



tspann 
LOG OUT 

0 53,639 / 153.08 MB

1

4

30

83

3

6

16

✓

3

0

0

0

C 22

28 EDT



Building IoT Applications with Open Source

Tim Spann, Senior Sales Engineer

AI + Streaming Weekly by Tim Spann



<https://bit.ly/32dAJft>

This week in Apache NiFi, Apache Polaris, Apache Flink, Apache Kafka, ML, AI, Streamlit, Jupyter, Apache Iceberg, Python, Java, LLM, GenAI, Vector DB and Open Source friends.

Tim Spann

Twitter: @PaasDev // Blog: datainmotion.dev

Senior Sales Engineer, Snowflake

NY/NJ/Philly - Cloud Data + AI Meetups

ex-Zilliz, ex-Pivotal, ex-Cloudera,

ex-StreamNative, ex-Hortonworks.

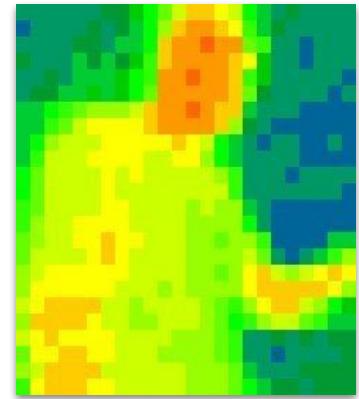
<https://medium.com/@tspann>

<https://github.com/tspannhw>





- Introduction
- Devices
- IoT Apps
- Messaging
- Edge AI
- Demos



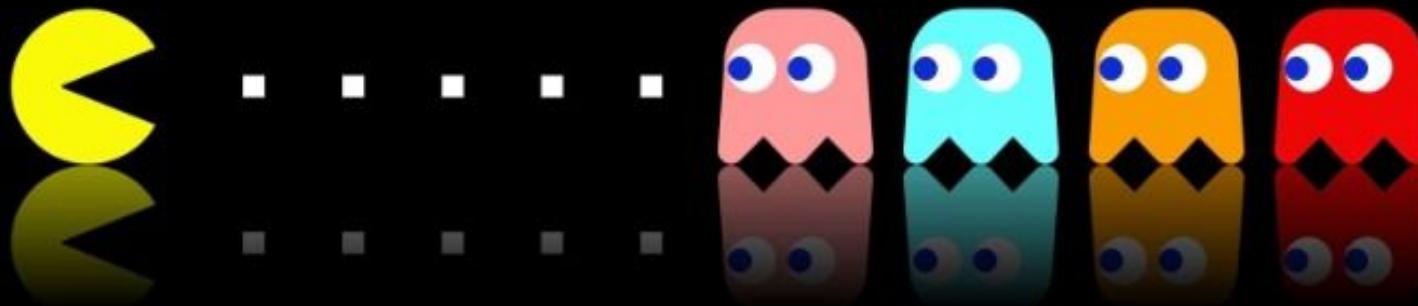
AI + All Data Weekly by Tim Spann



<https://bit.ly/32dAJft>

This week in Apache NiFi, Apache Flink, Apache Kafka, ML, AI, Streamlit, Jupyter, Apache Iceberg, Python, Java, LLM, GenAI, Vector DB and Open Source friends.

NIFI CONSUMING ALL THE DATA

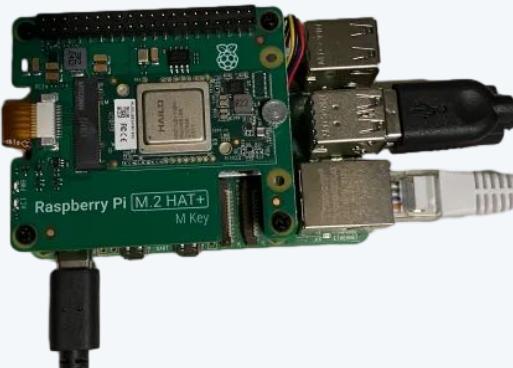


DEVICES

Raspberry Pi 5 + AI Kit

Raspberry Pi 5 with 8GB of RAM

The AI Kit adds a neural network inference accelerator capable of **13 tera-operations per second (TOPS)**, which is pretty good for \$70 US. Attached to this M.2 Hat is the Hailo-8L M.2 Entry-Level Acceleration Module which will give us our AI powers.



What is it?

<https://paperswithcode.com/task/pose-estimation>

1,431 papers with code

Human Pose Estimation is a computer vision technique that locates and estimates things like eyes, ears, shoulders, joints in motion.

It looks pretty cool and has some interesting applications for medical purposes and robotics. For me, it was one of the cool examples that runs on the AI Kit.

Pose Estimation by Hailo 8L

Each person is identified and represented by 17 keypoints

Examples

nose, eyes, ears, shoulders, elbows, wrists, hips, knees, and ankles.

We are tracking eyes and more (updated today)

<https://github.com/tensorboy/centerpose>

<https://softwaremill.com/human-pose-estimation-2023-guide/>

https://github.com/hailo-ai/hailo_model_zoo/blob/master/docs/public_models/HAILO8/HAILO8_pose_estimation.rst

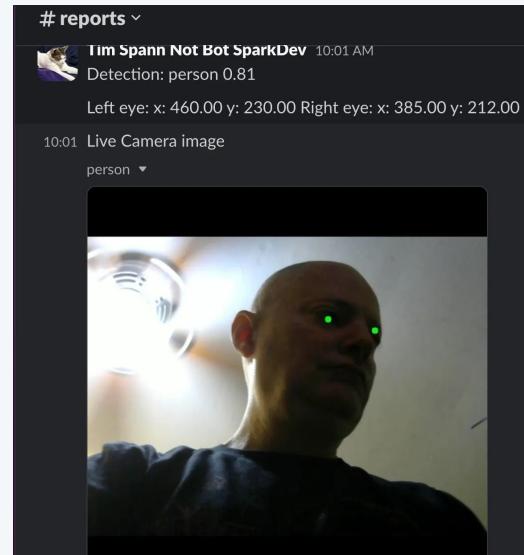
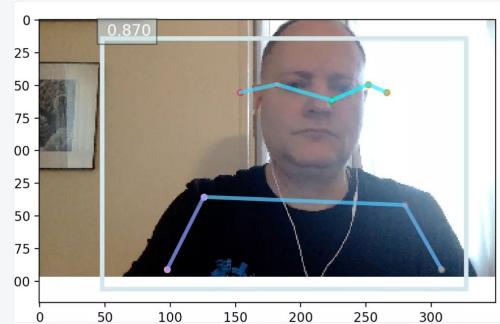
Pose Estimation on Hailo 8L

Pose Estimation COCO

Yolov8s_pose

Hailo-8L

<https://github.com/ultralytics/ultralytics>

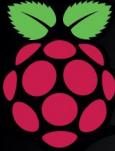


HAILO Raspberry Pi 5 Example Apps



Raspberry Pi 5 Example Apps

- Rich real world use cases
- Quick deployment
- A starting point for building your own AI applications



Raspberry Pi®

<https://github.com/hailo-ai/hailo-rpi5-examples>

New: CLIP Zero Shot Inference Application

Alternatives

- Just Released AI Kit+ with 26 Tops
- NVIDIA Jetson Series
- Smart Cameras like OAK-D
- Specialized Devices

Edge Vector Olympics

Gold - NVIDIA Jetson AGX Orin -
275 TOPS, 2048-core, 64 GB
RAM

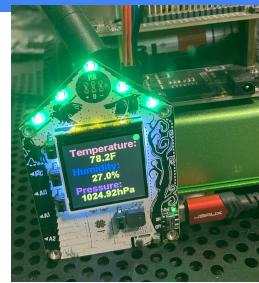
Silver - NVIDIA Jetson Xavier NX,
21 TOPS, 384-core, 8 GB RAM

Bronze - Raspberry Pi 5, 13
TOPS, 4-core, 8 GB RAM

Edge Computing Power - Edge Server



- Containers
- 64 bit processors and operating systems
- 8-64 GB Modern RAM
- Fast WiFi / Bluetooth
- 300+ Core GPUs
- eMMC Fast Storage
- TBs of SSD
- Examples: NVIDIA JETSON XAVIER NX,
NVIDIA JETSON ORIN AGX



Device 1 - AdaFruit Funhouse

- <https://github.com/tspannhw/pulsar-adafruit-funhouse>

(MQTT)

Raw JSON:

```
{"pressure": 1009.08,  
 "button_sel": "off",  
 "pir_sensor": "off",  
 "humidity": 36.0422, "temperature": 80.9526,  
 "button_down": "off", "captouch6": "off",  
 "captouch7": "off", "button_up": "off", "captouch8": "off",  
 "light": 6990}
```

Processor 240MHz / **RAM** 2+4MB

Device 2 - Raspberry Pi



- <https://github.com/tspannhw/FLiP-Pi-DeltaLake-Thermal>

Pulsar Protocol Raw JSON:

```
{"uuid": "thrml_zda_20220715182748", "ipaddress": "192.168.1.204",  
"cputempf": 108, "runtime": 0, "host": "thermal", "hostname": "thermal",  
"macaddress": "e4:5f:01:7c:3f:34", "endtime": "1657909668.7279365",  
"te": "0.0007398128509521484", "cpu": 1.8,  
"diskusage": "105078.0 MB",  
"memory": 9.0, "rowid": "20220715182748_fc4cbbb1-79da-4cla-8991-78bd23c9f221",  
"systemtime": "07/15/2022 14:27:53", "ts": 1657909673,  
"starttime": "07/15/2022 14:27:48",  
"datetimesecond": "2022-07-15 18:27:52.492469+00:00", temperature: 28.238,  
humidity: 29.61, co2: 992.0}
```

Processor 1.5 GHz, 64-bit quad-core / **RAM** 2-8 GB LPDDR4-3200 SDRAM

STREAMING



IoT to Cloud Data Platform



Ingest
Processors

Ingest
Gateway

UNIVERSAL
DATA DISTRIBUTION
(Ingest, Transform, Deliver)

Router, Filter &
Transform
Processors

Destination
Processors



Aggregate Data

Analyze Data

Source Data

Share Data



snowflake®

DATAVOLO



DATA
WAREHOUSE



DATA
LAKE



COLLABORATION



DATA
ENGINEERING



DATA
SCIENCE
& ML



APPLICATIONS



UNISTORE



CYBERSECURITY

Azure

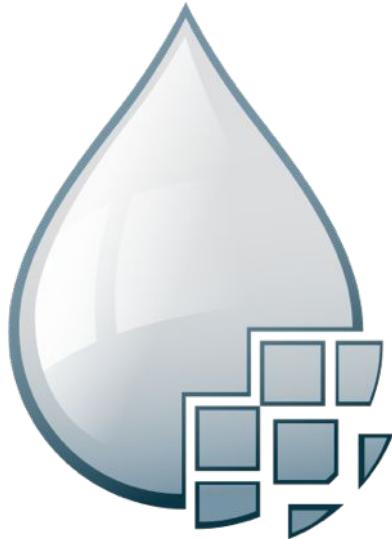
aws

Google Cloud



© 2024 Snowflake Inc. All Rights Reserved

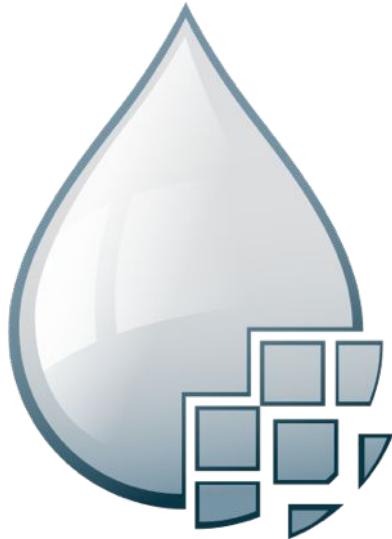
DataFlows for Data Ingest, Movement and Routing



- Guaranteed delivery
- Data buffering
 - Backpressure
 - Pressure release
- Prioritized queuing
- Flow specific QoS
 - Latency vs. throughput
 - Loss tolerance
- Data provenance
- Supports push and pull models
- Hundreds of processors
- Visual command and control
- Over a 200 sources
- Flow templates
- Pluggable/multi-role security
- Designed for extension
- Clustering
- Version Control



The Power of Apache NiFi



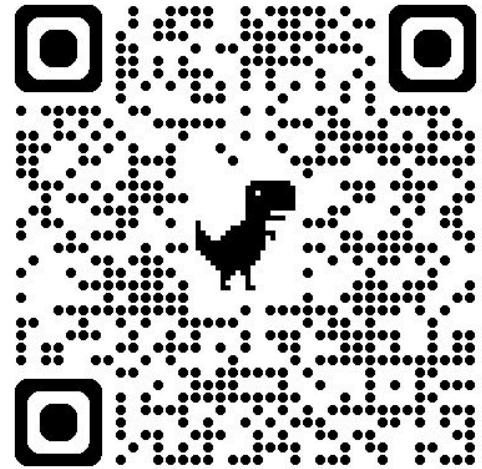
- Moving Binary, Unstructured, Image and Tabular Data
- Enrichment
- Universal Visual Processor
- Simple Event Processor
- Routing
- Feeding data to Central Messaging
- Support for modern protocols
- Kafka Protocol Source/Sink
- Pulsar Protocol Source/Sink

NIFI 2.0 FEATURES

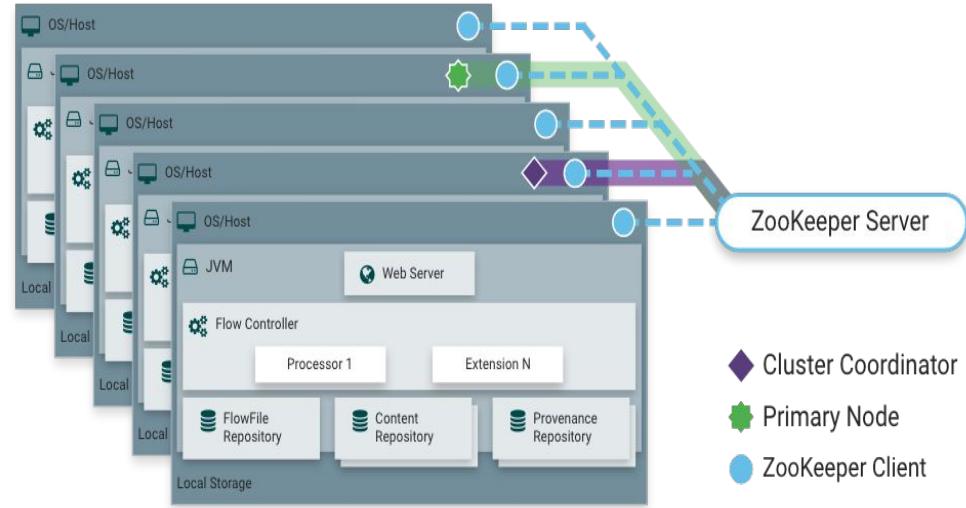
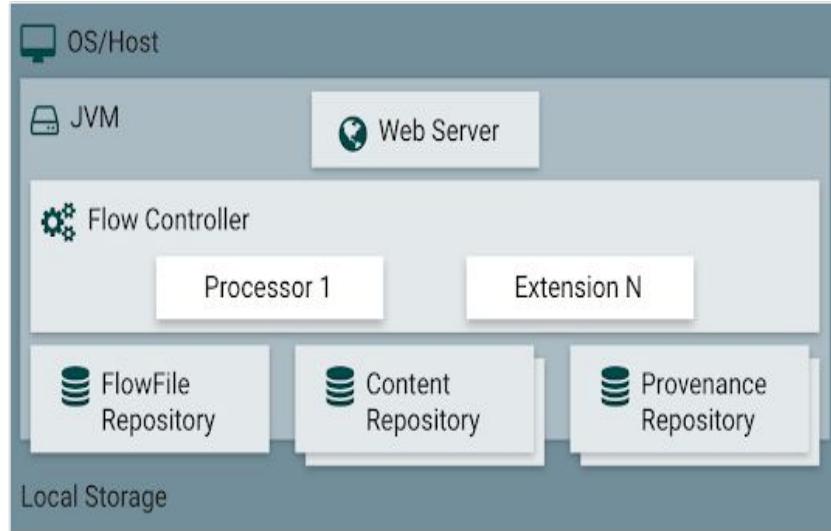
DataFlow is built for Real-Time Integration and AI

Major Updates:

- Python Integration
- ParameterIZATION
- JDK 21+
- Provenance / Data Lineage
- Rules Engine for Development Assistance
- Additional Azure Processors
- Integration with Zendesk, Slack,
- Database Tables as Schemas
- Amazon Glue Schema Registry
- OpenTelemetry Support



Architecture



- ◆ Cluster Coordinator
- ★ Primary Node
- ZooKeeper Client

<https://nifi.apache.org/docs/nifi-docs/html/overview.html>



What is Apache NiFi?

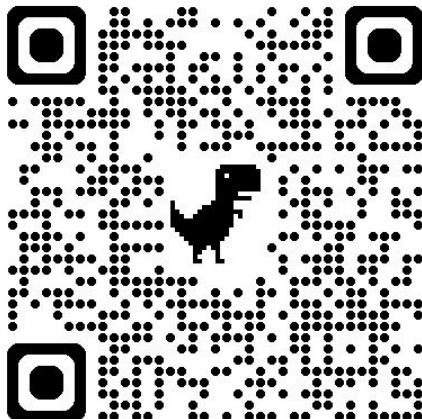
Apache NiFi is a scalable, real-time streaming data platform that collects, curates, and analyzes data so customers gain key insights for immediate actionable intelligence.





Address To Lat/Long

- Python 3.10+
- geopy Library
- Nominatim
- OpenStreetMaps (OSM)
- openstreetmap.org/copyright
- Returns as attributes and JSON file
- Works with partial addresses
- Categorizes location
- Bounding Box





Flink SQL

- Streaming Analytics
- Continuous SQL
- Continuous ETL
- Complex Event Processing
- Standard SQL Powered by Apache Calcite

The screenshot shows the Apache Flink Dashboard interface. On the left, there's a sidebar with options like Overview, Jobs, Task Managers, and Job Manager. The main area is titled 'Apache Flink Dashboard' and shows a 'Running' job named 'xenodochial_noocyte' with ID 'aa443658ed280787e73d8f8903676b4'. It indicates a start time of '2021-04-07 10:08:37' and a duration of '3h 6m 21s'. Below this, there are tabs for Overview, Exceptions, Timeline, Checkpoints, and Configuration. The Overview tab shows a detailed flow diagram of the job's execution. The flow starts with a 'Source: kfsdbsource: weather2 -> Kafka 10 assigner -> SourceConnector[weather2]', followed by a 'Filter[one]' step, then a 'Map[filter]' step, and finally a 'Sink: Webhook Process -> Sink: Webhook http sink'. The 'Parallelism: 1' is indicated. To the right of the flow diagram, there are sections for 'Detail', 'SubTasks', 'TaskManagers', 'Watermarks', 'Accumulators', 'BackPressure', and 'Metrics'. The 'Metrics' section includes a table with columns: Name, Status, Bytes Received, Records Received, Bytes Sent, Records Sent, Parallelism, Start Time, and Tasks. The table shows three entries: 'Source: kfsdbsource: weather2 -> Kafka 10 assigner -> SourceC...', 'Kafka Consumer Source', and 'Webhook Process -> Sink: Webhook http sink', all in 'RUNNING' status.

Name	Status	Bytes Received	Records Received	Bytes Sent	Records Sent	Parallelism	Start Time	Tasks
Source: kfsdbsource: weather2 -> Kafka 10 assigner -> SourceC...	RUNNING	0.0	0	8.33 MB	8,031	1	2021-04-0 10:08:37	1
Kafka Consumer Source	RUNNING	0.0	0	0.0	0	1	2021-04-0 10:08:37	1
Webhook Process -> Sink: Webhook http sink	RUNNING	8.34 MB	8,002	0.0	0	1	2021-04-0 10:08:37	1

Edge Models

HuggingFaceTB/SmoILM2-1.7B-Instruct

Small language models (SLMs)

NVIDIA Edge AI / Physical AI Edge Models

Raspberry Pi Edge AI



DEMO

IS THIS ENOUGH DATA?

RESOURCES AND WRAP-UP

Raspberry Pi AI Kit - Hailo
Edge AI

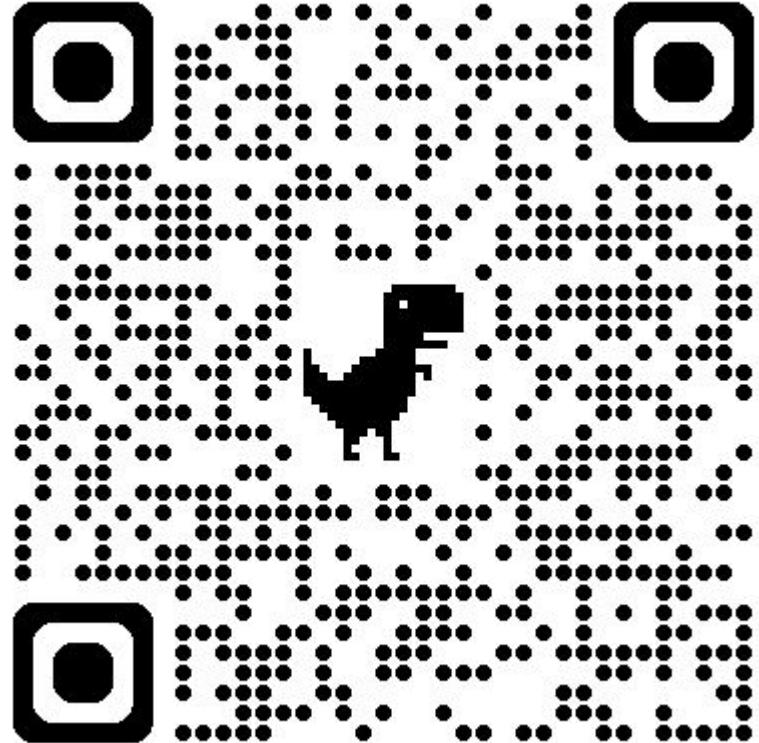
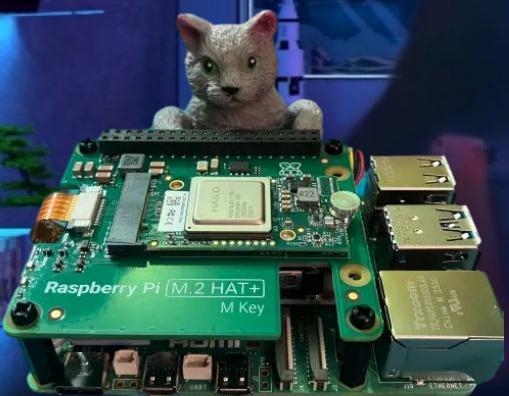


Milvus

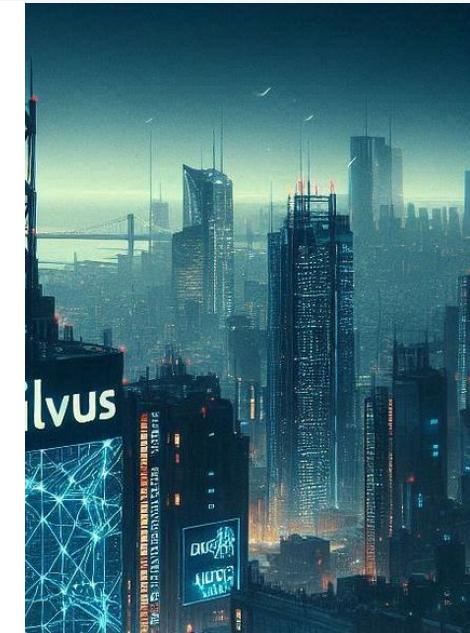
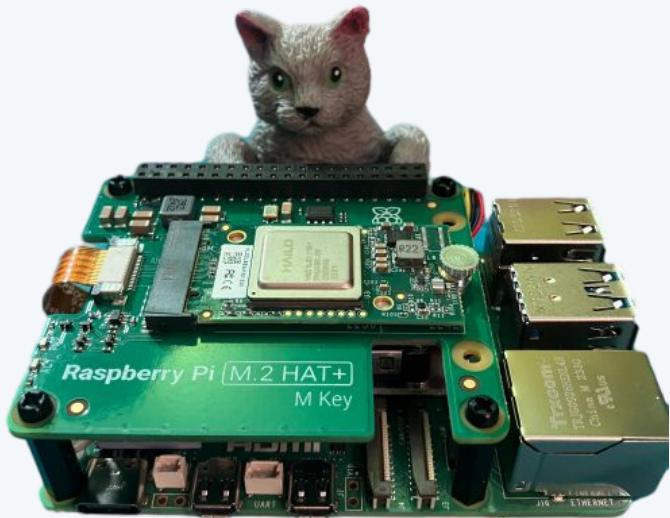


<https://medium.com/@tspann/unstructured-data-processing-with-a-raspberry-pi-ai-kit-c959dd7fff47>

Raspberry Pi AI Kit Hailo Edge AI Pose Estimation

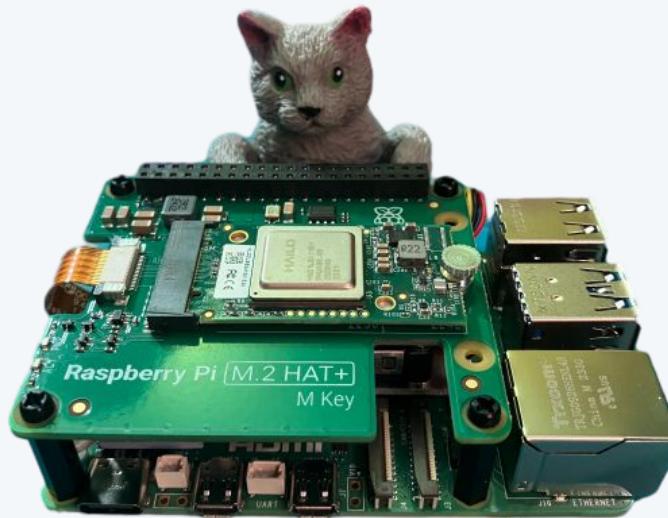


Code -



<https://bit.ly/4ftn04t>

Code - Pose Estimation



<https://bit.ly/4ebEPUJ>

Walk Through Article



<https://bit.ly/4hxjvvF>



/tspannhw/AIM-BecomingAnAIEngineer: AIM - Becoming An AI Engineer

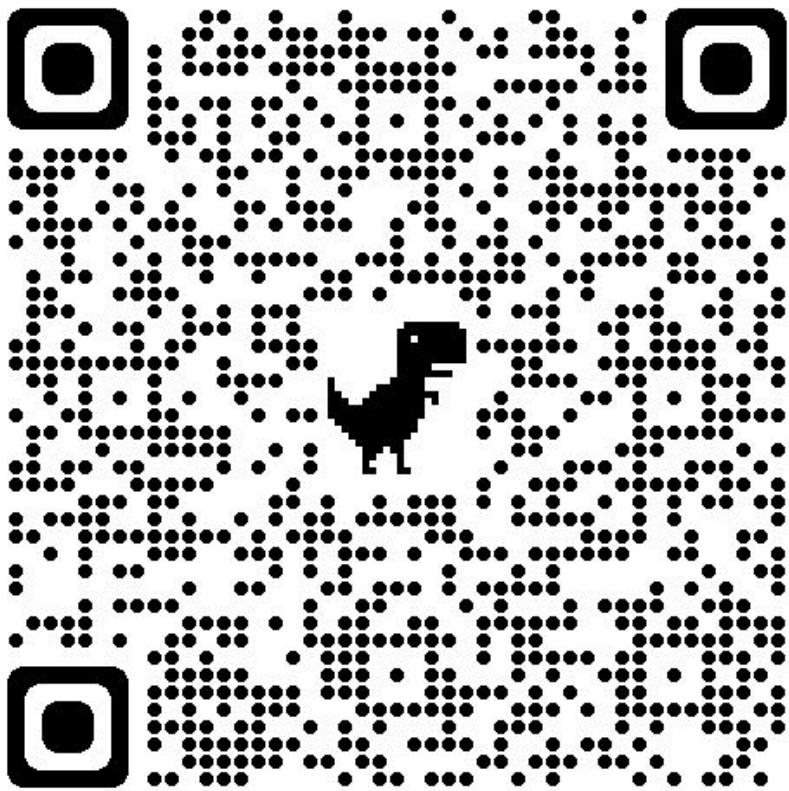


<https://bit.ly/3BV4IKX>

What's in the Air Tonight Mr. Milvus?



<https://bit.ly/4fQhBog>



Street Cameras



<https://medium.com/cloudera-inc/streaming-street-cams-to-yolo-v8-with-python-and-nifi-to-minio-s3-3277e73723ce>

Tim **SPANN**

<https://github.com/tspannhw>

<https://www.datainmotion.dev/>



