



LLMs and VLMs for Regulatory Compliance, Quality Control and Safety Applications

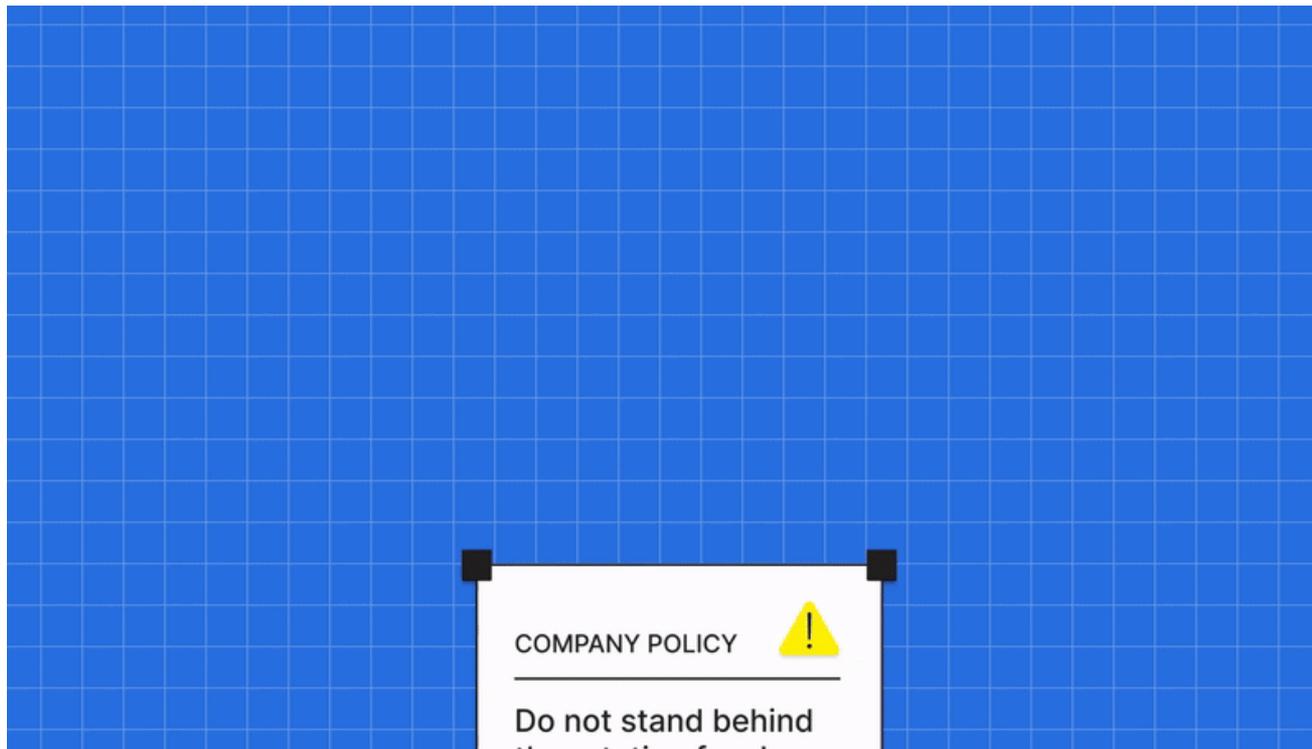
Lazar Trifunovic
Solutions Architect
Camio



LLMs & VLMs: A Step Towards Artificial General Intelligence

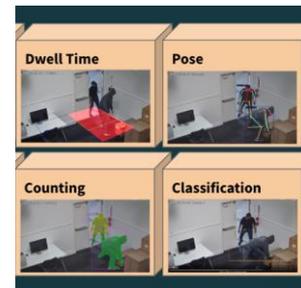
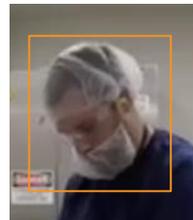
- LLMs are expanding and can process multiple types of data
 - Text, Images, Audio, Video, Sensor Data
 - LLM can be given tools (web search, API call→computer vision models) → VLM
- With enough context, VLMs can become better than humans
 - Similar detection and error rate (if not better)
 - Never needs to take a break
 - Can be quickly repurposed
- Make AI easy to use, deploy, and optimize

Text: An Easy Way to Deploy Complex Scenarios



Before: Individual Models with Manual Heuristics

- Policy Compliance Goals
 - PPE compliance
 - At least 2 employees present
 - Vials picked up with ONLY tongs
 - Vials set down with either tongs or hands
- Historical Solution
 - Manually deploy models
 - Define logic/heuristics to solve objectives
 - Program triggers and alert responses



After: Streamlined Deployment and Reconfiguration

Assessment

There is only one staff member present. Whenever the conveyor is moving, there must be at least two staff members present. Please call for backup.

The staff member is wearing latex gloves, a white hairnet, and a face covering, so this policy is compliant.

The staff member is selecting a vial using only metal tongs, so this policy is compliant.

Recommended Message

There is only one staff member present. Whenever the conveyor is moving, there must be at least two staff members present. Please call for backup.



Use Case: People and Process

Relevant Text

Every forklift operator must be wearing a white hard hat.

Every forklift must have its lights turned on when moving.

Cranberry crates must not be stacked more than 2 crates high.

Recommended Message

Assessment

The forklift operator is wearing a white hard hat.

The forklift is moving and its lights are on.

The crates are stacked 2 crates high.



Use Case: People and Process

Relevant Text

The floor must be mopped in an S-shape motion.

The person mopping the floor must never step onto the area that was just mopped.

Every person in the room must be wearing a white cleanroom suit.

Assessment

The person is mopping the floor in an S-shape motion.

The person is not stepping onto the area that was just mopped.

The person is wearing a white cleanroom suit with pink accents.

Recommended Message



Capable of Ranking Deviations

Assessment

LOW deviation - It is likely that the weapons are part of the job of the people shown in the video. They appear to be police officers. The weapons are not being used in a threatening manner.

Relevant Text

No weapons are allowed. Firearms, knives, and any other dangerous weapons are prohibited.



Leveraging VLMs to Triage Events Automatically

- Actions can be configured and associated with plain text language
- The VLM can select the action based on text association
- Rankings can be taken into consideration

The screenshot shows the Camio Actions configuration page on the left and a security event notification on the right.

Actions Configuration:

- Name:** Discord Notification
- Description:** Use this to notify security of significant policy violation
- Method:** POST
- URL:** `https://discord.com/api/webhooks/124651347389526`
- Body Parameters:**

```
JSON Template
{
  "title": "{{title}}",
  "url": "{{image_url}}",
  "description": "{{description}} [View it now]({{content_url}})",
  "color": 15258703,
  "image": {
    "url": "{{image_url}}"
  }
}
```

Security Event Notification:

- Channel:** # security
- Sender:** camio [APP] 4/17/25, 9:01PM
- Event:** 1. Aisles are obstructed due to the... at Search 2025-04-18T04:00:57.680556+00:00
- Content:**
 - 1. Aisles are obstructed due to the...
 - 1. Aisles are obstructed due to the presence of stocking carts or wire-caged items and boxes. This is a deviation from the policy.
 - 2. The shelves appear to be stocked with products. Unable to verify facing.
 - 3. Cardboard is present on the floor, suggesting that trash and cardboard are not properly managed. This is a deviation from the policy.
- Image:** A photograph of a store aisle with a sign "16" and a yellow wall in the background.
- Footer:** Recommended message: Please clear the aisles of any obstructions such as stocking carts, fallen items, or overstock. Ensure shelves are stocked and items are faced. Please manage trash and cardboard by removing the excess cardboard from the aisles. The image shows the interior of a Dollar General store. Aisles are visible, and some appear to have stocking carts or wire-caged items in them. Shelves are stocked with products. There's a sign with "16" on it. Some boxes are on the floor. A yellow wall is visible in the background. [View it now](#)

That's All Good, But What's the Bad and the Ugly?

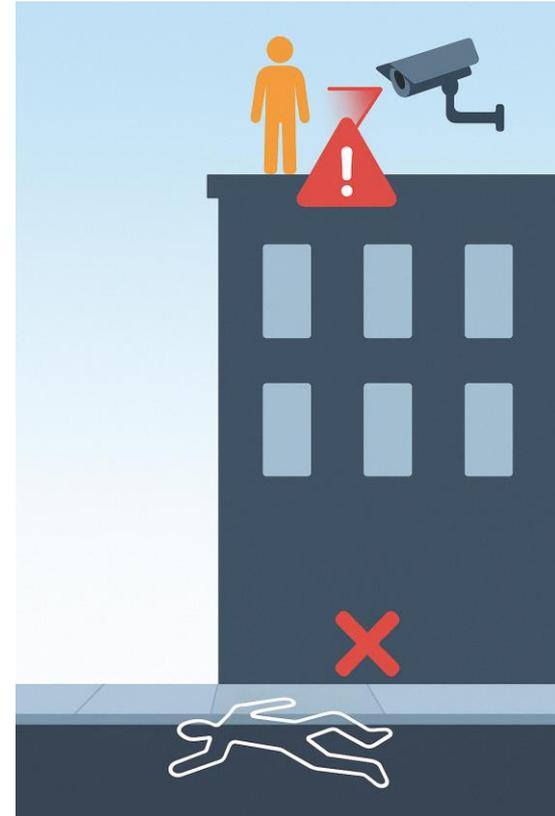
Early Stages of AI: Latency, Cost, and Quality

- AI projects are under the same three constraints
 - Time: Latency of the system
 - Cost: CapEx vs OpEx
 - Quality: What's acceptable?
- You get to pick 2 of 3 to optimize
- Building around restrictions
 - Latency: Queue management, scaling architecture, increased response window
 - Cost: Edge vs cloud and filtering
 - Quality: What's the actual level of complexity needed?



Making AI Proactive: “I Don’t Want Chalk Outlines”

- Reactive AI is troublesome
 - Latency is currently an issue for LMMs
 - Lower accuracy, lower cost solutions dominate the space
- Proactive AI provides value and ROI
 - Eliminating accidents before they occur by targeting bad practices
 - Detecting QA/QC issues before products leave the facility
 - Can require addressing subtleties via further training or tool sets



Extrapolating with Vectors: From Horses to Zebras

- VLMs can extrapolate using contextual clues
 - Striped animal
 - CV for animals returns horse
 - LLM concludes striped horse ⇒ zebra
- Vector mapping
 - Takes qualities and assigns weights to determine similarity
 - “Strange Shirt” → Skull design
- What about unique objects?
 - Unique local vehicles
 - Uniforms and PPE
 - Proprietary or unique parts of an assembly

```
{
  "query": {
    "text": "man in a strange black shirt"
  },
  "results": [
    {
      "events": "[\n {\n   \"camera\": \"Chaplin Ave\", \n   \"timestamp\": \"2024-05-05T12:35:41.135-0700\", \n   \"text\": \"A man wearing a black t-shirt with a skull design and khaki pants is standing in a driveway.\", \n   \"probability\": 0.92\n }, \n {\n   \"camera\": \"UVA SALA CAMBIO\", \n   \"timestamp\": \"2024-05-15T09:56:37.298-0700\", \n   \"text\": \"A man in an orange jumpsuit is visible\", \n   \"probability\": 0.85\n } \n]"
    }
  ],
  "usage": {
    "total_tokens": 1491
  }
}
```



horse → zebra

User Guided Model Creation: Teaching VLMs New Tricks



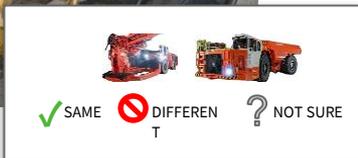
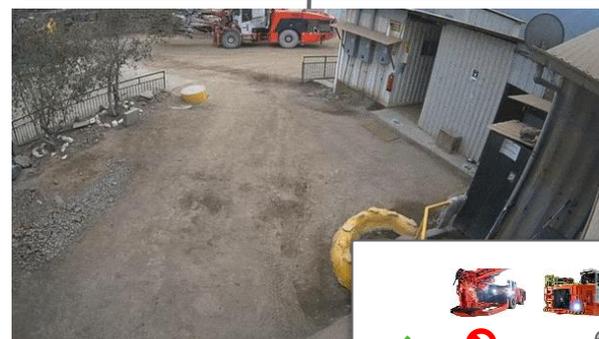
User Guided Model Creation: Optimization and Tuning

300+ positive **and** negative training samples extracted automatically from 1 click



Auto model training with reliable quality metrics published with each revision of the AI models

sandvik drill rig 5pm to 6am



feedback via usage



Auto published AI models optimized for edge but allowing for overflow to cloud, can be used for filtering

- LLMs and VLMs are powerful tools for leveraging video, audio, and data in our environment
 - These types of solutions have limitations, but can be applied to a wide variety of use cases
- LLMs and VLMs make AI easy to deploy and reconfigure
 - Text based policies are easy to deploy and update
 - Since the model is wired for multiple inputs, additional data can be provided for context (CAD drawing, access control information, etc)
- VLMs have shortcomings that can be engineered around
 - Filtering, vector embedding, and custom AI models

Camio's Visual Agents

Our Home Page

<https://camio.com/home>

Video Overview

<https://www.youtube.com/watch?v=JZeN3m3nnfA>

Industry Case Studies

<https://camio.com/resources/case-studies>

Additional VLM Learning Materials

Overview of Vision Language Models

<https://huggingface.co/blog/vlms>

Introduction to Vision-Language Modeling

<https://arxiv.org/html/2405.17247v1>