The Internet-of-Thing and Big Data ecosystem

Société Française de Thermique
May 31st, 2018

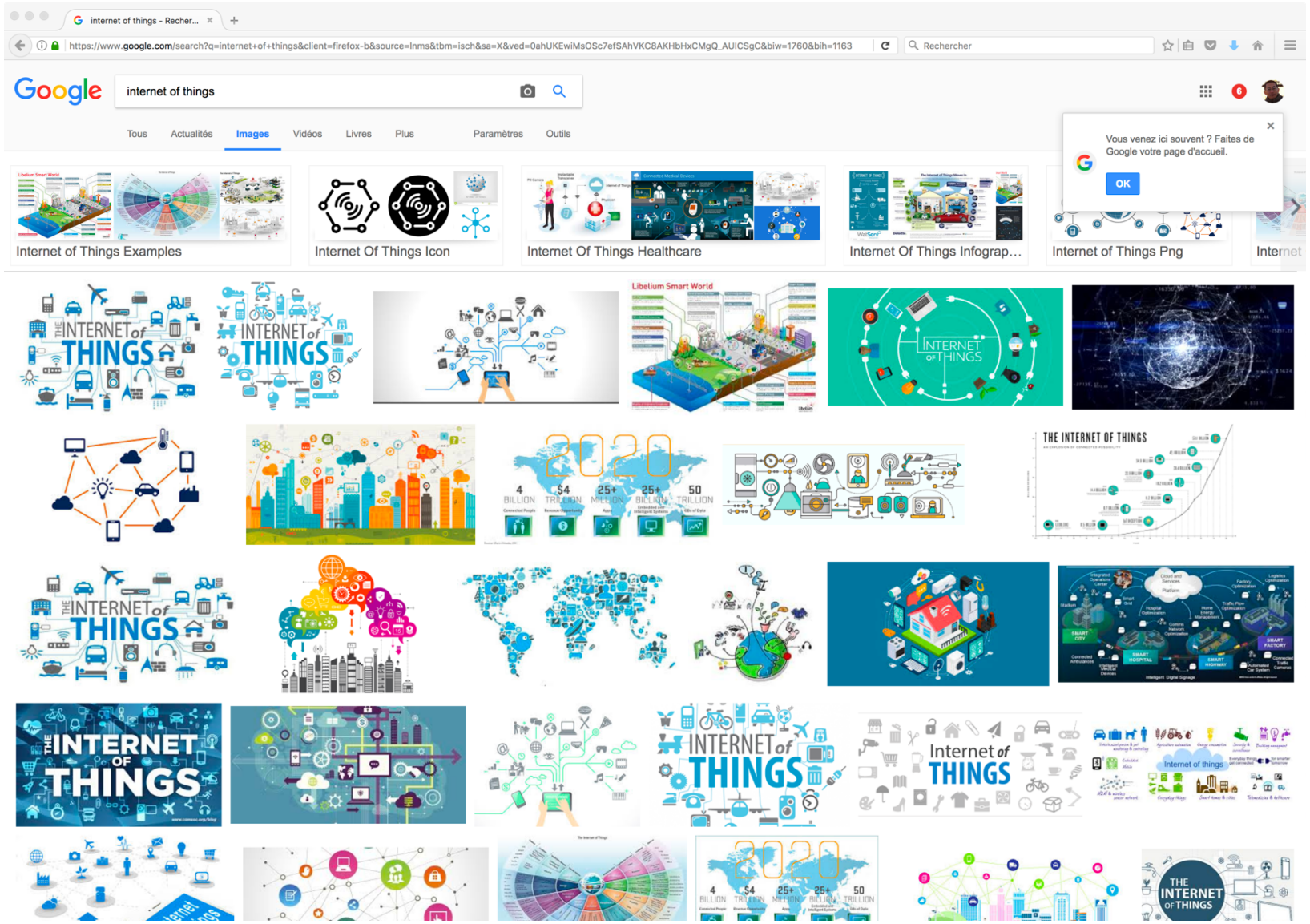
Prof. Congduc Pham
http://www.univ-pau.fr/~cpham
Université de Pau, France

L'Internet-des-Objets pour la collecte et l'analyse de données

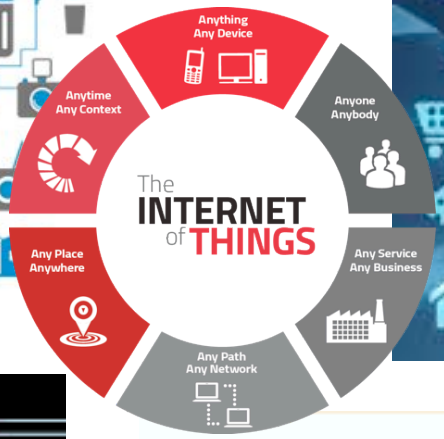
Société Française de Thermique
May 31st, 2018

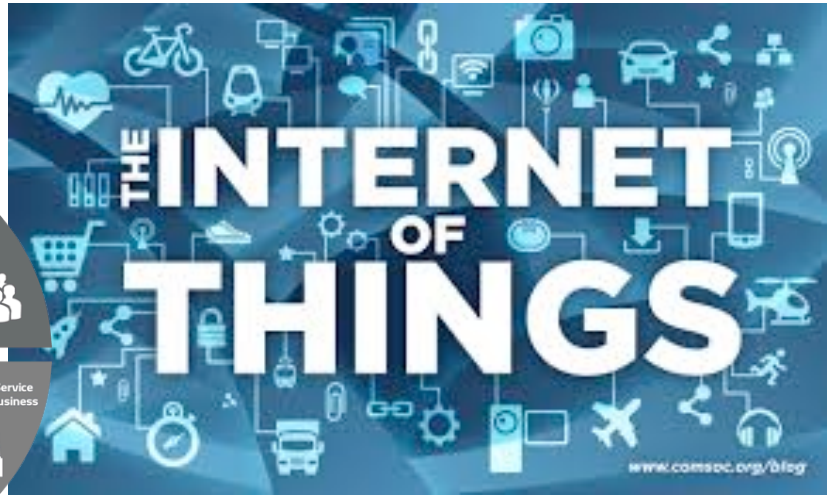
Prof. Congduc Pham
http://www.univ-pau.fr/~cpham
Université de Pau, France

Googling for « Internet of Things »…

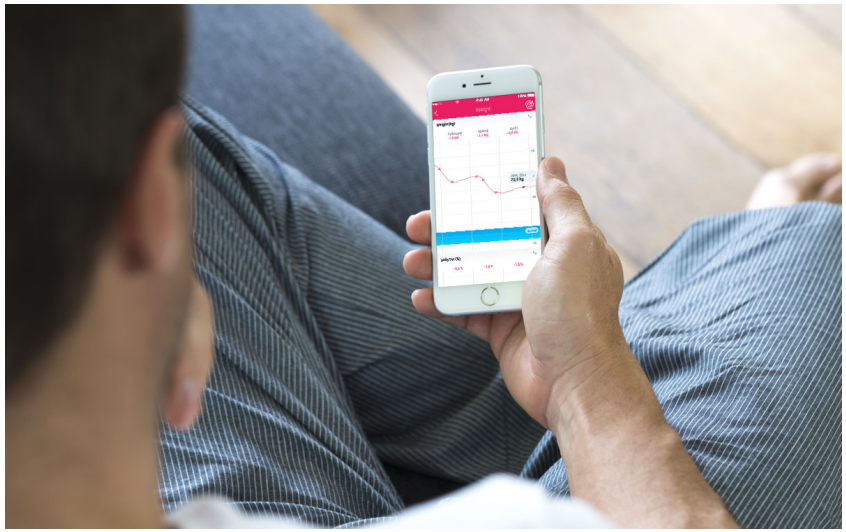


typically shows communicating objects



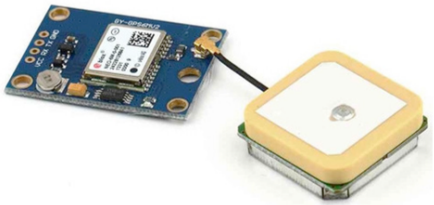


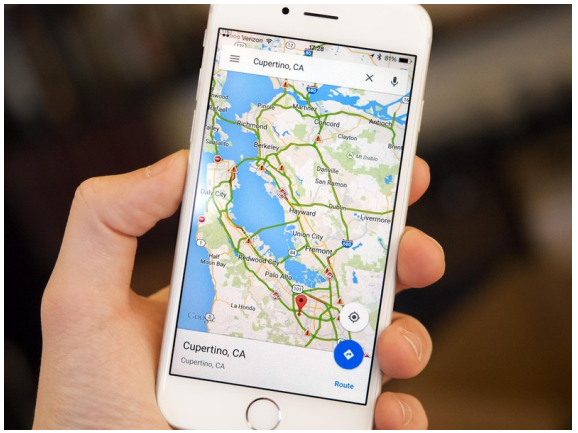
Home/consumer IoT products



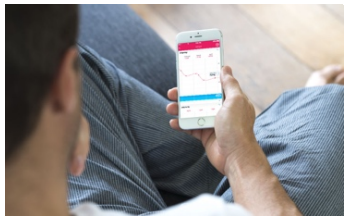
Pictures from WiThing, https://www.withings.com/eu/fr/products/body

IoT & physical world



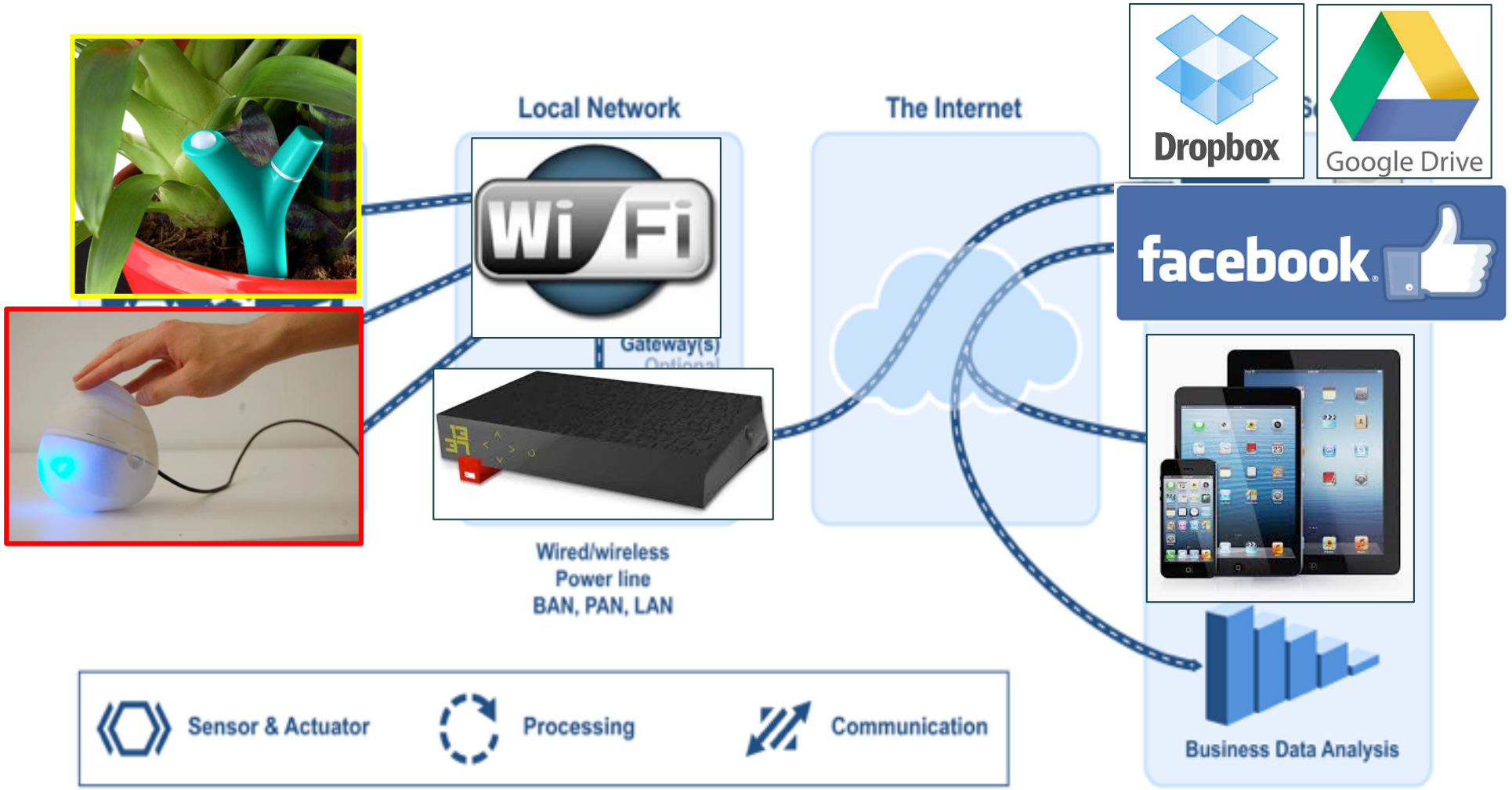


Local interaction is possible



But IoT usually means cloud data

General public IoT architecture

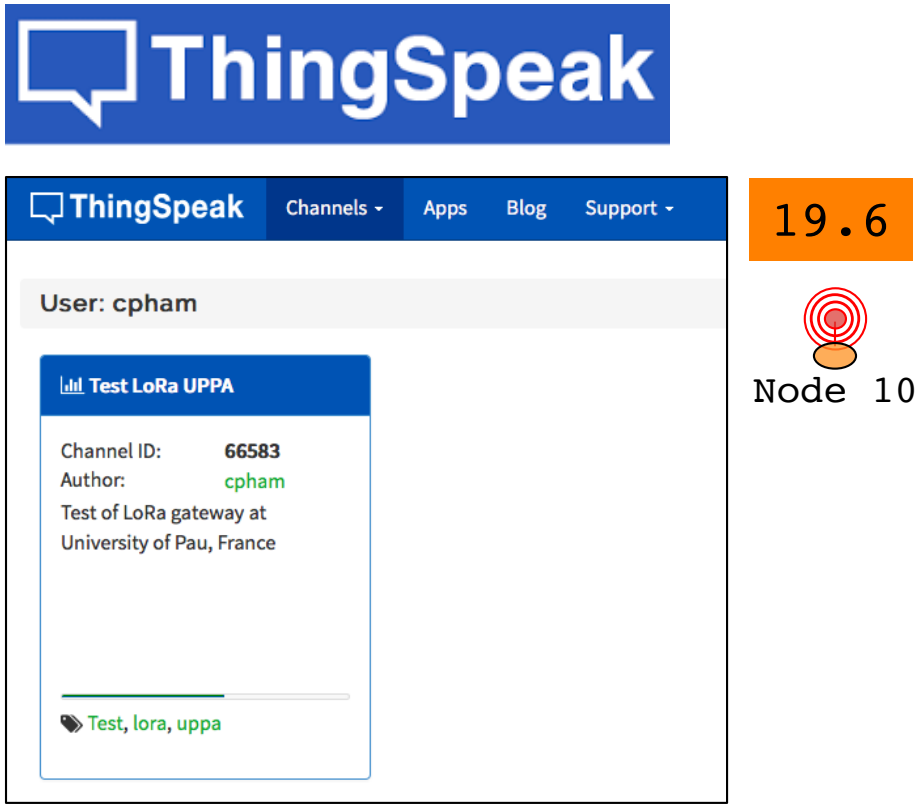


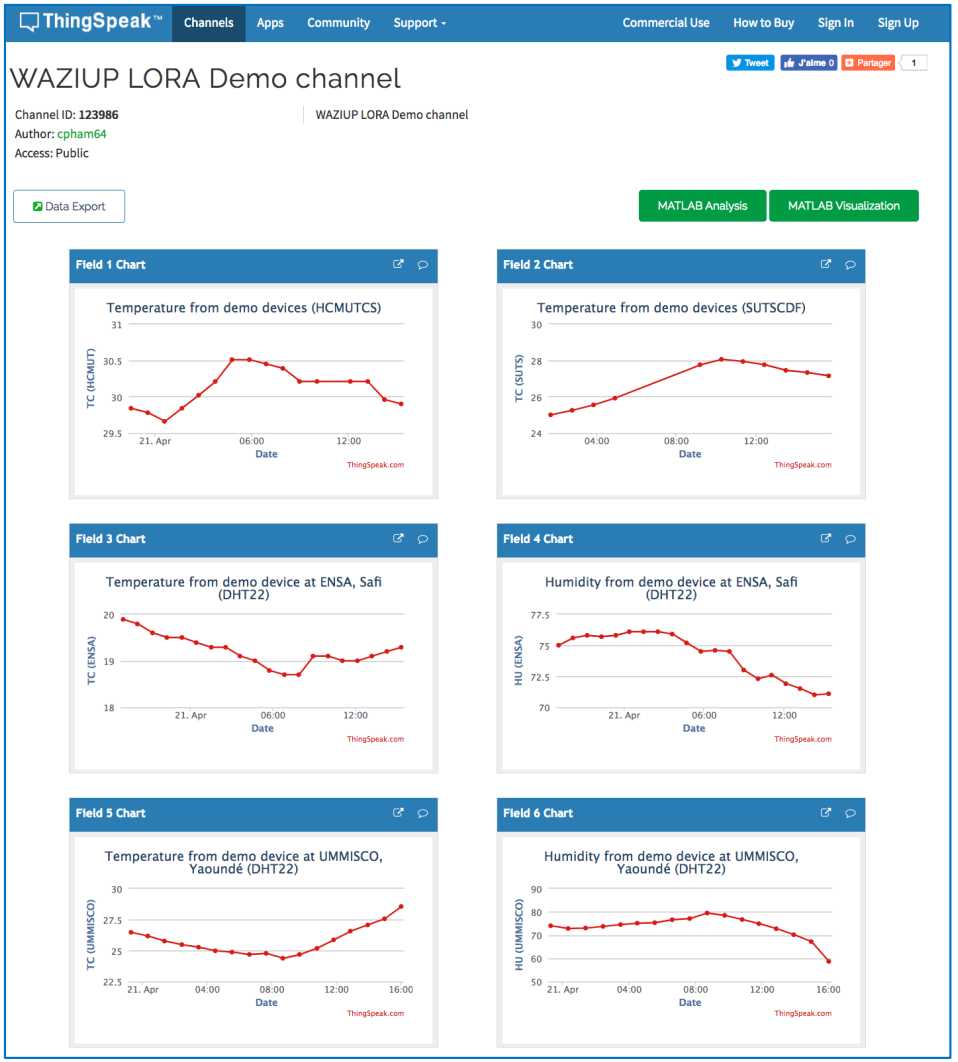
Pictures from ArchitectCorner

Dedicated IoT cloud

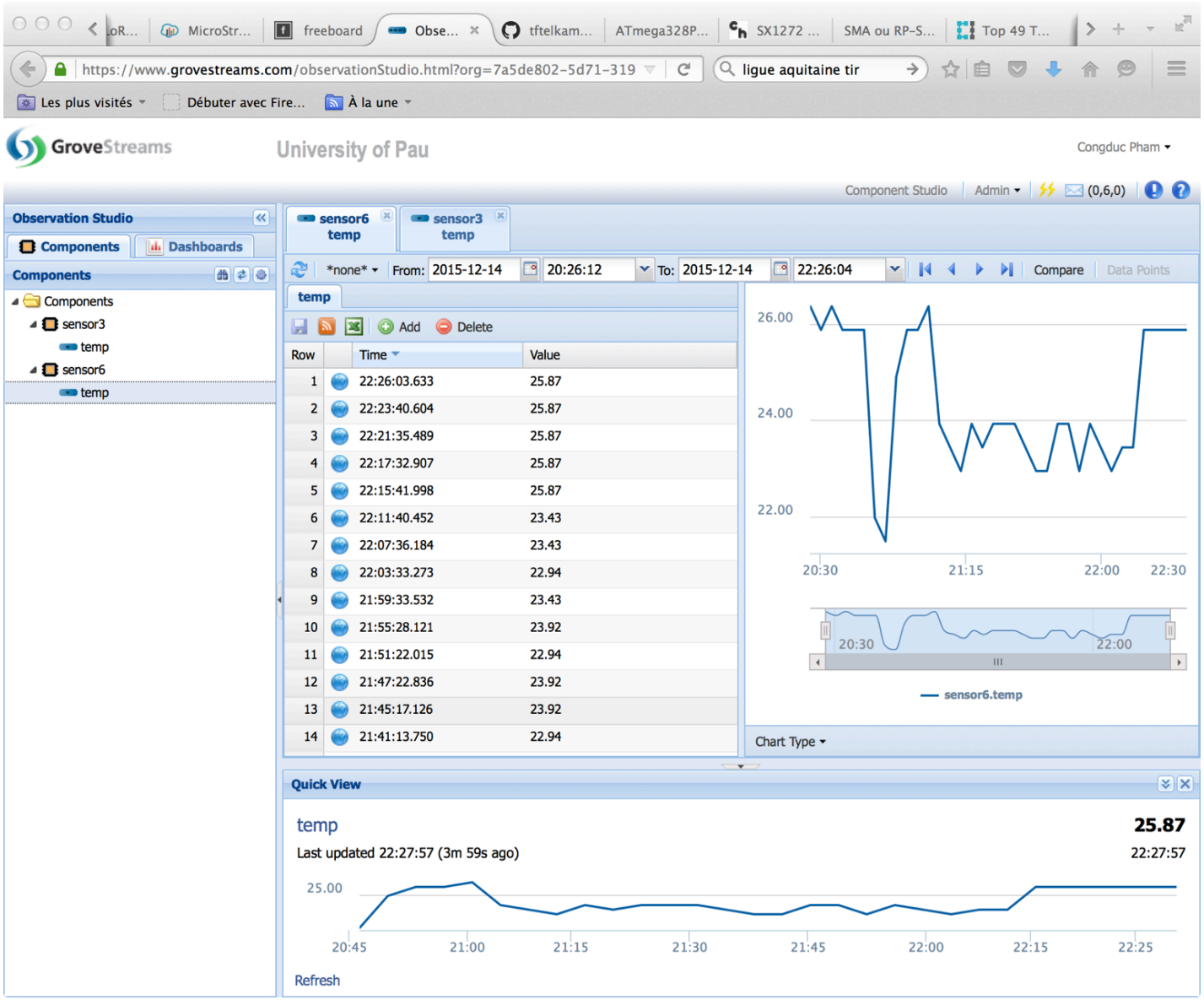
- **Most of them use HTTP POST/GET**

Using ThingSpeak

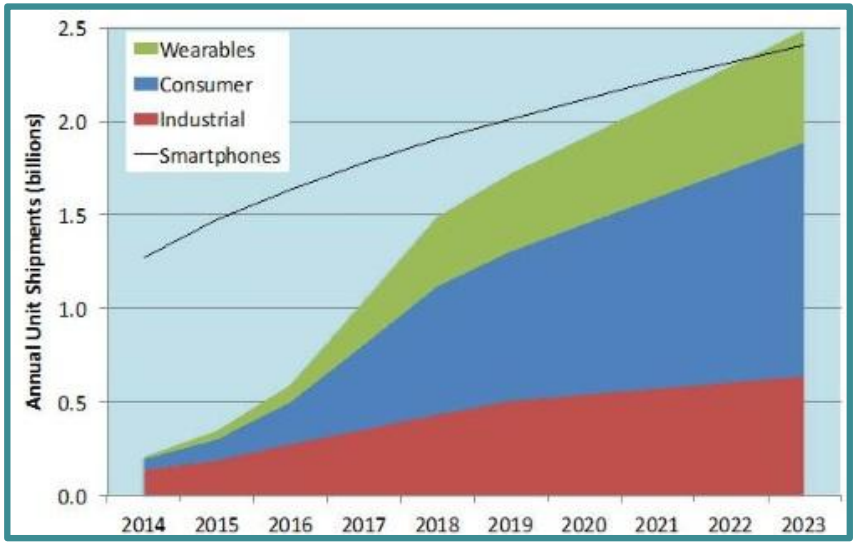


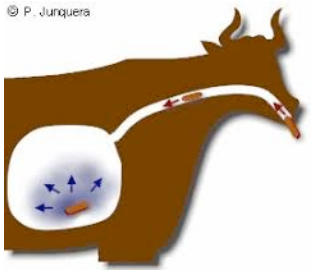


Using GroveStreams

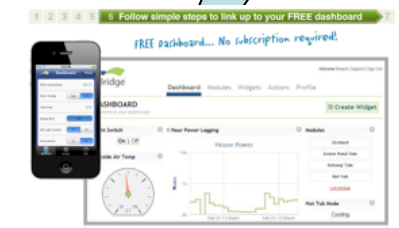


One of the most promising market is IoT!





Control, Optimize & Instrument !



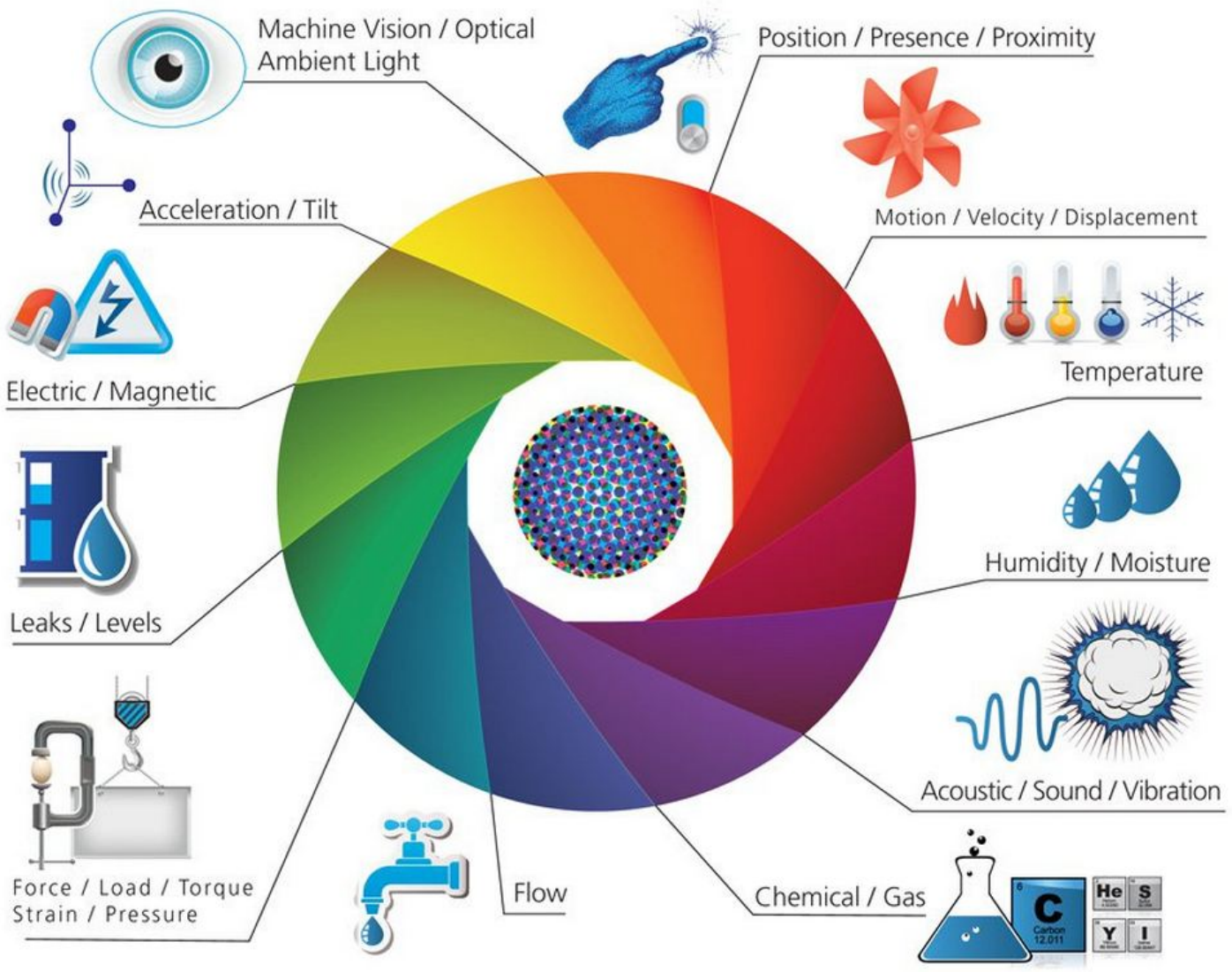
DATA ANALYSIS, OPTIMIZATION & CONTROL

PERVASIVE SYSTEMS

SENSING

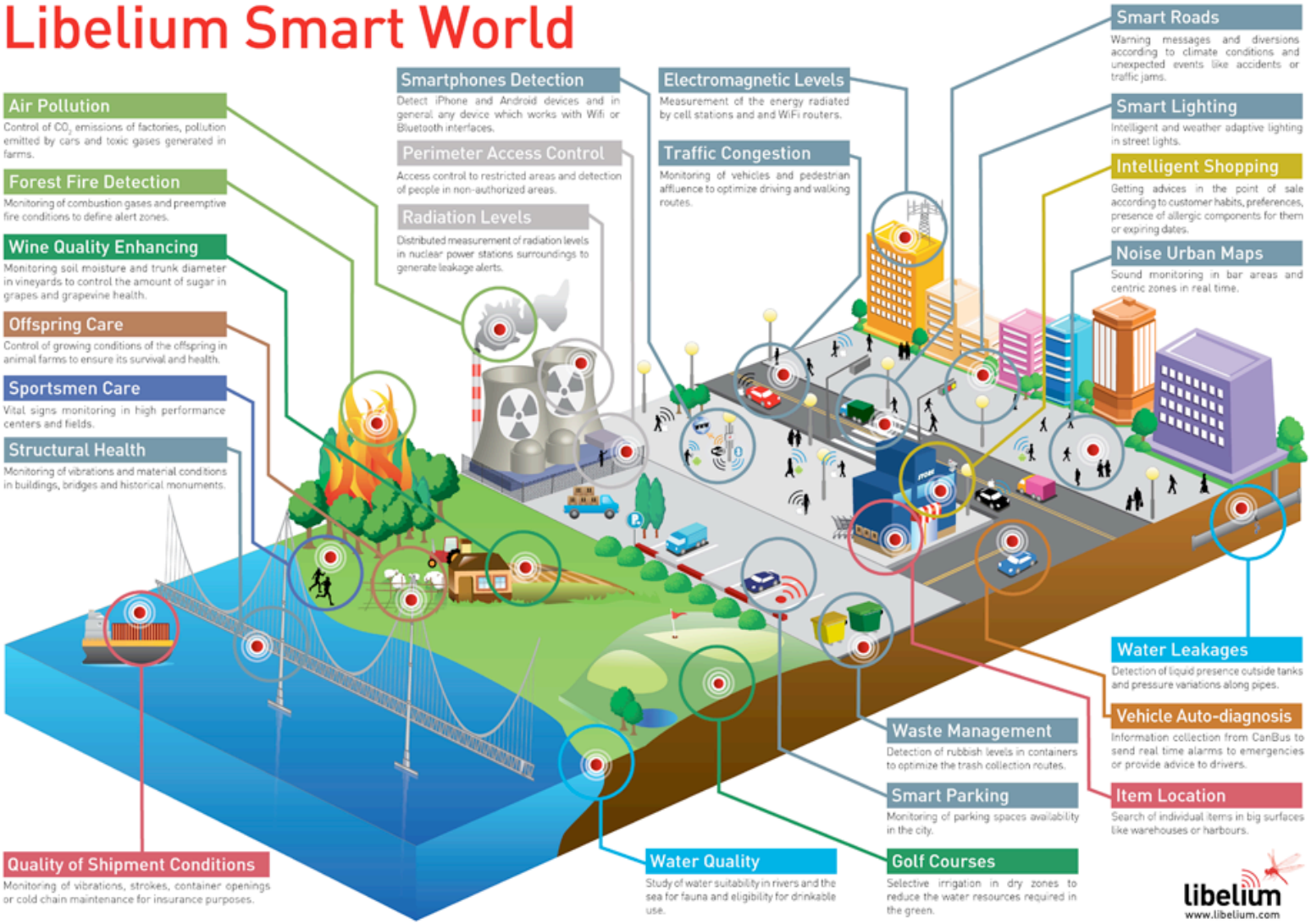
PEOPLES, INFRASTRUCTURES, BUILDINGS, VEHICULES

Large variety of sensoring needs



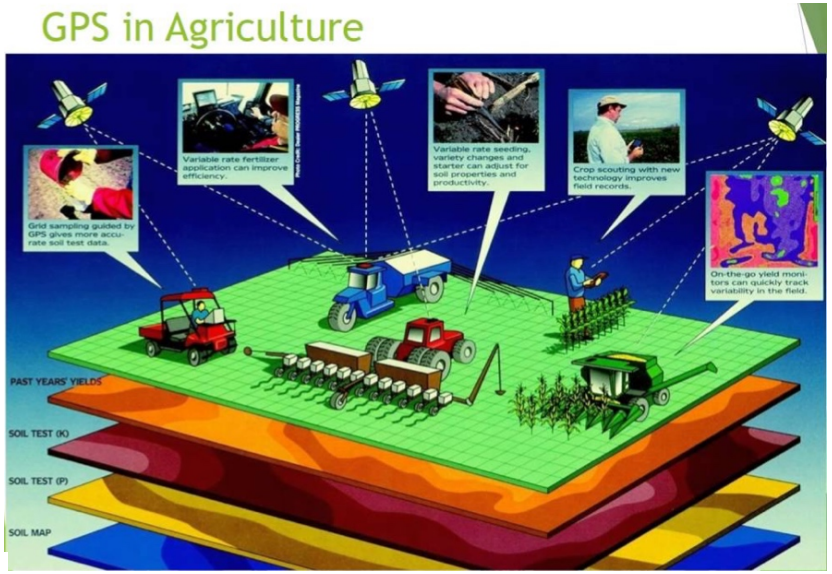
infographic made by Postscapes in collaboration with Harbor Research

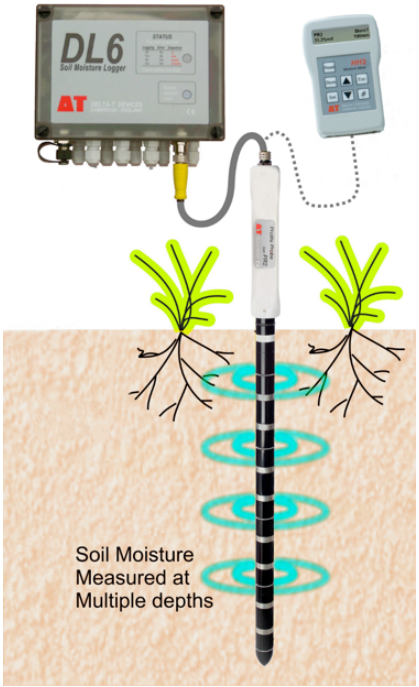
Example 1: Smart Cities



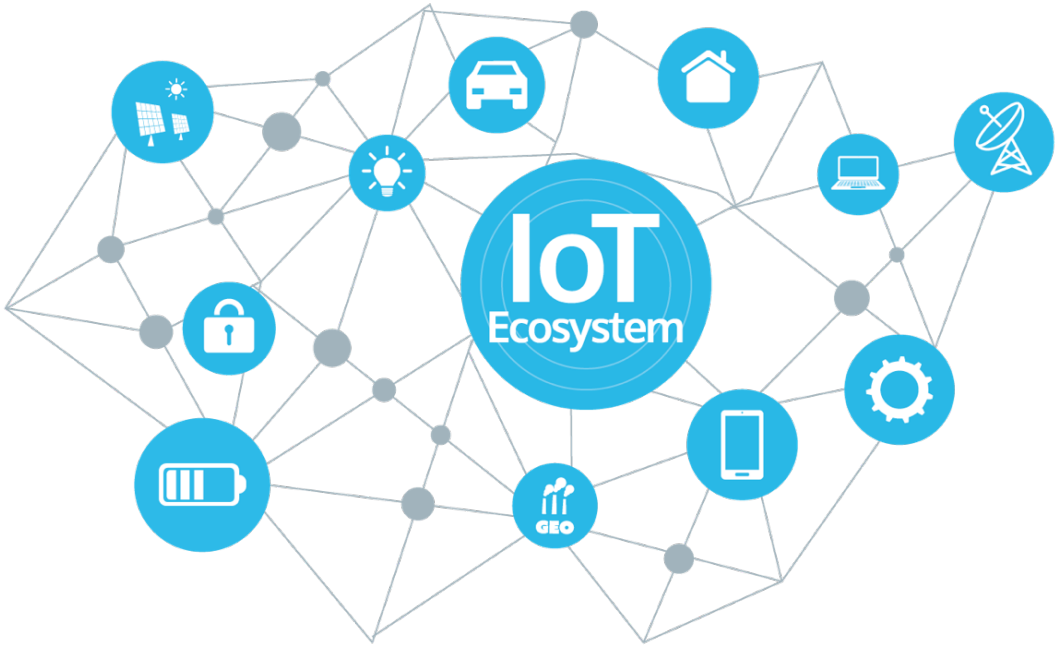
HTTP://WWW.LIBELIUM.COM/TOP_50_IOT_SENSOR_APPLICATIONS_RANKING/#SHOW_INFOGRAPHIC

Example 2: Farming & Agriculture



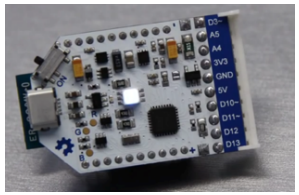


The IoT ecosystem



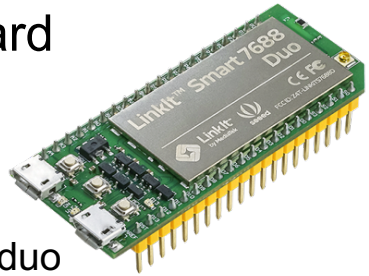
1st issue: IoT are small devices

- ANSWER: Smaller and more powerfull boards are now available!



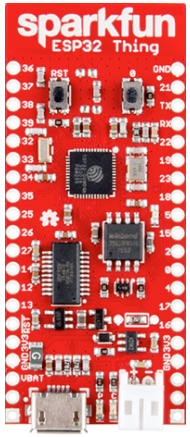
Theairboard

LoPy



LinkIt
Smart7688 duo

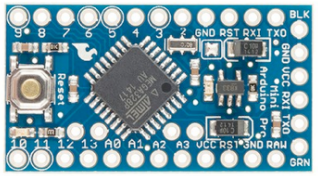
Adafruit Feather



Sparkfun ESP32 Thing

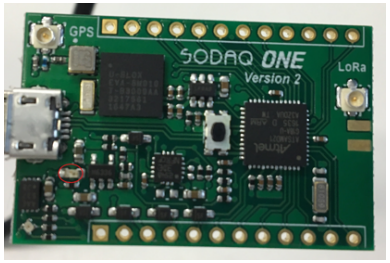
Expressif ESP32

Tessel



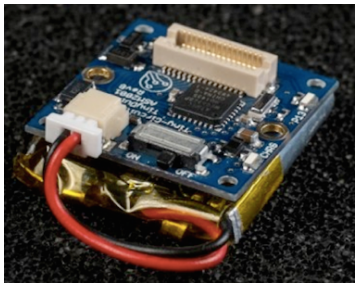
Arduino Pro Mini

Teensy 3.2



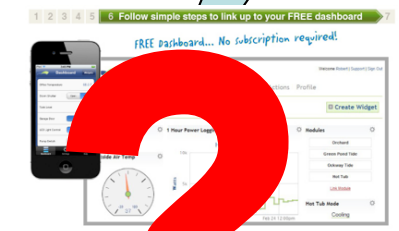
SodaqOnev2

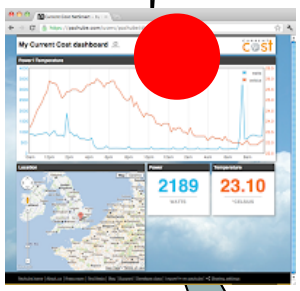
STM32 Nucleo-32



Tinyduino

2nd issue: collect data





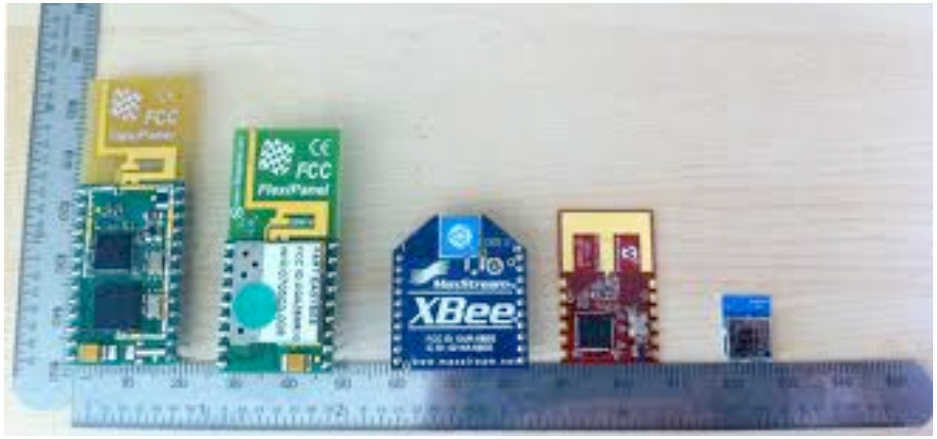
DATA ANALYSIS, OPTIMIZATION & CONTROL

PERVASIVE SYSTEMS

SENSING

PEOPLES, INFRASTRUCTURES, BUILDINGS, VEHICULES

Wireless Communication made easy



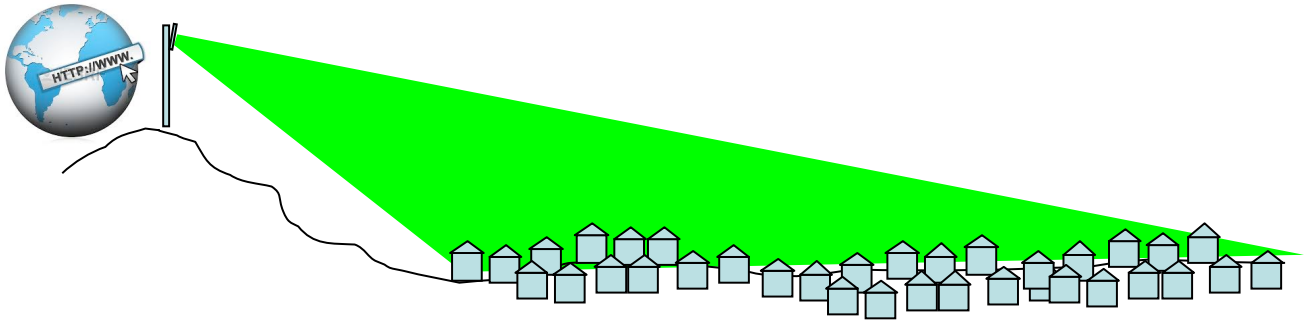
IoT=wireless+battery

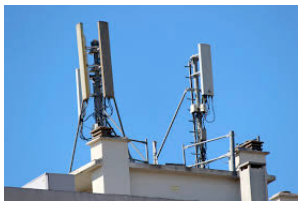
Telemetry and Transmission cost

Moisture/
Temperature of
storage areas

10-15kms

GSM GPRS

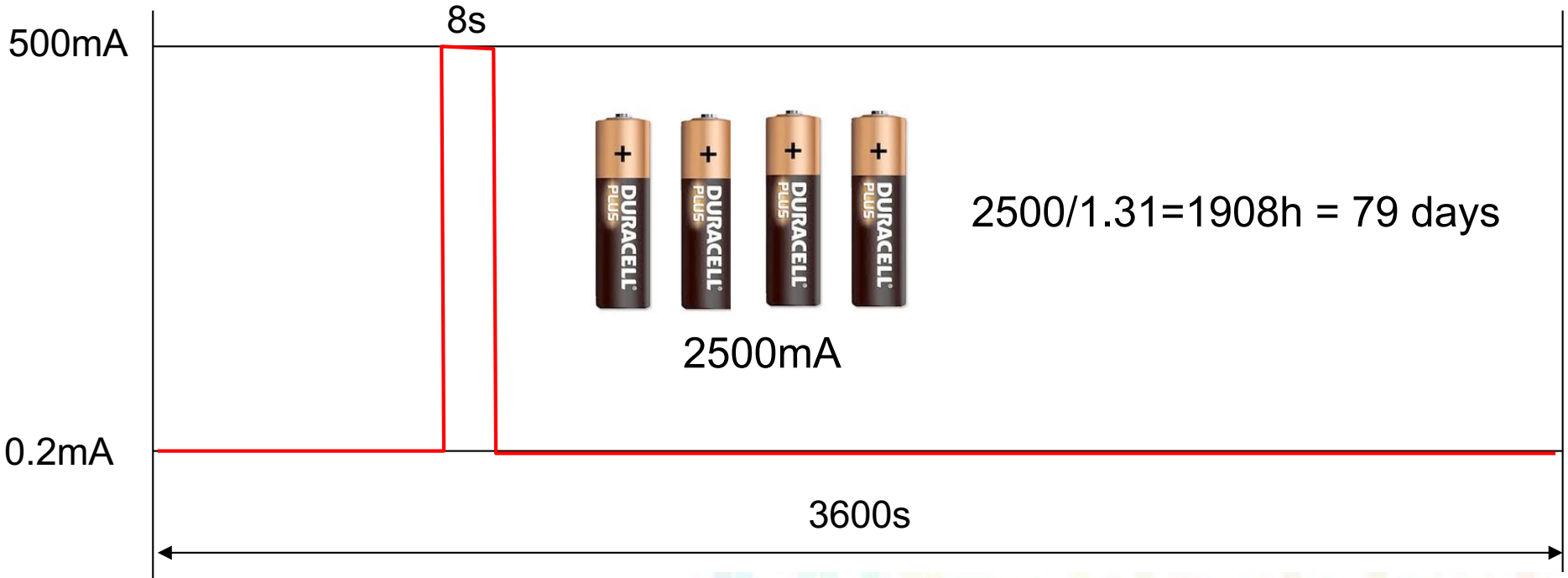


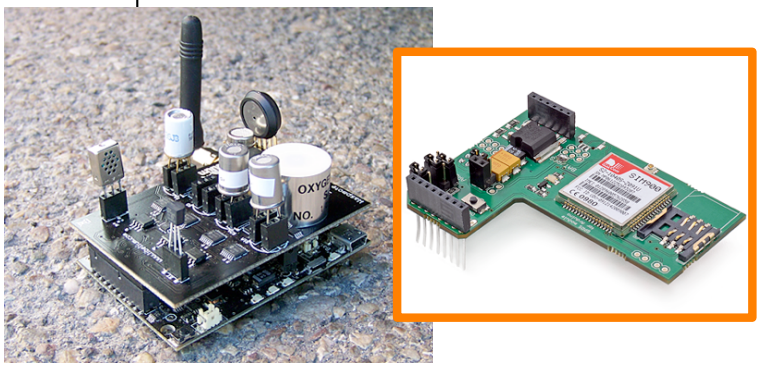


Technology	2G	3G	LAN
Range (I=Indoor, O=Outdoor)	N/A	N/A	O: 300m I: 30m
Tx current consumption	200-500mA	500-1000mA	100-300mA
Standby current	2.3mA	3.5mA	NC

Energy consideration

TX power: 500mA. Mean consumption: (8x500+3592x0.2)/3600=1.31mA

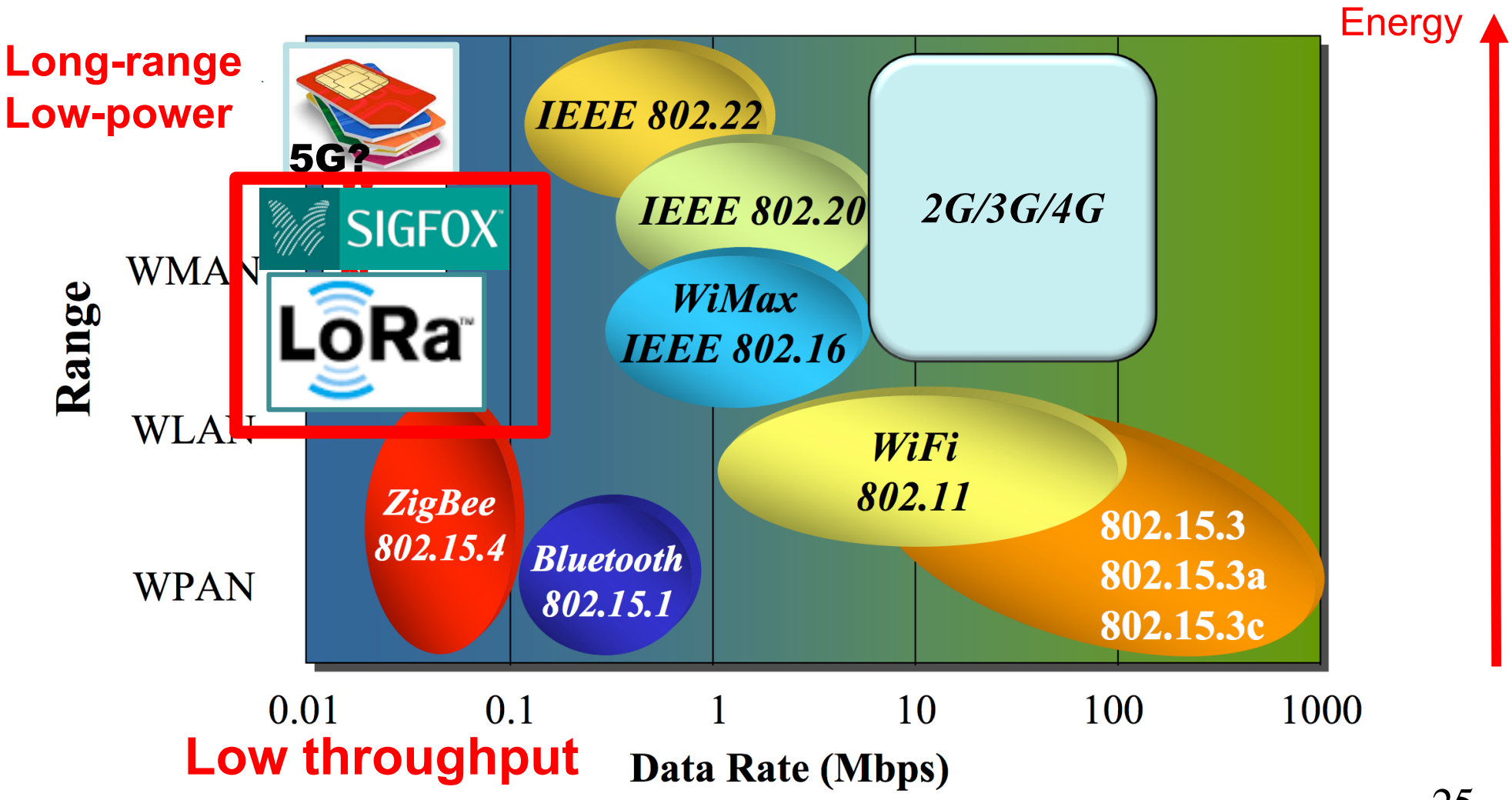




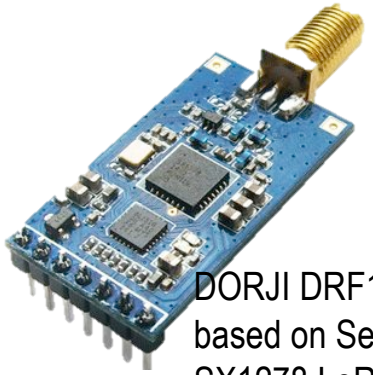
Technology	2G	3G	LAN
Range (I=Indoor, O=Outdoor)	N/A	N/A	O: 300m I: 30m
Tx current consumption	200-500mA	500-1000mA	100-300mA
Standby current	2.3mA	3.5mA	NC

Low-power & long-range radio technologies

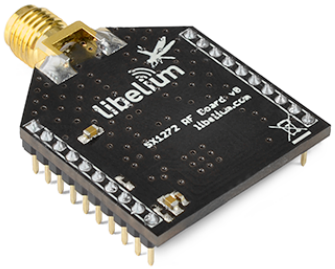
Energy-Range dilemma



LoRa modules from Semtech's SX127x chips



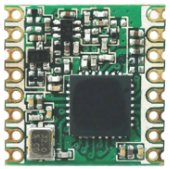
DORJI DRF1278DM is based on Semtech SX1278 LoRa 433MHz



Libelium LoRa is based on Semtech SX1272 LoRa 863-870 MHz for Europe

inAir9 based on SX1276

Froggy Factory LoRa module (Arduino)



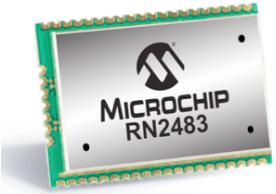
HopeRF RFM series

HopeRF HM-TRLR-D

LinkLabs Symphony module

IMST IM880A-L is based on Semtech SX1272 LoRa 863-870 MHz for Europe

Embit LoRa

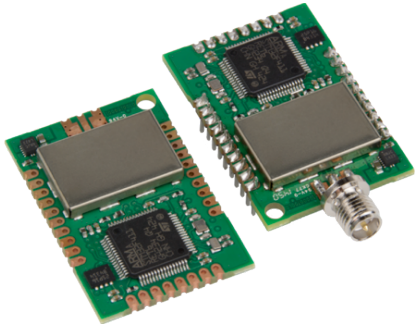


LoRa™ Long-Range Sub-GHz Module (Part # RN2483)

Microship RN2483

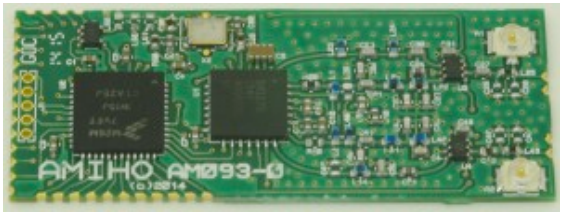
habSupplies

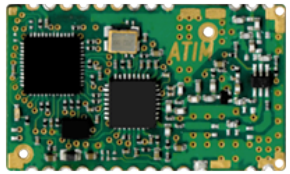
Adeunis ARF8030AA- Lo868



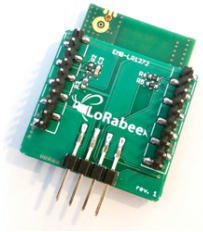
Multi-Tech MultiConnect mDot

AMIHO AM093

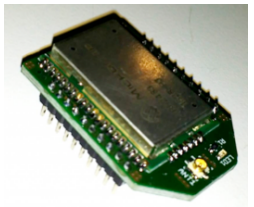




ARM-Nano N8 LoRa module from ATIM



SODAQ LoRaBee Embit



SODAQ LoRaBee RN2483

Energy consumption comparaison

Tables from Semtech

Technology	2G	3G	LAN	ZigBee	Lo Power WAN
Range (I=Indoor, O=Outdoor)	N/A	N/A	O: 300m I: 30m	O: 90m I: 30m	Same as 2G/3G
Tx current consumption	200-500mA	500-1000mA	100-300mA	18mA	18mA-40mA
Standby current	2.3mA	3.5mA	NC	0.003mA	0.001mA
Energy harvesting (solar, other)	No	No	No	Possible	Possible
Battery 2000mAh (LR6 battery)	4-8 hours(com) 36 days(idle)	2-4 hours(com) X hours(idle)	50 hours(com) X hours(idle)	60hours (com)	120 hours(com) 10 year(idle)

TX power: 30mA. Mean consumption: (8x30+3592x0.2)/3600=0.266mA

2500/0.266=9398h = 391 days = 13 months

LPWAN networks roll-ou

SIGFOX and Glen Canyon Corp. to 1 Million Smart Meters to of Things

LoRa™ technology to be integrated into FLASHNET's street lighting management solution

Tata Communication world's largest IoT n India

Semtech and STMicroelectronics aborate to Scale LoRa nology to Meet High-Volume Demands of rnet of Things Applications

OTIO to Connect 1 Million Devices to SIGFOX

From www.businesswire.com - February 22, 4:59 PM

OTIO, a subsidiary of Groupe HBF specializ electrical equipment, lighting and home automation, has chosen the SIGFOX netwo deploy its new international offer of connec devices.

Swisscom sets up a Swiss-wide network for the Internet of Things

From www.swisscom.ch - March 14, 7:49 PM

The Internet of Things has long connected millions of objects and devices to one another and to people. In the future, this number will reach into the billions worldwide. Swisscom is the first provider in Switzerland to set up an additional network dedicated to the Internet of Things: the Low Power Network, designed for the transmission of small amounts of data independently of the electrical network.

From www.sudo 16, 2015 2:24 PM

" Gestionnaire les Landes, e Dordogne, So ondes radio de Sigfox pour relever les compteurs et surveiller l'état des canalisations Le compteur..."

From www.enevo.com - March 6, 4:12 PM

network, a narrow-band technology which guarantees connectivity at a reduced energy consumption rate and at a lower cost. Orange has chosen to rely on LoRa (Long Range) technology to deploy this network that will cover the whole of metropolitan France.

T-Mobile to cover Republic with the for the Internet of

From www.theinternetofthin September 10, 4:41 PM

Following a pilot opera Czech Republic that ex expectations, T-Mobile SimpleCell Networks w SIGFOX's Internet of Th network throughout the country.

Rise from the Ashes?

" French Telecom LoRa radio technology for its own domestic IoT and M2M network."

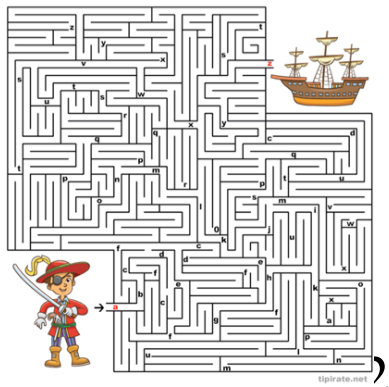
News taken from scoop.it (N. Ayuso)

3rd issue: finding the information you need

- Searching for information is a tough issue
  - Web search engine: Google,…
- Most IoT clouds uses HTTP request (GET, POST, PUT, …) to push/store data to web platforms/servers

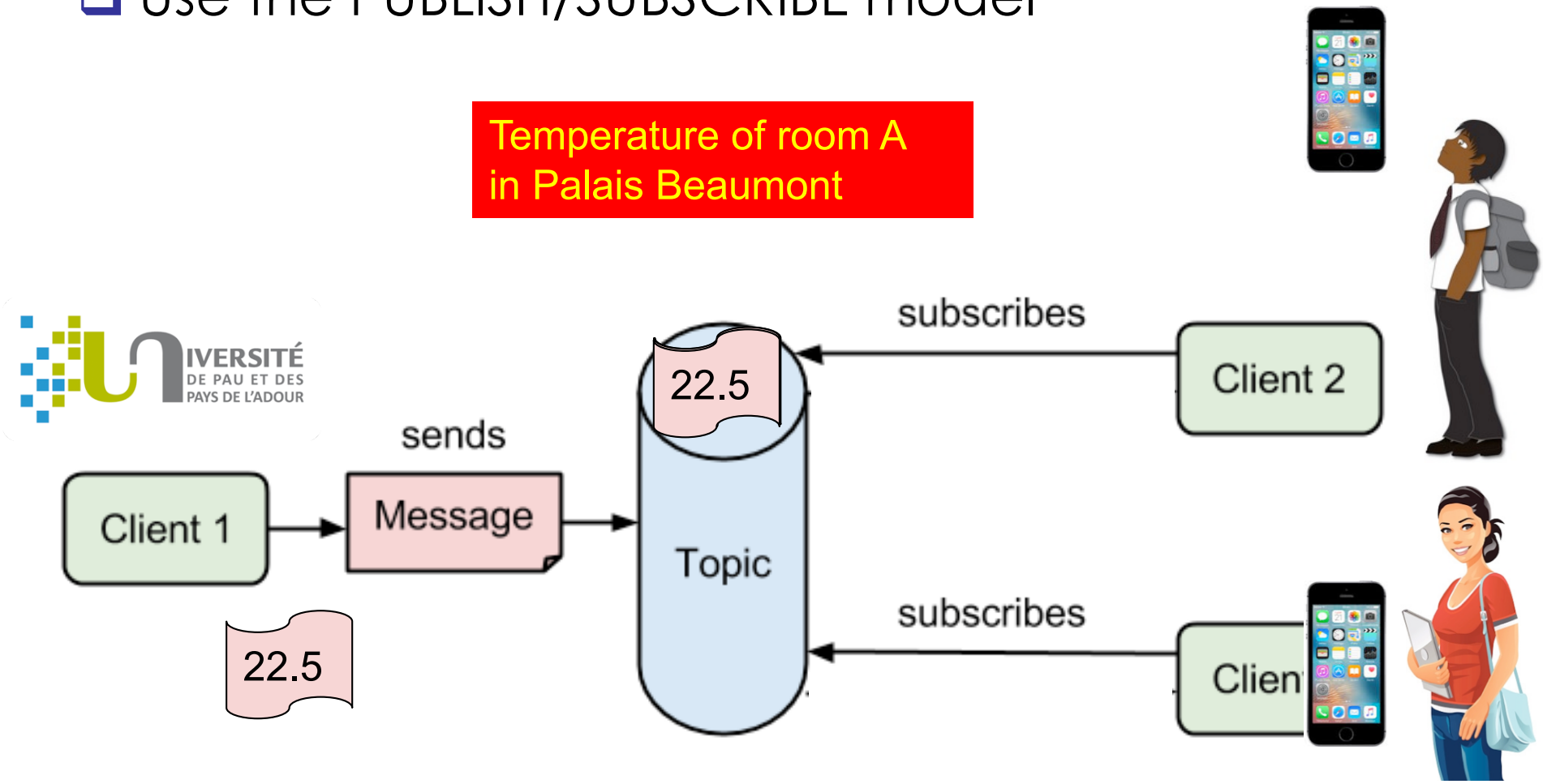
ThingSpeak

- If you need an information, for instance the temperature in room A of Palais Beaumont, then you have to go to the right web page
- When there can be millions of IoT nodes providing large variety of data, it is difficult to find your way!



From *search for info* to *get the info*

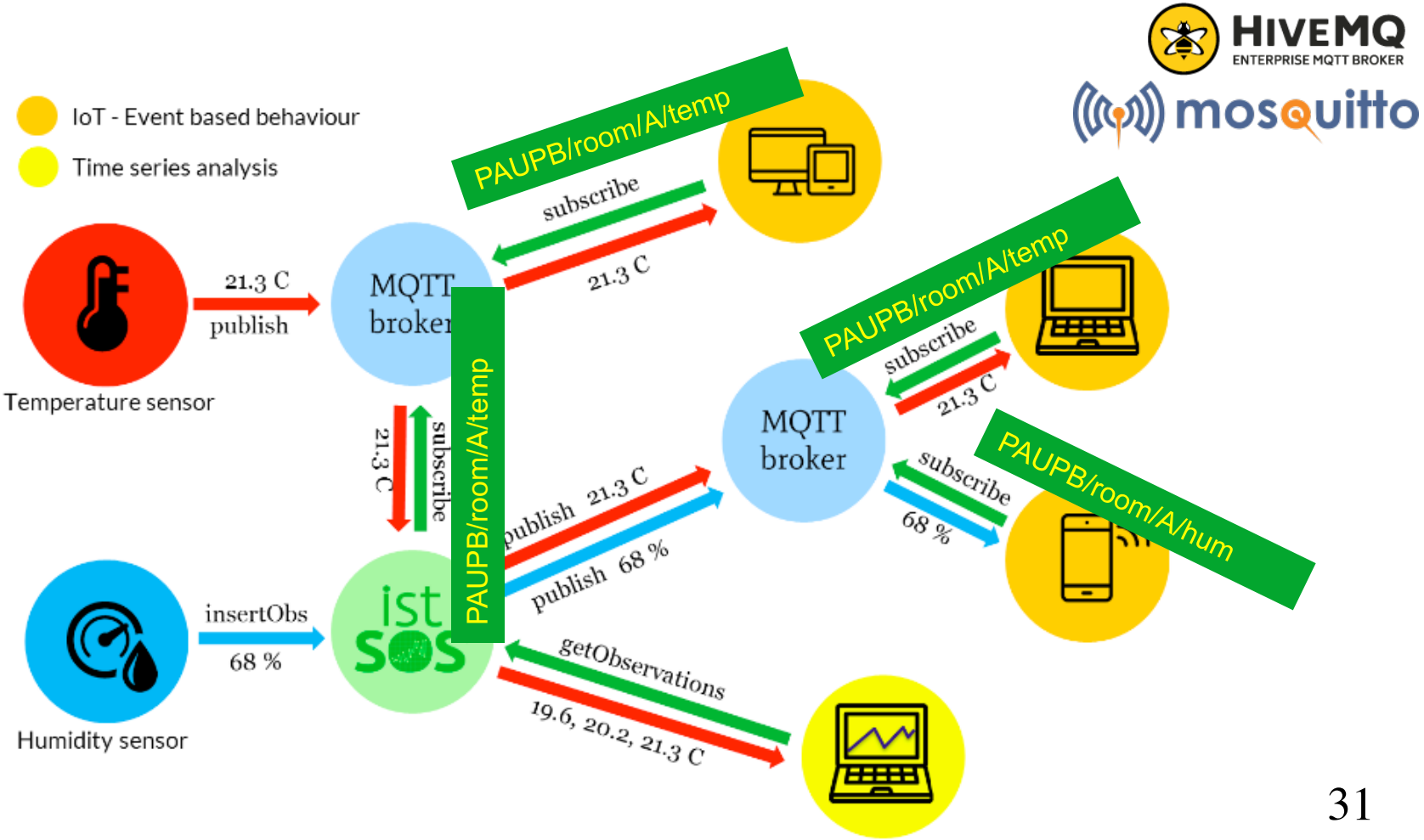
- Use the PUBLISH/SUBSCRIBE model



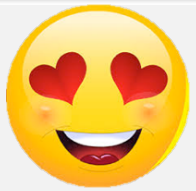
MQTT

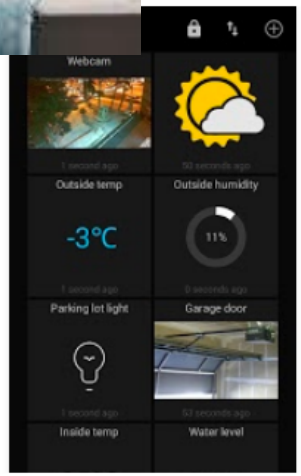
Message Queue Telemetry Transport

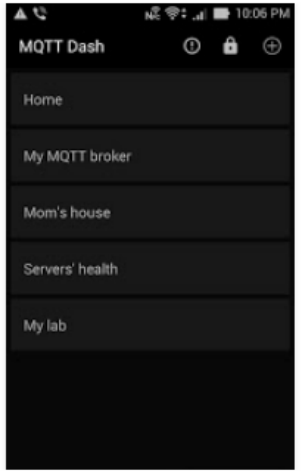
- Use broker nodes to manage topics
  - PAUPB/room/A/temp, PAUPB/room/A/hum

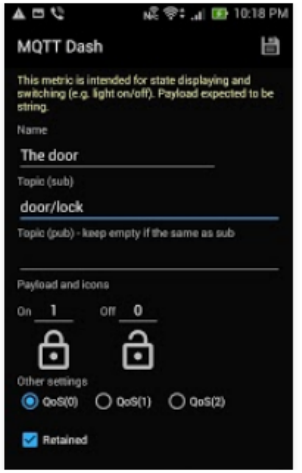


MQTT+smartphone=



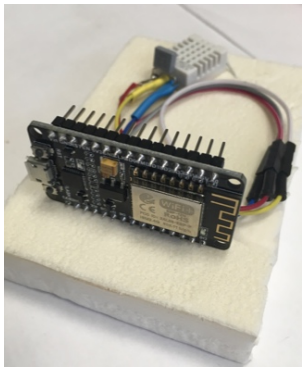






- Towards open data
  - PAUPB/room/#
  - PAU/CITY/WEATHER/#
  - SFT/CONGRESS/#

4th issue: make it simpler?



FIWARE

ioBridge® Connect things.

dweet.io

freeboard

GroveStreams

Firebase

TempoIQ

ThingSpeak

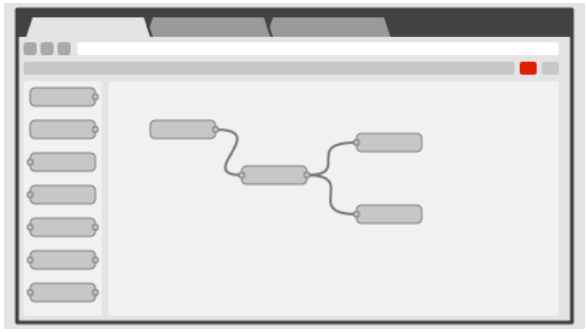
SensorCloud™

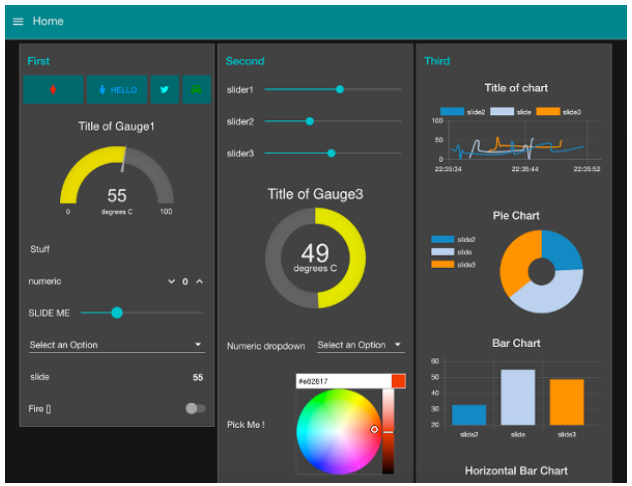
OpenRemote

- End-users are not necessarily computer science experts nor high-skilled programmers
- Use graphical tools to build data processing flows, allowing intuivive connection from data producers to data consumers

Node-Red

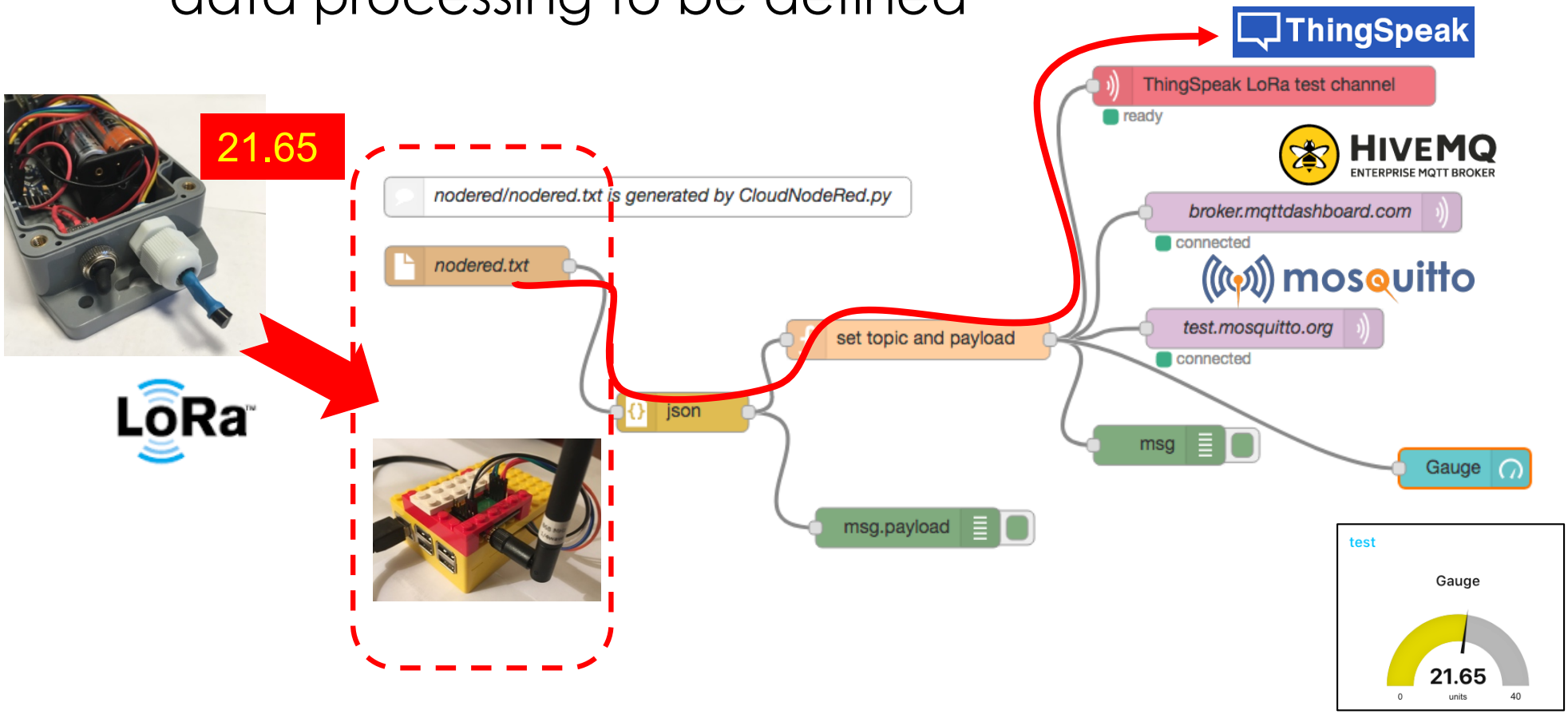
- Node-RED is a programming tool for wiring together hardware devices, APIs and online services, e.g. clouds of various types
- provides a browser-based flow editor to wire together flows with a wide range of nodes



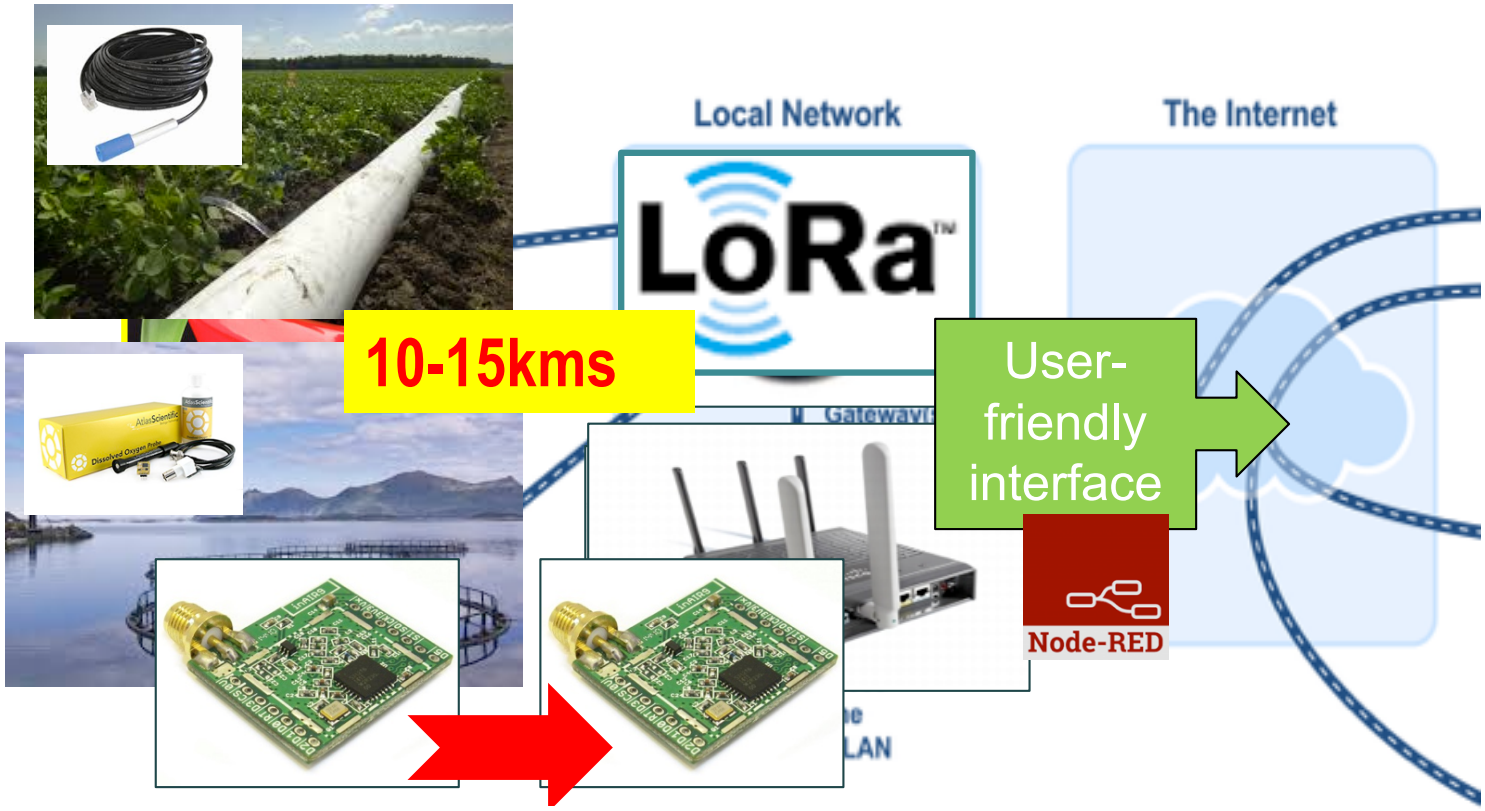


Node-red enabled IoT gateway

- Messages received on the IoT gateway can be injected into a Node-Red flow, allowing complex data processing to be defined



Global picture of long-range IoT ecosystem



FIWARE
dweet.io
ioBridge
Connect things.
freeboard
GroveStreams
Firebase
TempoIQ
ThingSpeak
SensorCloud
OpenRemote

The IoT BackOffice

IoT usually means

HORIZON 2020

WAZIUP

40 ZETTABYTES

[43 TRILLION GIGABYTES]

of data will be created by 2020, an increase of 300 times from 2005

6 BILLION PEOPLE

have cell phones

WORLD POPULATION: 7 BILLION

Volume

SCALE OF DATA

It's estimated that

2.5 QUINTILLION BYTES

[2.3 TRILLION GIGABYTES]

of data are created each day

Most companies in the U.S. have at least

100 TERABYTES

[100,000 GIGABYTES]

of data stored

The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015

4.4 MILLION IT JOBS

will be created globally to support big data, with 1.9 million in the United States

As of 2011, the global size of data in healthcare was estimated to be

150 EXABYTES

[161 BILLION GIGABYTES]

Variety

DIFFERENT FORMS OF DATA

By 2014, it's anticipated there will be

420 MILLION WEARABLE, WIRELESS HEALTH MONITORS

4 BILLION+ HOURS OF VIDEO

are watched on YouTube each month

30 BILLION PIECES OF CONTENT

are shared on Facebook every month

400 MILLION TWEETS

are sent per day by about 200 million monthly active users

The New York Stock Exchange captures

1 TB OF TRADE INFORMATION

during each trading session

Modern cars have close to

100 SENSORS

that monitor items such as fuel level and tire pressure

Velocity

ANALYSIS OF STREAMING DATA

By 2016, it is projected there will be

18.9 BILLION NETWORK CONNECTIONS

– almost 2.5 connections per person on earth

1 IN 3 BUSINESS LEADERS

don't trust the information they use to make decisions

Poor data quality costs the US economy around

$3.1 TRILLION A YEAR

27% OF RESPONDENTS

in one survey were unsure of how much of their data was inaccurate

Veracity

UNCERTAINTY OF DATA

Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTEC, QAS

IBM

But also how to analyse the data

- What is the meaning of the collected data?
- Example with farming
  - What is interesting for farmers?
    - Fertility detection
    - Eating/Ruminating time for welfare
  - What data can be easily obtained?
    - accelerometer data with neck-mounted collar
  - How to detect relevant event from these data?

Advanced data analysis

Need of experts from the domain!

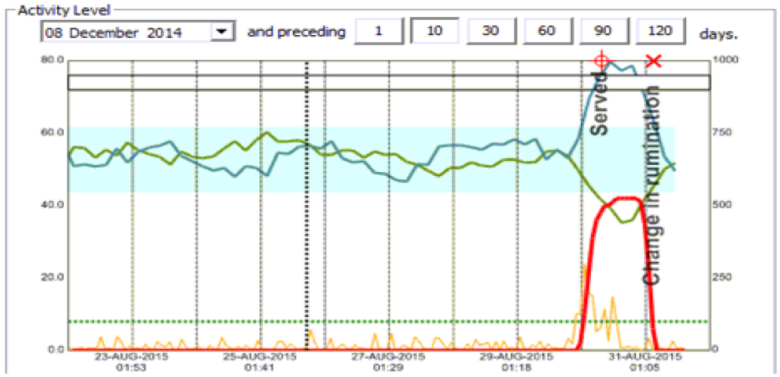
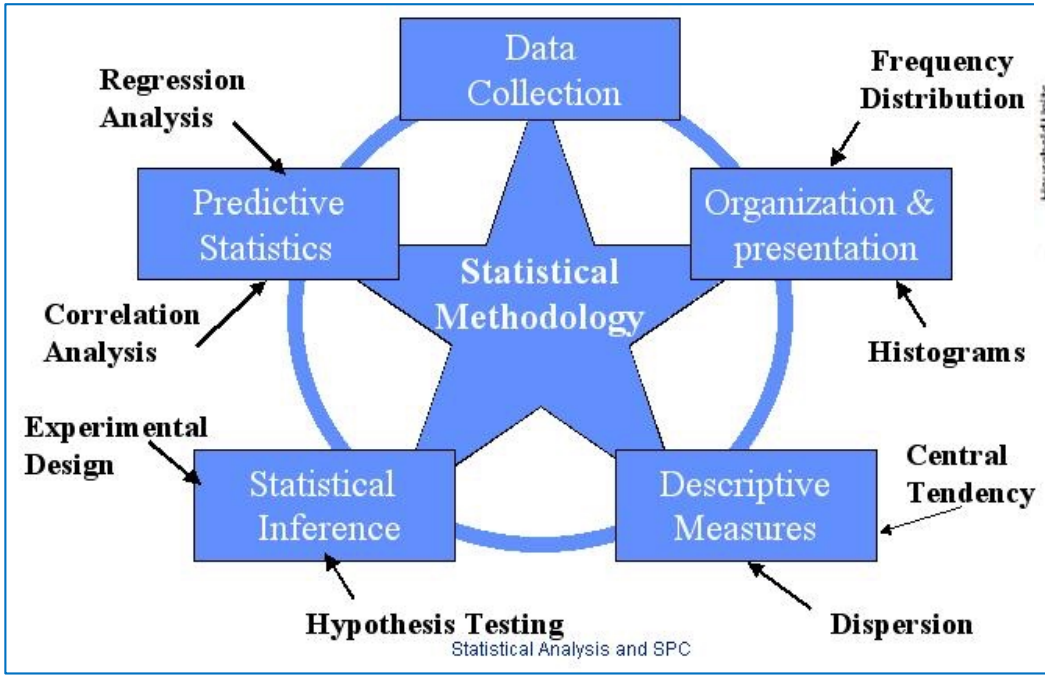
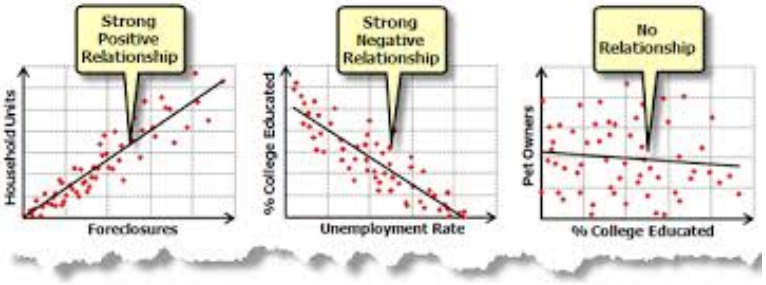


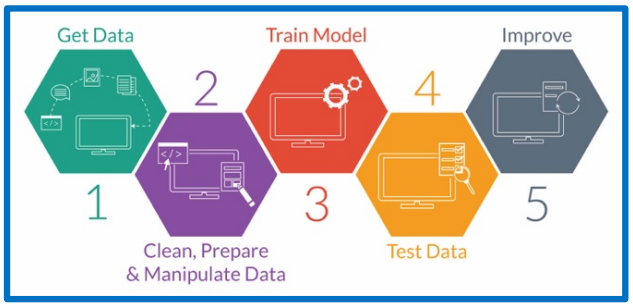
Fig. 3. Illustration of a rise in activity accompanied by a fall in rumination at the point of oestrus

Analysis techniques

- Traditional statistic methods still valid, and useful!

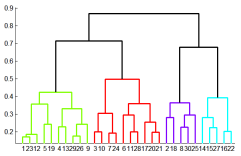


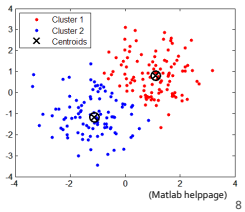




Clustering Analysis

- Definition
  - Grouping unlabeled data into clusters, for the purpose of inference of hidden structures or information
- Dissimilarity measurement
  - Distance : Euclidean(L_2), Manhattan(L_1), …
  - Angle : Inner product, …
  - Non-metric : Rank, Intensity, …
- Types of Clustering
  - Hierarchical
    - Agglomerative or divisive
  - Partitioning
    - K-means, VQ, MDS, …



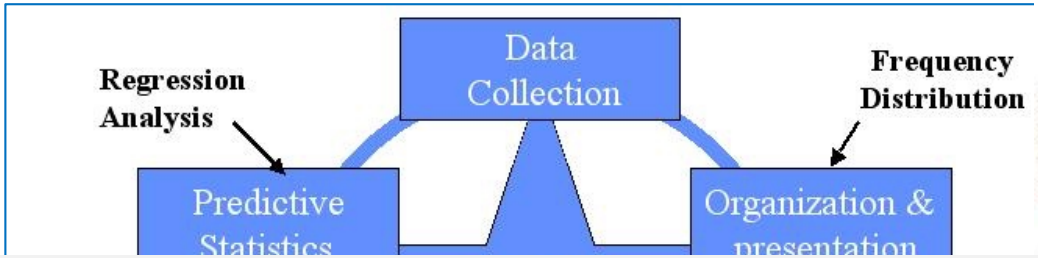


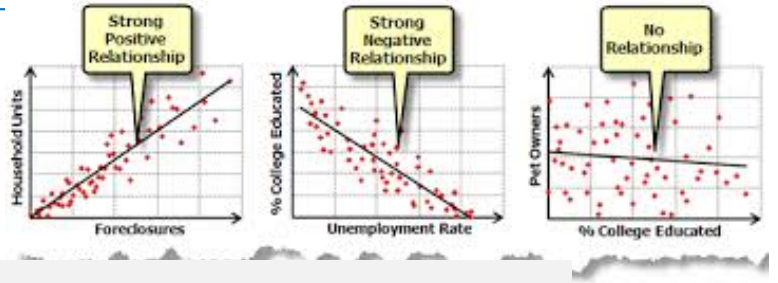
8

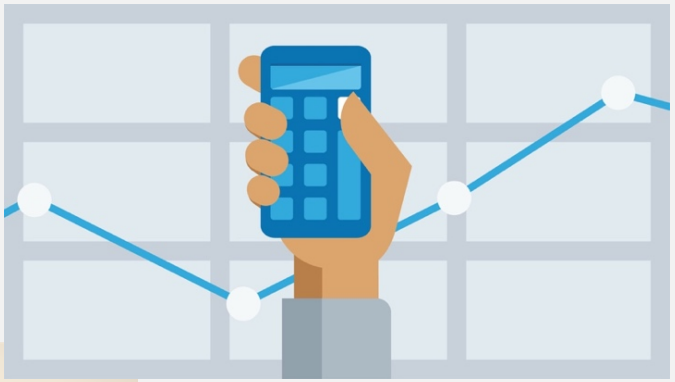
From Jong Youl Choi

Analysis techniques

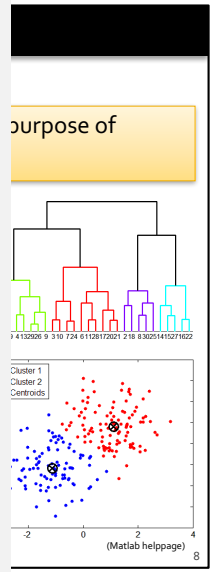
- Traditional statistic methods still valid, and useful!







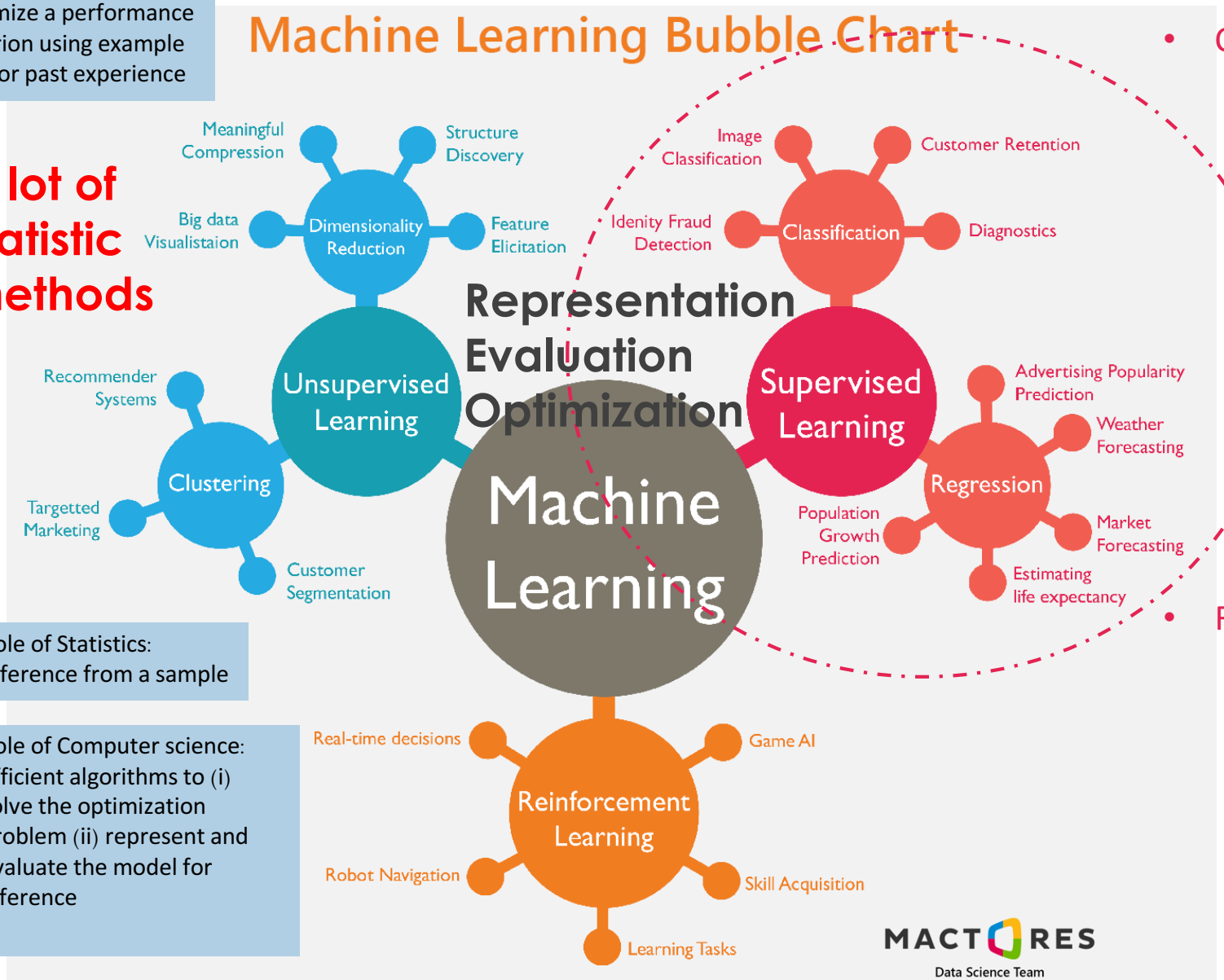
Going old school ?



Machine Learning Techniques

Optimize a performance criterion using example data or past experience

A lot of statistic methods

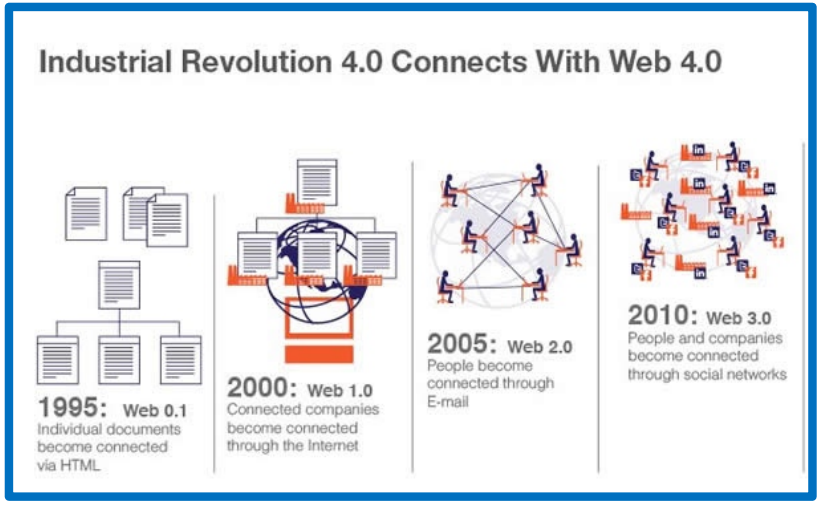


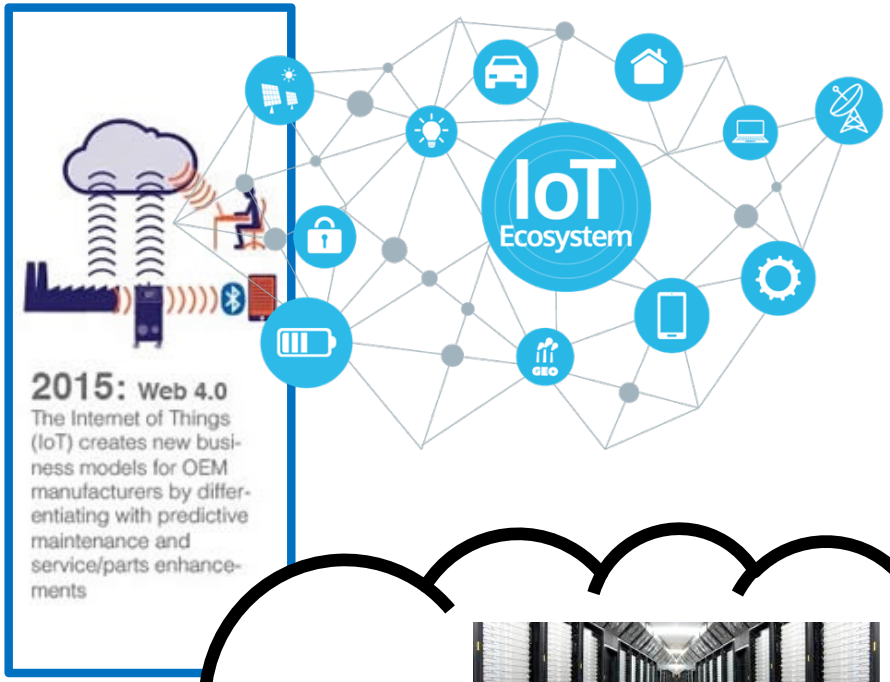
Role of Statistics: Inference from a sample

Role of Computer science: Efficient algorithms to (i) solve the optimization problem (ii) represent and evaluate the model for inference

- Classification
  - Logic
  - SVM
  - Random Forest
  - Hidden Markov
  - …
- Regression
  - Lasso
  - Ridge
  - Loes
  - KNN
  - Spline
  - XGBoost
  - …

Use the full power of the Internet!

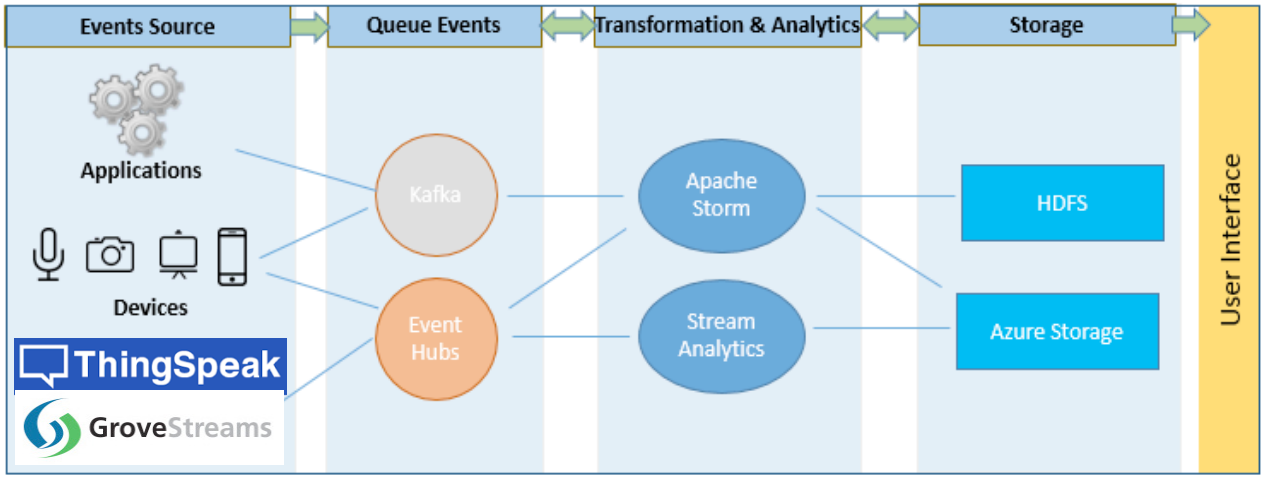




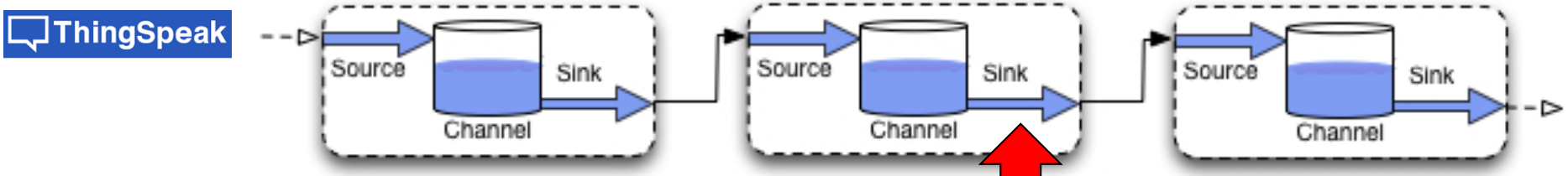
- IoT data are pushed on Internet data clouds
- Computing resources using Virtual Machines are obtained from Internet Computing clouds
- Parallel processing
- Optimized libraries
- Web tools to orchestrate

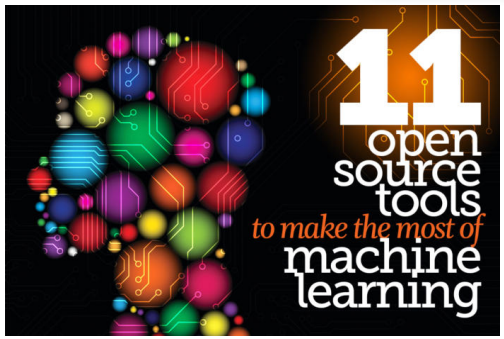
The Big Data landscape

Example: the APACHE ecosystem



SensorCloud kafka FLUME

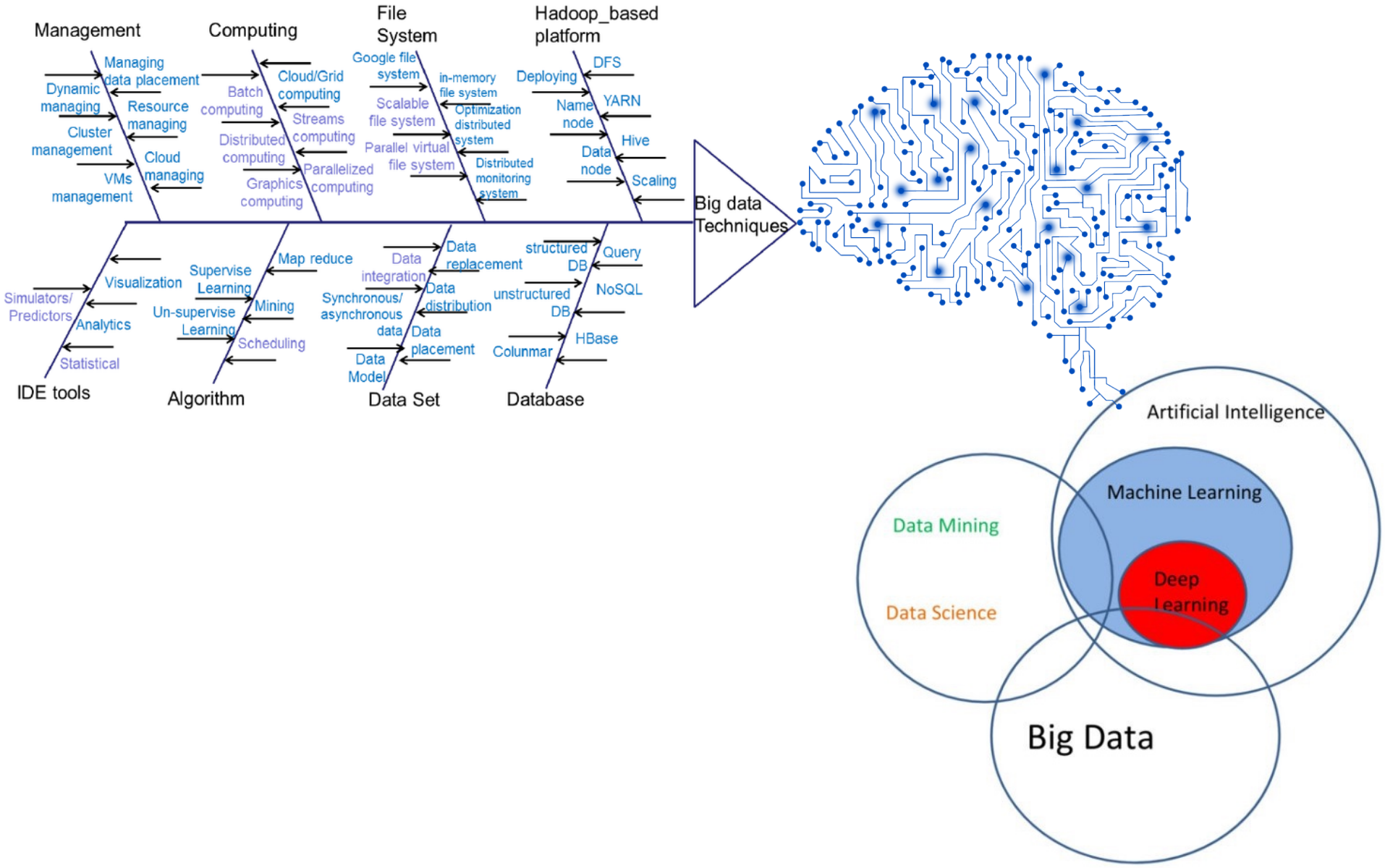




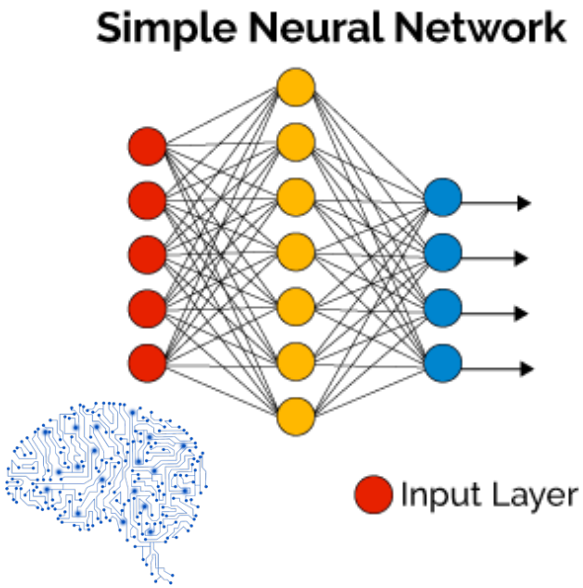
Google
Amazon
Microsoft
IBM
…

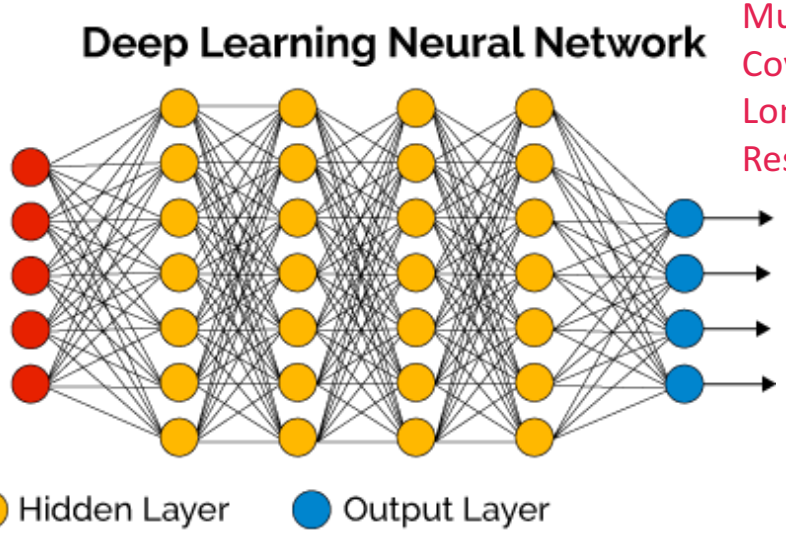
Advanced & customized data management

Beyonds Machine Learning?

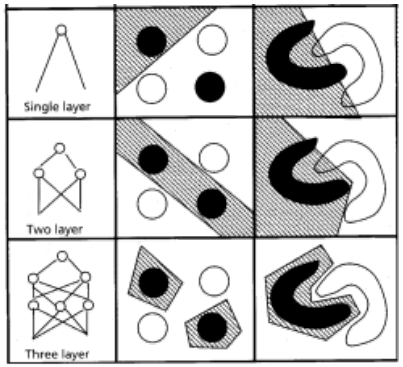


Deep Learning is essentially NN





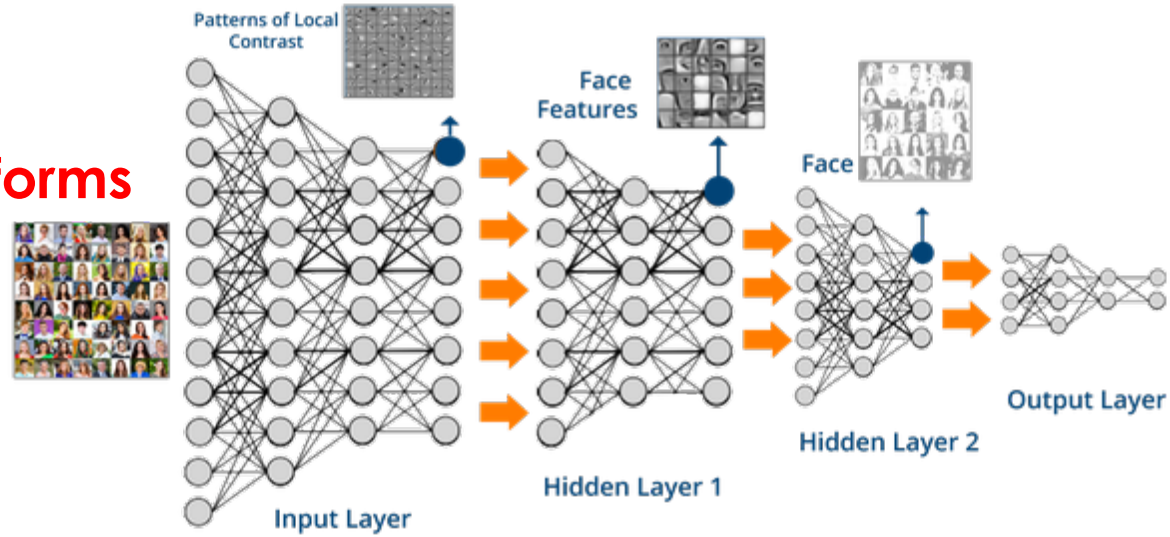
Multilayer Perceptron
Covolutional Neural Nets
Long Short–Term Memory
Restricted Boltzmann Machine



(Jain, 1996)

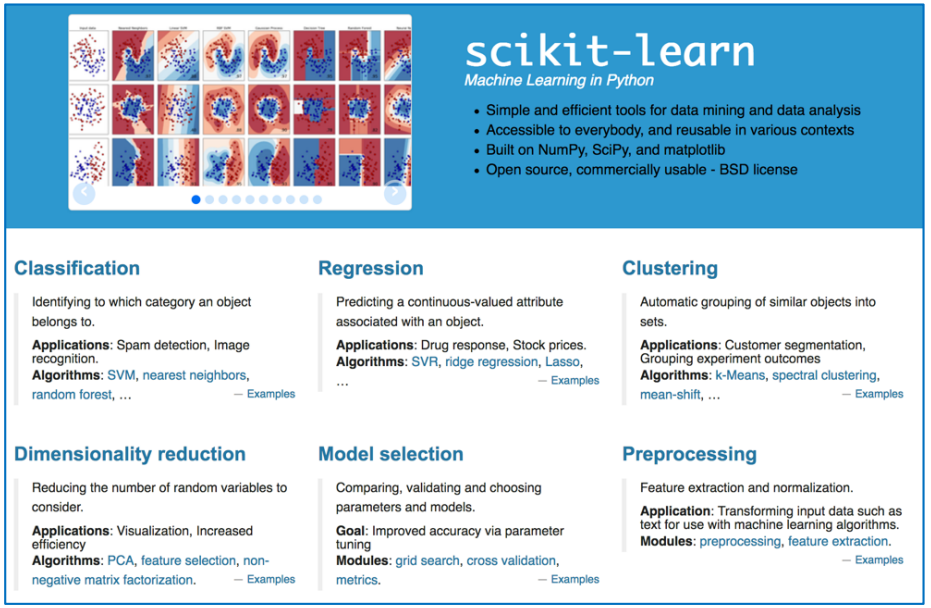
Voice/Face/Patterns recognition on many platforms

- **Facebook**
- **Google Photos**
- **Twitter**
- **Siri**
- **…**



Machine/Deep Learning for scientists

- Large variety of supported languages
  - Python, R, C++, Java, Scala, Javascript, Go, …
- Many statistical methods/algorithms are implemented in libraries
- Examples
  - Scikit-learn
  - Google TensorFlow
  - Microsoft Distributed Machine Learning Toolkit
  - Apache Mahout
  - …
- But, beware
  - There are hundredth of tools…
  - …and new tools every months!



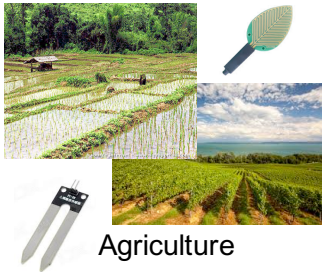
IoT for Development

Irrigation

Livestock farming

Fish farming & aquaculture

Storage & logistic



Agriculture

Environment

Needs, constraints, cost, design approach, control mechanism

Challenge 2: Bridging the digital divide

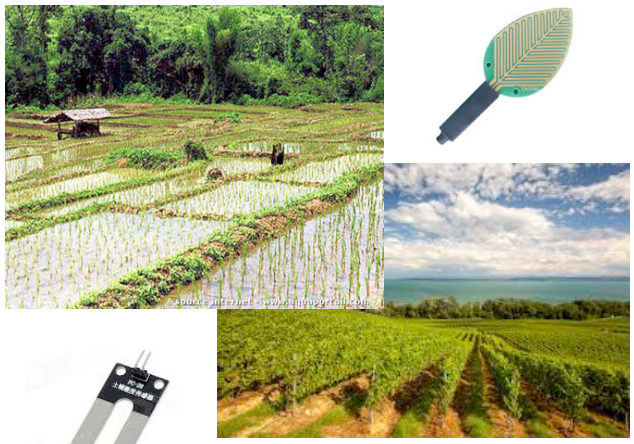
IoT4D: development for rural areas

Irrigation

Livestock farming

Fish farming & aquaculture

Storage & logistic



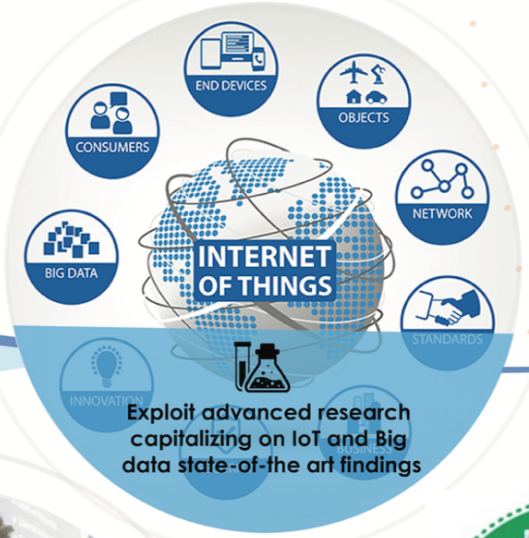
Agriculture

Environment

WAZIUP Open IoT and Big data platform for Africans, by Africans

HORIZON 2020

Affordable technologies to empower rural economics



Exploit advanced research capitalizing on IoT and Big data state-of-the art findings

Develop IoT solutions and applications meeting African needs

DO MORE with LESS

www.waziup.eu

Waziup IoT

Waziup IoT

Waziup

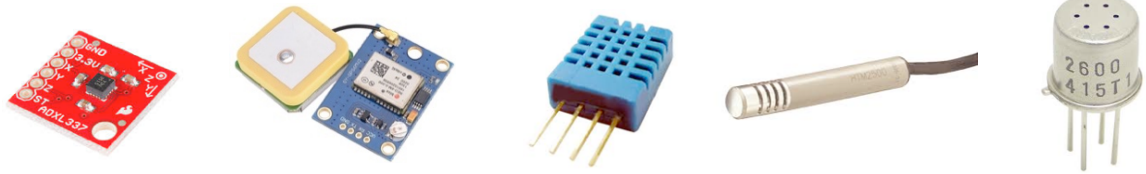
Waziup

CREATE-NET · CTIC · Farmerline · iSpace · coders africa · INNOTEC

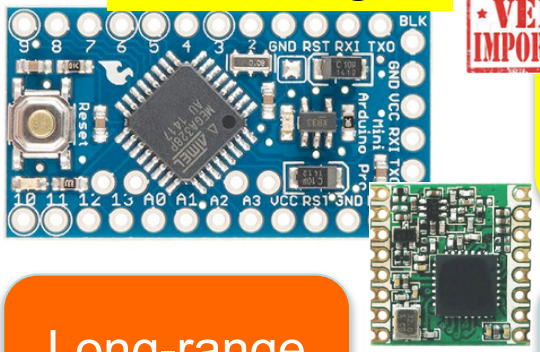
waziup.community@create-net.org

Ready-to-use templates

Physical sensor | Physical sensor | Physical sensor



Arduino Pro Mini @3.3V



Physical sensor mgmt

VERY IMPORTANT

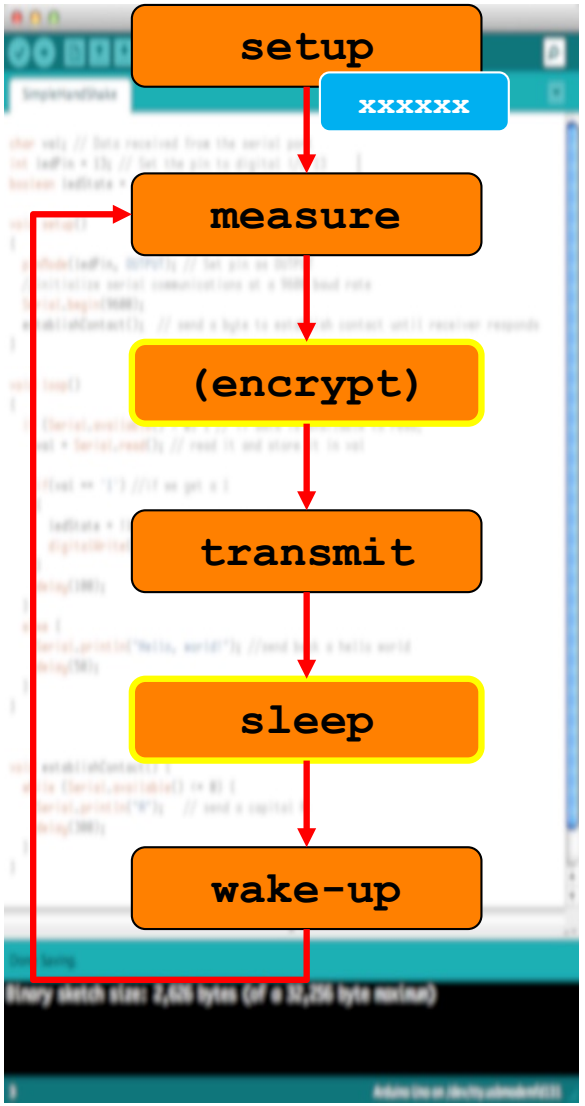
Activity duty-cycle, low power

VERY IMPORTANT

AES encryption

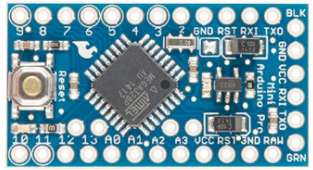
Long-range transmission

Logical sensor mgmt



A simple temperature sensor example

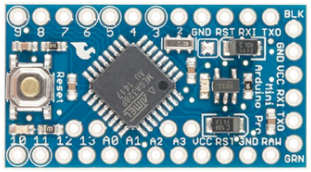
Arduino Pro Mini @3.3V



Modtronix inAir9

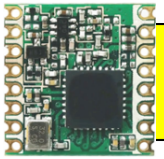
TMP36

Low-cost integration

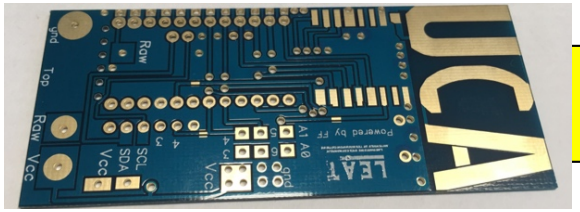


1.5€

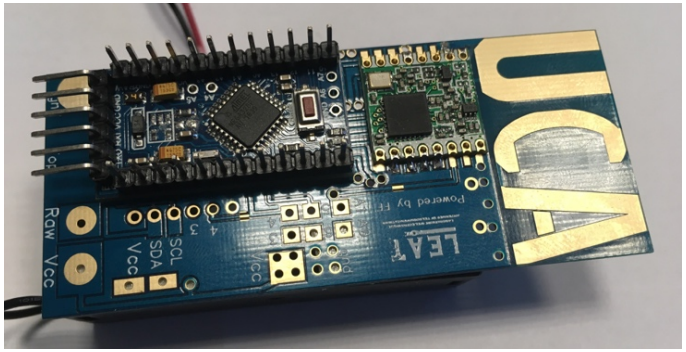
https://github.com/FabienFerrero/UCA_Board

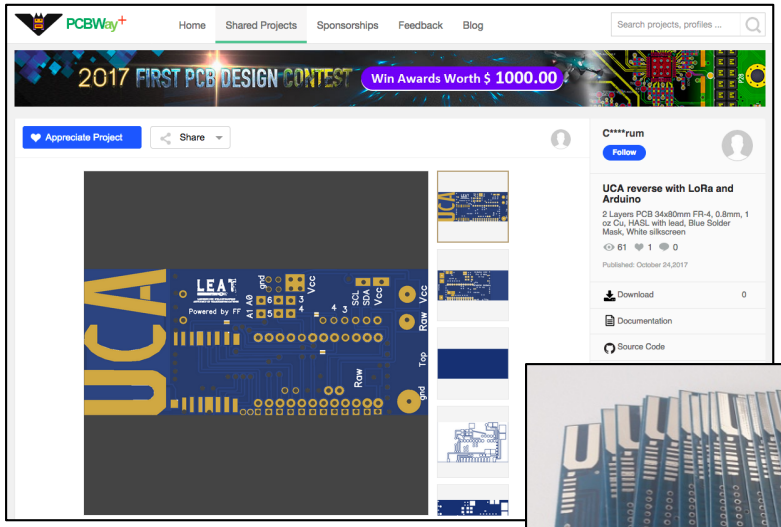


5€



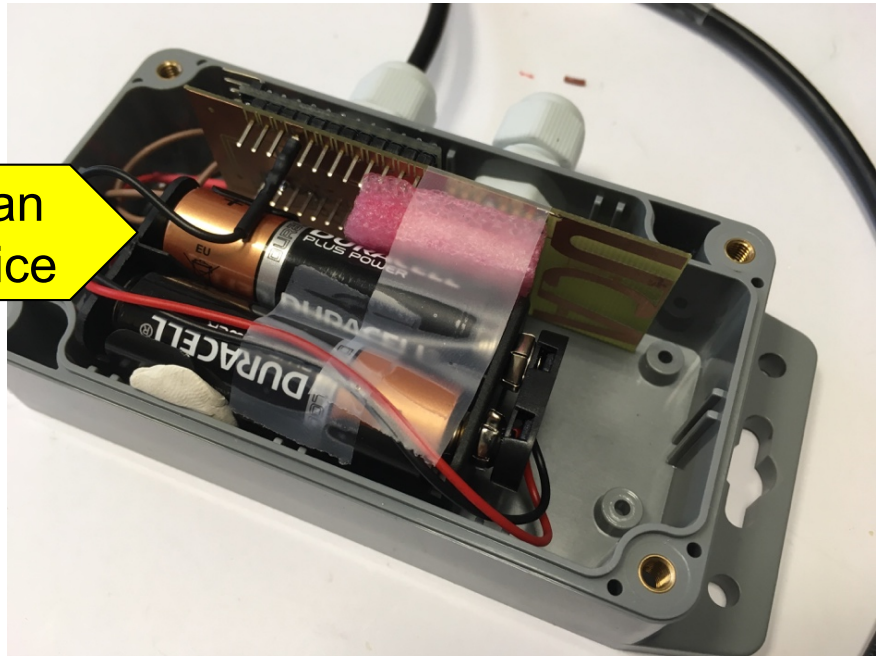
1€





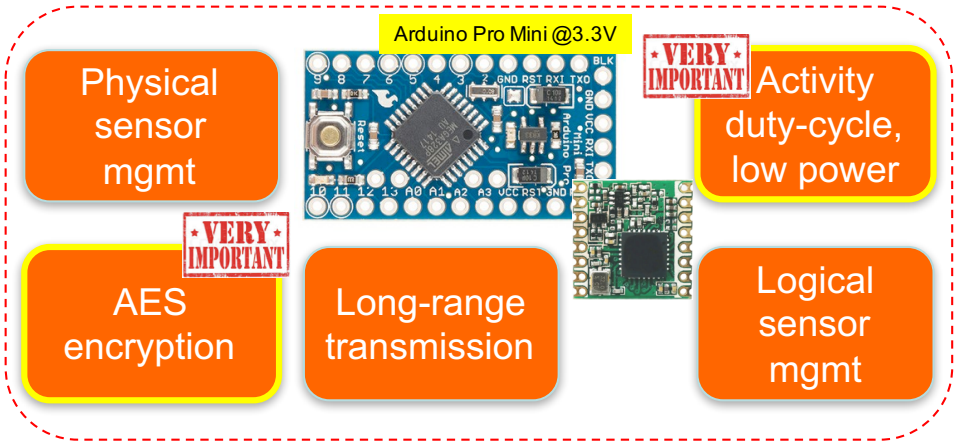
1-click order

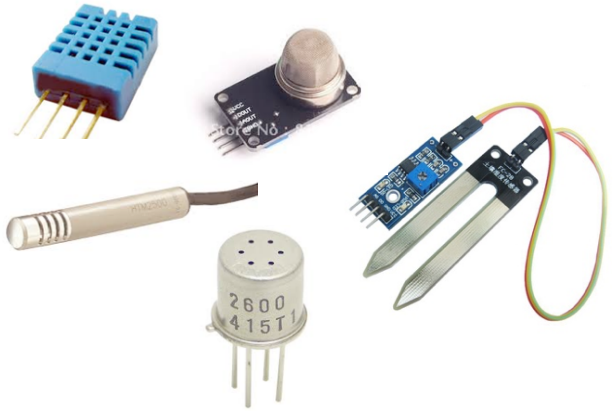
Less than 10€/device



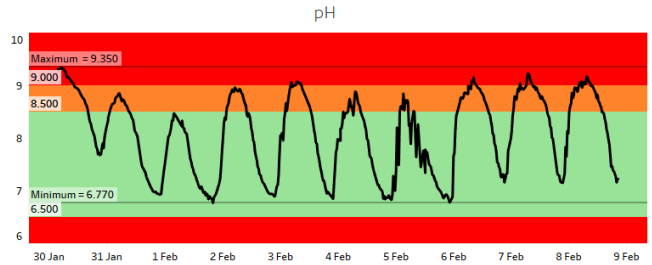
Generic sensing IoT device v.s. Highly specialized

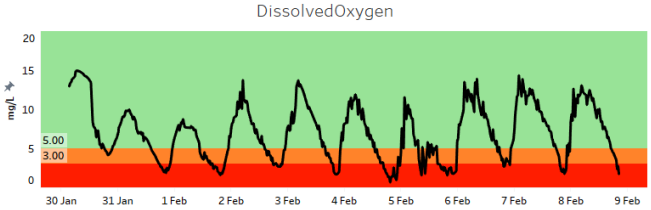
- Build low-cost, low-power, **long-range** enabled generic platform
- Methodology for low-cost platform design
- Technology transfers to user communities, economic actors, stakeholders,…

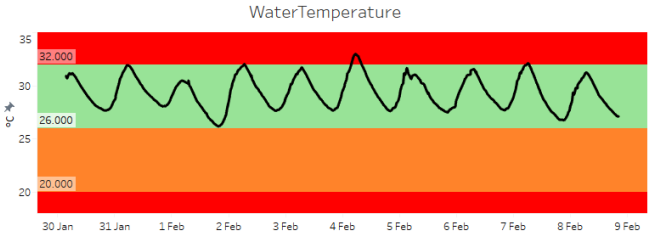




Low-cost buoy for fish farming MVP







Physical sensor reading

Credit: EGM

Physical sensor management

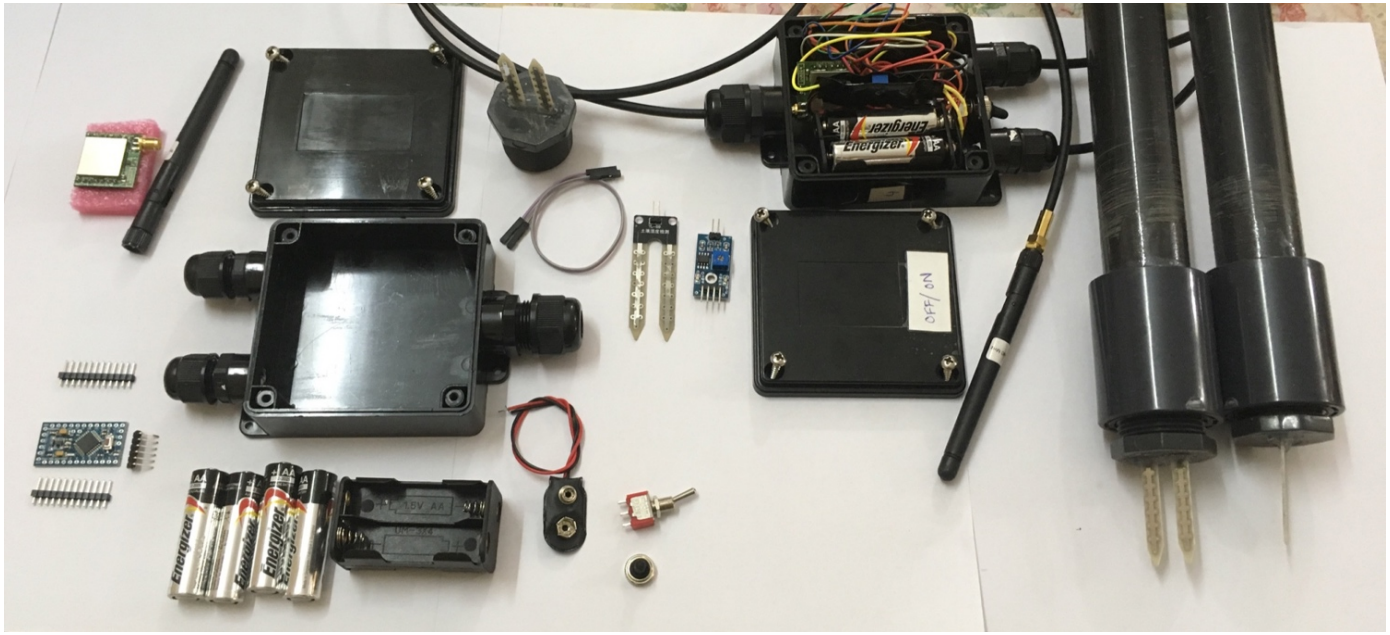
Activity duty-cycle, low power

Security

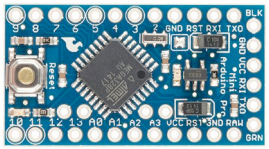
Long-range transmission

Logical sensor management

Soil humidity sensors for agri MVP



Physical sensor management



Activity duty-cycle, low power

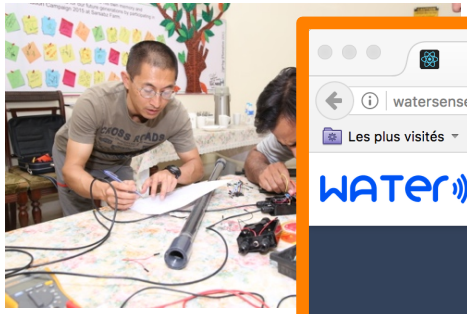
Security

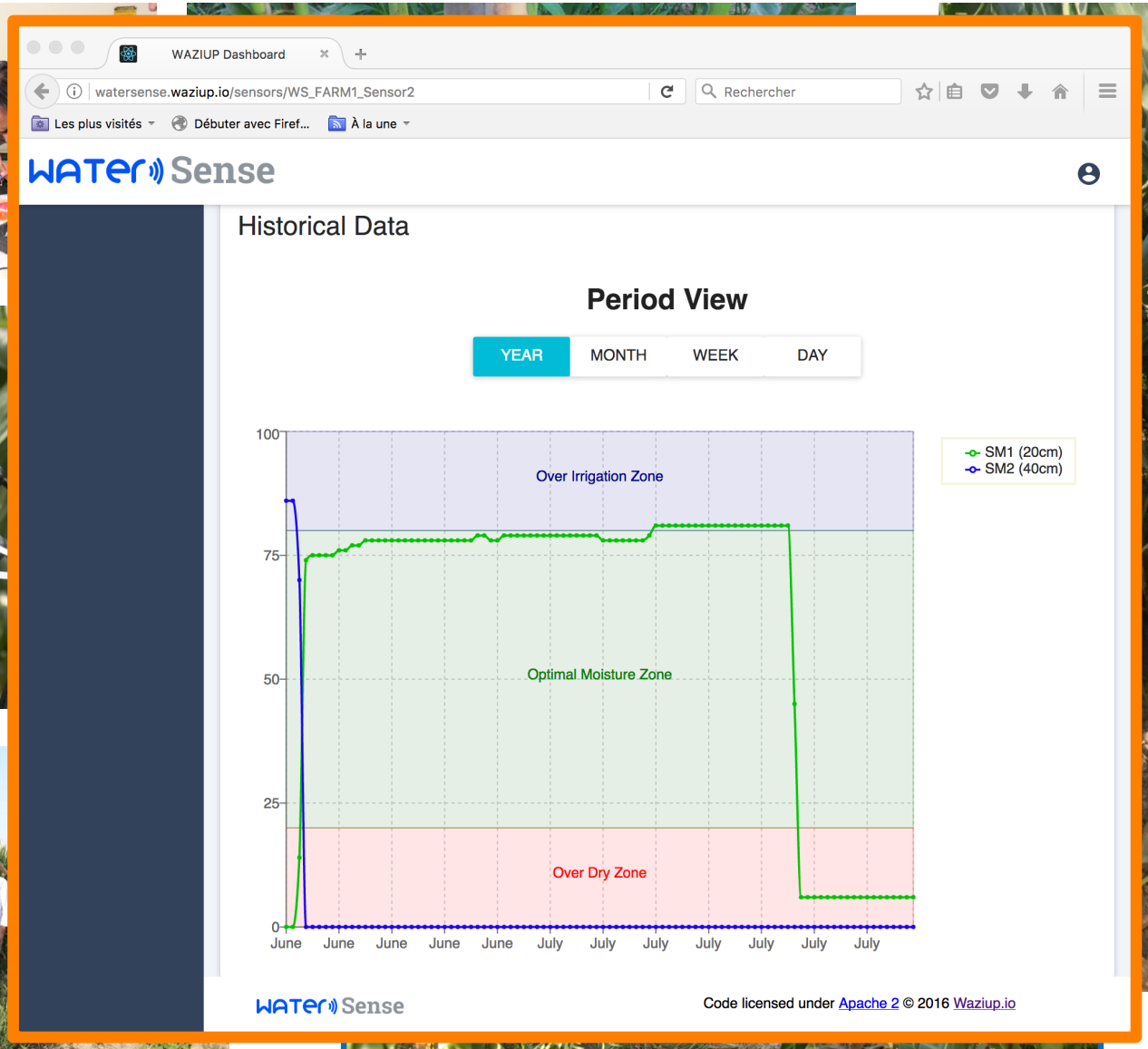
Long-range transmission

Logical sensor management

Deployment for Nestlé's WaterSense project

Deployment for Nestlé's WaterSense project





Local weather station for AGRI MVP

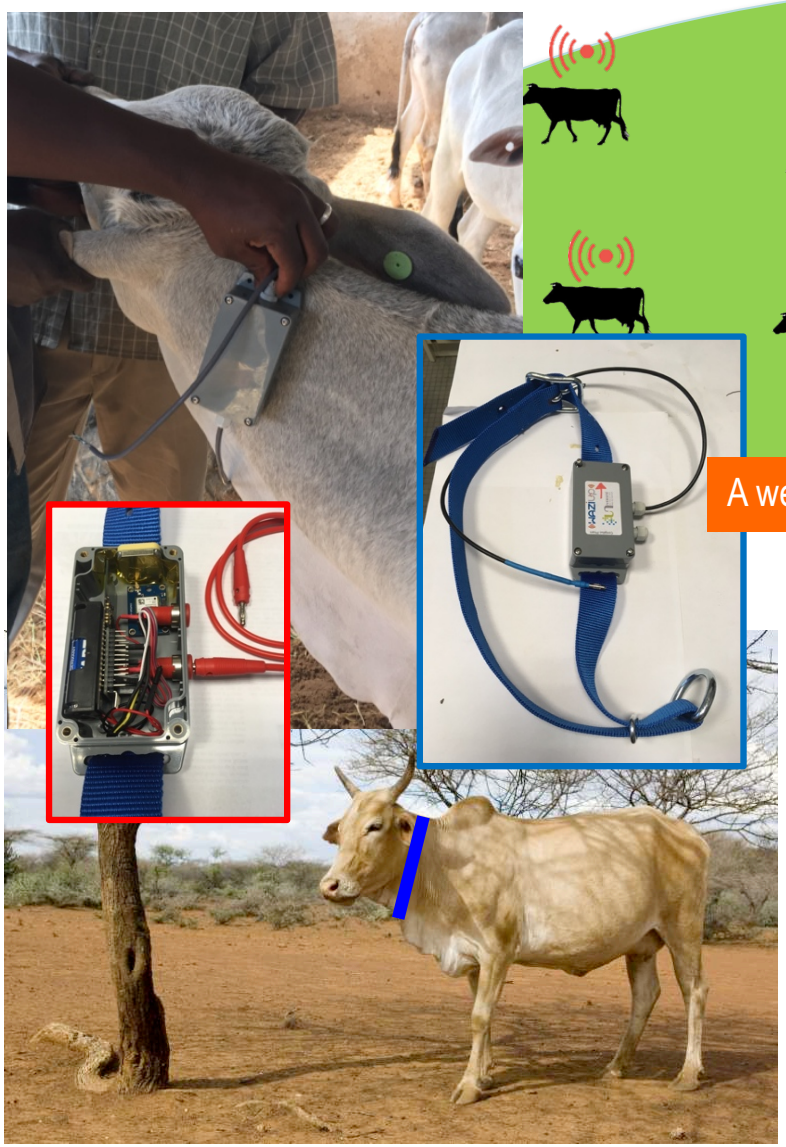
https://openweathermap.org/

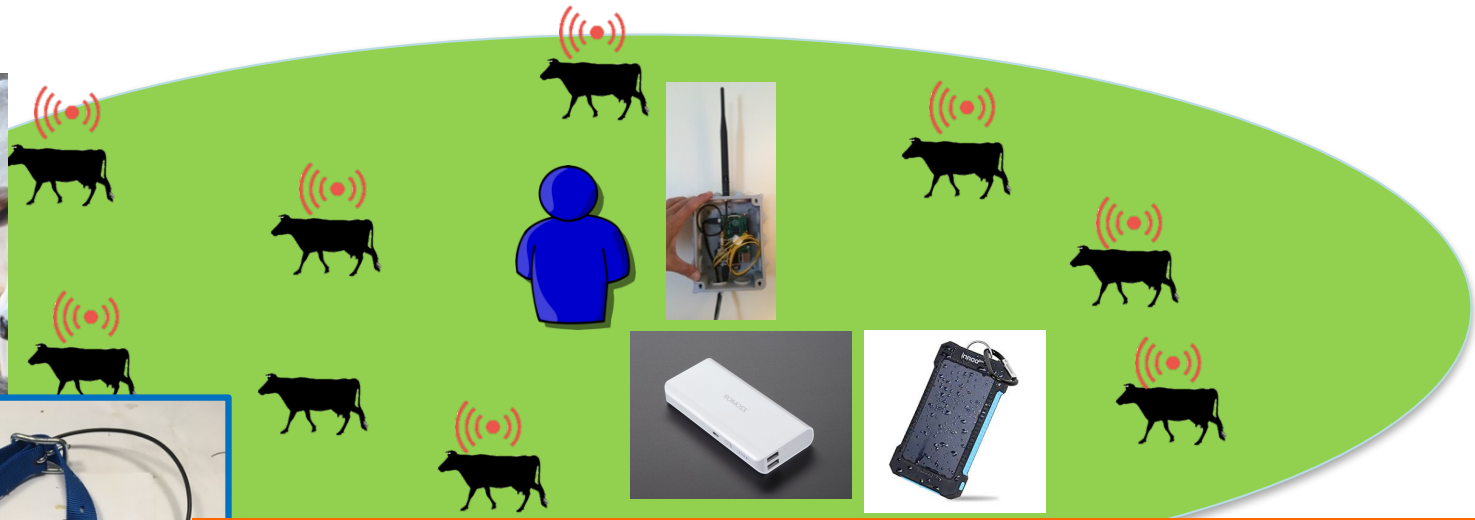
Photo from Unparallel

Get local weather measuments

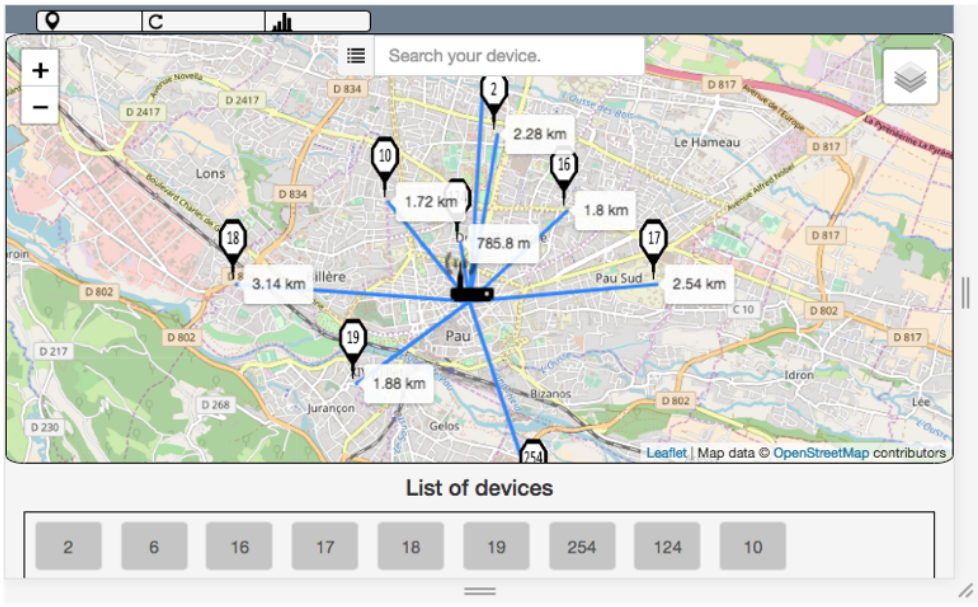
Combine with open weather data to get more accurate predictions

Collar for Cattle Rustling MVP



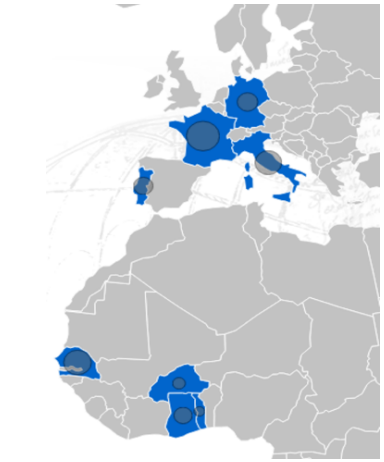


A web interface displays the position of the gateway those of the remote GPS devices



Scaling up!

Feb 2016 - 2019



May 2018 - 2021

Thanks.

Let's keep in touch

Soukeyna Wilma SOKENG
Communication & Event Manager

wilma.sokeng@cticdakar.com
www.cticdakar.com

contact@cticdakar.com

HORIZON 2020

WAZIUP

facebook.com/waziuploT

twitter.com/waziuploT

linkedin.com/groups/8156933

github.com/waziup