NECTEC
a member of NSTDA

NECTEC-ACE 2017 ANNUAL CONFERENCE & EXHIBITION

BigStream

แพลตฟอร์มสำหรับการบริหารจัดการข้อมูลแบบเรียลไทม์

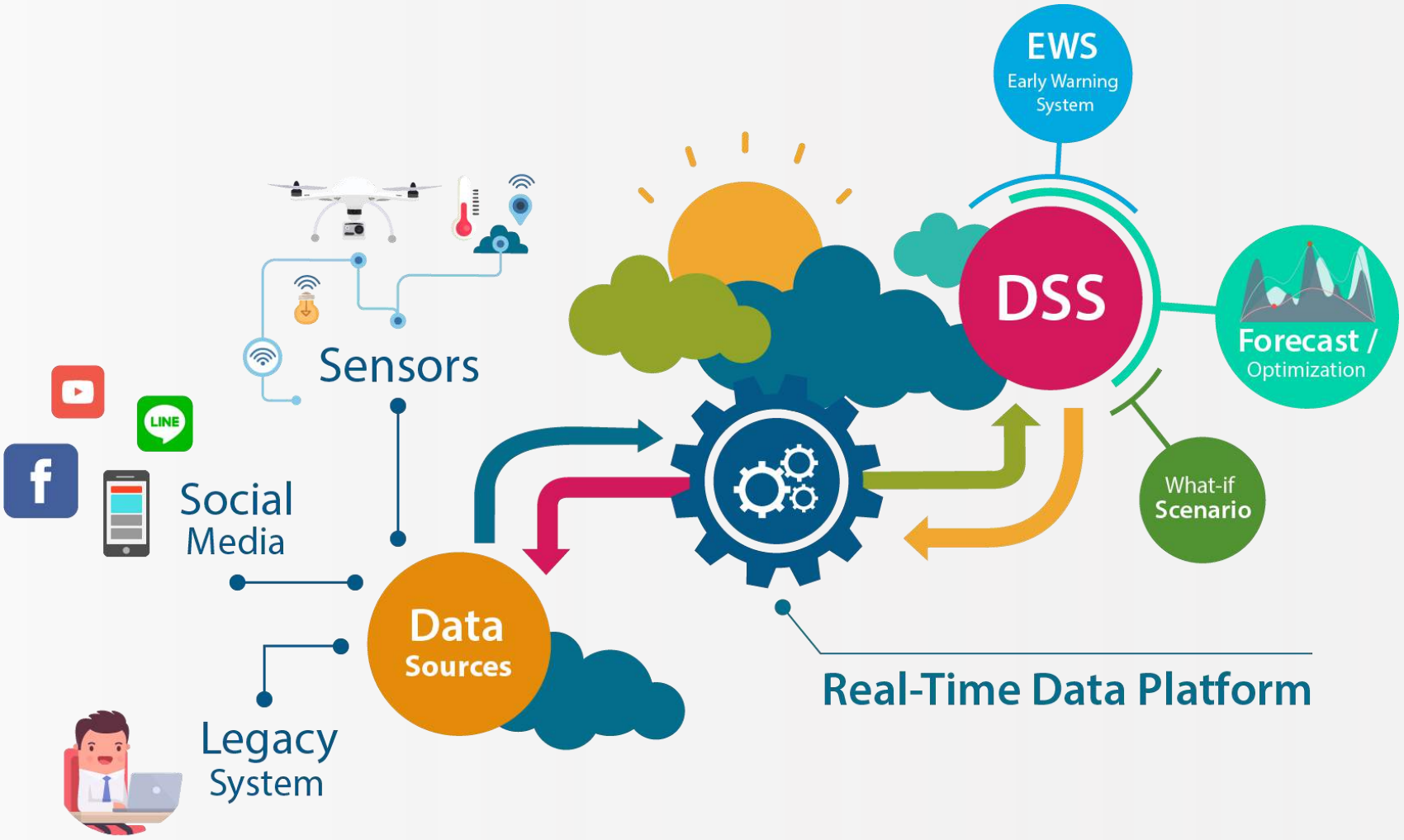
คำรณ อรุณเรื่อ

ผู้ช่วยวิจัย

ห้องปฏิบัติการวิจัยการจำลองขนาดใหญ่

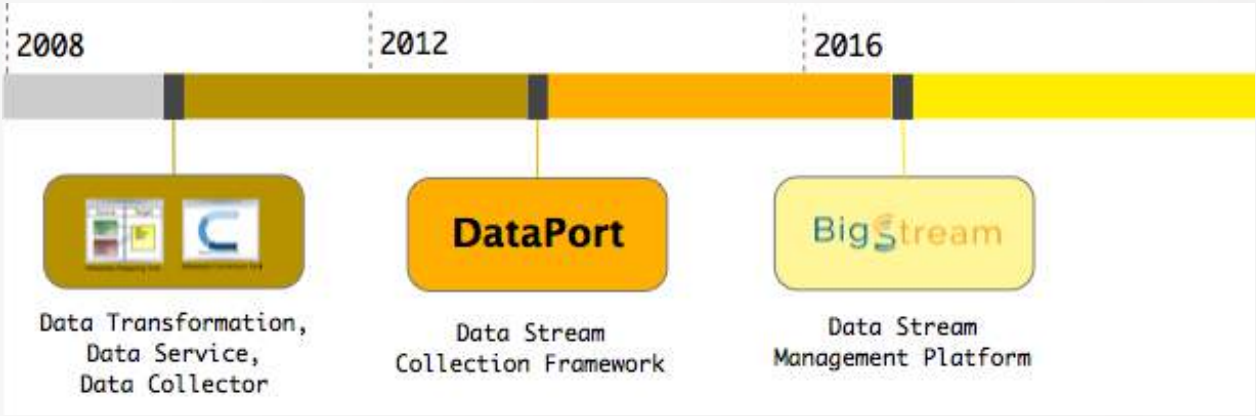
หน่วยวิจัยการวิเคราะห์ข้อมูลและการคำนวณ

ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

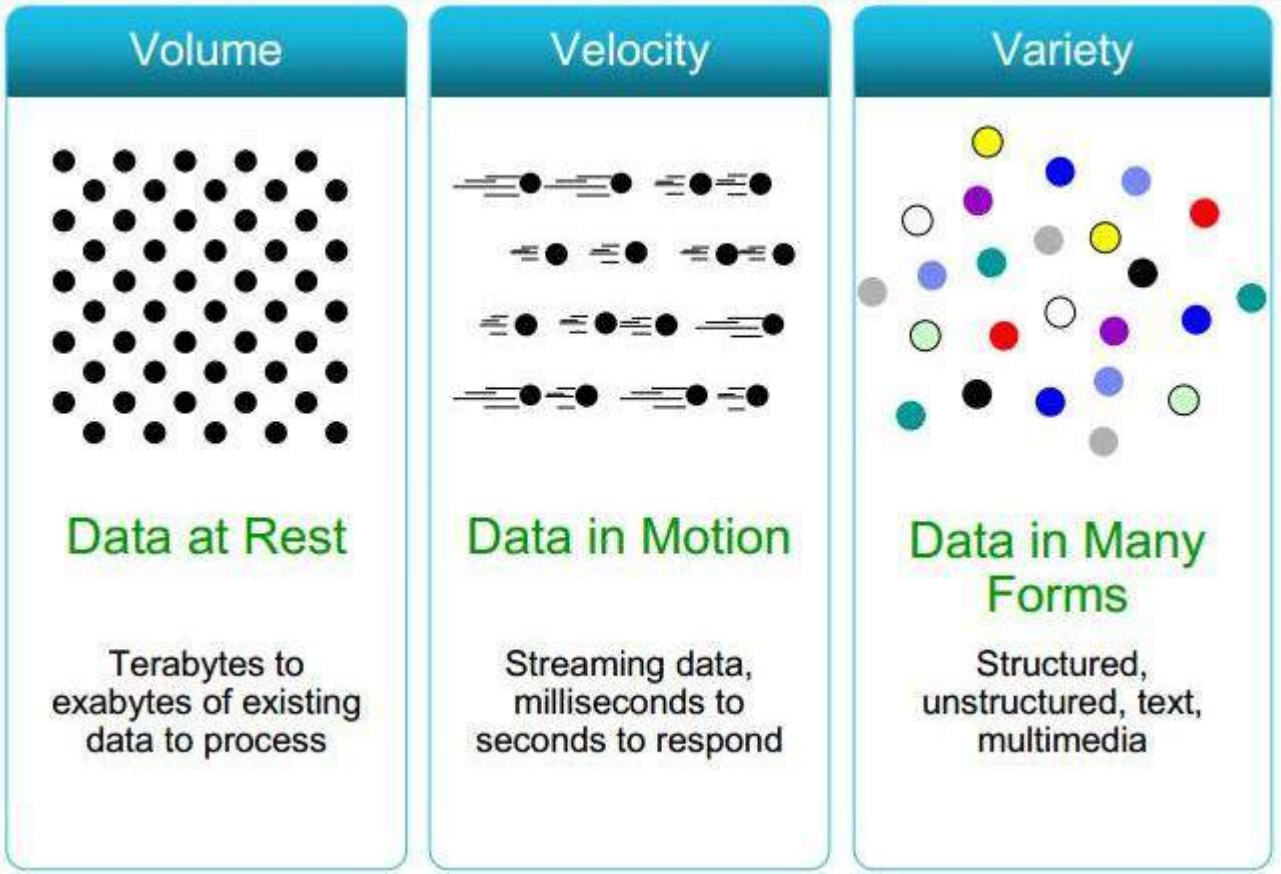

EWS
Early Warning System
DSS
Forecast / Optimization
What-if Scenario
Sensors
Social Media
LINE
Data Sources
Legacy System
Real-Time Data Platform

History of Data-Driven Platform

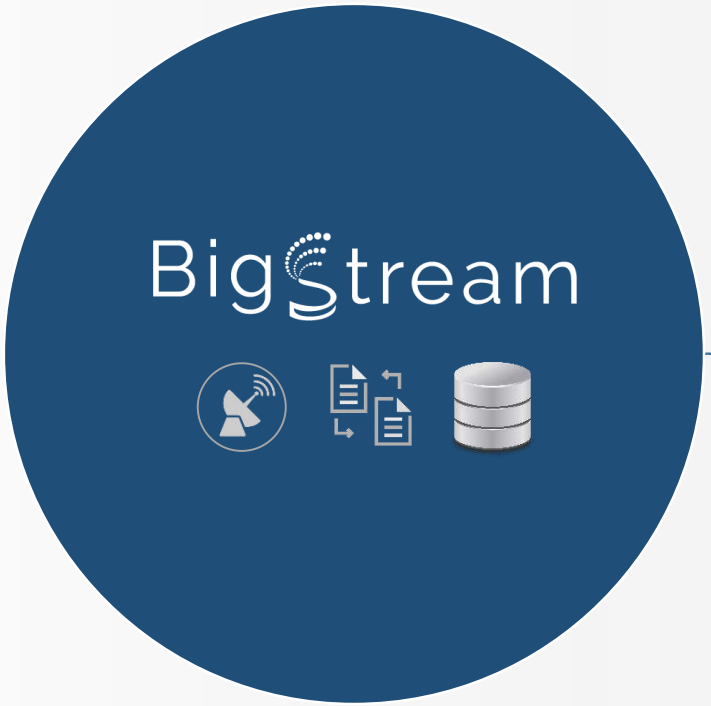
- Free style!
- Metadata Conversion Tool & Data Service
- DataPort
- BigStream



Characterization of Big Data



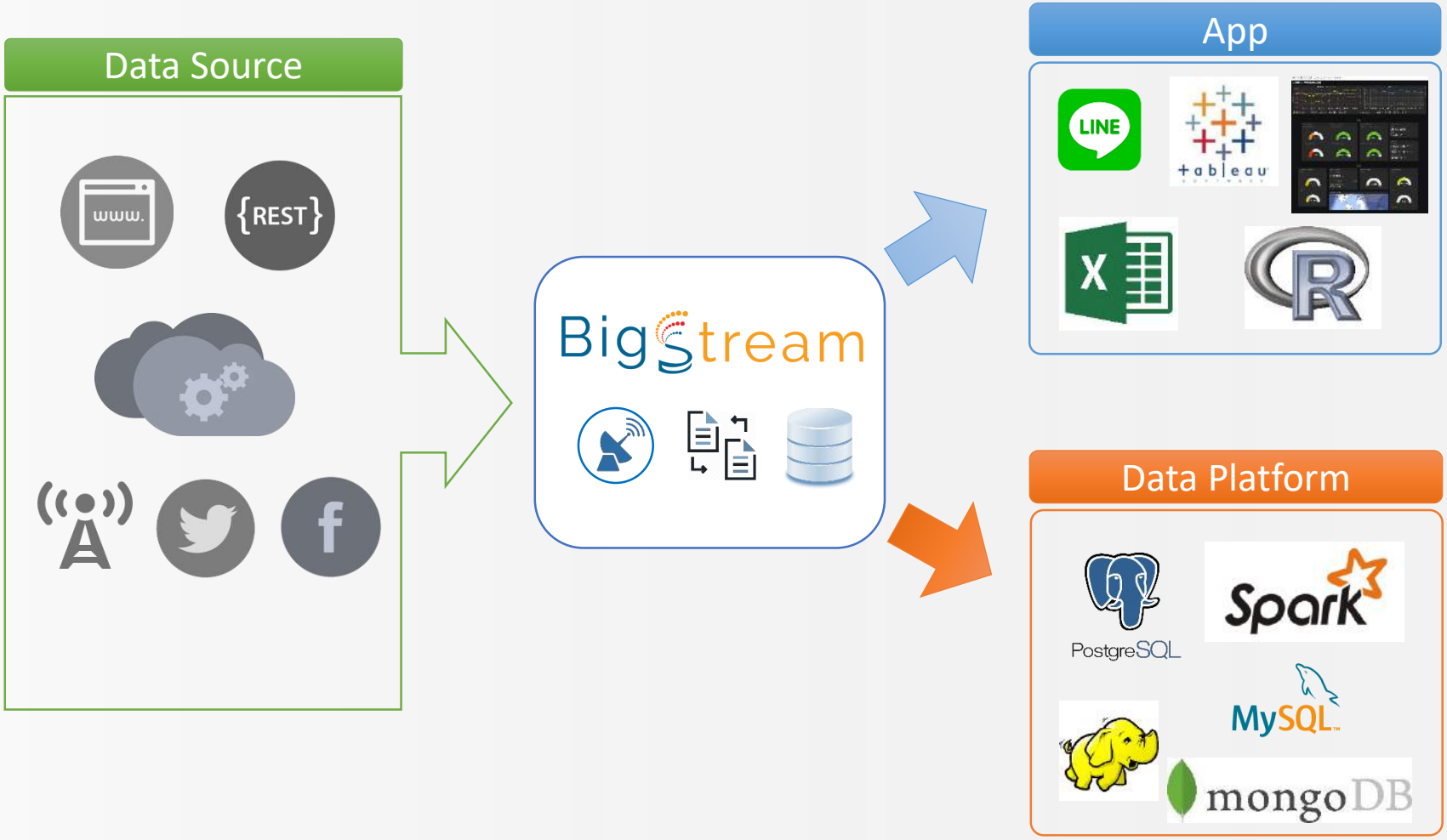
What is BigStream?



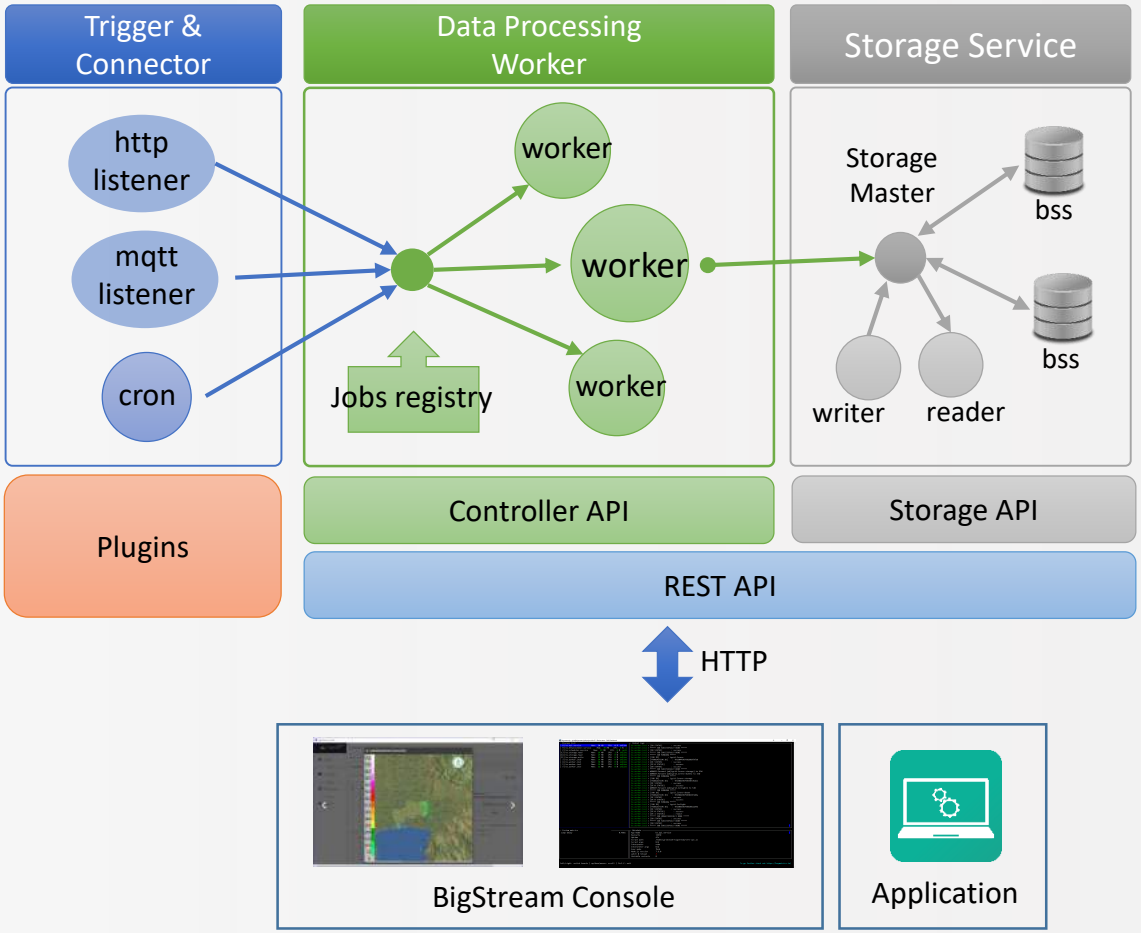
Real-time Data Management Platform

- Data Stream Collector
- Various Data Types Support
- Distributed Data Processing
- Built-in Data Storage Service
- Extendibility & Scalability
- Multi-purpose Data App

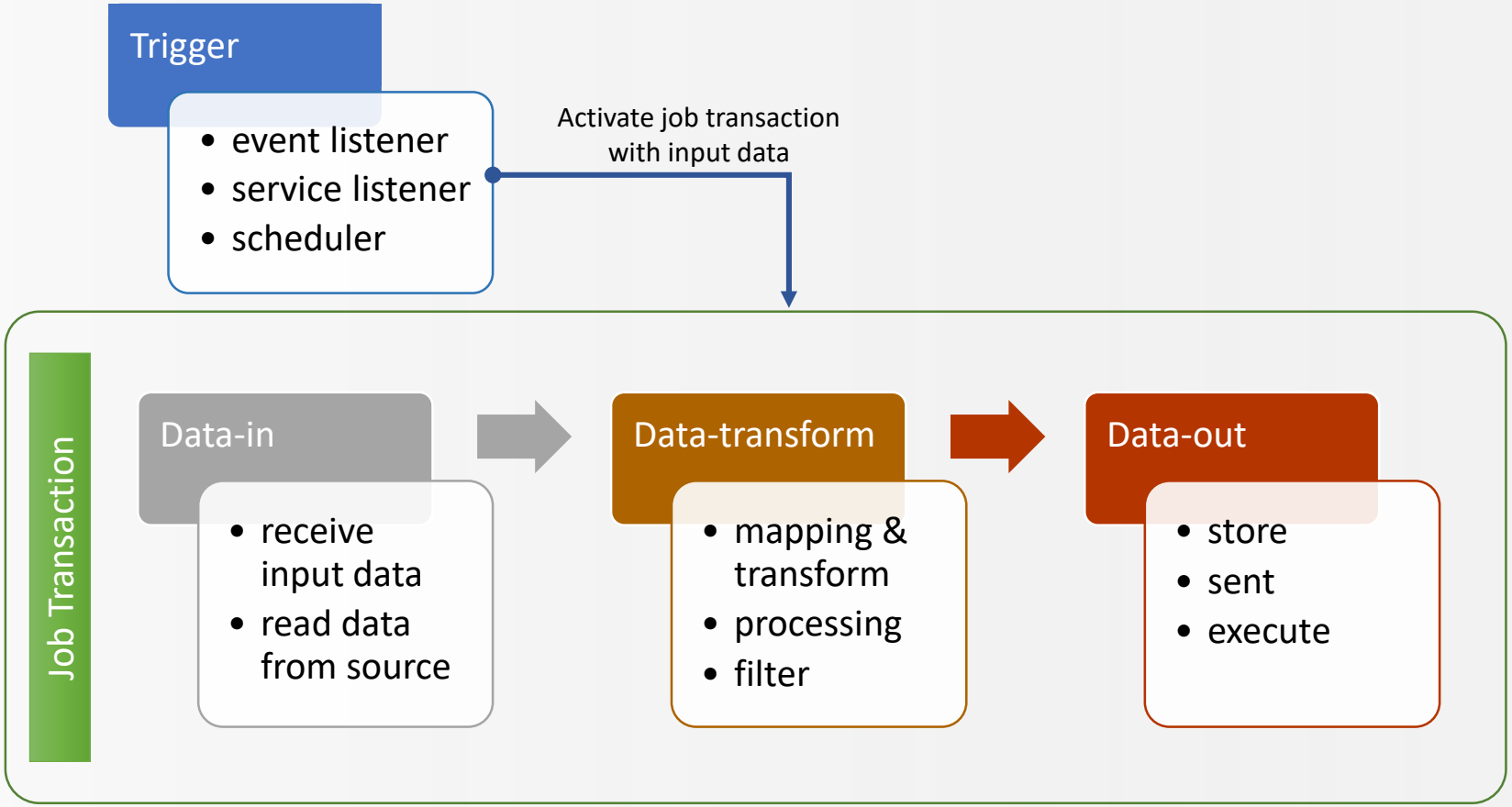
Where BigStream fits?



BigStream Architecture



BigStream Job Framework



Plugins (job modules)

Trigger	Data-in	Data-transform	Data-out
• http-listener • cron • storage-event • netpie	• input • filereader • http-request • ftp • agritronics	• map • reduce • transform • filter-condition • filter-duplicate • windows-event • windows-time • noop	• bss • dir • storage • pgsql • mysql • hdfs • dweet • line-notify • netpie • call

Job configuration

```
{
  job_id          :  unique id
  active          :  true | false
  job_timeout     :  millisecond
  trigger
    type          :  trigger type
    cmd           :  trigger command
  data_in
    type          :  plugin name
    param         :  plugin parameter
  data_transform
  [
    type          :  plugin name
    param         :  plugin parameter
  ]
  data_out
    type          :  plugin name
    param         :  plugin parameter
}
```

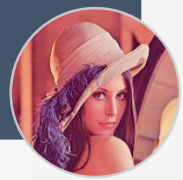
BigStream Storage

- Multi-purpose Stream Storage
- NoSQL Database (batch, real-time)
- Feature
  - large-scale data store (billions records per storage)
  - fast read/write, fast access
  - light weight & portable
  - various data types

Supported Data Types

- Image
- Wave Data
- Video Stream

Binary



- Web data
- Social network
- Encrypted text

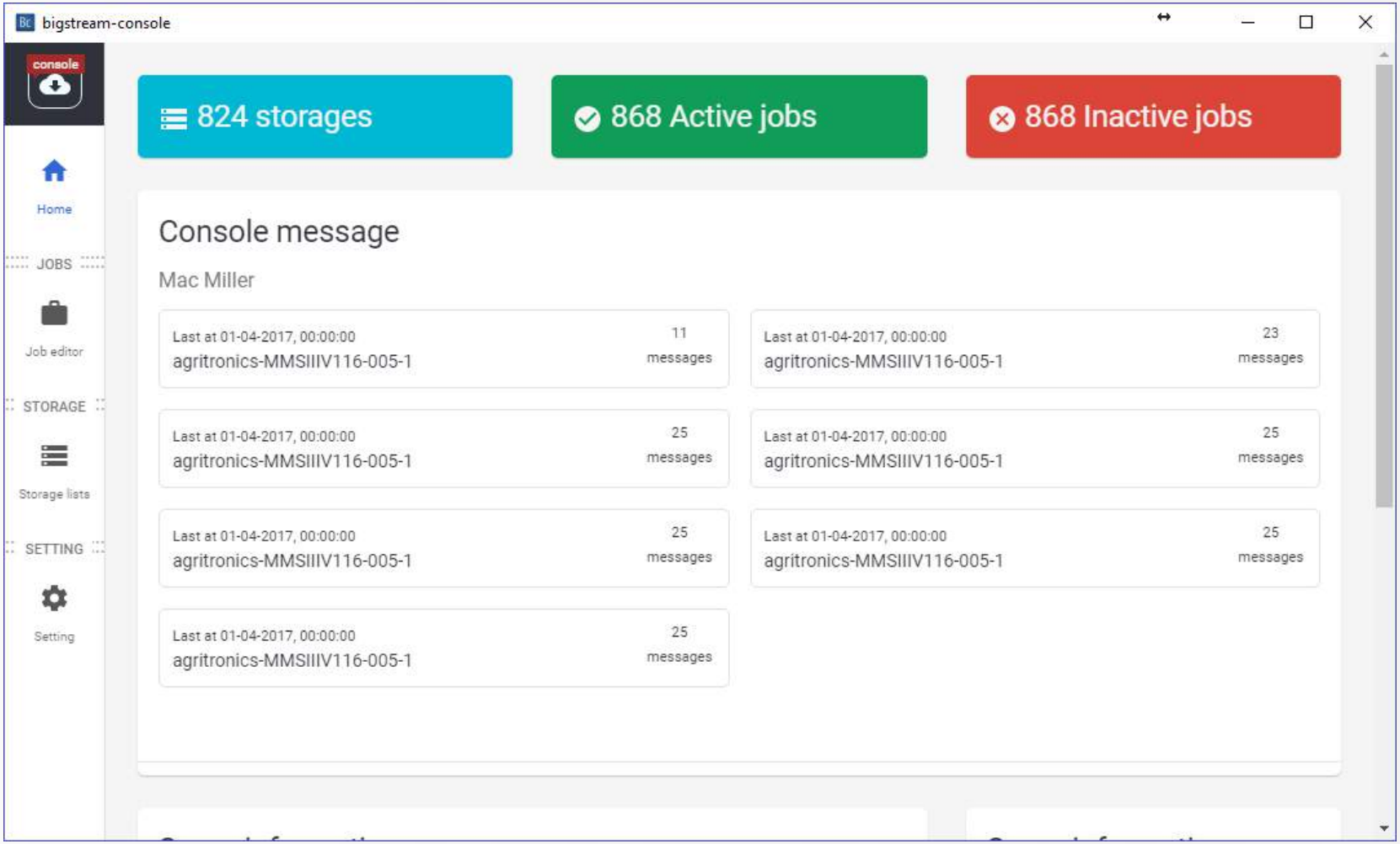
Text

- Structured Data Set
- Tabular Data
- JSON

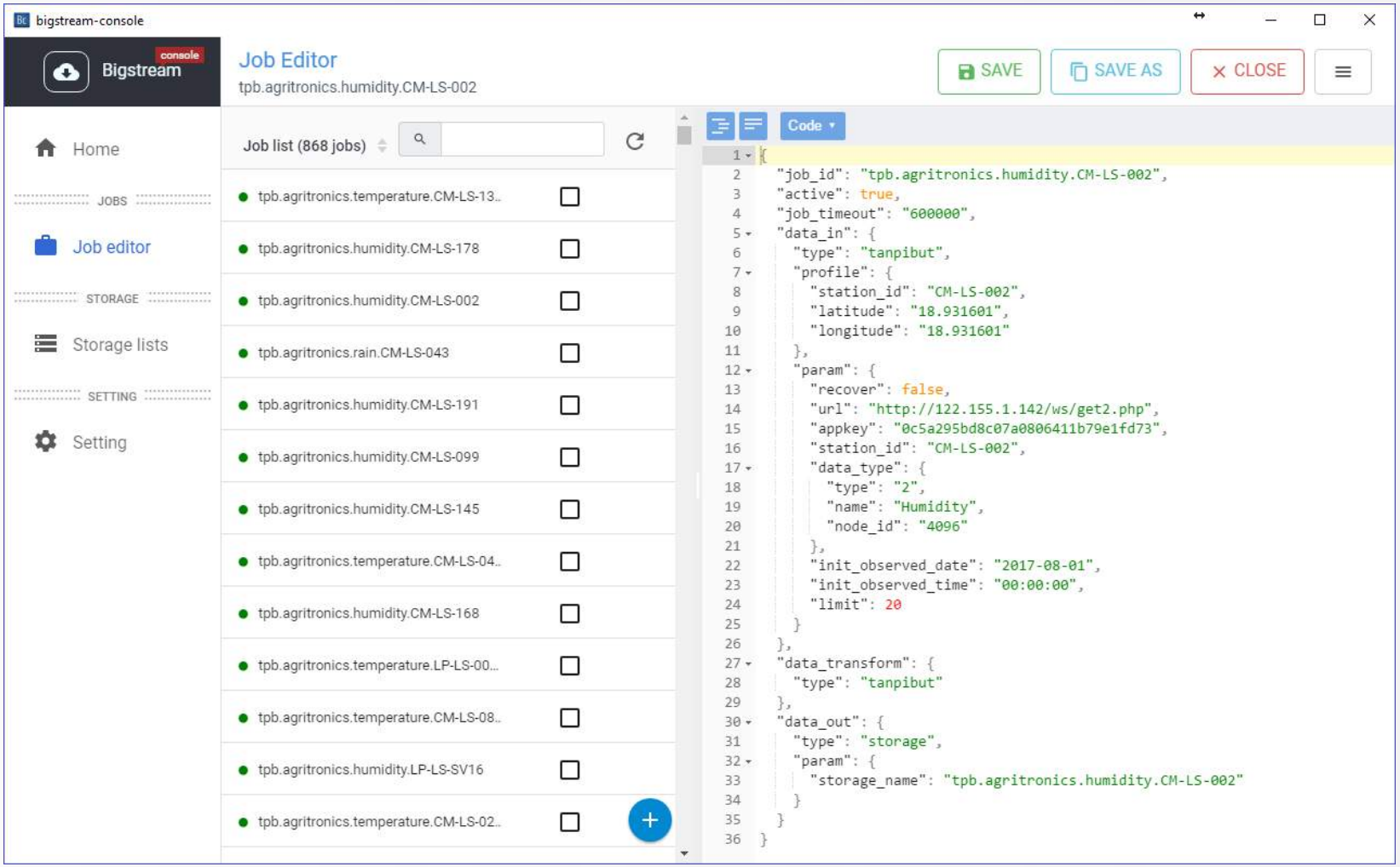
Object

3rd Party Software & Tool

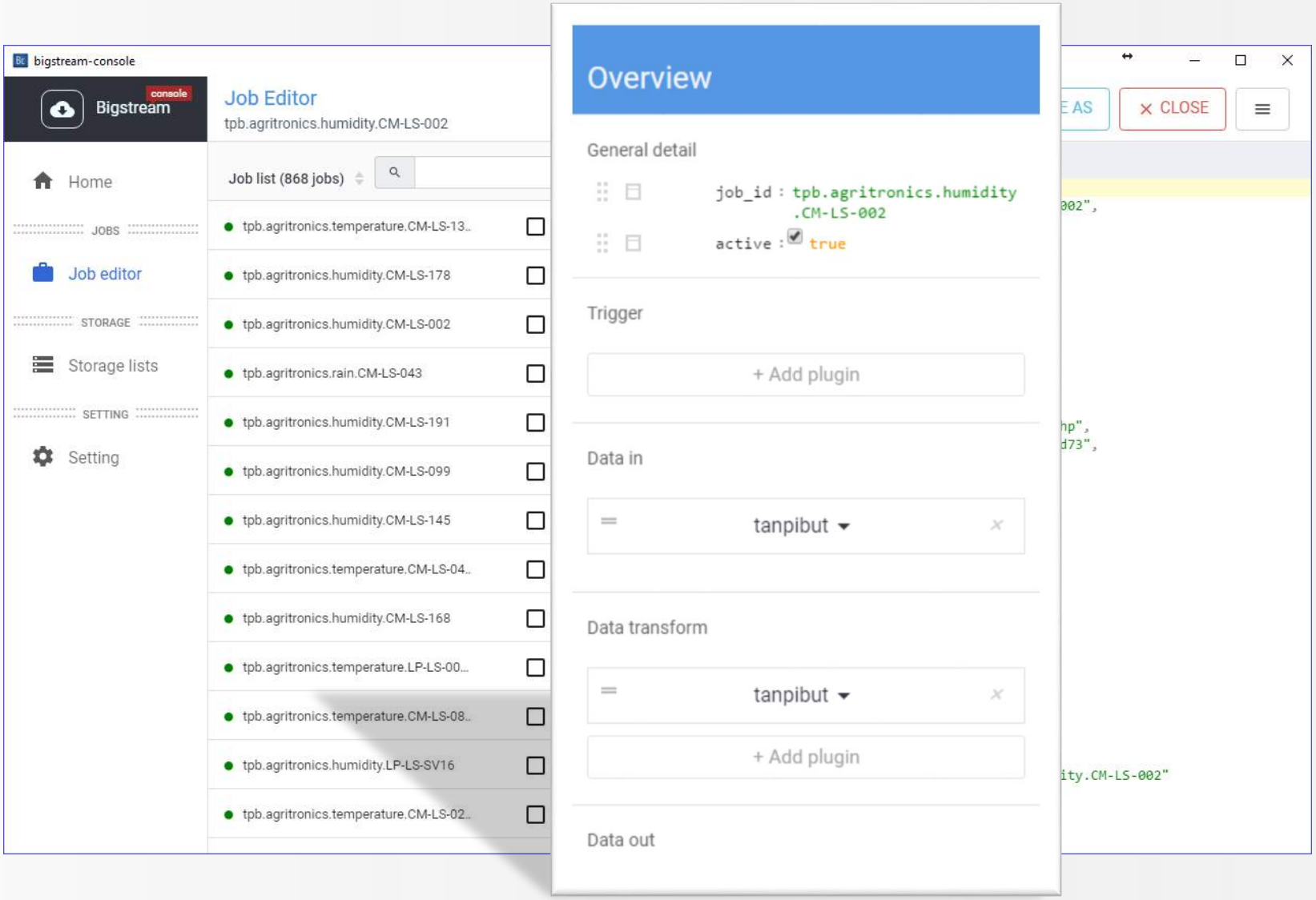
BigStream Console



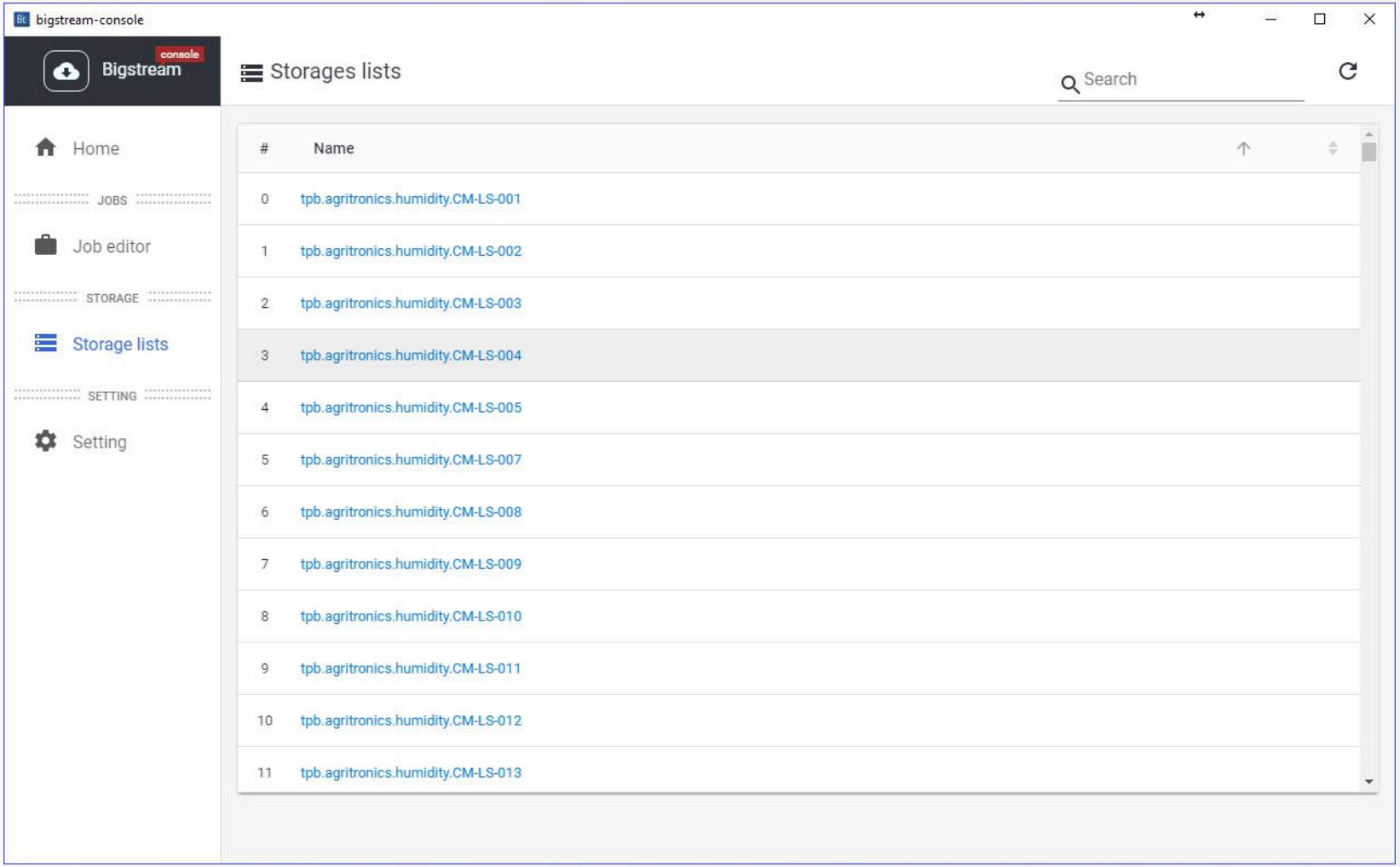
Job Editor


bigstream-console
Bigstream
console
Job Editor
tpb.agritronics.humidity.CM-LS-002
SAVE
SAVE AS
CLOSE
Home
JOBS
Job editor
STORAGE
Storage lists
SETTING
Setting
Job list (868 jobs)
tpb.agritronics.temperature.CM-LS-13..
tpb.agritronics.humidity.CM-LS-178
tpb.agritronics.humidity.CM-LS-002
tpb.agritronics.rain.CM-LS-043
tpb.agritronics.humidity.CM-LS-191
tpb.agritronics.humidity.CM-LS-099
tpb.agritronics.humidity.CM-LS-145
tpb.agritronics.temperature.CM-LS-04..
tpb.agritronics.humidity.CM-LS-168
tpb.agritronics.temperature.LP-LS-00...
tpb.agritronics.temperature.CM-LS-08..
tpb.agritronics.humidity.LP-LS-SV16
tpb.agritronics.temperature.CM-LS-02..
Code
{
"job_id": "tpb.agritronics.humidity.CM-LS-002",
"active": true,
"job_timeout": "600000",
"data_in": {
"type": "tanpibut",
"profile": {
"station_id": "CM-LS-002",
"latitude": "18.931601",
"longitude": "18.931601"
},
"param": {
"recover": false,
"url": "http://122.155.1.142/ws/get2.php",
"appkey": "0c5a295bd8c07a0806411b79e1fd73",
"station_id": "CM-LS-002",
"data_type": {
"type": "2",
"name": "Humidity",
"node_id": "4096"
},
"init_observed_date": "2017-08-01",
"init_observed_time": "00:00:00",
"limit": 20
}
},
"data_transform": {
"type": "tanpibut"
},
"data_out": {
"type": "storage",
"param": {
"storage_name": "tpb.agritronics.humidity.CM-LS-002"
}
}
}

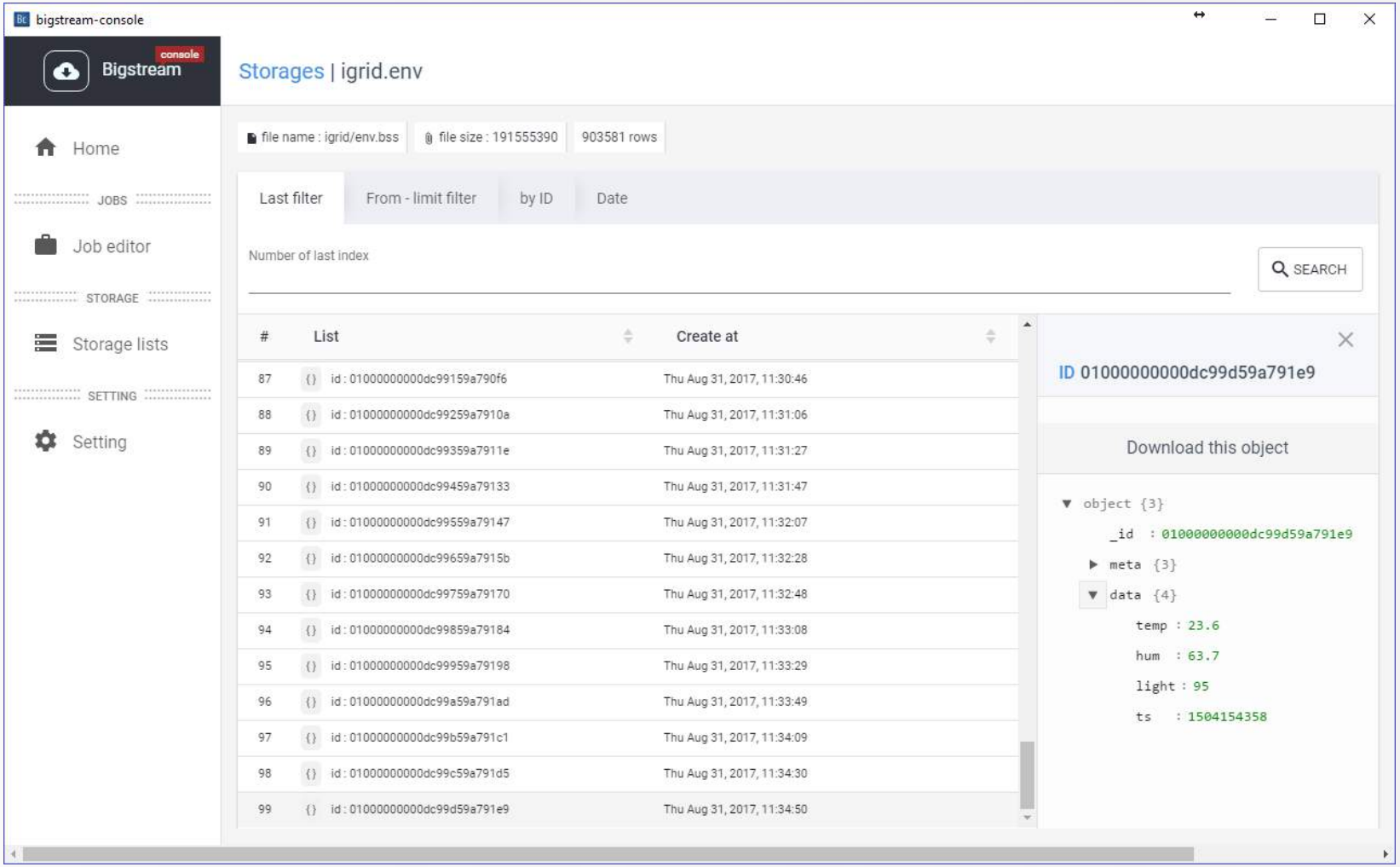
Job Editor


bigstream-console
Bigstream
console
Job Editor
tpb.agritronics.humidity.CM-LS-002
Home
JOBS
Job editor
STORAGE
Storage lists
SETTING
Setting
Job list (868 jobs)
tpb.agritronics.temperature.CM-LS-13..
tpb.agritronics.humidity.CM-LS-178
tpb.agritronics.humidity.CM-LS-002
tpb.agritronics.rain.CM-LS-043
tpb.agritronics.humidity.CM-LS-191
tpb.agritronics.humidity.CM-LS-099
tpb.agritronics.humidity.CM-LS-145
tpb.agritronics.temperature.CM-LS-04..
tpb.agritronics.humidity.CM-LS-168
tpb.agritronics.temperature.LP-LS-00...
tpb.agritronics.temperature.CM-LS-08..
tpb.agritronics.humidity.LP-LS-SV16
tpb.agritronics.temperature.CM-LS-02..
Overview
General detail
job_id : tpb.agritronics.humidity .CM-LS-002
active : true
Trigger
+ Add plugin
Data in
tanpibut
Data transform
tanpibut
+ Add plugin
Data out
CLOSE

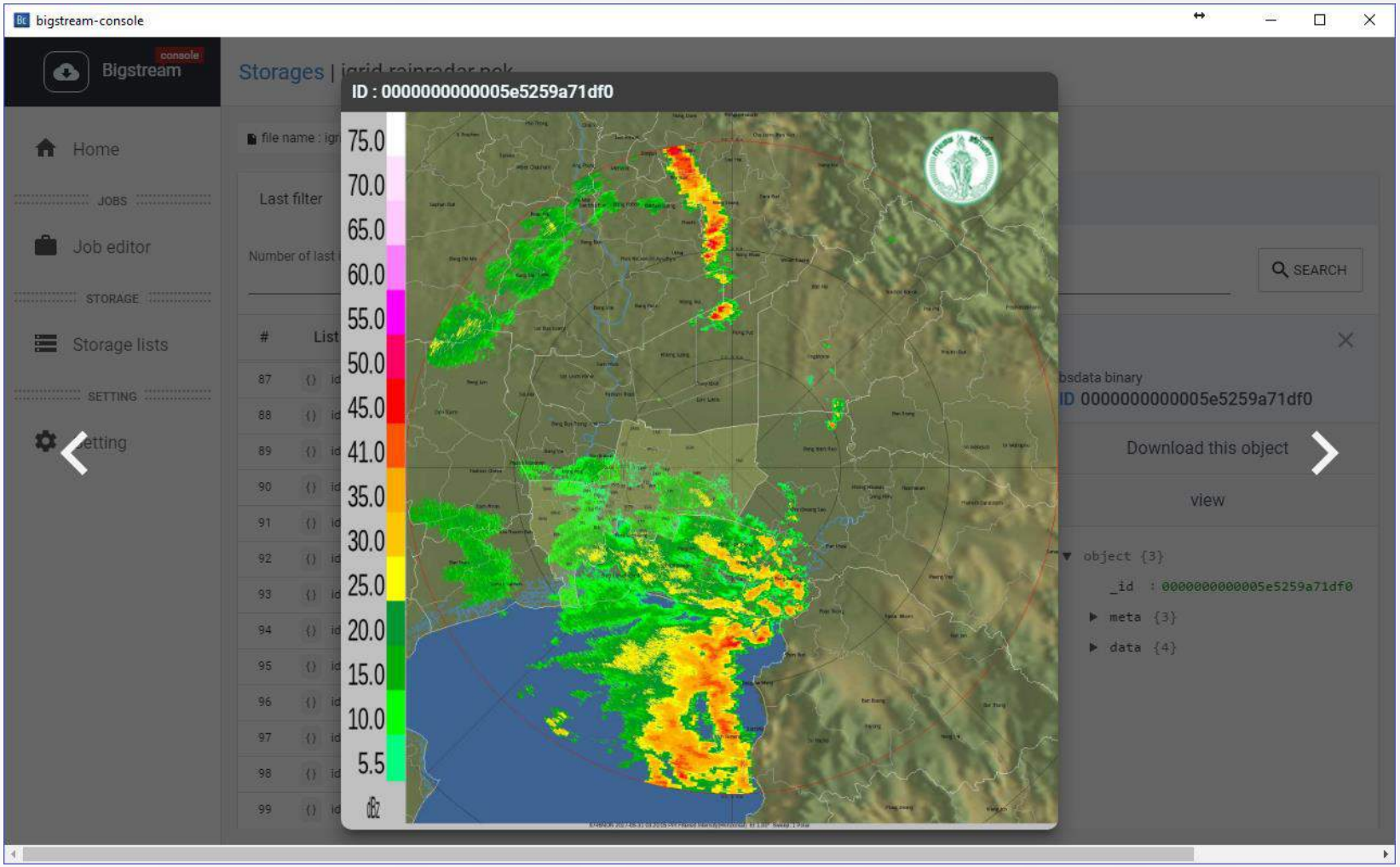
Storage Viewer



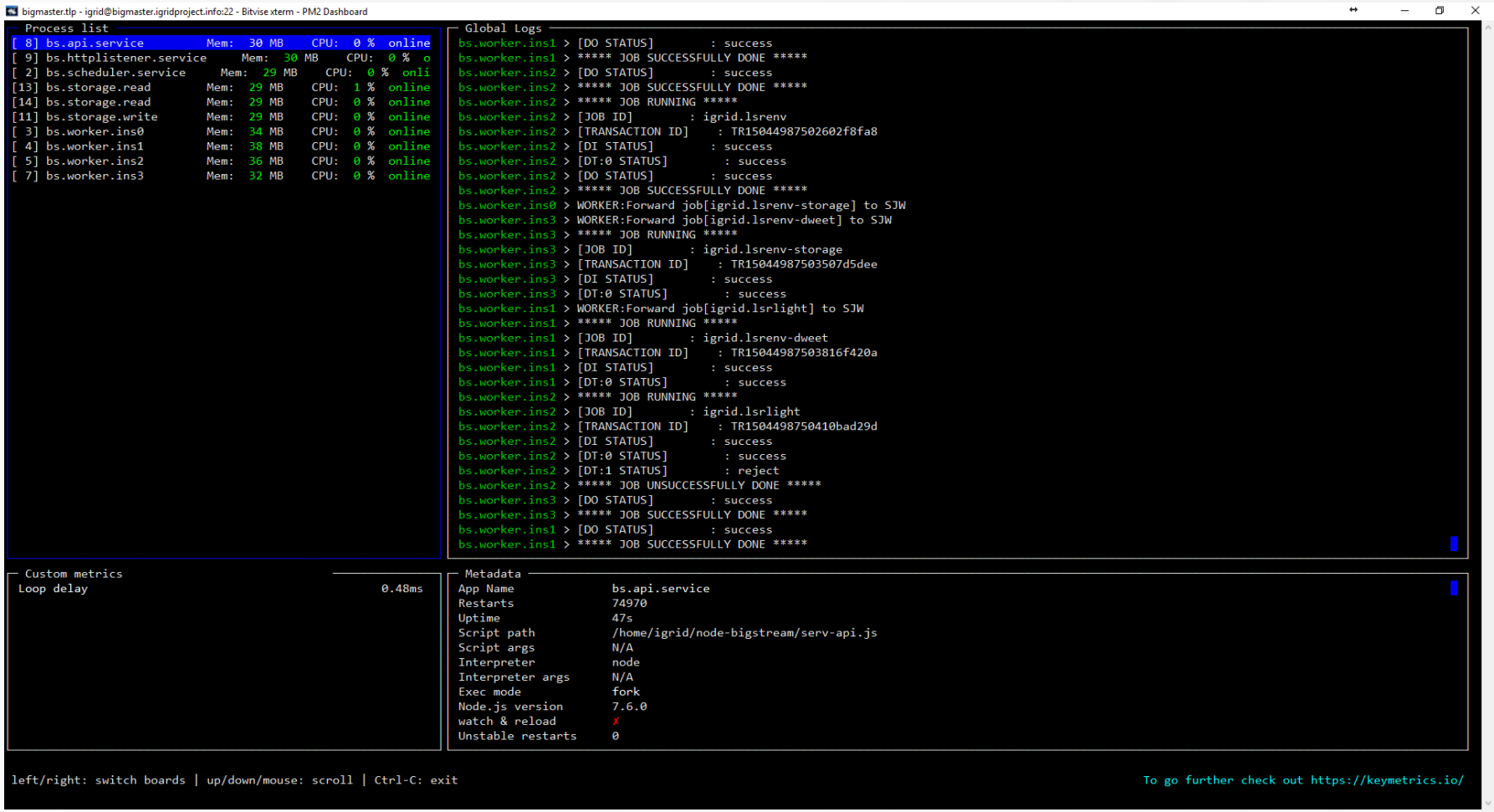
Storage Viewer



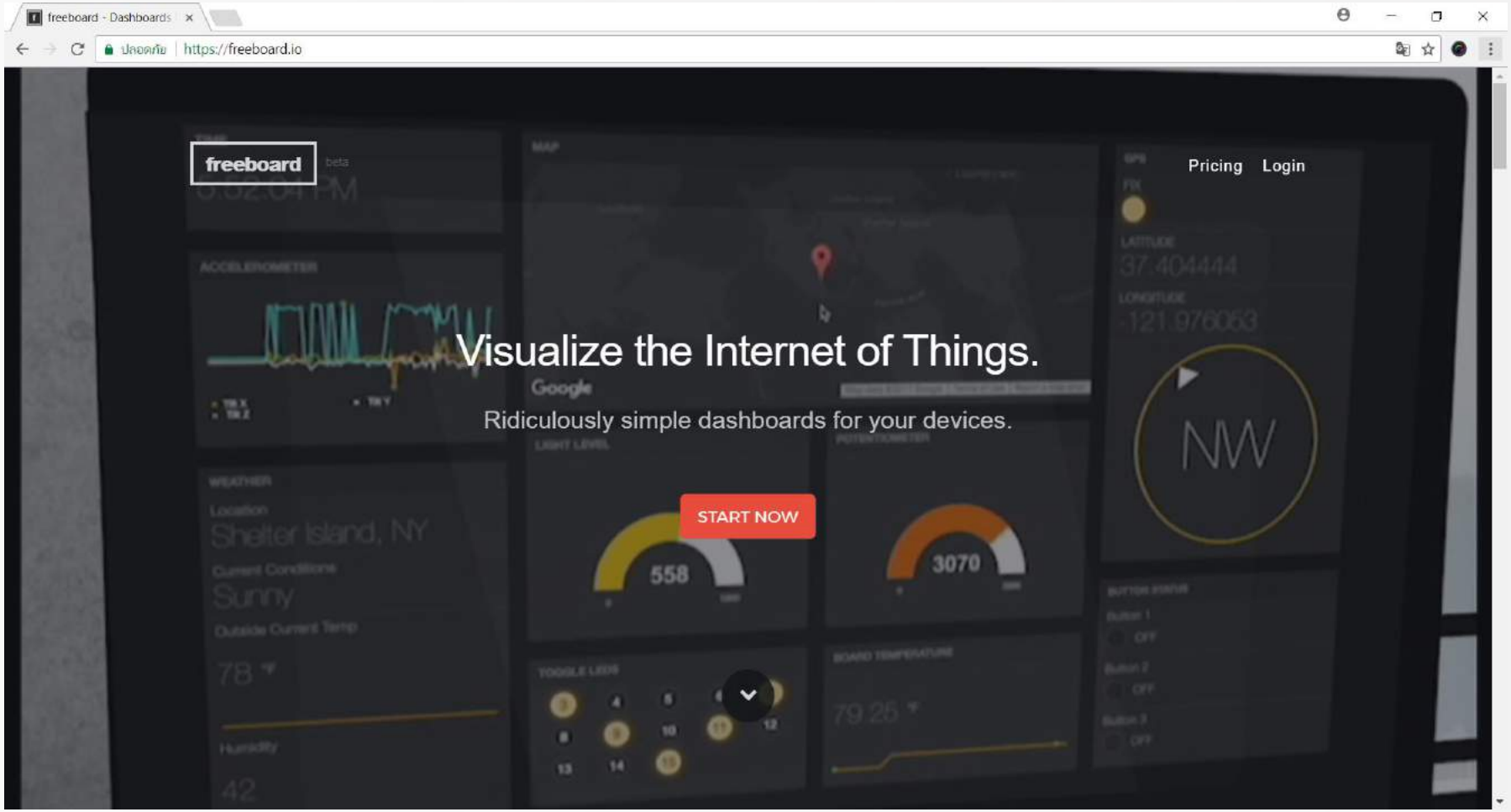
Storage Viewer



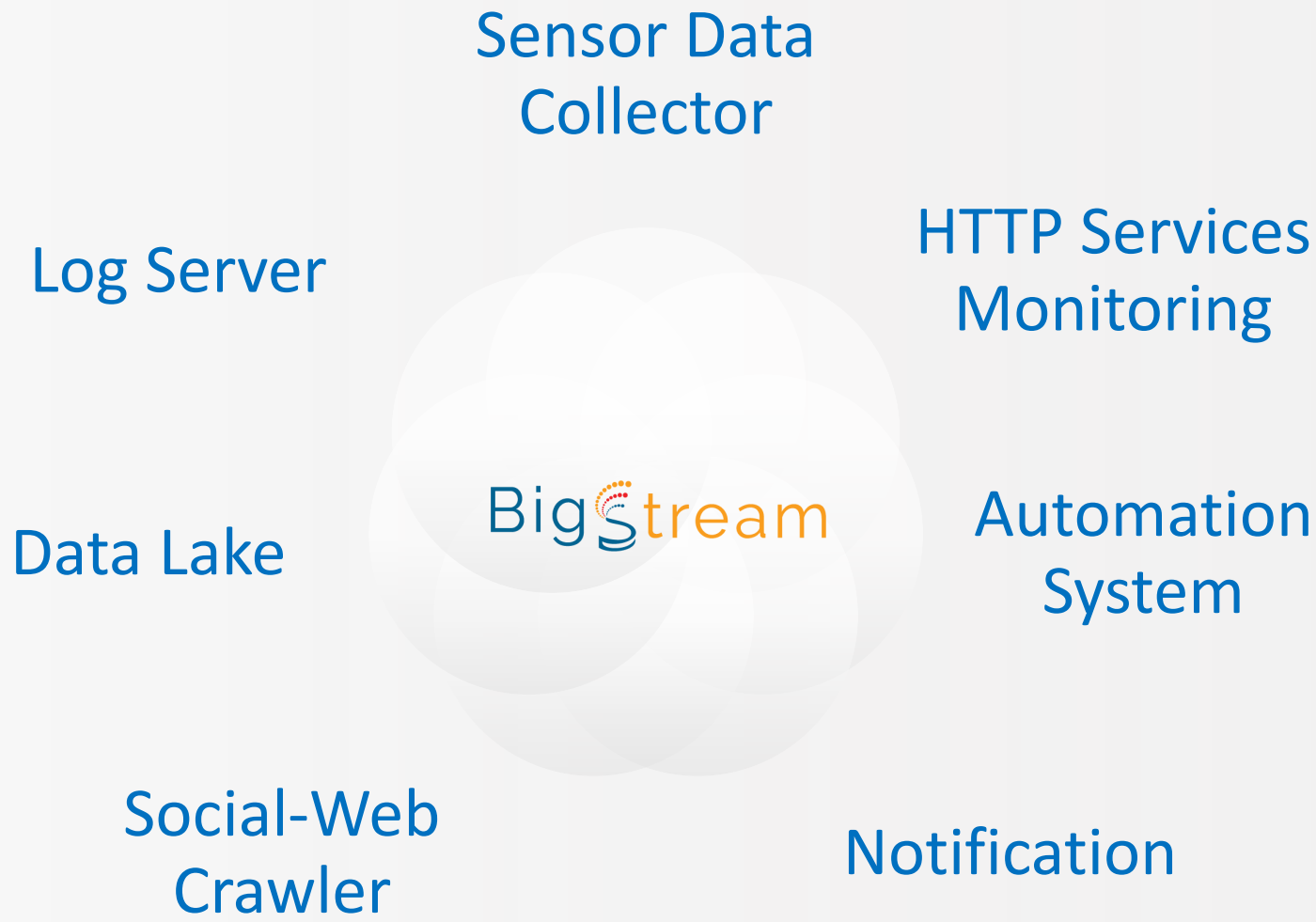
Monitoring Console



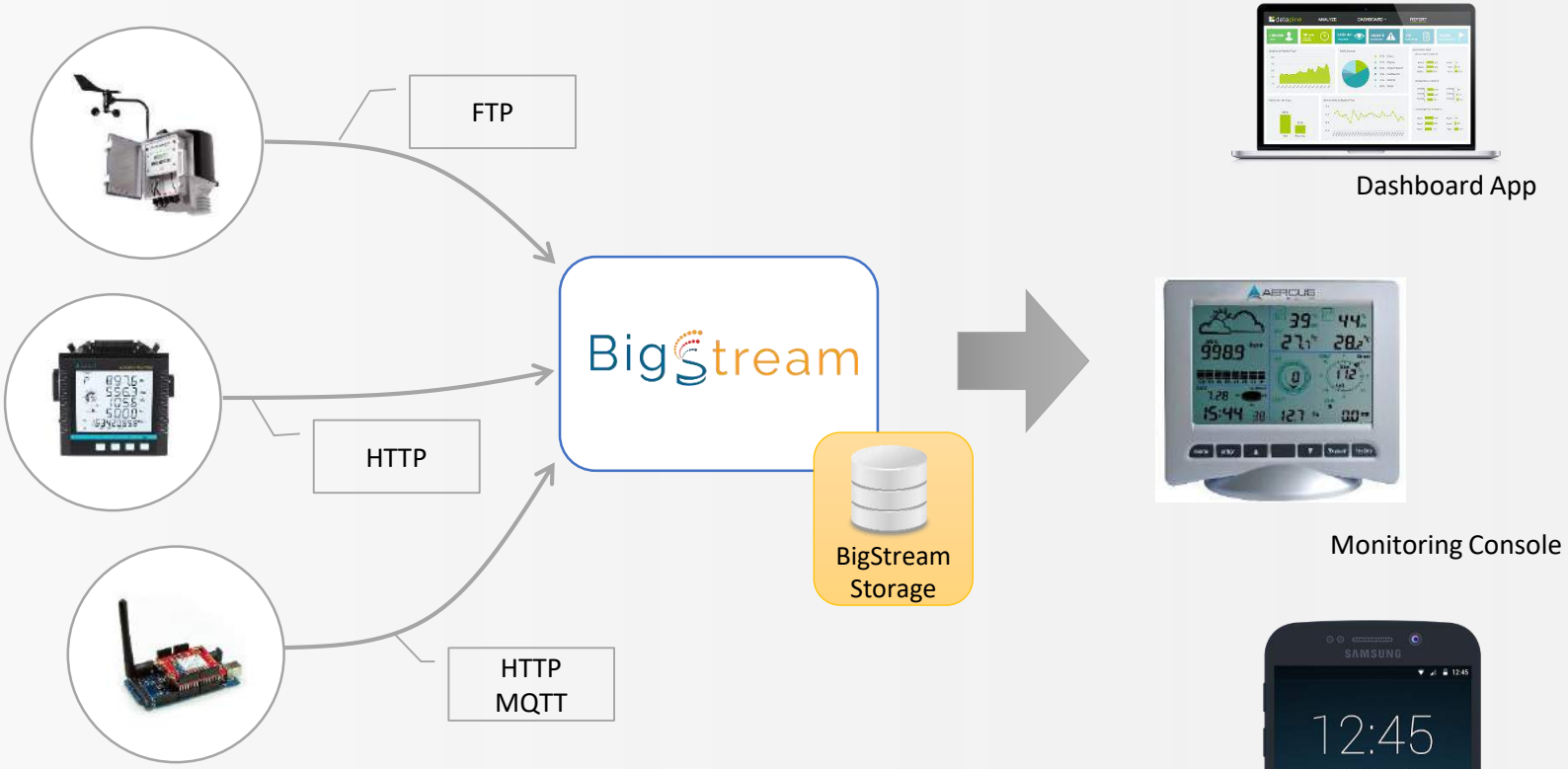
Freeboard.io



Potential Applications of BigStream



Sensor Data Collector



Sensor Data Collector

ESP8266 web client

ARDUINO SKETCH

construct HTTP request to be sent

```
String cmd;
cmd = "GET /appkey?data=value&...
cmd += temp;
cmd += "\r\n";
```

send the length of the command

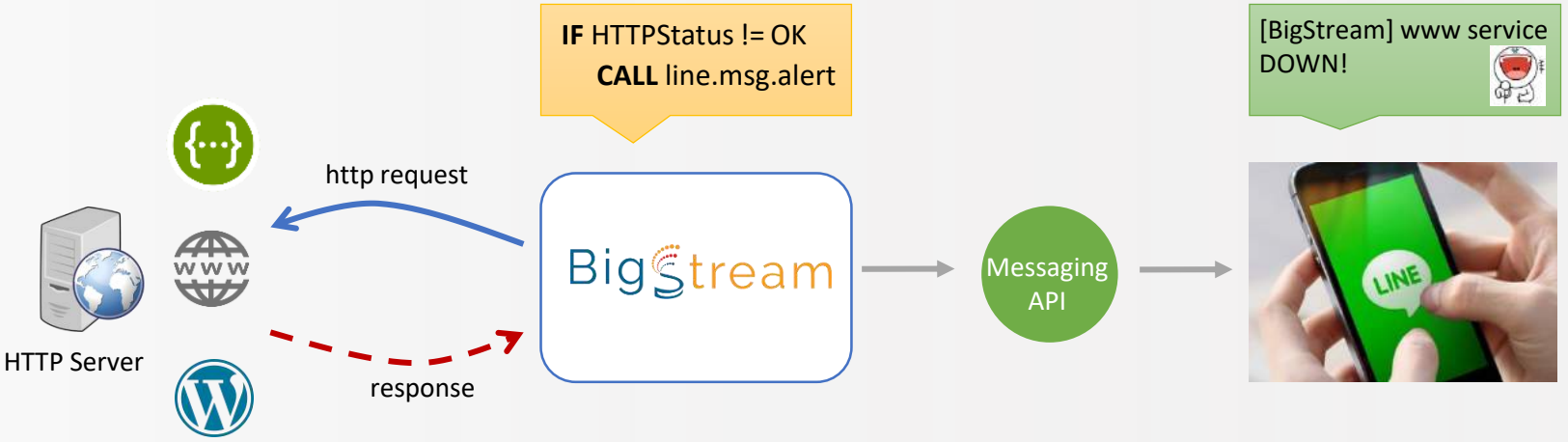
```
espSerial.print("AT+CIPSEND=");
espSerial.println(cmd.length());
```

wait for ">" with `espSerial.find(">")`

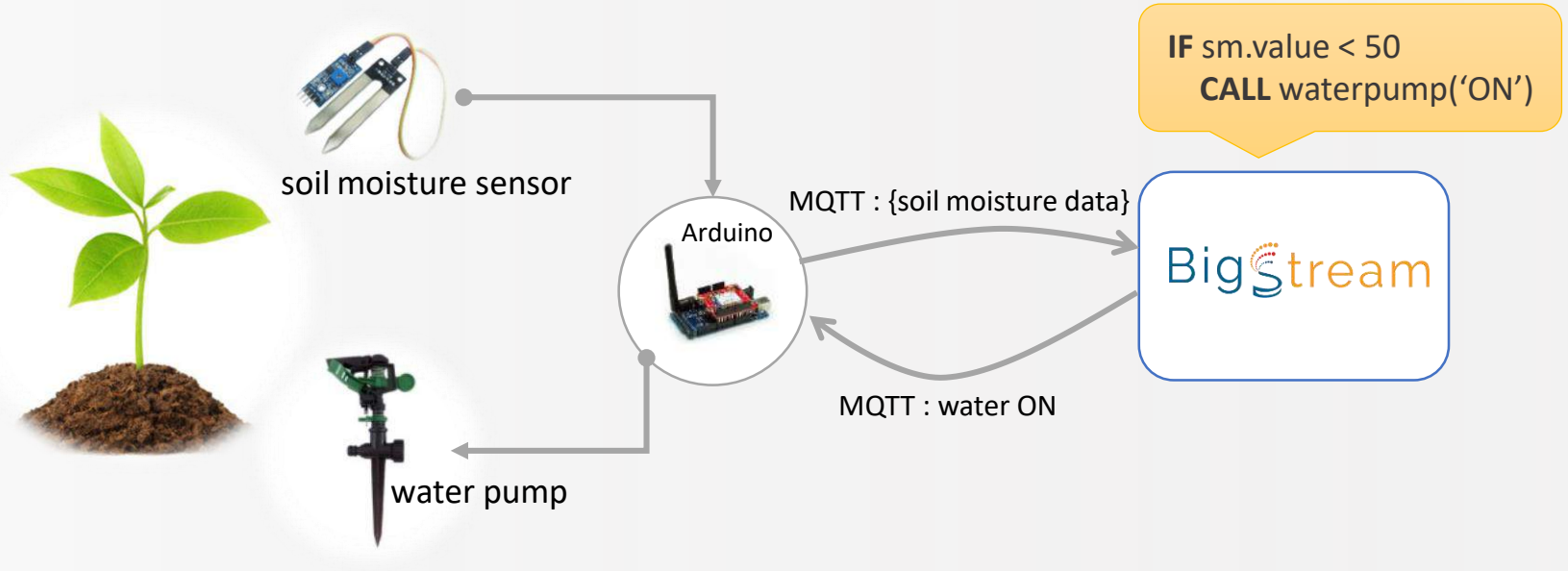
```
espSerial.println(cmd);
```

decode the reply with `espSerial.find("SEND OK")`

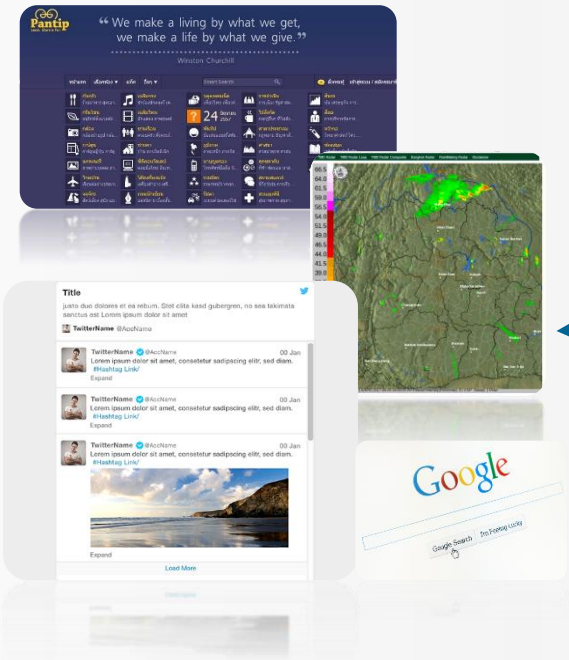
HTTP Service Monitoring



Automation System



Social Network - Web Crawler



feed

schedule job every 30 minute

BigStream

Storage API

Image Processing, DL ,ML App

Text Mining, NLP App

Data Analytics Platform

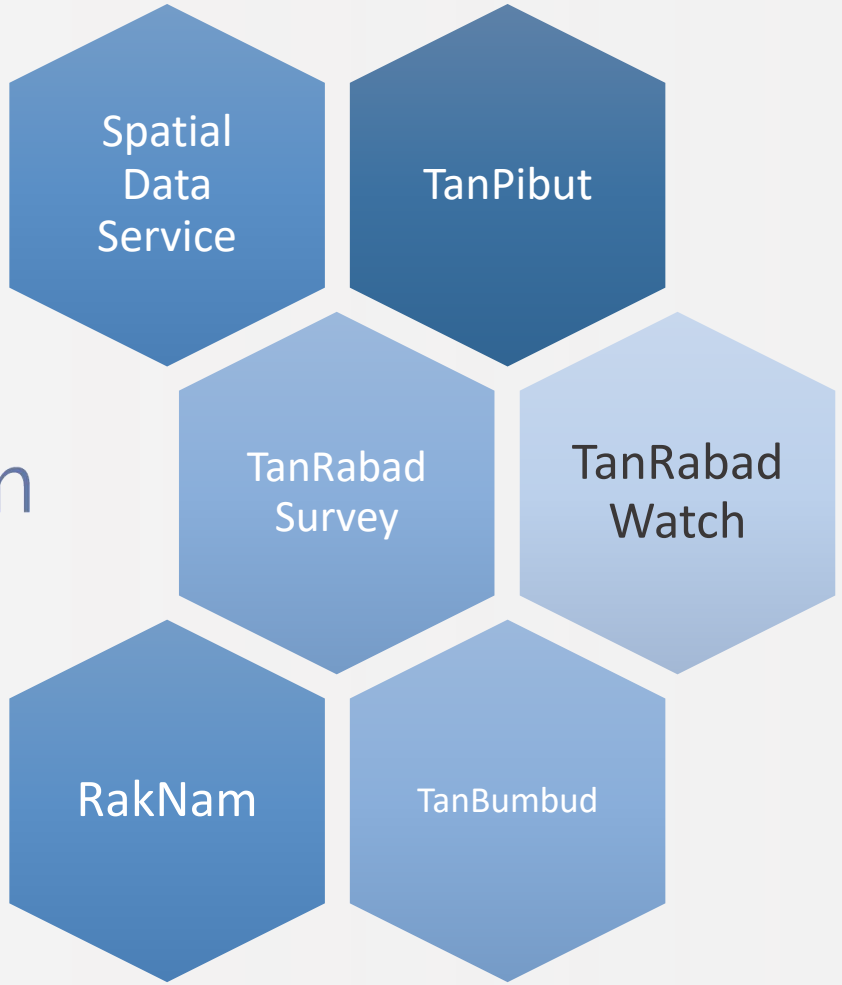
hadoop

Spark

KitWai

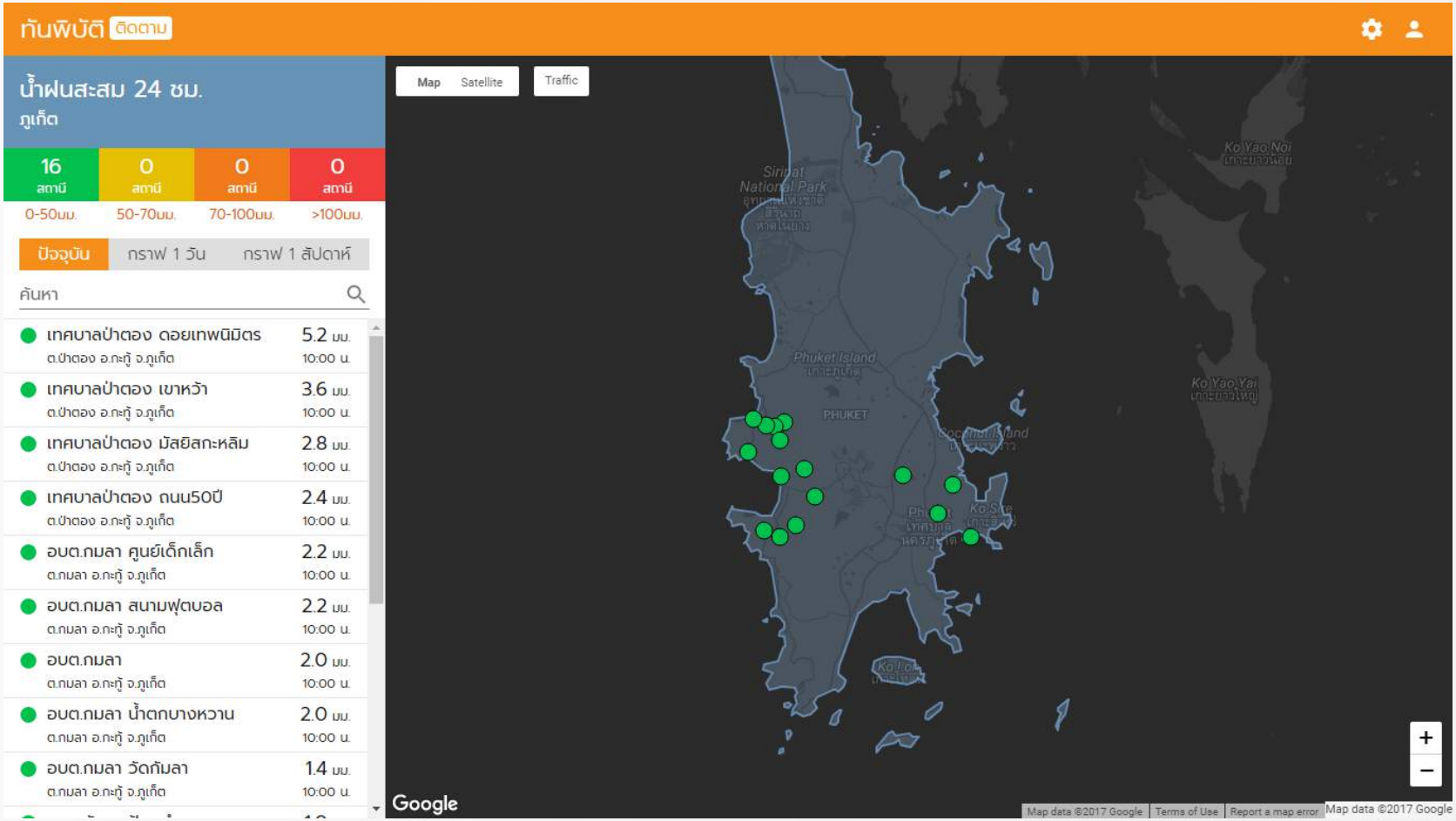

LOCAL TIME
2:49:36 PM
BIGSTREAM MONITOR
Agritronics Data
110898 objects
Gistda-water
255919 objects
RADAR
KLONG LUANG
Temp
33 C
Humidity
63 %
2.6
Wind
LSR ENVIRONMENT
Room Light
ON [brightness = 95%]
Temp
23.6 C
Humidity
67.3
%
0
100
LAAN MAN PARKING
Total
299 cars
68
%
0
100
TANRABAD SURVEY
Service Status
UP [8/31/2017, 2:49:33 PM]
Total
42034 surveys
Last Survey
Surveyor : jidapa
Time : 8/31/2017, 11:08:49 AM
Google

Projects with BigStream

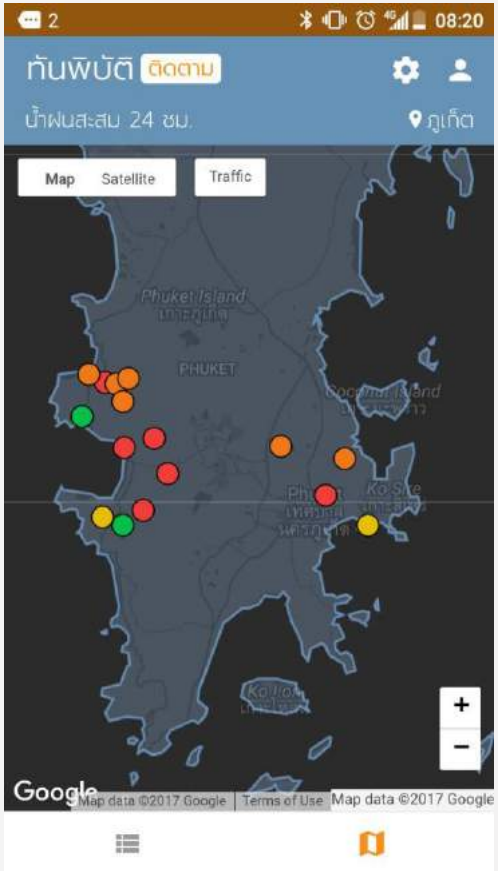


TanPibut (ทันพิบัติ)

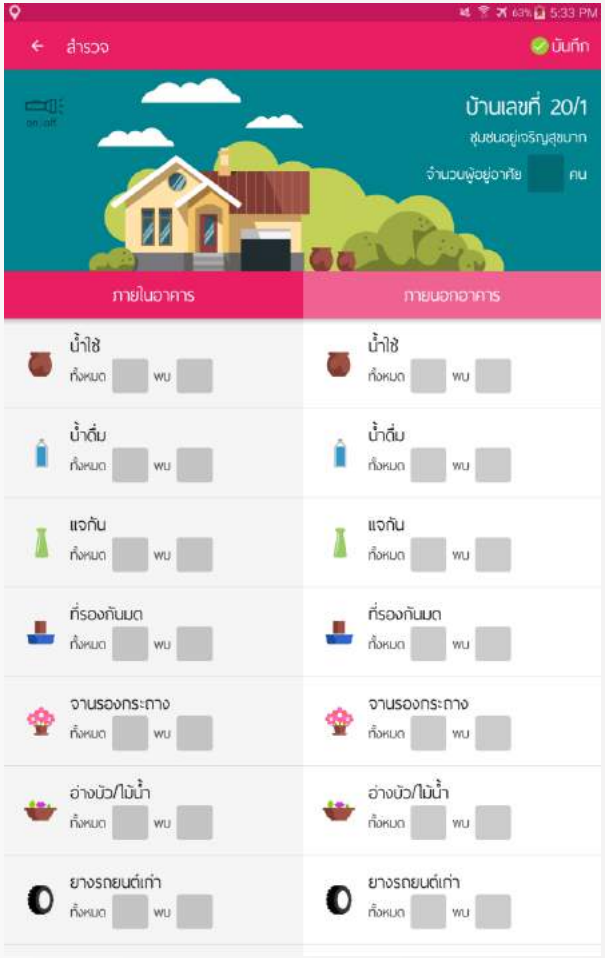
TanPibut (ทันพิบัติ)

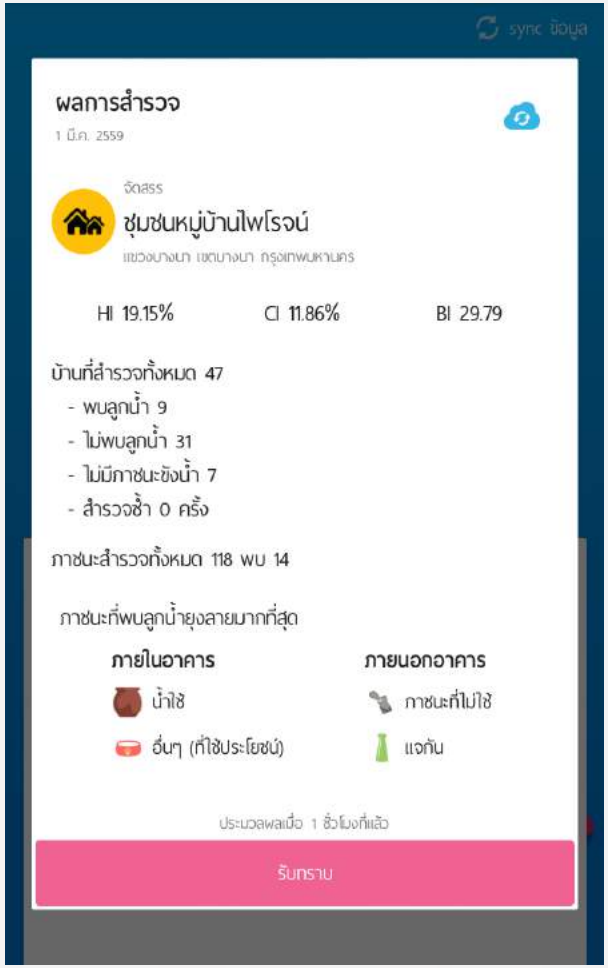


TanPibut (ทันพิบัติ)

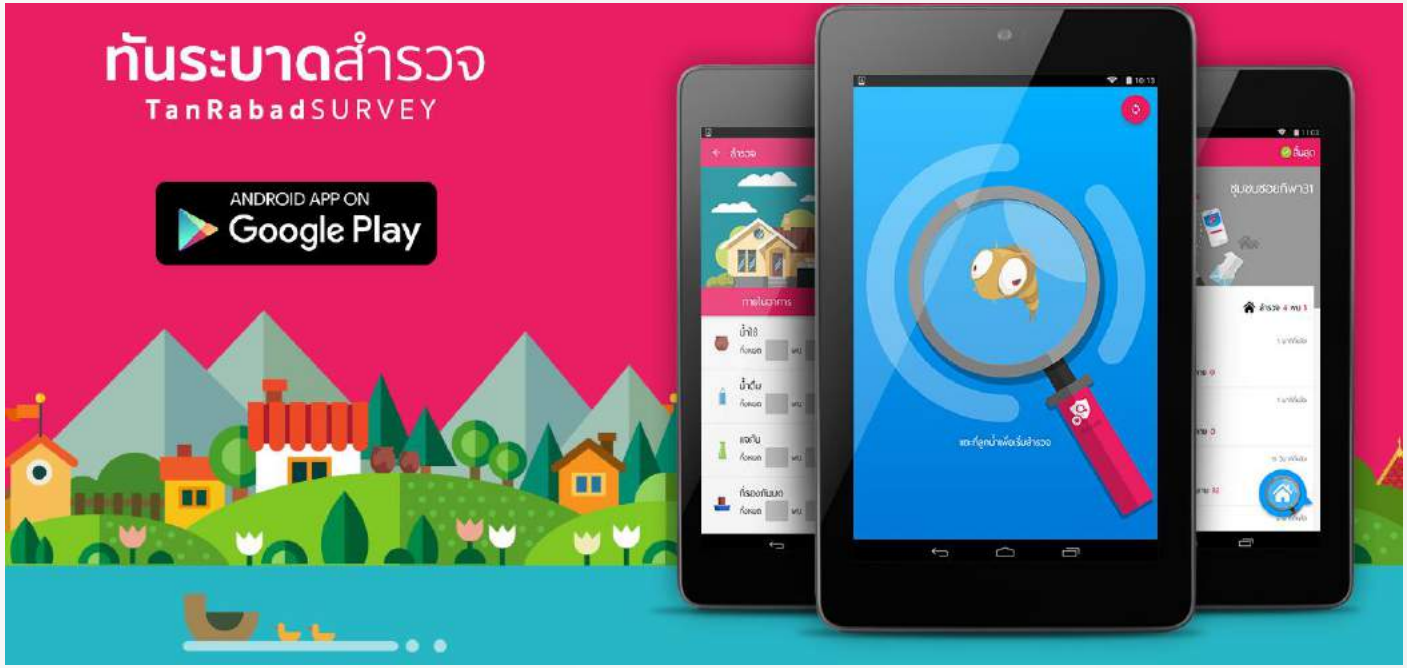


TanRabad (ทันระบาด)





TanRabad-SURVEY

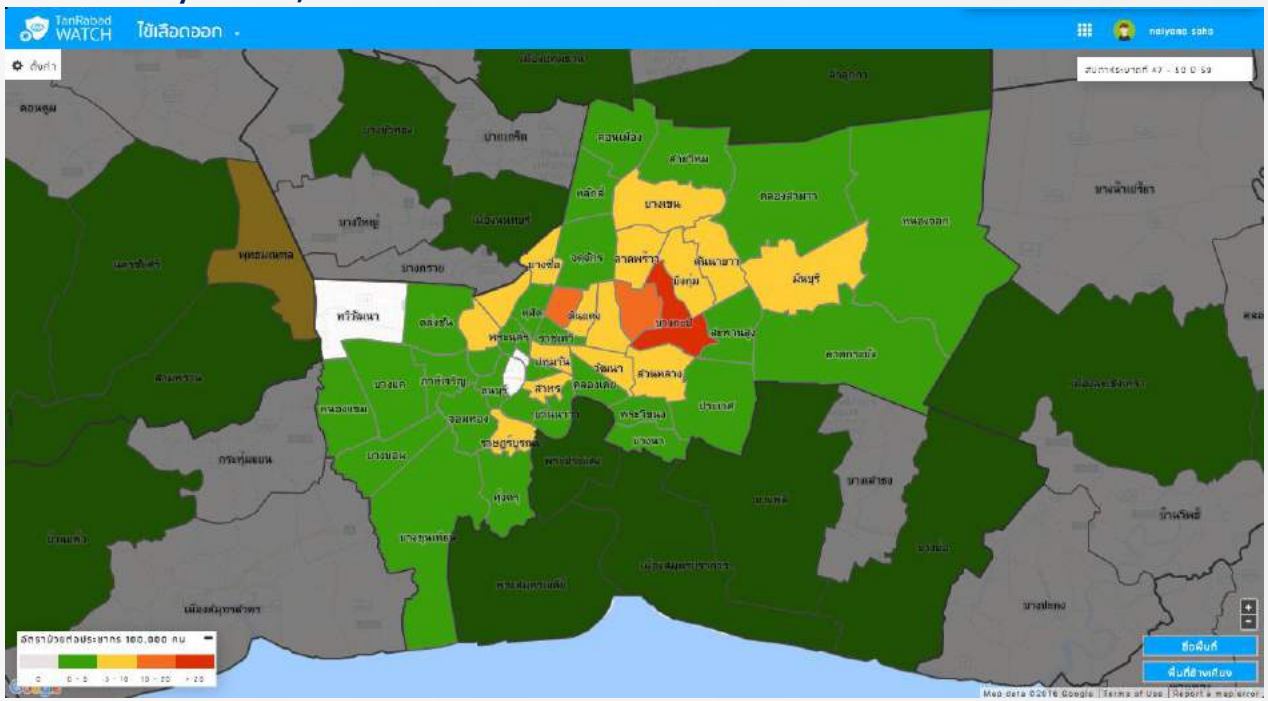


- collect larval survey data via on-line & off-line modes
- specify geo-location of places with GPS via Google Maps
- automatically transfer larval survey data to cloud
- real-time report entomology indices once the survey is completed

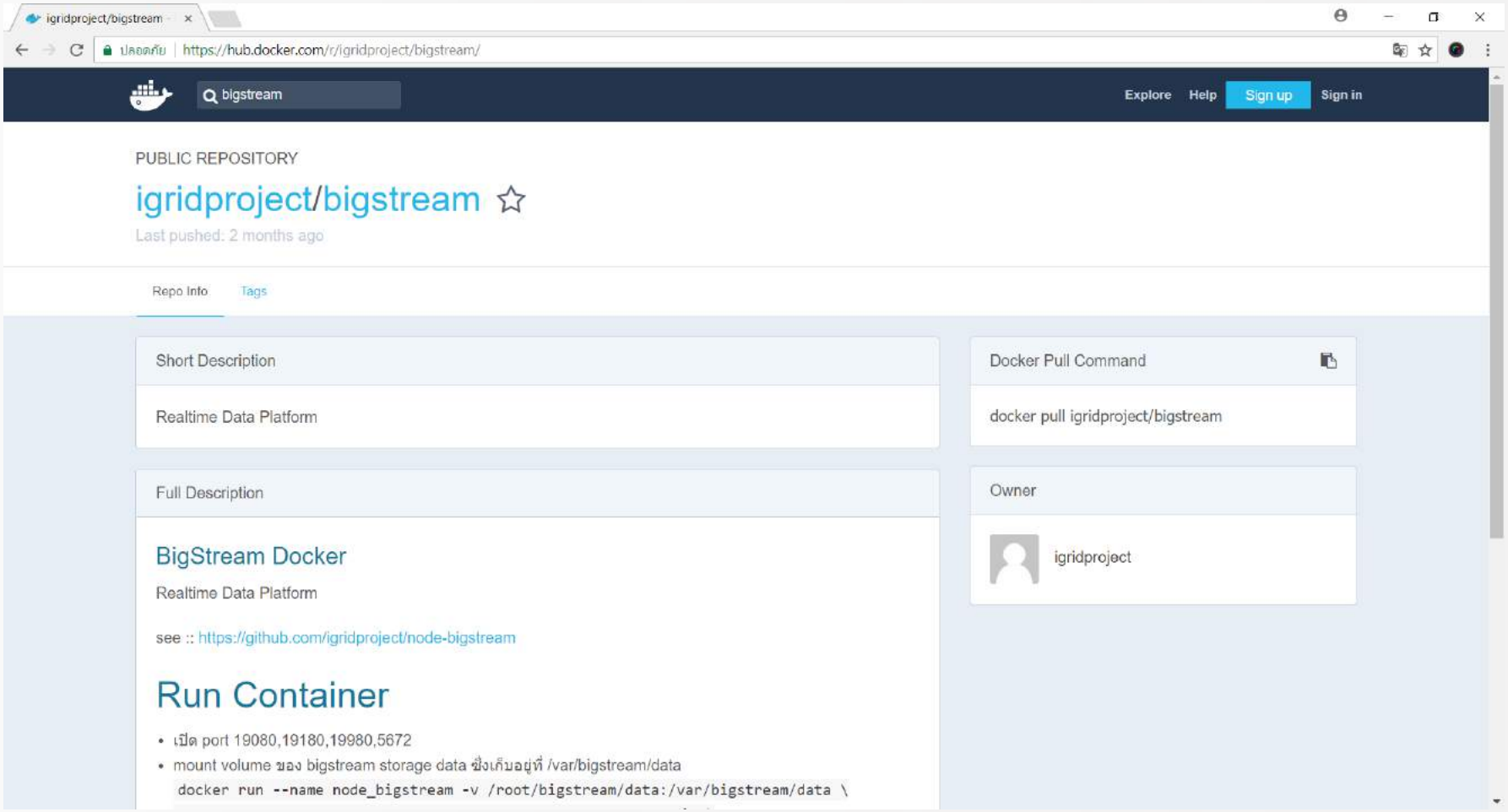
Real-time Epidemic Monitoring

TanRabad-**WATCH**

Morbidity Rate, Number of Patients with Median

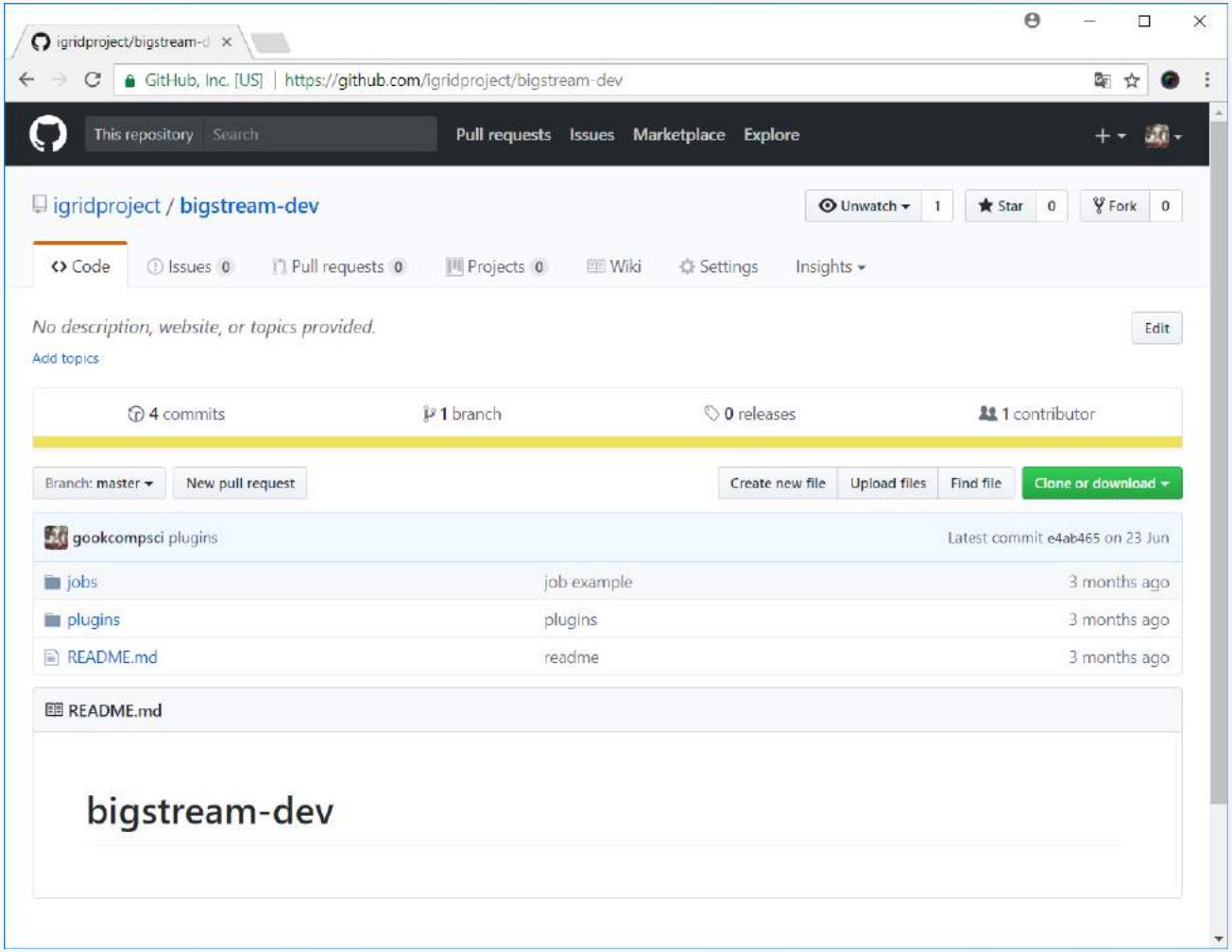


Try BigStream



https://hub.docker.com/r/igridproject/bigstream/

BigStream Dev



https://github.com/igridproject/bigstream-dev

“ งานวิจัยใช้ได้จริง สปริงบอร์ดนวัตกรรม ”

สนใจข้อมูลเพิ่มเติม กรุณาติดต่อ

ทีมวิจัย

คำรณ อรุณเรื่อ

kamron.aroonrua@nectec.or.th

ดร.นัยนา สหเวชชภัณฑ์

naiyana.sahavechaphan@nectec.or.th

Group
“NECTEC BigStream”

Session 7 “Dynamic Data Driven Application Systems (DDDAS) เพื่อสนับสนุนการบริหารจัดการสิ่งแวดล้อมในยุค Big Data”

วันจันทร์ที่ 11 กันยายน 2560 | Meeting room 3 ศูนย์ประชุมแห่งชาติสิริกิติ์