Internet of Things and Big Data for Smart Community Sensing(SCS),

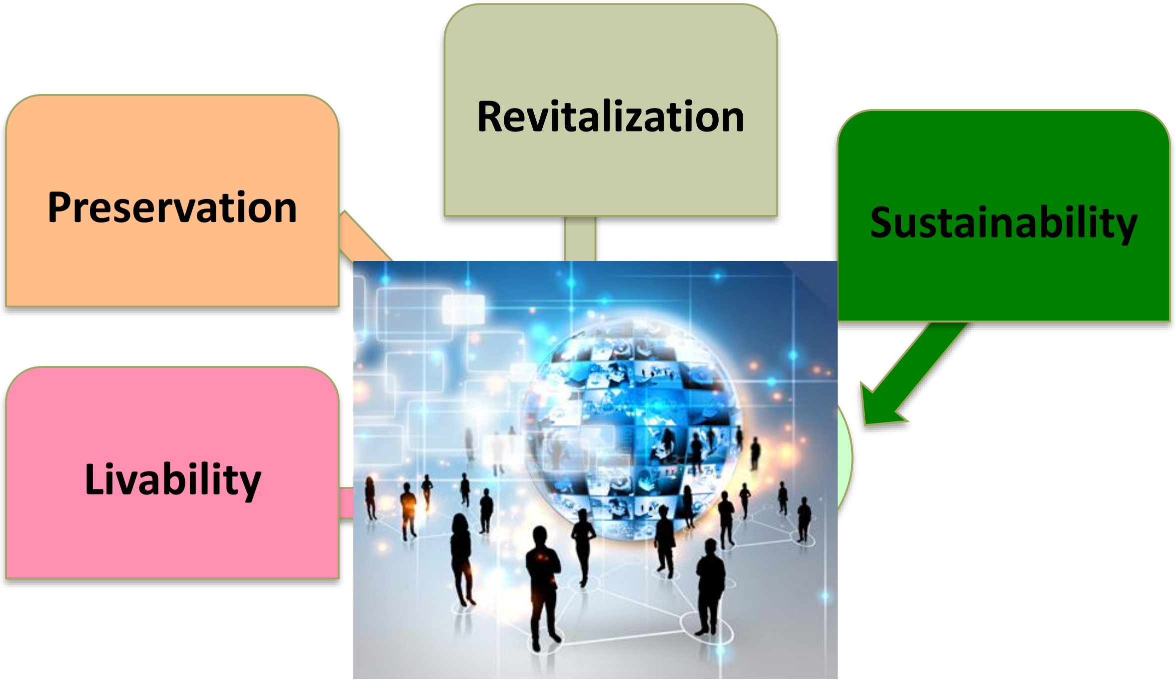
Stockholms universitet

Outline

- Characteristic of smart & community sensing
- Opportunities and challenges of IoT in smart & community sensing
- Opportunities and challenges of Big Data analytics in smart & community sensing
- Opportunities in Mobile Crowd sensing
- Internet of Things Data Taxonomy
- Sensing-as-Service
- A Social Sensors Service representation
- IoT Architecture Requirements
- IoT Broker

Characteristic of smart & community sensing

- Livability
- Preservation
- Revitalization
- Sustainability



LIVABILITY

- Provide more transportation choices

To decrease household transportation costs, reduce our dependence on oil, improve air quality and promote public health.

- Promote equitable, affordable housing

people of all ages, incomes, races and ethnicities to increase mobility and lower the combined cost of housing and transportation.

- Enhance economic competitiveness

by giving people reliable access to employment centers, educational opportunities, services and other basic needs

- Support existing communities

Through transit-oriented and land recycling – to revitalize communities, reduce public works costs, and safeguard rural landscapes

- Coordinate policies and leverage investment

To remove barriers to collaboration, leverage funding and increase the effectiveness of programs to plan for future growth

- Value communities and neighborhoods

By investing in healthy, safe and walkable neighborhoods, whether rural, urban or suburban.

Reference: https://www.transportation.gov/livability/101

Preservation

- Classified into two:
  - *Cultural heritage*
  - *Natural heritage*

  Constitutes a source of identity and cohesion for communities disrupted by bewildering change and economic instability.

  Tangible documents, paintings, sculptures and historic buildings and sites are just like intangible assets like folk traditions or languages essential elements of our cultural heritage and identity.

  The challenges of the 21st century in protecting the tangible and intangible, cultural and natural heritage of the world are global in character and need to be also addressed on a global level. Cultural and natural heritage are increasingly threatened with destruction not only by the physical and chemical mechanisms of decay, but also by changing social and economic conditions.

Reference: http://ipch.yale.edu/news/company-scholars-are-we-losing-our-past-or-our-future-sustainable-preservation-cultural

Stockholms universitet

Revitalization

- Many rural communities and small towns are facing challenges:
  - -Including declining rural populations
  - -Rapid growth at metropolitan edges
  - -Loss of farms and working lands
- small towns more closely with surprisingly complex
  - Economic development challenges
  - Analytical needs, to complex land use and transportation issues
  - Suffer from a chronic unemployment,

Reference: http://rural-design.org/resource/revitalizing-small-towns-resolving-downtown-challenges

Sustainability

- **Sustainability of a community depends** on creating and maintaining its economic and environmental health, promoting social equity, and fostering broad-based citizen participation in planning and implementation
- Communities that engage citizens and institutions to develop sustainability principles and a collective vision for the future and that apply an integrative approach to environmental, economic, and social goals are generally likely to be more successful.
- Job creation, energy use, housing, transportation, education and health are considered complementary parts of the whole. Since all issues are interconnected they must be addressed as a system.

Sustainability *as a system*

The process includes:

- broad and diverse involvement of citizens
- the creation of a collective vision for the future
- the development of principles of sustainability
- an inventory of existing assets and resources and additional assets that would benefit the community
- clear, measurable goals
- the development of community indicators to evaluate progress
- open and transparent communication
- early, visible results
- celebration of success

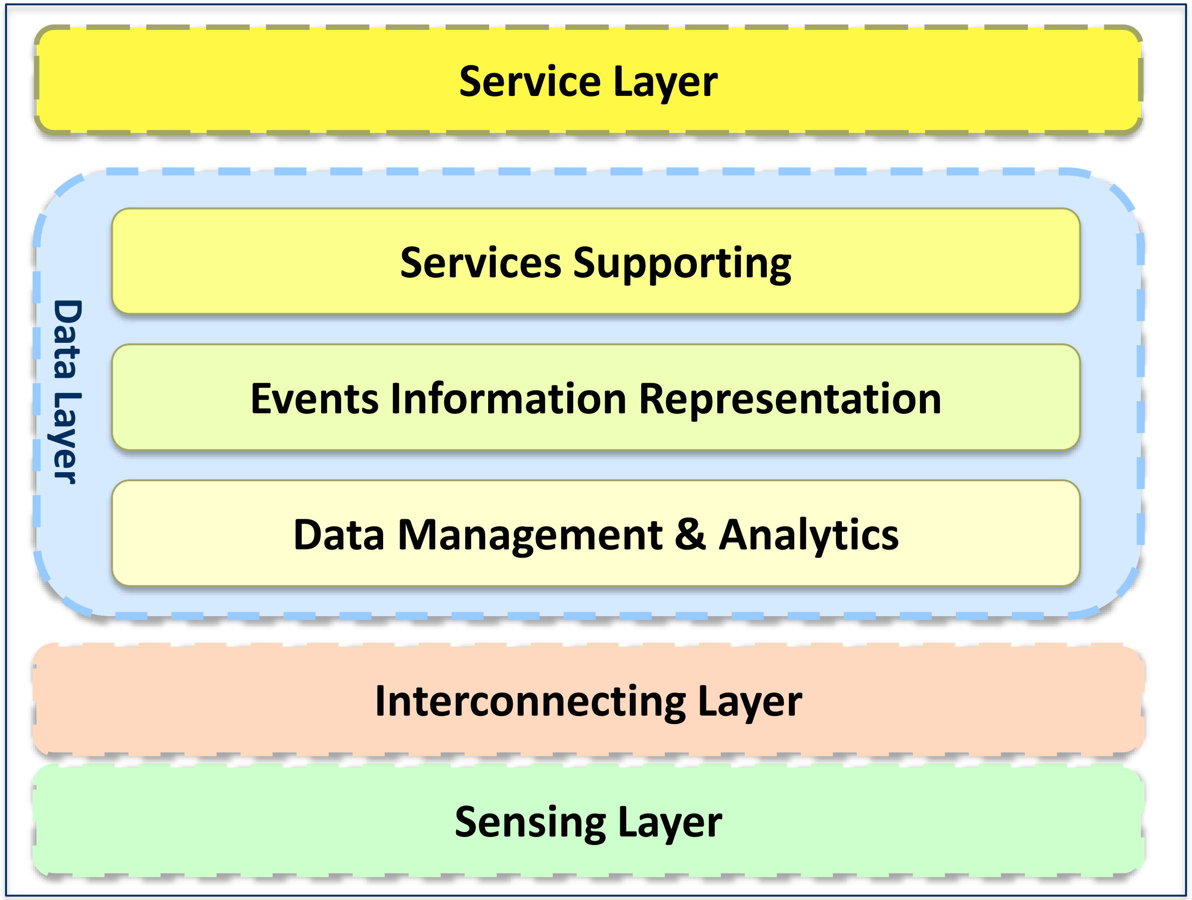
Stockholms universitet

Cont. Sustainability as a system

- Sustainability is a process of continuous improvement so communities constantly evolve and make changes to accomplish their goals.
- ways to make your community healthier, safer, greener, more livable, and more prosperous.

Opportunities and Challenges of IoT in Community sensing

- **Sensing Layer**
  - To realize ubiquitous sensing.
  - People-centric urban sensing – 3 categories
    - Personal sensing, Social sensing and public sensing
- **Interconnecting Layer**
  - is data transmission and information exchange among different devices and different domains
- **Data Layer**
  - Data Management & Analytics
  - Events information Representation
  - Services Supporting
- **Service Layer**
  - or application layer, is to provide various services for communities



Stockholms universitet

Opportunities of IOT in Community Sensing

- **M**obile **C**rowd **S**ensing (MCS)
  - A category of IoT applications relay on **data collection** from large number of mobile sensing devices such as smartphones.
  - Categorized into **three different classes** based on the type of phenomenon being measured or mapped***: environmental***, ***infrastructure***, and ***social***
- Cyber-Physical Cloud Computing
  - IoT is a networking infrastructure for Cyber-Physical System (CPS)
  - Smart networked systems with **embedded sensors**, **processors** and **actuators**
  - Designed **to sense** and **interact** with **physical world** and **support real-time**, guaranteed **performance in safety-critical** applications.

Challenges of IOT in Community Sensing

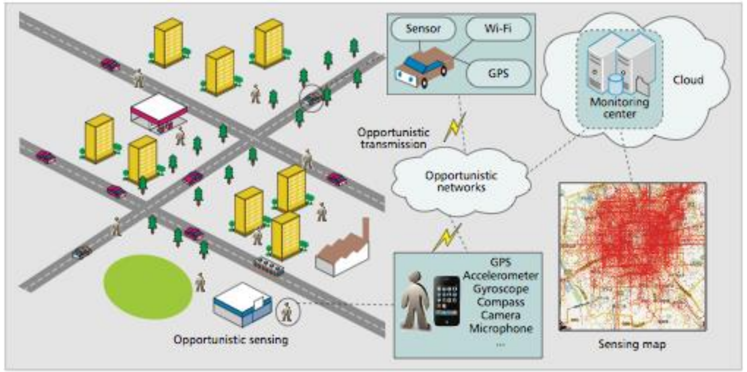
- Cyber Security and Privacy
  - Human-centered systems
  - It is important to preserve the security and privacy of an individual
  - Mobile **crowd sensing** application potentially **collect sensitive sensor** data pertaining to individuals.
- Resource Limitations
  - Set of mobile devices that are collecting sensor data are **highly dynamic** in **Availability** and **Capabilities.**
  - Due to this **dynamic nature** modeling and **predicting** the energy, bandwidth requirements in community sensing

Opportunities and Challenges of Big Data Analytics in community Sensing

- Opportunities of Big Data in Community sensing
  - Big Data Analytics enables the **move** from **IoT** to **real-time control**, which is desired for many community sensing applications.
  - In the era of IoT, Smart Community Sensing (SCS) feature the deployment of multitude of wireless sensors and **agents spanning many application** domains including: environmental, healthcare, smart interconnected automobiles and trucks, and smart buildings.
- Challenging of Big Data in Community sensing
  - **Data Heterogeneity:** How to translate **physical**, **biological**, or **social** variables into a meaningful electrical signal is a challenge.
    - How to improve quality (accuracy) of data in real time, i.e., sampling and filtering.
    - how to unify data representation and processing models to accommodate heterogeneous or new types of data.
    - how to improve intelligent data interpretation and semantic interoperability.
    - how to implement inter-situation analysis and prediction.
    - how to implement knowledge creation and reasoning.
    - how to conduct short-term and long-term storage
  - Decision making under Uncertainty:
    - Many sources of uncertainty must be considered in decision-making.
    - our understanding of many issues facing cities is incomplete.
    - we often lack the data needed to specify the boundary conditions with sufficient accuracy
    - We must improve decision making under uncertainty by understanding assessment, representation and propagation of uncertainty, developing robust-optimization methods, and designing optimal sequential decision making.

Section II:

Opportunities in Mobile Crowd sensing



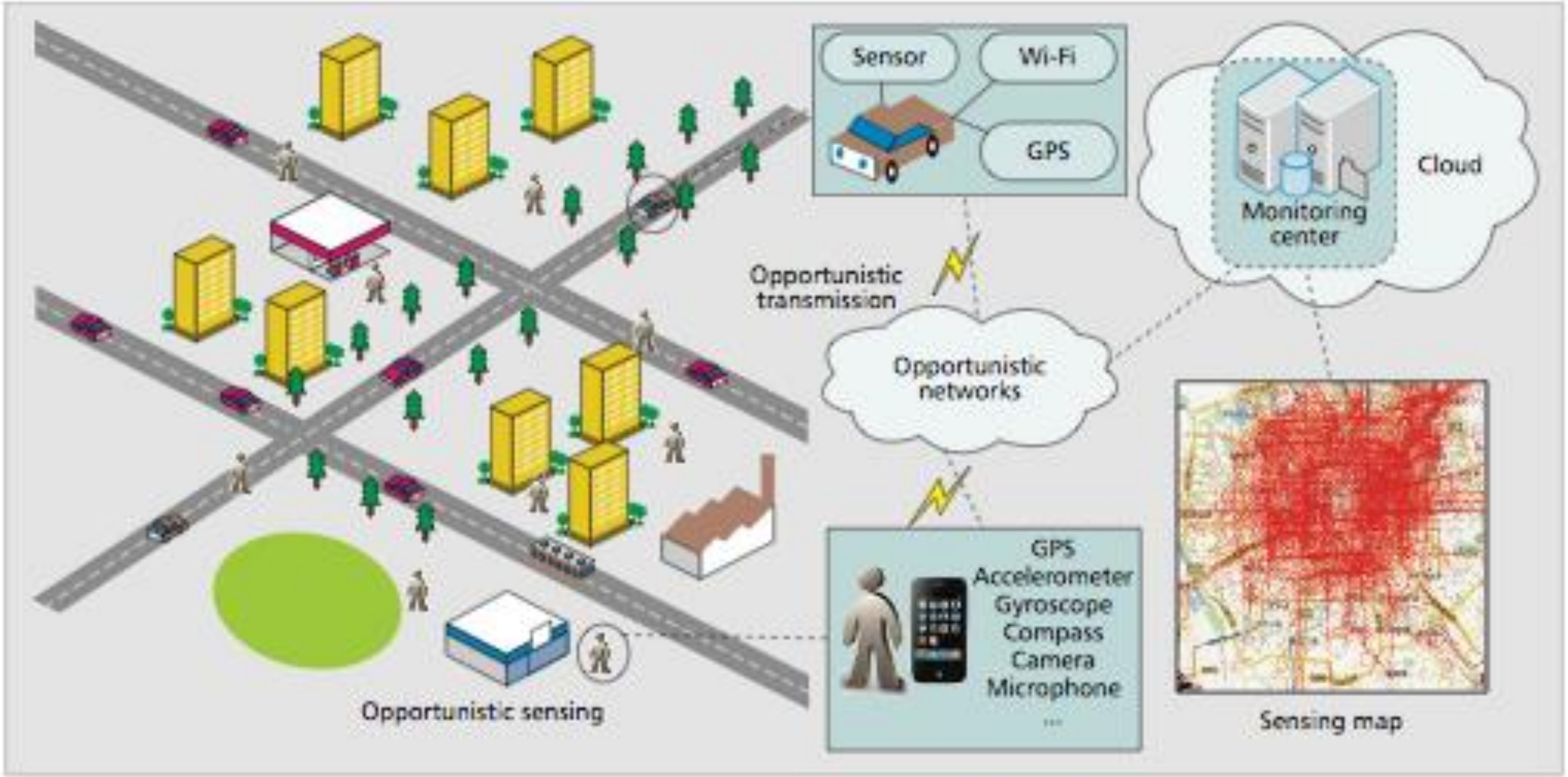
Mobile Crowd Sensing (MCS) or People/human- centric sensing

- Human involvement is one of the most important **Characteristics** of MCS :
  - Compared to traditional sensor networks, human mobility offers **unprecedented opportunities** for both **sensing coverage** and **data transmission**.
- Two classes of sensing paradigms in MCS:
  - Participatory Sensing (Participants to consciously opt to meet the application requests by deciding **When**, **what** and **How** to Sense)
  - Opportunistic Sensing (fully unconscious, namely the application may run in the back- ground and opportunistically collect data)
- **Opportunistic sensing** more support Large-scale deployments and application diversity
- Two classes of transmission paradigms
  - Infrastructure-based transmission (users reporting and accessing sensory data through the Internet by 4G/5G)
  - Opportunistic transmission (opportunistic data forwarding among mobile users through intermittent connections Wifi or Bluetooth)

Opportunistic Transmission

- It works well without requiring any centralized server or infrastructure for communication and management.
- It **reduce the workload** of cellular networks in dense areas.
- **Energy-efficient**: less expensive, which is very important because most users hope to save battery energy and data usage on their mobile devices.

- **Characteristics of human mobility** from the **perspectives** of both **sensing** and **transmission.**



Source: Opportunistic Transmission

HUMAN INVOLVEMENT: A DOUBLE-EDGED SWORD

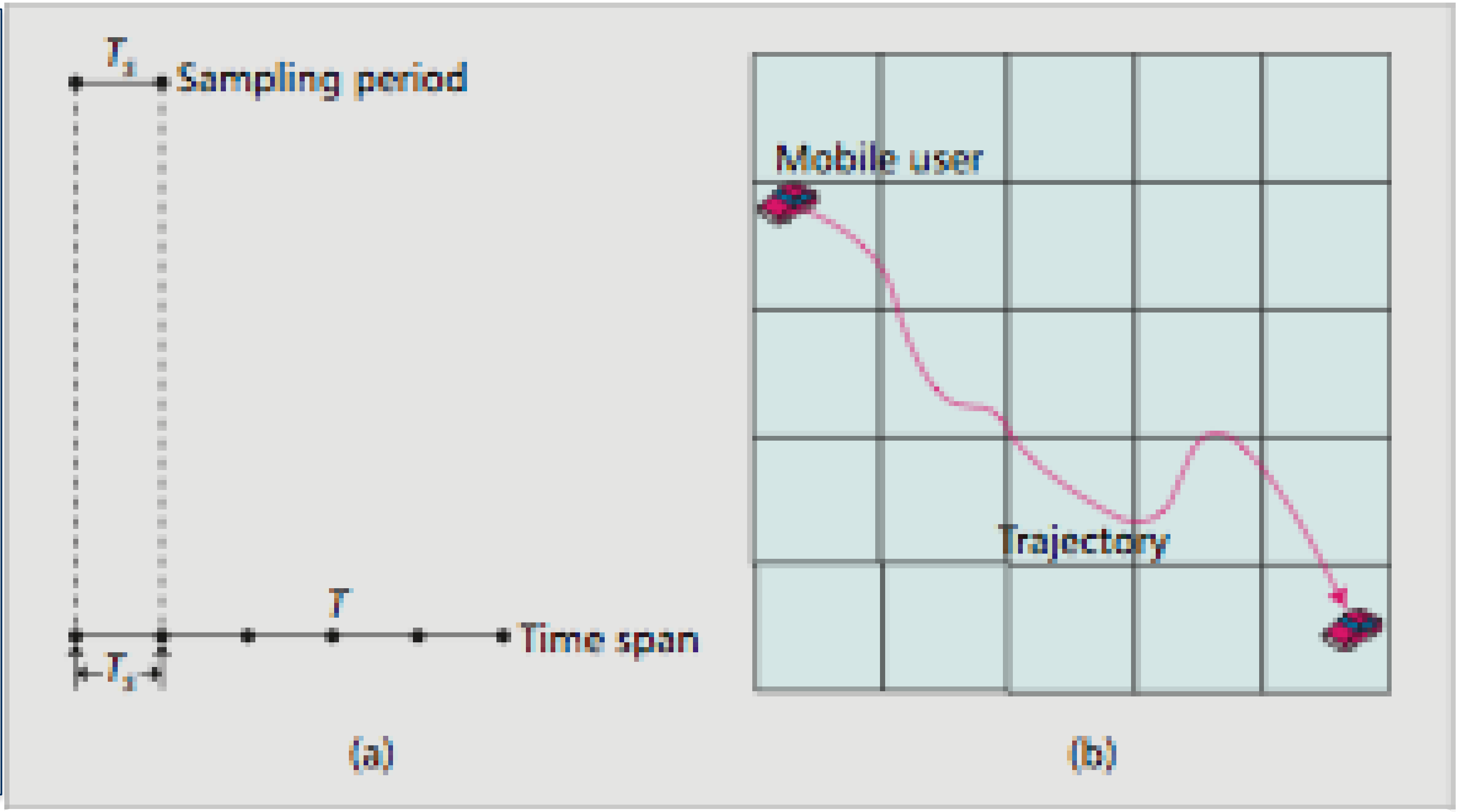
- One of the important characteristics of MCS is deeper involvement of humans in:
  - Data-to-decision process, Sensing, Transmission, Analysis of big sensory data and Decision making
- Double- edged sword
  - **Easier to deploy the network at lower cost** –› human mobility can be exploited to improve sensing coverage and data transmission.
  - **Easier to maintain the network** -› mobile nodes are always managed and maintained in good condition by their holders.
  - **Extensible and flexible** -› users to adapt to the expansion of the system scale.
  - **Availability of sensors**, and the **data quality always change** over time due to the randomness of human mobility and the dynamics of human contexts.
  - Human involvement naturally brings privacy concerns.
  - participating in MCS, mobile users consume their own resources.

OPPORTUNISTIC CHARACTERISTICS OF HUMAN MOBILITY

- **Unique characteristics** such as: Spatiotemporal correlation, hotspots' effects, and sociality. (uncontrolled mobility)
- **Identifying these characteristics** is **beneficial** to **estimate**: sensing quality, perform network planning, design efficient sensing and transmission protocols, and develop accurate mobility models.
- **OPPORTUNITIES IN OPPORTUNISTIC SENSING**
  - How can the sensing opportunities and sensing quality be measured?
    - **inter-cover time** to characterize the opportunity with which a sub-region is covered (in time domain).
  - How many mobile users can provide enough sensing opportunities to achieve the required sensing quality?
    - The opportunistic coverage ratio to characterize the relationship between the sensing quality and the number of users.
- **OPPORTUNITIES IN OPPORTUNISTIC TRANSMISSION**
  - The transmission opportunities of human mobility and their impacts on the data delivery performance have been intensively studied and relatively well understood in opportunistic networks or delay-tolerant networks (**DTNs**).
  - The **inter-contact time** is one key metric to characterize the transmission opportunities of the same couple of mobile users.
  - Human sociality has a key impact on human mobility, since it decides the spatial properties of human mobility (i.e., where people move). human sociality has been considered an important factor affecting the performance of opportunistic forwarding protocols

Inter-cover time

- The coverage in MCS is time-variant due to human mobility.
- *T* into multiple sampling periods of *Ts*.
- In the space domain, we divide the monitoring region into a set of grid cells (b).
- A grid cell is said to be covered by a mobile user only when a new sampling period arrives and the location of the mobile user is just within the area of the grid cell.
- The inter-cover time is defined as the time elapsed between two consecutive periods of coverage of the same grid cell.
- Shorter inter-cover time results in better sensing quality for a grid cell.



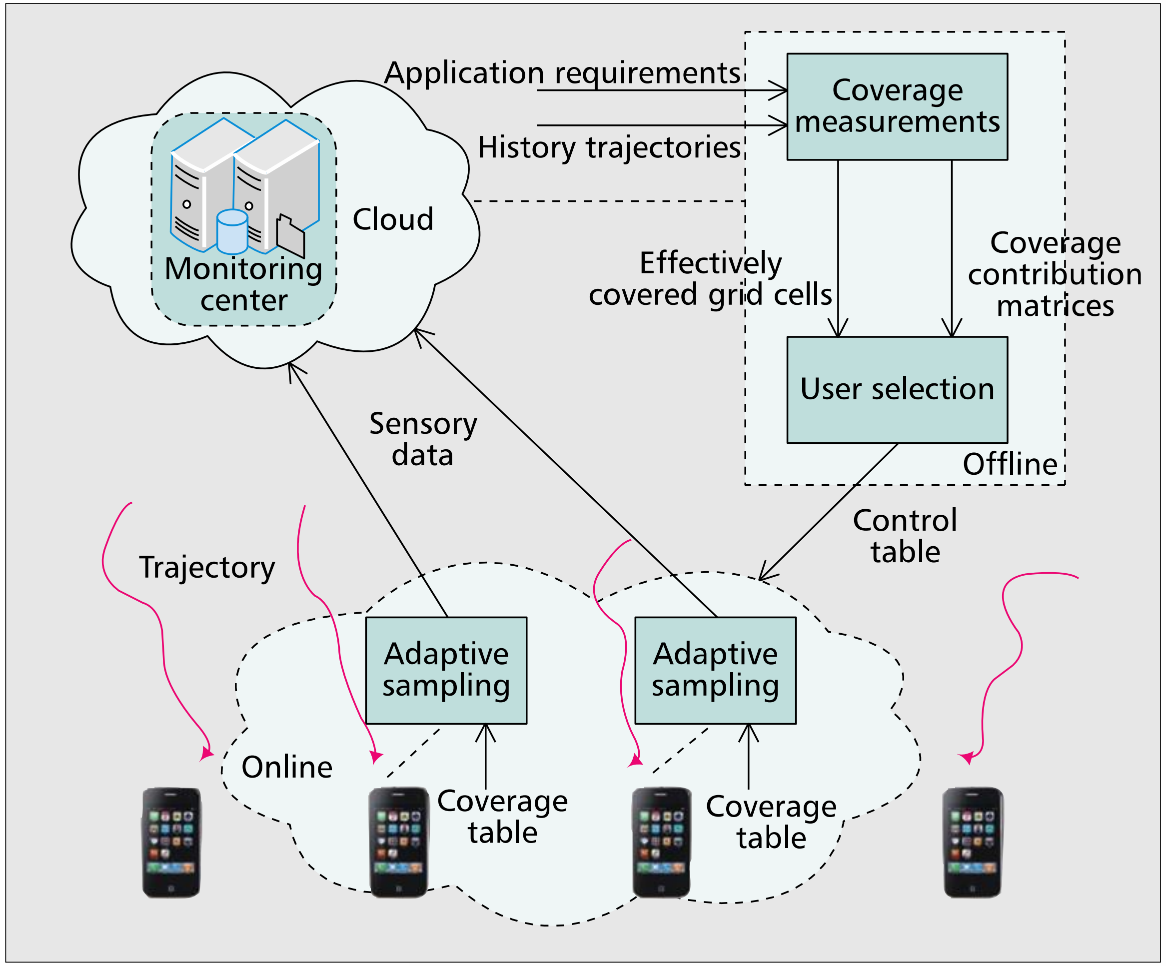
Source: Opportunistic Transmission

Existing social-based opportunistic forwarding protocols

Social Metrics	Meaning	Typical Protocols
Centrality	The Social Position of a user	SimBet, PeopleRank,BUBBLE
Similarity	The social distance between two users	SimBet, MobiSpace
Social Relationship	Acquaintance, friend, or stranger	SMART
Social Structure	Virtual community or physical hotspots	BUBBLE, Hotent

Opportunistic Sensing

- A commonly used method for opportunistic sensing is to make every mobile user sense **periodically**
- Many redundant data samples may be produced by a large number of mobile users
- In order to reduce data redundancy and improve energy efficiency, it is necessary to design a **cooperative sensing method** to control sensing activities of mobile users such that they produce just enough data samples for the application



Source: Opportunistic Sensing

Stockholms universitet

OPPORTUNISTIC TRANSMISSION

- **Many opportunistic forwarding protocols have been proposed in opportunistic networks**
- **Epidemic Routing**
  - Without knowing anything about the mobility of users.
  - Grasp each forwarding opportunity, thus resulting in the minimum delivery delay under ideal conditions (unlimited resources such as bandwidth and buffer space) at high cost.
  - In order to reduce the cost, many variants of epidemic routing have been proposed (*k*- hop schemes, probabilistic forwarding, spray- and-wait, etc)
  - Some protocols : identify the best relay node to achieve better trade-off between delivery delay and transmission overhead by exploiting the context information (e.g., the mobility of a node and its current energy level). Several works have exploited human sociality for improving opportunistic transmission performance. *Most existing social-based opportunistic forwarding protocols use traditional approaches in social networks or* ***ego networks*** *to evaluate social metrics*. These approaches have high spatiotemporal complexity due to transient user contacts and intermittently connected environment, and hence **cannot** be applied in large-scale opportunistic scenarios.
  - **It is necessary to develop a lightweight approach to exploiting human sociality for improving opportunistic transmission performance.**

THE IMPACT OF DATA FUSION

- **Data fusion** (or data aggregation) for two reasons:
  - Users may be interested only in the aggregated results of sensory data (e.g., the average temperature or noise level).
  - Sensory data collected in close proximity or time periods may be highly correlated, and data fusion can effectively eliminate redundancy and hence reduce network overhead.

Forwarding Sensory data without (a) and with(b) Data Fusion(DF)

Source: Opportunistic Sensing

Section III:

Internet of Things Data Taxonomy

4 Phases of the Internet

Phase 1:
Connectivity

Digitize Access to Information

- Email
- Web Browser
- Search

Phase 2:
Networked Economy

Digitize Business Process

- E-commerce
- Digital Supply Chain
- Collaboration

Phase 3:
Collaborative Experience

Digitize Interaction (Business & Social)

- Social
- Mobility
- Cloud
- Video

Phase 4:
Internet of Things

Digitize the world Connecting:

- People
- process
- Data
- Things

Small Data ***More Data*** ***Big Data***

Stockholms universitet

Data-centric Internet of Computers

Data-centric IoT

Physical World

People Generate | Things Generate

Web & Web of Data

People Gain | Things Gain

Information

People Discover | Things Discover

Knowledge

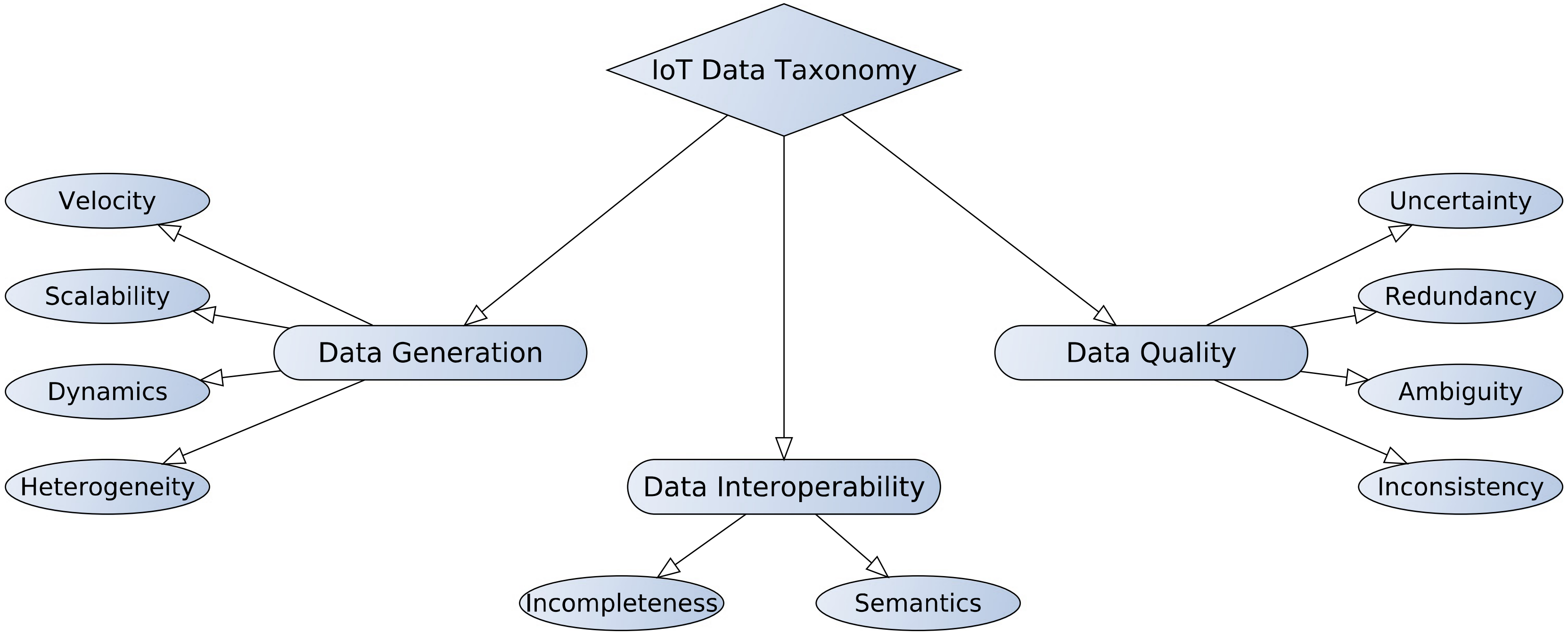
People Propose | Things Propose

Solution

Stockholms universitet

IoT Data Taxonomy

- Data Generation
  - Velocity
  - Scalability
  - Dynamic
  - Heterogeneity
- Data Quality
  - Uncertainty
  - Redundancy
  - Ambiguity
  - Inconsistency
- Data Interoperability
  - Incompleteness
  - Semantics



Source: Data Taxonomy

Stockholms universitet

IoT & Data Stream

- A data stream
  - A sequence of data objects
  - Continuously generated at a rapid rate
  - Data object can be described by a multi-dimensional attribute vector (Continuous, categorical or mixed attribute space).
- Characteristics of data streams:
  - Continuous arrival of data objects.
  - Disordered arrival of data objects.
  - Potentially unbounded size of a stream.
  - Normally no persistence of data objects after being processed.
  - Changing probability distributions of the unknown data generation process.

Cont. IoT & Data stream: General data stream processing

- Data stream can be generated in:
  - **Various scenarios** (Network of sensor nodes, stock market and network monitoring)
  - **Communication across sensor networks** or with a **centralized server** requires the largest amount of energy as sensing consumes less energy.
  - Consequently, **sensor data processing techniques**, including data aggregation, data compression, modeling and online querying, should be performed on-site or in-network to reduce communication cost
  - Demands on efficient data processing algorithms for sensor systems arise due to the limitations of computational power of sensor node.
  - Stock market and network monitoring systems, there also exist challenges in processing high-rate data streams.

Stockholms universitet

Cont. IoT & Data stream: Query Processing

- Aggregate Queries
  - Import class of queries in sensor system (*MIN*, *COUNT* and *AVG* operators)
  - Techniques have been proposed to efficiently process these aggregate operators in sensor systems to effectively reduce power consumption.
  - the properties of the aggregate functions in-network partial data could be preprocessed first, then be utilized to produce the final results for the issued queries.
- Join queries
  - To evaluate the query, stream readings from the sensors in regions R1 and R2 should be joined first before determine an object was detected in the two designated regions. Return the objects that were detected in both regions R1 and R2.
- Top-k Monitoring
  - most recently communicated top-k answers by maintaining some specified arithmetic constraints at the stream sources.
  - User specified error tolerance is also considered in order to provide high-quality answers. This technique can help reduce the overall communication cost between different sources.
- **Continuous queries:** To monitor designated changes in an environment, sensors are typically required to answer queries in continuous manner.

Stockholms universitet

Cont. IoT & Data stream: Stream mining

- Clustering
  - Grouping a set of objects in such a way that objects in the same group.
  - typically continuously cluster objects on memory constrained devices
  - Requirements to consider when designing algorithms for clustering data streams :

    1. providing clustering results via fast and incremental processing of data objects. 2. rapidly detecting new clusters or changes of existing clusters. 3. scaling to the potentially unbounded number of objects in data streams. 4. providing a model representation that is consistently compact regardless of the number of data objects. 5.rapidly detecting the presence of outliers and acting accordingly; and 6. dealing with different data types, such as XML trees, DNA sequences, GPS temporal and spatial information.
- Classification:
  - Uses prior knowledge to guide the partitioning process to construct a set of classifiers to represent the possible distribution of patterns. **Classification** is a **supervised learning process** whereas **clustering is an unsupervised learning process**
- Outlier and anomaly detection
  - Main task is to find data points that are most different from the remaining points in a given data set.
- Frequent item set mining
  - To find sets of items or values that co-occur frequently, or in other words, to find co- occurrence relationships in a transactional data set.

Cont. IoT & Data stream: RFID data stream processing

- RFID data cleaning (uncertainty and unreliability)
  - The false reads in RFID streams can be classified into two categories:
    - Missing-reads
    - Cross-reads
- RFID data inference and compression
  - Techniques will need to clean RFID data first and then infer to the high level information about the tagged objects, i.e., location and containment relationships.
- RDF triple stream processing
  - Linked Data is **a method** for publishing structured data and interlink such data to make it more useful
  - Data from different sources can be **connected** and **queried** in the form of Linked Data
  - Stream reasoning can provide the abstractions, foundations, methods and tools required to **integrate data streams**
  - Substantial research efforts have been put forward, **focusing on how** to apply reasoning on streaming data, how to publish raw streaming data and connect them to the existing data sets on the Semantic Web, and how to extend the SPARQL query language to process streaming data.
- Linked stream processing and reasoning
  - Efforts to apply the linked data principles to stream (sensor) data have been initiated and this wealth of information could be easily included in the Linked Data cloud.
  - Continuous SPARQL
- Extracting RDF triples from unstructured data streams
  - Although the current Linked Open Data (LOD) cloud has tremendously grown over the last few years, it delivers mostly encyclopedic information (such as albums, places, and kings) and fails to provide up-to-date information.

Stockholms universitet

IoT & Data storage models

- New Architecture
  - Data in C-store (column-store) is not physical stored using its related relational logical data model
  - Most row stores implement physical tables directly and then add various indexes to speed access, C-store implements only projections
  - Projections are sorted subsets of the attributes of a table.
- Large-scale storage in distributed environments
  - Users from different locations may have different data consumption needs
  - Propose a selective replica strategy which supports replica of tables in the Web databases at record level to alleviate the overly replicated issue.
  - Replica strategy, each replica location stores a full or partial *copy of the replicated table depending on the data needs. Selective replica strategy*: ***Consistency, Availability*** and ***Partition tolerance***.
- Storage on resource-constrained devices
  - Storage issues also arise in resource-constrained scenarios in IoT. Sensor networks, communication activity normally plays a more important role than storage. Batch data collection, delay-tolerant mobile applications, and disconnected operations in static networks, the storage-centric paradigm becomes more critical. Decreasing costs and increasing capacity of storage hardware.
    - » ***SolarStore, TinyDB and Antelope***

IoT & Search techniques

Relevant areas such as the Deep Web, Semantic Web and then searching things in the IoT environments.

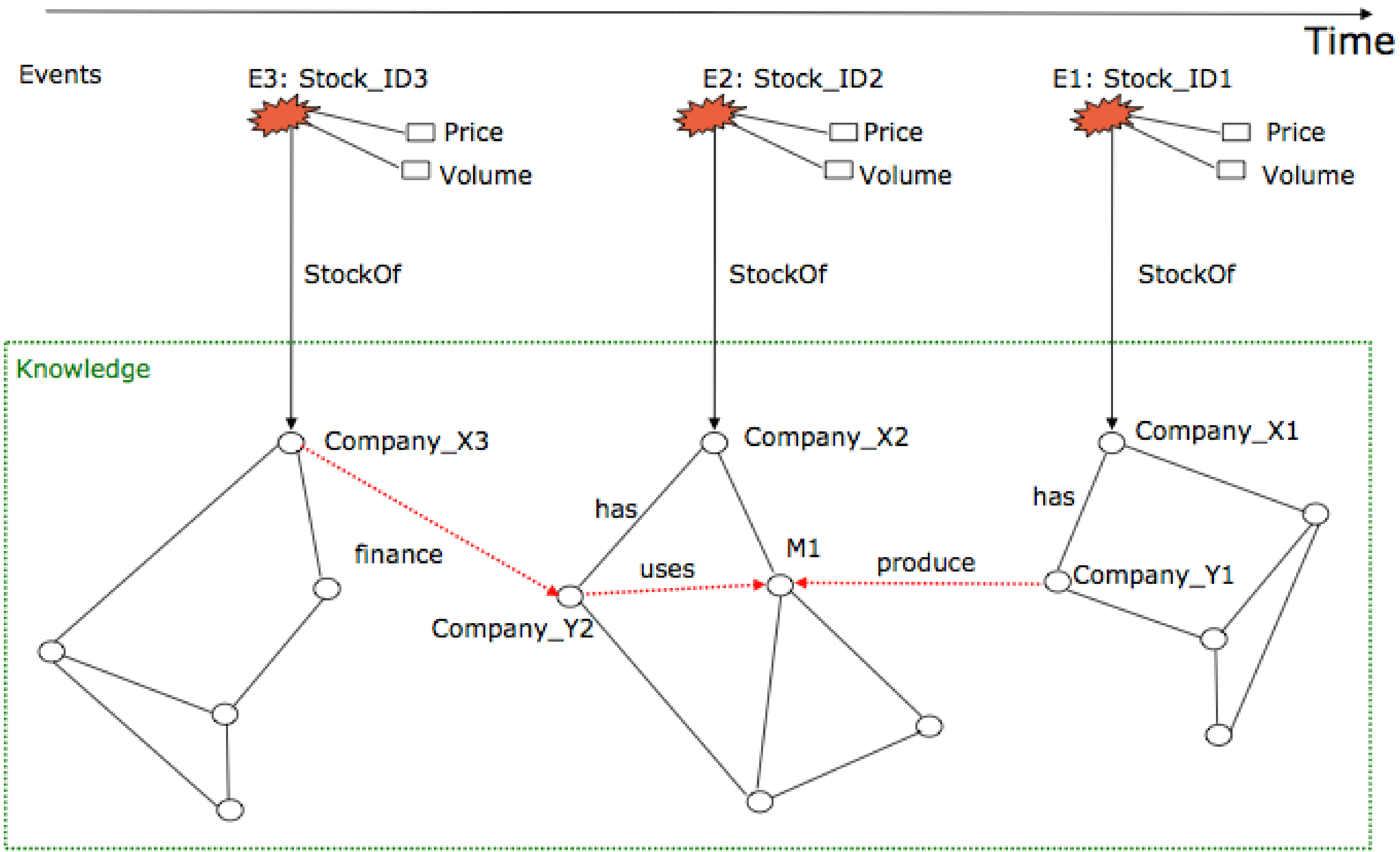
- Deep Web (DW) and Semantic Web(SW)
  - DW: *That is not indexed by standard search engines.*
  - *DW is estimated to be* ***several orders of magnitude larger*** *than that of the so-called Surface Web (the Web that is accessible and indexable by text search engines).*
  - *DW provides a* ***wealth of hidden data*** *in* ***semi-structured form***, ***accessible through Web forms*** *and* ***Web services***.
  - *SW grows as a common structured data source.*
  - *SW is standard (W3C) with Resource Description Framework (RDF) and Web Ontology Language (OWL).*
  - *SW aims to unify the way semantic information is* ***stored*** *and* ***exchanged***.
  - *SW makes it possible for machines themselves* ***to not just read***, *but also "****understand****" the data from* ***data sources***, *which enables machine to machine communication*
- Web Search (**Virtual Twitter User (VTU)** :which can actively post tweets on the Web. VTU monitor conditions such as temperature, humidity and radiation levels. It then automatically tweets the resulting data via Twitter
  - Real-time web search
    - Twitter handled more than 85 million tweets per day in 2015 ,
    - real-time search service is indeed very challenging in such Large-scale micro-blogging systems
  - Searching information over RDF data
    - Semantic Web community uses a very different data model, which is RDF, MIDAS-RDF, a distributed P2P RDF/S repository that is built on top of a distributed multi-dimensional index structure
  - Collaborative web search
- Search of Things in IoT
  - Keywords based search of Things
  - Collaborative search of Things

IoT & Complex Event Processing(CEP)

Data streaming techniques process incoming data through a **sequence of transformation based on Common SQL operators.**

- The Complex Event processing model **views** the **information in** the streams as events in the physical world.
- Systems for event processing and in particular **event recognition** (event pattern matching) **accept** a stream of **time-stamped,** simple or low-level events as input.
- A low-level event is the result of applying a **computational derivation process** to some other event, such as an event coming from a sensor.
- Using low-level events as input, (complex) event processing systems identify composite or high- level events of interest.
- They are also collections of events that satisfy certain patterns.
- SASE (a complex event processing system) and other new query evaluation model is designed for processing pattern matching over RFID streams, employing a new type of automaton that comprises a nondeterministic finite automaton (NFA) and a match buffer, named NFA.
- Nested CEP language called NEEL is proposed to support the flexible nesting of AND, OR, Negation and SEQ operators at any level

Example of CEP

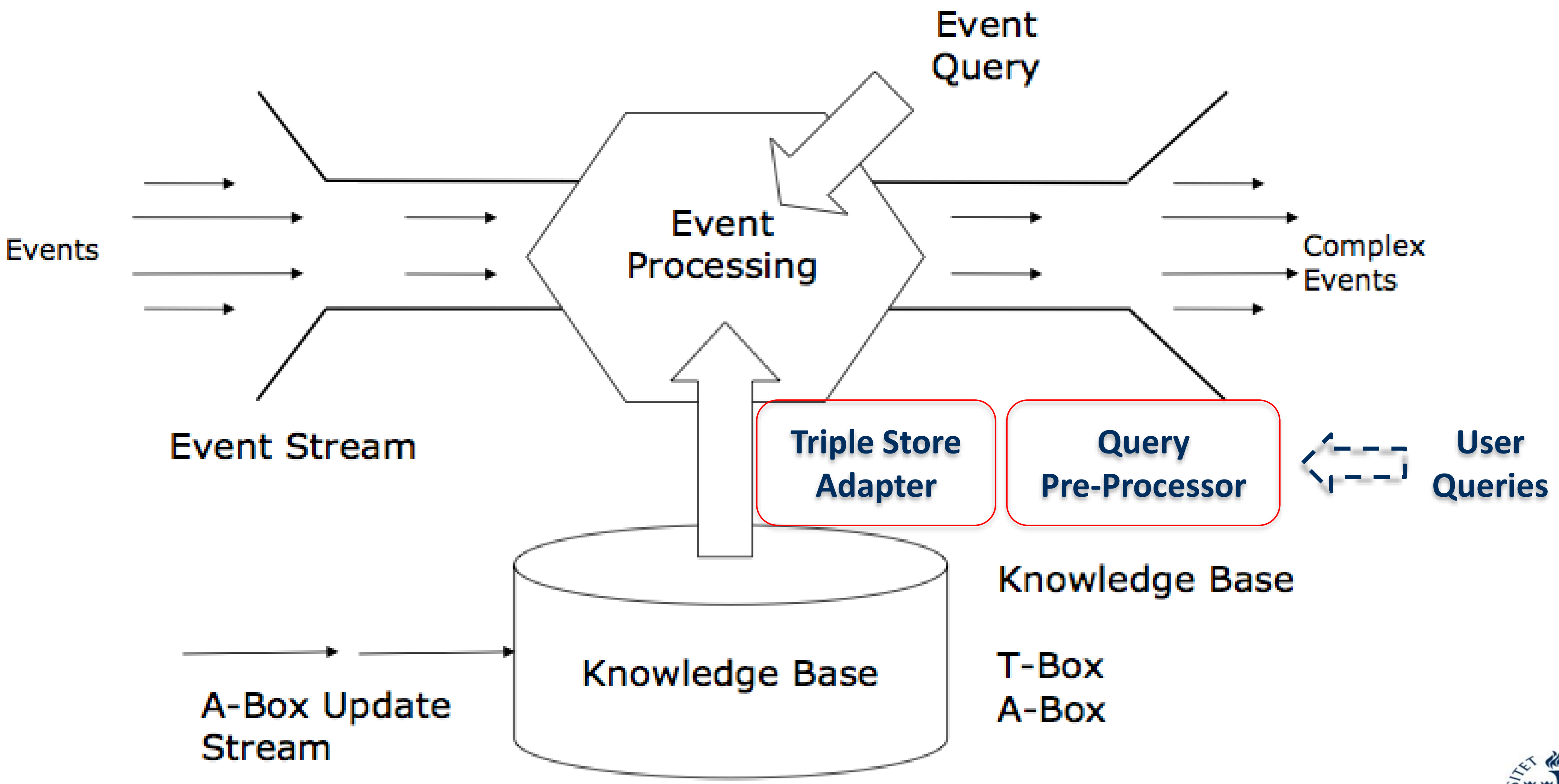


IoT & Semantic Complex Event Processing

The **combination of event processing** and **knowledge representation** can lead to **novel semantic-rich event processing engines.**

- Semantic Complex Event Processing
  - Semantic models of events can improve event processing quality by using event meta- data in combination with ontologies and rules (i.e., knowledge bases).
  - The fusion of background knowledge with data from an event stream can help the event processing engine to know more about incoming events and their relationships to other related concepts.
- Semantic event enrichment
  - The usage of background knowledge about events and their relations to other concepts in the application domain can improve the expressiveness and flexibility of CEP systems.
  - The process of reducing information incompleteness is called event enrichment.
  - Several challenges are identified for event enrichment, including determination of the enrichment source, retrieval of information items from the enrichment source, finding complementary information for an event in the enrichment source and fusion of complementary information with the event.
- Approximate semantic matching
  - To achieve approximate matching, semantic selection and inexact selection are used.
  - The semantic selection evaluates pattern constraints based on the semantic equivalence of attribute meanings captured by the event ontology instead of syntactic identical attribute values.
  - The inexact selection selects events and allows a limited number of mismatches to detect relevant patterns.

Knowledge-based Event Processing



Source: Fusion of Event Stream and Background Knowledge

Potential IoT Applications

- Smart Cities and homes
  - IoT can connect billions of smart things and can help capture information in cities
  - In a modern city, lots of digital data traces are generated there every second via cameras and sensors of all kinds
  - IoT can facilitate resources **management issues** for **modern cities**
  - static resources (e.g., fire stations and parking spots) and mobile resources (e.g., police cars and fire trucks) in a city can be managed effectively using IoT technologies.
- Environment Monitoring
  - IoT technologies can also help to monitor and protect environments thereby improving human's knowledge about environments.
  - IoT could improve water management in a city.
  - A device has been designed and built to monitor residential noise pollution to address and determine what kind of sound is noise.

Cont. Potential IoT Applications

- Health
  - Healthcare environment such as a large hospital or aged care could tag all pieces of medical equipment .
  - As personal health sensors become ubiquitous, they are expected to become interoperable.
  - Mobile technology and sensors are creating ways to inexpensively and continuously monitor people's health.
  - Consumers would be able to self-track their pulse and studies show heart rate monitoring could be useful in detecting heart conditions and enabling early diagnosis.
  - include data on blood sugar levels, and other biometrics collected via mobile devices.

Cont. Potential IoT Applications

- Energy
  - Home heating is a major factor in worldwide energy use IoT, home energy management applications could be built upon embedded Web servers
  - IoT applications would allow drivers to find the most fuel-efficient routes for their vehicles between arbitrary end-points.
  - Smart Grid is the concept of integrating information technologies to the legacy power grids for optimizing its operation. IoT concept can be used to provide embedded intelligence to the existing grids in developing countries. The different segments of power grids which require monitoring or control includes electricity generation, transmission, distribution and its consumption.

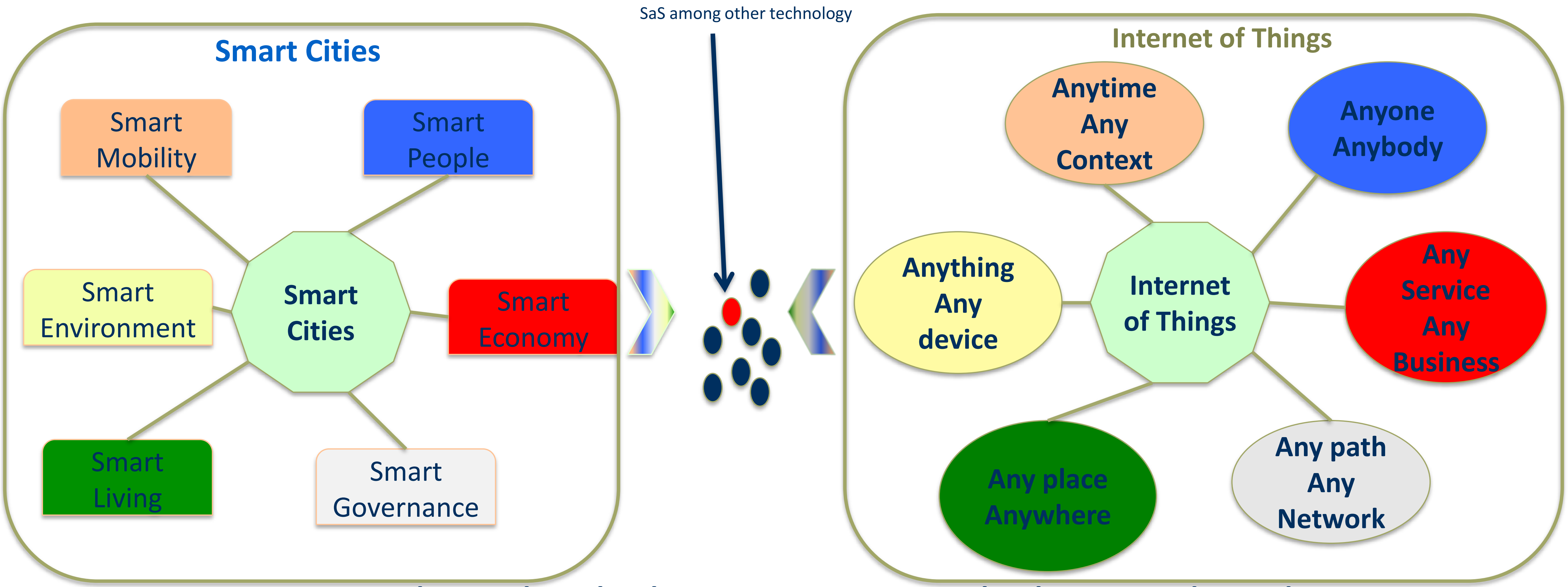
Stockholms universitet

Cont. Potential IoT Applications: **Business**

IoT technologies would be able to help to improve efficiency in business and bring other impacts on business:

- Commercial point of view
- Social and political point of view
- Personal point of view
- a set of tools and standards for tracking and sharing RFID-tagged products in IoT.
- IoT technologies could be used to make it easier to use all this data, to integrate it into various applications, and to build more flexible, scalable, global application for better (even real-time) logistics.

Sensing-as-Service (SaS)

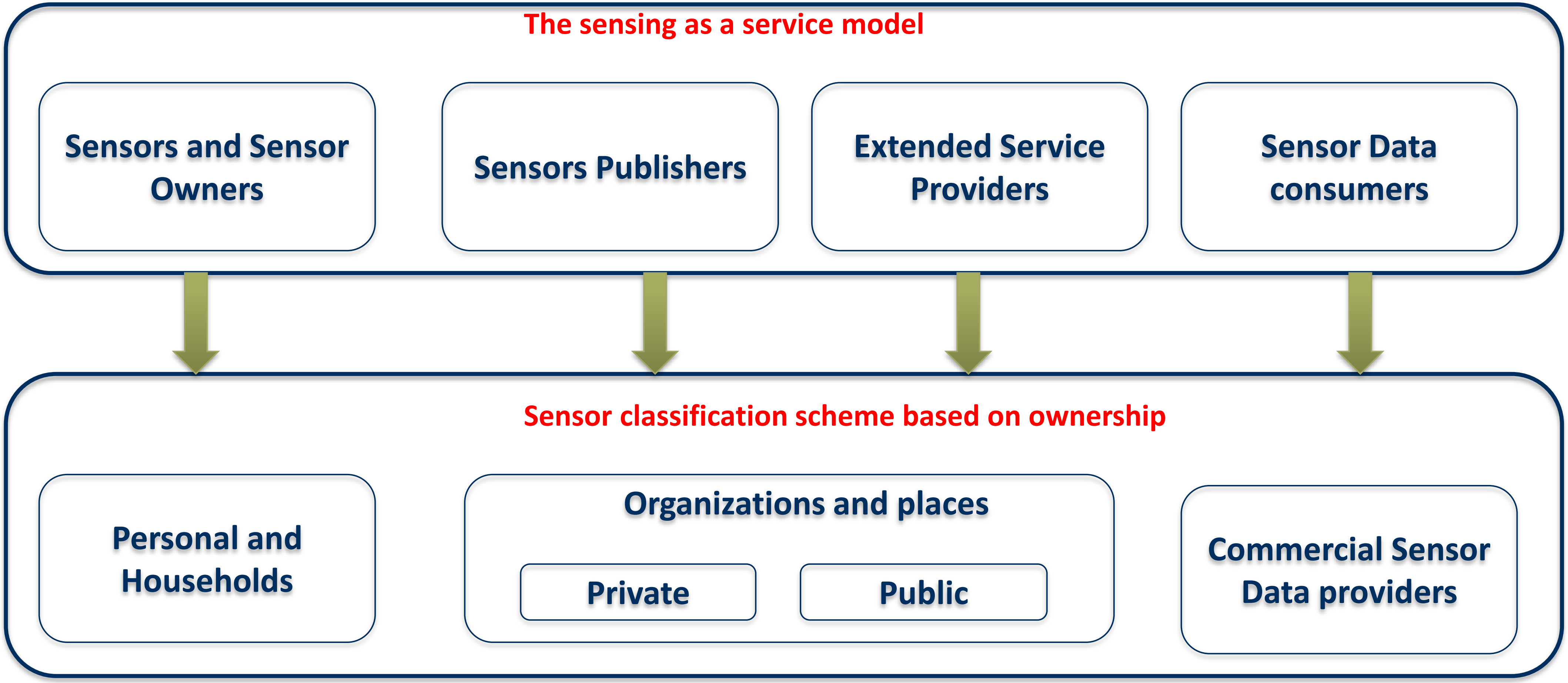


From Need towards Technology

From Technology towards Need

Stockholms universitet

Sensing as a Service model

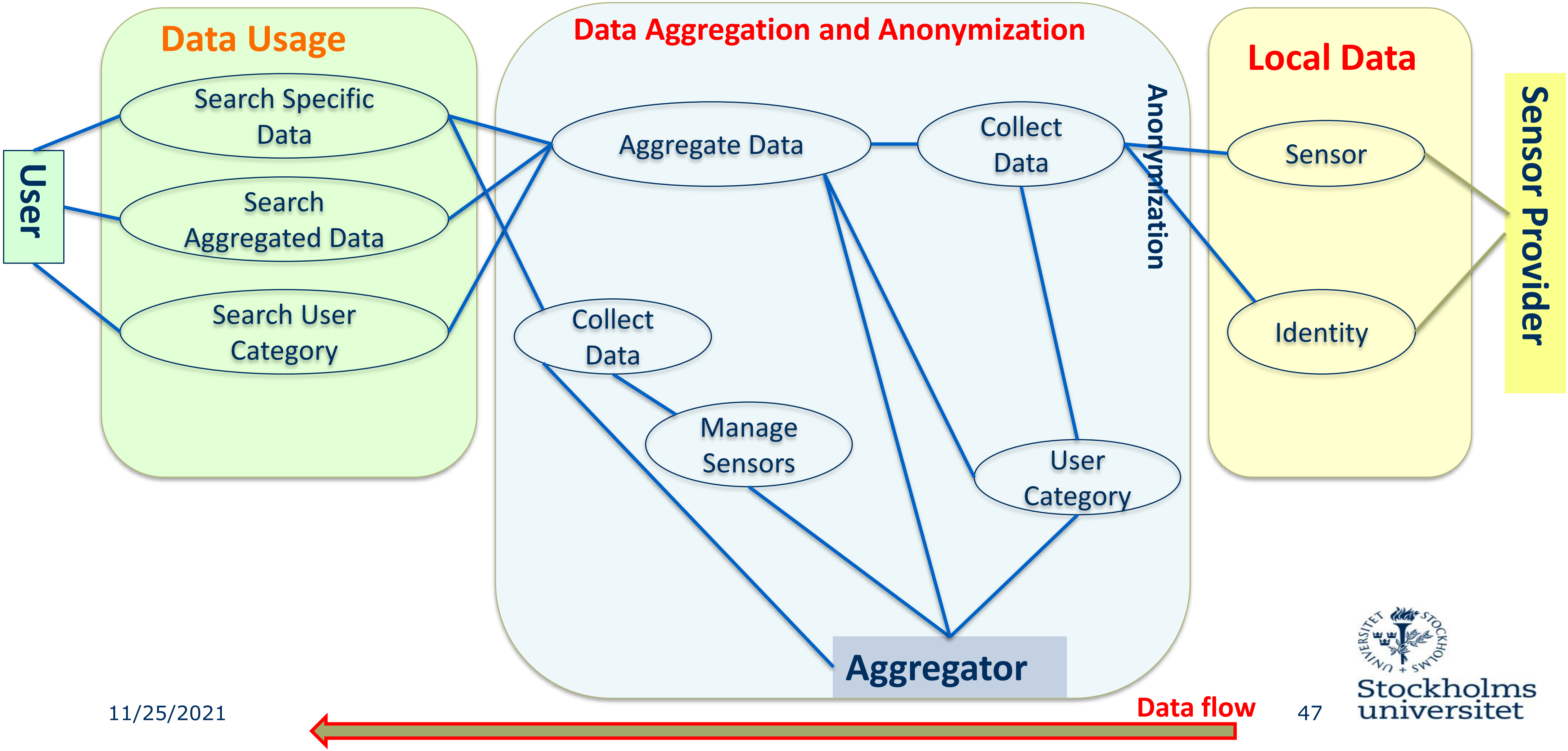


A Social Sensors Service representation

Social Sensors Services aims to aggregate measures and information collected by sensors in a **specific environment** (e.g. home) and to share them **in a bigger context (** e.g. neighborhood).

- Social Sensors services ***applicability areas*** is *Environment*, *E-government* and *Intelligent home*.
- Data measurements and information can be used for deriving knowledge (e.g *pattern analysis*) –› Related to how an **environment is operating**.
- The service is characterized by the need and possibility to access and use data stemming from sensors in different administrative domain (e.g homes, companies, public administration and government, social networks).
- Social Sensors Services sets an example of how large and independent sensor networks can cooperate in order to serve larger communities.

Social Sensors Service : *The Participants*

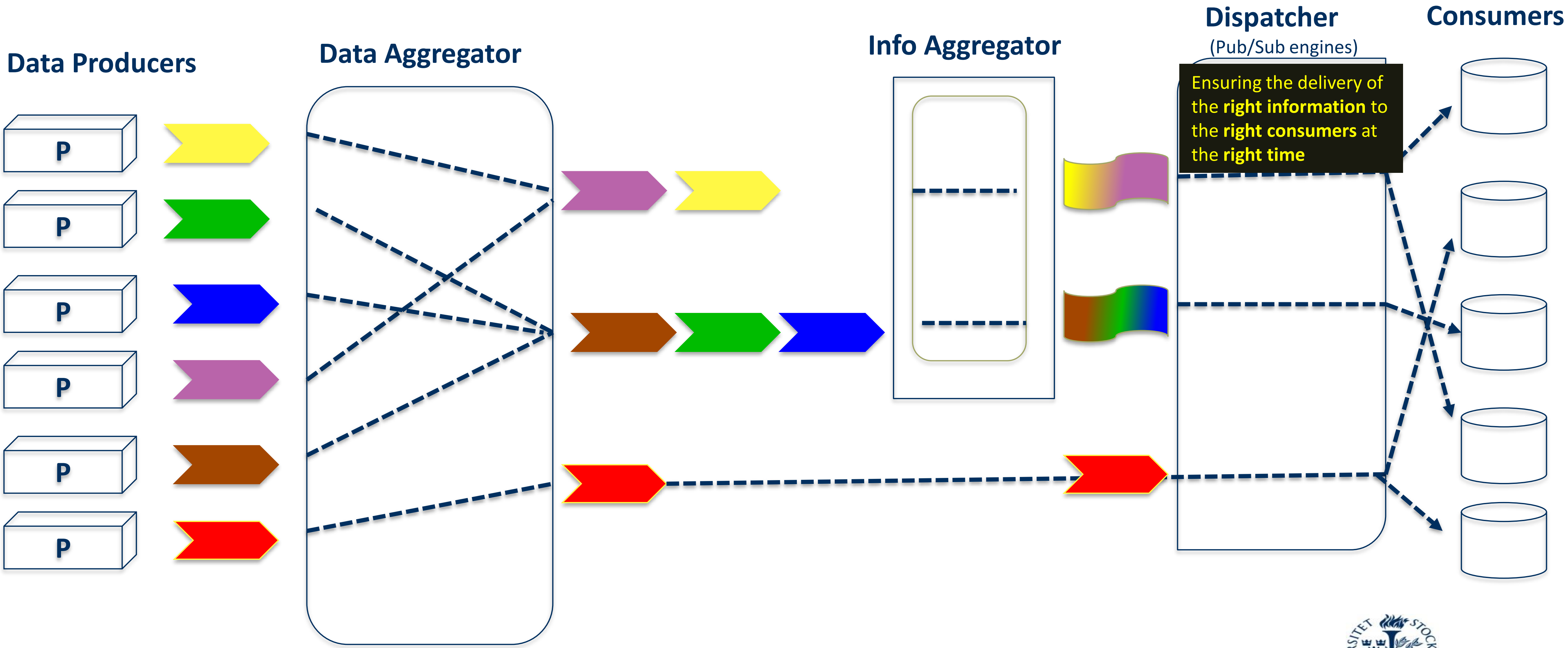


Stockholms universitet

The Participants

- The Sensor Provider
  - Actor that manages a home sensor network that monitors many aspects of daily life (http://bwired.nl/).
  - Many other activities centered on the “home”.
  - Data represents the behavior of the people living in a house, and storing and analysis of the data.
- The Aggregator
  - The actor that actually collects and properly deals with a wealth of data
  - the differences between the data representation and to try to organize data in a meaningful manner after it has been anonymized.
  - The Aggregator could be also a provider of a distributed sensor network.
- The User
  - A company or programmer that is using the sourced data in order to determine the social behavior of citizens, or a very specific subset of individuals

Social Sensors Business Model



Stockholms universitet

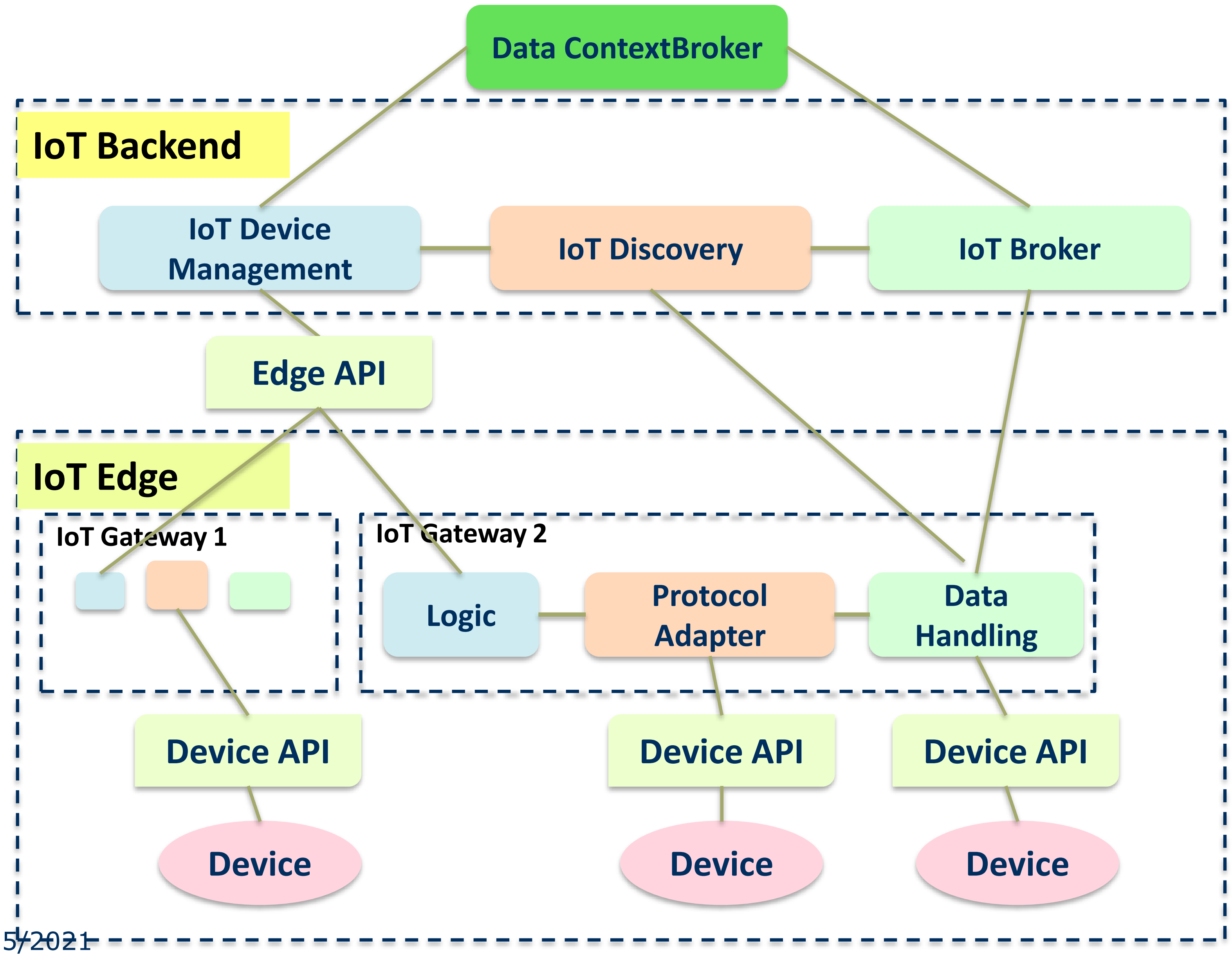
IoT Architecture Requirements

- The way Sensor data collection
- A **Pub/Sub** approach will be adopted
- Information to be intelligently performed in order to satisfy consumers requirements.
- The **storage** and **organization** of information
- ***Similar to that adopted by Twitter***: Each single sensor can be virtualized in the infrastructure of the service provider.
- Each sensor individually identified and will generate a timeline of messages (very similar to tweets) that will be passed to the consumer and **store** in the provider infrastructure.
- Virtual Sensors can be grouped to set of sensors.
- Much like Twitter allows for sending and receiving of private messages between users.

Cont… IoT Architecture Requirements

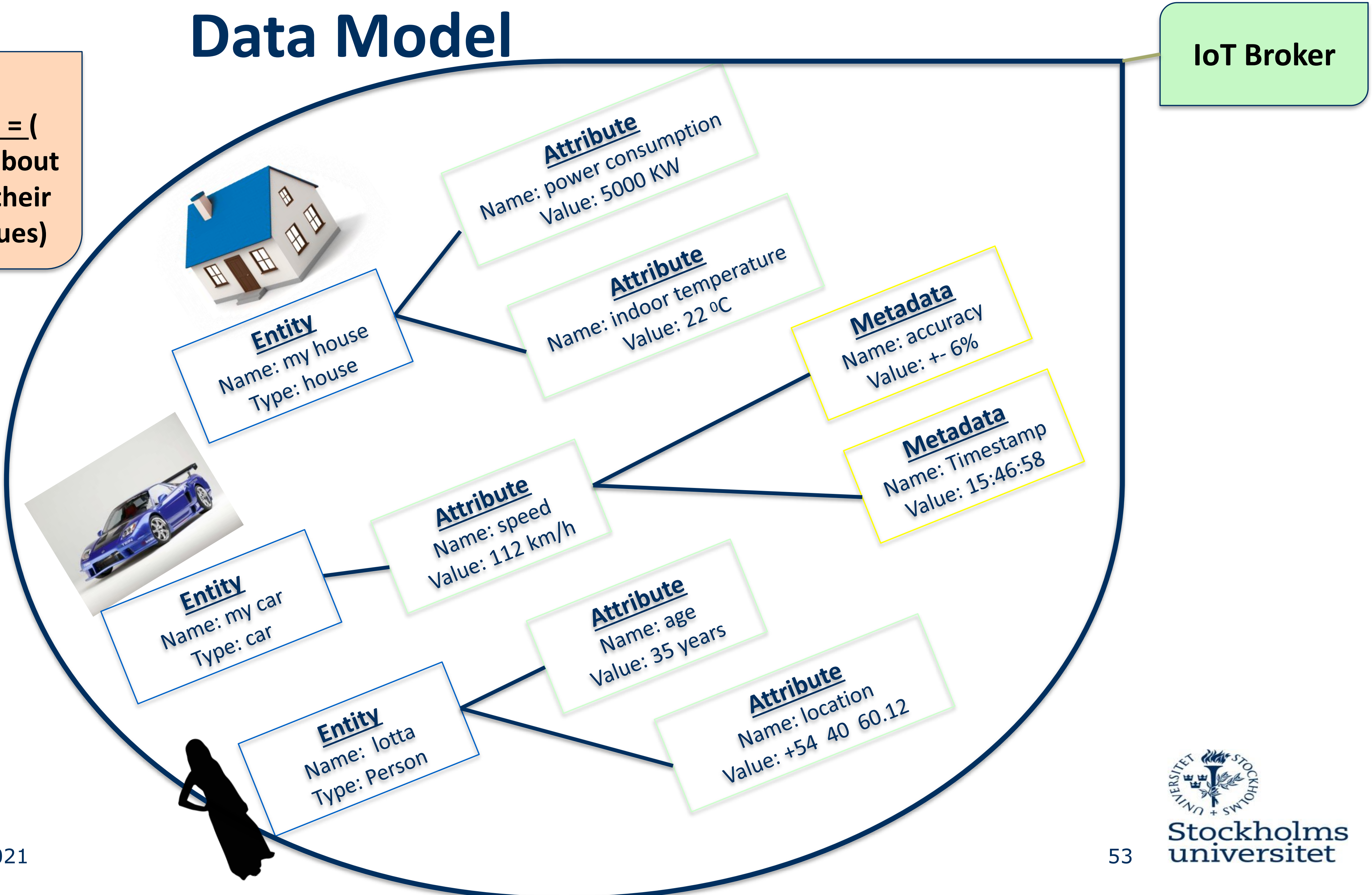
- Some of virtual sensor can be anonymized , the par of their timeline can be made available to aggregators who can use this information for running data mining algorithms to organize data or to extract information from different timelines
- The **aggregator** has the need to access data in different “pages” (e.g. to check data in different sensors timelines) in order to dynamically extract the relevant information.
- Quit similar to Twitter possibly using Hadoop or similar solutions.
- The coupling of **Publish/Subscribe** and Hadoop extended with cognitive platform capabilities seems to be a strong indication for a workable IoT architecture.

IoT Complete Architecture



Data Model

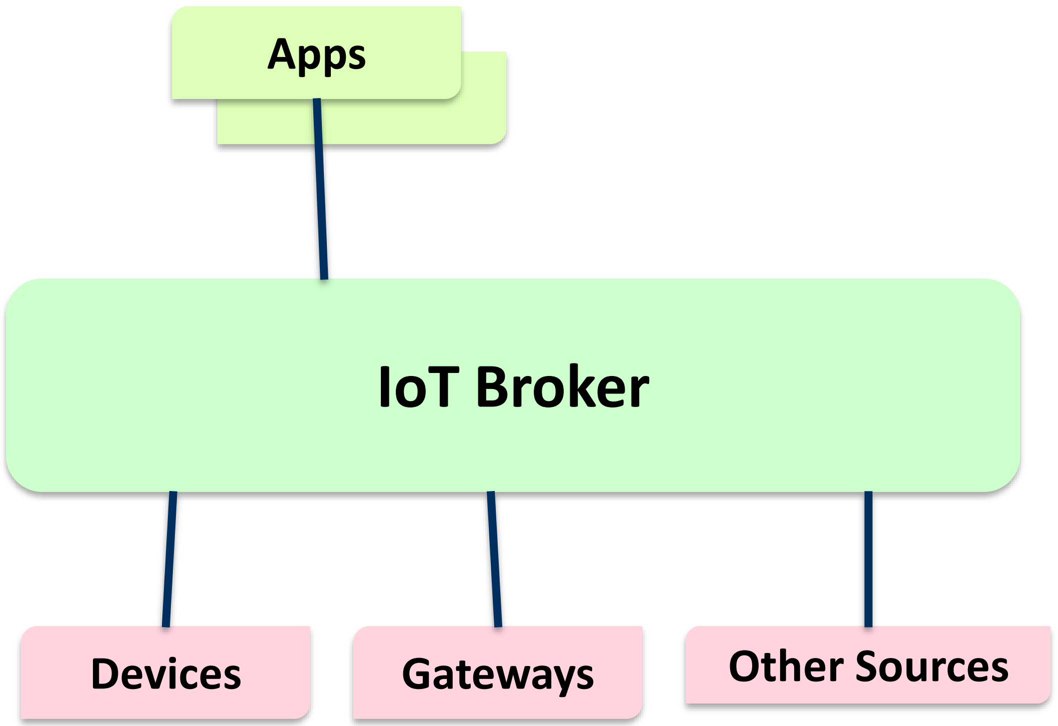
Context Information = (Information about Entities and their attribute values)



What does the IoT Broker?

Organizing the information flow for thing-level interaction

- Execute information queries on behalf of application
- Finding the resources providing the requested information
- **Collecting** & **aggregating** the received information



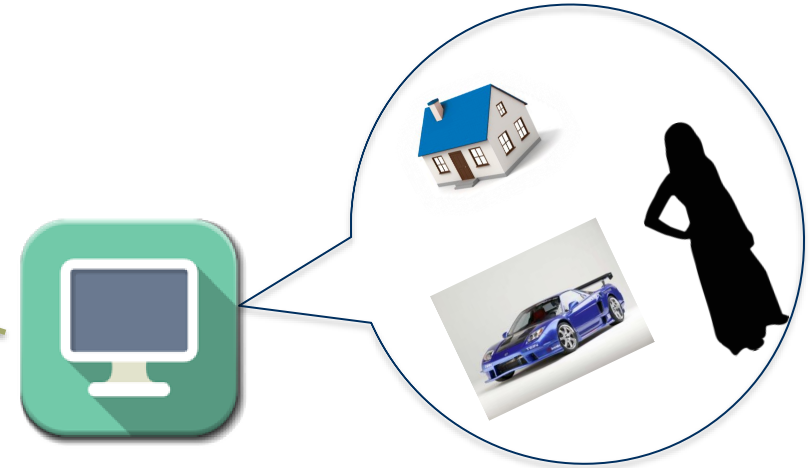
Stockholms universitet

Interface for exchanging information

Abstract Interface

Interface for exchange information about **entities** and **their attribute values**

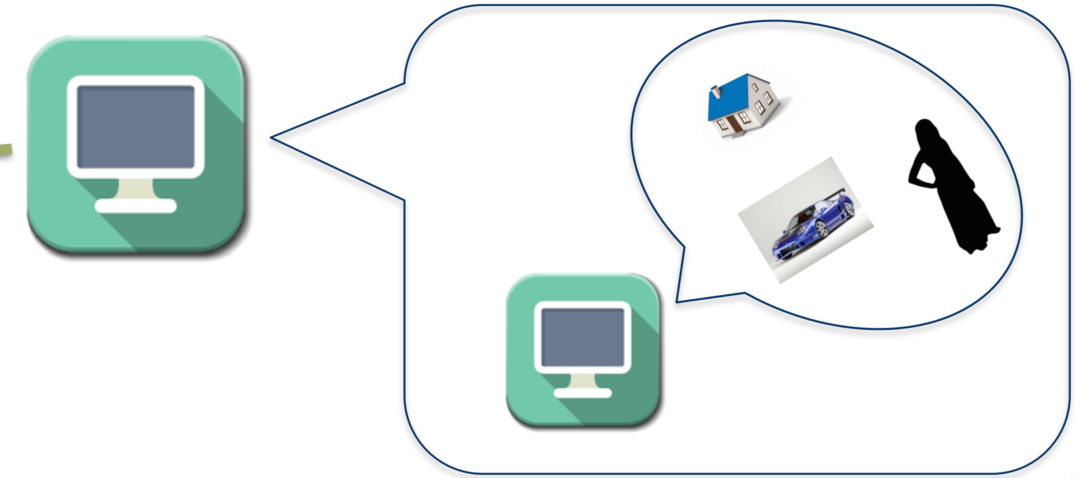
- Querycontext
- SubscribeContext + notifyContext
- UpdateContext



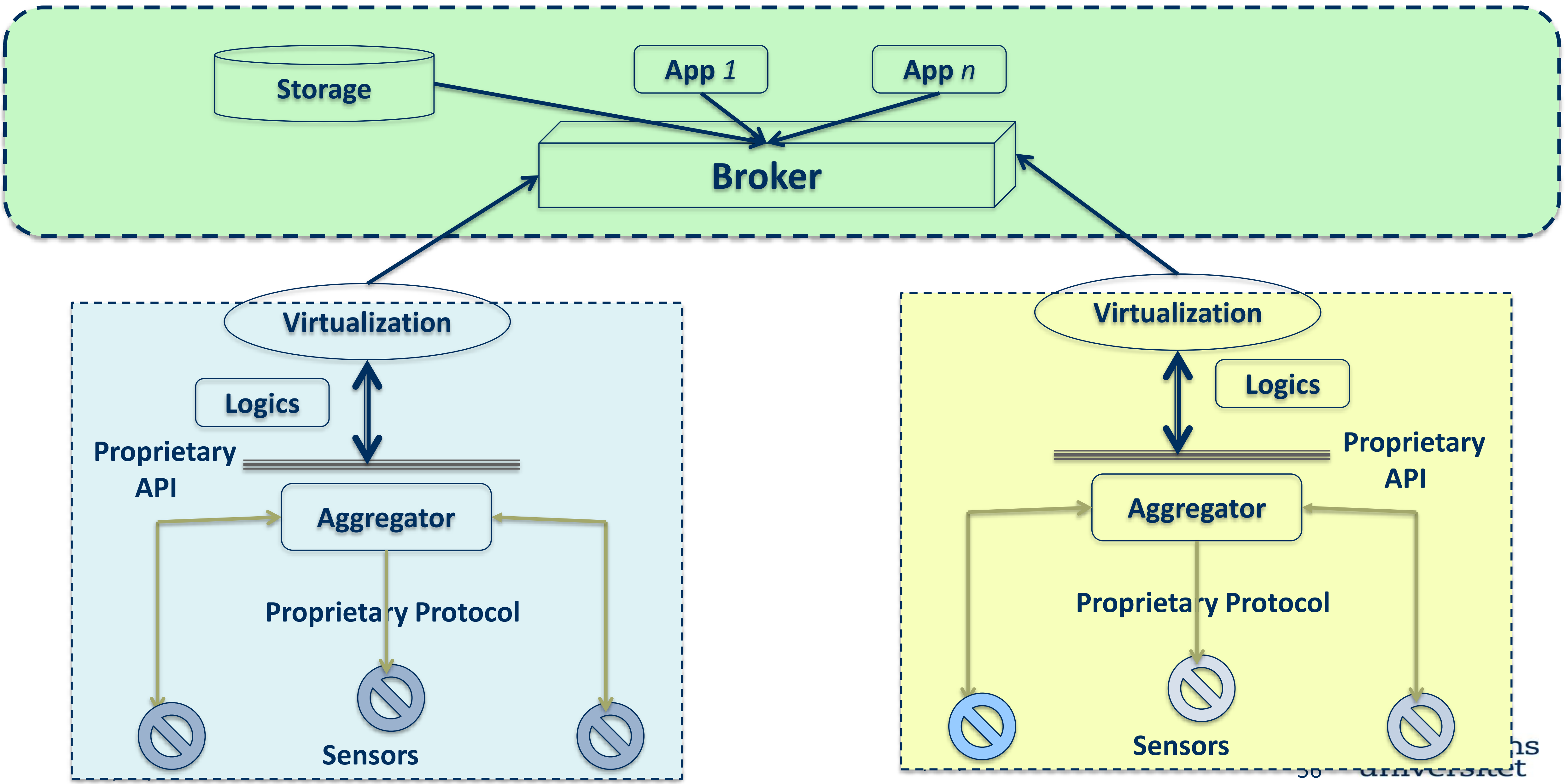
Offical standards of open Mobile Alliance

Interface for exchange information **about availability of information** about enitities and their attribute values

- DiscoverContext Availability
- SubscribeContext Availability+ notifyContext Availability
- RegisterContext



Broker of Social Sensor Derived Information



Broker of Social Sensor Derived Information

Basic Functions supported by an IoT infrastructure

- Data collection from different sources.
- Intelligent Event Dispatcher (e.g., Pub/Sub based)
- Integration of different formats and data representations
- Storage and ability to retrieve historical data from a single source (a single sensor or a set of sensors)
- Addition of sensors (i.e., a new resource) on a dynamic basis Aggregation of sensors and monitoring functions for areas (e.g., in order to understand a phenomena in a certain area)

Broker of Social Sensor Derived Information

Additional functions supported by an IoT Architecture

- Anonymization of users associated with objects.
- Data Mining and Information Inference.
- Return of meaningful data to the single user and comparison of the single user with specific users (anonymous users or the "average" user).
- Ability to allow access to historical usage data and creation of a sort of SPIME archive.

An indication of the functions/requirements that this use case is stressing

- Representation of proprietary objects into observable and normalized virtual objects.
- Anonymization of objects.
- Creation of a sort of social network for objects and ability to open up interfaces for data mining.
- Interoperability of objects pertaining to different administrative domains.
- Self-management with no user intervention.

Stockholms universitet

Section III:

Open Source tools and resources for IoT

Drivers for open source

- There are many drivers behind the popularity of open source:
  - Consumers want to use a large variety of consumer technology devices, and they don't want to be limited to using devices from one specific vendor (some smart watches can only be paired with the same vendor's smartphone).
  - Vendors of IoT devices want to increase the number of technology ecosystems into which their devices can be integrated without major effort.
  - Application developers want to support a broad range of devices without having to develop vendor-specific code.
- **Open source solutions** allow for this scale, velocity of innovation, and flexibility.

Scale

- In order to scale to connect and support 50 to 70 billion **sensors** in the next decade, potentially millions of routers, gateways, and data servers will be required. There's no way to achieve these levels of scale without relying on open source frameworks and platforms within that infrastructure. Any other approach will simply be unaffordable.

Velocity of innovation

- Open source technologies support rapid innovation through several advantageous characteristics, allowing for a more natural adoption approach within the enterprise. It's free and generally easy to download, install, and launch.
- This allows easy exploration of and experimentation with new technologies and enables enterprises to get comfortable with the software on smaller, non-mission-critical projects before any financial commitment is required.
- open source software enables permissionless innovation, easing concerns over royalties or lawsuits. In particular, open source promotes innovation by integration, where developers create new systems by combining freely available open source components.

Vibrant developer community

- Open source software projects tend to promote innovation faster than proprietary solutions because they draw contributions from a large community of developers. The accumulation of this community participation accelerates the delivery of the key features and ecosystem that enterprises need. The result is that open source usually delivers on-target capabilities faster than proprietary alternatives.

Interoperability

- The best possible way to have a new technology achieve rapid adoption is by combining open standards with a robust open source implementation. Open source implementations provide an easy adoption path with interoperability, and they reduce the cost of entering the market. Developers can spend less time implementing a standard and focus instead on building software that provides the firm with the product-differentiating features customers value.

Stockholms universitet

No lock-in

- A key advantage of open source solutions is that they don't lock your organization in to a proprietary provider. The cost of switching solutions and their vendors is typically high, which holds enterprises and startups hostage to proprietary service providers. Open source mitigates the risk of proprietary solutions being discontinued or no longer supported. How many firms have purchased proprietary software and invested in costly customizations, only to be stuck with an unsupported system or a costly upgrade path?

Open source projects in study

Industry consortiums

- AllSeen Alliance
- Open Interconnect Consortium (OIC)
- COMPOSE
- Eclipse
- Open Source Hardware Association (OSHWA)

Protocol

- Advanced Message Queuing Protocol (AMQP)
- Constrained Application Protocol (CoAP)
- Extensible Messaging and Presence Protocol (XMPP)
- OASIS Message Queuing Telemetry Transport (MQTT)
- Very Simple Control Protocol (VSCP)

Stockholms universitet

Cont. Open source projects ……

Operating systems (OS)

- ARM mbed
- Canonical Ubuntu and Snappy Ubuntu Core
- Contiki
- Raspbian
- RIOT
- Spark
- webinos

APIs

- BipIO
- Qeo Tinq
- Zetta
- 1248.io

Middleware

IoTSyS

OpenIoT

OpenRemote

Kaa

Cont. Open source …. - Horizontal platforms

- Canopy
- Chimera IoT
- DeviceHive
- Distributed Services Architecture (DSA)
- Pico Labs (Kynetx open source assigned to Pico Labs)
- M2MLabs Mainspring
- Nimbits
- Open Source Internet of Things (OSIOT)
- prpl Foundation
- RabbitMQ
- SiteWhere
- Webinos
- Yaler

Cont. Open source....

Water
Oxford Flood Network

Node flow editors
- Node-RED
- ThingBox

Toolkits
- KinomaJS
- IoT Toolkit

Search
- Thingful

Data visualization
- Freeboard
- ThingSpeak

Hardware
- Arduino Ethernet Shield
- BeagleBone
- Intel Galileo
- OpenPicus FlyportPro
- Pinoccio
- WeIO
- WIZnet

Home automation
- Home Gateway Initiative (HGI)
- Ninja Blocks
- openHAB
- Eclipse SmartHome
- PrivateEyePi
- RaZberry
- The Thing System

Robotics
Open Source Robotics Foundation

Mesh networks
- Open Garden
- OpenWSN

In-memory data grids
- Ehcache
- Hazelcast

Health
e-Health Sensor Platform

Air pollution
HabitatMap Airbeam

What you learned

- Characteristic of smart & community sensing
- Opportunities and challenges of IoT in smart & community sensing
- Opportunities in Mobile Crowd sensing
- IoT Data Taxonomy
- Sensing-as-Service
- A Social Sensors Service representation
- IoT Broker

Thank you !

Rahim Rahmani
Dept. of Computer & System Sciences
Stockholm University, Sweden
rahim@dsv.su.se