New Value Chains and Technical Issues for Future Data Eco-Society

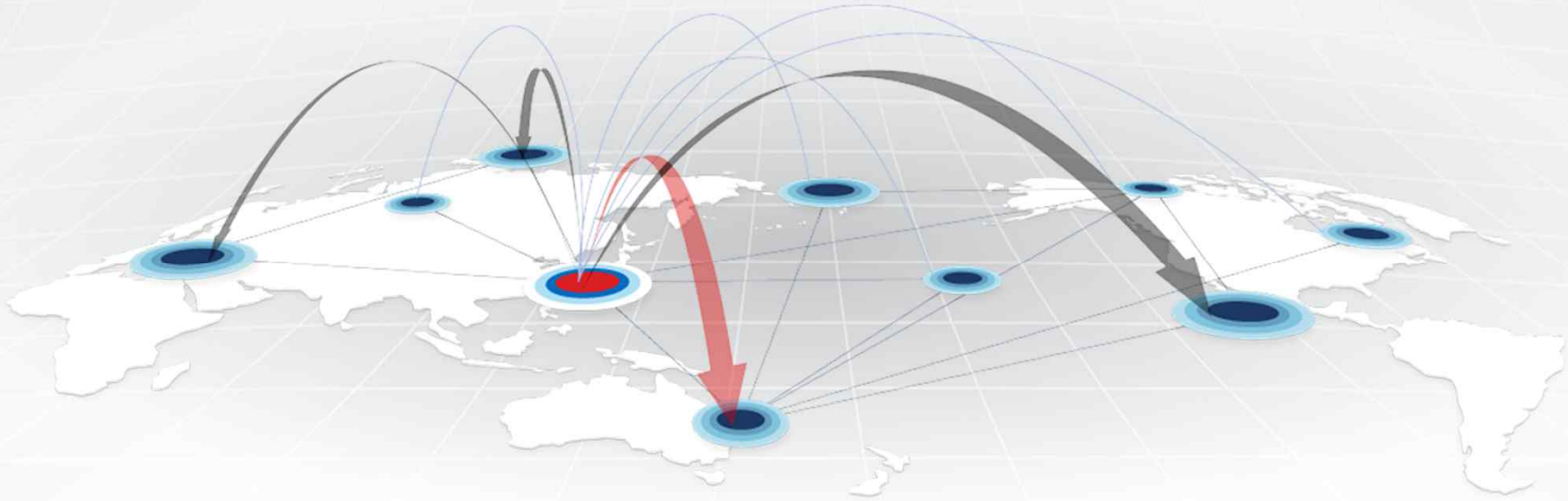
Jun Kyun Choi

Professor, Korea Advanced Institute of Science and Technology (KAIST)
jkchoi59@kaist.ac.kr

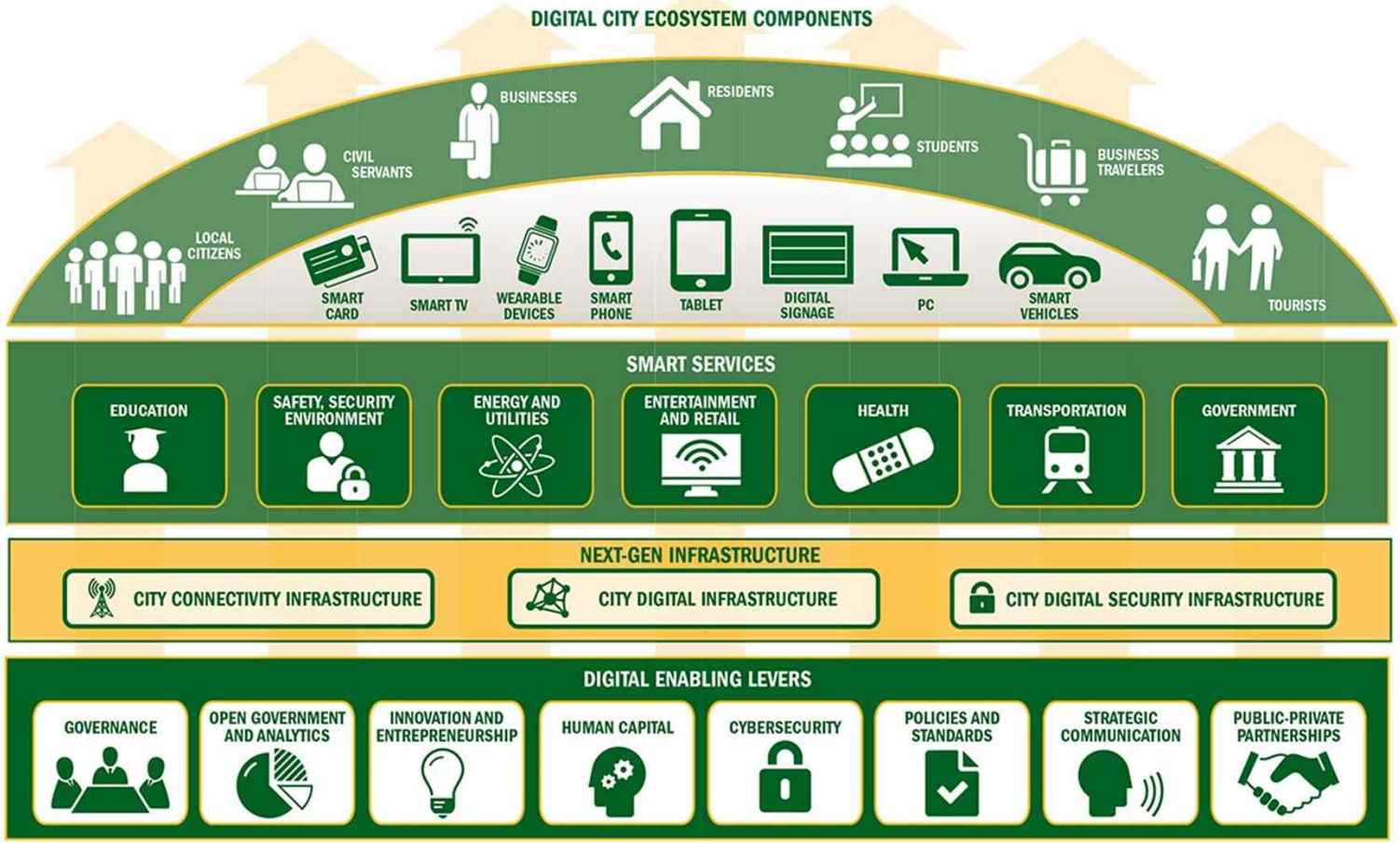
KAIST 한국과학기술원
Korea Advanced Institute of Science and Technology

Contents

Value Chains of Data Eco-Society



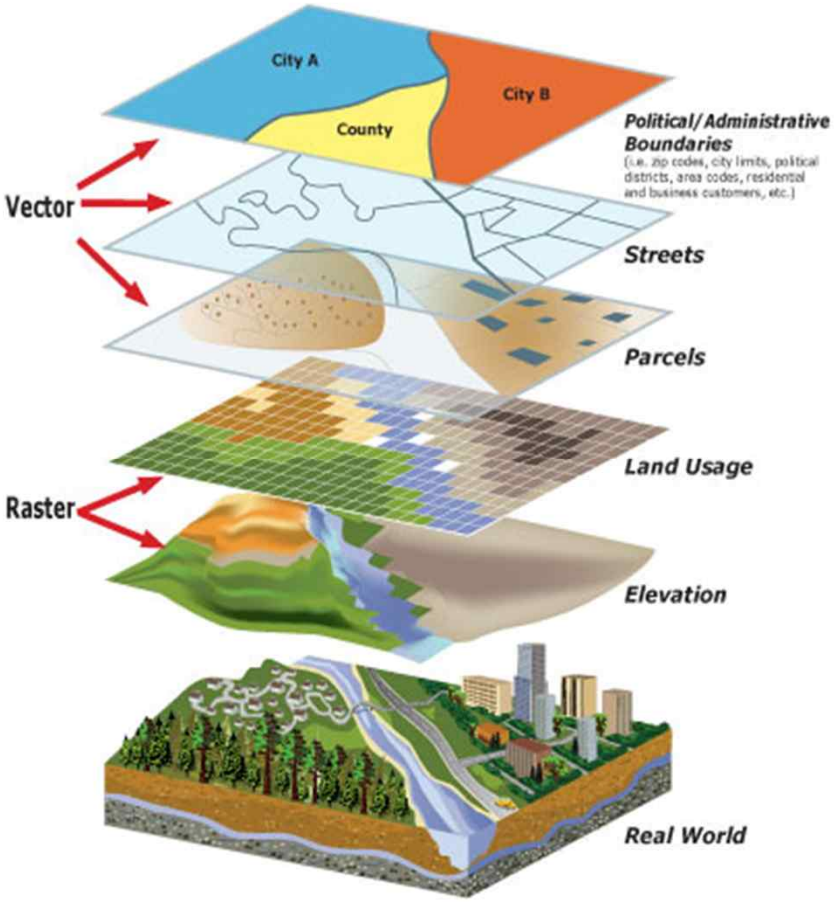
Smart City Eco-System

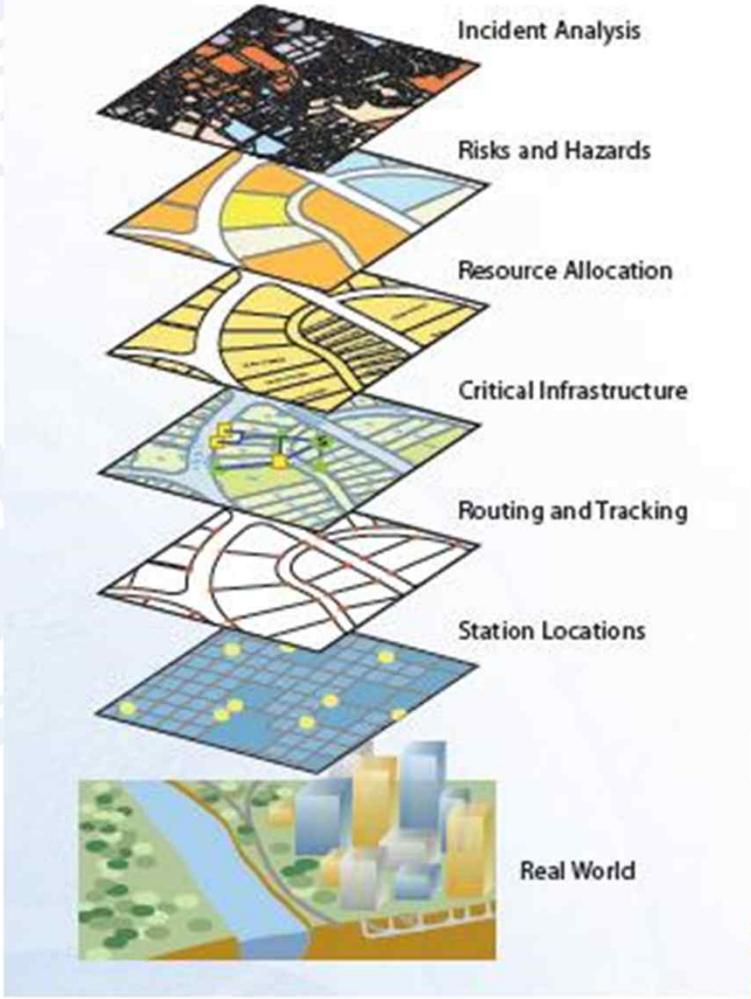


http://www.urenio.org/2015/12/20/booz-allen-hamilton-on-smart-cities/

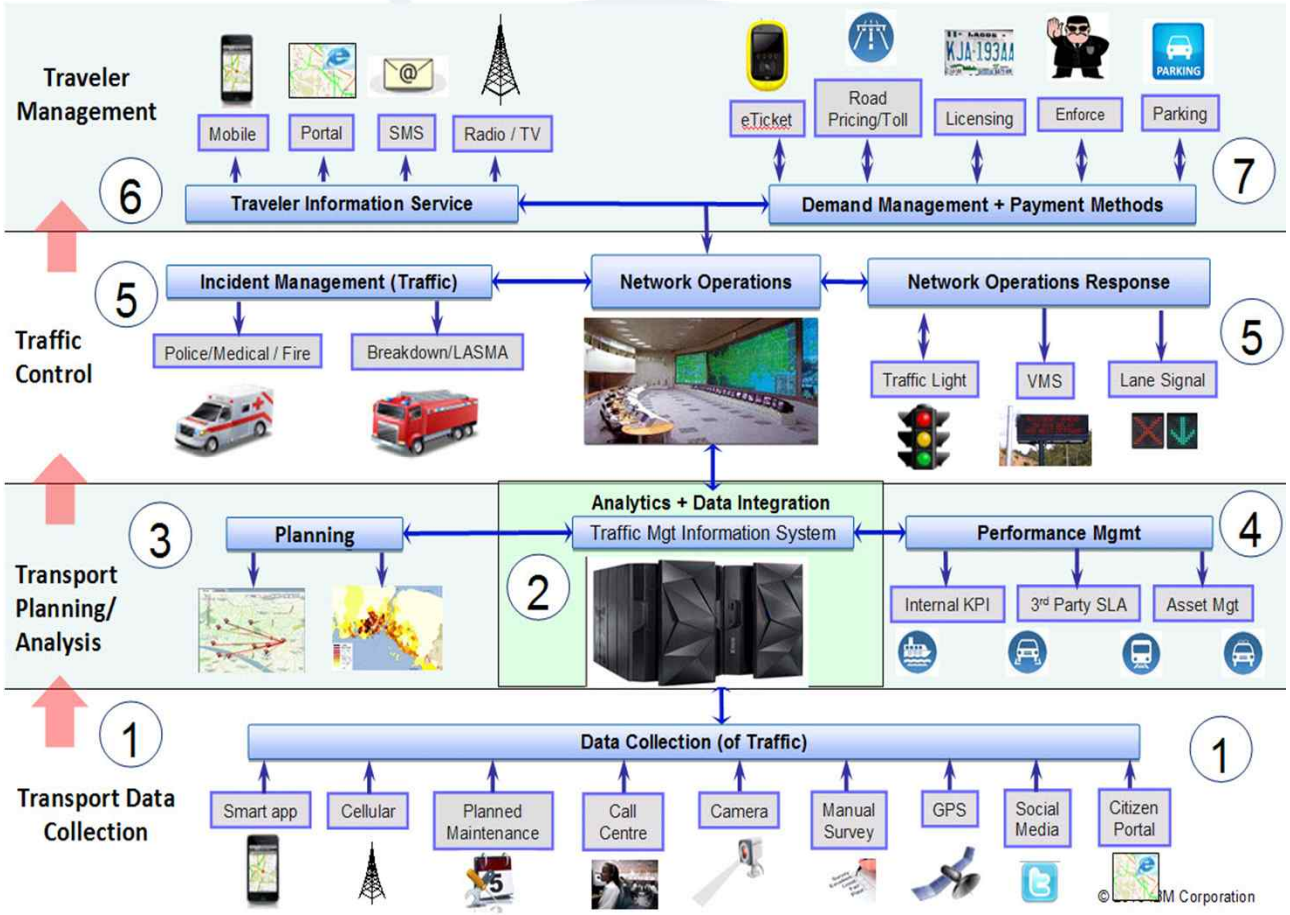
KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

Geographical Data



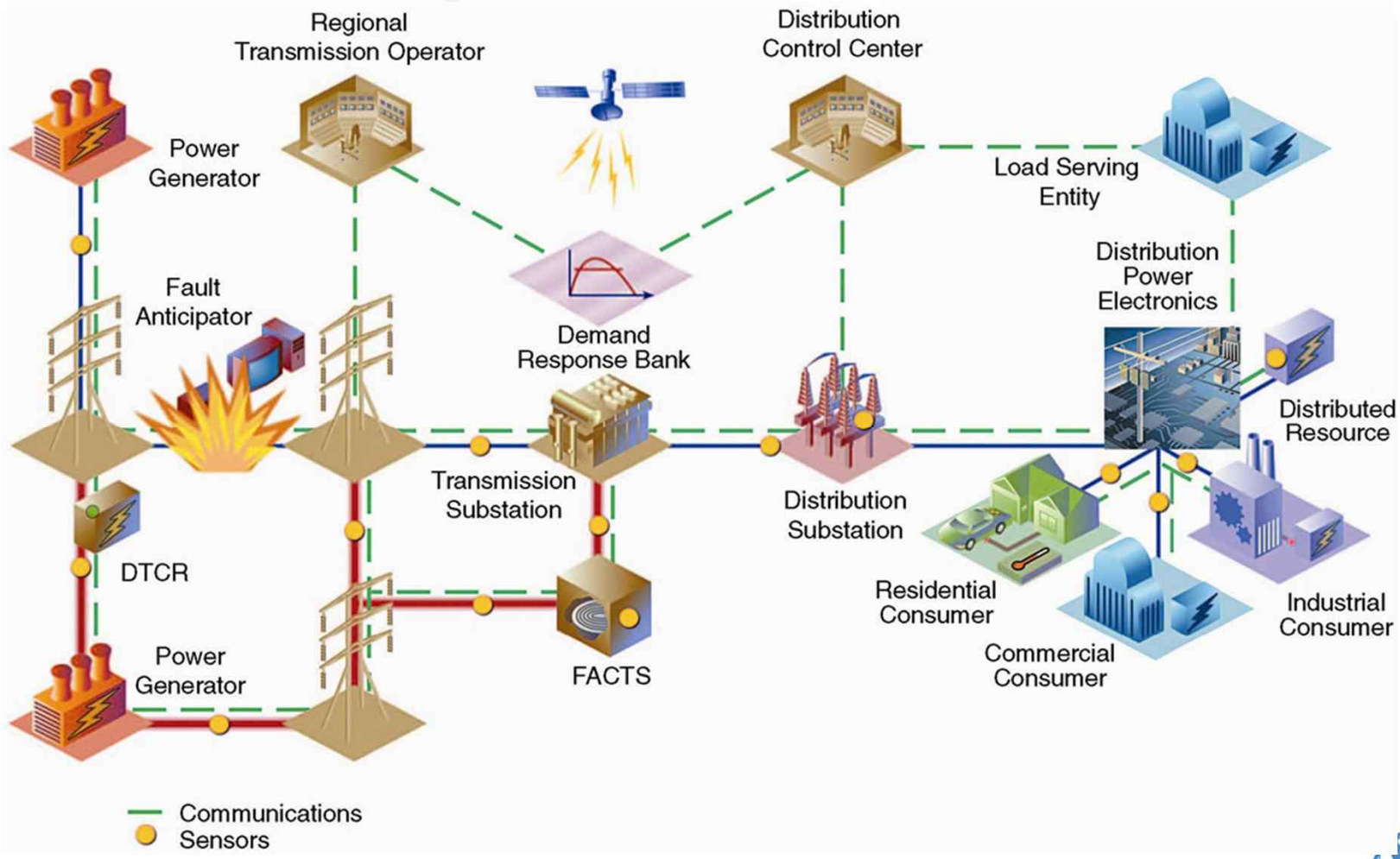


Data of Intelligent Transport System



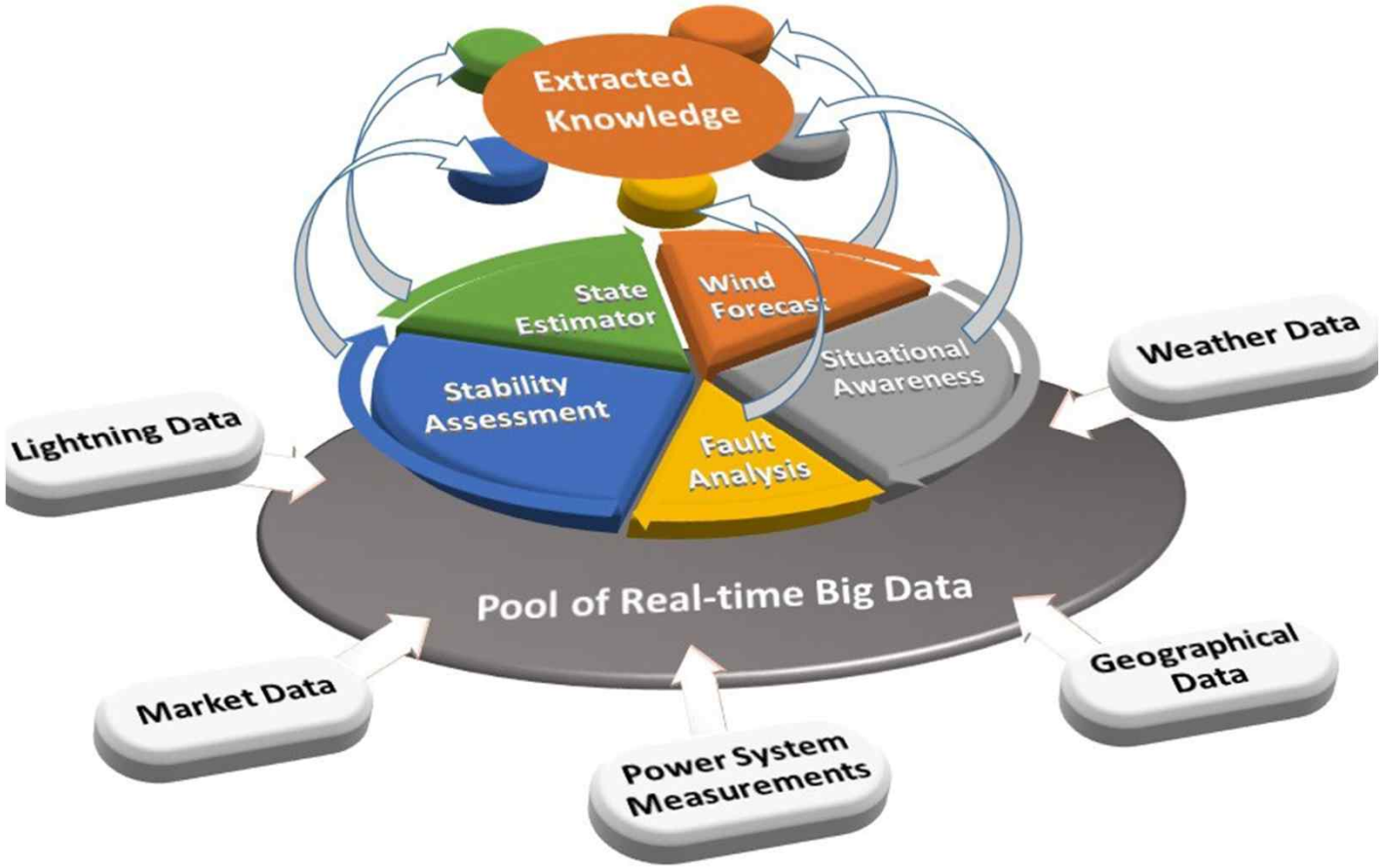
(ref) https://smartercitieschallenge.wordpress.com/category/lagos-nigeria/

Overview of Smart Grid

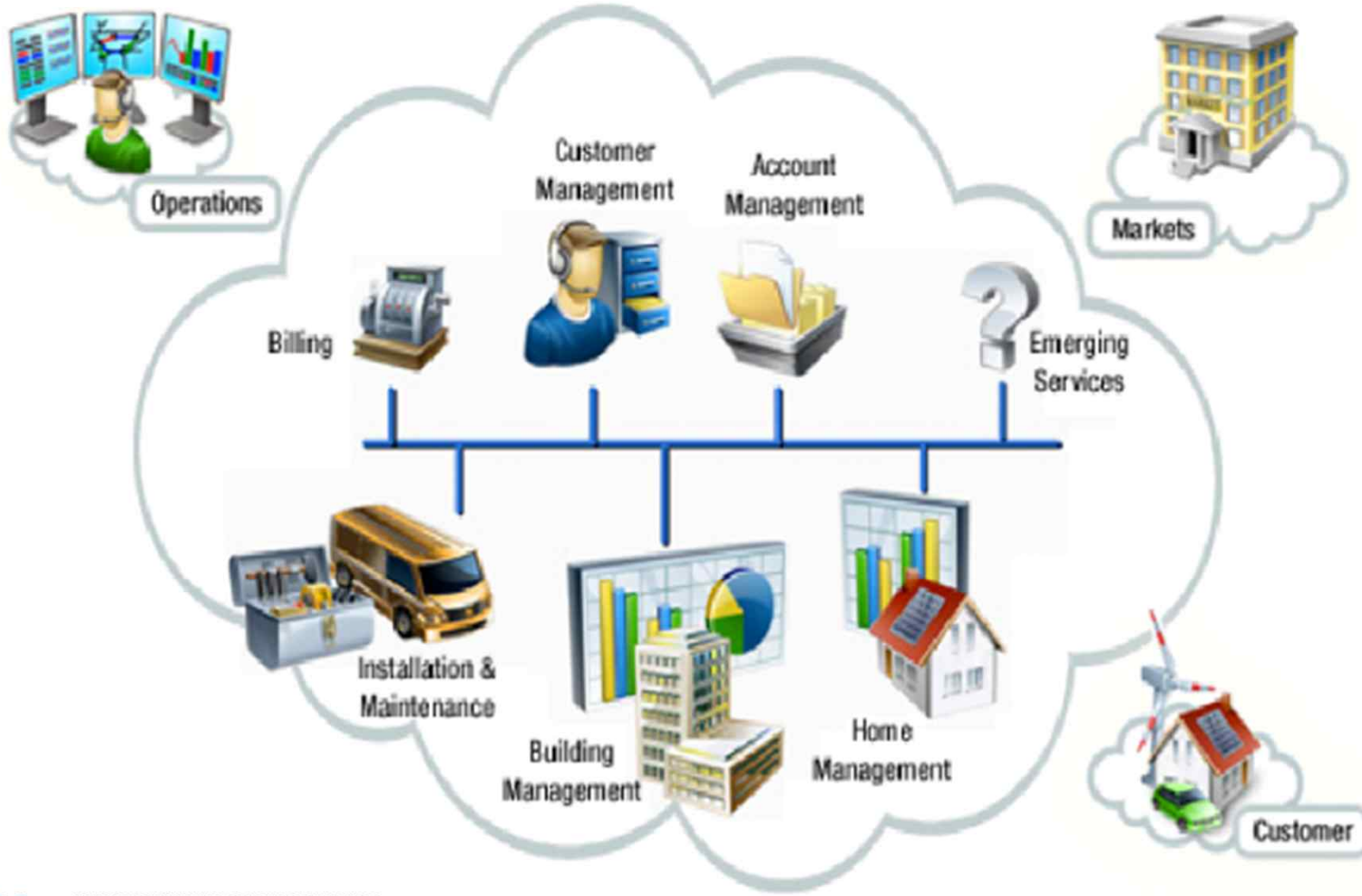


KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

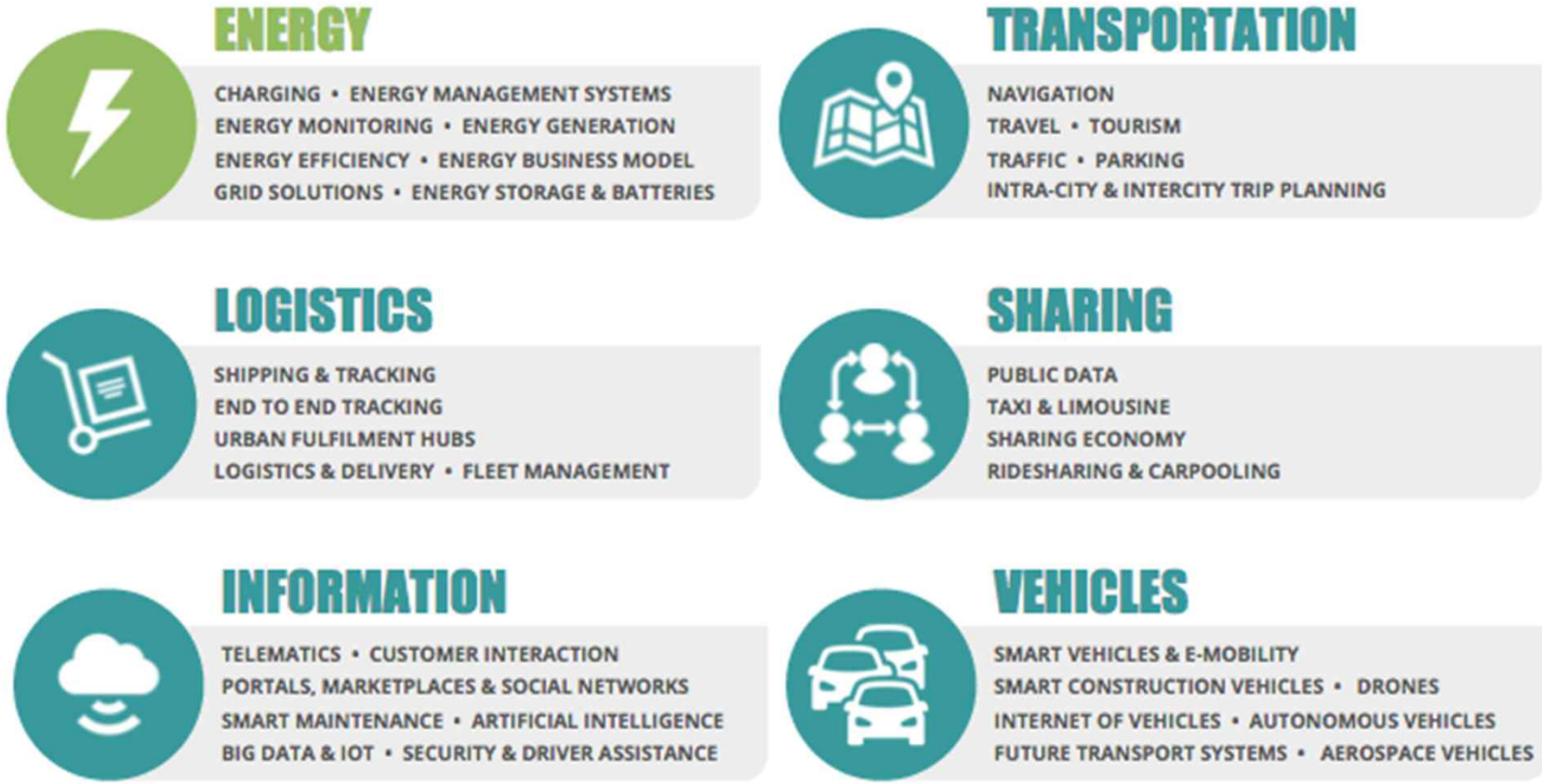
Data of Smart Grid



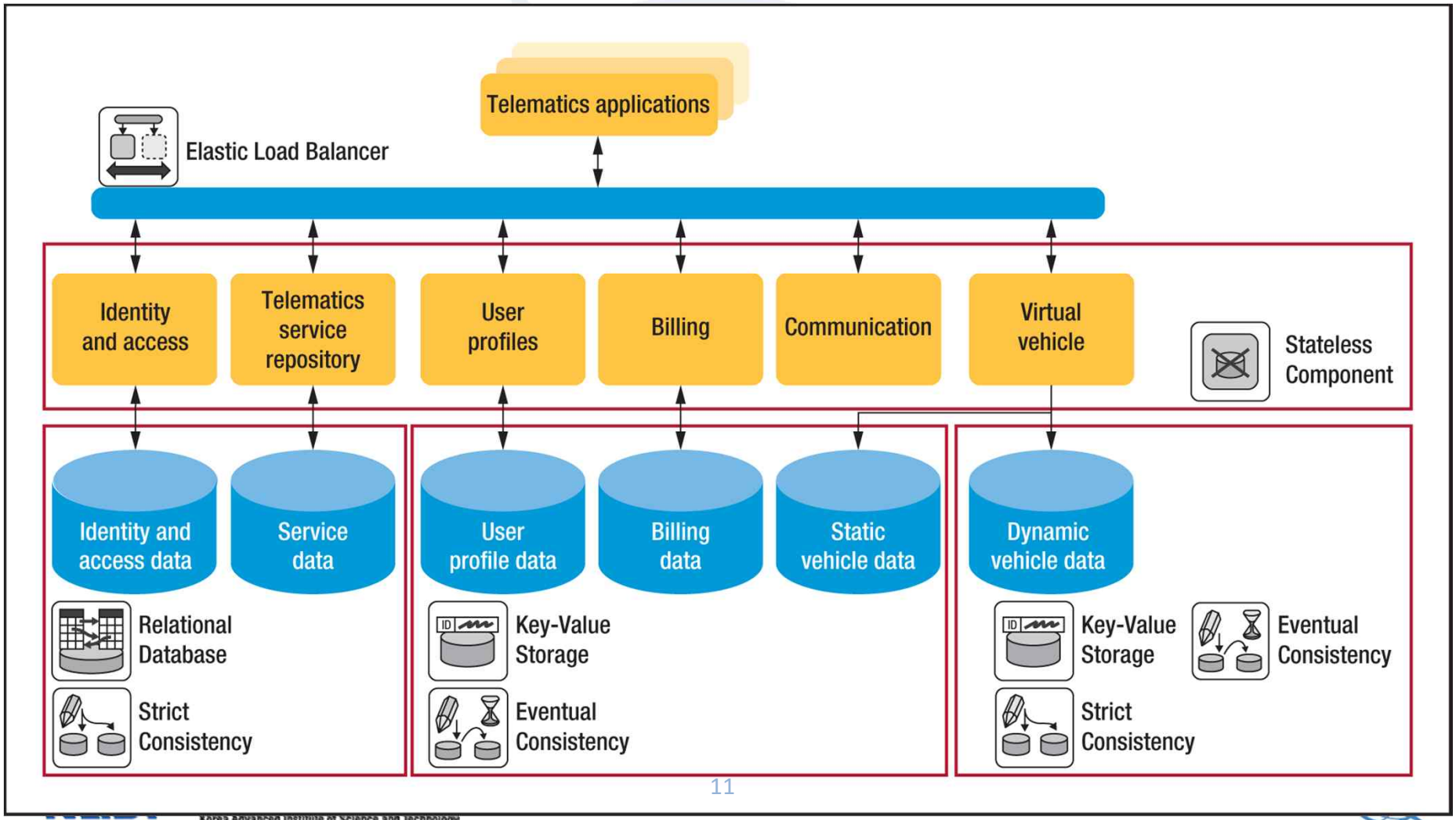
Data Business for Smart Grid



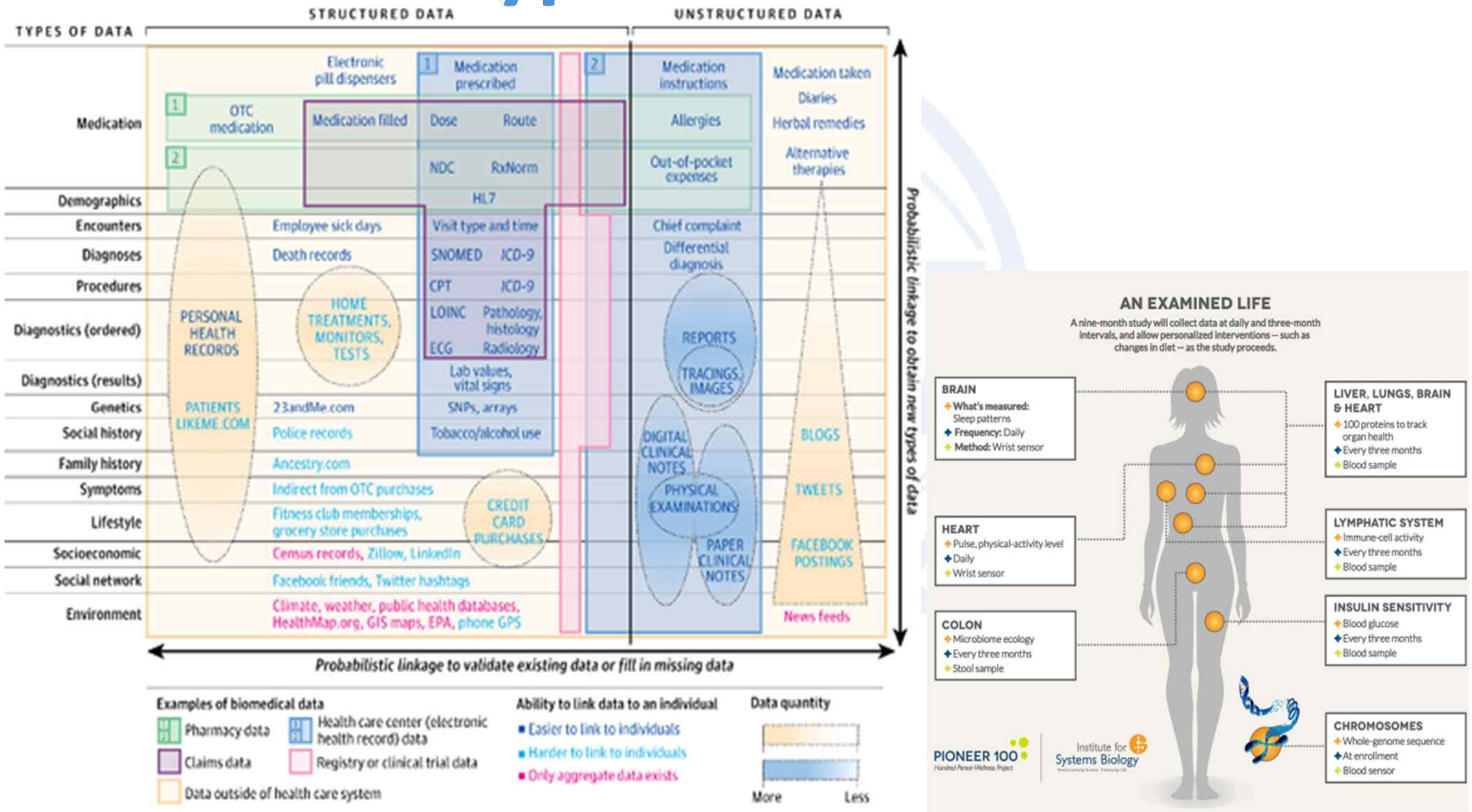
Data Types for Smart Grid



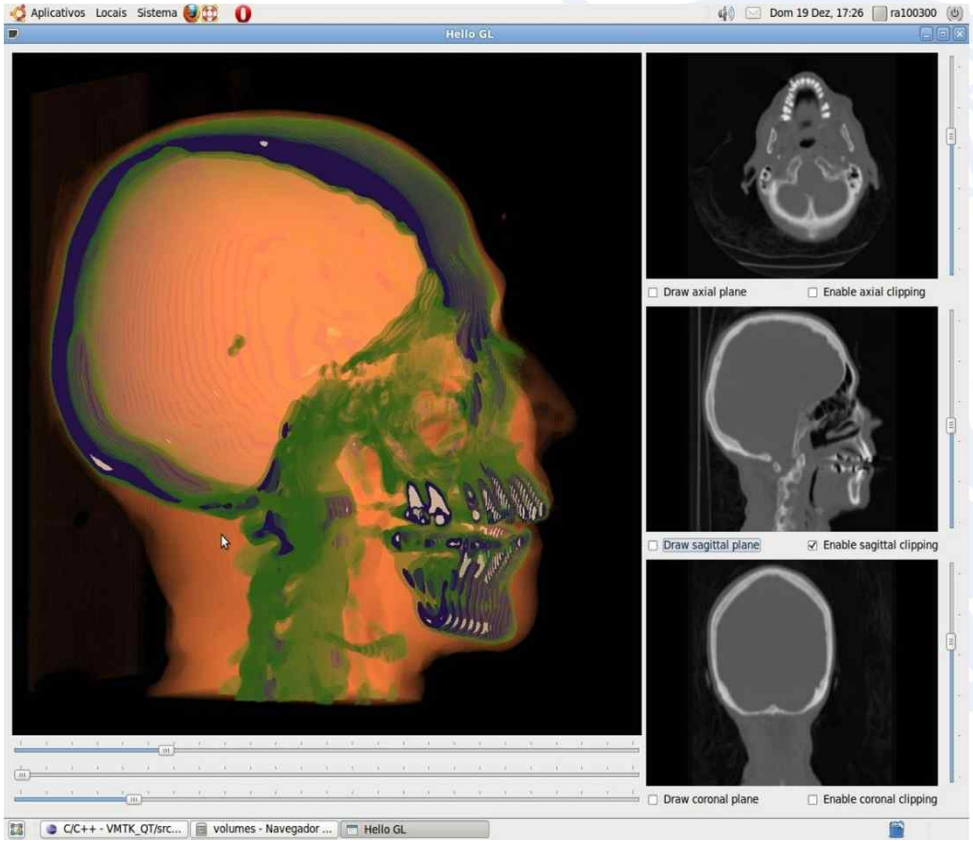
Telematics Data

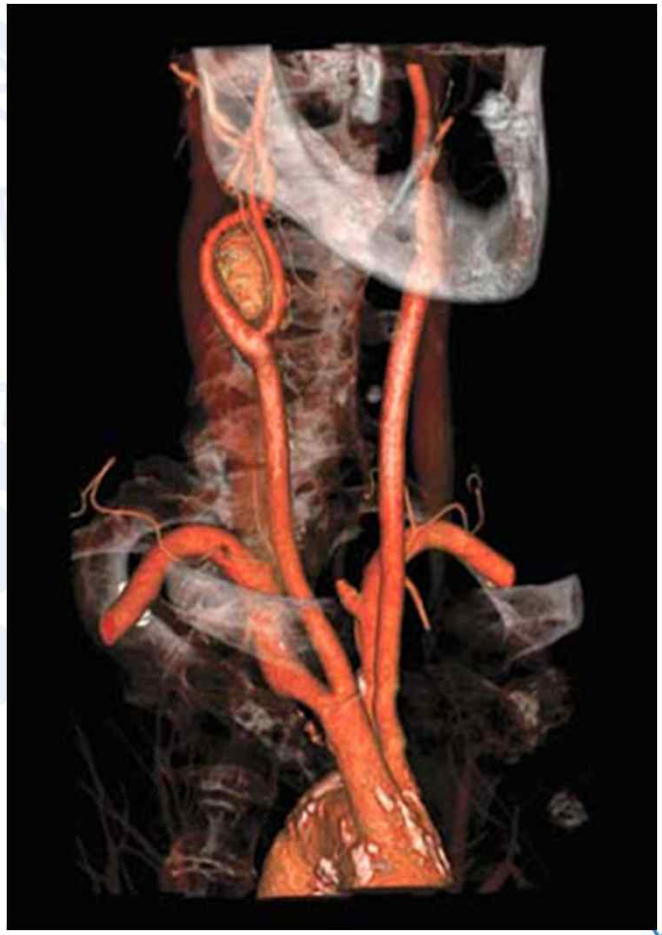


Data Types of Healthcare



Medical Data Visualization



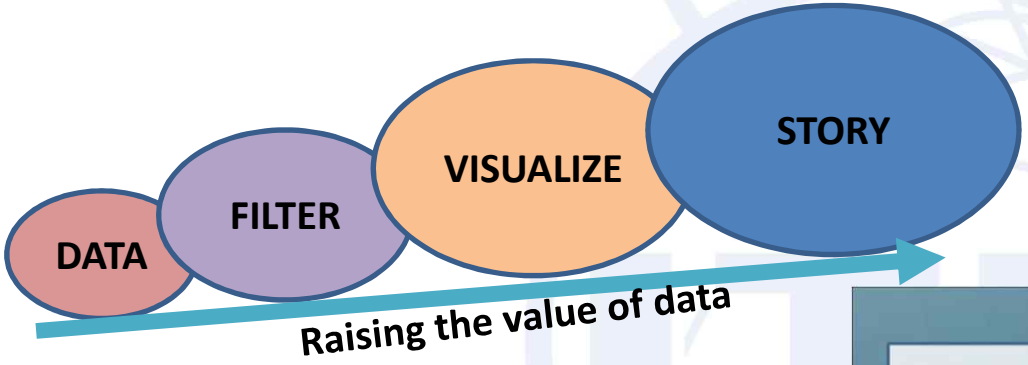


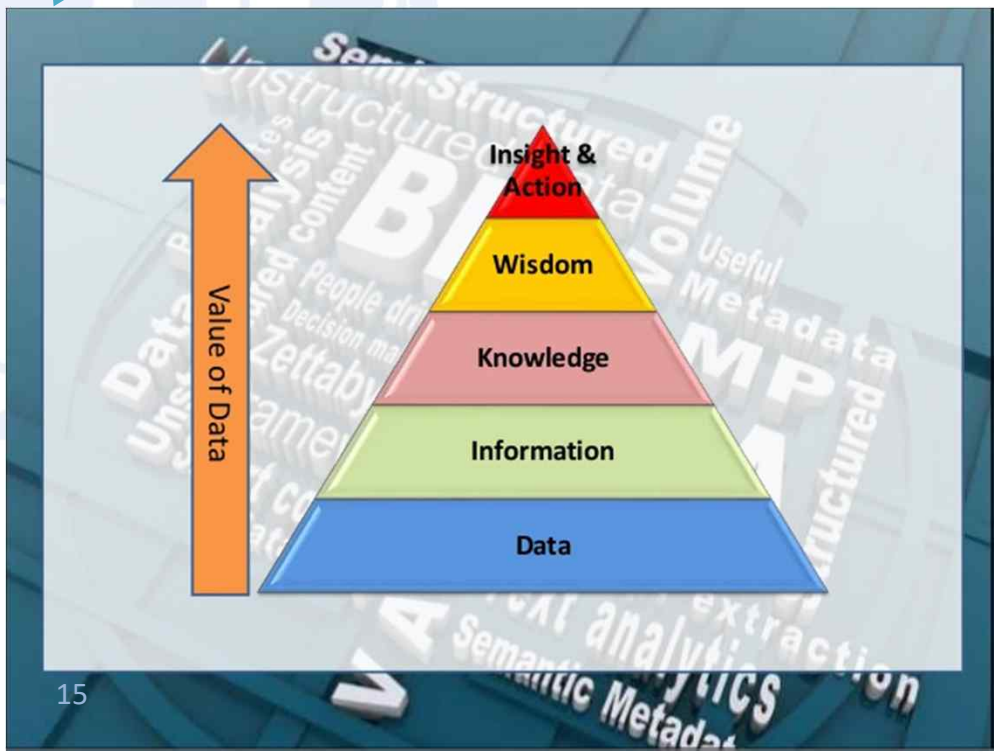
What is the Language of Cyber World ?

Cyber Language
Language for Semantic Knowledge
Language for Platform
Language for System and Application
Operating System (OS)
Machine Language (e,g., device driver)

Physical Language	Natural Language
Human Language	Human Language
Language for Building, Road, Station, Airport, Hospital	Language for Monkey
Language for TV, Car, and Airplane	Language for Dolphin, Dog, Pig, Elephant
Language for Hardware and software	Language for Mouse, Chicken, Bird
Language for Components and Sensors	Language for Bee, Ant, Insect

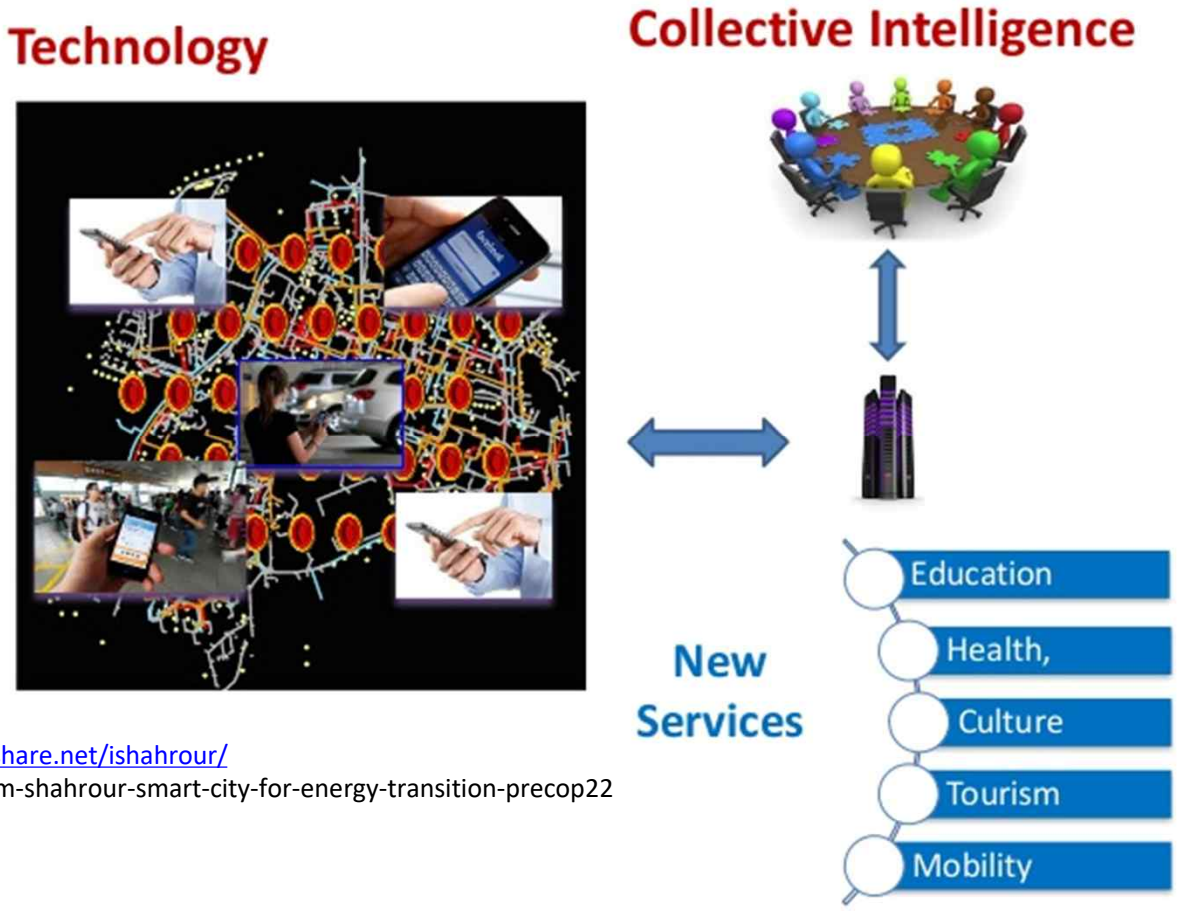
Value of Data - 1





KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

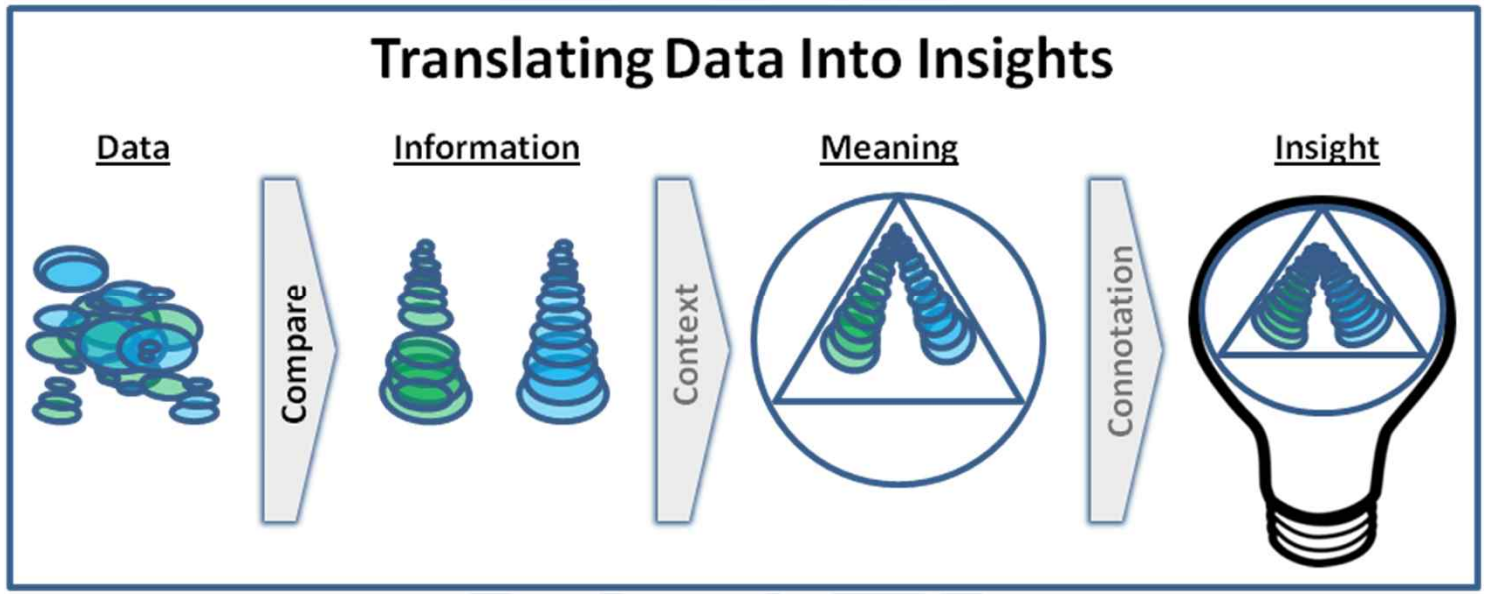
Value of Data - 2



http://www.slideshare.net/ishahrour/
conference-of-isam-shahrour-smart-city-for-energy-transition-precop22

KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

Value of Data - 3



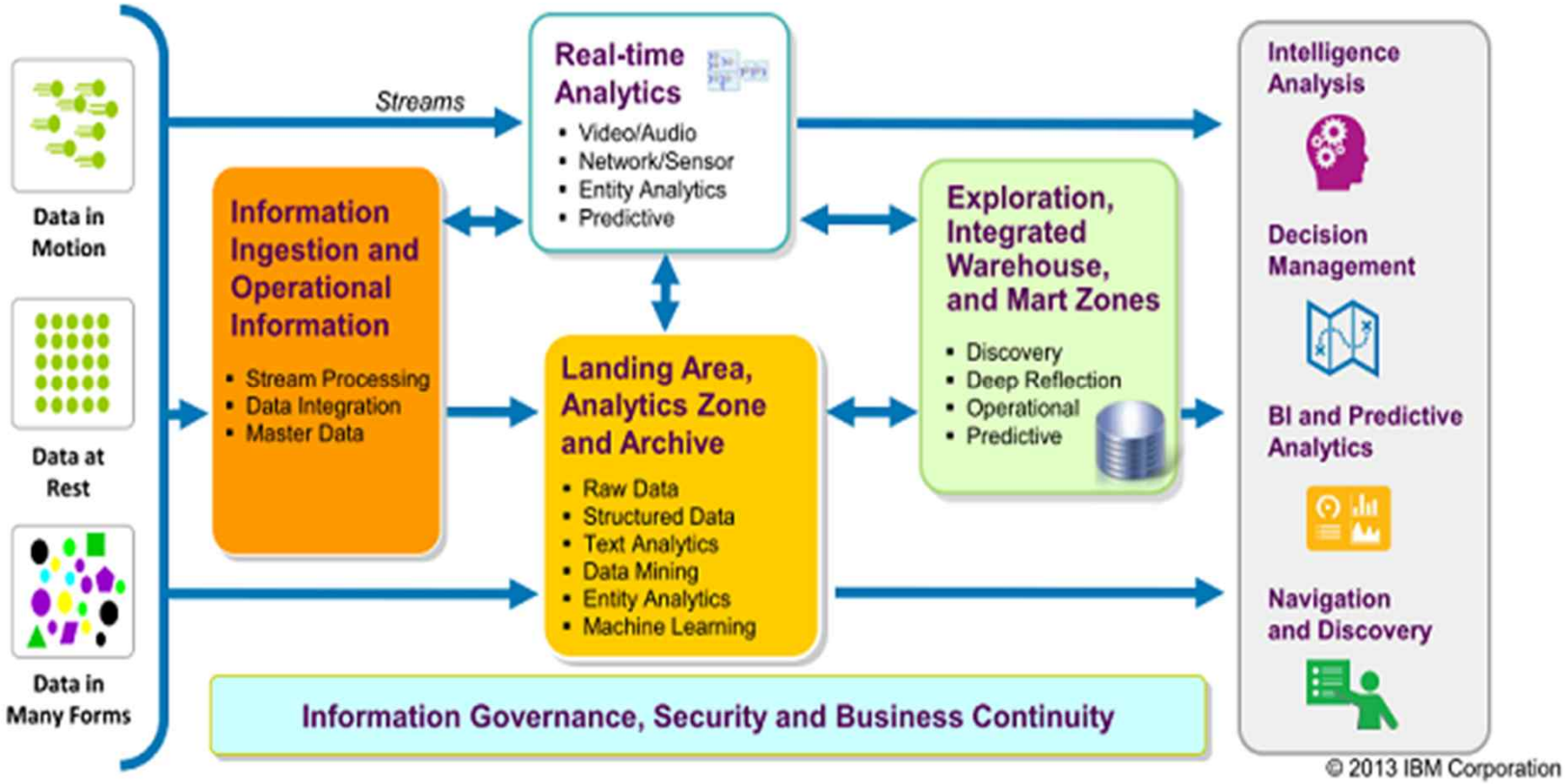
Decision = Data + Rules

"Big Data"

Data Science

KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

Big Data and Analytics Process

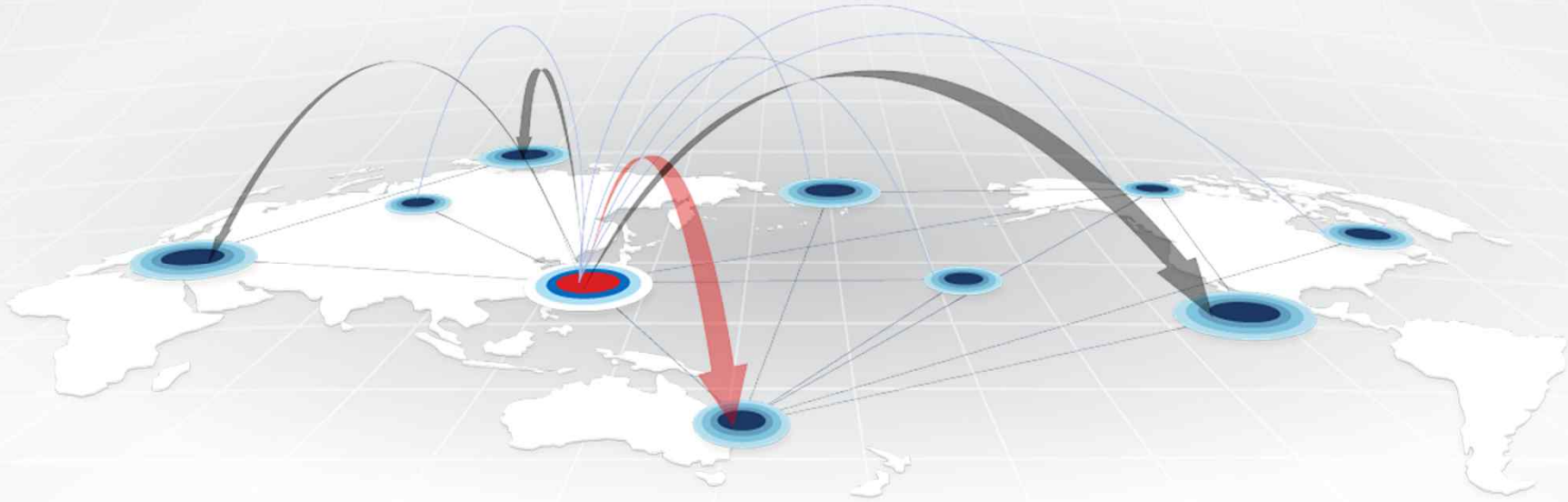


http://jtonedm.com/2013/06/05/big-data-and-analytics-fueling-competitive-advantage/

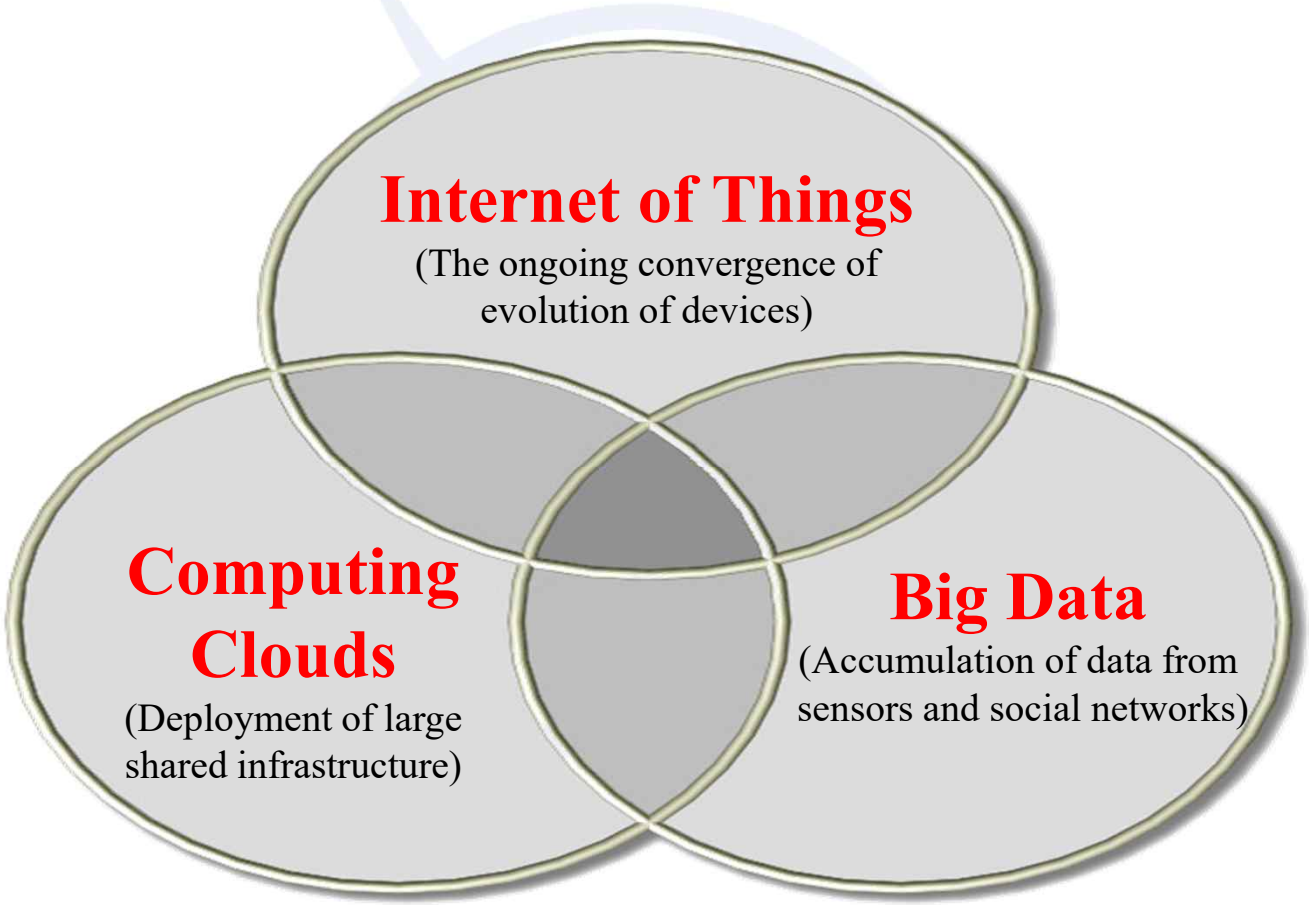
Who is the Winner at Data Eco-society ?

- **Maximize data eco-environments or eco-domains**
  - Maximize coverage and applications
  - Minimize conversion or interpretation overheads
- **Win-Win approaches among stakeholders**
  - Get synergy effects of all the stakeholders
- **Get the values from collective intelligences**
  - Find new values by knowledge accumulation of heterogeneous domains

Technical Issues for Data Technologies



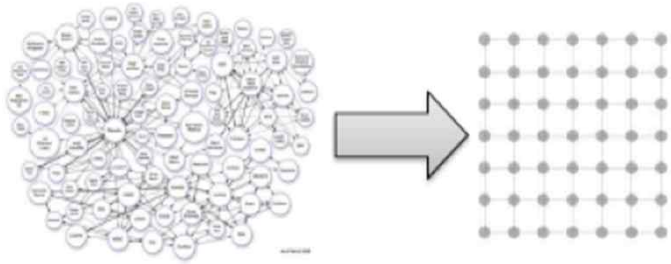
Key Research Challenges for ICT



Data as a Service

Web Scale Data Acquisition

http://www.thetechbulletin.com/promptcloud-big-data-crawlers-18384/



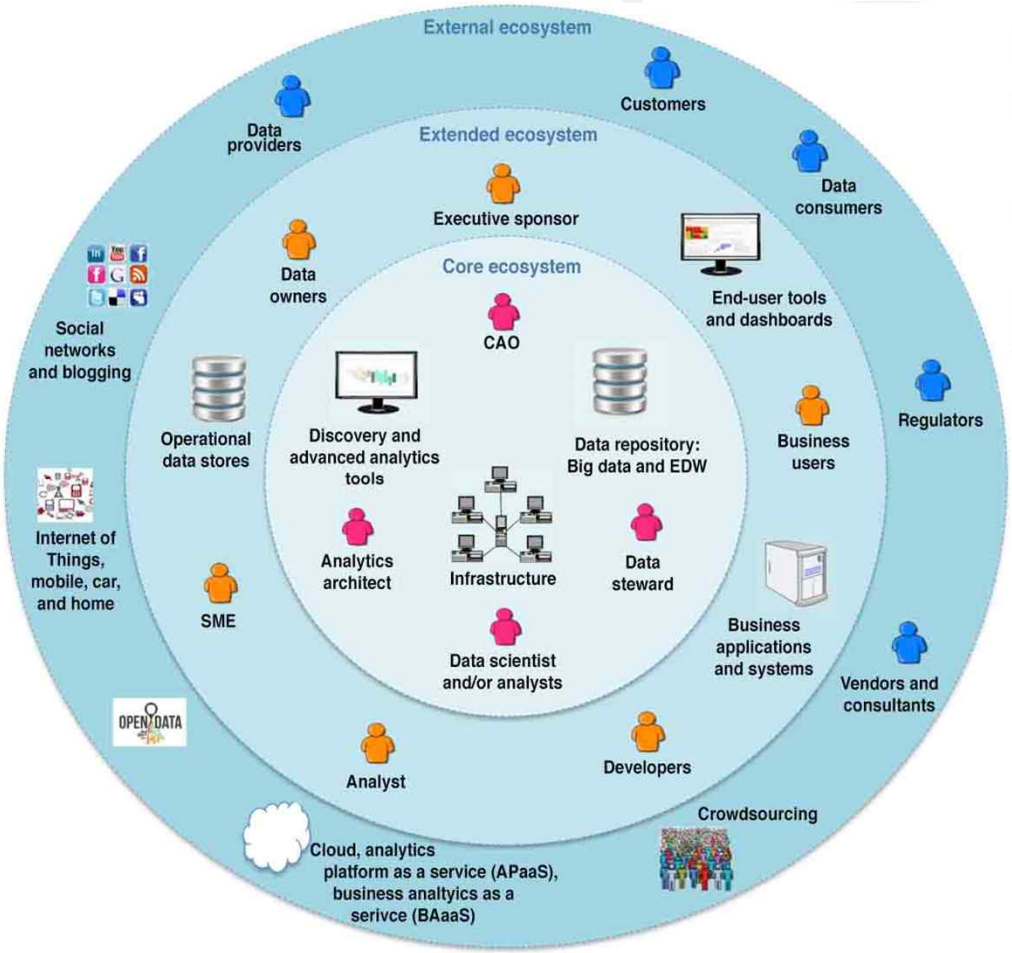
Converting Unstructured Data to Structured Data

Big Data Analytics

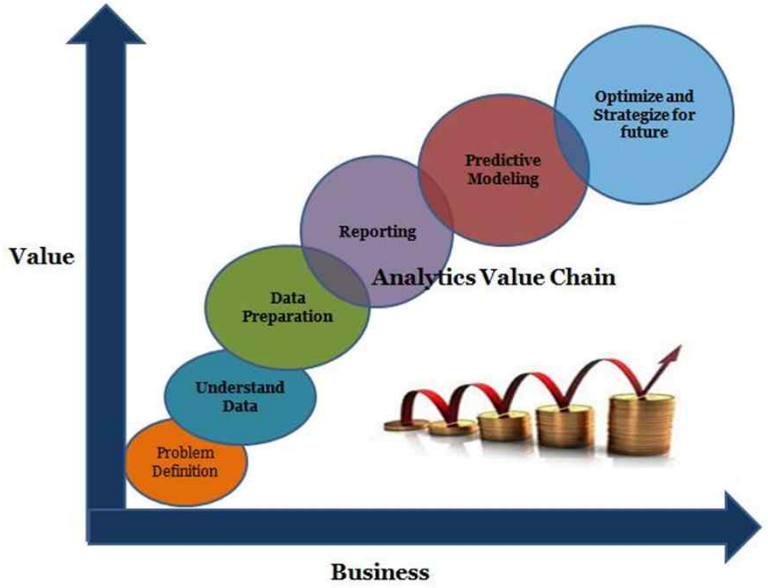
Hosted Indexing

Low Latency crawls

Data Analytics Eco-system



http://planetsystems.in/blog/preductive-analysis/

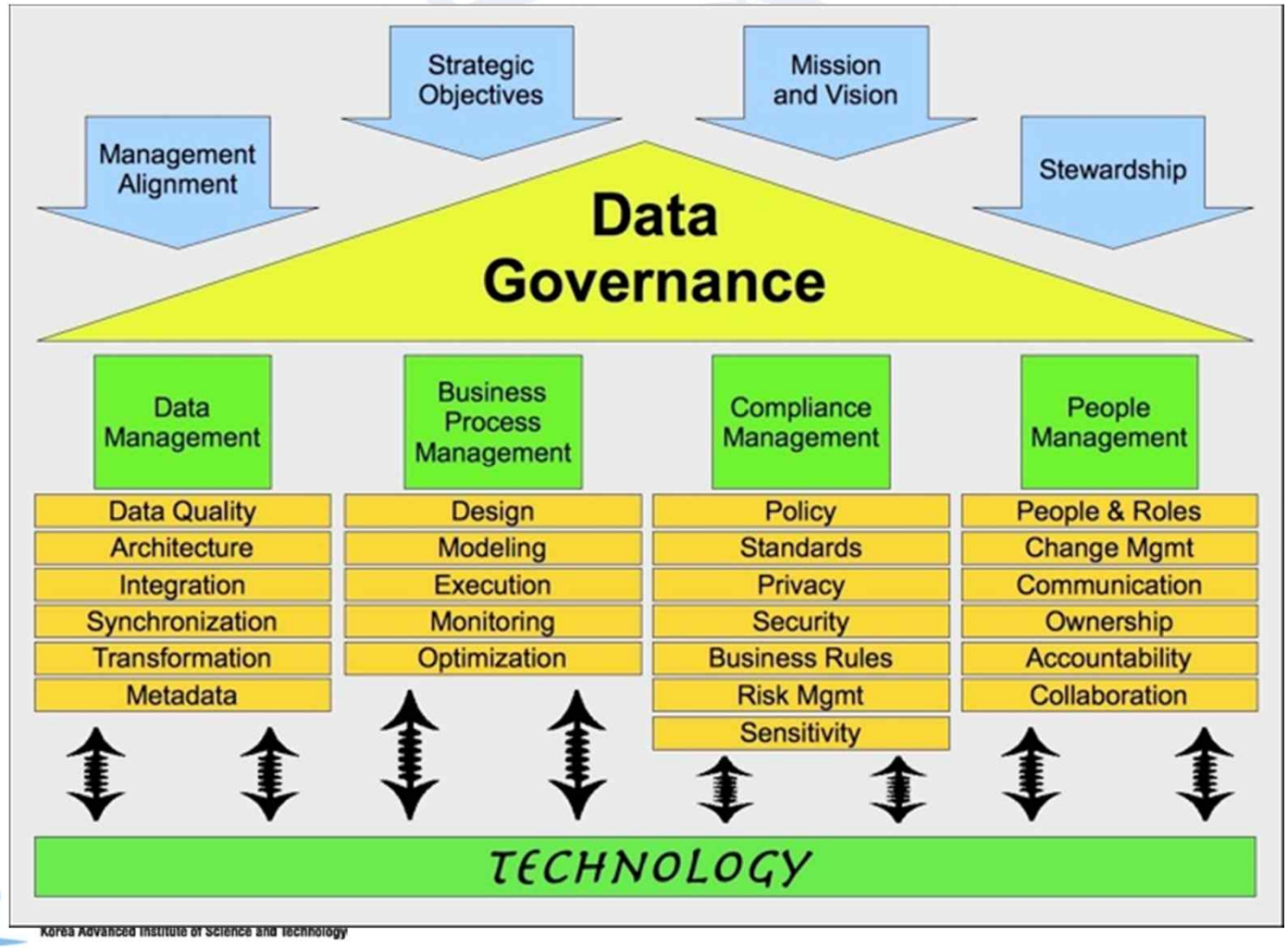


http://www.ibmbigdatahub.com/blog/going-beyond-data-science-toward-analytics-ecosystem-part-2

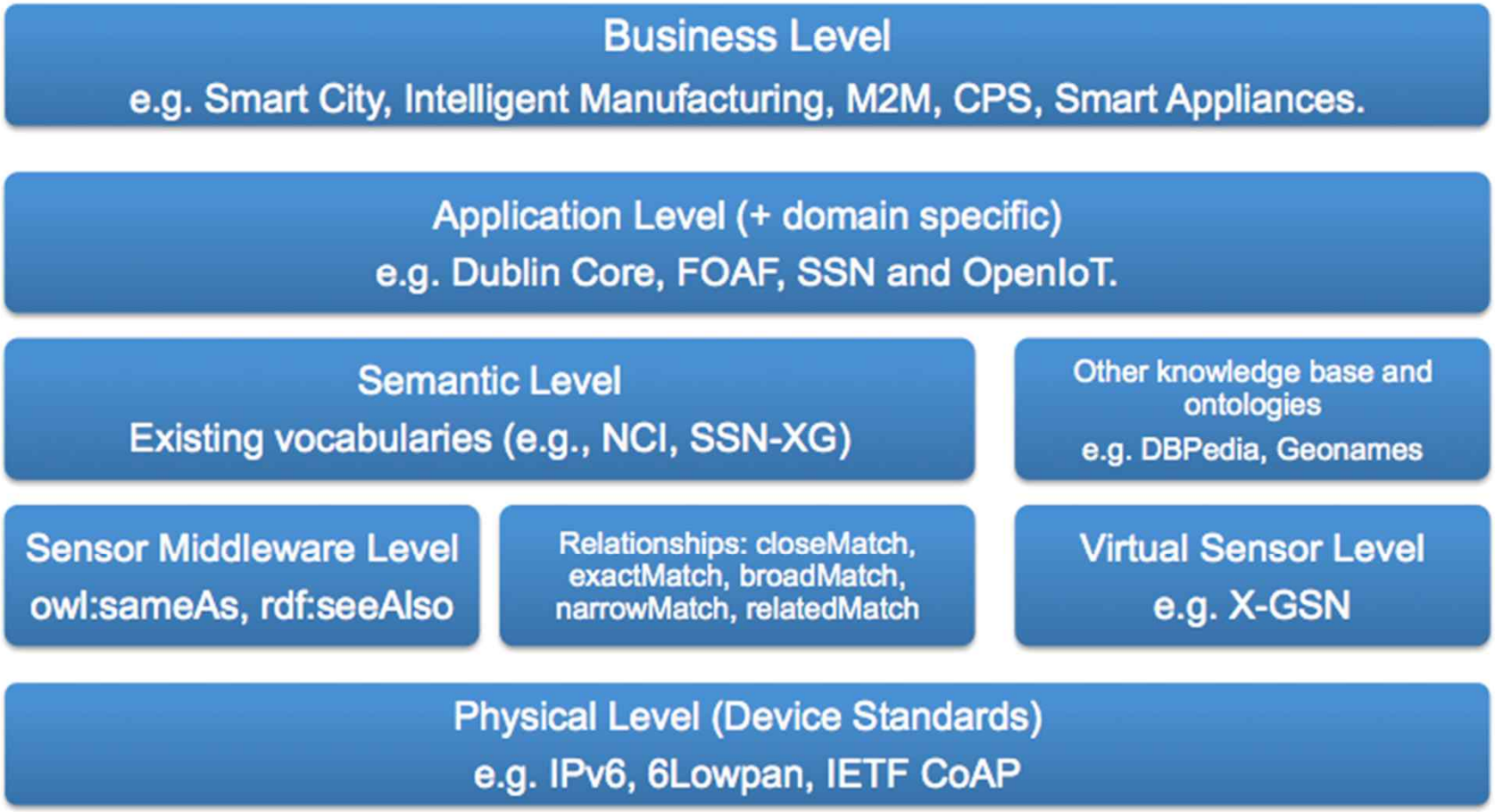
KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

Data Governance

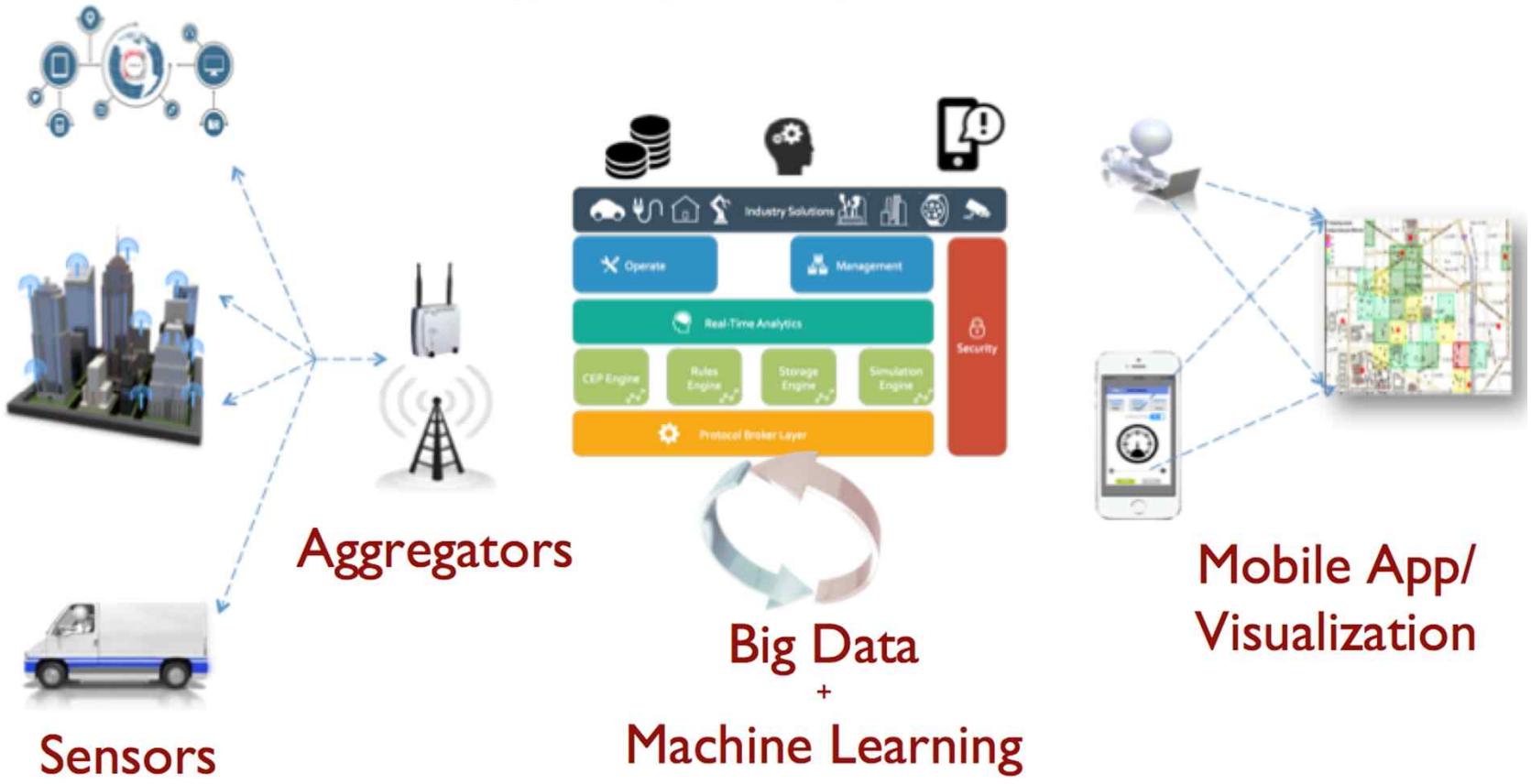
https://www.linkedin.com/pulse/nailing-down-data-governance-strategy-development-vimal-mani



Data Intelligence

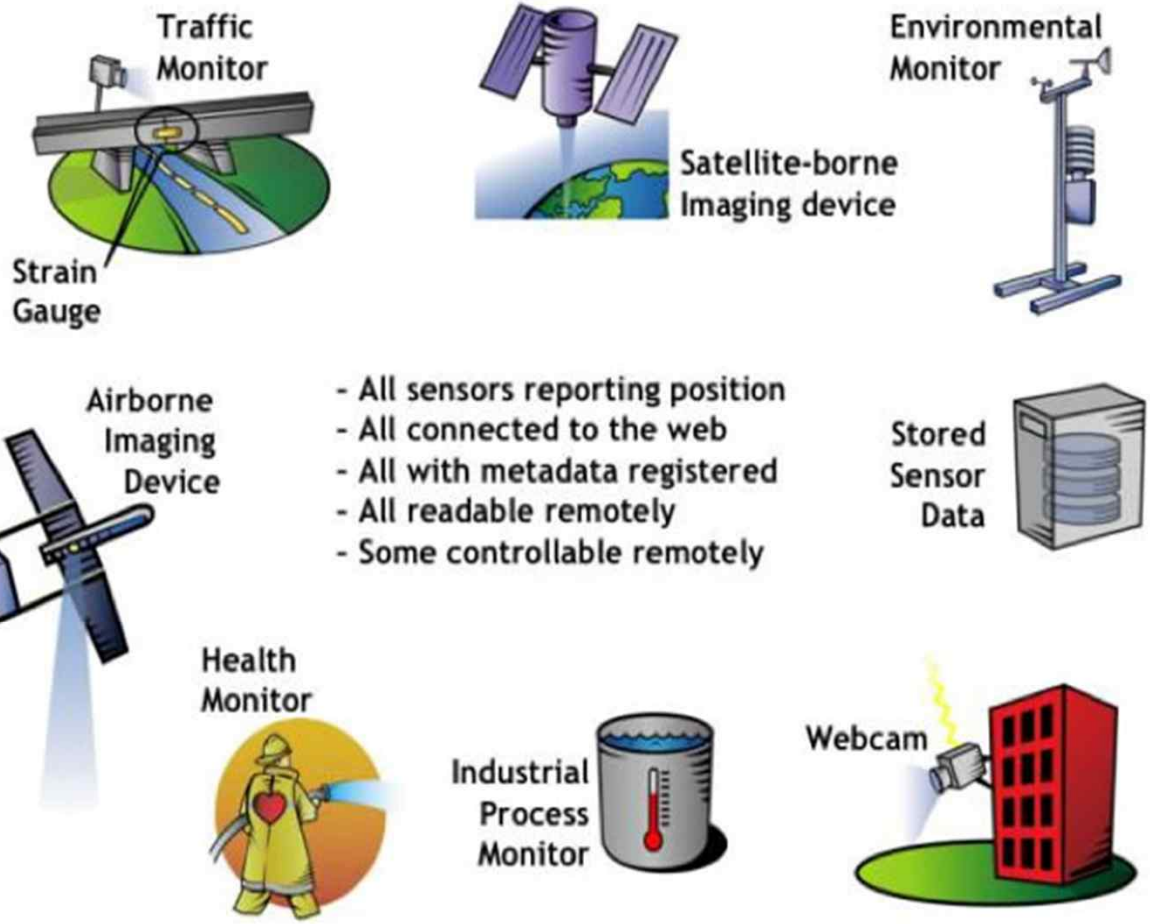


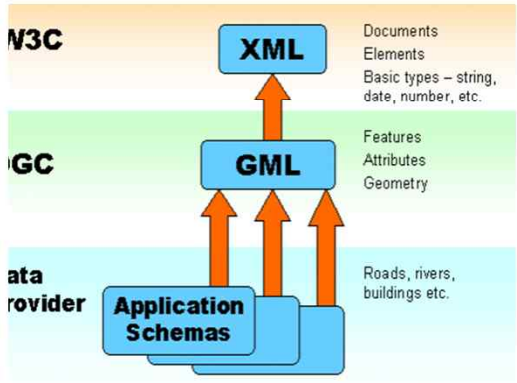
IoT Data Analytics



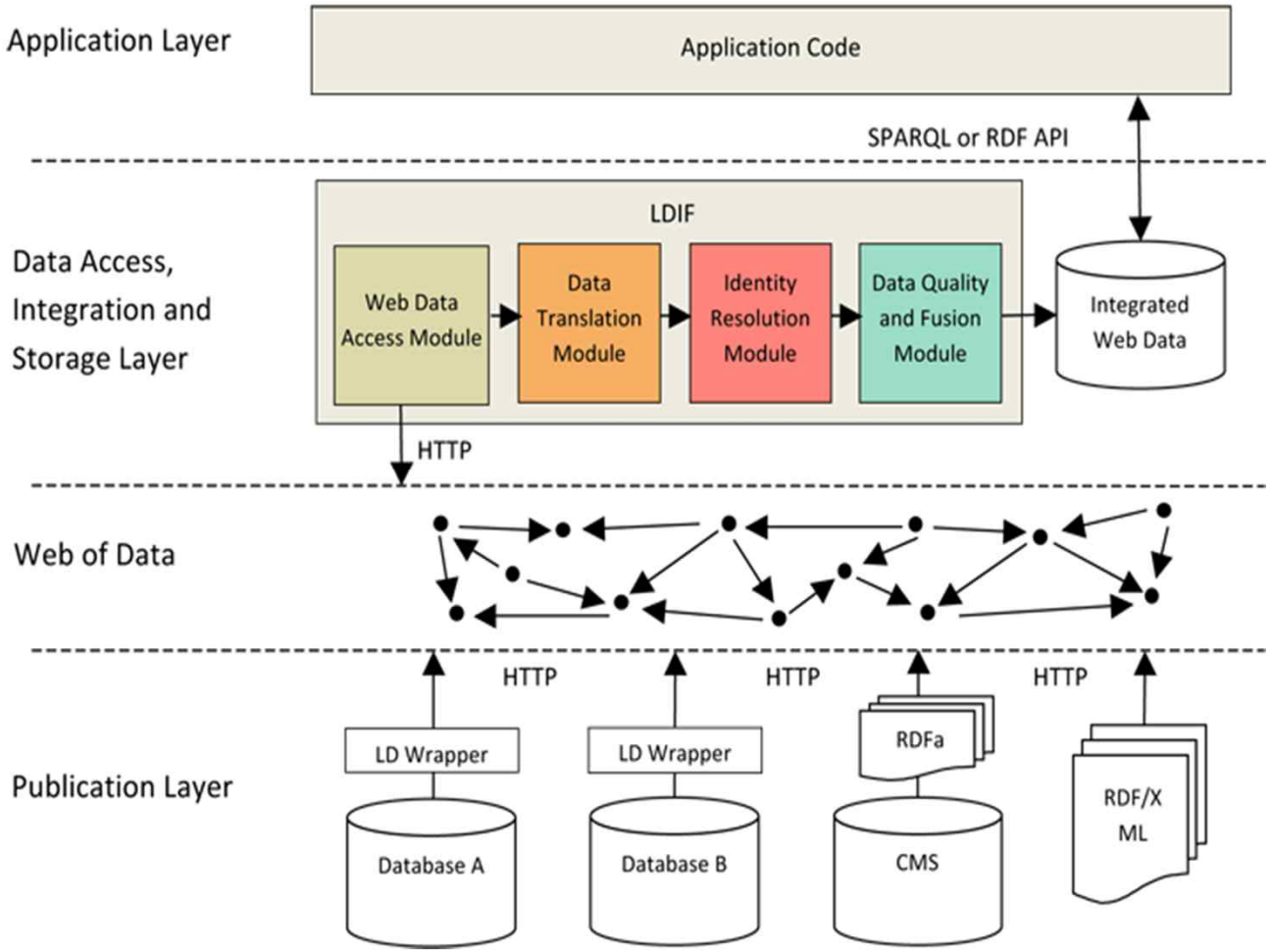
http://www.telecom-cloud.net/iot-big-data-networked-programmable-and-machine-learning/

Geospatial Data of OGC





Linked open Data (LoD) of Web



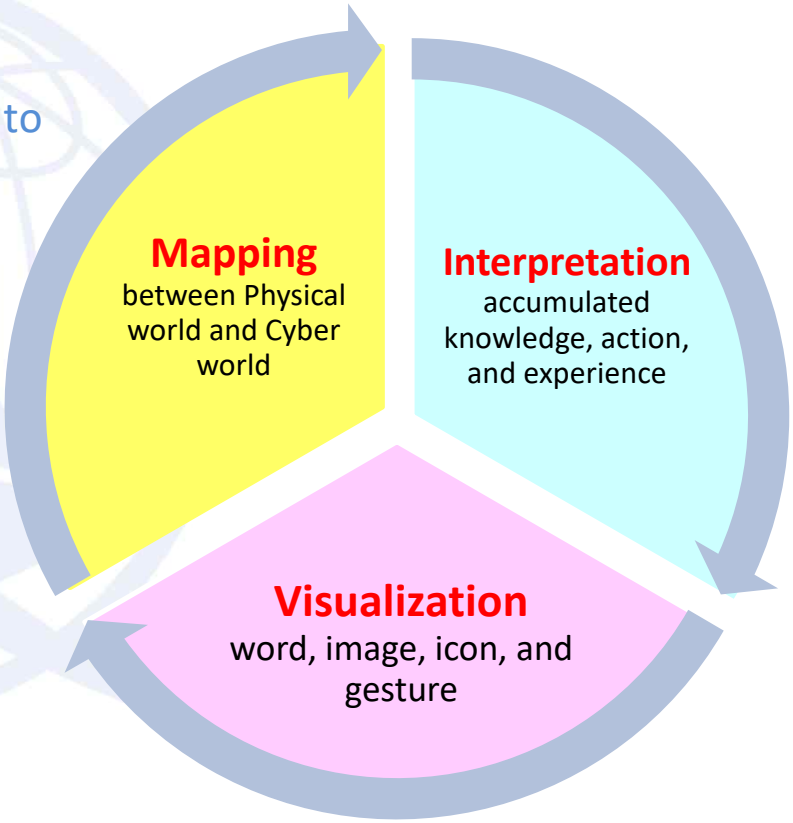
KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

Approaches toward New Data Society

- **Key Issues to realize New Data Society**
  - **How to utilize interactive power of computer and communications technology ?**
- **General Approaches toward New Data Society**
  - **Behavior Cycle** for future human life and business culture
  - **Open Environments** for Network, Software, and Device
  - **New data format** to create, deliver, and consume
  - **Eco-systems** between physical society and cyber society
  - **Intelligent emerging devices** including smart phone, smart TV, smart car, smart building, and smart things, etc.

Creating New Data World

- **Need new data types for knowledge society**
  - (**Mapping**) Data types connecting physical world to cyber world
    - Identification/classification, Location, Status, Role/Function, etc.
    - (Example) GPS, address, serial no. etc.
  - (**Interpretation**) Data types for human understanding, meaning, interpretation, and translation, perception, recognition based on accumulated knowledge, reasoning, learning, action, behaviors, and experience
    - (Example) Pythagorean Theorem, know-how of medical treatments, etc.
  - (**Visualization**) Data types for communication, sharing, visualization, rendering, expression of word, image, gesture, etc.
    - (Example) icon, logo, graphic image, character sets,



Data Types depending on Applications

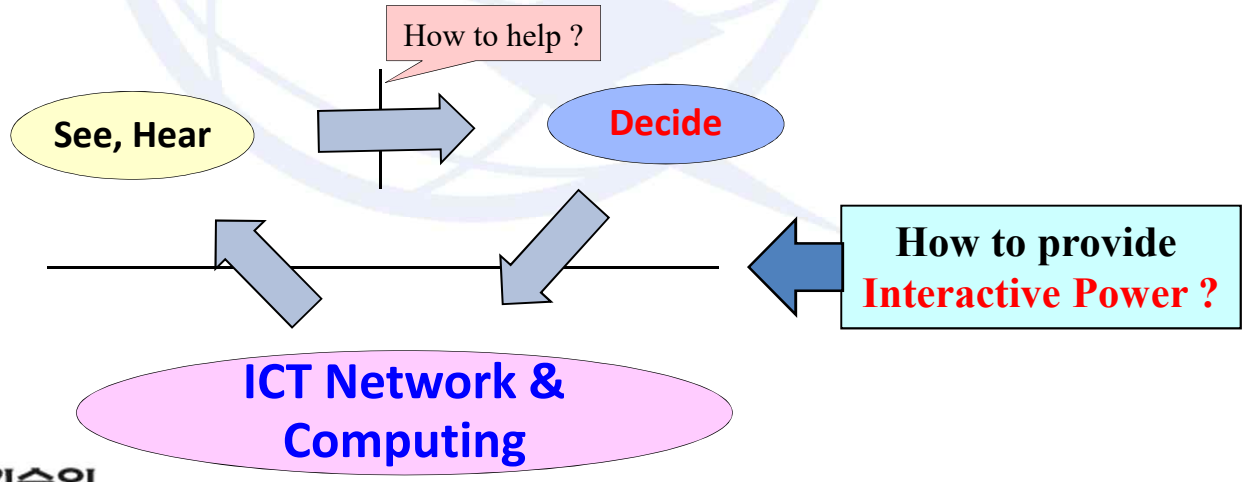
- **Telecommunication and Broadcast Industry**
  - Telephony, SMS, AV/Multimedia stream, and AR/VR, etc.
- **Internet and Web Industry**
  - File, image, documents, and social media, etc.
  - Virtual/Object Data, and web of data
- **Location related Industry**
  - GPS, CPS, and physical 2D/3D geographic location
  - Transport and Logistics (Geolocation Map)
- **Identification related Industry**
  - Sensor/RFID code, product code, bar code, and blockchain, etc.
  - Trade, copyright, and ownership (Shipping code, product code, watermark, etc.)
- **Data Intensive or Contexts related Industry**
  - Big data analytics, Healthcare and medical applications

Requirements for Future Data Format

- **(Basic) Recognize by human organ and their supporting utilities**
  - Simple audio/sound, and visual image
  - Language including translation tools
- **(Extended) Accumulation, Filtering, and Processing**
  - Linked chains among related data or sensing data from IoT equipment
  - Accumulation by collective intelligence and crowdsourcing
  - Filtering based on experience, preference, and accumulated know-hows
  - Creation of new information and knowledge (like big data processing)
- **(Applications) IoT Media, Energy Avatar, Traffic Guider**
  - Context-aware information based on IoT sensors and AV devices
  - Navigation assistant (or Guider) for road and traffic conditions
  - Energy Avatar for analysis and prediction of energy consumption

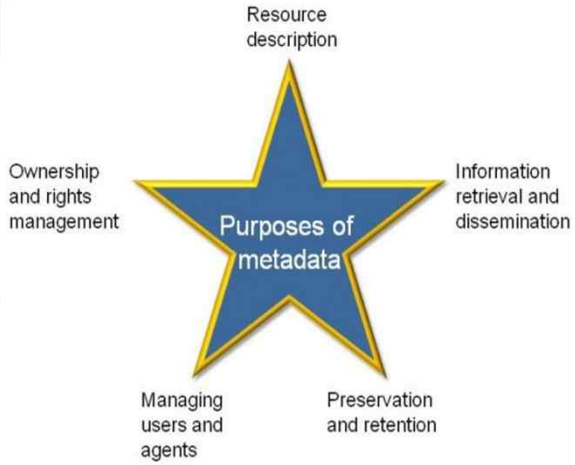
Technical Issues for Data Platform

- **Usage Behavior Analysis**
  - Behaviors for Human Relationship (chat, discuss, share, etc.)
  - Behaviors for Entertainments (Game, TV, Drama, Film)
  - Behaviors for Works (collect, search, analyze, decide, etc.)
  - Behaviors for Life (Shopping, Education, Dining, etc.)
- **Interaction Process between Human and ICT Environments**



Open Data Platform - 1

- **(Metadata) Data could not stand alone without metadata or descriptors**
  - **Re-define Data and Metadata** → Metadata is not only descriptive information of data
    - used for handling, sharing, and processing data
  - **New data format including metadata is needed**
    - Active Hyperlink or JavaScript at web → like computer virus
    - New Metadata is different from existing metadata standards
    - Recursive data format according to levels of perception and intelligence



(ref) http://aspiresquared.co.uk/

KAIST 한국과학기술원 Korea Advanced Institute of Science and Technology

Open Data Platform - 2

- **(Platform) new model for data processing including database**
  - **"Data + Processing + Storage" in harmony → web-based common platform?**
  - **New version of web platform → HTML5-based IoT/WoT world?**
    - How to contain location, status, behaviors information that is not descriptive from existing web standards (such as text script/binary-based web)
    - Recursive data format to support complex and iterative algorithm or logics
    - New Markup language to adopt new UI/UX tools (e.g., 2D/3D drawing, gesture, expression, etc.)
  - **Data platform for IoT/WoT application !**
    - Web platform to reflect physical world (e.g., new organic sensors, etc.)
    - Semantics for experience/knowledge accumulation from IoT devices
  - **How to build Cyber Physical System for future flexibility?**
    - 3D virtual space, location, and depth/granularity/tier/level/attribute, etc.

Open Data Platform - 3

- **(Next Generation Web) good for future open data world ?**
  - Common platform for data creation, delivery, share, and consumption
  - Keep Simple User Interface and allow billions of software/applications
    - Utilize existing wireline/wireless network, computing/software, and database
    - No download and no installation → Just access and use it !
  - Support flexible data platform for energy, transport, medical/health, education, and safety, etc.
    - Don't steal data from owners and customers → good data governance !
- **(Cloud) sharing data by using cloud platform**
  - Performance, security, availability, and manageability of cloud platform are suitable for future data eco-society
  - Open ecosystem for data sharing and good data governance

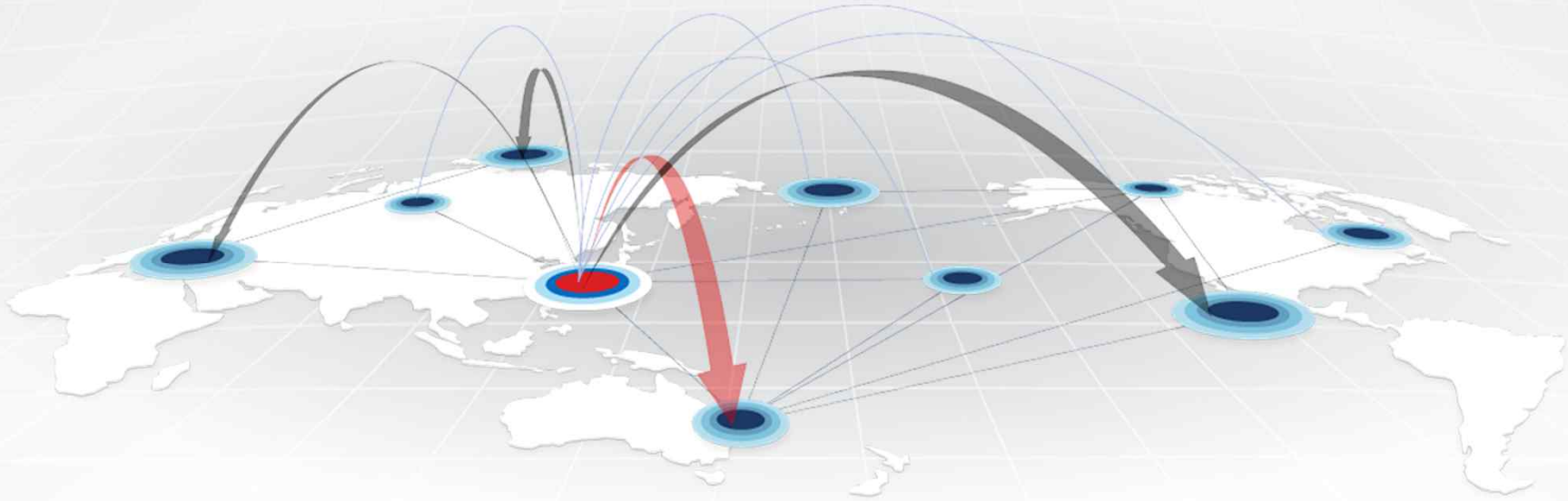
Technical Issues of New Data Format

- **HTML5-based UI/UX to help human perception**
  - Device API, RESTful interface, Javascript, etc.
- **Data and Metadata Together**
  - Media object and media resource model
  - Media Ontology, Media Annotation
  - File Format, Microformat, ATOM/RSS
  - Device Data Format for Mashup
- **Data format for Web-based open Platform**
  - Open, Auto-configurable, and future flexible
  - But, Securable and manageable is in question ?

Web-based IoT Data Format

- **XML/RDF Schema**
  - Well-known data format ?
    - JSON, Microformat (e.g., vCard, hCalendar), ATOM/RSS
  - DTD syntax, schema, and semantics, etc.
    - Tag, Index, Summary, Thumbnail, Preview, etc.
- **IoT Data Format for future life and business**
  - Continue to URI/URL/URN-based ?
  - Microservice, microdata for semantics
  - Mashup format for heterogeneous data applications

IoT Data Standards for Smart City



IoT Data Standardizations for Smart City

- IoT data cloud system for Smart City
- Real-time data analytics for Smart City
- IoT metadata for Smart City
  - Syntax, semantics, context-aware for unstructured IoT data
- Data security and trust for Smart City
- Data applications for Smart City (including visualization)
- Interoperability and Deployment scenarios for IoT-based Smart City


Thank you!

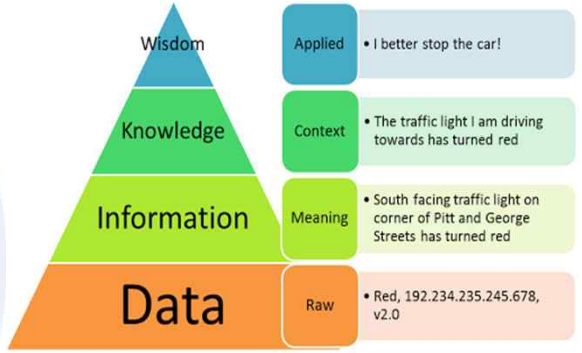
Appendices

Meanings of Data, Information, and Knowledge - 1

- **(Value) "Data is King" or "Data is new Oil"**
  - **New Value when data is processed and accumulated**
    - Added new value when combined with other data and different environments
    - Different meaning and understanding when the environments is changed
  - **Various form of data when it is created, shared, processed, and utilized**
    - Depending on applications (energy, transport, health, education, safety, etc.)
    - Supported by **metadata** and descriptor
    - Depending on activated conditions and environments
      - Vitalized with the related information and filtering conditions
    - May have self-proliferation phenomenon of biological cell at future
  - **But, Data is very "Dangerous" like Nuclear Power**
    - Imagine radioactive contaminated water

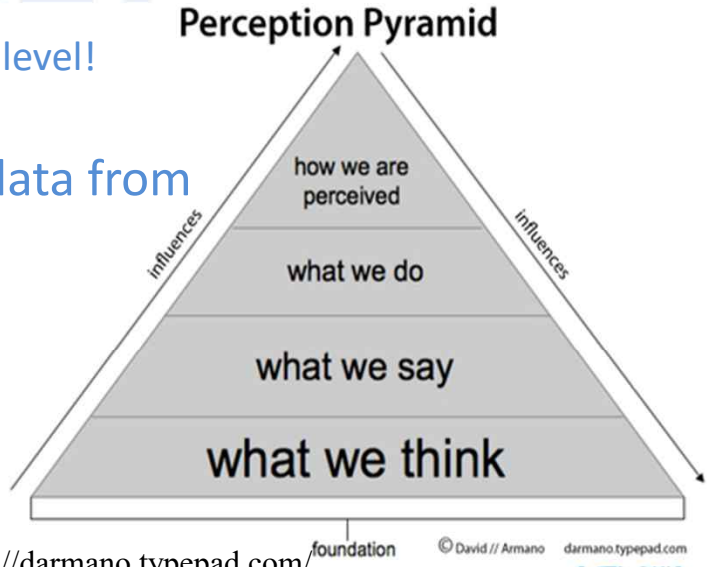
Meanings of Data, Information, and Knowledge - 2

- **(DIKW) Value chains among Data, Information, and Knowledge**
  - **Linked Structure** among data, information, and knowledge
  - How to make a **process** of data to get useful information
    - No meaning like "01100101100110…"
  - Tier-x data extracted from **intelligent filtering** process
    - Based on previous or background information, and statistics, etc.
  - **Decision making** by aggregating the related data
    - Depending on know-how, experience, and context-awareness
- **(Environments) What environment for data is valuable**
  - At **right** time, **right** place, and **right** condition
  - Location-, time-, and context-**awareness**: 5W1H rule!



Data Sciences for Knowledge Society - 1

- **(Human Perception) How to make a process from data acquisition to get knowledge?**
  - What percentage of data is perceived by human?
    - Could not monitor all the CCTV cameras during 24 hours/day
  - What amounts of information are extracted from same image and sound?
    - Depending on his/her experience and intelligence level! (e.g., X-ray image, intonation of sound, smell, etc.)
  - If new tools create, share, and utilize open data from IoT devices
    - Identify location and time
    - Recognize screen image and situation!



(ref) http://darmano.typepad.com/

Data Sciences for Knowledge Society - 2

- **(Type/Format) Cyber physical space for Energy, Transport, Health, Environment, and Surveillance, etc.**
  - Various **Standards** for 3D physical space!
    - 3D scalable geo-graphic, MPEG, OGC, Web3D, OpenGL, etc.
  - **Resolution and Accuracy** of cyber physical 3D information
    - Depending on applications, data volume, and processing time, etc.
  - **Data Type and Format** for IoT/WoT applications
    - Data format for energy, transportation, health, surveillance/monitoring, etc.
    - 3D game, 3D image for navigation, 3D simulation for war and building construction, etc. → data format may be different!

Data Sciences for Knowledge Society - 3

- **(Accumulation) Learning/Experience for Cognition Process of Knowledge**
  - **How to accumulate human knowledge collectively!**
    - Collectively intelligence by mechanisms of perception – see, hear, touch, smell, taste, and attention
    - New platform for knowledge sharing and accumulation!
  - **Sharing tacit knowledge and real/virtual experience!**
    - Experiences of driving, cooking, and gaming by simulation at virtual space

KAIST 한국과학 Korea Advanced Institu

ITU

Linked Chain of Data

- **New values of Linked Open Data (LoD)**
  - **New Values when Data are linked**
    - H + H + O → Water [H_2O]
    - Dynamic Hyperlink among similar and/or heterogeneous data
  - **Environments/Conditions such that Data is meaningful**
    - vegetable + salt + pepper + **pot [Environment]** → Kimchi
    - "On the Origin of Species" → survival plan of live data
    - CCTV camera + location + status → meaningful data/information

Data Model of Resource Description Framework (RDF)

- **Entity–Attribute–Value model**
  - Making statements about resources
    - (Examples) XML DTD (Document Type Definition), JSON, tag, name, address, etc.
- **Data Serialization Model**
  - File, memory buffer, packets of communication protocol, time-varying data (A/V), etc.
  - Text-based/binary-based, structured/unstructured, hierarchical/non-hierarchical, scalar/vector/graph, class/object, etc.
    - (Examples) Binary/Integer/Real/Exponent/Character/String/Boolean/Time, Vector/Matrix/Array, 2D/3D Graphics, Recursive, Audio/Video Stream, etc.
- **Markup/Metadata/Schema/Semantic Model**
  - Specify the processing to be performed or the related actions (i.e., activate, trigger, invoke, etc.)
  - Create the values of the data
    - (Examples) metadata, semantic ontology for IoT/CPS, OWL, etc.

KAIST 한국과학기술원
Korea Advanced Institute of Science and Technology