DigitalWorld 2018 International Expert Panel:

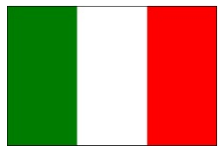
# Quality of Data and Services:

## The Essences From Geo-computing, Society/Crowd Sensing, and Mobility/Service-related Sensing

March 27, 2018, Rome, Italy

DigitalWorld / The Tenth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2018) / ALLSENSORS / ICDS


GEOProcessing, ALLSENSORS, ICDS / DigitalWorld
March 25–29, 2018 - Rome, Italy



# DigitalWorld Expert Panel: . . . Contemporary Views . . .

## Panelists and Contributors

- *Claus-Peter Rückemann* (Moderator),
  Westfälische Wilhelms-Universität Münster (WWU) /
  Leibniz Universität Hannover /
  North-German Supercomputing Alliance (HLRN), Germany
- *Yerach Doytsher,* *[Panelist and Contributor]*
  Technion - Israel Institute of Technology, Israel
- *Paulo E. Cruvinel,* *[Panelist and Contributor]*
  Embrapa Instrumentation, São Carlos, Brazil
- *Muhammad Haris,* *[Contributor]*
  GIS Center, PUCIT University of the Punjab, Pakistan
- *Claus-Peter Rückemann,* *[Panelist and Contributor]*
  WWU Münster / Leibniz Universität Hannover / HLRN, Germany

DigitalWorld / GEOProcessing 2018 / ALLSENSORS / ICDS:

`http://www.iaria.org/conferences2018/GEOProcessing18.html`
`http://www.iaria.org/conferences2018/ALLSENSORS18.html`
`http://www.iaria.org/conferences2018/ICDS18.html`

Program: `http://www.iaria.org/conferences2018/ProgramGEOProcessing18.html`

# DigitalWorld Expert Panel: . . . Contemporary Views . . .

## Panel Statements and Preview:

- **Knowledge view and value:**
  - Quality of Data (and Services) is basically conjoint with the being of knowledge.
  - Grade of value is closely associated with knowledge/data quality and implementation.
  - In contemporary context, knowledge/data/service value can be exploited by realisation.
- **Volunteered Geographic Information and Internet of Things:**
  - Internet of Things (IoT) is contributing to volunteered geographic information/data.
  - IoT/IoP (Internet of People) are used by industry for communication and creating connects.
  - Spatial crowdsourcing is contributing to very promising projects.
- **Agricultural data, geo-computing, IoT, and decison Making:**
  - Data quality is relevant for decision making, knowledge mining, . . . (modeling and crop management, control of inputs, automation, analysis, . . .).
  - Approaches depend on scenario and amount and level of "data with quality" required.
  - Facing "Industry 4.0", needs for digital development, geo-computing, sensor networks, Internet of Things (IoT), Deploy technical opportunities, increase productivity, sustainability, security, safety, precision, faster response to incidents, achieve consumers's confidence, . . .
- **Spatial data from social networks:**
  - Spatial data and its quality are vital parts of social networking platforms.
  - "Spatial data" from social networks can be improved.
  - Object/attribute models can foster the integration of spatial data and its quality.

# DigitalWorld Expert Panel: . . . Contemporary Views . . .

## Pre-Discussion-Wrapup:

- **Knowledge:** What does the understanding of knowledge have to do with data and services and their value and quality.
- **Centricity:** Does centricity matter for Quality of Data and Services?
- **Implementation and realisation:** What is their impact on value and quality?
- **Internet of Things:** How is the IoT contributing?
- **Data quality:** Why/how is data quality relevant for decision making . . .?
- **Level:** Which "data quality" levels are required for different scenarios?
- **Purpose:** How can technical opportunities be deployed, productivity be increased, sustainability, consumer's confidence . . . be achieved?
- **Tackling:** How to tackle security, safety, precision, faster response . . .?
- **Improvement:** How can "spatial data" (e.g., from social networks) be improved?
- **Object/attribute models:** What is their contribution fostering spatial data integration and quality?
- **Views:** Different academia/industry applications/views?
- **Recommendations:** Which general and special recommendations?
- **Networking:** Discussion! Open Questions? Suggestions for next Expert Panel?

# DigitalWorld Expert Panel: Conclusions / Post-Panel-Discussion Summary

## Post-Panel-Discussion Summary (2018-03-28):

Resulting from the discussed topics, the essences of quality of data and services are:

- **Knowledge view and value:** It is crucial to solidly understand the essentials of knowledge (especially, Aristotles' ethics and modern definition and concept of factual, conceptual, metacognitive, procedural knowledge). Significant value is in the knowledge and its quality. Implementation and realisation can contemporarily exploit data / service value. Education and practice are essential for application and decisive management. (Rückemann)
- **Volunteered Geographic Information and Internet of Things:** IoT is contributing to volunteered (geographic) information/data. Crowdsourcing can deliver data quality suitable for many scenarios. Examples and endeavours: Industry "generations", Web "generations", spatial crowdsourcing contributing to OSM, . . . (Doytsher)
- **Agricultural data, geo-computing, IoT, and decison Making:** Data quality is relevant for advanced application scenarios as well as for decision making. Productivity, sustainability, security, safety, industry "generations" require many contributing technologies, e.g., geo-computing, sensor networks, and IoT in order to achieve best results and consumers's confidence. Solutions to challenges depend on "data with quality". (Cruvinel)
- **Spatial data from social networks:** Spatial data and its quality are vital parts of social networking platforms. Object based models can help improving "spatial data" from social networks. (Haris)
- **Services are based on knowledge / data. Quality of knowledge / data is most relevant.**
- **Honouring the value, the panel proposes to establish "Knowledge Level Agreement (KLA)" and "Data Level Agreement (DLA)" as counterparts for services.**
- **Requirements regarding knowledge / data quality are different for application scenarios.**
- **Knowledge / data quality can justify both "automated' processes like crowdsourcing AND manual processes in hands of specialised experts.**

# DigitalWorld Expert Panel: Table of Presentations, Attached

## Panelist Presentations: (presentation order, following pages)

- **A Knowledge View on Quality of Data and Services:**
  **The Grade of Value and its Contemporary Invigoration** (*Rückemann*)
- **Contribution of Internet of Things to**
  **Volunteered Geographic Information** (*Doytsher*)
- **Agricultural data for geo-computing,**
  **sensors' network, and advances in IoT:**
  **quality requested in the decision making processes** (*Cruvinel*)
- [
  **Contribution topic and case / no presentation:**
  **An object based model for improving spatial data from social networks** (*Haris*)
  ]

# DigitalWorld 2018 International Expert Panel: Quality of Data and Services:

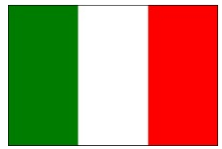
## The Essences From Geo-computing, Society/Crowd Sensing, and Mobility/Service-related Sensing

# A Knowledge View on Quality of Data and Services: The Grade of Value and its Contemporary Invigoration

DigitalWorld / Tenth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2018) / ALLSENSORS / ICDS

March 27, 2018, Rome, Italy

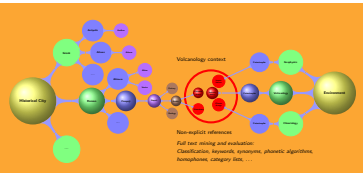

Dr. rer. nat. Claus-Peter Rückemann[1,2,3]

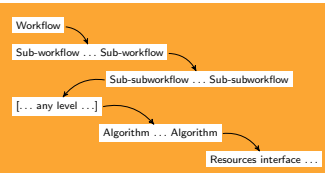


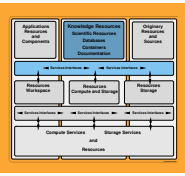
[1] Westfälische Wilhelms-Universität Münster (WWU), Münster, Germany
[2] Leibniz Universität Hannover, Hannover, Germany
[3] North-German Supercomputing Alliance (HLRN), Germany

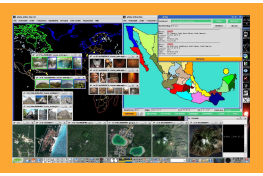
ruckema(at)uni-muenster.de









## Status:

- **Knowledge and data:**
  Isolated terms, marketing driven changes of paradigms, buzz context of data-knowledge-intelligence-wisdom-information-. . ., inconsistency of contextualisation (DIKW "Data - Information - Knowledge - Wisdom", x.x "versioning" of abstruse "achievements").
- **Data and complexity:**
  "Inexhaustible (natural, free) resource", "complexity must be reduced", "has to be done economically".
- **Value:**
  Economic interests, e.g., expressed by services.
- **Driving forces:**
  Off-topic management strategies, cheap success.
- **Consequence:**
  "Fuzzy" understanding and actions.
- **Quality is a matter of abstruse "measuring" and "promotion" approaches.**

## Vision of understanding:

- **Knowledge and data:**
  Defined terms, solid and holistic concept and implementations.
- **Data and complexity:** Has to be invested in, should be deployed/used.
- **Value of data:** Not depending on economic aspects of services.
- **Driving forces:** Scientists and researchers.
- **Consequence:** Fuzzy understanding and actions.
- **Quality is a matter of education AND experience based expertise.**

## Vision regarding knowledge and quality:

- **Understanding:** The complementary nature of precise sciences and data science.
- **Defining:** Data, knowledge, ...
- **Implementing:** Applications based on appropriate means and measures.
- **Preserving:** Essential data, knowledge, and value (for long-term).

## Knowledge and Ethics:

- **Nicomachean Ethics** (Aristotle, 350 B.C.E.)

## Knowledge Complements:

- **Factual knowledge**, **conceptual knowledge**, **procedural knowledge**, **metacognitive knowledge** (Anderson, Krathwohl, 2001)

## Knowledge, defining (Summit delegates and contributors)

- **"Knowledge is created from a subjective combination of different attainments as there are intuition, experience, information, education, decision, power of persuasion and so on, which are selected, compared and balanced against each other, which are transformed, interpreted, and used in reasoning, also to infer further knowledge. Therefore, not all the knowledge can be explicitly formalised. Knowledge and content are multi- and inter-disciplinary long-term targets and values. In practice, powerful and secure information technology can support knowledge-based works and values."**

**Citation:** *Rückemann, C.-P., Gersbeck-Schierholz, B., and Hülsmann, F., Przemysław Skurowski, Michał Staniszewski (2015): Post-Summit Results, Delegates' Summit: Best Practice and Definitions of Knowledge and Computing; September 23, 2015, The Fifth Symposium on Advanced Computation and Information in Natural and Applied Sciences (SACINAS), The 13th International Conference of Numerical Analysis and Applied Mathematics (ICNAAM), September 23–29, 2015, Rhodes, Greece.*

**Delegates and contributors:** *Claus-Peter Rückemann, Friedrich Hülsmann, Birgit Gersbeck-Schierholz, Knowledge in Motion / Unabhängiges Deutsches Institut für Multi-disziplinäre Forschung (DIMF), Germany;Przemysław Skurowski, Michał Staniszewski, Silesian University of Technology, Gliwice, Poland; International EULISP post-graduate participants, ISSC, European Legal Informatics Study Programme, Leibniz Universität Hannover, Germany*

## Quality and Knowledge

- **Understanding of knowledge as a key factor to value and quality**. It is crucial to solidly understand and apply the essentials of knowledge. (especially, holistic context, modern knowledge definitions, and implementation concepts).
  Significant value is in knowledge and its quality.
  *Knowledge is used to generate (new) knowledge.*
- **Centricity concepts can foster strategic developments**.
- **Long-term, multi-disciplinary work creating knowledge resources**.
  Implementation and realisation can contemporarily exploit data/service value, e.g., deploying long-term and multi-disciplinary knowledge resources.
- **Implementation and realisation of services based on knowledge resources**.
  Impact on value and quality at implementation/realisation time interval.
- **Deployment of appropriate means:** Education and practice are essential for decisive management, integration, and documentation.
- **Means, examples:** Data-centric and computing-centric implementations,
  Universal Decimal Classification (UDC), e.g., codes for spatial featues and place:
  UDC:(1) place and space, UDC:(2) physiographic designation,
  UDC:(3) place of ancient and medaevial world,
  UDC:(4/9) countries and places of the modern world,
  Unified Modeling Language (UML),
  Knowledge discovery, High End Computing, Standardisation, ...

## Conclusions:

- **Quality of Data (and Services) is basically conjoint with the being of knowledge.**
- **Value of data is not depending on economic aspects of services.**
- **Grade of value is closely associated with knowledge/data quality and implementation.**
- **Knowledge/data/service value can be exploited by realisation, in contemporary context.**
- **Investments in knowledge and quality need to be done.**
- **Education is required on understanding of knowledge and quality, implementation and realisation, architectures, scenarios, standardisation, . . .**
- **Education on knowledge is required for decisive management and staff.**

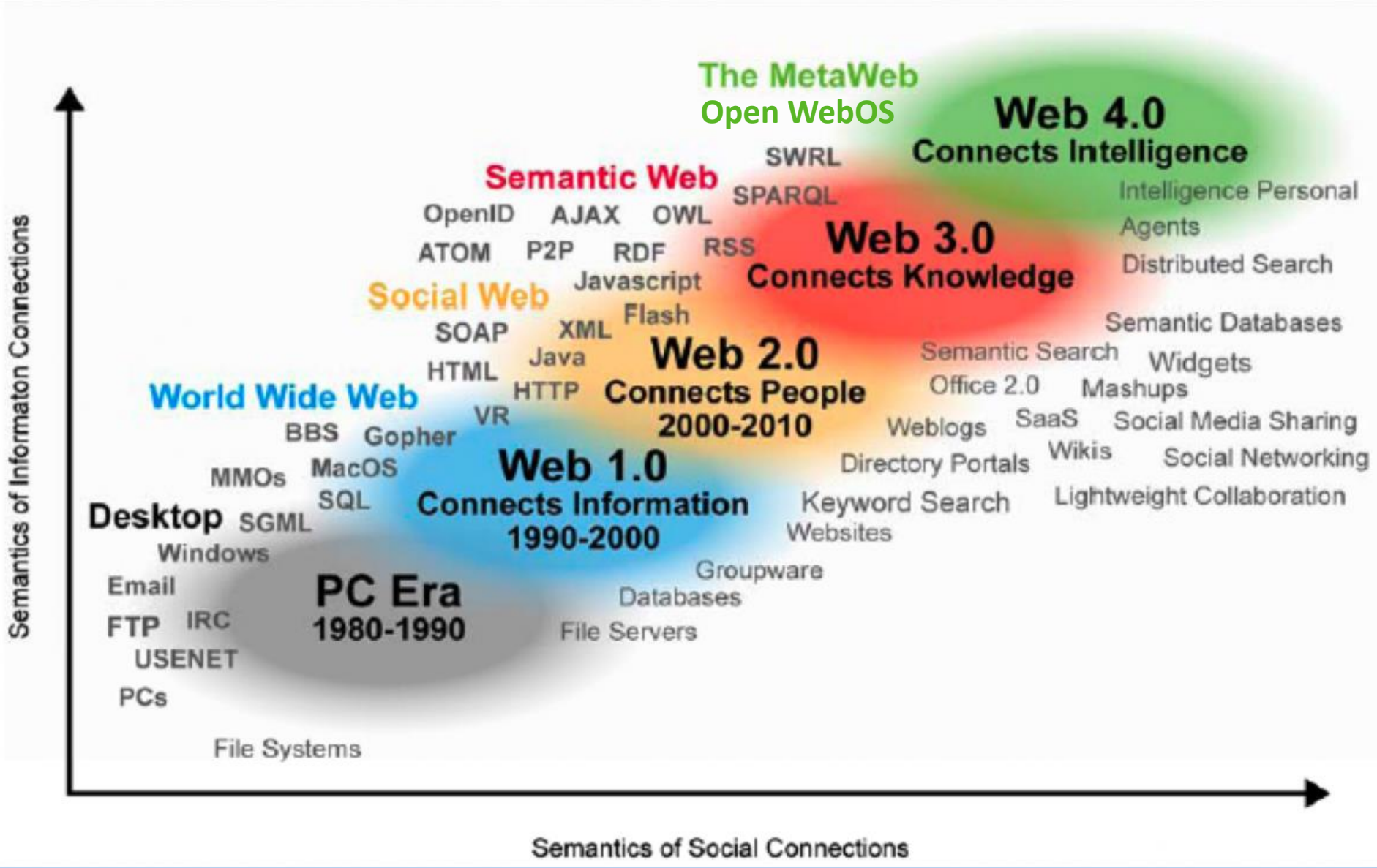
**Thank you for your attention!**

**Wish you an inspiring conference and a pleasant stay in Rome!**

# Contribution of Internet of Things to Volunteered Geographic Information

Prof. Dr. Yerach Doytsher
Mapping and Geo-Information Engineering,
Technion, Israel

# WEB Generations

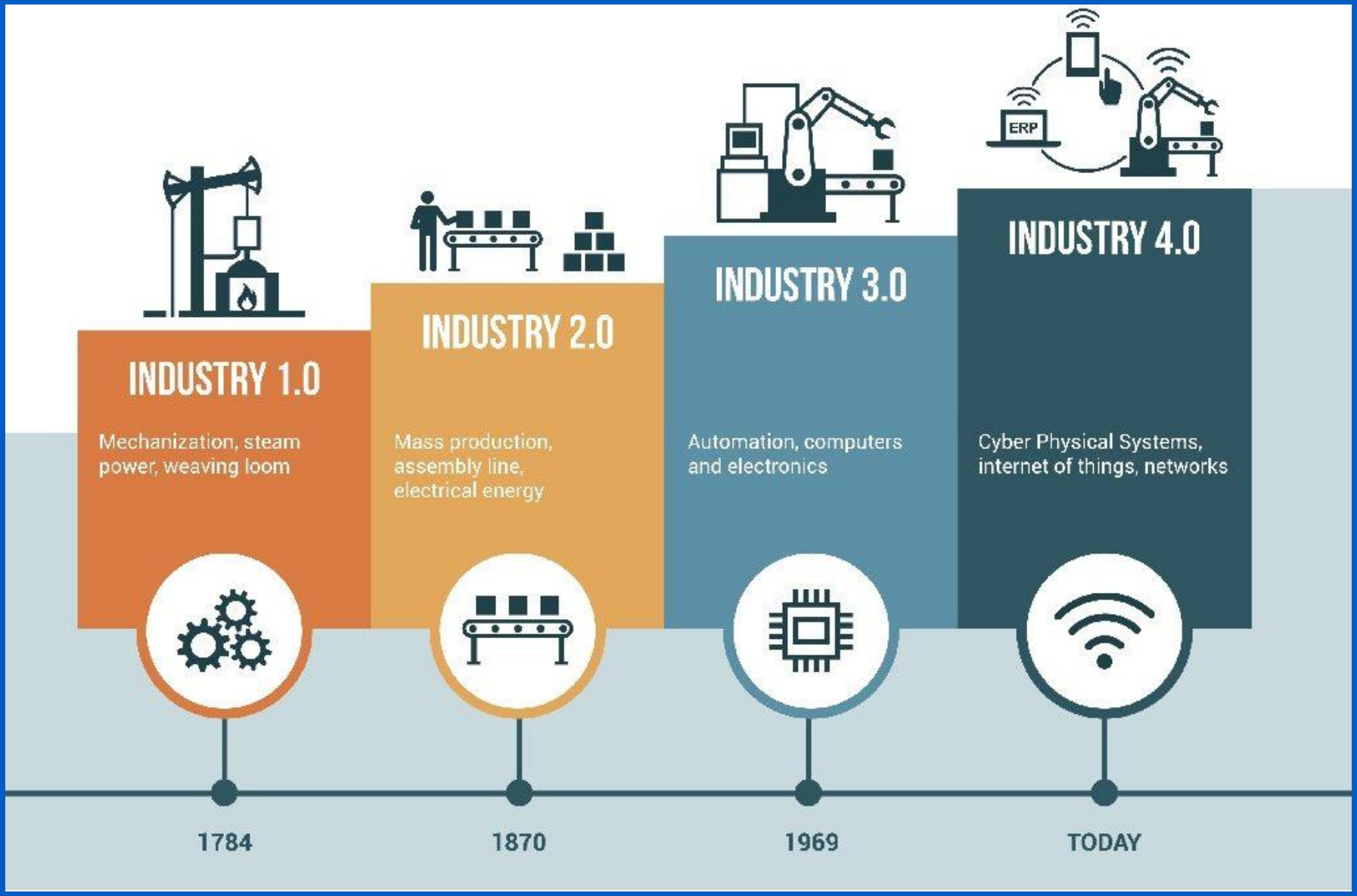


# Generations

## Talking a different language

| | Maturists (pre-1945) | Baby boomers (1945-1960) | Generation X (1961-1980) | Generation Y (1981-1995) | Generation Z (Born after 1995) |
|---|---|---|---|---|---|
| Formative experiences | Wartime rationing<br>Rock'n'roll<br>Nuclear families<br>Defined gender roles - particularly for women | Cold War<br>'Swinging Sixties'<br>Moon landings<br>Youth culture<br>Woodstock<br>Family-orientated | Fall of Berlin Wall<br>Reagan/Gorbachev/ Thatcherism<br>Live Aid<br>Early mobile technology<br>Divorce rate rises | 9/11 terrorists attacks<br>Social media<br>Invasion of Iraq<br>Reality TV<br>Google Earth | Economic downturn<br>Global warming<br>Mobile devices<br>Cloud computing<br>Wiki-leaks |
| Attitude toward career | Jobs for life | Organisational - careers are defined by employees | "Portfolio" careers - loyal to profession, not to employer | Digital entrepreneurs - work "with" organisations | Multitaskers - will move seamlessly between organisations and "pop-up" businesses |
| Signature product | Automobile | Television | Personal computer | Tablet/smartphone | Google glass, 3-D printing |
| Communication media | Formal letter | Telephone | E-mail and text message | Text or social media | Hand-held communication devices |
| Preference when making financial decisions | Face-to-face meetings | Face-to-face ideally but increasingly will go online | Online - would prefer face-to-face if time permitting | Face-to-face | Solutions will be digitally crowd-sourced |

# Industry Revolutions



# Industry 4.0

- The term "Industry 4.0" originates from a project in the high-tech strategy of the German government, which promotes the computerization of manufacturing.

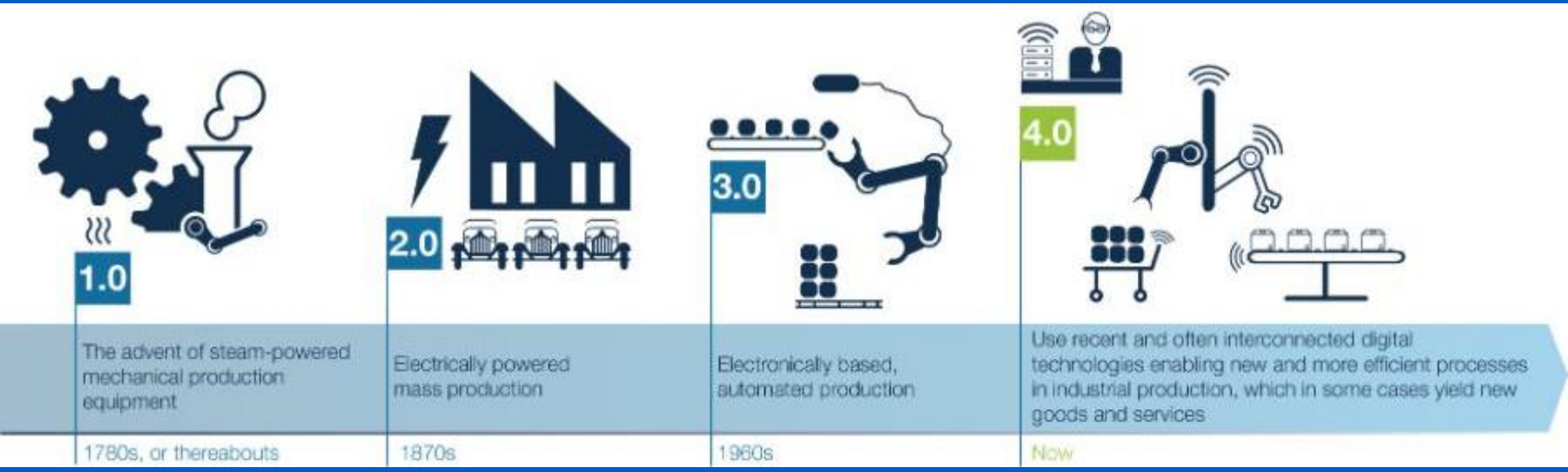
- The term "Industry 4.0" was revived in 2011 at the Hannover Fair. In October 2012 the Working Group on Industry 4.0 presented a set of Industry 4.0 implementation recommendations to the German federal government. The Industry 4.0 workgroup members are recognized as the founding fathers and driving force behind Industry 4.0.

- Unlike the other three industrial revolutions the name of that revolution was declared at the very beginning, as 4$^{th}$ Industrial revolution.

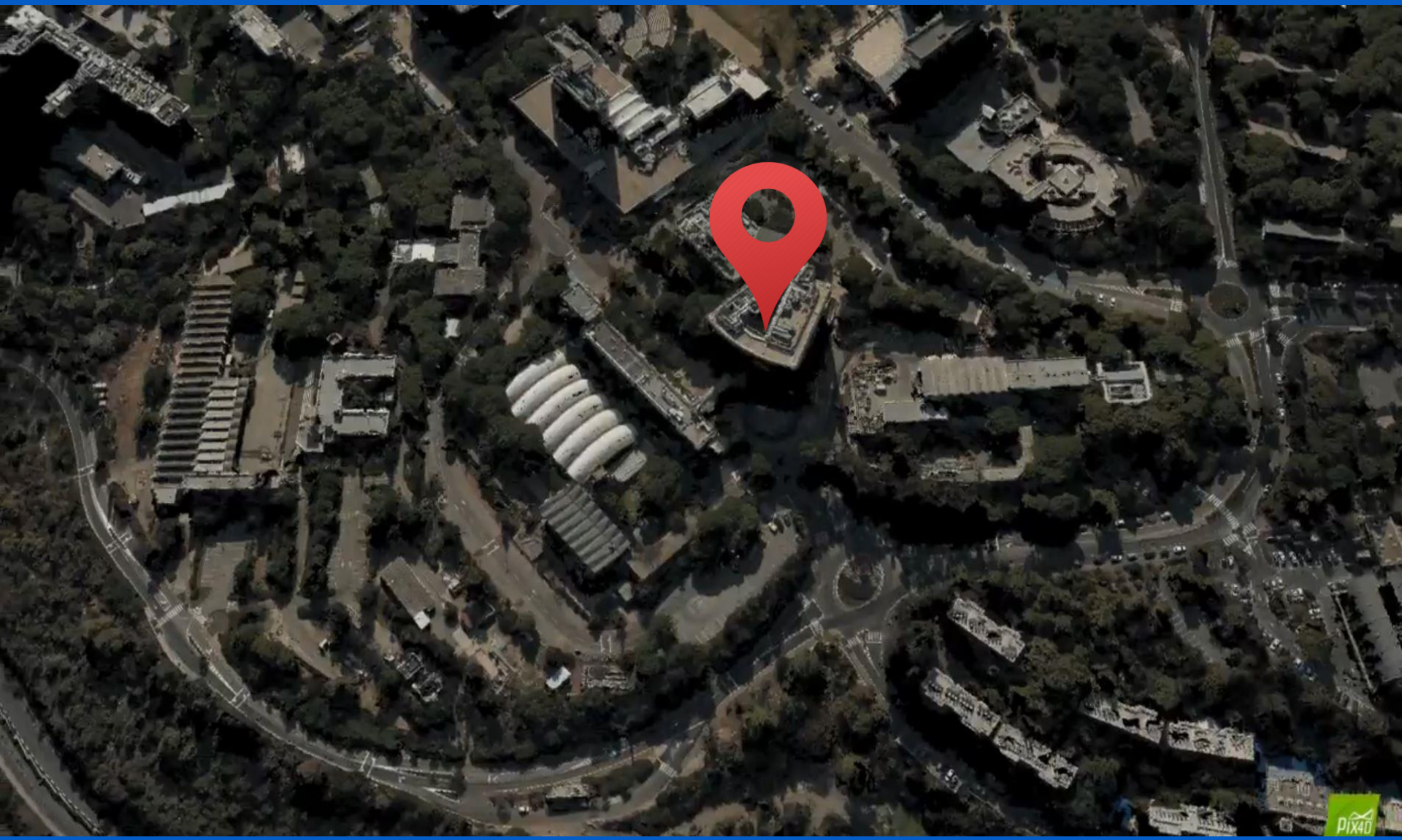
# Design Principles of Industry 4.0

- **Interoperability:** The ability of machines, devices, sensors, and people to connect and communicate with each other via the Internet of Things (IoT) or the Internet of People (IoP)
  - Adding IoT will further automate the process to a large extent
- **Information transparency:** The ability of information systems to create a virtual copy of the physical world by enriching digital plant models with sensor data. This requires the aggregation of raw sensor data to higher-value context information.
- **Technical assistance:** First, the ability of assistance systems to support humans by aggregating and visualizing information comprehensibly for making informed decisions and solving urgent problems on short notice.
- **Decentralized decisions:** The ability of cyber physical systems to make decisions on their own and to perform their tasks as autonomously as possible.

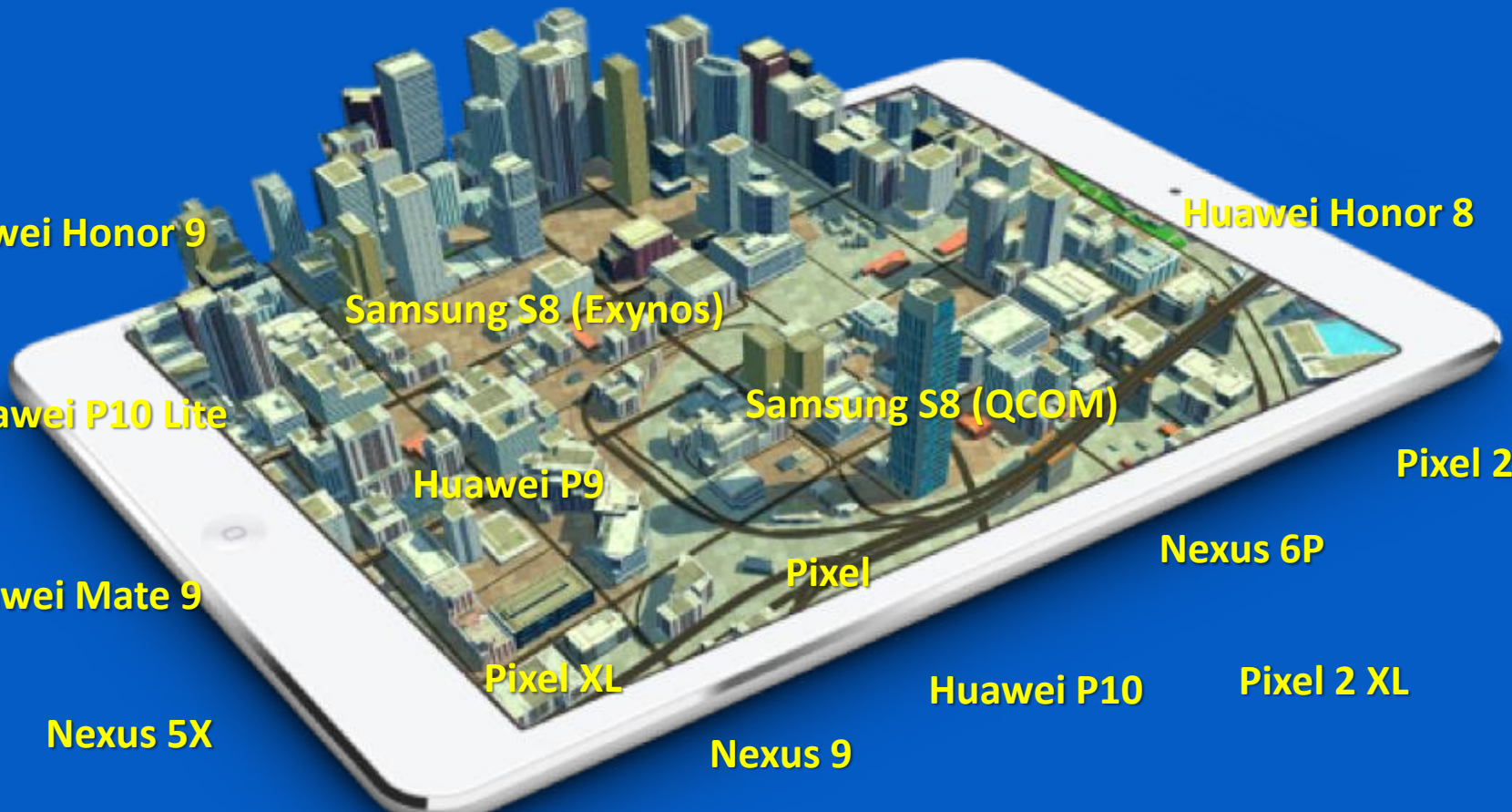
# Industry 4.0 – Internet of Things -IoT

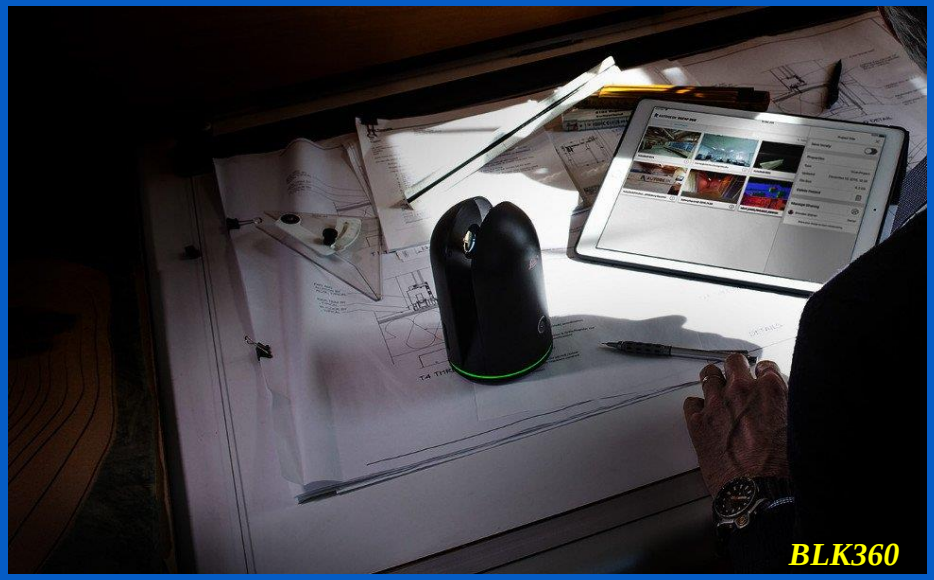


# Positioning / Spatial Information



# Mobile Devices GNSS Raw Data

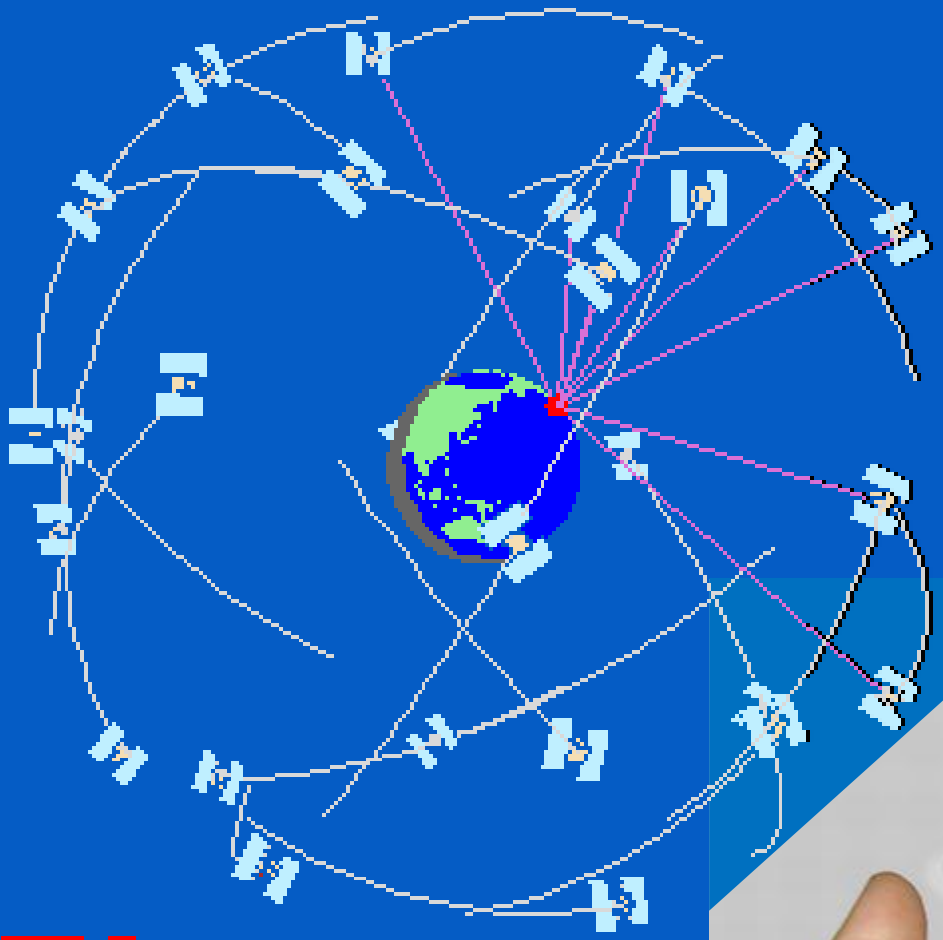




*BLK360*

*iphone 8*

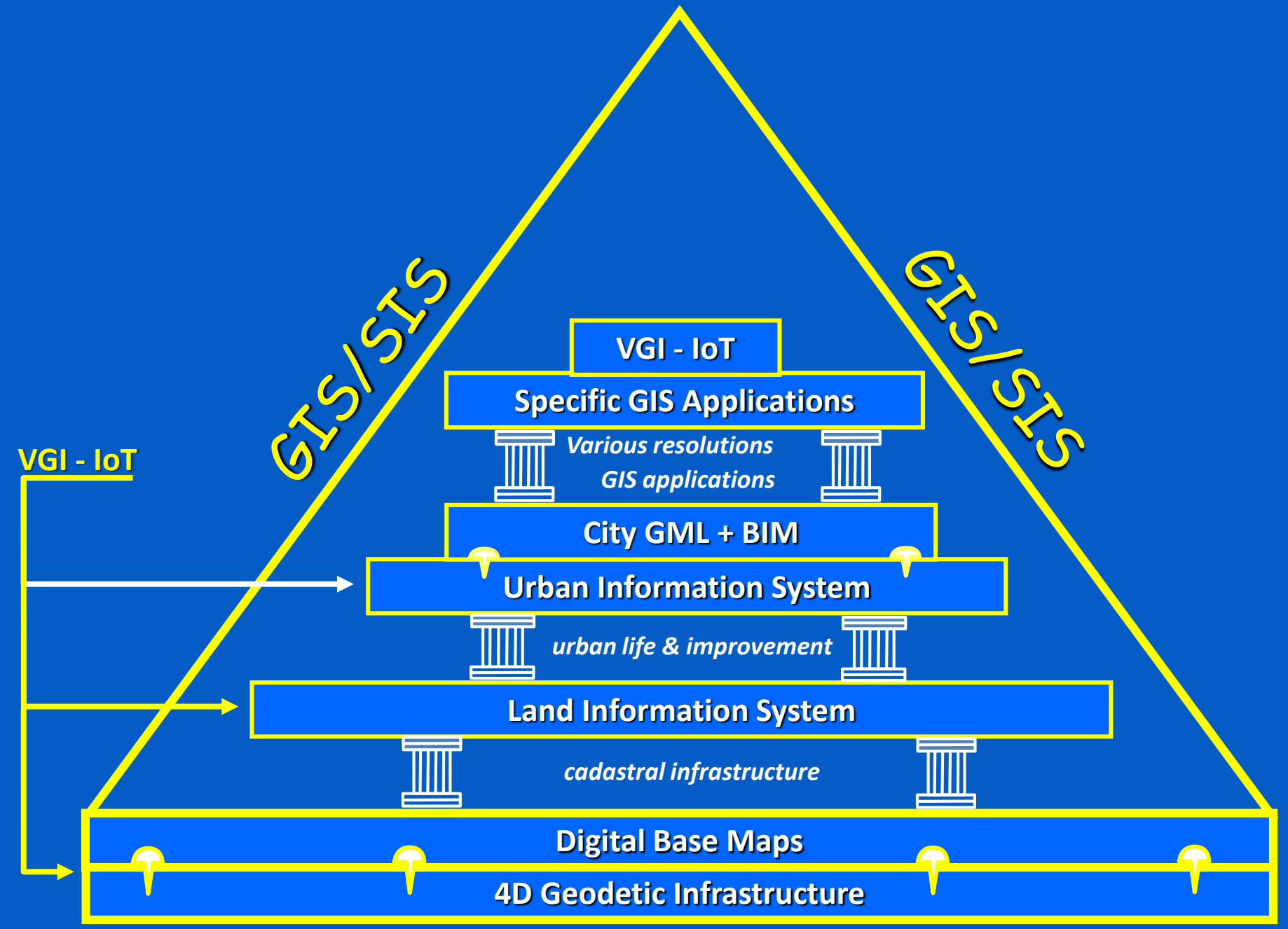
*drons*




JPL


LEGAL
LEGAL
ETHICS
RULES
INTEGRITY
INTEGRATION


ARMUSEMENT
AUGMENTED REALITY


GIS/SIS
GIS/SIS
VGI - IoT
VGI - IoT
Specific GIS Applications
Various resolutions
GIS applications
City GML + BIM
Urban Information System
urban life & improvement
Land Information System
cadastral infrastructure
Digital Base Maps
4D Geodetic Infrastructure

# OpenStreetMap
## (as a sample of spatial crowdsourcing)

**One of the most promising Volunteered Geographic Information (VGI) projects**

- No. of registered OSM members is – approximately 4.9 million
- No. of active members last month – close to 55,000
- No. of OSM Nodes in the database – near 4.4 billion
- Number of uploaded GPS points – more than 6.0 billion
- Nodes created daily – near 2.0 million
- No. of OSM Ways in the database – 480 million
- Ways created daily – approximately 250,000
- Accuracy: ~ 6.0 meter (2010)
  ~ 3.5 meter (2018)

# Thank You

**Prof. Dr. Yerach Doytsher**

**Mapping and Geo-Information Engineering, Technion, Israel**

*doytsher@technion.ac.il*

DigitalWorld 2018

The Third International Conference on Advances in Sensors, Actuators, Metering and Sensing

ALLSENSORS 2018

March 25, 2018 to March 29, 2018 - Rome, Italy

# Panel

## Quality of Data and Services: The Essences From Geo-computing, Society/Crowd Sensing, and Mobility/Service-related Sensing

### Agricultural data for geo-computing, sensors 'network, and advances in IoT: quality requested in the decision making processes

Paulo E. Cruvinel
paulo.cruvinel@embrapa.br

# Data Quality

The term "**data quality**" is used with reference to a set of characteristics that data should own:

## Accuracy & Currency

Addressed in different research areas, including statistics, management, computer science, geo-computing, sensors ‘network, and the use of IoT in decision making processes.

***data quality problem: how to define, measure and improve the quality of electronic data, stored in databases, data warehouses and legacy systems.***

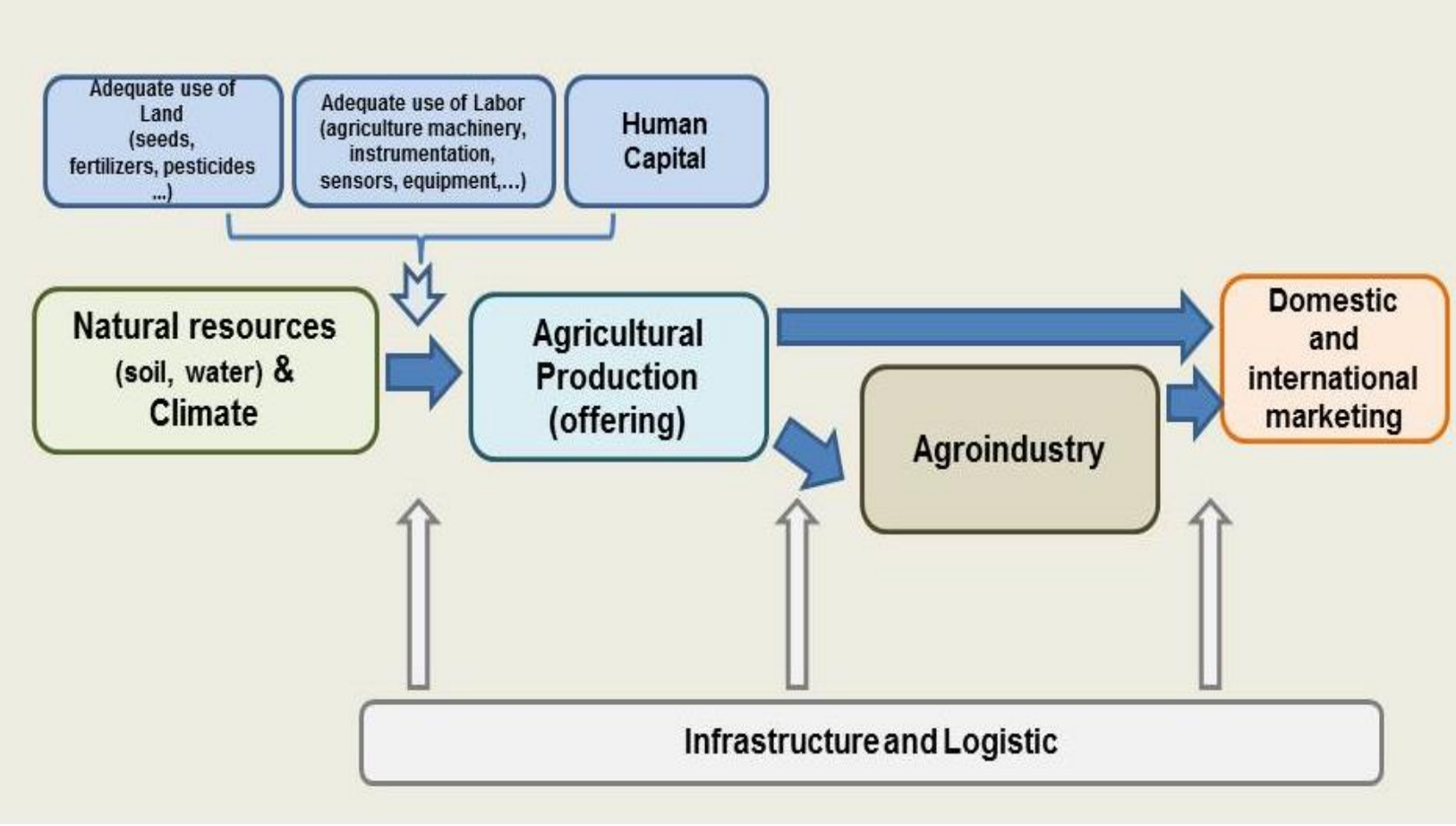
# Data Quality

Data quality is also related with metrology, physical concepts of the phenomena, practical aspects of measurements and analyses, knowledge, such as:

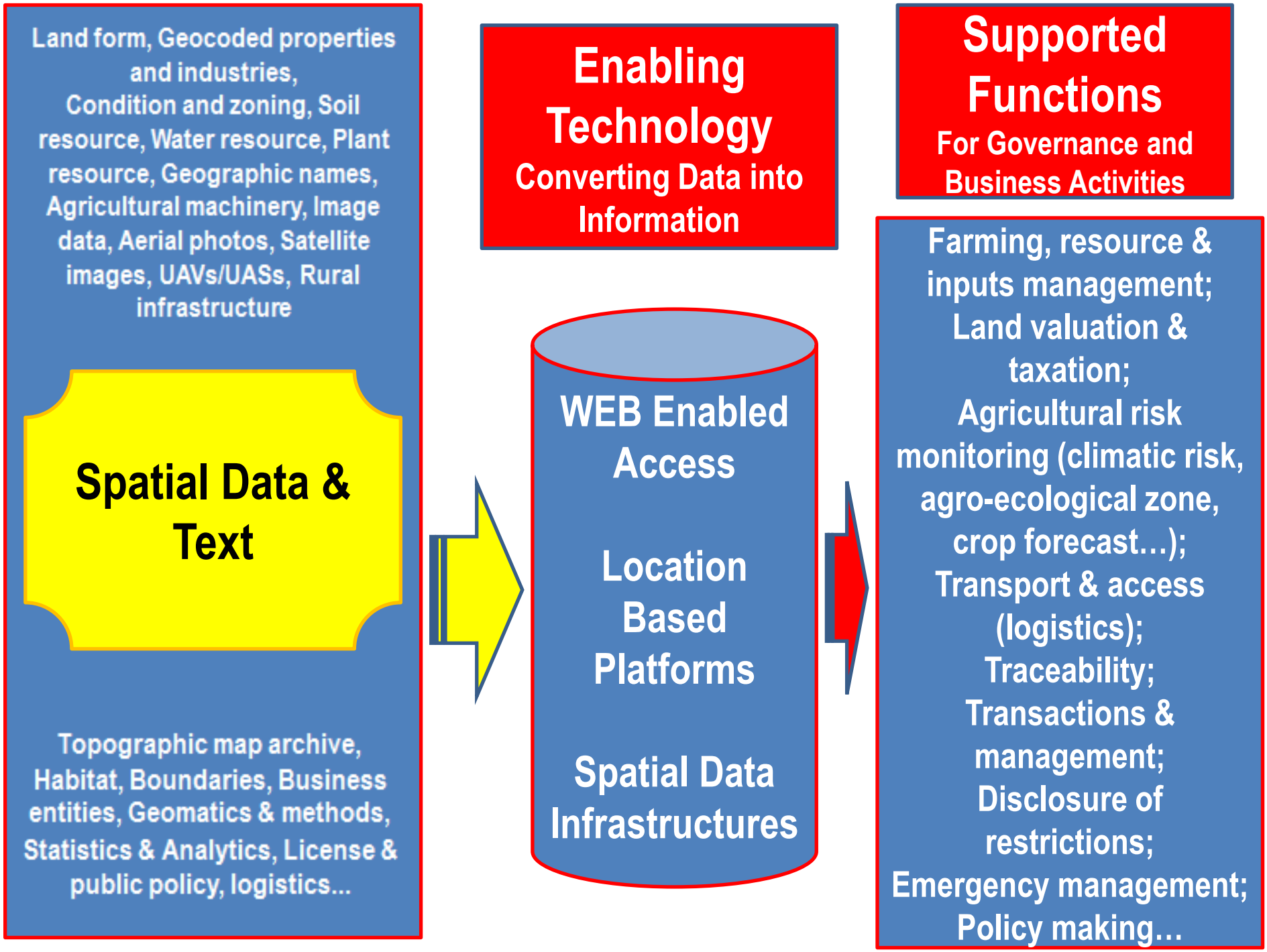
1. Units of measurement, standards, and methods;
2. Calibration, errors in measurements, periodical evaluation;
3. Infrastructure: Sensors, measuring instruments, networks and topologies, connectivity, computer architectures and their algorithms;
4. Accuracy and completeness in the decision models;
5. Inspection and techniques, i.e., the known “field truth” or “ground truth” (physical aspects);
6. Design, Manufacturing aspects (gauges of all kinds and data fusion);
7. Education and Training, among others.

# Agriculture and data quality requested in the decision making processes

Agriculture plays an important role for life and it is crucial to have data with a proper level of quality to support correctly the decision making process into the value chain.




Source: Flows in the agricultural value chain, Embrapa, 2013)


Land form, Geocoded properties and industries, Condition and zoning, Soil resource, Water resource, Plant resource, Geographic names, Agricultural machinery, Image data, Aerial photos, Satellite images, UAVs/UASs, Rural infrastructure
Spatial Data & Text
Topographic map archive, Habitat, Boundaries, Business entities, Geomatics & methods, Statistics & Analytics, License & public policy, logistics...
Enabling Technology
Converting Data into Information
WEB Enabled Access
Location Based Platforms
Spatial Data Infrastructures
Supported Functions
For Governance and Business Activities
Farming, resource & inputs management;
Land valuation & taxation;
Agricultural risk monitoring (climatic risk, agro-ecological zone, crop forecast…);
Transport & access (logistics);
Traceability;
Transactions & management;
Disclosure of restrictions;
Emergency management;
Policy making…

# Dimensions of Data Quality for decision making

- **Validity**: is a characteristic of measurement in which a tool actually measures what the researcher intended to measure;
- **Reliability**: is a characteristic of measurement concerned with consistency;
- **Timeliness**: is the relationship between the time of collection, collation and reporting to the relevance of the data for decision making processes;
- **Precision**: is the depth of knowledge encoded by the data. Precision comes in many forms such as the resolution of images, audio, and video and the degree of dis-aggregation of statistics;
- **Accuracy**: is the likelihood that the data reflect the truth;
- **Integrity**: is related with the ‘truthfulness’ of the data.

**Threats to Validity**: Definitional issues; Proxy measures; Inclusions /Exclusions; Data sources

## **Validity:** Questions to ask …

- Is there a relationship between the activity or program and what you are measuring?
- What is the data transcription process? Is there potential for error?
- Are steps being taken to limit transcription error (e.g., double keying of data for large surveys, built in validation checks, random checks)?

**Threats to Reliability:** Collection methodologies; Collection instruments; Personnel issues; Analysis and methodologies

**Reliability:**
Questions to ask …

- Is the same instrument used from year to year, site to site?
- Is the same data collection process used from year to year, site to site?
- Are there procedures in place to ensure that data are free of significant error and that bias is not introduced (e.g., instructions, indicator information sheets, training, etc.)?

**Threats to Timeliness**: Collection frequencies; Reporting frequencies; Time dependency

**Timeliness:**
Questions to ask …

- Are data available on a frequent enough basis to inform program management decisions?
- Is a regularized schedule of data collection in place to meet program management needs?
- Are data from within the reporting period of interest (i.e., are the data from a point in time after the intervention has begun)?
- Are the data reported as soon as possible after collection?

**Threats to Precision**: Source error / bias; Instrumentation error; Transcription error; Propagation of errors

**Precision:**
Questions to ask …

- Is the margin of error less than expected change being measured?
- Are the margins of error acceptable for program decision making?
- Have issues around precision been reported?
- Would an increase in the degree of accuracy be more costly than the increased value of the information?

**Threats to Integrity**: Corruption, intentional or unintentional; Personal manipulations; Technological failures; Lack of audit verification and validation

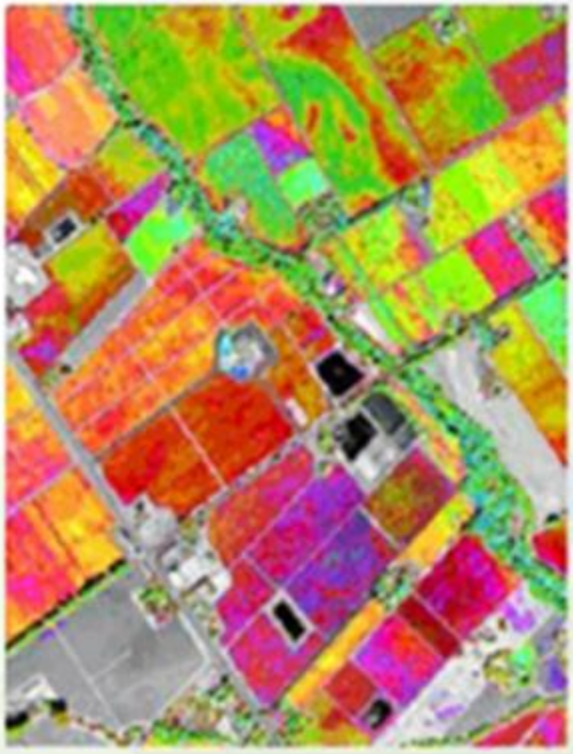
## A framework for agricultural data quality assessments in decision making

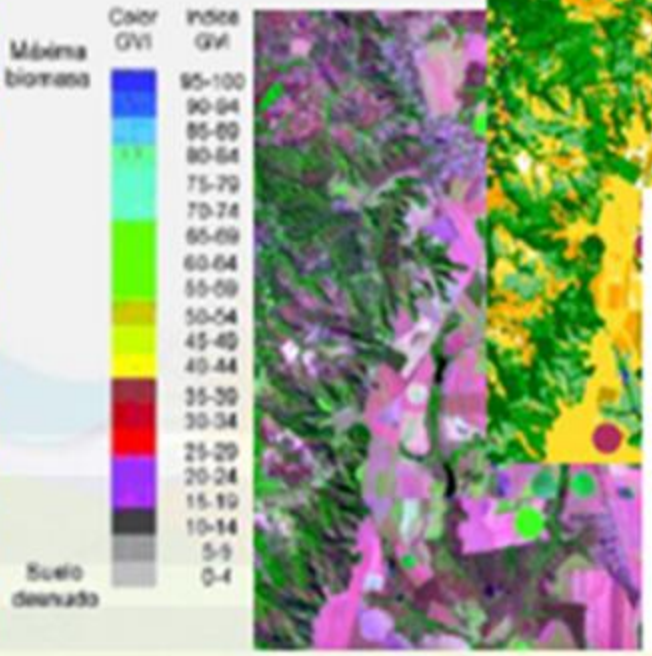
| **Data Management System** Processes / Procedures | **Data Quality System** Processes / Procedures | **Auditable System & Serviceability** | **Risk Verification** |
|---|---|---|---|
| Source | **Validity** | verification of the entire process, i.e., data acquisition, architectures involved, data processing, information retrieval, decision making models, among others | |
| Collection | **Reliability** | | |
| Collation | **Integrity** | | |
| Models and analyses | **Precision** | | |
| Reporting | **Timeliness** | | |
| Usage | | | |

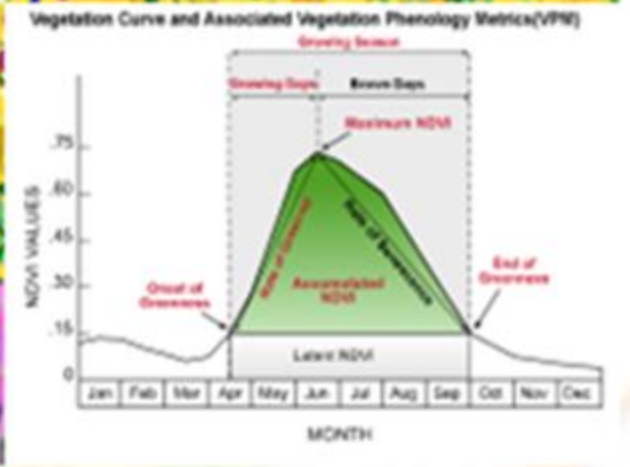
Risk Monitoring

Modeling

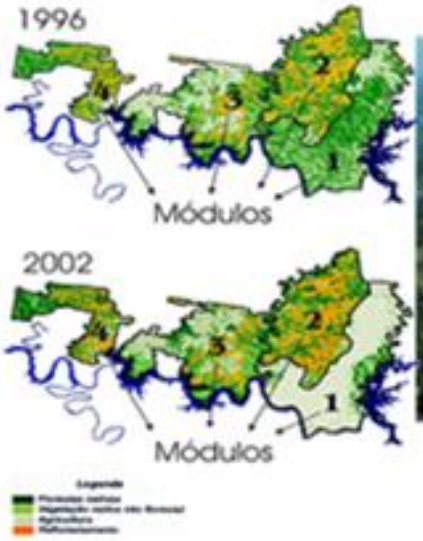
Plant, soil & environmental variables monitoring

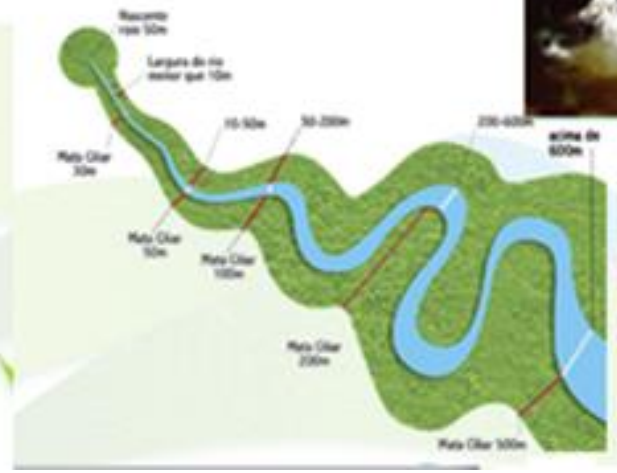




# Sustainability
## Modeling



# Logistical Risk Monitoring System

## Modeling of Variables

## Transportation and Storage







# Conclusion

There are many challenges in ongoing data quality related with processes for decision making in agriculture. Such processes are related with modeling and crop management, control of the inputs, quality control and assurance, automation, environmental analyses, storage, logistics, traceability, and when one is supplying information and products for the consumers. The approach used to tackle each one of these issues depends on the application scenario, as well as from the amount and the level of data with quality required for the intended use.

# Thanks for your attention!

Paulo E. Cruvinel
paulo.cruvinel@embrapa.br