

# Operationalising Municipal Sovereign AI

Decoupled Architectures, Local Inference Stacks, and Public Sector Algorithmic Governance

## THE CIVIC UTILITY PRINCIPLE IN MUNICIPAL AI:

Artificial intelligence deployed in municipal governance is not an outsourced commercial subscription; it is an automated extension of statutory public authority. Relying on foreign proprietary API endpoints exposes citizen records, spatial telemetry, and municipal decisions to extraterritorial warrants (US CLOUD Act), unpredictable API deprecations, and unverified algorithmic bias. A sovereign municipal AI stack enforces complete local custody over hardware compute, open-weights foundation models, vector databases, and civic interfaces.

## THE FOUR DECOUPLED ARCHITECTURAL LAYERS OF SOVEREIGN MUNICIPAL AI

|                                                                                                                                                                                            |                                                                                                                                                                                      |                                                                                                                                                                |                                                                                                                                                                                    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Layer 1: Compute Infrastructure &amp; Edge</b><br><br>Domestic sovereign cloud zones (SecNumCloud, EU NIS2-aligned)<br><br>Edge nodes process urban transit & utility telemetry locally | <b>Layer 2: Inference Open-Weights Engine</b><br><br>Containerised foundation models (Mistral, LLaMA, Gemma)<br><br>Zero data exfiltration; citizen prompts never leave jurisdiction | <b>Layer 3: Retrieval Local Context (RAG)</b><br><br>Municipal bylaws, welfare statutes & case law<br><br>Role-based access control guarantees confidentiality | <b>Layer 4: Gateway Civic Service Interface</b><br><br>Web and mobile service portals providing explainable AI<br><br>Immutable audit logs feeding public algorithmic transparency |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

## COMPARATIVE MUNICIPAL AI DEPLOYMENT MODELS

| Deployment Model                                                        | Data Residency & Subpoena Exposure                                                                        | Technical Autonomy & Portability                                                                | Cost Profile & Licensing Risks                                                                          |
|-------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| <b>Public Hyperscale API</b><br>(Commercial Cloud Endpoints)            | Severe exposure: Prompts and telemetry processed in foreign jurisdictions (US CLOUD Act, GDPR Article 45) | Zero portability: Proprietary vendor lock-in; no access to model weights or inference logs      | High operational costs; unpredictable pricing; vendor lock-in; potential for biased or censored outputs |
| <b>Sovereign Cloud Zone</b><br>(Domestic European/National IaaS)        | Protected: Hardware housed in domestic facilities; data residency enforced                                | High portability: Open standards (OCI, OPA) and open-weight models; audit logs accessible       | Medium costs; standardized licensing; no vendor lock-in; transparent pricing                            |
| <b>Inter-Municipal IT Consortium</b><br>(Shared Regional Hardware Pool) | Maximum sovereignty: Fully air-gapped regional data centers; no cross-border data flow                    | Complete autonomy: Customised open-weight models; full control over hardware and software stack | Lowest costs; shared infrastructure; no vendor lock-in; transparent pricing                             |
| <b>Localized Edge Deployment</b><br>(Traffic & Public Transit Nodes)    | Zero transit risk: Video streams, sensor feeds, and metadata processed locally; no network egress         | High autonomy: Standardised open-weight models; full control over hardware and software stack   | Lowest costs; no network egress; no vendor lock-in; transparent pricing                                 |

# A Phased Transition Framework for Local Councils and Metropolitan Authorities

## THE 4-STAGE SOVEREIGN AI ADOPTION ROADMAP

### 1 Algorithmic Inventory & Exposure Audit (Months 1–3)

Identify all active algorithmic models and external API dependencies across council departments. Classify systems under EU AI Act risk tiers (Minimal, Specific Transparency, High Risk). Map data transit pathways to identify cross-border exfiltration.

### 2 Open-Weights Pilot & Sovereign Vector Store (Months 4–6)

Deploy a containerised open-weights foundation model (e.g. Mistral 7B/8x7B) on dedicated domestic infrastructure. Connect to an internal Retrieval-Augmented Generation (RAG) store indexing municipal bylaws and zoning documents.

### 3 Regional Public IT Compute Consortium (Months 7–9)

Establish public-public partnerships (PuPs) with neighbouring local councils (replicating Germany's Dataport or France's GIP RECIA models) to co-finance dedicated GPU compute nodes and negotiate sovereign hosting agreements.

### 4 Public Algorithmic Register & Continuous Audit (Months 10–12)

Establish an open municipal AI transparency registry (following Amsterdam and Helsinki standards) documenting model parameters, training data lineage, human oversight mechanisms, and annual algorithmic bias audit results.

## MUNICIPAL AI PROCUREMENT & TENDER GOVERNANCE CHECKLIST

#### ✓ Zero Data Ingestion & Model Training Prohibition

Mandate strict contractual clauses confirming that citizen queries, municipal inputs, and uploaded documents are never retained or utilized to train or fine-tune commercial vendor foundation models.

#### ✓ Algorithmic Explainability & Model Lineage Cards

Vendors must provide comprehensive documentation detailing model architectures, training dataset compositions, known bias limitations, and deterministic decision parameters before deployment.

#### ✓ Public Algorithmic Register Compliance

Every AI system impacting citizen services or welfare eligibility must be registered in the public municipal algorithmic register, ensuring full democratic transparency and public scrutiny.

#### ✓ Local Inference & Domestic Jurisdiction Mandate

Require model execution, inference weights, and embeddings to reside strictly within domestic or sovereign cloud environments, with explicit legal immunity from foreign extraterritorial warrants.

#### ✓ Vendor Exit & Portability Guarantees

Contracts must ensure full portability of fine-tuned weights, RAG vector embeddings, and prompt libraries in standard open formats, allowing migration to alternative runtimes within 30 days without penalty.

#### ✓ Continuous Bias Auditing & Human Oversight

Require annual third-party fairness audits and mandate that high-impact municipal decisions (such as social housing allocation or benefits triage) maintain binding human-in-the-loop oversight.

## ECOSYSTEM COLLABORATION & SUPPORT NETWORKS FOR MUNICIPAL AI

### Inter-Municipal IT Cooperatives (Dataport / RECIA)

Join existing regional public sector IT consortia to access shared high-performance computing clusters, eliminating the high individual cost of sovereign GPU infrastructure.

### Municipal Algorithmic Registries (Helsinki / Amsterdam)

Leverage established open-source schema templates for civic AI registers to satisfy EU AI Act transparency obligations and build enduring public trust.

### Open-Source AI Foundations & OpenUK

Collaborate with the Linux Foundation, OpenUK, and Open Source Initiative to adopt verified open-weights models and standardized public sector model deployment blueprints.

### Academic & National Research Institutes

Partner with local universities and public supercomputing centres for independent model auditing, ethical impact assessments, and local language fine-tuning.