V5.0 Unified Specification

This is the definitive, final V5.0 specification. By integrating all these components, you are creating the **Complete Multimodal Gateway** that eliminates all remaining technical debt and activates the full PAD vision, ensuring the system is both production-hardened and infinitely scalable via transfer learning.

Part 1: Strategic & Functional Requirements (SRS)

1.1 Core Reliability & Tracing (New V5.0 Focus)

- **FR-01 (Idempotency Check):** The Worker **MUST** check a shared state store (e.g., Firestore/Redis) upon receiving a new task_id. If the task is completed, it **SHALL** immediately ACK and exit.
- **FR-02 (Distributed Tracing):** The Dispatcher **MUST** generate a unique OpenTelemetry trace_id for every task and inject it into the Pub/Sub message attributes.
- **FR-03 (Trace Propagation):** The Worker **SHALL** propagate this trace_id across all downstream calls, including the Marketplace API and the Agent Topic.

1.2 Model Adaptability (Transfer Learning)

- **FR-10 (Tokenizer Persistence):** The initial training run (on computer data) **MUST** save the Tokenizer object alongside the model weights. The Temporal Encoder (Service C) must load this Tokenizer to ensure token IDs are consistent across all domains.
- **FR-11 (Domain Fine-Tuning Mode):** Service C's model pipeline **MUST** support a method (via configuration) to load the initial V4 weights and allow fine-tuning on a new, smaller domain-specific training dataset.
- **FR-12 (Full Vocabulary Inclusion):** The fine-tuning process **MUST** include new, domain-specific vocabulary (e.g., 'Suture', 'Preheat') and update the model's Embedding layer accordingly, preventing data corruption.

1.3 Execution, Multimodality & Monetization

- **FR-04 (Execution Trigger):** The Worker **MUST** publish the final Pathway.json URI to the **Agent-Execution Pub/Sub Topic** (pad-agent-tasks) to trigger the Risk and Simulation Agents.
- **FR-05 (Marketplace Delivery):** The Worker **SHALL** make an authenticated POST call to the Pathway Marketplace Ingestion API to register the new asset.
- **FR-06 (Multimodal Audio):** The Worker **SHALL** implement logic to extract and transcribe the video's audio track, appending the transcript to the LLM prompt for fusion (FR-07).
- **FR-07 (Pixel Accuracy):** The refinement step **MUST** integrate a dedicated **Object Detection Model** (Service D) to predict and verify pixel-accurate ui_region bounding boxes.

- **FR-08 (Compliance Lock):** The Worker **MUST** enforce the mandatory presence of all v0.5 metadata fields (license_tier, compliance_tag) before submission (FR-07).
- **FR-09 (IoT Bridge Hook):** The Worker **SHALL** implement logic to process and link optional **DED Machine Telemetry** data into the pathway's telemetry_context block.

Part 2: Technical Design Document (TDD)

2.1 Final Architecture Update: Five Services, Two Topics

- **Service A: Dispatcher (V2)**: Frontend API and Trace Injector.
- **Service B: Worker (V3)**: Core Orchestrator.
- **Service C: Temporal Encoder (V4)**: The **LSTM Memory Engine**.
- **Service D: Object Detector (New)**: The **Pixel Accuracy Engine** (YOLO/Vision Transformer).
- **Service E: Risk Agent (New)**: The **Execution/Consumption Engine** (subscribing to pad-agent-tasks).

2.2 Worker Service (Service B) Execution Flow (The V5 Pipeline)

- **Deployment/Setup:** The Worker must load the **Tokenizing Model and Weights** from the deployment package during startup (FR-10).
- **Ingestion:** Run ffmpeg to extract audio track and call STT API (FR-06).
- **Visual Processing:** Call **Gemini 2.5 Pro** with video +transcript (FR-06) to get $\text{semantic\_description}$ and target_text.
- **Pixel Accuracy:** Call **Object Detector (Service D)** with the target_text (FR-07).
- **Sequential Encoding:** Call **Temporal Encoder (Service C)** to get the $\text{temporal\_context\_vector}$.
- **Submission:** Perform **Marketplace API Call** (FR-05) and publish URI to **pad-agent-tasks Topic** (FR-04).

2.3 Temporal Encoder Deployment & Adaptability

- **Deployment Package Requirement (FR-10):** The Service C Docker container **MUST** package the following files: lstm_model.h5 and **tokenizer.pickle**.
- **Service C Startup Logic (FR-10):** $\text{encoder\_app/main.py}$ **MUST** load the tokenizer.pickle file during startup before model weights are loaded.
- **Fine-Tuning Logic (FR-11/FR-12):** The local training script (train_lstm.py) must implement a function to:
    - Load existing lstm_model.h5 weights.
    - Load the tokenizer.pickle file.

    - **Resize the Embedding Layer** to accommodate new domain vocabulary while retaining existing weights (enabling transfer learning).

2.4 Final Schema (PAD v0.5)

- **Metadata Block (Business):** Now includes mandatory compliance_tag and license_tier fields (FR-08).
- **Node Block (ActionNode):** Includes $\text{full\_audio\_transcript\_segment}$ and the telemetry_context block (FR-09).
- **Audit Block (Reliability):** Includes trace_id and $\text{idempotency\_check\_passed}$ (FR-01/FR-02).