SI-Watcher:

a Real-Time Video-to-Knowledge Engine on the Edge

NOTE: most of the images in this presentation have been generated with Microsoft's Image Creator

Maurizio Martignano
Spazio IT – Soluzioni Informatiche s.a.s
Via Manzoni 40
46051 San Giorgio Bigarello, Mantova
https://spazioit.com

January 2026

Vision: Transforming Video into Actionable Intelligence

- Real-time understanding of live or recorded video streams
- Context-aware visual analysis using Generative AI
- Structured knowledge extraction — locally, securely, efficiently

What is SI-Watcher?

- An integrated pipeline that:
  - **Ingests** live **video** or recorded streams
  - **Samples frames** at configurable intervals
  - **Uses multimodal Generative AI** for visual understanding
  - **Extracts structured insights** in real time on edge devices

Core Capabilities

- **Video Ingestion** — files, live streams, webcams
- **Frame Sampling** — time-based intervals
- **Contextual Understanding** — Vision-capable GenAI
- **Structured Extraction** — events, objects, situations
- **Output / Integration** — logs, APIs, downstream systems

Key Innovations

- 🎥 Multimodal GenAI applied directly to video frames
- ⚡Non-blocking pipeline with frame dropping for responsiveness
- 🖼️ Edge-optimized JPEG encoding and resizing
- 🔁 Seamless live-stream reconnection handling
- 🧰 CLI-first, automation-friendly design

Why Edge Deployment?

- 🚫 No dependency on cloud
- 🔐 Privacy-first: sensitive data stays local
- ⚡ Ultra-low latency for real-time decisions
- 🌐 Works in offline or low-connectivity environments

Potential Domains of Application

- Industrial monitoring & inspections
- Smart surveillance and safety systems
- Manufacturing quality control
- Research and laboratory observation
- Remote site supervision

SI-Watcher in Action (Use Case)
Industrial Monitoring Scenario:

- Camera observes an industrial process
- SI-Watcher samples a frame every N seconds
- GenAI interprets visual context (e.g. anomaly detected)
- Structured output logged or forwarded via API
- Human operator remains in the loop

Comparison to Market

🧩 Capability	🏢 Typical Market Solutions	🚀 SI-Watcher
🎥 Real-time edge video analysis	Partial or cloud-assisted	**Yes — fully edge-based**
☁️ Cloud dependency	High (cloud-first or hybrid)	**None**
🧠 Multimodal GenAI	Add-on or external	**Native**
🔓 Open-source core	Rare	**Yes**
⚙️ Configurable triggers	Limited or fixed	**Yes**

Competitive Landscape & Differentiation

Aspect	Market Landscape	What Makes SI-Watcher Different
Domain focus	Broad Edge Video Analytics	Semantic-driven Edge Video Analytics
Typical platforms	NVIDIA Metropolis / DeepStream	Lightweight, purpose-built engine
Cloud-first solutions	Azure AI Video Indexer, SaaS tools	Fully cloud-independent
Vision approach	Object detection–centric	Multimodal GenAI reasoning
Processing model	Continuous heavy pipelines	Periodic frame sampling
Architecture	Resource-intensive, monolithic	Edge-first, efficient
Interface	GUI-centric platforms	CLI-driven, automation-friendly
Extensibility	Closed or limited	Open-source and easily extensible

Technical Highlights

- OpenCV-based capture and rendering
- Multithreaded pipeline (capture, inference, UI)
- JPEG compression with adaptive resizing
- OpenAI-compatible vision API integration
- Production-grade error handling

Contacts

- GitHub: github.com/mmartign/Video-to-Knowledge

- **Contact**: info@spazioit.com
  Website: https://spazioit.com