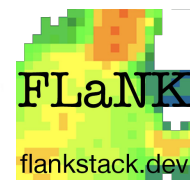


# Using the FLaNK Stack for Edge AI

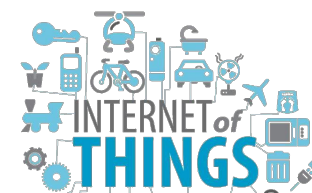


Timothy Spann  
Principal DataFlow Field Engineer  
Cloudera  
**@PaasDev**



AI and IoT Bulgaria Summit, 2021

June 26<sup>th</sup>



Sofia,  
Bulgaria

*Tim*

# SPANN

<https://github.com/tspannhw>

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

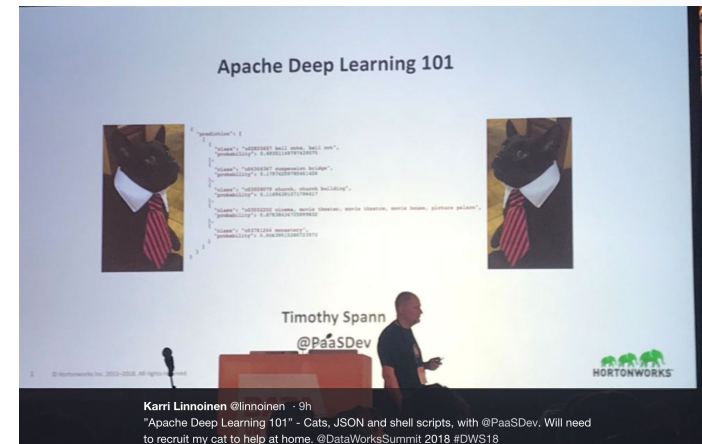
<https://www.meetup.com/futureofdata-princeton/>



# Speaker Bio

Principal Data Flow Field Engineer  
Cloudera

DZone Zone Leader and Big Data MVB;  
ex-Real-Time Streaming Engineer @ Energy Device Startup  
Princeton NJ Future of Data Meetup;  
@PaasDev  
ex-Hortonworks Senior Sales Engineer  
ex-Pivotal Field Engineer  
ex-HPE Systems Architect V  
<https://github.com/tspannhw> <https://www.datainmotion.dev/>  
<https://github.com/tspannhw/SpeakerProfile>  
<https://dev.to/tspannhw>  
<https://sessionize.com/tspann/>  
<https://www.slideshare.net/bunkertor>



# Welcome to Future of Data - Princeton - Virtual



<https://www.meetup.com/futureofdata-princeton/>

From Big Data to AI to Streaming to Containers to Cloud to Analytics to Cloud Storage to Fast Data to Machine Learning to Microservices to ...



# Agenda

---

- Who, What, Where, Why
- What Tools
- Deep Learning Integration
- Device Agents
- Agents, Server, Data, Action and Architecture
- What is the Data? What is the hardware?
- Continuous SQL
- Demo
- Takeaways
- Upcoming Events
- Thanks!



# FLaNK Stack for Cloud Data Engineers - Edge AI

Multiple users, frameworks, languages, clouds, data sources & clusters



CLOUD DATA ENGINEER

- Experience in ETL/ELT
- Coding skills in Python or Java
- Knowledge of database query languages such as SQL
- Experience with Streaming
- Knowledge of Cloud Tools



CAT

- Expert in ETL (Eating, Ties and Laziness)
- Edge Camera Interaction
- Typical User
- No Coding Skills
- Can use NiFi
- Questions your cloud spend



AI / Deep Learning / ML / DS

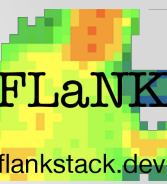
- Can run in Apache NiFi
- Can run in Kafka Streams
- Can run in Apache Flink
- Can run in MiNiFi Agents



# Apache Tools and Frameworks Used



# Apache MXNet Native Processor through DJL.AI for Apache NiFi



```
#workshop
11:30 AM =====
Deep Learning Class Label: person
File: cc0a469f-c108-42c7-95c6-10e5fda95006.person.png
Probability: 0.96
UUID: 32ef65a3-0650-42cd-965c-ba25597eb1ad
Rank: 1
Bounding Box (Height/Width, X,Y)
0.74 / 0.69
0.27, 0.25
Image (Height/Width, X,Y)
480 / 640
0, 0
=====
```

tspann 11:30 AM  
371bdb8f-35bc-4a2a-919c-bdeb609b726c.person.png



```
return Paths.get(uri).toAbsolutePath().getParent().toString();

private void runAndAssertHappy() {
    testRunner.setValidateExpressionUsage(false);
    testRunner.run();
    testRunner.assertValid();

    testRunner.assertAllFilesTransferred(DeepLearningProcessor.REL_SUCCESS);
    List<MockFlowFile> successFiles = testRunner.getFlowFilesForRelationship(DeepLearningProcessor.REL_SUCCESS);

    for (MockFlowFile mockFile : successFiles) {
        assertEquals("mockFile", mockFile.getAttribute("class"), "car");
        assertEquals("mockFile", mockFile.getAttribute("probability"), "1.00");
        Map<String, String> attributes = mockFile.getAttributes();
        System.out.println("Size: " + attributes.size());
        for (String attribute : attributes.keySet()) {
            System.out.println("Attribute: " + attribute);
        }
    }
}

@Test
public void testProcessor() throws Exception {
    java.io.File resourcesDirectory = new java.io.File("resources");
    System.out.println(resourcesDirectory.getAbsolutePath());

    testRunner.setProperty(DeepLearningProcessor.BACKGROUND_PROCESSOR, true);
    testRunner.setProcevt(DeepLearningProcessor.DATASET_PATH, resourcesDirectory);
    DeepLearningProcessorTest testProcessor()
}

Run: DeepLearningProcessorTest.testProcessor
Tests passed: 1 of 1 test - 4 s 818 ms
DeepLearningProcessorTest (con: 4 s 818 ms)
testProcessor 4 s 818 ms
Attribute: boundingbox_height_1 = 0.15
Attribute: probability_1 = 1.00
Attribute: image_min_y_1 = 0
Attribute: image_min_x_1 = 0
Attribute: class_1 = car
Attribute: rank_1 = 1
Attribute: uuid = e5953c52-75ab-4849-8876-a25796714984
Attribute: boundingbox_width_1 = 0.24
```

Attribute Values	
boundingbox_height_1	0.99
No value set	
boundingbox_width_1	0.90
No value set	
boundingbox_x_1	0.09
No value set	
boundingbox_y_1	0.01
No value set	
class_1	tvmonitor
No value set	
filename	2020-08-26_1330.jpg.tvmonitor.png
	2020-08-26_1330.jpg (previous)

This processor uses the DJL.AI Java Interface

<https://github.com/tspannhw/nifi-djl-processor>

<https://dev.to/tspannhw/easy-deep-learning-in-apache-nifi-with-djl-2d79>





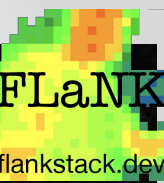
# MXNetm.. FLaNiFiKafka

- <https://www.slideshare.net/bunkertor/apache-deep-learning-101-apach econ-montreal-2018-v031>
- <https://www.slideshare.net/bunkertor/apache-deep-learning-202-washi ngton-dc-dws-2019>
- <https://www.slideshare.net/bunkertor/apache-deep-learning-201-barce lona-dws-march-2019>

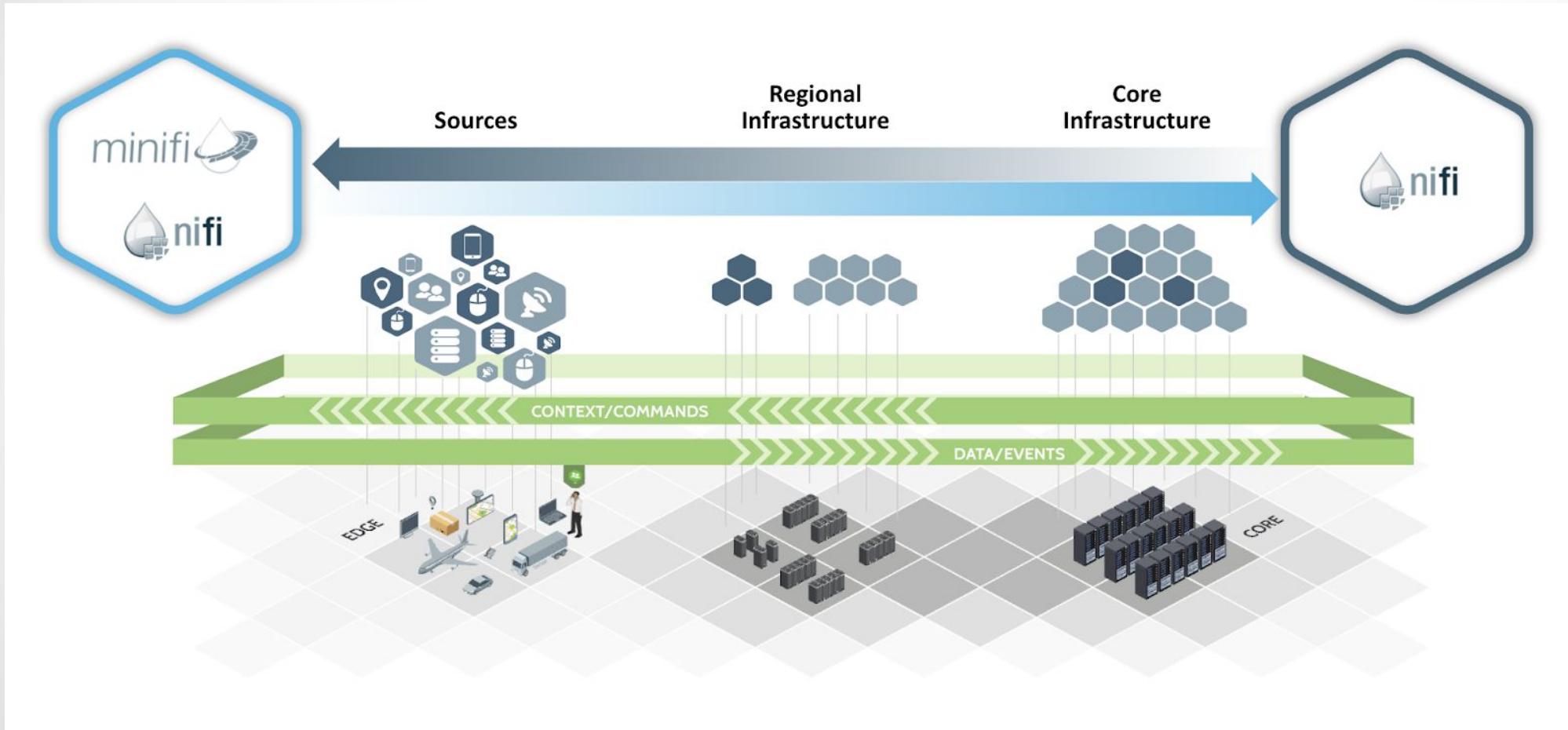


# MiNiFi.. FLaNK

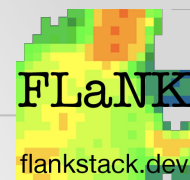
- <https://www.datainmotion.dev/2019/03/using-raspberry-pi-3b-with-apache-nifi.html>
- <https://www.datainmotion.dev/2019/05/cloudera-edge-management-introduction.html>
- <https://www.datainmotion.dev/2019/11/running-demo-apache-flink-application.html>
- <https://www.datainmotion.dev/2019/11/learning-apache-flink-19.html>
- <https://www.datainmotion.dev/2019/10/migrating-apache-flume-flows-to-apache-42.html>



# What is Apache NiFi and MiNiFi used for?



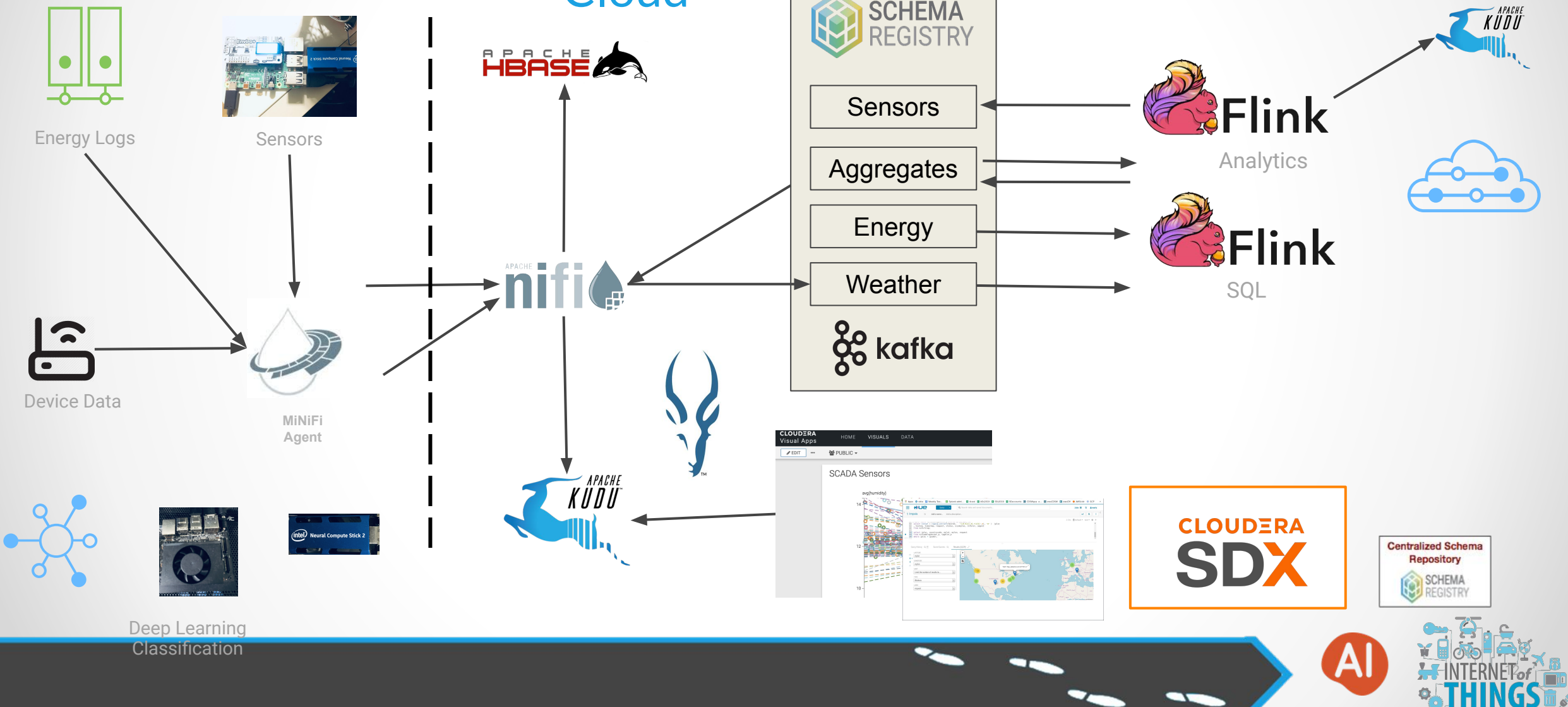
# Edge AI to Cloud Streaming Pipeline



## Edge

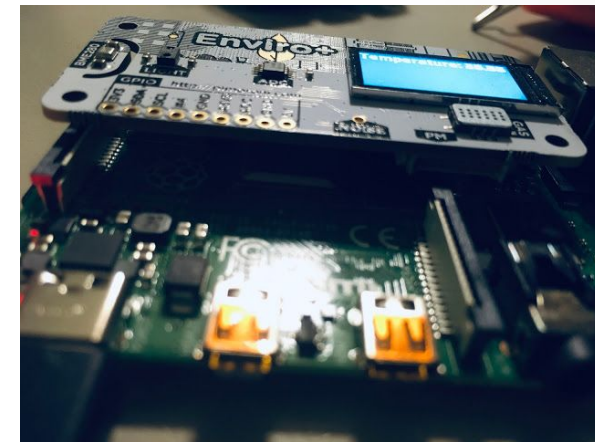
## Private Cloud

## Multi-Public Cloud



# SHOW ME THE DATA

```
{
  "uuid": "rpi4_uuid_jfx_20200826203733",
  "amplitude100": 1.2,
  "amplitude500": 0.6,
  "amplitude1000": 0.3,
  "lownoise": 0.6,
  "midnoise": 0.2,
  "highnoise": 0.2,
  "amps": 0.3,
  "ipaddress": "192.168.1.76",
  "host": "rp4",
  "host_name": "rp4",
  "macaddress": "6e:37:12:08:63:e1",
  "systemtime": "08/26/2020 16:37:34",
  "endtime": "1598474254.75",
  "runtime": "28179.03",
  "starttime": "08/26/2020 08:47:54",
  "cpu": 48.3,
  "cpu_temp": "72.0",
  "diskusage": "40219.3 MB",
  "memory": 24.3,
  "id": "20200826203733_28ce9520-6832-4f80-b17d-f36c21fd8fc9",
  "temperature": "47.2",
  "adjtemp": "35.8",
  "adjtempf": "76.4",
  "temperaturef": "97.0",
  "pressure": 1010.0,
  "humidity": 8.3,
  "lux": 67.4,
  "proximity": 0,
  "oxidising": 77.9,
  "reducing": 184.6,
  "nh3": 144.7,
  "gasKO": "Oxidising: 77913.04 Ohms\nReducing: 184625.00 Ohms\nNH3: 144651.47 Ohms"}
}
```



## WHERE DID THAT DATA COME FROM?

BME280 - temperature, pressure, humidity sensor

LTR-559 - light and proximity sensor

MICS6814 - analog gas sensor

ADS1015 ADC

MEMS - microphone

0.96-inch, 160 x 80 color LCD



# Flink SQL Examples

```
INSERT INTO global_sensor_events
```

```
SELECT scada.uuid, scada.systemtime ,  
scada.temperaturef ,  
scada.pressure , scada.humidity ,  
scada.lux , scada.proximity ,  
scada.oxidising ,  
scada.reducing , scada.nh3 ,  
scada.gasko,energy.`current` ,  
energy.voltage ,energy.`power` ,  
Energy.`total` ,energy.fanstatus
```

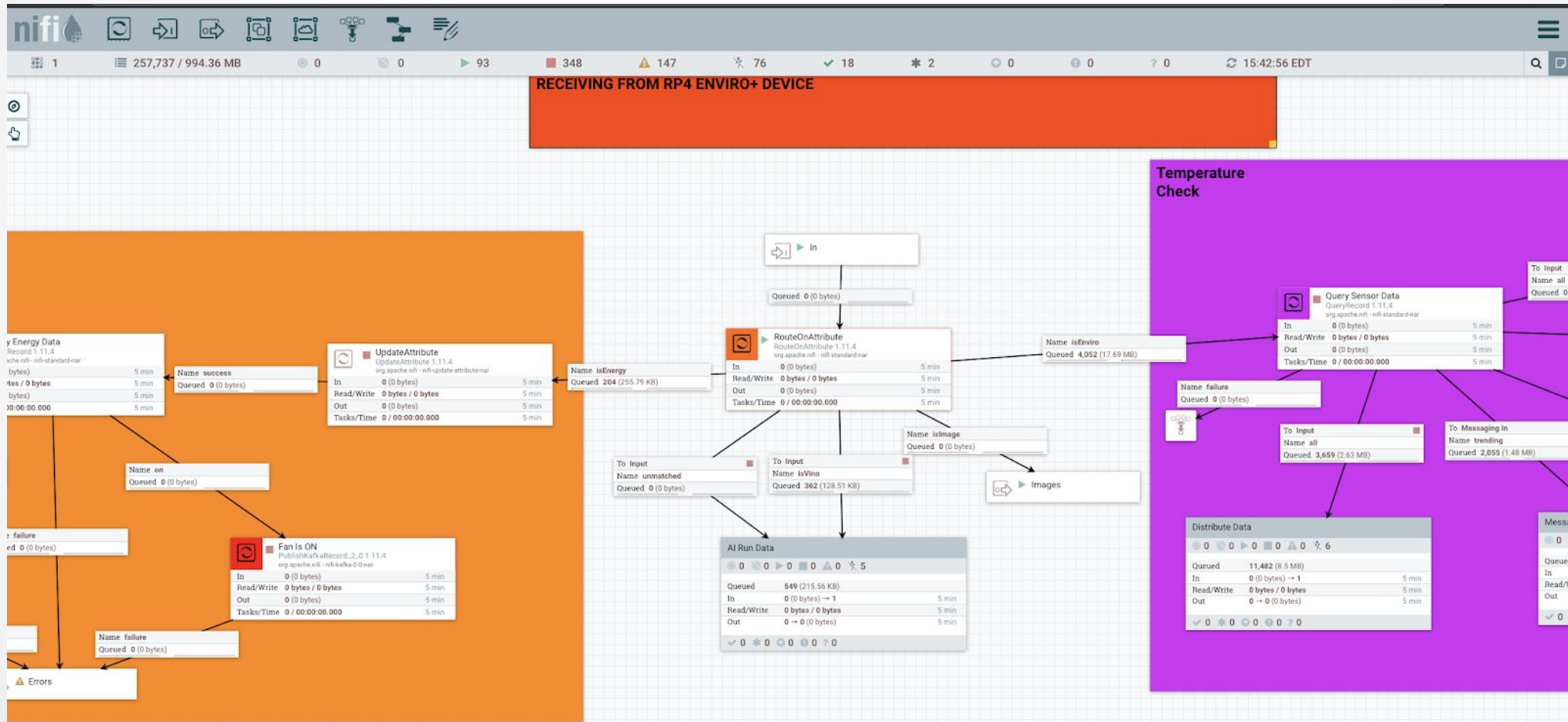
```
FROM energy, scada
```

```
WHERE
```

```
scada.system time >= energy.systemtime;
```



# Demo



The screenshot shows the NiFi web interface with the following components and metrics:

- Header:** nifi logo, navigation icons, and status bars showing 257,737 / 994.36 MB, 93 tasks, and 15:42:56 EDT.
- RECEIVING FROM RP4 ENVIRO+ DEVICE:** A large red banner at the top of the canvas.
- Flow Process:**
  - In** (Source) - Queued 0 (0 bytes)
  - RouteOnAttribute** (Router) - RouteOnAttribute 1.11.4, org.apache.nifi-nifi-standard-nar. Queued 4,052 (17.69 MB).
    - Name isEnviro:** Queued 4,052 (17.69 MB)
    - Name isEnergy:** Queued 204 (255.79 KB)
    - Name isVino:** Queued 362 (128.51 KB)
    - Name isImage:** Queued 0 (0 bytes)
  - AI Run Data** (Sink) - Queued 549 (215.56 KB), In 0 (0 bytes) → 1, Read/Write 0 bytes / 0 bytes, Out 0 → 0 (0 bytes).
  - Query Sensor Data** (Sink) - QueryRecord 1.11.4, org.apache.nifi-nifi-standard-nar. Queued 3,659 (2.63 MB).
    - Name failure:** Queued 0 (0 bytes)
    - To Messaging In:** Name trending, Queued 2,055 (1.48 MB)
  - Distribute Data** (Sink) - Queued 11,482 (8.5 MB), In 0 (0 bytes) → 1, Read/Write 0 bytes / 0 bytes, Out 0 → 0 (0 bytes).
  - UpdateAttribute** (Transformer) - UpdateAttribute 1.11.4, org.apache.nifi-nifi-update-attribute-nar. Queued 0 (0 bytes).
  - Other components:** Name success (Queued 0), Name on (Queued 0), Fan Is ON (PublishKafkaRecord 2.0.1.11.4, org.apache.nifi-nifi-kafka2-0.nar), Name failure (Queued 0), and Errors (Sink).





# Takeaways

---

- IoT can be easy with Open Source
- Download examples start today
- Start with a single device
- Capture data locally
- Send data to Azure securely
- Data provenance helps debugging, monitoring and metrics
- Share with the community



## DEMO SOURCE CODE

- <https://github.com/tspannhw/FlinkForwardGlobal2020>
- <https://github.com/tspannhw/ApacheConAtHome2020>
- <https://github.com/tspannhw/minifi-xaviernx>
- <https://github.com/tspannhw/minifi-jetson-nano>
- <https://github.com/tspannhw/minifi-enviropius>
- <https://github.com/tspannhw/EverythingApacheNiFi>
- <https://github.com/tspannhw/CloudDemo2021>
- <https://github.com/tspannhw/FlinkSQLWithCatalogsDemo>

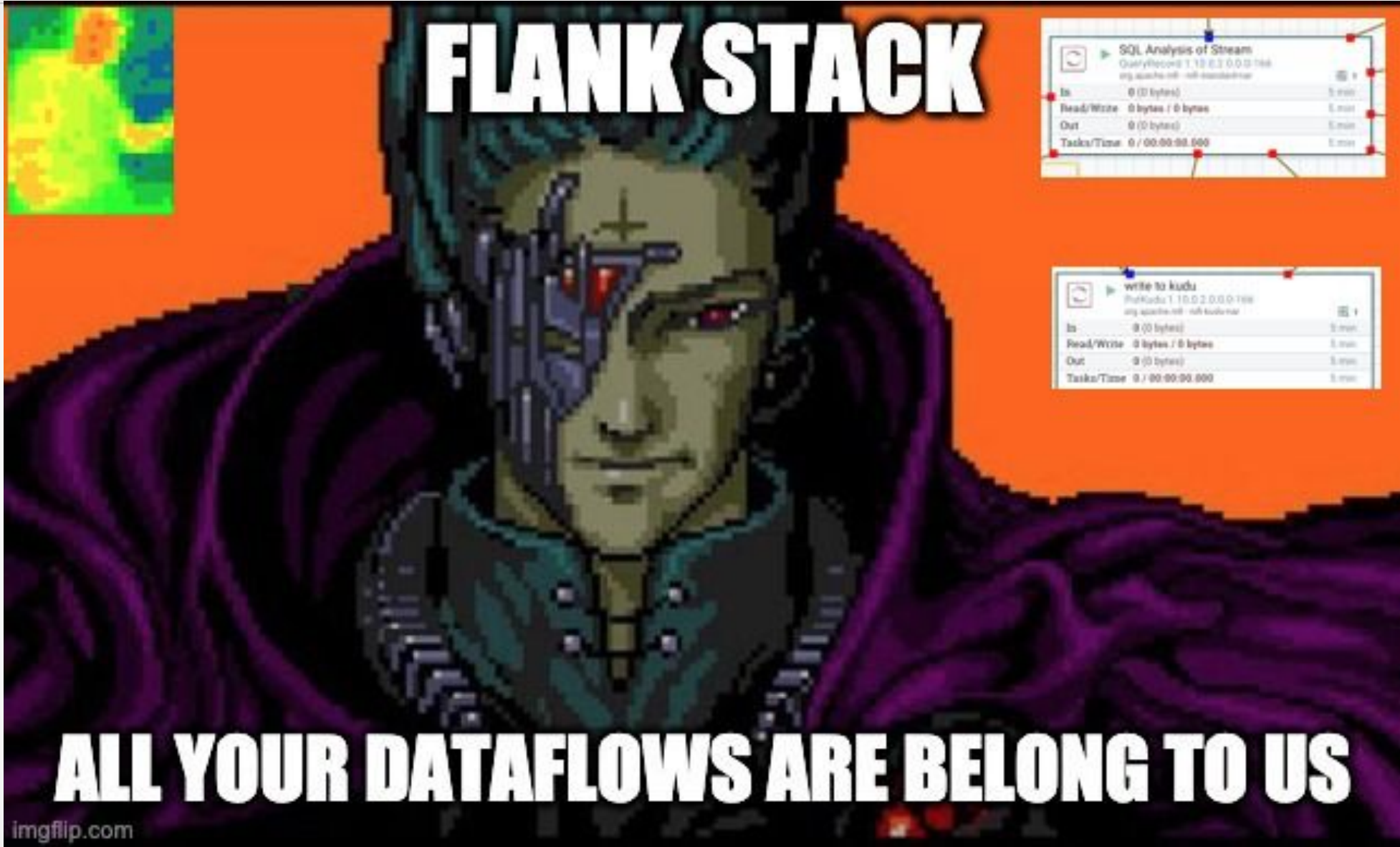
The code, build scripts, schemas, table DDL, Flink SQL, Kafka Connect configuration, NiFi flows, HBase tables, Kudu tables, Hive tables, HDFS directories, alerts, images, HTML, docs, links and all the goodies are here. Please **fork** and contribute.



## DEEPER CONTENT

- <https://www.datainmotion.dev/2020/10/running-flink-sql-against-kafka-using.html>
- <https://www.datainmotion.dev/2020/10/top-25-use-cases-of-cloudera-flow.html>
- <https://github.com/tspannhw/EverythingApacheNiFi>
- <https://github.com/tspannhw/CloudDemo2021>
- <https://github.com/tspannhw/StreamingSQLExamples>





# Upcoming Events

---



Data Saturday Plovdiv, 2021

August 28

[Eventbrite](#)



Data Saturday Sofia, 2021

October 09

[Eventbrite](#)

**js.talks();**

jsTalks (Bulgaria), 2021

Nov 19-20

<http://jstalks.net/>



# Thanks to our Sponsors

---



**BOSCH**

Invented for life



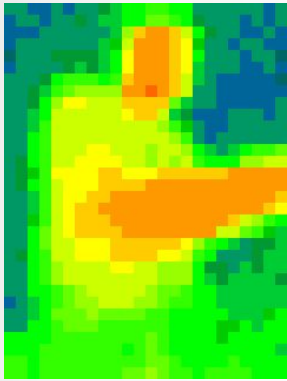
**MENTORMATE™**




Digital Ideas Accelerated



Microsoft





TH  N  Y  U

