Picca – Tech Architecture (Public Extract v0.1)

Purpose This document gives Codex-1 (and human collaborators) the minimum, non-confidential context needed to generate or navigate code inside the public repository. Proprietary optimisations, dataset IDs, and customer details are intentionally omitted.

High-Level Data Flow

```
flowchart TD
  subgraph Client
    A[Web (Next.js)] -->|HTTPS| B(API Gateway)
  end
  subgraph Cloud Run
    B --> C[Go Gateway]
    C --> D[Python AI Service]
    D --> E[PostgreSQL]
    D --> F[GCS Artifacts]
  end
  %% Future branch (dashed)
  C -.-> G((Depth-Pilot RealSense))
```

- Dashed edges represent planned extensions and are **out-of-scope for the 14-day MVP**.

Service Boundaries

Service	Runtime	Endpoint prefix	Key Responsibility
go-gw	Go 1.22 / Fiber	`/api/v1/*`	AuthN/Z, rate-limit, request fan-out
ai-core	Python 3.11 / FastAPI	`/core/*`	Pose extraction, DCV scoring
jobs	Cloud Run Jobs	*internal*	Batch SHAP audits, nightly data purges

---

Stack Summary (MVP)

- **Infra as Code** — Terraform 1.8
- Google Cloud Run (services above)
- Cloud SQL (PostgreSQL 15) – single region `asia-northeast1`
- Cloud Storage buckets: `picca-media-tmp`, `picca-artifacts`
- Pub/Sub topics: `inference-queue`, `audit-events`

- Secret Manager with 30-day rotation window
- **CI/CD** — Cloud Build trigger on `main`
- Unit ▶️Integration ▶️Deploy to **preview**
- **Observability** — Cloud Trace, Prometheus exporters auto-scraped via Managed Metrics

---

Security & Privacy Hooks

1. **24 h media purge** — Object lifecycle rule enforces TTL on RGB uploads.
2. **Face-less inference** — Only MediaPipe key-points kept; frames discarded.
3. **Explain-or-Reject** — Every 400/403 includes JSON body with rule ID.
4. **OAuth2** — Bearer tokens (30 min TTL) via Google Identity Platform.

Versioning Roadmap

Stage	Target	New Capabilities
`v1-alpha`	Day 0	Single-region, manual secret bootstrap
`v1-beta`	+14 d	Multi-region failover, KMS envelope encryption
`v1-rc`	+30 d	Depth-Pilot edge branch behind feature flag

Last updated 2025-07-08