



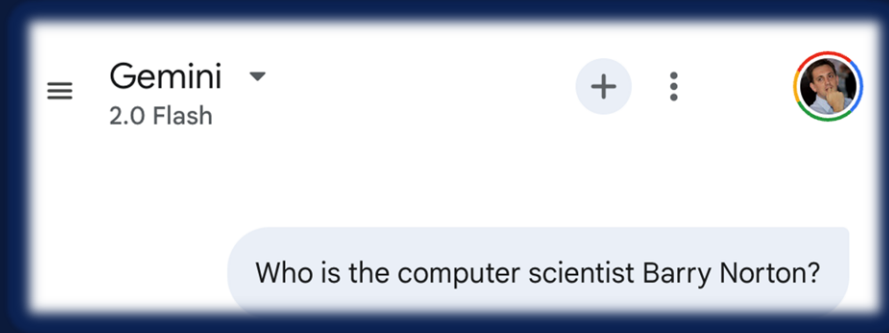
Learning from You:

The Promise of Generative AI in the Physical Security Industry

Dr. Barry Norton, Milestone Fellow



Large Language Models



Large Language Models

≡ Gemini ▾
2.0 Flash



Who is the computer scientist Barry Norton?

Barry Norton is a computer scientist specializing in artificial intelligence and computer vision. He has worked in research positions at various universities and technology companies. In 2018, he joined Milestone Systems, a video management software company, as Director of Research, and was promoted to Vice President of Research in 2021.

2,
7,
99,
1,
...

3,
18,
9,
7,
...

Large Language Models

The screenshot shows the Gemini 2.0 Flash chat interface. At the top, the Gemini logo and version '2.0 Flash' are visible. A user profile picture is shown on the right. The chat input area contains the question: 'Who is the computer scientist Barry Norton?'. The response area displays a paragraph of text about Barry Norton, followed by two numbered sources. Source 1 is 'Barry Norton - Google Scholar' with the URL 'scholar.google.com'. Source 2 is 'Milestone Systems promotes Dr. Barry Norton to Vice President of Research' with the URL 'git-sicherheit.de' and a GIT logo. The interface includes a hamburger menu, a plus icon, and a vertical ellipsis icon.

Gemini 2.0 Flash

Who is the computer scientist Barry Norton?

Barry Norton is a computer scientist specializing in artificial intelligence and computer vision. He has worked in research positions at various universities and technology companies. In 2018, he joined Milestone Systems, a video management software company, as Director of Research, and was promoted to Vice President of Research in 2021.

1. Barry Norton - Google Scholar
scholar.google.com

2. Milestone Systems promotes Dr. Barry Norton to Vice President of Research
GIT git-sicherheit.de

2,
7,
99,
1,
...

3,
18,
9,
7,
...

5,
11,
12,
10,
...

4,
17,
99,
7,
...

Large Language Models

The screenshot displays the Gemini 2.0 Flash web interface. At the top, the Gemini logo and '2.0 Flash' are visible. Below the header, there are two user queries in light blue bubbles. The first query is 'Who is the computer scientist Barry Norton?'. The second query is 'There are various areas of artificial intelligence — what other areas has he been involved in?'. To the right of each query, the model's response is shown in a white bubble, followed by a list of sources. The first response includes a paragraph about Barry Norton's background and two sources: '1. Barry Norton - Google Scholar' and '2. Milestone Systems promotes Dr. Barry Norton to Vice President of Research'. The second response includes a paragraph about his involvement with machine learning and one source: '1. Barry Norton - Google Scholar'. The interface also features a search bar at the top and a profile picture of a person.

Query 1: Who is the computer scientist Barry Norton?

Response 1: Barry Norton is a computer scientist specializing in artificial intelligence and computer vision. He has worked in research positions at various universities and technology companies. In 2018, he joined Milestone Systems, a video management software company, as Director of Research, and was promoted to Vice President of Research in 2021.

Sources for Query 1:

- 1. Barry Norton - Google Scholar
scholar.google.com
- 2. Milestone Systems promotes Dr. Barry Norton to Vice President of Research
git-sicherheit.de

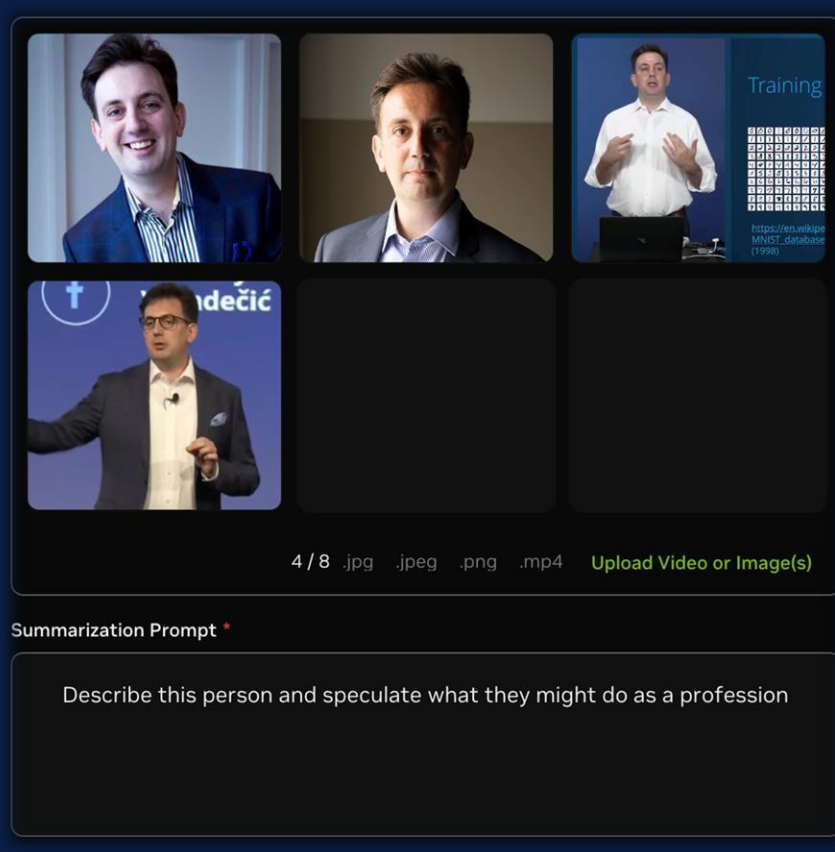
Query 2: There are various areas of artificial intelligence — what other areas has he been involved in?

Response 2: Barry Norton has been involved with machine learning, semantic web services, and knowledge graphs.

Sources for Query 2:

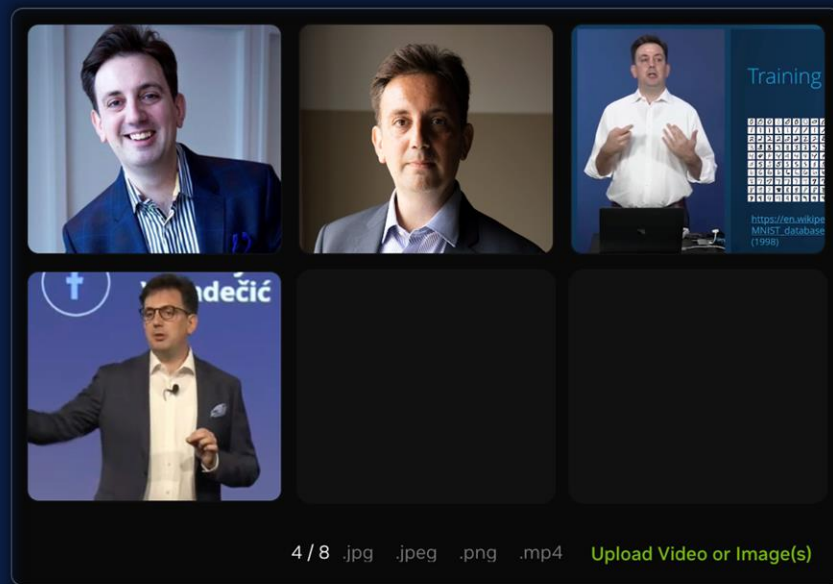
- 1. Barry Norton - Google Scholar
scholar.google.com

Vision-Language Models



The interface displays a grid of images. The top row contains three images: a man in a blue blazer, a man in a grey blazer, and a man in a white shirt standing next to a screen labeled 'Training' with a URL https://en.wikipedia.org/wiki/MNIST_database (1998). The bottom row contains one image of a man in a dark blazer and two empty slots. Below the grid is a file upload bar showing '4 / 8' and supported file types: .jpg, .jpeg, .png, .mp4, and a button 'Upload Video or Image(s)'. Below the upload bar is a 'Summarization Prompt' section with a text input field containing the prompt: 'Describe this person and speculate what they might do as a profession'.

Vision-Language Models



Summarization Prompt

Describe this person and speculate what they might do as a profession

nvidia / **cosmos-nemotron-34b** PREVIEW

Multi-modal vision-language model that understands text/img/video and creates informative responses

vlm

vision language model

image caption

image to text

Output Preview JSON

The person is a man with short, dark hair and a light beard. He's wearing a dark blue blazer over a white shirt with a striped tie, and a small blue pocket square. He appears to be well-dressed and confident, suggesting he might be a professional in a field that requires a certain level of presentation or public speaking, such as a business executive, a lawyer, or a consultant.



Hallucination



The person is a man with short, dark hair and a light complexion. He's wearing a dark blue blazer over a white shirt with a striped tie, and a small blue pocket square. He appears well-dressed and confident, suggesting he might be a professional in a field that requires a certain level of presentation or public speaking, such as a business executive, a lawyer, or a consultant.



Prompt Engineering



4 / 8 .jpg .jpg

Summarization Prompt *

The subject of these pictures is the computer scientist Barry Norton. He recently took on the role of Fellow at Milestone Systems, a global leader in data-driven video technology software. Describe what kind of work he is shown doing.

Summarization Prompt *

The subject of these pictures is the computer scientist Barry Norton. He recently took on the role of Fellow at Milestone Systems, a global leader in data-driven video technology software. Describe what kind of work he is shown doing.

nvidia / vila **PREVIEW**

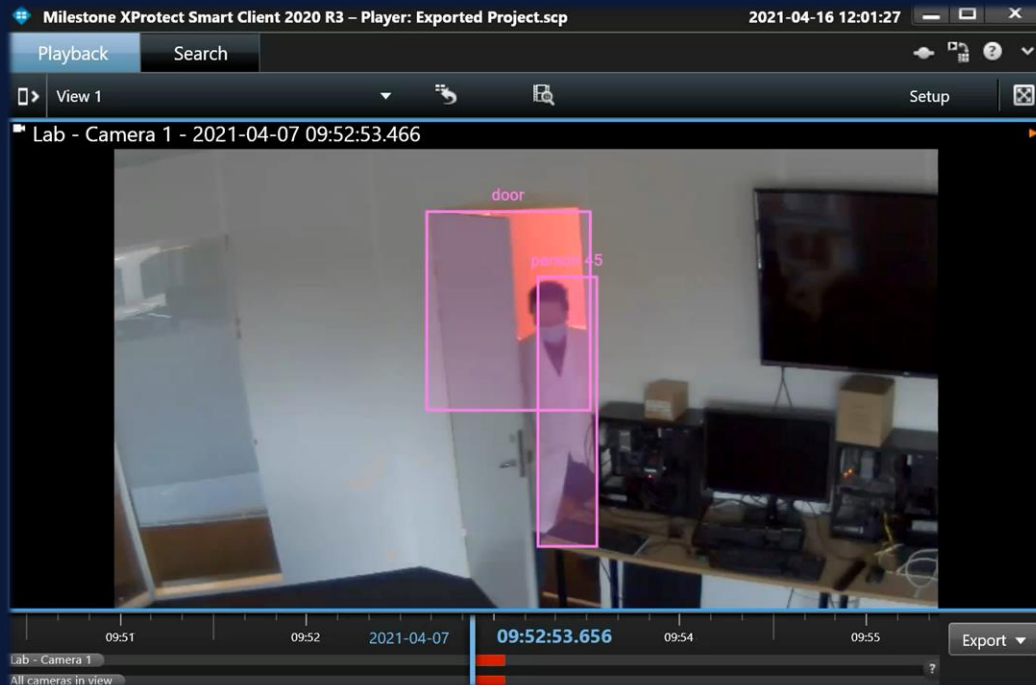
Multi-modal vision-language model that understands text/img/video and creates informative responses

vlm vision language model image caption image to text

Output **Preview** JSON

Barry Norton is a computer scientist who has been appointed as a Fellow at Milestone Systems, a company that specializes in data-driven video technology software. As a Fellow, he is likely involved in high-level research, development, and strategic planning for the company. The images show him in various settings, including a presentation or lecture, where he is speaking to an audience, and in a more formal setting, possibly a meeting or a conference.

Video Analytics



Demo



Alarm Management

Refresh

Active Alarms

100

High Priority

0

Total Alarms

100

Alarm List

All States

All Priorities

ID	Priority	Message	Name	Time	State
1398	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 12:46:03	
1397	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 12:43:59	
1396	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 12:43:17	
1395	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 12:07:31	
1394	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 10:22:19	
1393	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 09:45:58	
1392	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 09:27:12	
1391	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 09:24:18	
1390	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 09:23:48	
1389	Low	LinkedEventRising	VisEncoder Test	02/14/2025, 09:13:08	

Few-Shot / In-Context Learning



A man in a white open-necked shirt, and no jacket or tie, presenting on stage.



A man in a black t-shirt and grey jacket, with no tie, presenting on stage.



A man in graduation robes and a tie, posing in front of an academic building.

Presented as a
prompt, followed
by...



Describe this person and what they're doing...

Few-Shot & Security



A closed door with a low handle



An open door with a low handle



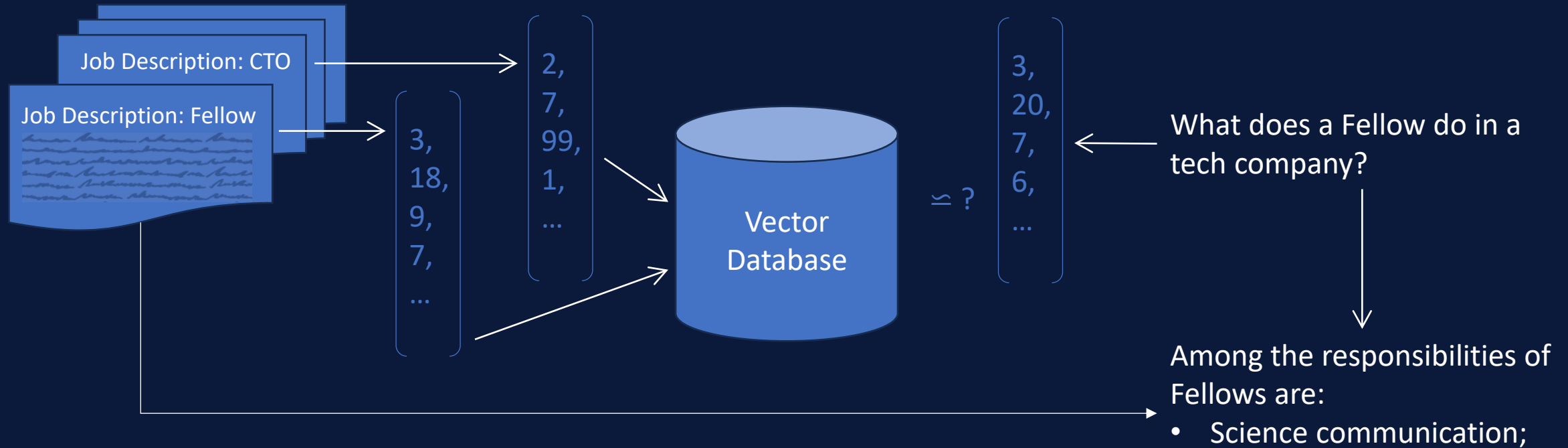
A box with mailing labels

Presented as a
prompt, followed
by...

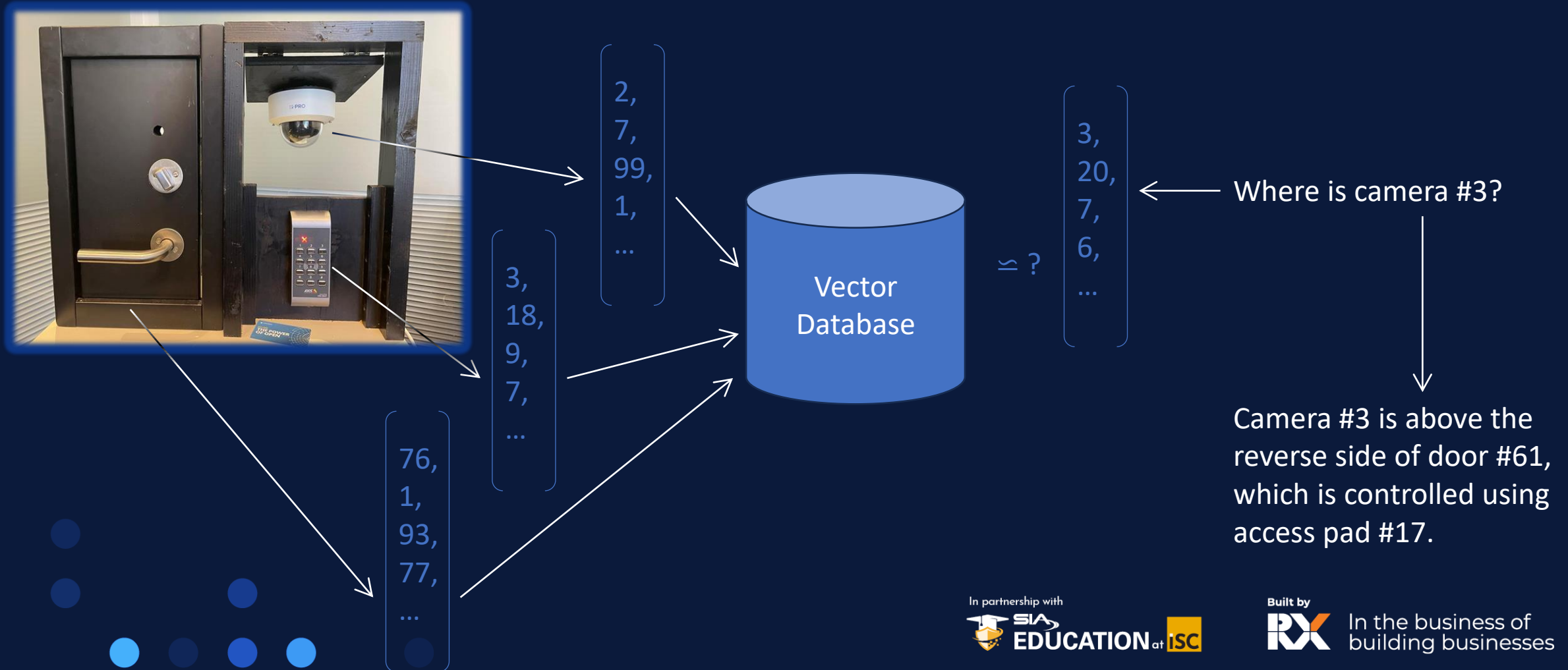


Why is the door alarm going off in this video?

Retrieval-Augmented Generation (RAG)



RAG & Security



Conclusions

- VLMs extend the capabilities of LLMs;
- Pre-trained VLMs recognize many objects and situations relevant to security;
- VLMs accuracy can be focused via:
 - prompt engineering,
 - In-context learning (few-shot examples),
 - retrieval augmented generation (RAG).



Thank you!

**Have thoughts about SIA
Education@ISC?**

Scan the QR Code on the left to provide your feedback
on SIA Education@ISC Sessions at ISC West

