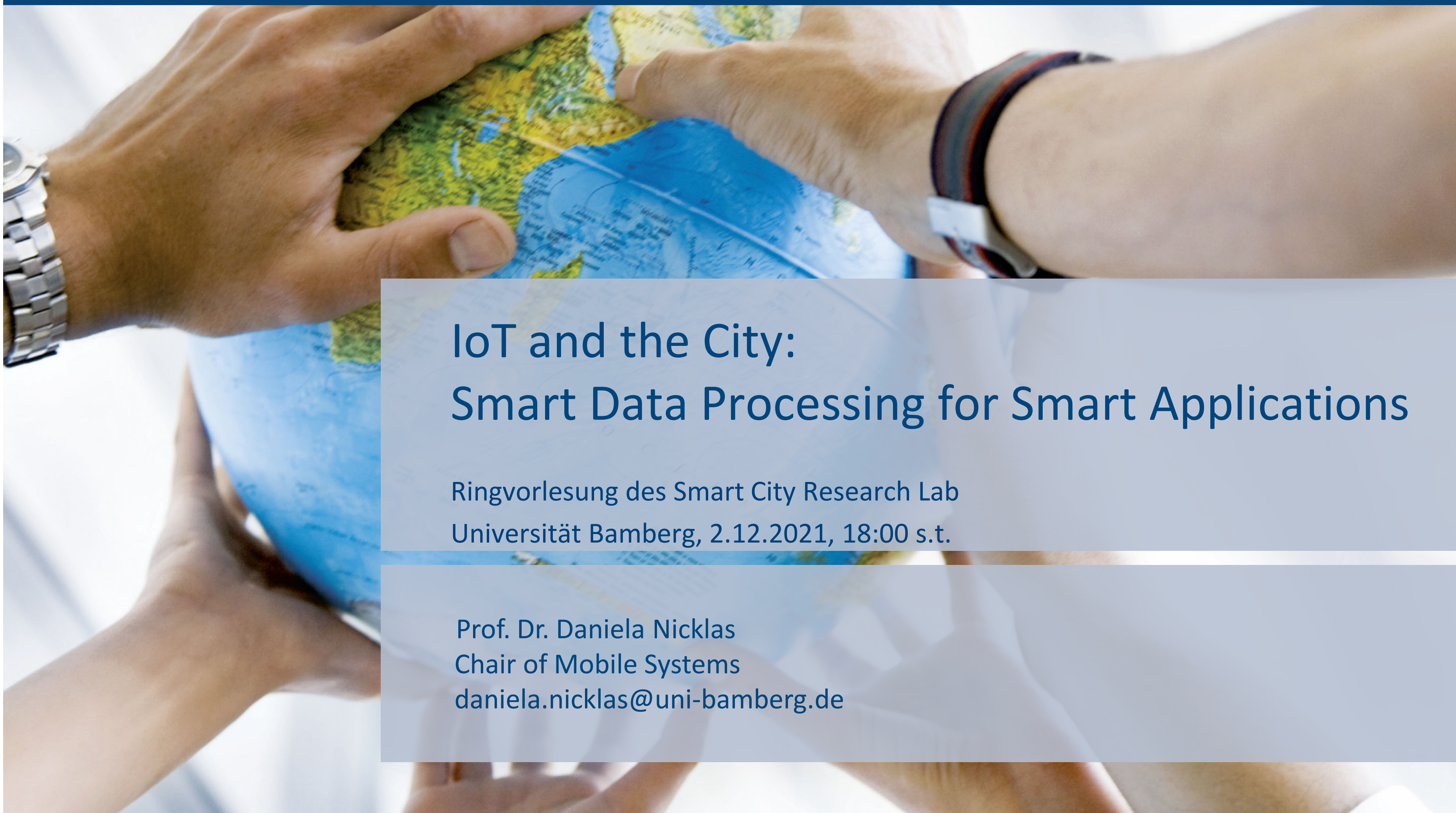
Universität Bamberg



IoT and the City: Smart Data Processing for Smart Applications

Ringvorlesung des Smart City Research Lab

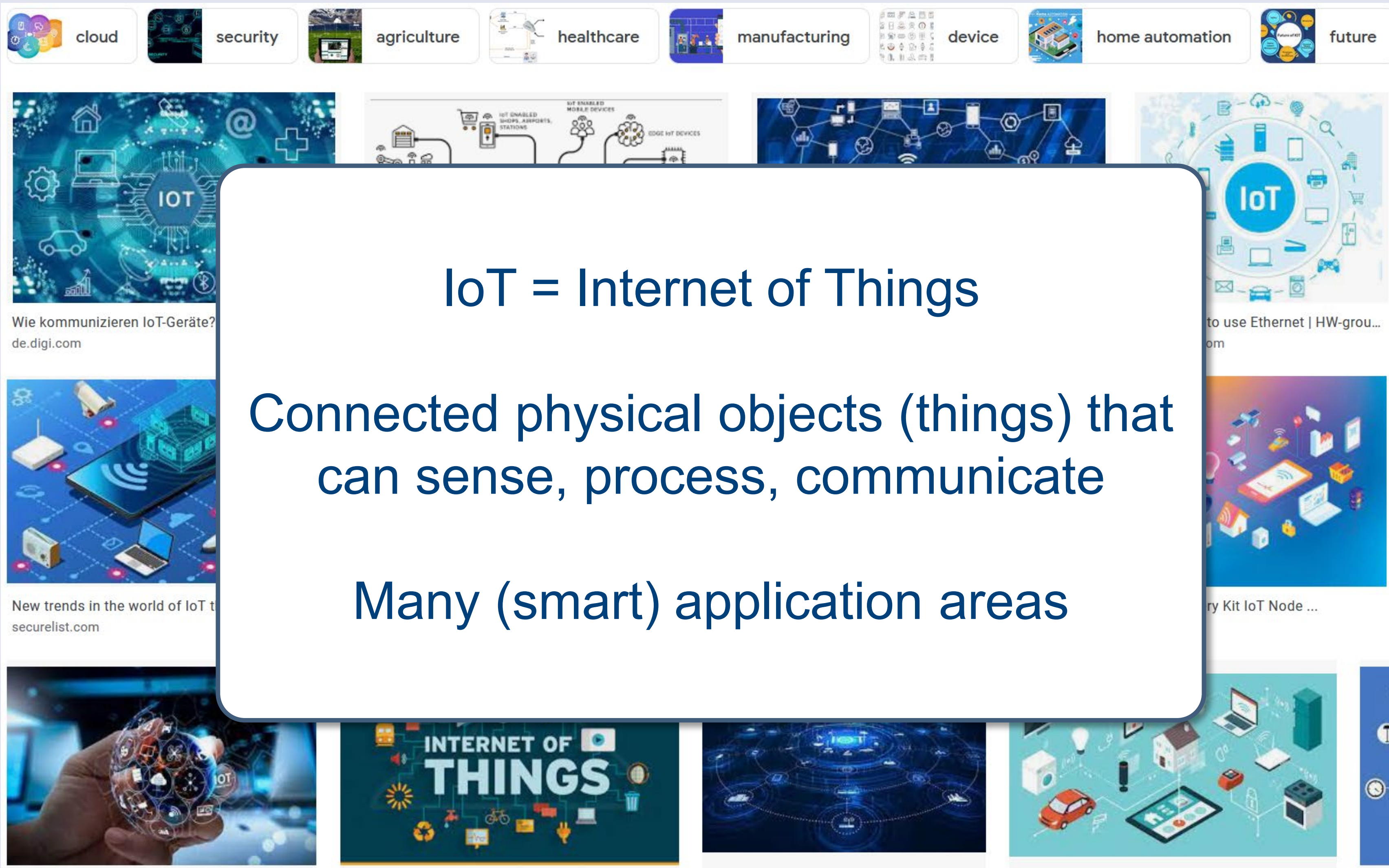
Universität Bamberg, 2.12.2021, 18:00 s.t.

Prof. Dr. Daniela Nicklas
Chair of Mobile Systems
daniela.nicklas@uni-bamberg.de

OK Google – What is IoT?

Universität Bamberg

https://www.google.com/search?q=iot (images)



IoT = Internet of Things

Connected physical objects (things) that can sense, process, communicate

Many (smart) application areas

IoT Applications

Universität Bamberg

for X in ...

- Meter
- Grid
- Factory
- Home
- City
- Phone
- Transportation

X

+ sensors

+ magic

= SmartX

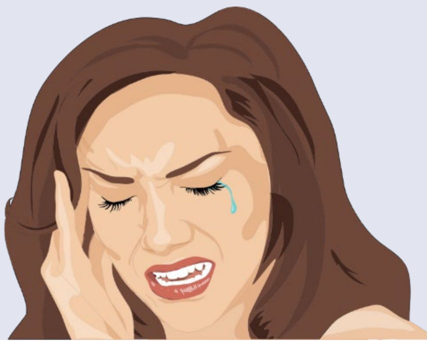
Realized by: IoT

Similar challenges in pervasive computing, ambient intelligence, physical computing, ...

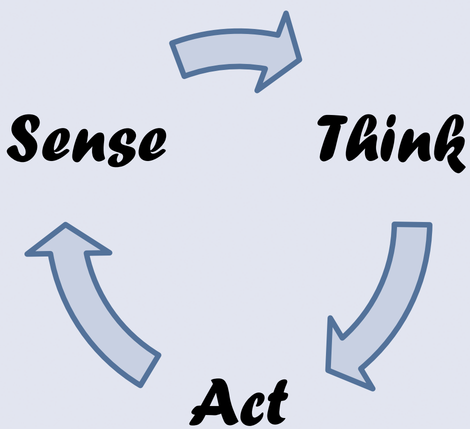
What is „smart“?

From the Wiktionary [1]

- Exhibiting social ability or cleverness
- Exhibiting intellectual knowledge
- Equipped with intelligent behaviour
  - Ex.: smart car, smartcard, smartphone
- … (Causing sharp pain; stinging)



Etymology:
From Middle English *smerten*, from Old English *smeortan* (“to smart”), from Proto-Germanic **smertaną* (“to hurt, ache”), from Proto-Indo-European **(s)merd-* (“to bite, sting”). [1]



[1] https://en.wiktionary.org/wiki/smart

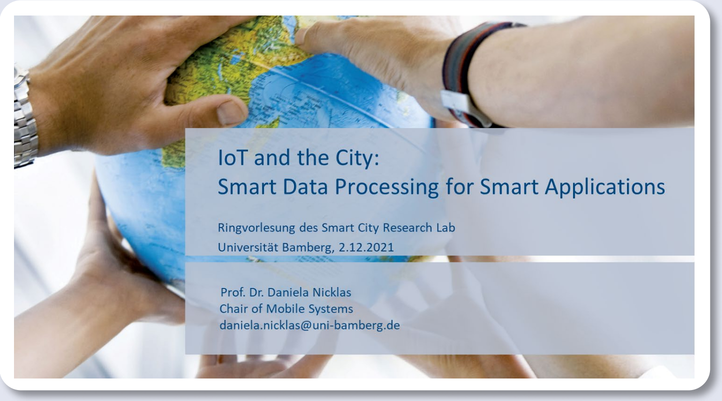
Overview

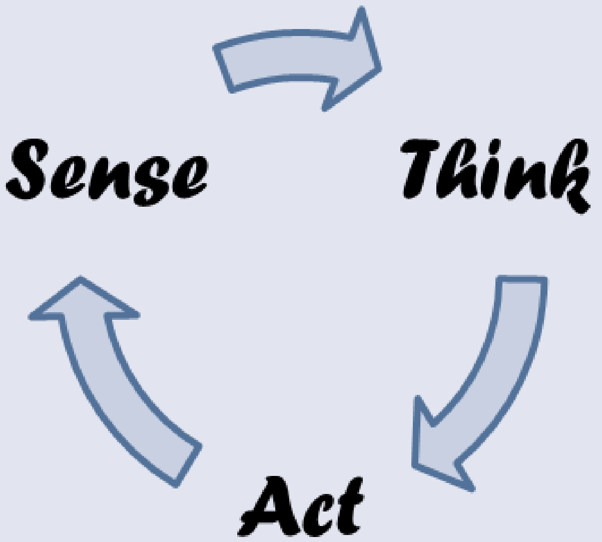
What is IoT? What is Smart?

Smart …

- City
- Applications
- Data Processing

- Conclusion
- Discussion!





Lasse Gerrits, „Chancen und Risiken der Smart City“, 19.11.2022, Highly recommended!
Ringvorlesung Smart City Research Lab
https://www.youtube.com/watch?v=9uPdvuFOqS0&t=1227s

Smart City Bamberg

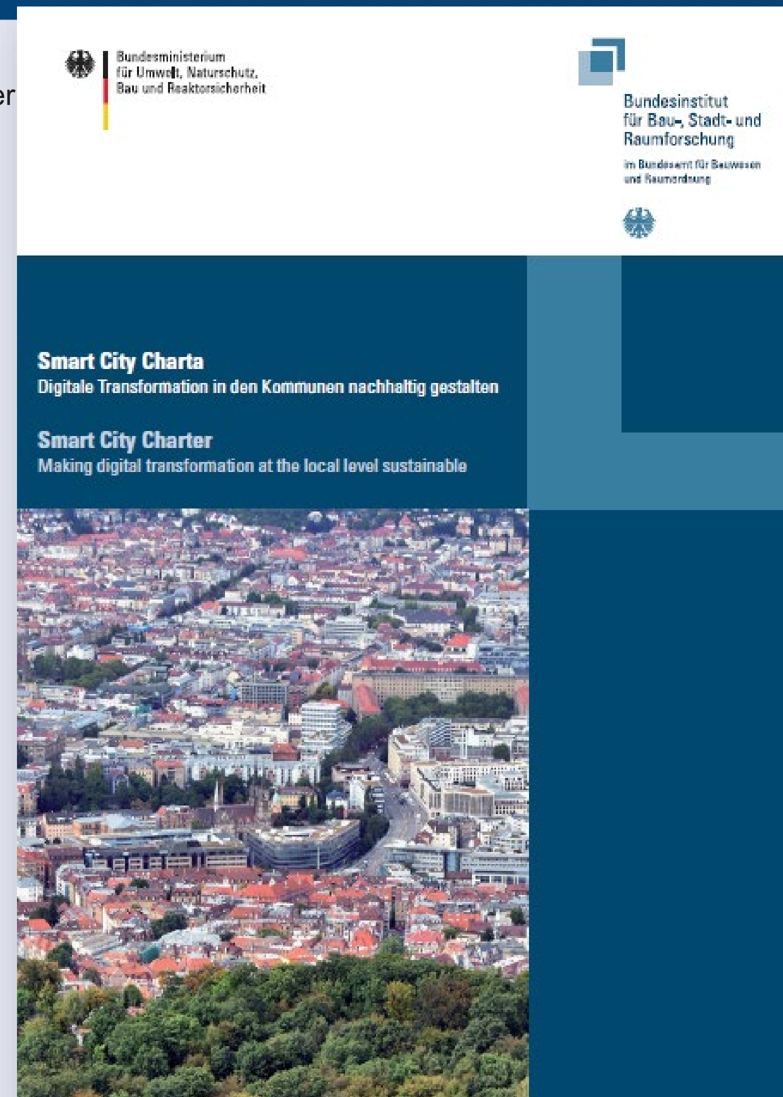
Universität Bamberg

- Smart City Bamberg: 7 year funding (2+5) according to the Smart City Charta (BMI)

A Smart City is …

Digital transformation requires cities, counties and municipalities to be **open to new technologies**, and to be aware of their broader values and goals in order to be able to apply those technologies with a **long-term and considered view**. Participants of the Smart Cities Dialogue Platform co-determine the normative image of the **intelligent, future-oriented** municipality.

- **liveable and lovable**
- **diverse and open**
- **participatory and inclusive**
- **climate-neutral and resource-efficient**
- **competitive and thriving**
- **open-minded and innovative**
- **responsive and sensitive:** it uses sensors, data acquisition and processing as well as new forms of interaction and learning to continuously improve local processes and services
- **safe and freedom-enhancing**



Smart City Bamberg and SC Research Lab

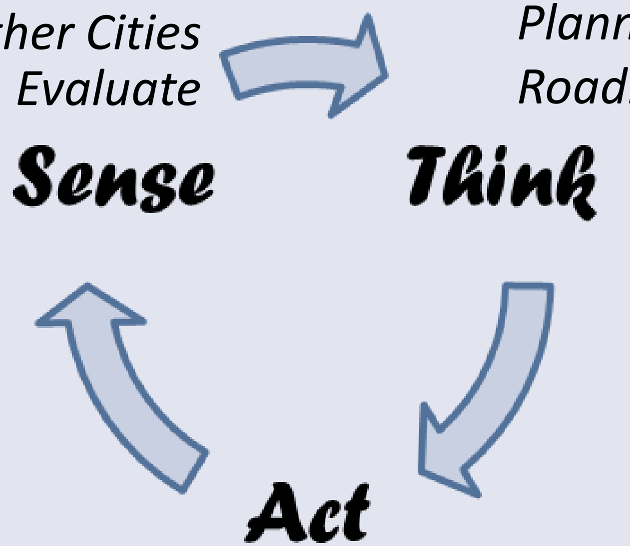
Universität Bamberg

- Smart City Bamberg: 7 year funding (2+5) according to the Smart City Charta (BMI)

- Smart City Research Lab: Interdisciplinary Research Network at the University of Bamberg

Citizen Participation!
Open Space
Interviews, Studies
Learn from other Cities
Evaluate

Discussion
Hackathon
Planning
Roadmapping



Funded Projects
Smart Applications

Overview

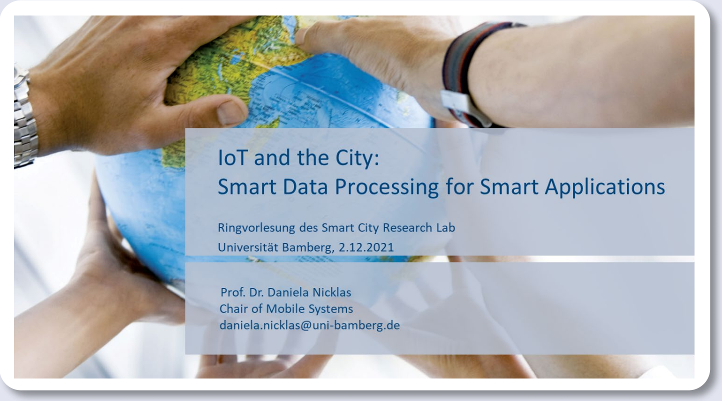
Universität Bamberg

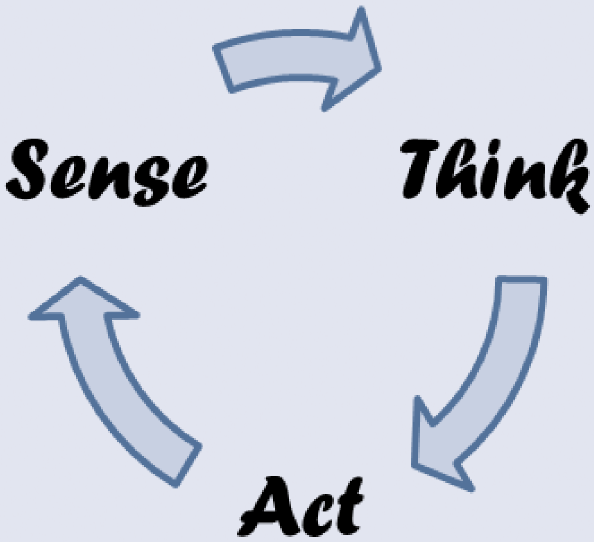
What is IoT? What is Smart?

Smart …

- City
- Applications
- Data Processing

- Conclusion
- Discussion!

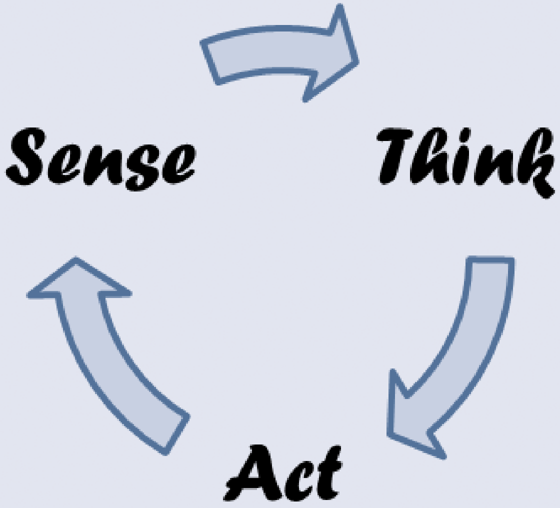




Smart Applications

Universität Bamberg

- From the charta: „**responsive and sensitive:** it uses sensors, data acquisition and processing as well as new forms of interaction and learning to continuously improve local processes and services“
- AKA „Context-aware application“: Know the context of the user (or the city) and adapt to it
- Example: Smart Traffic Lights
  - Sense: Traffic conditions of different road users (bicycles, cars, busses, pedestrians, …)
  - Think: Calculate optimal traffic flow
  - Act: Shorten or lengthen phases, synchronize traffic lights („grüne Welle“)

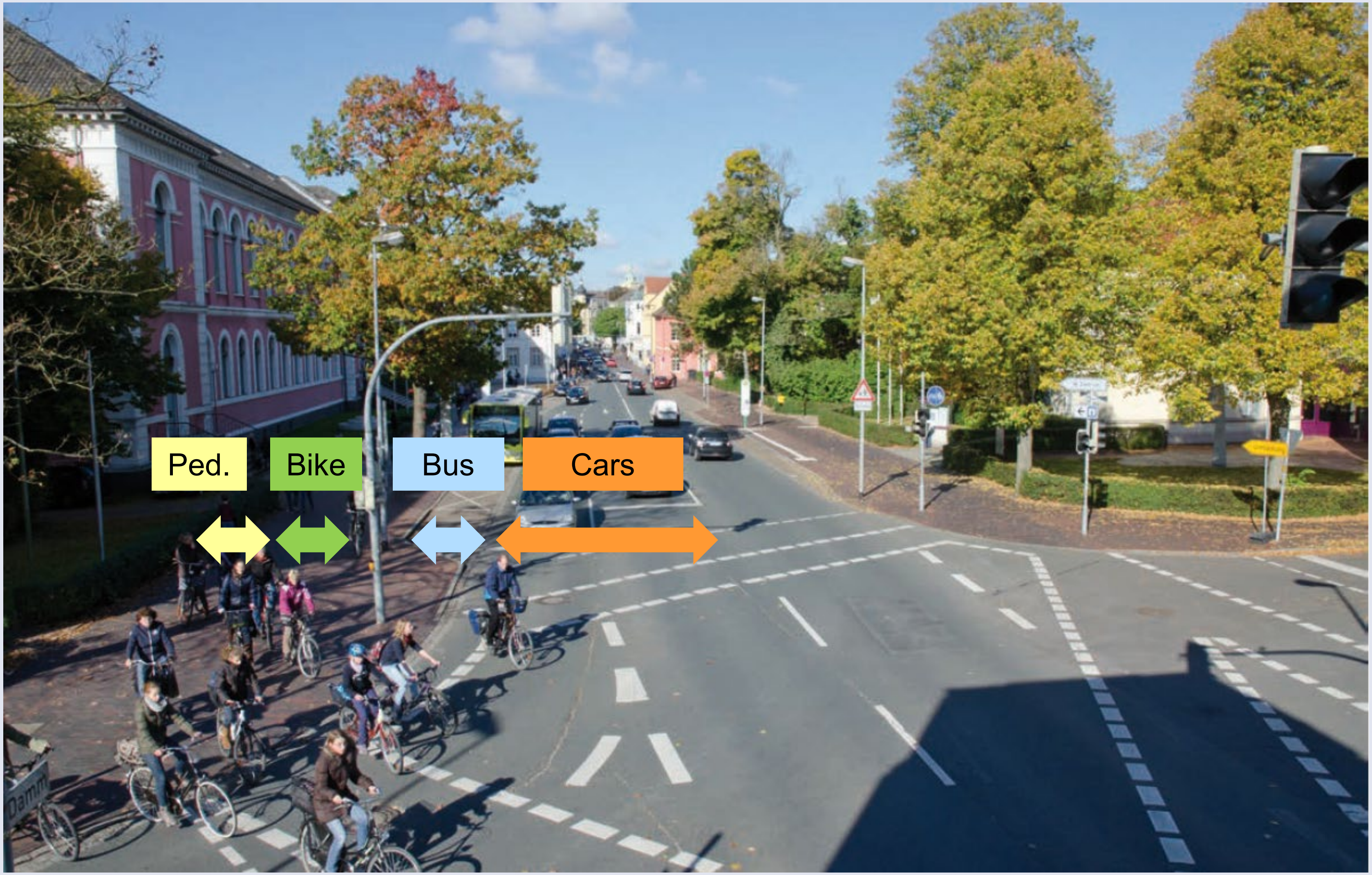


Smart Mobility: The Case of Oldenburg

nwzonline.de, Bild: Piet Meyer

Smart Mobility: The Case of Oldenburg

Universität Bamberg



[1] Strategieplan Mobilität und Verkehr 2025; Foto: www.peterduddek.de

Smart Mobility: The Case of Oldenburg

Universität Bamberg

Space
the final frontier

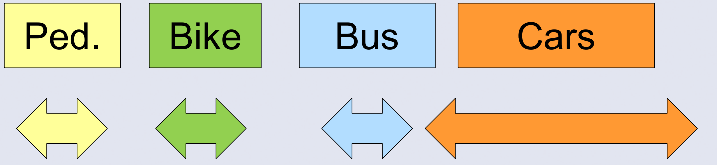
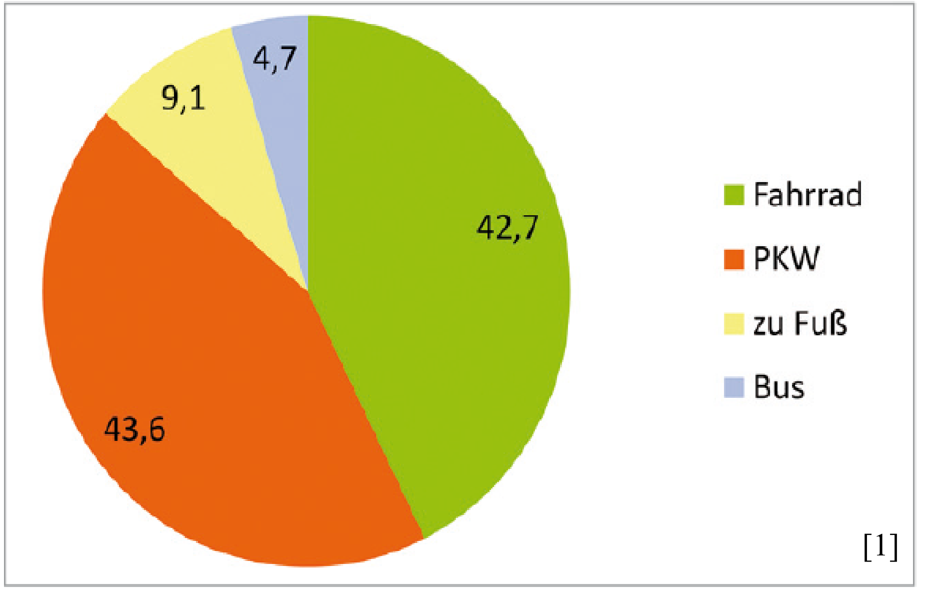


Abbildung 2: Verkehrsmittelwahl in Oldenburg
(% der Wege im Binnenverkehr, 2009)
Quelle: Stadt Oldenburg, Fachdienst Verkehrsplanung 2010



[1]

[1] Strategieplan Mobilität und Verkehr 2025; Foto: www.peterduddek.de

Space

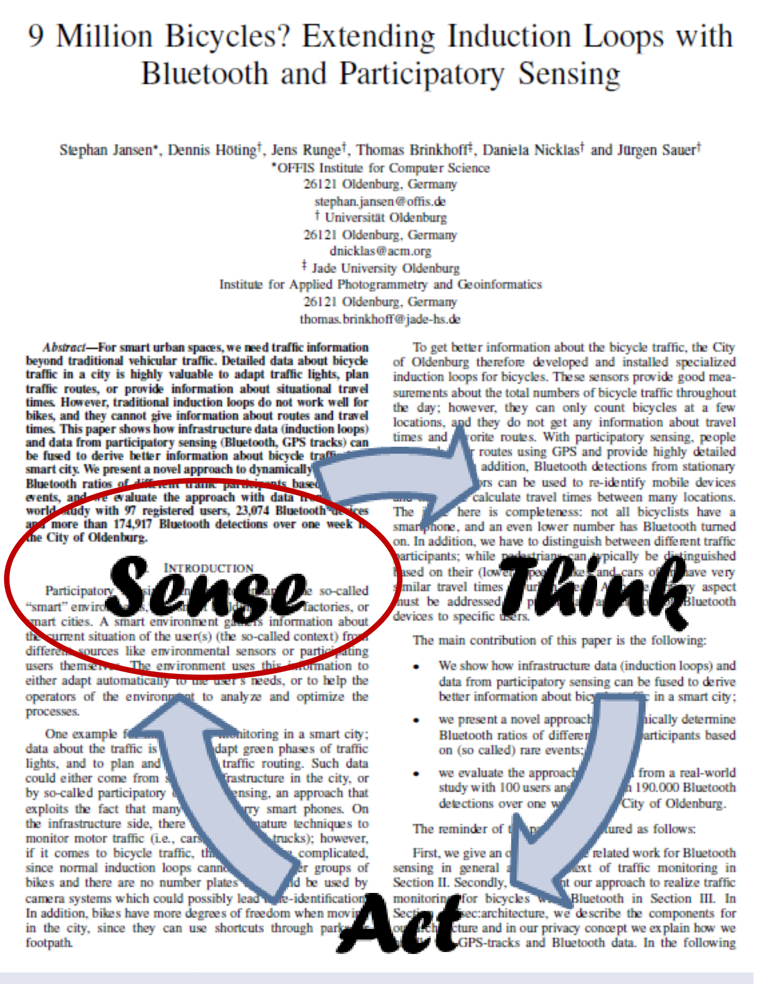
the final frontier

15.06.2018, www.infranken.de, Foto: Matthias Hoch

Smart Mobility: The Case of Oldenburg

Universität Bamberg

- Needed: More data about bicycle traffic!
  - Numbers (how many?)
  - Travel times (how fast?)
  - Routes (where?)
- Smart Sensing:
  - Combine information from different sources
    - Stationary Sensors
    - Mobile Sensors
    - „Donate your data“: Citizens, Participatory Sensing
  - Protect privacy
    - Consent + Remove sensitive parts
  - Low-energy footprint



2014 IEEE 15th International Conference on Mobile Data Management

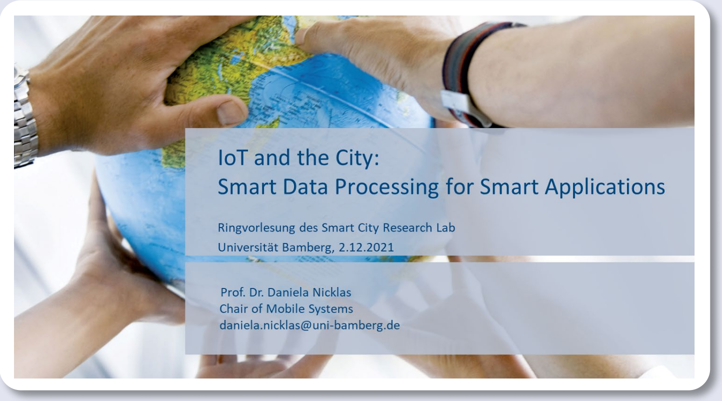
Overview

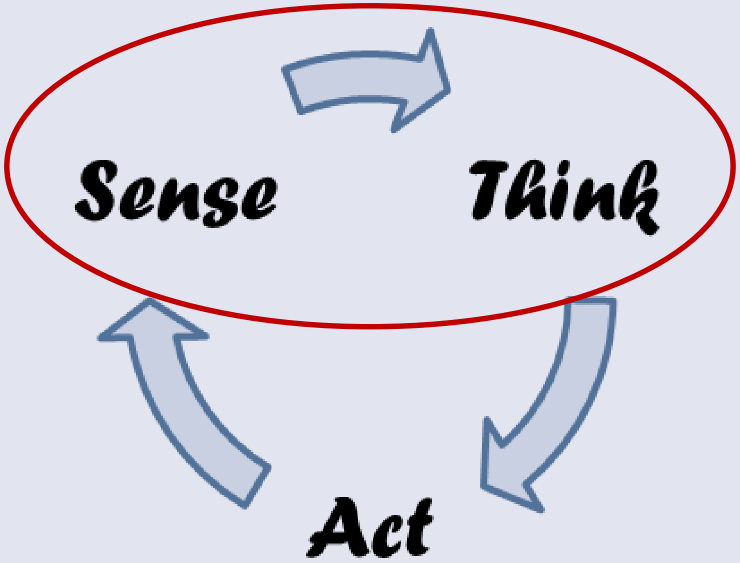
Universität Bamberg

What is IoT?

Smart …

- City
- Applications
- Data Processing
  - IoT Architectures
- Conclusion
- Discussion!

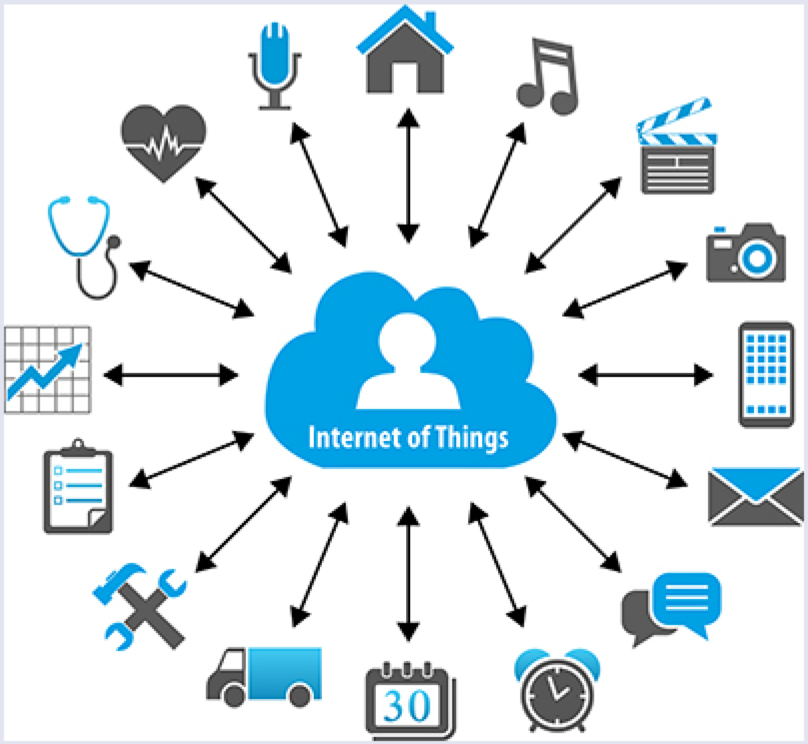




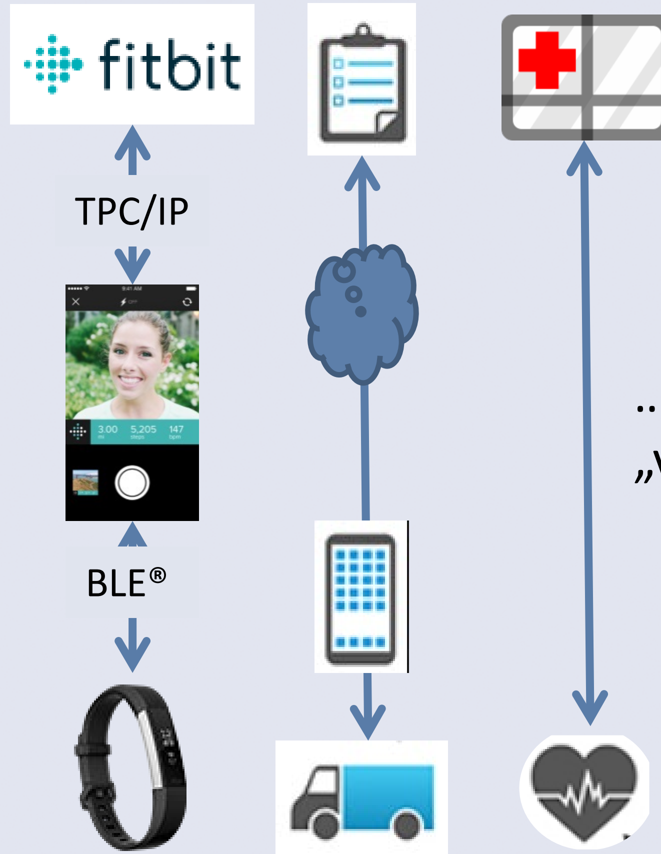
IoT Architecture

Universität Bamberg

Vision



Reality



… many „verticals“

Ok for single products,
not very smart for a city …

Building blocks of IoT

Universität Bamberg

- Identification: Naming and Adressing
  - Crucial for finding IoT systems and to match service with demands
- Sensing:
  - Awareness of the physical world
- Communication
  - Mostly wireless!
  - Tradeoff: Bandwidth / Energy consumption
- Computation
  - On device: „Edge“ and „Fog“ processing
- Service: Types of application
- Semantic: Understand data

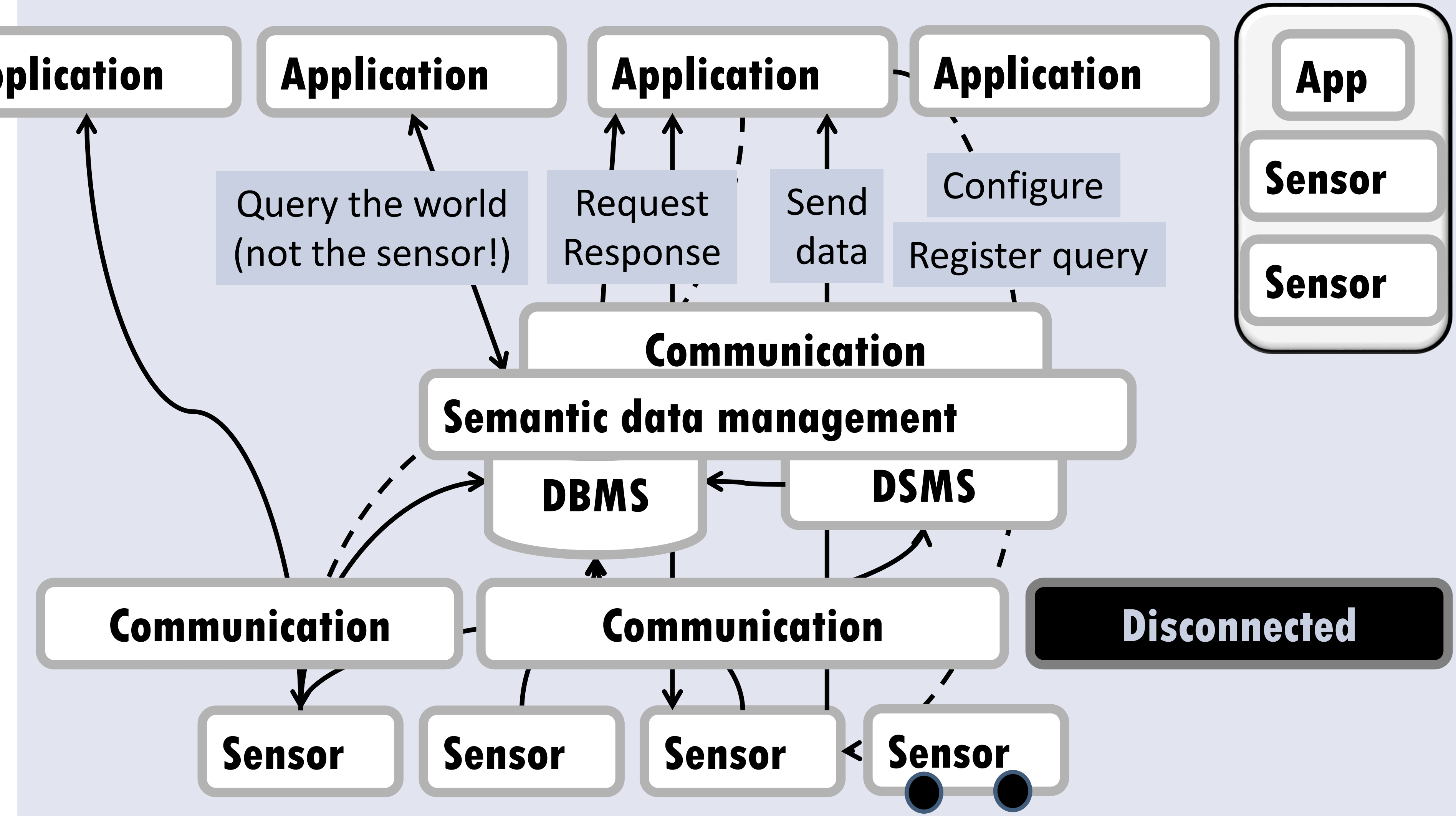
- Missing in [1]: Actuators
  - Do something! (e. g., open door, change traffic light …)

Sample technology for elements [1]

IoT Elements		Samples
Identification	Naming	EPC, uCode
	Addressing	IPv4, IPv6
Sensing		Smart Sensors, Wearable sensing devices, Embedded sensors, Actuators, RFID tag
Communication		RFID, NFC, UWB, Bluetooth, BLE, IEEE 802.15.4, Z-Wave, WiFi, WiFiDirect, , LTE-A
Computation	Hardware	SmartThings, Arduino, Phidgets, Intel Galileo, Raspberry Pi, Gadgeteer, BeagleBone, Cubieboard, Smart Phones
	Software	OS (Contiki, TinyOS, LiteOS, Riot OS, Android); Cloud (Nimbits, Hadoop, etc.)
Service		Identity-related (shipping), Information Aggregation (smart grid), Collaborative-Aware (smart home), Ubiquitous (smart city)
Semantic		RDF, OWL, EXI

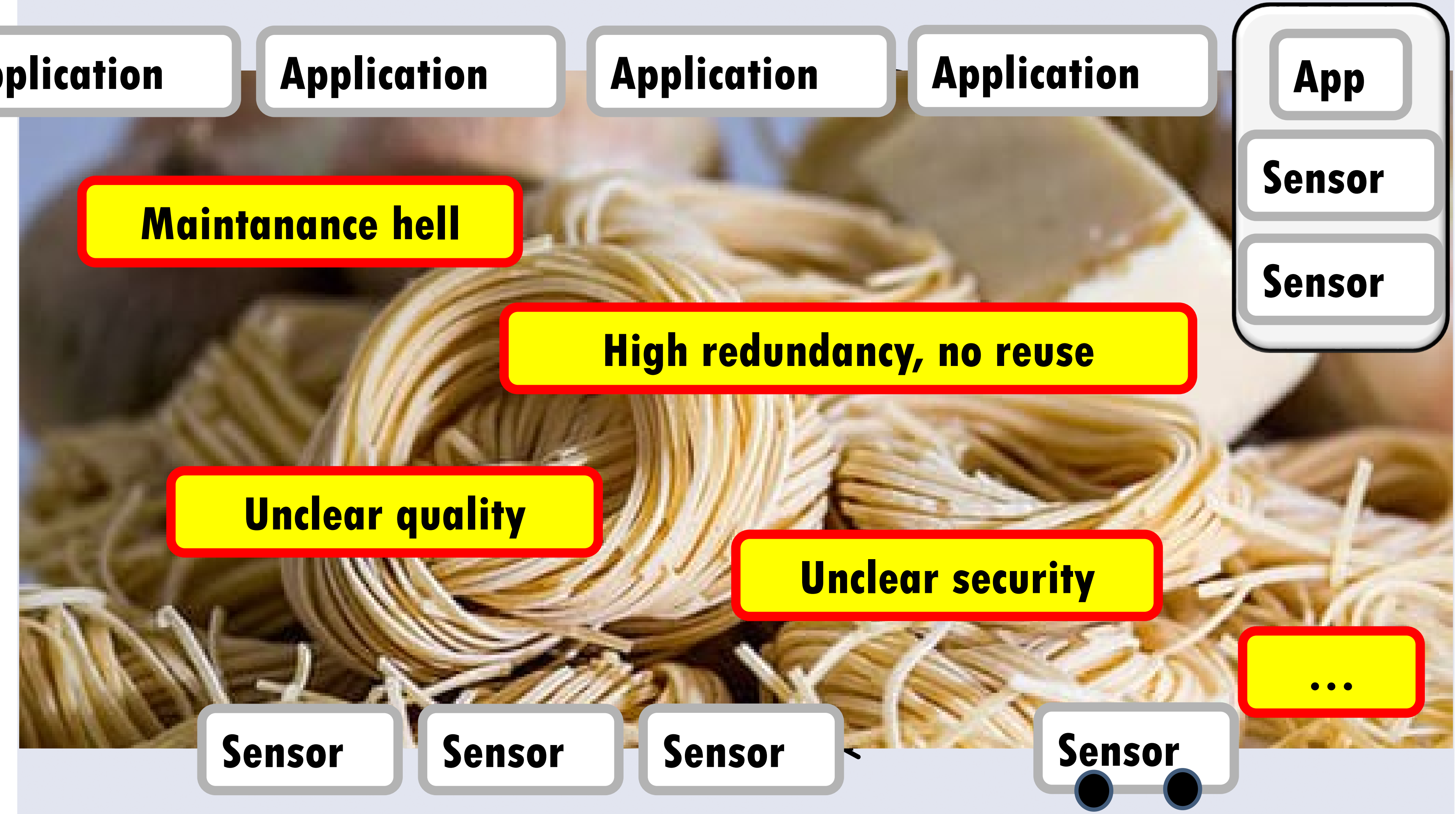
[1] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, und M. Ayyash, „Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications“, *IEEE Communications Surveys Tutorials*, Bd. 17, Nr. 4, S. 2347–2376, 2015

Evolution of a sensor-based system

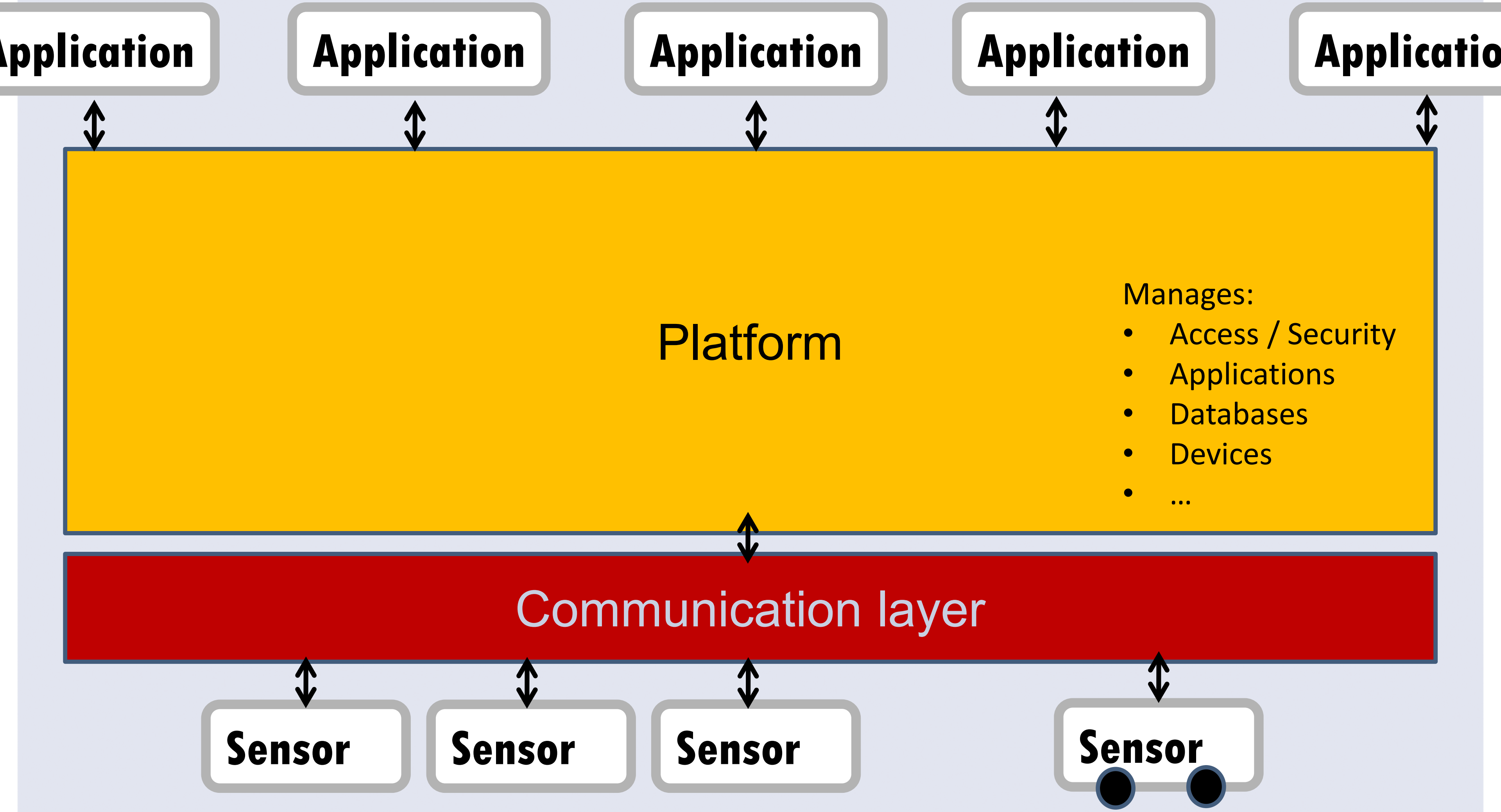


Evolution of a sensor-based system

Universität Bamberg



Use a platform architecture!



Example: City platform layer architecture

Universität Bamberg

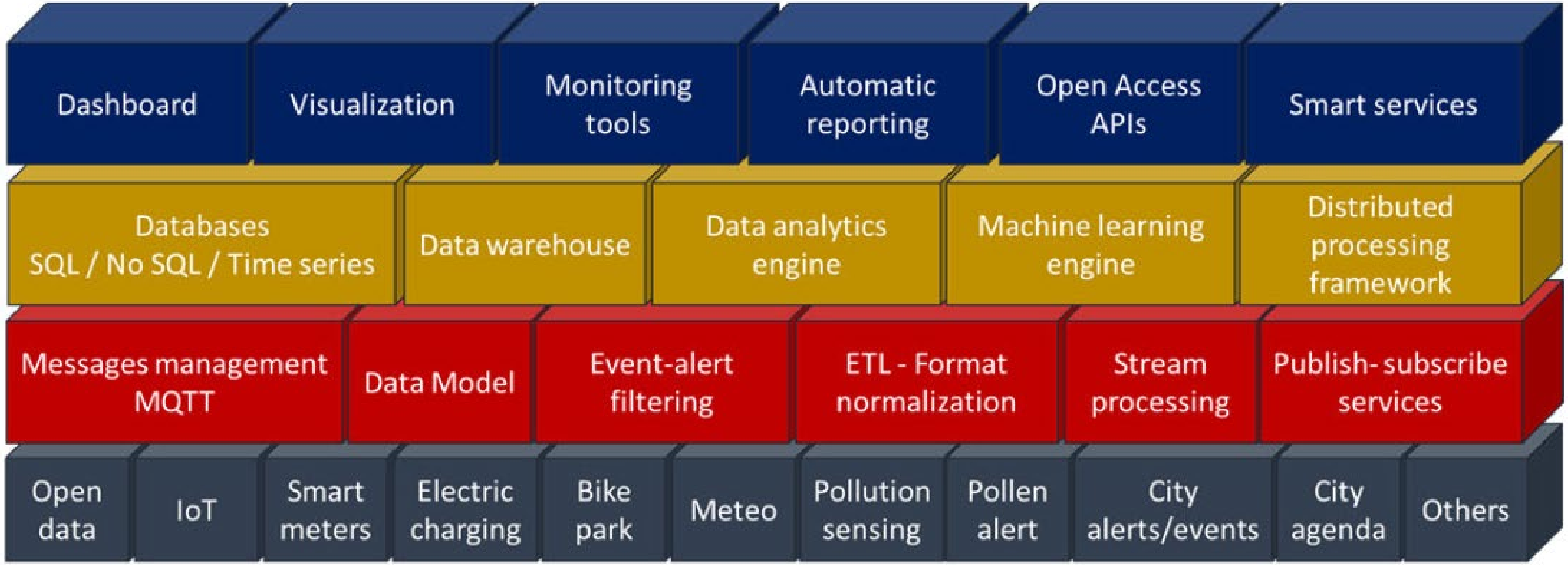


FIGURE 2. **Schematic description of the city platform layer architecture.**

A. Al-Rahamneh *u. a.*, „Enabling Customizable Services for Multimodal Smart Mobility With City-Platforms“, *IEEE Access*, Bd. 9, S. 41628–41646, 2021, doi: 10.1109/ACCESS.2021.3065412.

Next step in architecturess: Fog Computing

Universität Bamberg

- Sending all raw sensor data to the cloud cannot be the final solution:
  - Bandwidth
  - Energy comsumption
    - (computing needs less than communication)
  - Application needs, e.g., privacy or data cleansing tasks (online data quality)
- Edge computing:
  - Move the processing to the edge of the network
- Fog computing:
  - Utilize further processing nodes on the way

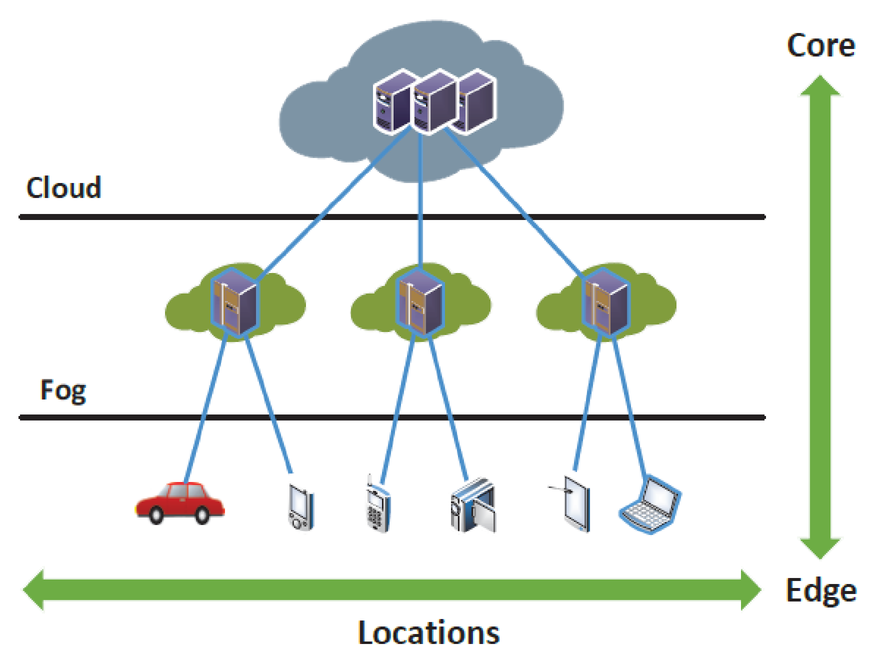


Fig. 1. Fog between edge and cloud. [1]

→ Our research: Distributed Data Stream Management for Smart Data Processing

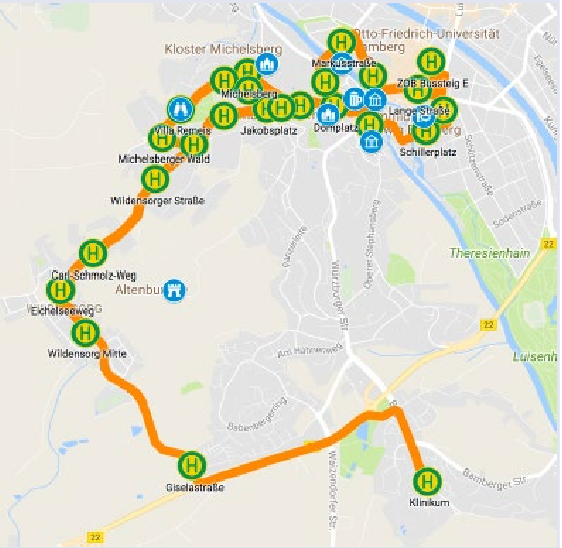
[1] I. Stojmenovic and S. Wen, "The Fog Computing Paradigm: Scenarios and Security Issues," 2014, pp. 1–8.

What would be Smart Data Processing?

- Know thy application:
  - Do not send everything in the cloud, but send only information that is needed
- Know thy resources:
  - Save energy and bandwidth
- Know thy privacy:
  - Aggregate or remove sensitive information
- Know thy quality:
  - Monitor factors that affect sensor data quality
  - Assessment or correction

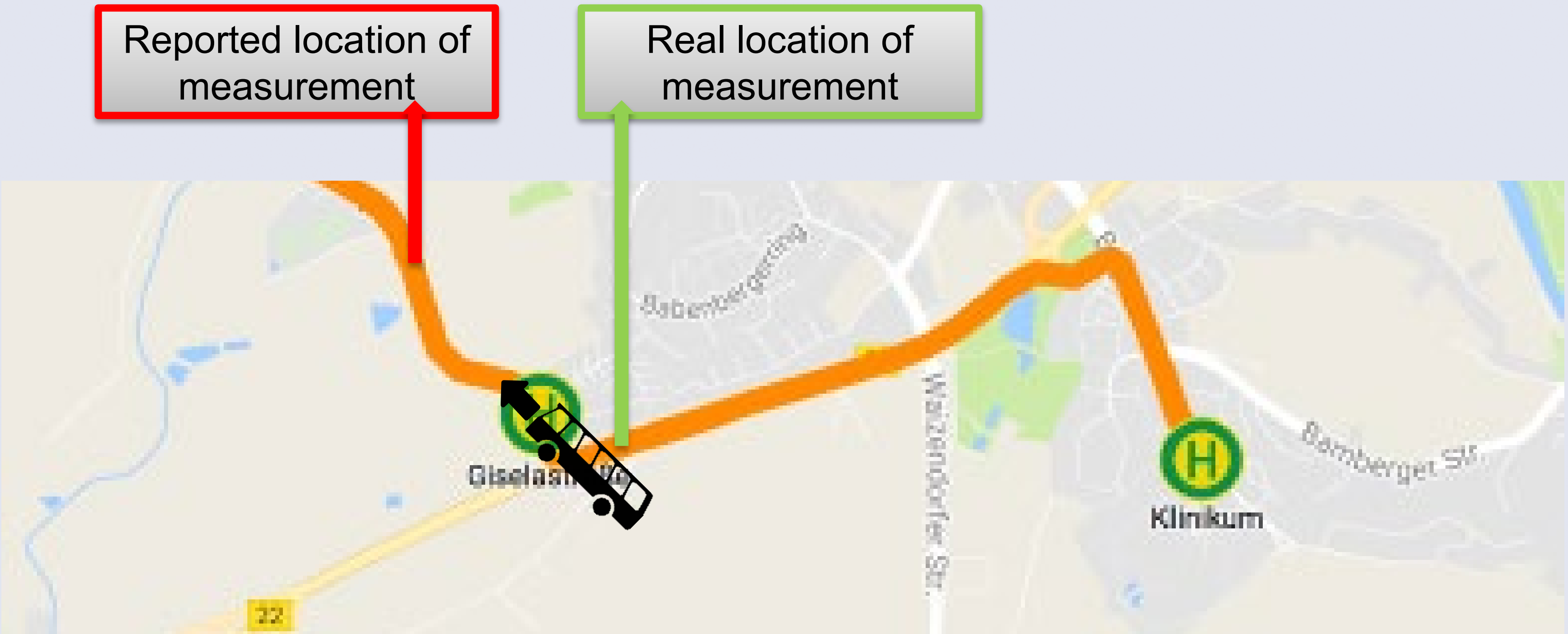
Example: Mobile air quality application

- Idea:
  - Measure air quality on city bus (from project FutureIOT)
  - Augment with crowd sensing when anomalies are detected
- IOT devices:
  - Mobile sensor system:
    - Air quality sensor with particulate matter
    - GPS location
  - Gateway device for sensors: communication over
    - LPWAN (online, limited data volume)
    - Bluetooth LE (only in vincinity)
    - WLAN (only in certain locations)
  - Smartphones of participating citizens
    - Should receive push notification when crowd sensing help is needed
- Challenges:
  - Limited connectivity of mobile sensor and gateway
  - Location privacy of participating citizens
  - Spatio-temporal data quality adjustment: Measurements have latency → Actual value is valid only behind the bus (depends on velocity)



Ongoing research: Spatio-temporal data quality adjustment

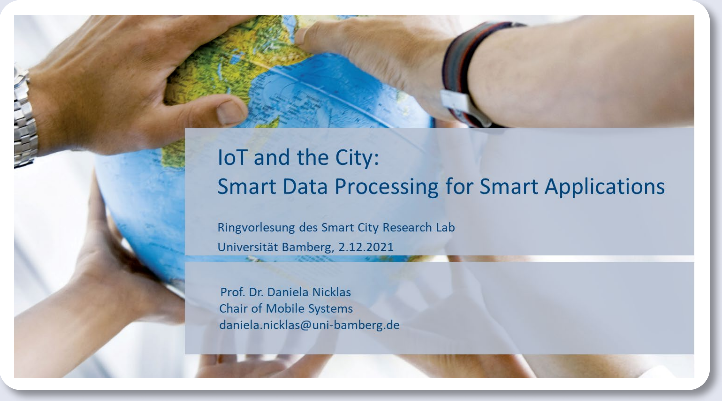
- PM measurement has latency
- Real measurement is behind the bus
- Needs correction based on speed and direction

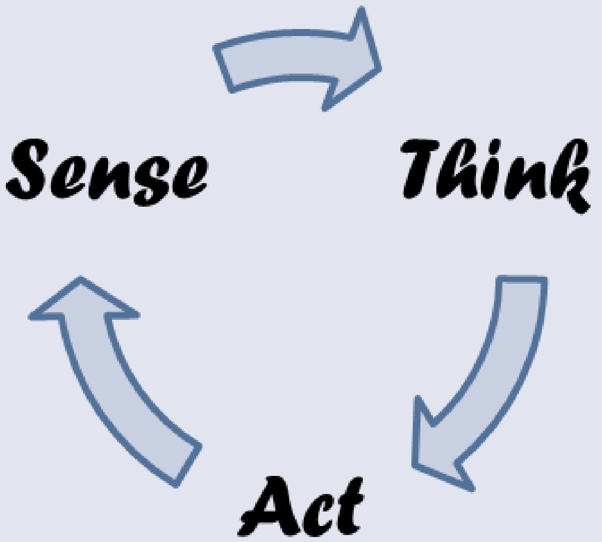


What is IoT?

Smart …

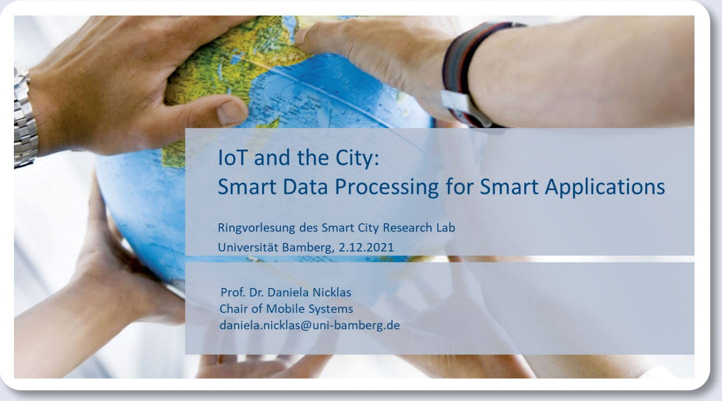
- City
- Applications
- Data Processing
  - IoT Architecture
- Conclusion
- Discussion!





Conclusion

- Smart Data Processing considers …
  - application (what?)
  - resources (energy, bandwith, memory, …)
  - privacy (cities are public spaces!)
  - quality (no data is perfect)

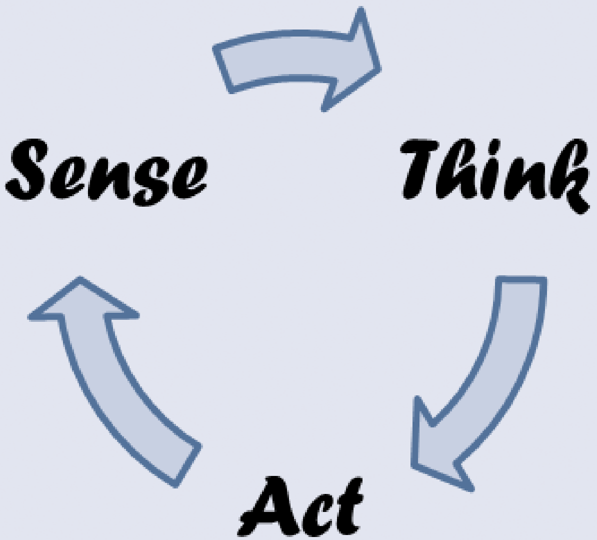


- Might avoid Big Data problems

Helps to build smart applications for a

smart

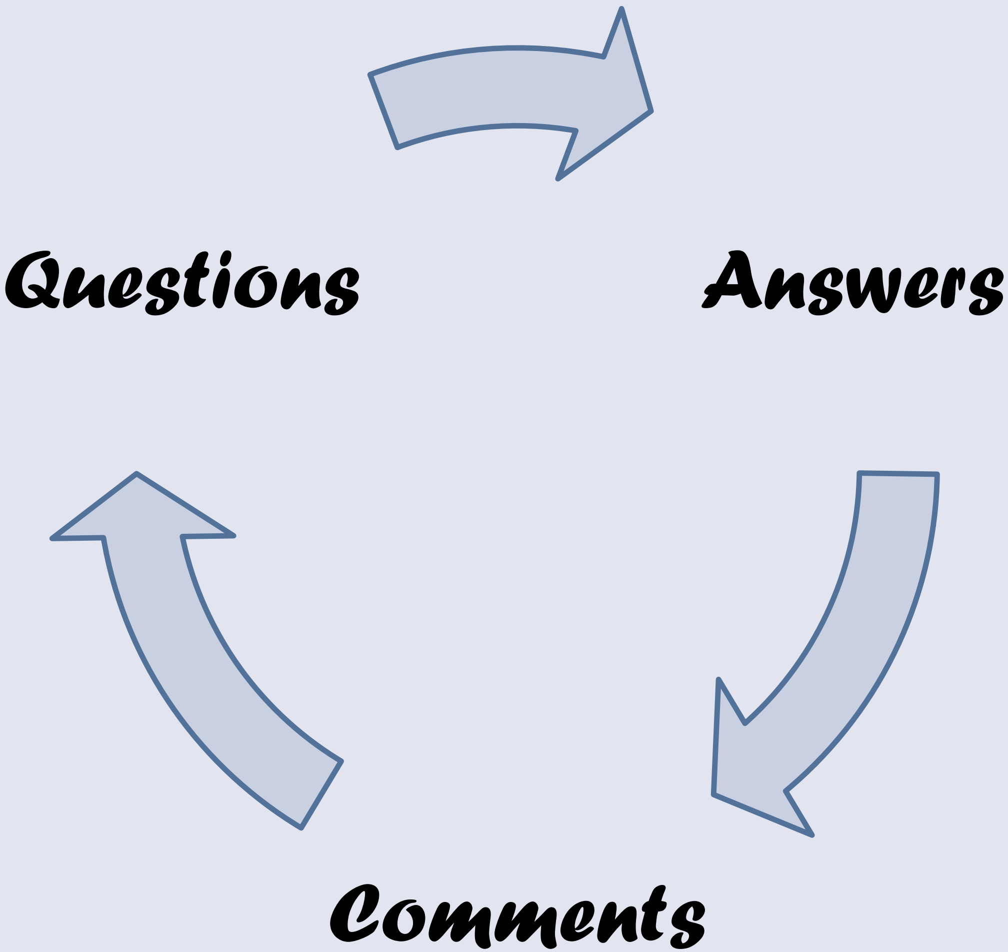
city



Discussion

Universität Bamberg

Thank you!



Overview

- IOT and IOT applicatations
- Data streams and mobility
- Sensor data quality
- Challenges for Data Management in IOT

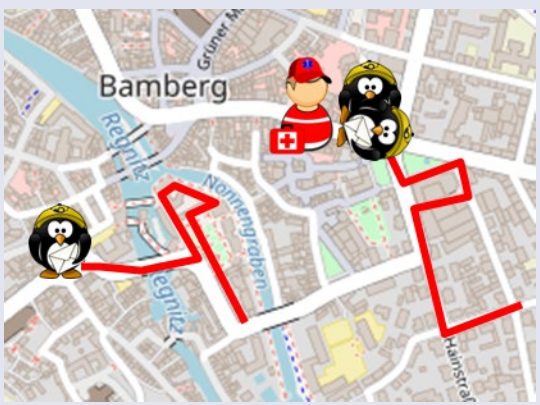
Who watches the watchmen?
Online data quality in future IOT applications

Blockveranstaltung „Informationssysteme in mobilen und drahtlosen Umgebungen“, Universität Jena, 19.10.2020

FutureIOT

- Pan-Bavarian research projects, 02/2018-09/2021
- 10 research partners, >20 industry partners
- Smart city and smart agriculture applications, e.g.:

Long-term evaluation and new services with inductive parking sensors



Privacy-aware delivery tracking (with safety and security features)

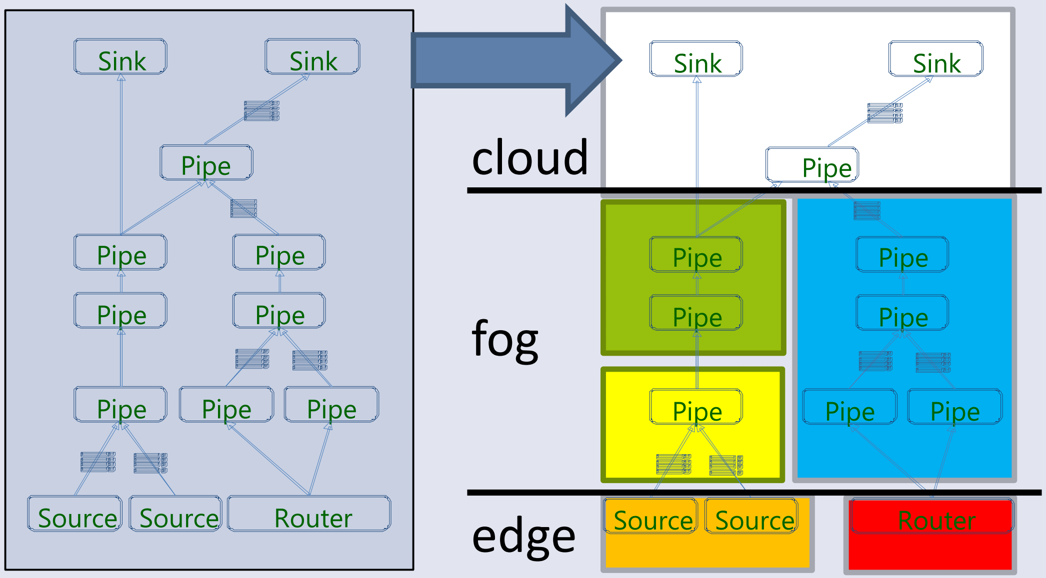
Environmental sensing – stationary and mobile

Activity recognition for cows (to detect anormal behaviour)

Fog computing and distributed data stream management

Universität Bamberg

- Data stream management:
  - Provides a higher-level abstraction to stream-based data processing
- Distributed stream management:
  - Distributes the execution of the data stream processing over nodes
  - Finds an optimized query execution plan
  - Can adapt to changing situations and migrate the execution



Can we use (distributed) DSMS to implement sensor data stream management in a fog-computing architecture?

Overview

- IOT and IOT applicatations
- Data streams and mobility
- Sensor data quality
- Challenges for Data Management in IOT

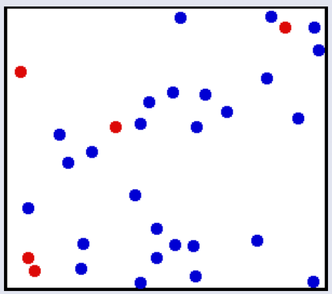
Who watches the watchmen?
Online data quality in future IOT applications

Blockveranstaltung „Informationssysteme in mobilen und drahtlosen Umgebungen“, Universität Jena, 19.10.2020

Data and mobility

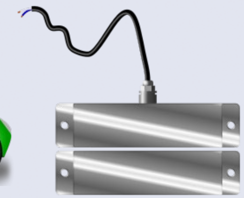
Universität Bamberg

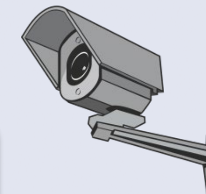
Thermal energy (heat)



[1]

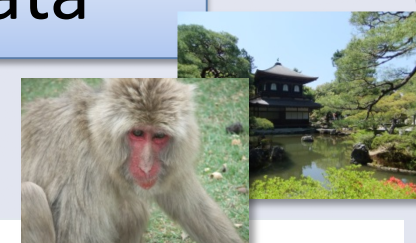
Mobility and movement

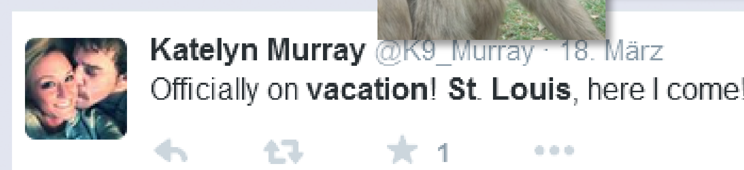




Data

Earthquake!





2006, http://ana.blogs.com/maestros/2006/11/data_is_the_new.html, retrieved 21.3.2015

Data is the New Oil

By Michael Palmer

"Data is the new oil!" Clive Humby, ANA Senior marketer's summit, Kellogg School.

Data is just like crude. It's valuable, but if unrefined it cannot really be used. It has to be changed into gas,

Data is the new crude oil …

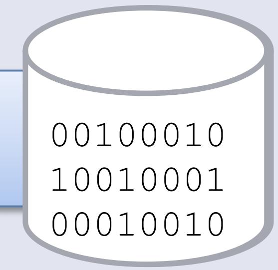
Universität Bamberg

… it needs to be refined to be valuable.

Crude oil

- Fuel oil → mobility
- Chemical products:
  pharmaceuticals → health
  fertilizers → increase growth
  pesticides → kill insects
- …

Data



Information

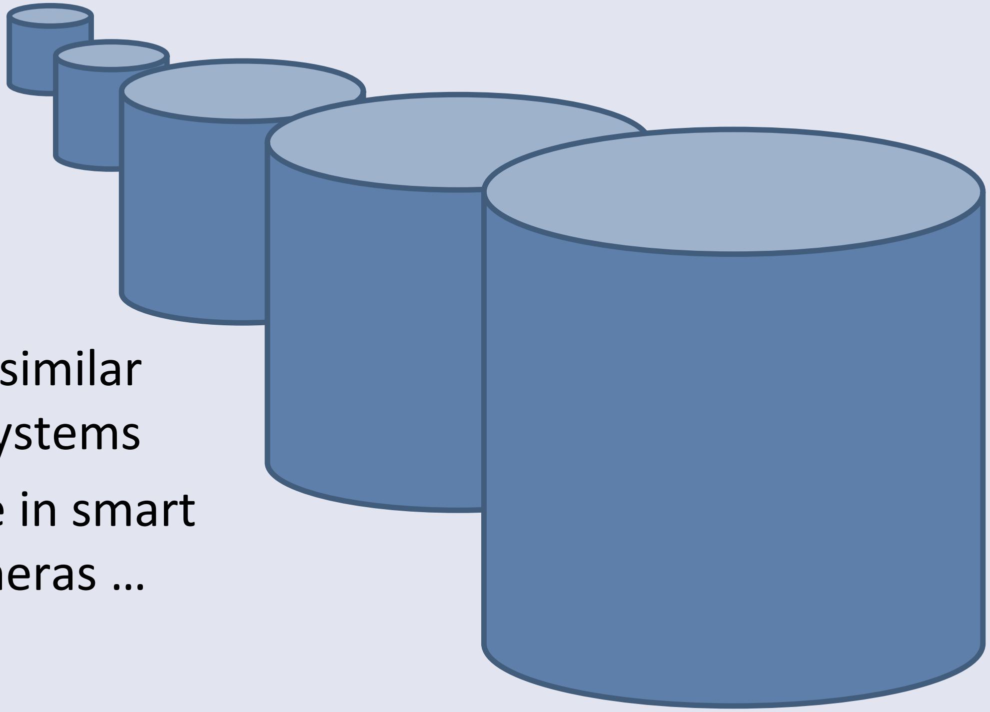
Knowledge

Action

DATA and mobility

Universität Bamberg

- Googles self-driving car: nearly 1MB data per second[1]
  - Per day: 85 GB
  - Per year: ~30 TB
  - If 10% of the cars would be like this, or 50%, or … (> 1 Billion cars on the world)
  - … not only by self-driving cars, similar for advanced driver assistant systems
  - … plus data from infrastructure in smart cities, like induction loops, cameras …



→Big Data!

[1]Bill Gross, Founder and CEO of Idealab
https://www.linkedin.com/today/post/article/20130502024505-9947747-google-s-self-driving-car-gathers-nearly-1-gb-per-second

Big Data Challenges

- Many definitions, often by a numer of 3-5 "V" challenges:

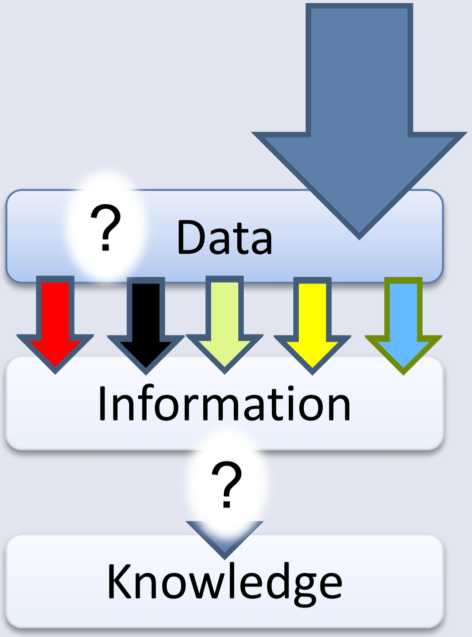
Volume — A lot of data (amount varies)

Variety — Data differs in structure

Variability — Structure changes

Velocity — Many updates

Veracity — Unclear source or quality

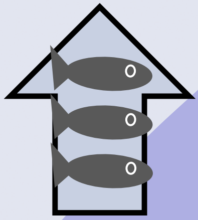


Not in list of challenges: Pricacy

(analysis of sensible data, how to adhere to legal / societal constraints?)

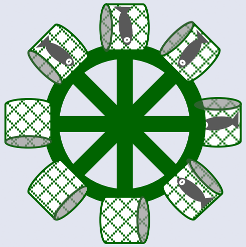
Data stream management and Big Data

Universität Bamberg



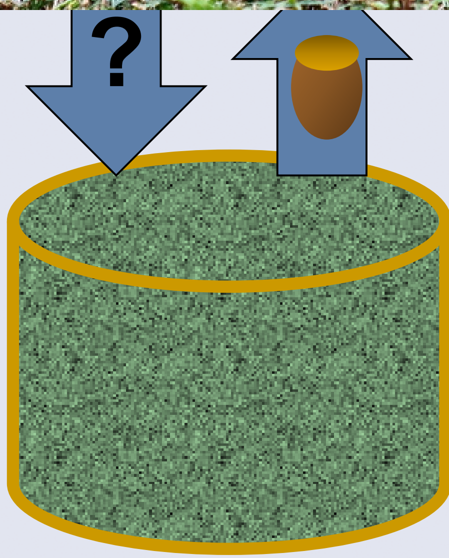
- More "velocity", less "volume"
- Direct processing
  - Online, (hard/soft) real time, "right time"
- More information, less data
  - Enrichment of data streams
    - E.g., product information for an RFID tag
  - Interpretation and reasoning
    - E.g., classification ("this is a car")
  - Data cleansing
    - Remove redundancy, anomaly detection
- Online quality assessment
- Enables built-in privacy methods
  - Online pseudonmization and anonymization
  - Data economy
  - Certify and/or publish your query plans

Bild: Ronny Senst / pixelio.de



Data management approaches

Bild: elli60 / pixelio.de



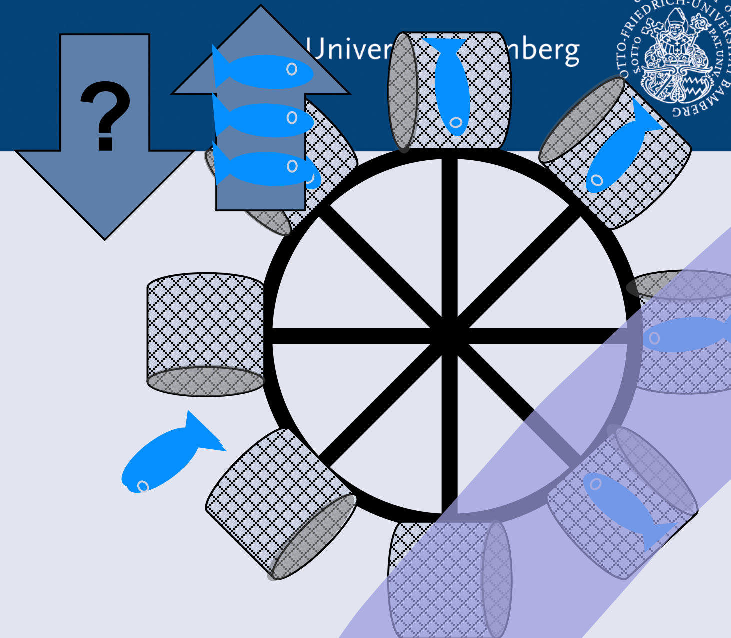
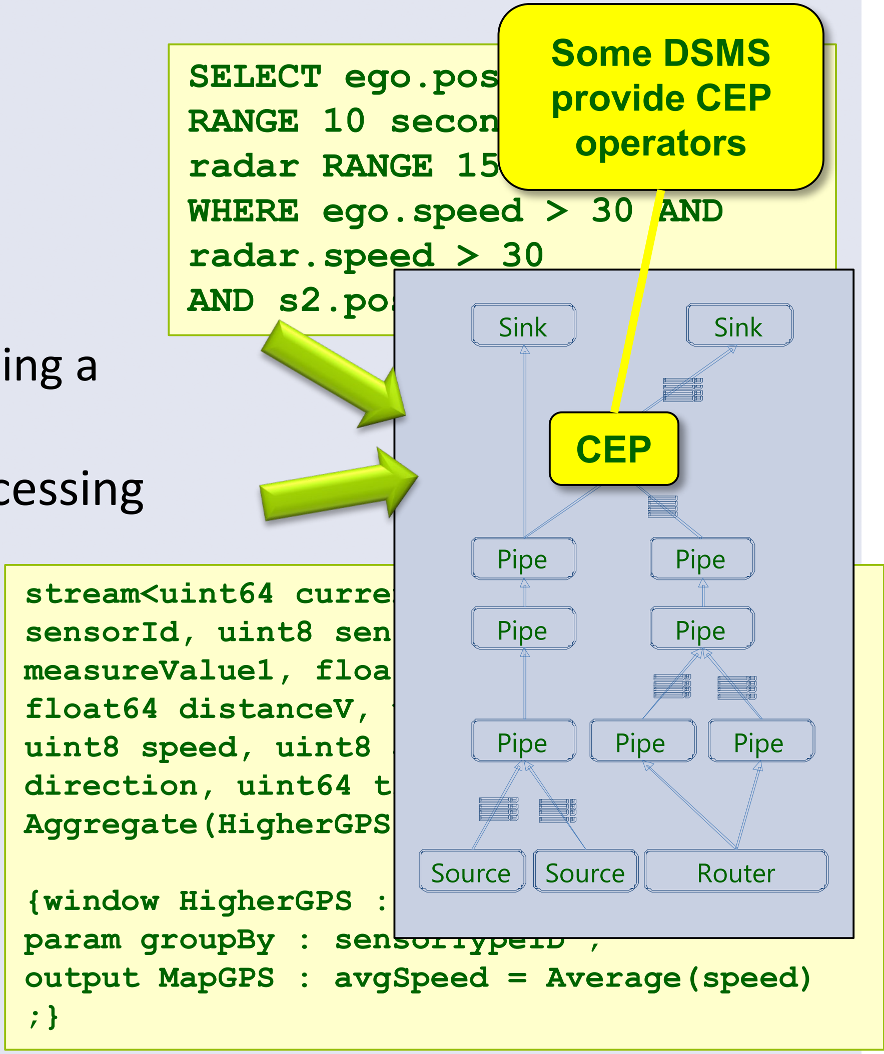


Bild: Ronny Senst / pixelio.de

Features of Data Stream Management

Universität Bamberg

- Programming abstraction
  - Declarative: Query
  - Functional: Data flow graph
    - Enables query optimizations
    - Better maintanance of systems
  - → Using a DSMS on data streams is like using a DBMS instead of files
- Easy to combine with complex event processing (CEP)
- Parallel execution of operators in graph → no shared memory
- Data streams can be unbounded: Issues with sorting, joins, aggregation
  - Approximate answers
  - Window semantics



Overview

- IOT and IOT applicatations
- Data streams and mobility
- Sensor data quality
- Challenges for Data Management in IOT

Who watches the watchmen?
Online data quality in future IOT applications

Blockveranstaltung „Informationssysteme in mobilen und drahtlosen Umgebungen“, Universität Jena, 19.10.2020

Sensors

A **sensor** is an electronic component, module, or subsystem whose purpose is to detect events or changes in its environment and send the information to other electronics, frequently a computer processor.

https://en.wikipedia.org/wiki/Sensor

- Technical systems can achieve situational awareness by using data from sensors
- However, sensor data is often …
  - incomplete (not everything can be sensed)
  - late (results do not not arrive in time)
  - inaccurate (values are not exact)
  - mobile (sensed by moving systems)
- To make things worse, sensor data needs to be interpreted

  … and interpretations can cause further errors

Sensor data challenges

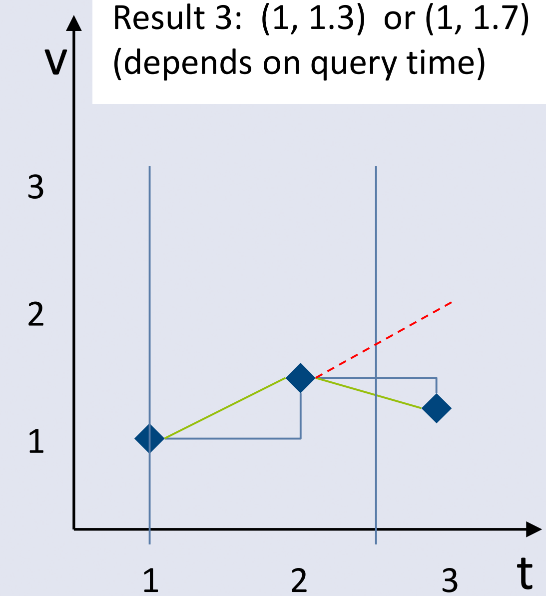
Universität Bamberg

- Format:
  - Structured (e.g. (Timestamp, Value), or (Value, Value, Value))
  - Unstructured (e.g. image stream (video) or audio stream)
  - Semi-structured (e.g. photo + DXF meta data (timestamp, location, resolution, …))
- Semantic levels:
  - Raw: just the signal
  - Feature: a typed attribute of anentity, e.g. the location
  - Object: multiple attributes grouped together for an object
  - Situation: a complex sitation was detected

  → Higher levels are often results of sensor data fusion
- Validity: How long is the sensor value valid?
  1. Only at timestamp
     - if sensor sends with fixed frequency
  2. Fixed until next data comes in
     - if sensor sends when value deviates from last value by threshold
  3. Changing according to model
     - if sensor sends when value deviates from a function of time
     - „dead reckoning“ → often used for moving objects (but can be applied to other phenomena)

select t, v from sensordata
where t = 1 or t = 2.5

Result 1: (1, NULL)
Result 2: (1, 1.5)
Result 3: (1, 1.3) or (1, 1.7)
(depends on query time)



Some common quality issues

- Data source
  - Measurement method, e.g. low frequency of sensor for fast moving objects
  - Environment, e.g., temperature too high for good measurements
  - Moving sensor with latency
- Data processing
  - Wrong training data for classifier
  - Over-simplified models or missing concepts
  - Not enough input data for algorithm
  - Stale models (due to concept drift)
  - …
- Some can be detected after installation of system, some occur later

→ Decisions based on inpresise data

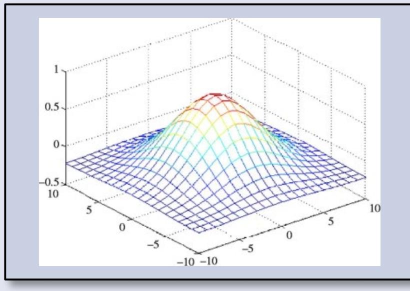
Features for a quality-aware DSMS

Goal: programming abstractions for dealing with non-perfect data

Approach:

1. develop unified data model to represent data quality
2. consider data quality in operators

→ data management can attach combined quality metadata to result



- How to determine data quality and correlations?
  - given by data source / sensor (e.g., accuracy)
  - given by algorithm (e.g., confidence)
  - learned by observation (requires redundancy)

→ store in sensor relationship model

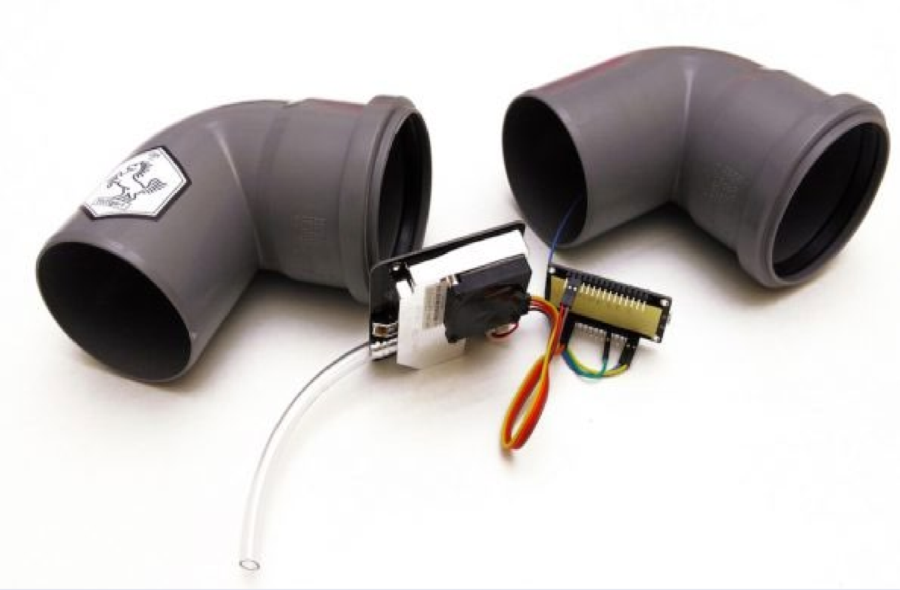
type (probability)
bicycle (0.8)
pedestrian (0.1)
other (0.1)

Quality Matters: Supporting Quality-aware Pervasive Applications by Probabilistic Data Stream Management, DEBS2014

Example 1: Citizen Science Project luftdaten.info

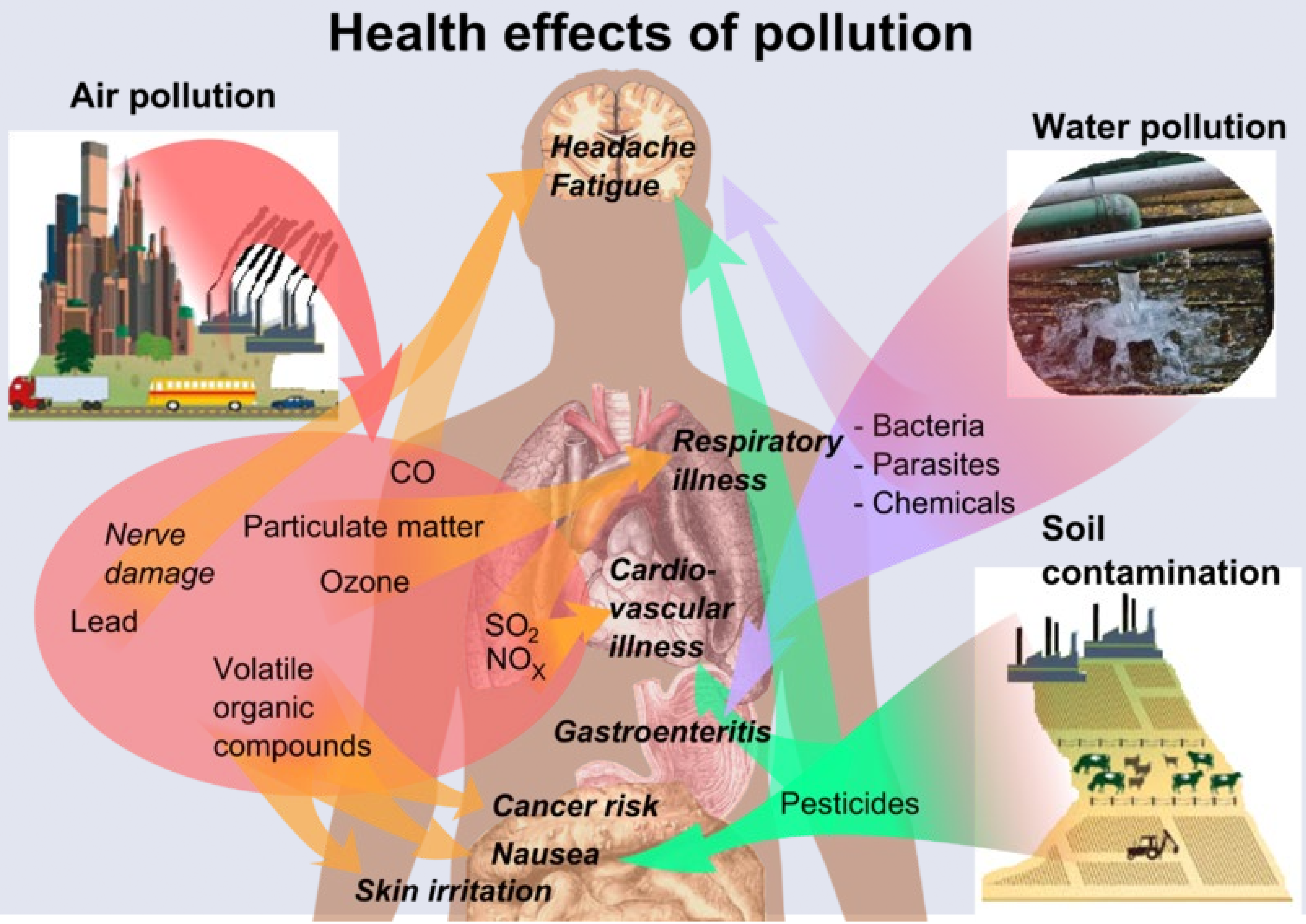
Universität Bamberg

- Founded by Open Knowledge Lab Stuttgart
  - A group of ten people working mostly on *Citizen Science* projects
- Start: June 2015
- Access for all people:
  - Feasible components
  - Easy assembly
  - Regular workshops and talks throughout Germany
- Goals:
  - Monitor the air quality in Stuttgart
  - Involve citizens in the process
  - Increase the coverage to other areas in Germany and other countries



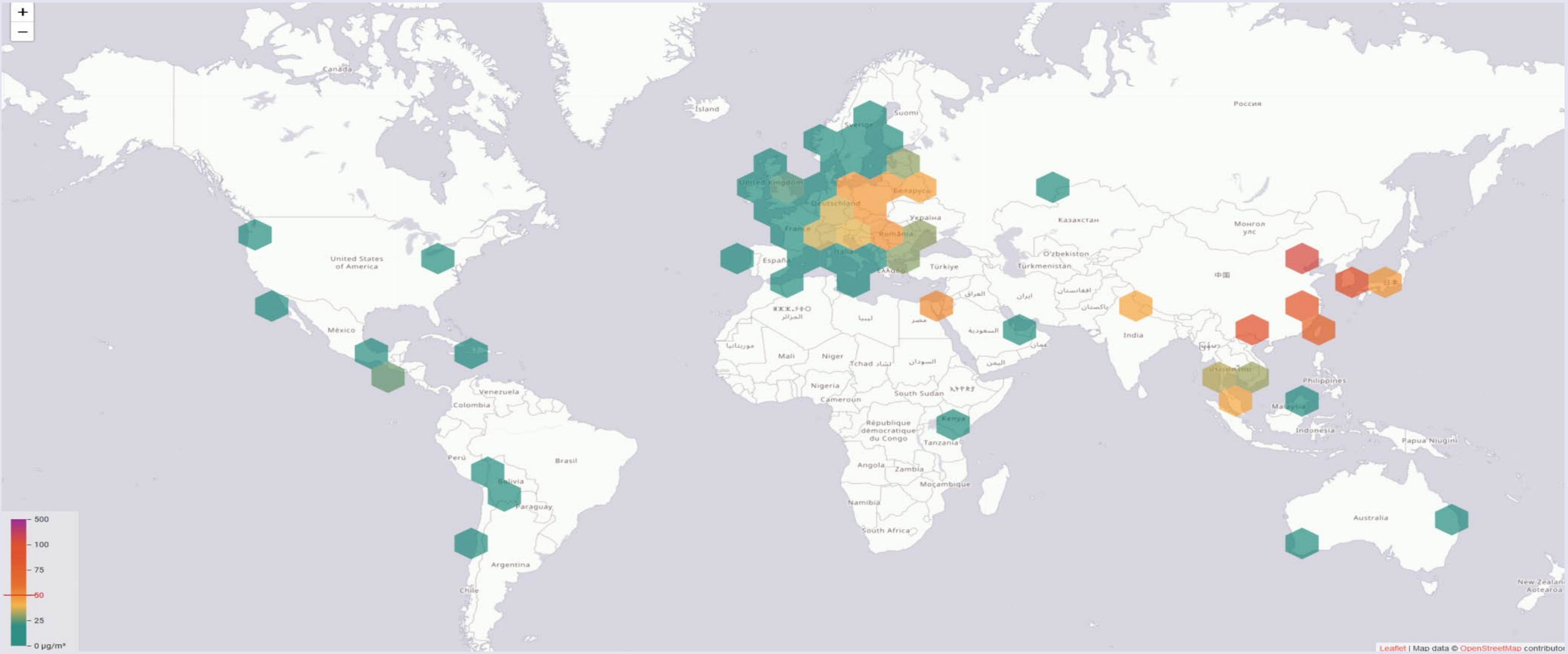
Why should we care?

Universität Bamberg



Current status of sensor installations

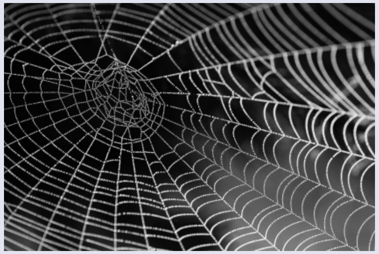
- 4230 registered sensors
- Data from over 15 countries
- Average sensor measurements produced per day: ~80000 readings (as of 02/2018)



Quality issues

Universität Bamberg

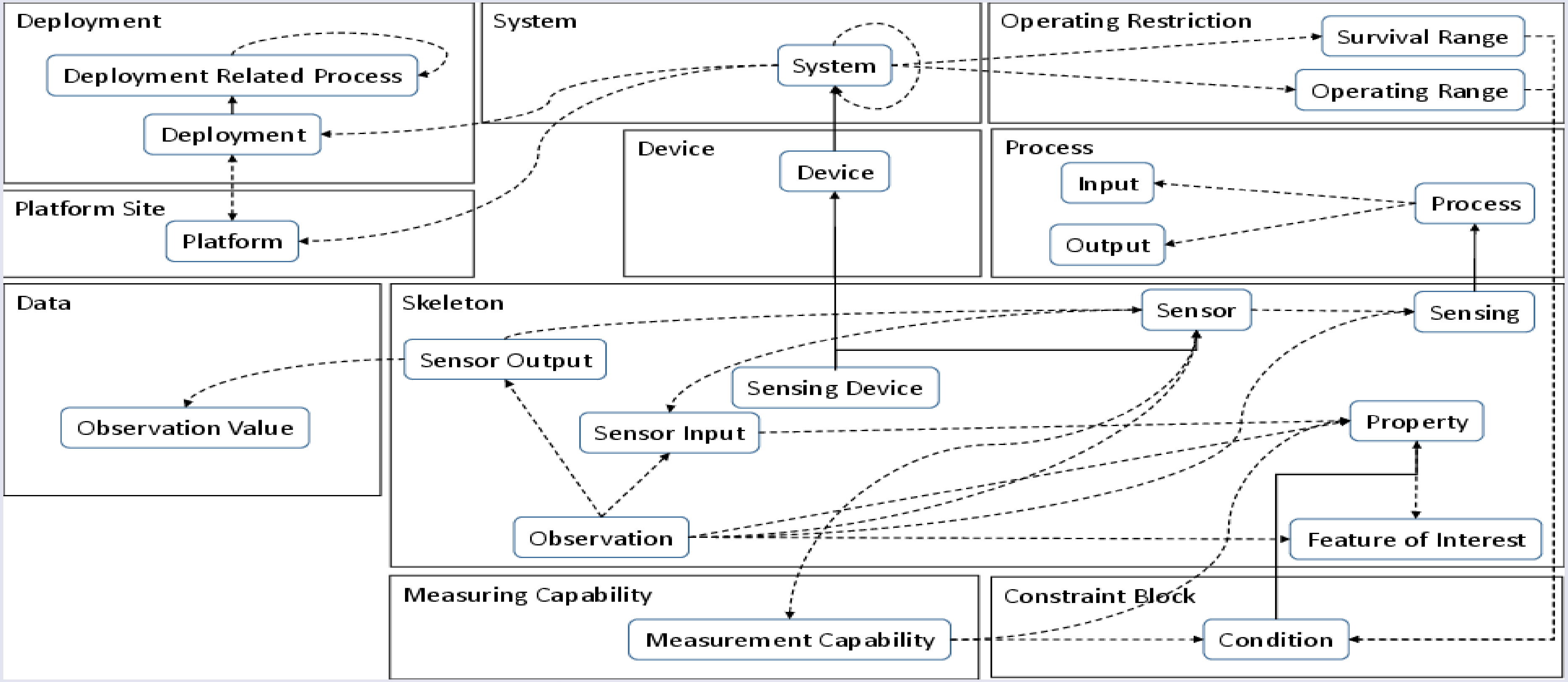
- Sensors can have faulty behaviour and anomalies
  - PM values are no longer reliable when the humidity rises above 70%
  - Some sensors are installed indoors → data leads to wrong conclusions
  - Strange effects might occur later (e.g., spiders in the sensor box)



- Offline detection of the PM quality issues
  - Fusion of the sensor data with data from the neighbouring weather stations
- Online detection of the PM quality issues
  - Use of the on board humidity sensor as a reference
  - Live Stream processing of the sensors with the related weather satations based on the underlying semantic model

SSN ontology: Overview

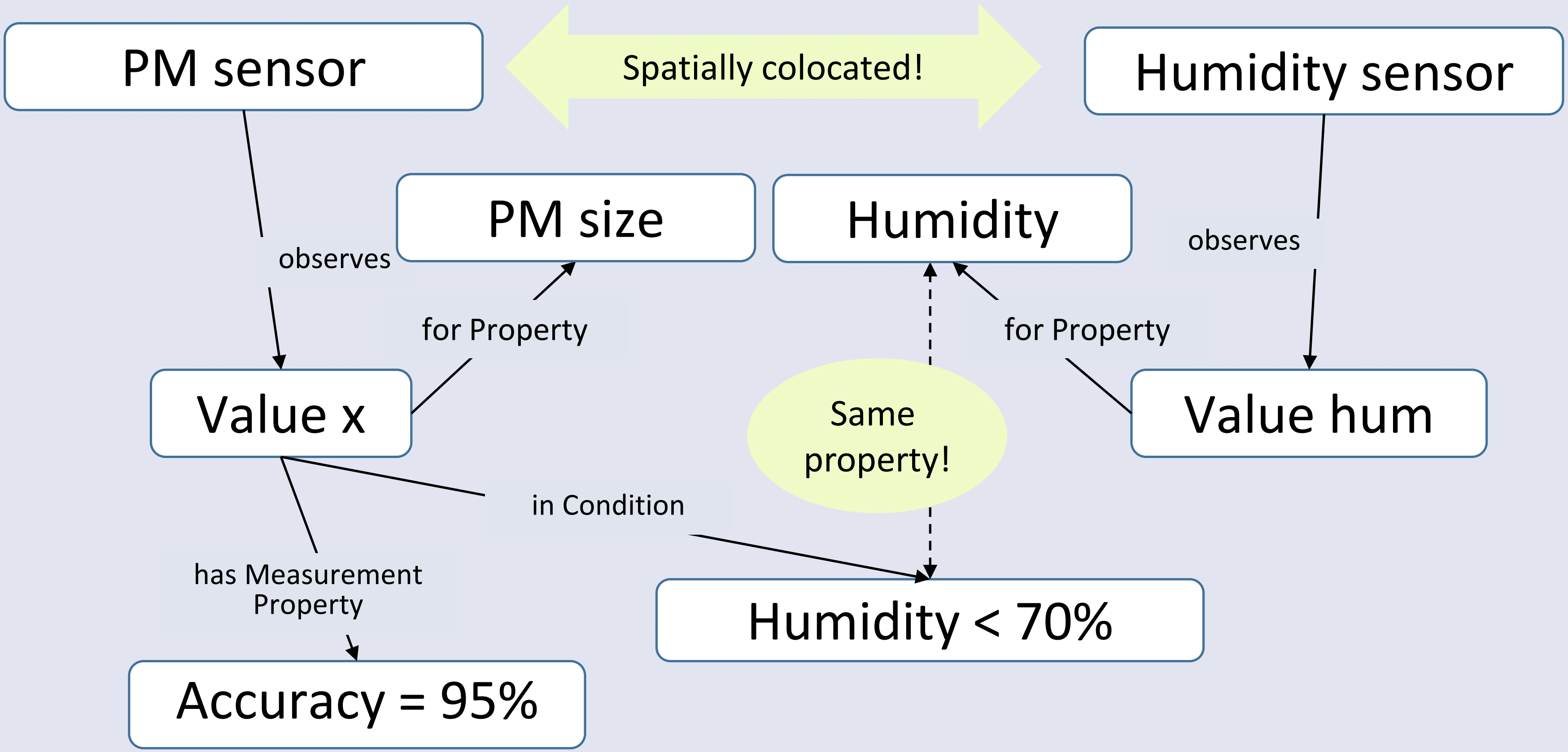
- Created by W3C Semantic Sensor Network Incubator Group (2011) for
  - Syntactic interoperability
  - Semantic compatibility of sensors and their measurements



Christian Kuka: Qualitätssensitive Datenstromverarbeitung zur Erstellung von dynamischen Kontextmodellen, PhD, Universität Oldenburg, 2015

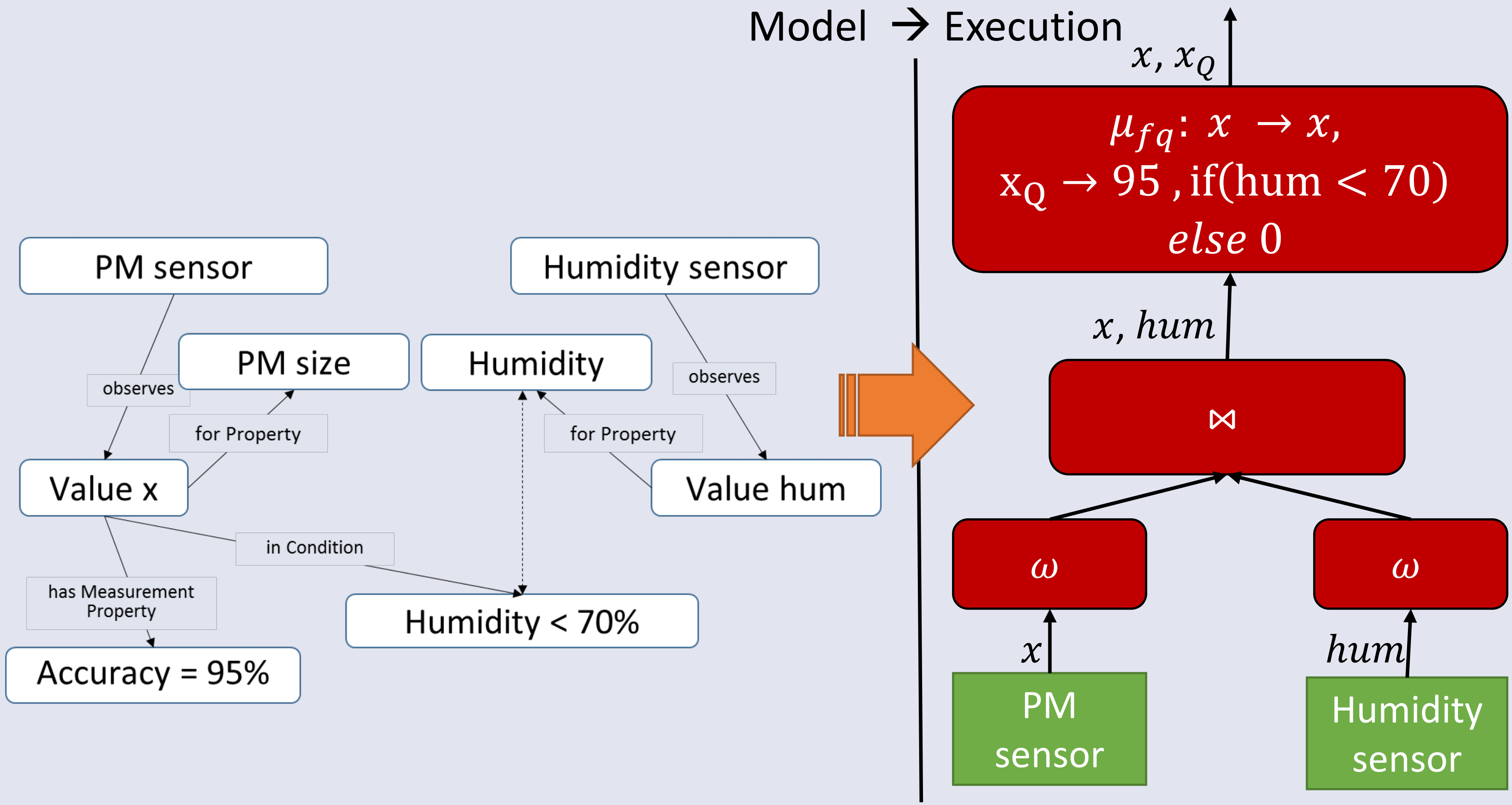
Modeling online sensor quality in SSN

- Measurement capabilities can have conditions
- Conditions can be measured / observed by other sensors!



From ontology to stream processing (3)

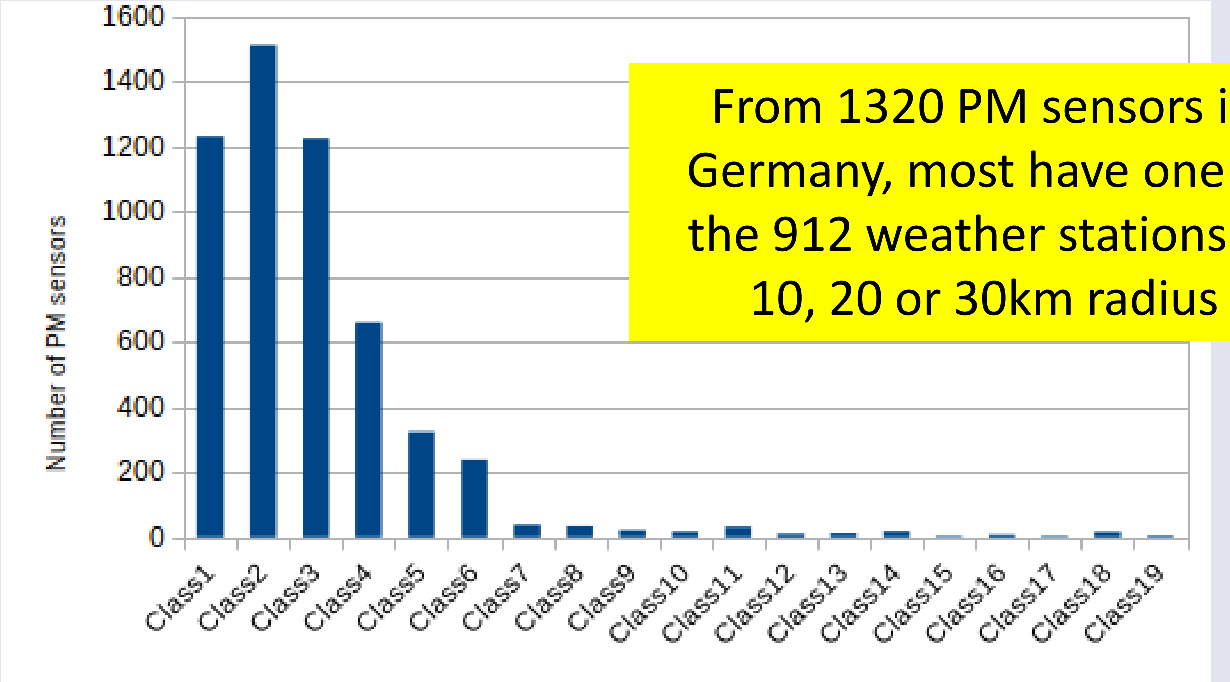
Universität Bamberg



Which humidity sensors can we use?

Universität Bamberg

- Can we trust an on-board huminidity sensors?
- Or: Use weather stations as data quality reference sources?
- Data from the weather stations used to assess the accuracy of sensors



From 1320 PM sensors in Germany, most have one of the 912 weather stations in 10, 20 or 30km radius

Number of sensors covered by at least one weather station in 10km radius classes

Overview

Universität Bamberg

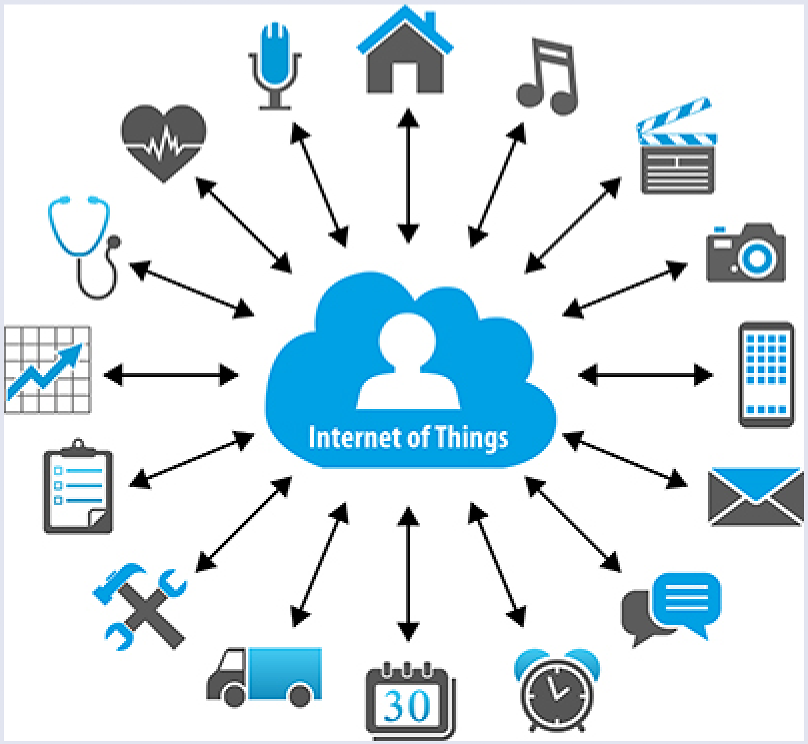
- IOT and IOT applicatations
- Data streams and mobility
- Sensor data quality
- Challenges for Data Management in IOT

Who watches the watchmen?
Online data quality in future IOT applications

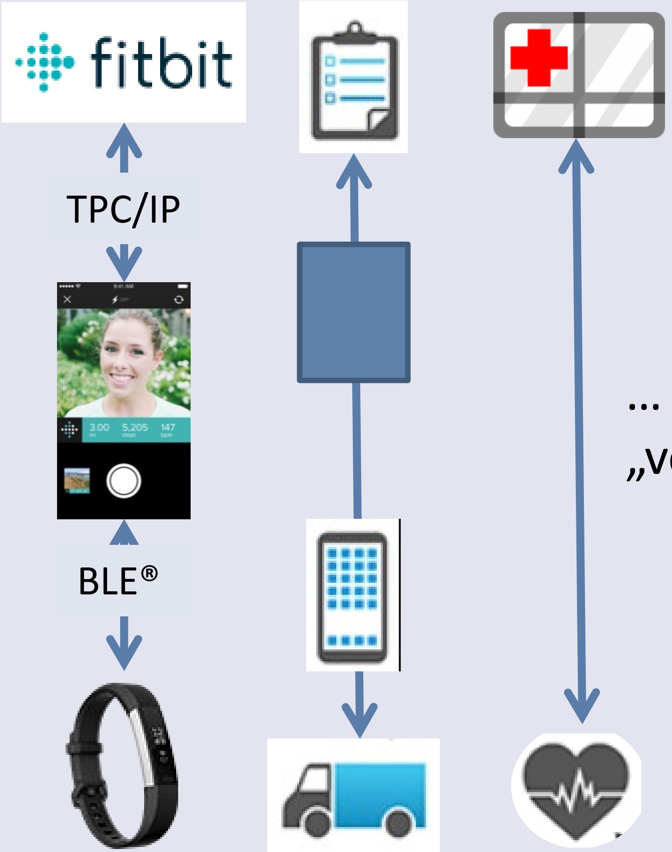
Blockveranstaltung „Informationssysteme in mobilen und drahtlosen Umgebungen“, Universität Jena, 19.10.2020

IoT Architecture

Vision



Reality



… many „verticals“

Next step in architectures: Fog Computing

- Sending all raw sensor data to the cloud cannot be the final solution:
  - Bandwidth
  - Energy comsumption
    - (computing needs less than communication)
  - Application needs, e.g., privacy or data cleansing tasks (online data quality)
- Edge computing:
  - Move the processing to the edge of the network
- Fog computing:
  - Utilize further processing nodes on the way

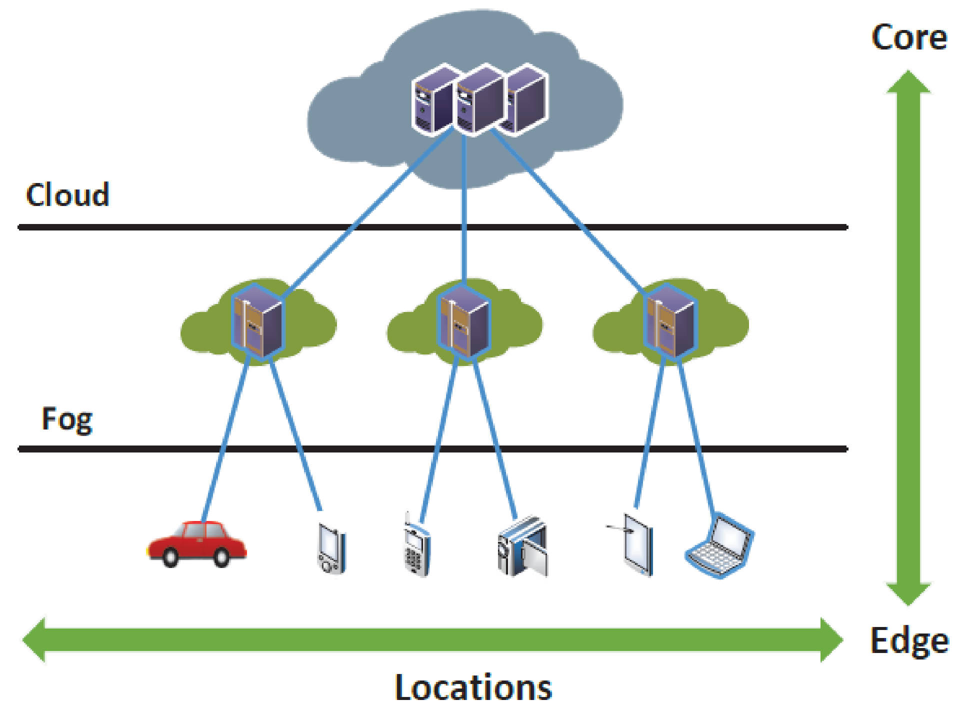


Fig. 1. Fog between edge and cloud.

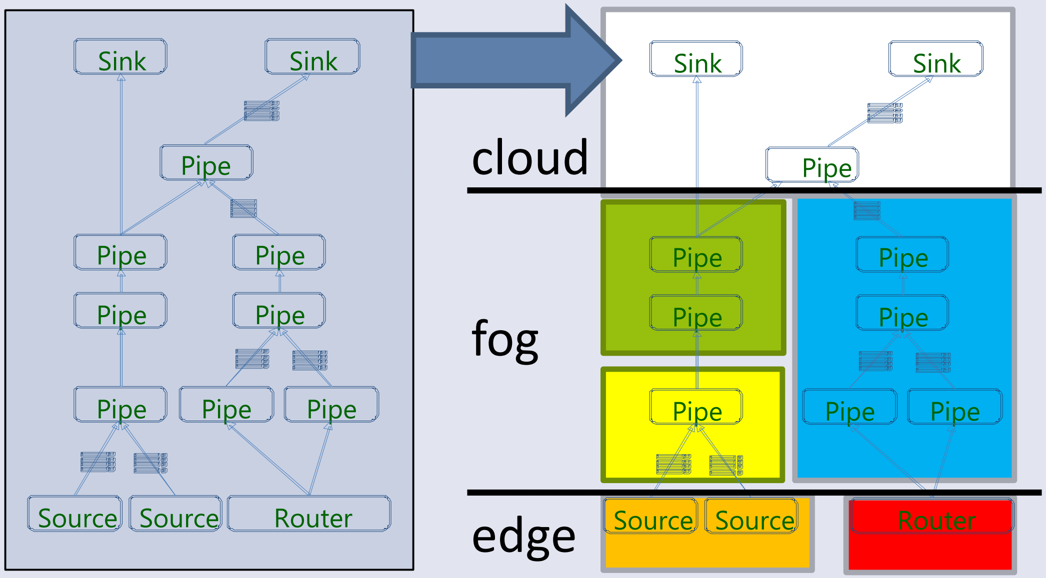
[1]

[1] I. Stojmenovic and S. Wen, “The Fog Computing Paradigm: Scenarios and Security Issues,” 2014, pp. 1–8.

Fog computing and distributed data stream management

Universität Bamberg

- Data stream management:
  - Provides a higher-level abstraction to stream-based data processing
- Distributed stream management:
  - Distributes the execution of the data stream processing over nodes
  - Finds an optimized query execution plan
  - Can adapt to changing situations and migrate the execution



Can we use (distributed) DSMS to implement sensor data stream management in a fog-computing architecture?

Distributed data management for IOT

Universität Bamberg

Distributed data stream management for IOT would require …

- CRUD for data streams
  - Create:
    - Integration with device management (known data sources)
    - Create with parameters, e.g., update rates or simple filters
  - Read:
    - Access data stream
    - Optional: Add filters and simple single-stream operations
  - Update:
    - Change parameters
  - Delete:
    - Drop data streams; keep integrity (e.g., RESTRICT if still in use)
- Complex queries for data streams
  - Window, Aggregate, Join, User defined oprations, Join with archived data
- Distributed query optimization
  - Operator placement, data shipping
- Data ownership and authorization

If it would be for stored data, this would be …

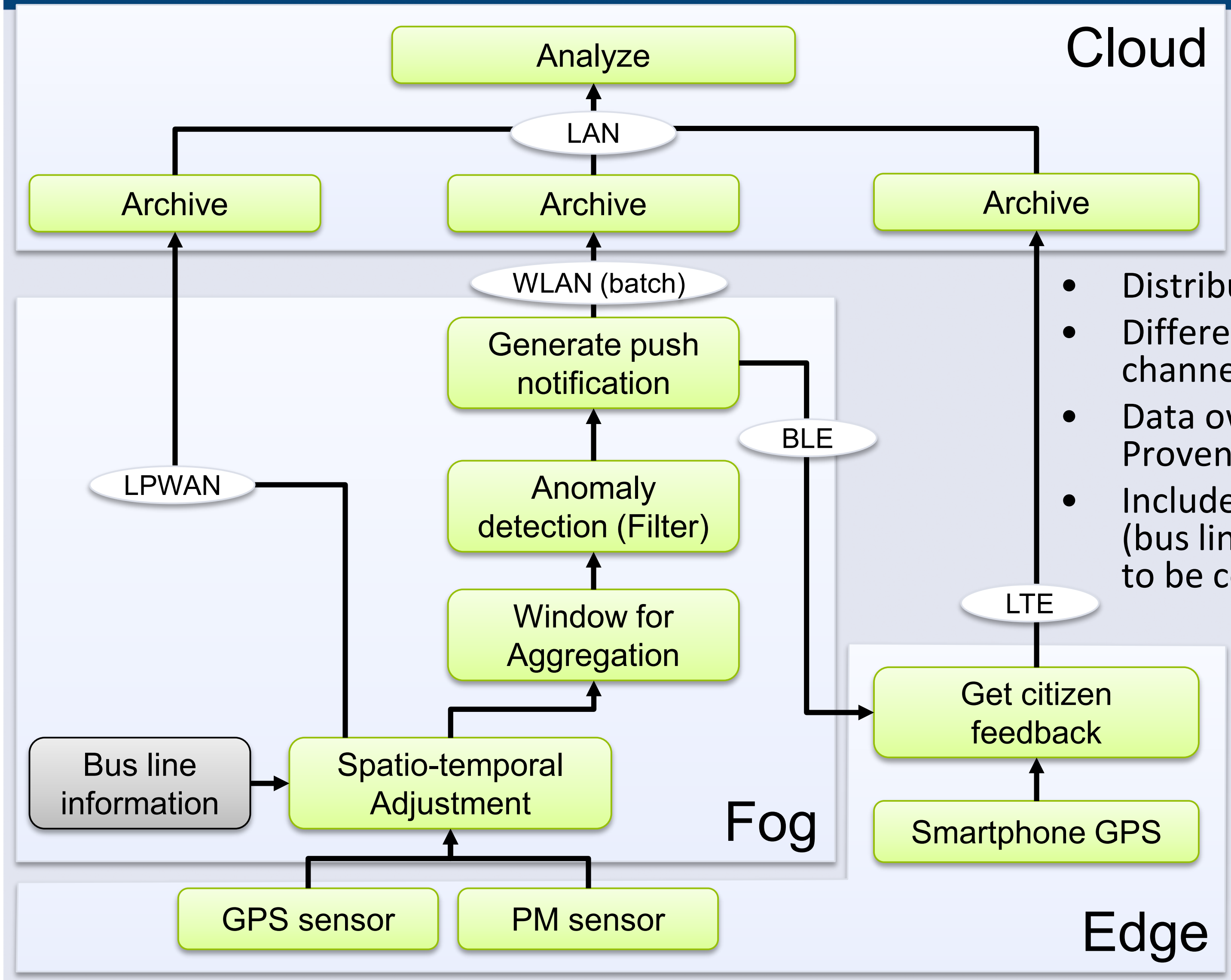
Spatio-temporal data quality adjustment

Universität Bamberg

- PM measurement has latency
- Real measurement is behind the bus
- Needs correction based on speed and direction

Data flow for example application

Universität Bamberg



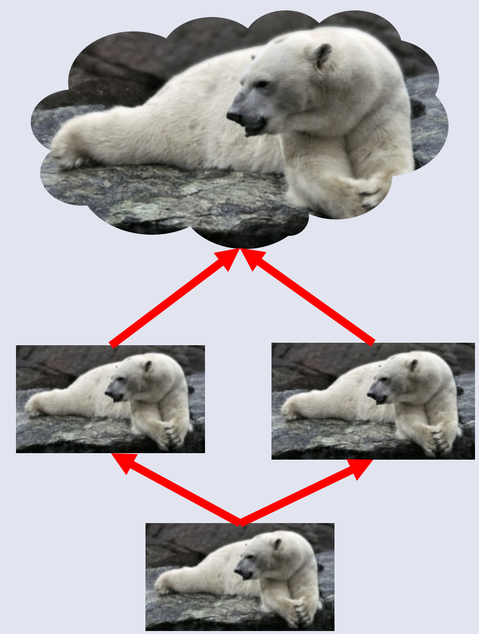
- Distributed data management
- Different communication channels
- Data ownership? Provenance?
- Includes non-stream data (bus line information needs to be configured)

Summary and outlook

Universität Bamberg

- IoT applications can lead to large-scale sensor data management systems
- Issues to solve:
  - The „V“ challenges → maybe you do not need to store everything in the cloud
  - The „P“ challenge → maybe you can anonymize or aggregate at the edge or in the fog
  - The „Q“ challenge → know thy quality, before and during operation
- IoT platforms can help, but are only slowly moving towards fog architectures – and they only provide limited data management support

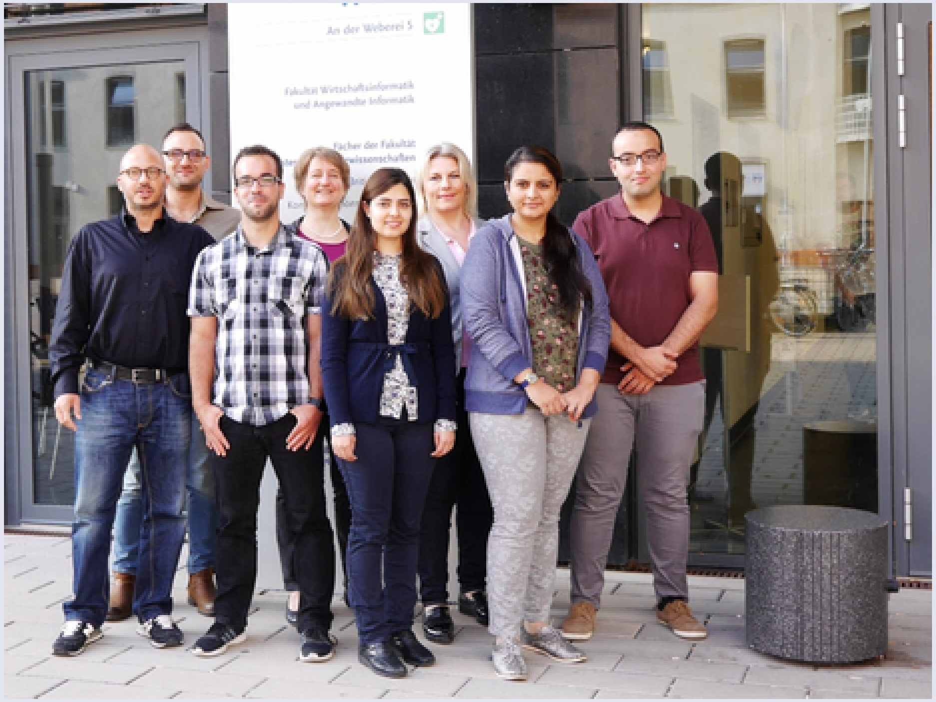
  → Distributed data stream processing revisited?



Ronny Senst / pixelio.de

Thank‘s for all the fish!

Universität Bamberg



WIAI | Fakultät Wirtschaftsinformatik und Angewandte Informatik

Prof. Dr. Daniela Nicklas
Lehrstuhl für Informatik,
insb. Mobile Software Systeme / Mobilität

Otto-Friedrich-Universität
An der Weberei 5
96047 Bamberg

Telefon +49 (0) 951/863 3670
Sekretariat: +49 (0) 951/863 3671
dnicklas@acm.org
www.uni-bamberg.de

Any Questions?

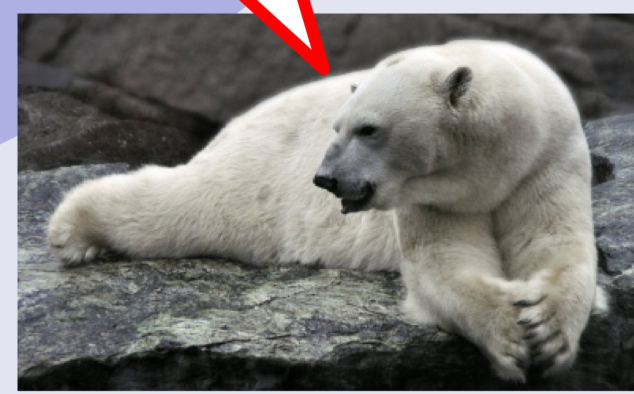


Bild: Ronny Senst / pixelio.de