

1 - The investigator loop (the agent loop, pointed at footage)

Decompose -> Plan -> Call a tool -> Observe -> Consult memory (repeat until answered)
Then: Write the finding (clip + timestamp) -> Human reviewer decides (never the agent)
Planning + tool use + memory: the three primitives fire on every turn of the loop.

2 - The six-tool belt (cheap & wide -> costly & precise)

Detector (YOLO)	what & where?	very low	scans whole nights
Tracker (ByteTrack)	where did it go next?	low	
Segmenter (SAM 2)	same exact object?	low-med	
Search index (RAG)	which clips match?	low/query	built once
VLM reader (Qwen-VL)	what's happening?	HIGH	a few clips only
Report writer (LLM)	summarize + evidence	low	

3 - Forensic search in two phases (never read every frame)

INGEST (cheap, continuous): detector notes + sampled keyframes -> CLIP embeddings
+ transcript + metadata -> searchable index
QUERY (only when asked): question -> embed -> retrieve short list -> rerank
-> ~10 clips -> VLM reads & answers
Math: 16 cams x 15 fps x 12 h ~ 10.4M frames/night. VLM-every-frame ~ \$5,000/night;
re-read the retrieved clips ~ 10 clips per hour ~ \$500/hour

4 - EU AI Act lanes (design the agent to clear the lanes by default)

PROHIBITED	Real-time face ID in public space (Article 5, in force since Feb 2025).
HIGH-RISK	Identifying WHO a person is, post-event. Heavy duties; applies 2 Dec 2027.
ANALYTICS	Searching objects / events / descriptions - NOT identities. Safe default.
OVERSIGHT	Human in the loop on every consequential action. Agent proposes; person decides.

5 - Scoping checklist (ask these before you build, buy, or wrap)

- ☐ Does the agent retrieve-then-read, or read every frame? It must retrieve, then read a few clips.
- ☐ Is the tool set small and sharp (~6 tools), not a kitchen drawer of thirty?
- ☐ Is there episodic memory, so the agent stops re-flagging the same shadow every night?
- ☐ Is scope kept to objects and events, not identity, by design rather than by configuration?
- ☐ Does a human approve every alert, dispatch, or lockdown? The agent never acts alone.
- ☐ Build, buy, or wrap a VMS? Start by buying/wrapping; build only what market tools get wrong.