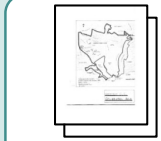
1. Reader (Agent)

VLM → PDFInfo · reads raw PDF (no OCR)



PDFInfo *(structured JSON)*

- Site · postcodes · grid ref. · scale
- Map page indices
- Road · place · parish names
- Is district wide? → district name

PDFInfo

3. Critic (Optional)

Independent LLM · Post-commit review of top-3 considered candidates

map + mask | tile + polygon

Metrics

- Count of inliers
- Road name agreement
- Scale consistency

Approve · **Switch** · **Retry locate** *(≤2×)*

After commit, Top-3 considered candidates

Switch / Retry locate

2. Worker (Agent)

Re-call with feedback

PDFInfo, map page

Lat., Lon., σ

2.1 Locate sub-agent (Agent)

Tool: OS Open Names · 2–4 queries → approximate site coordinates

Output:

- Latitude (Lat.) , Longitude (Lon.)
- Positional-uncertainty radius (σ)
- Confidence (high / med / low)

5x

Lat, Lon, σ, map page

Refined candidate, quality signals

2. 2 Map Match to Real World

MINIMA / LoFTR ↔ **OS tiles**

Refine approximate position to exact location

SAM3 + LoRA → boundary mask

RANSAC affine → GeoJSON

→ Refined candidate + quality signals (Count of inliers, Road name agreement, Scale consistency)

Best candidate

2. 3 Commit candidate

→ **Final GeoJSON**

2.4 Lookup district

For district wide docs, bypasses 2.1-2.3