

FIELD MANUAL · VOLUME 15

---

# Capture. Control. Confidential.

The room hears. The room understands. Nothing leaves the room.

---

PREPARED FOR

Operators · Project Managers · Designers · Integrators · Owner's Reps

---

CONTACT

**Restrepo Innovations**

Franklin Lakes, NJ · [mike@restrepoinnovations.com](mailto:mike@restrepoinnovations.com)  
[restrepoinnovations.com](http://restrepoinnovations.com) · Phone on file with PM

PARTNER

**Island Media Systems**

Audio · Video · Network · Lighting · Shades  
One team. One standard. One execution.



MORE · [RESTREPOINNOVATIONS.COM](http://RESTREPOINNOVATIONS.COM)

MANIFESTO

# The room hears. The room understands. Nothing leaves the room.

Vol 15 governs the design and installation of local artificial intelligence in private residential conference rooms, boardrooms, and home offices. It is a hardware, audio, network, and discipline standard. Every AI workflow specified in this volume runs on owner-controlled silicon, on an isolated VLAN, with no inference traffic crossing the residential WAN. The conference room is a private surface. Vol 15 keeps it that way.

SCOPE

Residential-only. Applies to private conference rooms, boardrooms, executive offices, and family-office meeting suites inside Restrepo Innovations luxury residential builds. Specifies hardware tiers, software stack, audio capture, network discipline, and pre-rack coordination. Cross-references Vol 14 (AI in the Field) for operator discipline and Vol 12 (HT Installation) for rack and acoustic principles.

CONTENTS

|    |                           |       |
|----|---------------------------|-------|
| 01 | Why Local AI              | p. 5  |
| 02 | The Listening Room        | p. 8  |
| 03 | Hardware Tiers            | p. 11 |
| 04 | The Software Stack        | p. 14 |
| 05 | Audio Capture             | p. 18 |
| 06 | The AI Appliance          | p. 21 |
| 07 | Network and Privacy Stack | p. 24 |
| 08 | Field Workflows           | p. 28 |
| 09 | Confidentiality Protocol  | p. 30 |
| 10 | Cross-System Integration  | p. 33 |



|    |                           |       |
|----|---------------------------|-------|
| 11 | Without Written Direction | p. 37 |
| 12 | Pre-Rack Coordination     | p. 40 |
| 13 | QC · Audit · Closeout     | p. 43 |

---



PART 1

---

# Foundation & Hardware

Sections 01–04 — Why Local AI, The Listening Room, Hardware Tiers, The Software Stack

## § 01

# Why Local AI

## PART I — FOUNDATION & HARDWARE

Cloud AI services route every spoken word, every uploaded document, every prompt across the open internet to a third-party data center. For a residential conference room used by attorneys, family-office principals, board members, or healthcare-adjacent advisors, that traffic profile is unacceptable. Vol 15 specifies a local-only AI appliance: an Apple Silicon Mac Studio or Mac Mini running open-source inference engines, isolated on a dedicated VLAN, with no outbound inference traffic.

### THE FOUR PRIVACY POSTURES

Cloud AI (Otter.ai, Zoom AI Companion, Microsoft 365 Copilot) processes conversation audio in vendor-controlled infrastructure. Otter.ai stores transcripts indefinitely and has been the subject of consent litigation under California Penal Code § 632 and similar two-party consent statutes. Microsoft 365 Copilot transmits prompts and selected document context to Microsoft Azure OpenAI tenancies under Microsoft's Data Protection Addendum. Zoom AI Companion processes meeting audio under Zoom's service terms, with summary data retained on Zoom infrastructure.

Hybrid AI (Apple Intelligence with Private Cloud Compute) processes lightweight tasks on-device and escalates complex tasks to Apple's Private Cloud Compute. Apple PCC offers strong cryptographic attestation and stateless processing, but the data still leaves the home for inference.

Local AI (Ollama, llama.cpp, Whisper.cpp on owner-controlled hardware) executes every inference on-premises. Audio never leaves the conference room. Transcripts are written to local storage on the AI appliance. The Restrepo standard.

#### NOTE

#### Regulatory floor

GDPR Chapter 5 prohibits transferring personal data outside the EEA without an approved transfer mechanism. HIPAA Security Rule (45 CFR 164.302–164.318) requires covered entities to limit PHI exposure. Local-only inference avoids both transfer and exposure entirely. NIST AI RMF MAP and MEASURE functions become tractable when the model is owner-controlled.

**BEST PRACTICE**

## **The Restrepo posture**

If audio enters the conference room, it stays in the conference room. Inference runs on the appliance. Transcripts are written to the appliance. Search runs against a vector database on the appliance. Nothing leaves the room.

## **PRIMARY SOURCES**

GDPR Chapter 5 — Transfers · HHS HIPAA Security Rule · NIST AI RMF · Apple Private Cloud Compute



§ 02

# The Listening Room

PART I — FOUNDATION & HARDWARE

A conference room that captures speech for AI transcription is an acoustic instrument. Reverberation, HVAC noise, and asymmetric speaker positioning degrade word error rate. Whisper-class transcription that scores 5 percent WER in studio conditions can degrade to 15-20 percent WER in an untreated room. Vol 15 specifies the acoustic envelope before the AI appliance is selected.

ACOUSTIC TARGETS

| Metric                          | Target            | Notes                                             |
|---------------------------------|-------------------|---------------------------------------------------|
| RT60 (reverberation)            | 0.4–0.6 seconds   | Mid-band; longer rooms drift past 0.7s untreated  |
| Background noise (NC)           | NC-25 or lower    | HVAC and AV-rack fans dominant; isolate or remote |
| Speech transmission index (STI) | ≥ 0.65            | Required for clean Whisper transcription          |
| Ceiling height                  | 9–11 ft preferred | Drives ceiling-array mic spacing                  |
| Wall treatment coverage         | ≥ 30 percent      | Diffusion behind speakers; absorption side walls  |

ROOM ARCHETYPES

Tier 1 boardroom (16-24 ft, 8-14 seats): ceiling-array microphone with beamforming, ceiling speakers on a separate zone, in-wall acoustic treatment behind the display wall, rack housed in adjacent IT closet. Sennheiser TCC 2 or Shure MXA920 specified.

Tier 2 family office (12-16 ft, 6-8 seats): ceiling-array microphone or premium boundary array, ceiling or in-wall speakers, in-rack AI appliance acceptable if rack is isolated. Shure MXA710 boundary or MXA920 specified.

Tier 3 home office (under 12 ft, 1-4 seats): high-quality desk microphone or boundary array, near-field monitoring, on-desk or under-desk AI appliance. Shure MV7 or Audio-Technica AT2020USB-X specified for small rooms.

**FIELD RULE**

**Acoustic before AI**

Treat the room before specifying the appliance. A Mac Studio M2 Ultra cannot recover speech that the microphone never captured cleanly. Acoustic treatment, microphone selection, and gain structure precede model selection on every Restrepo job.

**DO NOT**

**Without written direction**

Do not deploy ceiling-array microphones on rooms with RT60 above 0.8 seconds untreated. Do not place the AI rack inside the conference room without rack-room acoustic isolation. Do not power microphones from the same circuit as the AI appliance.



§ 03

# Hardware Tiers

PART I — FOUNDATION & HARDWARE

Three hardware tiers cover the residential conference-room landscape. Tier 1 is specified for principals, attorneys, and family-office boardrooms running 70-billion parameter models with concurrent transcription. Tier 2 is specified for executive offices and small boardrooms running 13-30B models. Tier 3 is specified for personal home offices running 7B models for summarization and lightweight search.

THE TIER MATRIX

| Tier   | Platform            | RAM      | Models                              | Concurrency                |
|--------|---------------------|----------|-------------------------------------|----------------------------|
| Tier 1 | Mac Studio M2 Ultra | 192 GB   | Llama 3.3 70B Q4 + Whisper Large-v3 | 2-3 simultaneous workflows |
| Tier 2 | Mac Mini M4 Pro     | 64 GB    | Llama 3.1 30B Q4 + Whisper Medium   | 1-2 simultaneous workflows |
| Tier 3 | Mac Mini M4 base    | 16-24 GB | Llama 3.2 8B + Whisper Small        | 1 workflow at a time       |

WHY APPLE SILICON

Apple Silicon Mac Studio and Mac Mini deliver unified memory architecture: the GPU and CPU share the same memory pool. A 70B parameter model that requires 40 GB of model weights at Q4 quantization runs natively on a 192 GB M2 Ultra without GPU-CPU memory transfers. Equivalent x86 + RTX 4090 builds require model splitting, active cooling, and 600-800 W power draw under load. The Mac Studio draws 295 W maximum and runs silent.

ALTERNATIVE: NVIDIA WORKSTATION

Where the owner requires CUDA-only frameworks (vLLM tensor-parallel, specific research libraries), specify a single-node 2U workstation with one RTX 4090 24 GB or one RTX 6000 Ada 48 GB. Acoustic isolation cabinet required. 493 W maximum power draw. Not the Restrepo default.

SPECIFICATION

Specification rule

Default to Apple Silicon. Specify NVIDIA only with written design direction citing a CUDA dependency. Document the dependency in the project knowledge base.

PRIMARY SOURCES

Apple Mac Studio Specs · Apple Mac Mini Specs · NVIDIA RTX 4090



§ 04

# The Software Stack

PART I — FOUNDATION & HARDWARE

The Restrepo AI stack is open source. No vendor lock-in. No license renewals. No telemetry. Six components are specified by name: Ollama for model serving, Open WebUI for the user interface, Whisper.cpp for transcription, Piper for text-to-speech, llama.cpp as the inference engine, and ChromaDB for vector search.

STACK MATRIX

| Component   | Purpose                                            | License    | Port     |
|-------------|----------------------------------------------------|------------|----------|
| Ollama      | Model server REST API on localhost                 | MIT        | 11434    |
| Open WebUI  | Browser-based chat interface                       | MIT        | 3000     |
| Whisper.cpp | Local speech-to-text (Apple Silicon Neural Engine) | MIT        | n/a      |
| Piper TTS   | Local text-to-speech for voice playback            | MIT        | n/a      |
| llama.cpp   | Inference engine (used by Ollama)                  | MIT        | internal |
| ChromaDB    | Vector database for project knowledge search       | Apache 2.0 | 8000     |

RECOMMENDED MODELS BY TIER

| Tier | Chat model               | Embeddings                | Whisper           |
|------|--------------------------|---------------------------|-------------------|
| 1    | Llama 3.3 70B Q4 (40 GB) | BGE-M3 (568 MB)           | Large-v3 (1.5 GB) |
| 2    | Llama 3.1 30B Q4 (18 GB) | nomic-embed-text (274 MB) | Medium (1.4 GB)   |
| 3    | Llama 3.2 8B Q4 (4.7 GB) | nomic-embed-text (274 MB) | Small (483 MB)    |

NOTE

Model lifecycle

Models are pinned at the version installed at commissioning. Owner-approved updates follow the Vol 14 verification protocol. The AI appliance does not auto-update models.

## PRIMARY SOURCES

Ollama GitHub · Open WebUI · Whisper.cpp · Piper TTS · llama.cpp · ChromaDB



PART 2

---

# Audio & Network Infrastructure

Sections 05–07 — Audio Capture, The AI Appliance, Network and Privacy Stack

§ 05

# Audio Capture

PART II — AUDIO & NETWORK INFRASTRUCTURE

Microphone selection determines transcription quality more than model selection. Vol 15 specifies four approved capture devices and the room conditions each is matched to. All four output Dante or AES67 to a network audio bridge that feeds the AI appliance over the audio VLAN.

APPROVED MICROPHONES

| Device             | Coverage           | Output                    | Application                              |
|--------------------|--------------------|---------------------------|------------------------------------------|
| Sennheiser TCC 2   | 9 m × 9 m          | Dante / analog            | Tier 1 boardroom — premier ceiling array |
| Shure MXA920       | 8.5 m × 8.5 m      | Dante AES67               | Tier 1/2 — square boardroom              |
| Shure MXA710       | Linear, table-edge | Dante AES67               | Tier 2 — long table or boundary          |
| Biamp Parlé TCM-XA | 6 m × 6 m          | AVB / Dante via processor | Tier 2 — Biamp ecosystem                 |

MOUNTING DISCIPLINE

Ceiling-array microphones mount flush in 9-11 ft ceilings. Hard-lid mount preferred over T-bar for acoustic isolation. Cable enters from above through a backbox; no exposed cable in the room. PoE+ powering specified for TCC 2 and MXA series; verify switch supports IEEE 802.3at on each AI VLAN port.

GAIN STRUCTURE

Set unity gain at the microphone, with output level calibrated to -18 dBFS at normal speaking volume. Headroom for raised voices preserves Whisper accuracy. Auto-gain on the microphone DSP is disabled; gain is set once at commissioning and locked.

INSPECTION STANDARD

Commissioning verification

Speak from each seat at conversational volume. Verify peak meter reads -18 to -12 dBFS at the AI appliance audio interface. Capture a 60-second test recording and run Whisper Large-v3. Word error rate ≤ 8 percent passes; above 10 percent requires acoustic remediation before signoff.

**DO NOT**

**Without written direction**

Do not deploy lavalier or wireless handheld microphones for primary transcription capture. Do not chain ceiling arrays to consumer USB audio interfaces. Do not run Dante audio over the same VLAN as AI inference traffic.



§ 06

# The AI Appliance

PART II — AUDIO & NETWORK INFRASTRUCTURE

The AI appliance is the physical unit that runs the model server, transcription engine, and vector database. Vol 15 specifies appliance form factor, rack mounting, thermal envelope, and storage configuration. Pre-Rack AI Standard sidecar covers pre-arrival pathway, power, and cooling provisioning.

APPLIANCE PROFILES

| Tier | Form factor         | Rack                     | Power                     | Storage                            |
|------|---------------------|--------------------------|---------------------------|------------------------------------|
| 1    | Mac Studio M2 Ultra | 3U Sonnet xMac shelf     | 295 W max, 20A circuit    | 8 TB internal                      |
| 2    | Mac Mini M4 Pro     | 1U vented shelf or niche | 140 W max, 15A circuit    | 2 TB internal + 4 TB Thunderbolt   |
| 3    | Mac Mini M4 base    | Desk-mounted or VESA     | 75 W typical, 15A circuit | 512 GB internal + 2 TB Thunderbolt |

THERMAL ENVELOPE

Mac Studio M2 Ultra: 1,007 BTU/h maximum, 6 dBA idle, ambient under 95°F sustained. Mac Mini M4 Pro: 478 BTU/h maximum, inaudible at 1 m, ambient under 95°F sustained. Rack room thermostat setpoint 72°F. Above 80°F sustained, M-series Macs throttle single-core performance; transcription latency increases.

STORAGE LAYOUT

Internal SSD holds the operating system, model weights, and current-quarter transcripts. Thunderbolt enclosure (or Synology NAS on the AI VLAN) holds the long-term project knowledge base. Daily snapshot to encrypted external drive stored offline in the rack room — no cloud backup.

BEST PRACTICE

Rack the appliance

Even on Tier 3 home offices, prefer rack mount or VESA mount over desk placement. Stable thermal environment, secure cabling, and locked physical access are non-negotiable for an appliance that holds attorney-client privileged audio.

UPS REQUIREMENT

All tiers require online double-conversion UPS. Tier 1 specified APC Smart-UPS SUA1500RM2U or equivalent (1500 VA / 1000 W, 2U). Tier 2/3 specified APC Back-UPS Pro BR1500MS or equivalent. Runtime target 15 minutes — long enough to complete the active transcription, save state, and shut down cleanly.



§ 07

# Network and Privacy Stack

PART II — AUDIO & NETWORK INFRASTRUCTURE

The AI appliance lives on its own VLAN. No exceptions. Vol 15 specifies the VLAN schema, firewall posture, DNS telemetry blocking, and the internal TLS certificate that secures Open WebUI to client devices. The privacy posture is engineered, not promised.

VLAN SCHEMA

| VLAN   | Purpose                              | Default route                               | Inter-VLAN access                             |
|--------|--------------------------------------|---------------------------------------------|-----------------------------------------------|
| AI-LAN | AI appliance, ChromaDB, vector store | Block to internet                           | Receive from User VLAN only on TCP 3000, 8443 |
| AUDIO  | Dante or AES67 microphone bridge     | Block to internet                           | Audio bridge to AI appliance only             |
| USER   | Client laptops, tablets              | Standard internet egress                    | Initiate to AI-LAN on 3000, 8443 only         |
| MGMT   | Switch and AP management             | Block to internet, allow to monitoring host | Isolated                                      |

FIREWALL POSTURE

AI-LAN and AUDIO VLANs default-deny outbound. Explicit allowlist for NTP (one internal NTP server), package updater (paused except during owner-approved maintenance windows), and certificate revocation lists. No allow-all. Logged at the firewall and reviewed monthly.

DNS TELEMETRY BLOCKING

DNS for AI-LAN and AUDIO VLANs points to an internal resolver. Apple, Microsoft, Google, and OpenAI telemetry domains are blackholed. The AI appliance is signed in with a non-Apple-ID local account; Apple Intelligence is disabled at the OS level.

INTERNAL TLS

Open WebUI is served over HTTPS using an internal CA. The CA root is installed on every owner-approved client device. Self-signed certificates are not used. The owner sees a green lock; the system is not training the owner to ignore warnings.

#### INSPECTION STANDARD

### VLAN verification at signoff

From a User VLAN client, ping 8.8.8.8 — succeeds. From the AI appliance terminal, ping 8.8.8.8 — fails. From the AI appliance terminal, dig openai.com — returns 0.0.0.0 or NXDOMAIN. Capture screen recordings to the project knowledge base.

#### PRIMARY SOURCES

NIST AI RMF · GDPR Chapter 5 · HIPAA Security Rule



PART 3

---

# Workflow & Integration

Sections 08–10 — Field Workflows, Confidentiality Protocol, Cross-System Integration

## § 08

# Field Workflows

## PART III — WORKFLOW & INTEGRATION

The AI appliance supports four standard workflows in a residential conference room. Each is specified end-to-end: trigger, capture, processing, output, and consent treatment. Workflows that fall outside this list require written design direction before deployment.

### WORKFLOW 01 — MEETING TRANSCRIPT

Owner presses the Crestron room button labeled RECORD. Visible signage on the table and the in-room display indicates recording is active. Audio captures over the ceiling array to Whisper.cpp on the appliance. At meeting end, transcript is filed to the project knowledge base under the meeting date with a confidentiality flag. No upload, no email auto-send, no cloud sync.

### WORKFLOW 02 — POST-MEETING SUMMARY

After transcript completion, owner triggers SUMMARIZE. Llama 3.3 70B (Tier 1) or Llama 3.1 30B (Tier 2) generates an action-item list, decisions made, and open questions. Summary appears in the Open WebUI session window. Owner reviews and approves before any redistribution.

### WORKFLOW 03 — KNOWLEDGE BASE QUERY

Owner asks a natural-language question about prior meetings or documents. ChromaDB or Qdrant retrieves relevant excerpts via BGE-M3 or nomic-embed-text embeddings. Llama composes the answer with inline source citations to the original transcripts. All retrieval and generation runs on the appliance.

### WORKFLOW 04 — DOCUMENT DRAFTING

Owner dictates or types a draft request. Llama generates the draft. Owner edits in Open WebUI or exports to Pages or Word. Vol 14 verification protocol applies before any external distribution.

#### NOTE

#### Consent and signage

Per New Jersey one-party consent (N.J.S.A. 2A:156A-3), the owner can record their own conversations. For California two-party consent (Penal Code § 632), all participants must consent. Restrepo posture: visible table-top signage on every recording session, verbal consent at session start, regardless of jurisdiction.





## § 09

# Confidentiality Protocol

## PART III — WORKFLOW & INTEGRATION

A residential conference room AI appliance routinely processes attorney-client privileged communications, family-office investment decisions, healthcare-adjacent discussions, and SEC-regulated advisory content. Vol 15 specifies the four-pillar confidentiality protocol: storage, access, retention, and provenance.

### STORAGE

Audio recordings, transcripts, and embeddings are stored on the AI appliance and the appliance-attached encrypted drive only. FileVault enabled on macOS. APFS volume encryption keys held only by the owner. No cloud sync, no Time Machine to network shares, no iCloud Drive.

### ACCESS

Open WebUI requires per-user authentication. Restrepo recommends the appliance run in single-user mode for the owner principal, with delegated read-only access via scoped API tokens for assistants or family-office staff. Authentication events are logged on-appliance.

### RETENTION

Default retention: transcripts 24 months, audio 90 days, embeddings indefinite. Owner can override per-meeting at recording start. Quarterly retention review with the owner. Hard-delete via APFS secure erase, not file move.

### PROVENANCE

Every AI-generated output is tagged with the model name, model version, source transcripts, and generation timestamp. C2PA content credentials applied to AI-generated images per Vol 14 § 08. Summaries shared externally include a machine-readable provenance footer.

#### SPECIFICATION

##### Privilege posture

Restrepo will execute a confidentiality addendum with the owner specifying that the integrator does not access the AI appliance content. Service access is limited to system administration on a scheduled basis, never content review.

### PRIMARY SOURCES



N.J.S.A. 2A:156A-3 · Cal. Penal Code § 632 · C2PA v2.4



## § 10

# Cross-System Integration

## PART III — WORKFLOW & INTEGRATION

The AI appliance does not stand alone. It integrates with the room's Crestron control system, Lutron lighting, and where present, Control4 distributed audio. Vol 15 specifies integration patterns that respect the privacy posture: control only flows into the appliance, not out.

### CRESTRON

Crestron 4-Series processor sends room-state events (record start, record stop, summarize, silence room) to the AI appliance via SIMPL# Pro REST module to the Ollama API on localhost (when the processor is racked with the appliance) or over the AI VLAN (when the processor is in a separate rack). The appliance never publishes back to the Crestron system. Status visible only in Open WebUI.

### LUTRON

Lutron HomeWorks integration is single-direction: when RECORD is active, the Crestron system commands a Lutron preset for recording-friendly lighting (warm, even, no flicker). The Lutron processor does not query the AI appliance. The AI appliance does not write to the Lutron API.

### CONTROL4

Where Control4 is the primary platform, the C4 Driver SDK provides a custom driver that calls the Ollama API for room-level summaries displayed on the C4 touchpanel. Driver runs in single-direction mode: read prompt result, display, do not log or transmit beyond the room.

#### NOTE

#### Cross-volume reference

Vol 12 (HT Installation) governs rack mechanical and acoustic standards. Vol 14 (AI in the Field) governs operator-facing AI discipline including verification and disclosure. Vol 15 governs the conference-room appliance specifically. The Pre-Rack AI Standard sidecar governs pre-arrival pathway and power.

**DO NOT**

**Without written direction**

Do not connect the AI appliance to the residential ISP-router DMZ. Do not expose the Ollama API to the WAN. Do not enable remote SSH from outside the property. Do not integrate with Siri, Alexa, or Google Assistant.



PART 4

---

# Discipline & Closeout

Sections 11–13 — Without Written Direction, Pre-Rack Coordination, QC · Audit · Closeout

## § 11

# Without Written Direction

## PART IV — DISCIPLINE & CLOSEOUT

These items are not prohibited. They are not deployed without written direction from the owner or owner's representative, recorded in the project knowledge base with date, scope, and acceptance signature. The list is residential, conference-room specific, and may extend with future revisions.

### DO NOT

#### Cloud AI services

Otter.ai, Zoom AI Companion, Microsoft 365 Copilot, Google Duet AI for the conference room. The Restrepo default is local-only inference; cloud inference deviates from the Vol 15 privacy posture.

### DO NOT

#### Always-on transcription

Continuous open-microphone capture without an explicit RECORD trigger. Owner may opt in for specific use cases with documented retention policy.

### DO NOT

#### Smart-speaker assistants

Amazon Echo, Google Nest, Apple HomePod with Siri active. These devices stream wake-word audio to vendor cloud and are not paired with the conference-room AI appliance.

### DO NOT

#### Auto-email summaries

Automatic transmission of meeting summaries to email recipients without owner review. Owner reviews every summary before any external distribution.

### DO NOT

#### WAN-exposed Ollama API

Port-forwarding the Ollama API to the public internet for remote access. Use a Tailscale or WireGuard mesh on the User VLAN with explicit owner-approved devices.



**DO NOT**

**Mixed-use VLAN**

AI appliance sharing a VLAN with consumer IoT, Alexa devices, or guest Wi-Fi. AI appliance VLAN is dedicated.

**DO NOT**

**Internet-facing OpenWebUI**

Open WebUI exposed to the internet via reverse proxy without owner-approved authentication and certificate posture.



## § 12

# Pre-Rack Coordination

## PART IV — DISCIPLINE & CLOSEOUT

The AI appliance does not arrive on a Restrepo project as a surprise. The Pre-Rack AI Standard sidecar (companion to this volume) covers nine pre-arrival categories: conduit and pathway, power and circuit, cooling and acoustics, network and VLAN, rack and mount, audio capture rough-in, control bridge prep, without-written-direction, and the pre-arrival checklist.

### PRE-ARRIVAL CHECKLIST

#### INSPECTION STANDARD

##### Conduit

1" EMT or 3/4" smurf tube from rack location to network closet. One Cat6A pulled, one spare smurf left in place. Both ends labeled RED 'AI-VLAN'.

#### INSPECTION STANDARD

##### Power

Dedicated 20A circuit to rack location for Tier 1 (15A acceptable for Tier 2/3). Isolated ground recommended. Hospital-grade NEMA 5-20R receptacle. Circuit breaker labeled 'AI APPLIANCE'.

#### INSPECTION STANDARD

##### Cooling

Rack-room ambient 65-75°F (18-24°C). For Tier 1 in a closet, passive vent grille high and low or active 80mm fan exhaust controlled by thermostat at 78°F setpoint.

#### INSPECTION STANDARD

##### Network

AI VLAN provisioned on the network switch. Access port to rack location, configured with DHCP reservation for the appliance MAC. Audio VLAN provisioned to ceiling-array microphone location.

**INSPECTION STANDARD****Rack**

Tier 1: 3U Sonnet xMac shelf. Tier 2: 1U vented shelf. Cable management: rear hook-and-loop bundles, no zip ties on power; 5 percent spare in patch panel.

**INSPECTION STANDARD****Audio**

Ceiling-array microphone backbox installed. PoE+ verified at switch port. Cat6A homerun from microphone to network closet. Acoustic treatment installed before commissioning.

**INSPECTION STANDARD****Control**

Crestron 4-Series Ollama bridge module loaded; SIMPL# Pro project compiled and tested with stub responses on the bench before site delivery.

**INSPECTION STANDARD****UPS**

APC Smart-UPS SUA1500RM2U (Tier 1) or BR1500MS (Tier 2/3) racked, charged, and burn-in tested for 24 hours before delivery.

**INSPECTION STANDARD****Documentation**

Project knowledge base seeded with site address, owner principals, rack location photographs, network schema, and acoustic measurements.

**INSPECTION STANDARD****QC gate**

All ten Pre-Rack AI Standard QC items signed by lead operator before AI appliance is shipped to site.



§ 13

# QC · Audit · Closeout

PART IV — DISCIPLINE & CLOSEOUT

Closeout is the moment Restrepo is accountable. The QC gate is the operator's signature that the room hears, the room understands, and nothing leaves the room. Twelve items must pass before the system is turned over to the owner.

QC GATE — TWELVE ITEMS

- ☐ **VERIFY** Acoustic measurement:  $RT60 \leq 0.6s$ ,  $NC \leq 25$ ,  $STI \geq 0.65$  documented.
- ☐ **VERIFY** Microphone gain locked at -18 dBFS unity; auto-gain disabled; gain calibration recording archived.
- ☐ **VERIFY** Whisper Large-v3 (Tier 1) or Medium (Tier 2/3) WER  $\leq 8$  percent on signoff transcript.
- ☐ **VERIFY** VLAN isolation verified: AI appliance ping to 8.8.8.8 fails; client ping succeeds.
- ☐ **VERIFY** DNS telemetry blocked: openai.com, anthropic.com, mistral.ai, telemetry.apple.com all return NXDOMAIN.
- ☐ **VERIFY** Internal CA root installed on owner's primary devices; Open WebUI green lock verified.
- ☐ **VERIFY** Crestron RECORD button triggers transcription; in-room signage and display indicator illuminate.
- ☐ **VERIFY** Lutron preset for recording-friendly lighting recalls correctly when RECORD active.
- ☐ **VERIFY** Knowledge base ingestion of three test documents indexed and queryable.
- ☐ **VERIFY** Snapshot to encrypted external drive completed; owner has the recovery key.
- ☐ **VERIFY** UPS runtime tested at full appliance load: 15 minutes minimum.
- ☐ **VERIFY** Owner walkthrough documented with all four standard workflows demonstrated.

BEST PRACTICE

Manifesto closeout

The room hears. The room understands. Nothing leaves the room. Capture, control, confidential — three words on the cover are the contract on closeout.

CITATIONS

NEC 2023 (NFPA 70) — National Electrical Code · Apple Mac Studio Specs · Apple Mac Mini Specs · Sennheiser TCC 2 · Shure MXA920 · Shure MXA710 · Biamp Parlé TCM-XA · Ollama · Open WebUI · Whisper.cpp · Piper TTS · llama.cpp · ChromaDB · Qdrant · APC Smart-UPS SUA1500RM2U · GDPR Chapter 5 · HIPAA Security Rule · NIST AI RMF · N.J.S.A. 2A:156A-3 · Cal. Penal Code § 632 · C2PA v2.4 · Apple PCC



# Captured. Controlled. Confidential.

Restrepo Innovations · Island Media Systems

