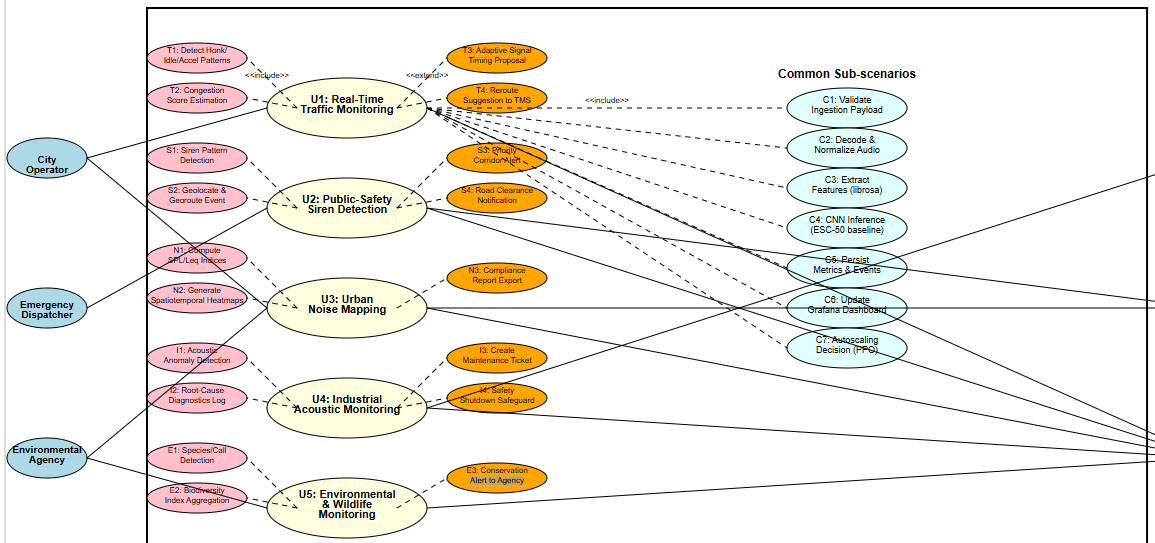
## **2.4 Data flow**

Provide some information about data in transit and rest. If DB available – please provide the DV schema including its description (15-20 sentences)

.

Appendix A

**Use Case Specification**



Picture 1. Use Case Specification Diagram

**Use Case 1: Real-Time Traffic Monitoring**

|  |  |
| --- | --- |
| Field | Description |
| Actors | City operators, traffic management systems |
| Preconditions | Microphones deployed at intersections and road segments stream audio to the IoT Hub. |
| Basic Flow | The serverless pipeline ingests audio, extracts spectral features using librosa, and classifies events such as honking or congestion indicators via pretrained CNN models. Metrics are aggregated into dashboards that provide near real-time situational awareness. |
| Postconditions | Operators receive alerts and visualizations that support immediate traffic control actions, such as rerouting or adjusting signal timing. |
| Primary Benefit | Reduced congestion and enhanced planning through evidence-based traffic management. |

**Use Case 2: Public-Safety Siren Detection**

|  |  |
| --- | --- |
| Field | Description |
| Actors | Emergency dispatch centers, law enforcement agencies |
| Preconditions | Continuous audio ingestion from urban microphones with connectivity to the serverless pipeline. |
| Basic Flow | Audio snippets are processed for siren patterns, and the CNN model classifies the presence of emergency vehicles. A geolocation service annotates events and triggers real-time alerts. |
| Postconditions | Dispatchers obtain actionable notifications to optimize response routing and prioritize critical interventions. |
| Primary Benefit | Faster emergency response and improved situational awareness during crises. |

**Use Case 3: Urban Noise Mapping**

|  |  |
| --- | --- |
| Field | Description |
| Actors | Environmental agencies, urban planners |
| Preconditions | Continuous collection of audio data across multiple city zones. |
| Basic Flow | The architecture processes incoming streams, estimates sound pressure levels (SPL) and aggregated measures such as Leq, and generates spatiotemporal heatmaps. Data are stored in cloud repositories for long-term trend analysis. |
| Postconditions | Agencies gain access to regulatory compliance reports and interactive dashboards for policy design. |
| Primary Benefit | Evidence-driven strategies to mitigate noise pollution and improve quality of life. |

**Use Case 4: Industrial Acoustic Monitoring**

|  |  |
| --- | --- |
| Field | Description |
| Actors | Industrial plant operators, maintenance engineers |
| Preconditions | Microphones or sensors installed in factory floors or machinery environments. |
| Basic Flow | Real-time streaming detects abnormal acoustic signatures via CNN inference. Detected anomalies trigger proactive alerts and logs are archived for diagnostic analysis. |
| Postconditions | Maintenance teams are informed about potential failures before they escalate. |
| Primary Benefit | Reduced downtime and improved operational efficiency through predictive maintenance. |

**Use Case 5: Environmental and Wildlife Monitoring**

|  |  |
| --- | --- |
| Field | Description |
| Actors | Ecologists, conservation agencies |
| Preconditions | Edge devices deployed in natural habitats with connectivity to the cloud. |
| Basic Flow | The system ingests environmental sounds, identifies species or specific ecological events, and stores annotated data in cloud repositories. Periodic analytics generate biodiversity indicators. |
| Postconditions | Conservation teams receive validated data that support ecological studies and conservation strategies. |
| Primary Benefit | Non-intrusive monitoring of biodiversity and evidence-based conservation actions. |