EE 5104 - Data Science

- Today
  - Data Science : Introduction
  - Course Overview

- Turn your video on 

- Check out lecture materials posted in the syllabus

Up-to-date Syllabus: http://tiny.cc/y3wouz

Data Science: Introduction

Young Tae Noh

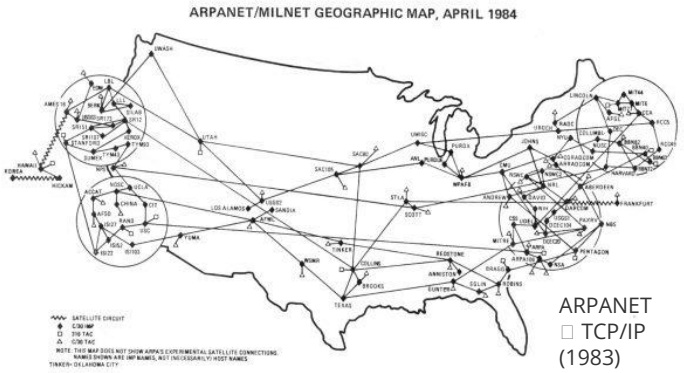
KENTECH

Syllabus URL: http://tiny.cc/y3wouz

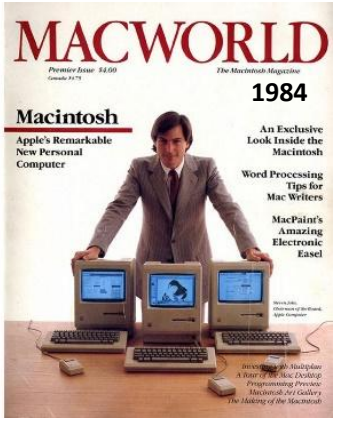
The majority of slides are borrowed from Prof. Uichin Lee @ KAIST – KSE80

Computing Trends in Late 1980s

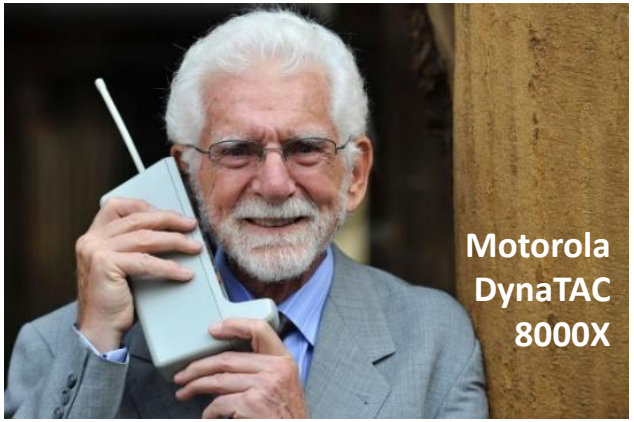
Mainframe computers



Before Internet was widespread



Macintosh had just come out



Cellphones were bulky & expensive

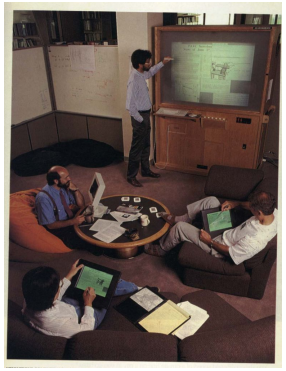
Computing Trends in Early 1990s

The Computer for the 21st Century

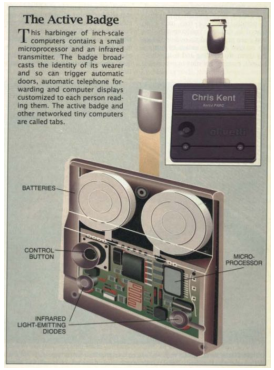
Specialized elements of hardware and software, connected by wires, radio waves and infrared, will be so ubiquitous that no one will notice their presence

by Mark Weiser

"Embedding computers into the physical worlds"

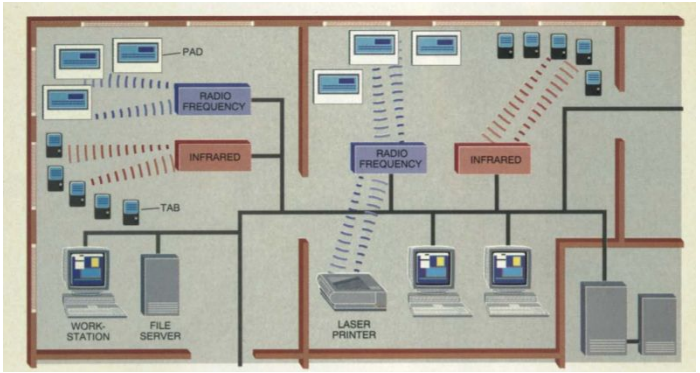


Tabs, Pads, Boards



Badges

Radio



Wireless Networking

Computing Trends in Early 2000s

COVER FEATURE

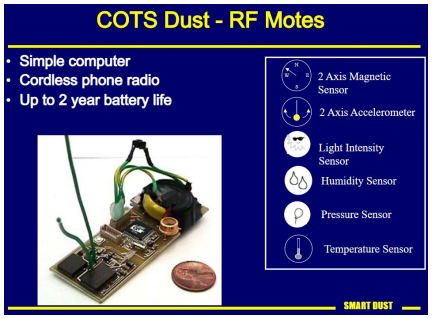
Brett Warneke
Matt Last
Brian Liebowitz
Kristofer S.J. Pister
University of California, Berkeley

Smart Dust: Communicating with a Cubic-Millimeter Computer

The Smart Dust project is probing microfabrication technology's limitations to determine whether an autonomous sensing, computing, and communication system can be packed into a cubic-millimeter mote to form the basis of integrated, massively distributed sensor networks.

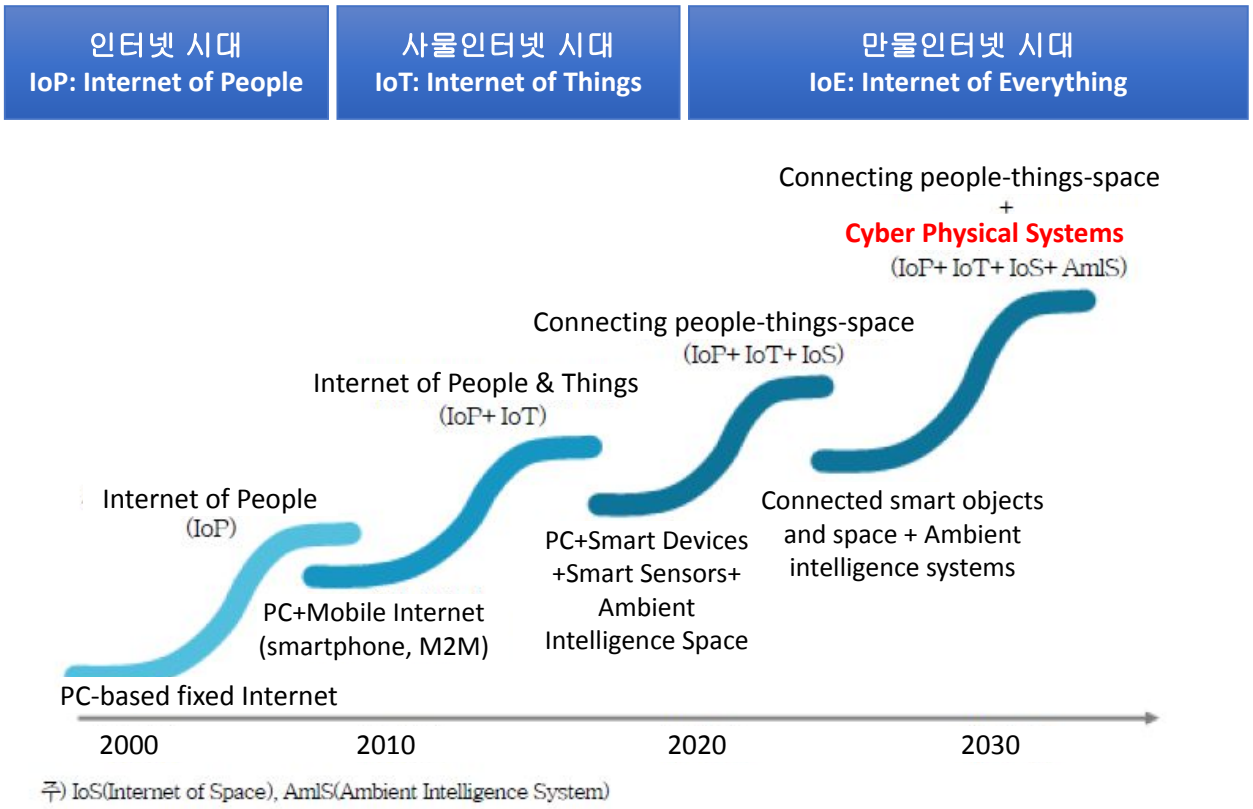
"Sensors everywhere + wireless networking"

TRSS Node | Crossbow | Ember | Sensoria | Dust, Inc.



https://slideplayer.com/slide/5328596/

What We Observe



만물지능인터넷 패러다임과 미래창조IT 신전략, 2013

And Now

1. Devices in diverse “form factors”
2. Sensors everywhere
3. Rich variety of inputs and outputs
4. All wirelessly connected
5. Data-driven intelligence

Computation, communication, and sensing
all integrated into physical world

1. Devices in diverse “form factors”

Smartphones

Activity Trackers

Withings, Fitbit, iHealth, Fitbug
Weight Scales

MonBaby
Baby
Monitor

Belkin WeMo
(smart switch)

Smart Bulb
(LiFX, Hue)

Nest Learning Thermostat
(NEST Labs)

2. Sensors everywhere

Quantified Self

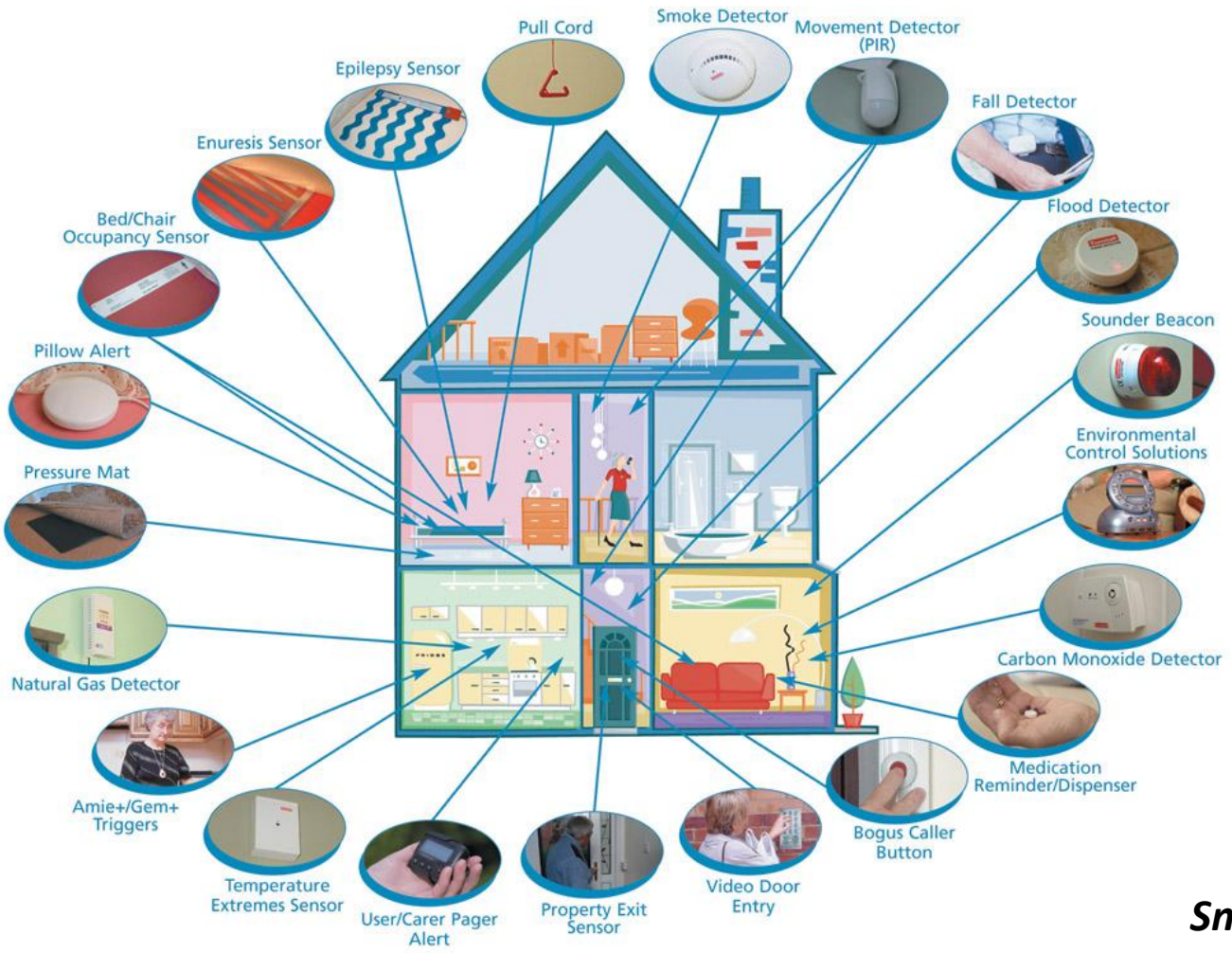
self knowledge through numbers

https://www.ted.com/talks/gary_wolf_the_quantified_self

Elon Musk says Neuralink will be like a ‘Fitbit in your skull with tiny wires’

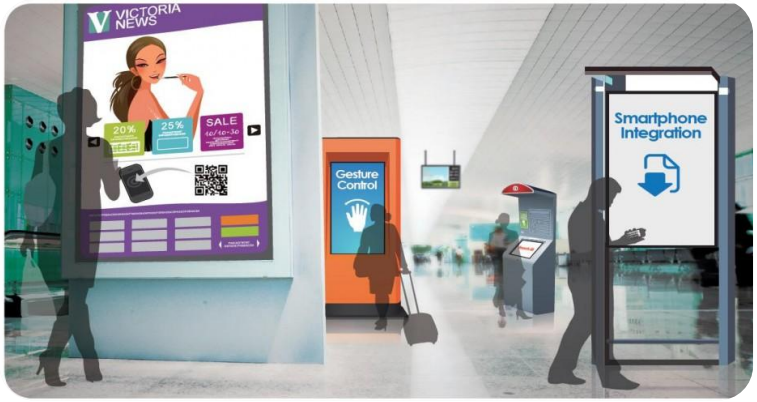
https://www.digitaltrends.com/news/neuralink-progress-update-2020/

2. Sensors everywhere

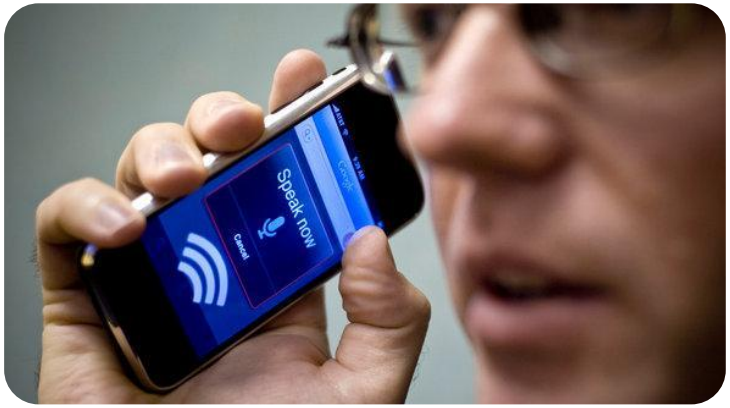


Smart home

3. Rich variety of inputs and outputs



Digital signage

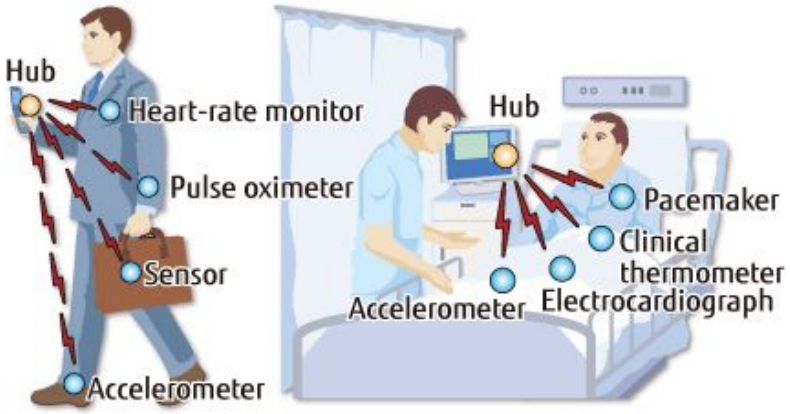


Voice command

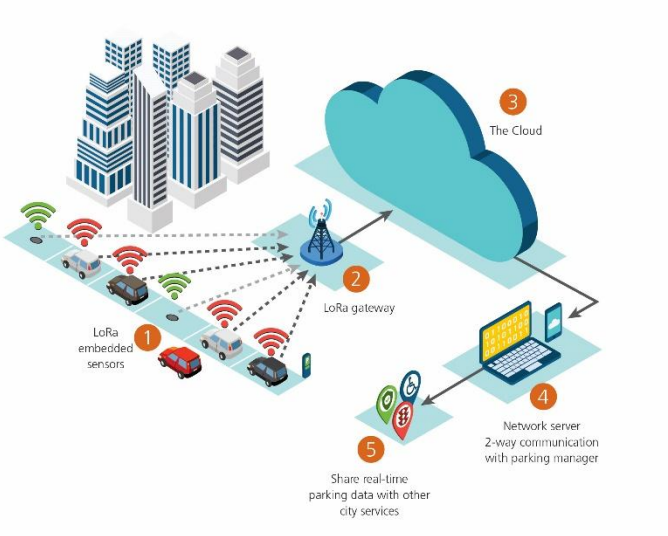
Physical exertion input

Gesture input

4. All wirelessly connected



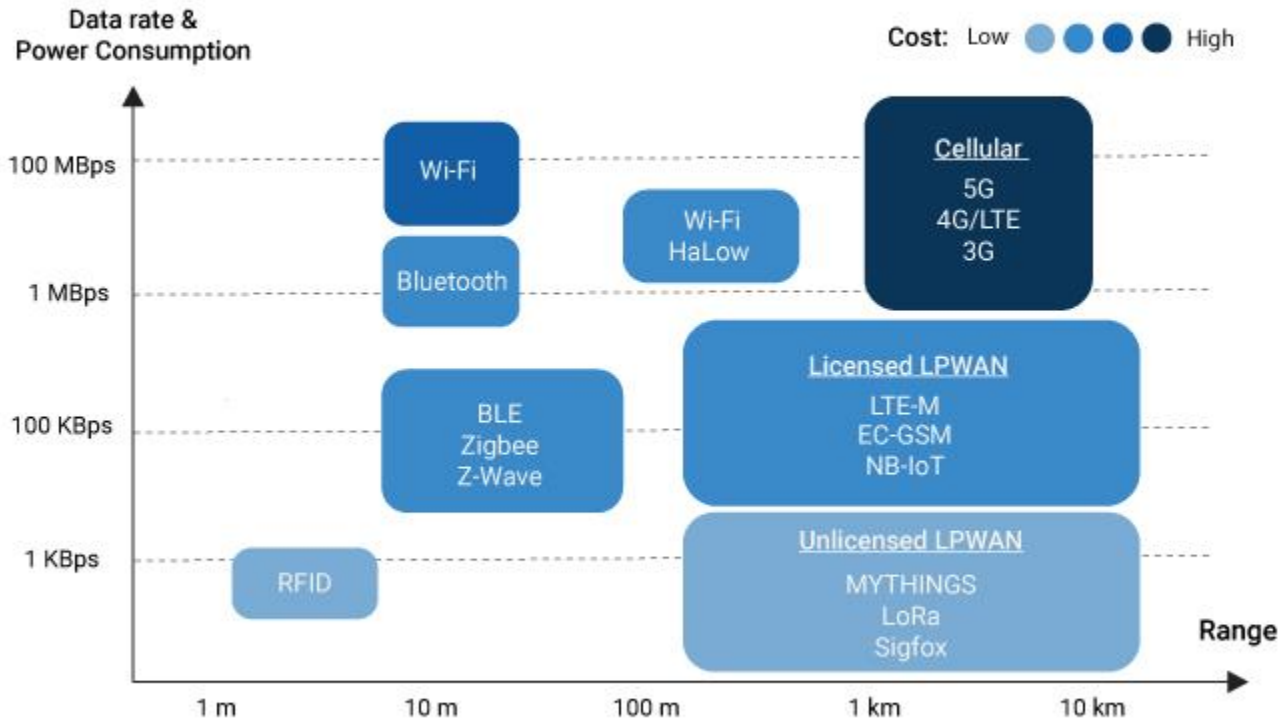
BAN: Body-area networks
(e.g., Zigbee, Bluetooth)



LPWAN: Low-power wide-area networks
(e.g., LoRa, NBIoT)

https://www.electronicsweekly.com/news/iot-smart-cities-the-long-range-forecast-for-wireless-connectivity-2018-08/

4. All wirelessly connected

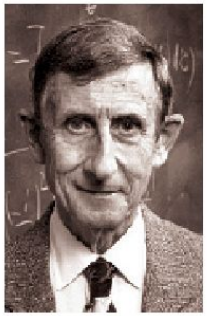


https://behrtech.com/blog/6-leading-types-of-iot-wireless-tech-and-their-best-use-cases/

5. Data-driven intelligence

Ubiquitous sensing and data analysis as a new tool

" New directions in science are launched by new tools much more often than by new concepts. The effect of a concept-driven revolution is to explain old things in new ways. The effect of a tool-driven revolution is to discover new things that have to be explained. "



Freeman Dyson (1997) *Imagined Worlds*

Harvard University Press, Cambridge, MA

Slide by Thomas Insel, 2015

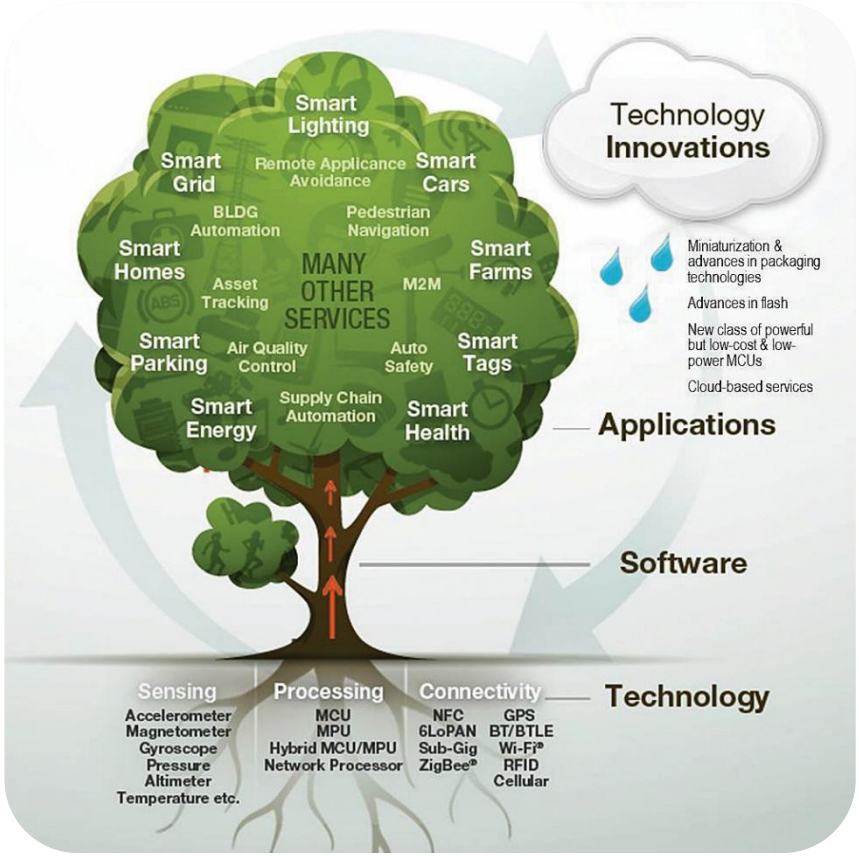
5. Data-driven intelligence

New opportunities for data-driven intelligence with sensor data analysis

https://www.youtube.com/watch?v=I2l3e1oNwUU

Why now?

- Ecosystem is ready

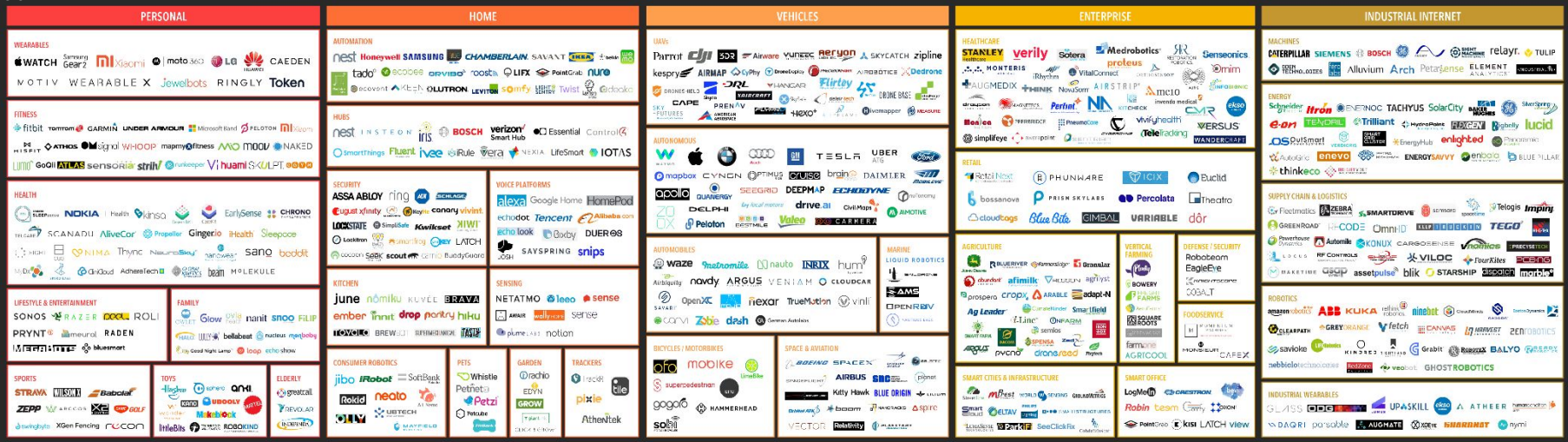


Underlying sensing, processing, and networking technology & data availability

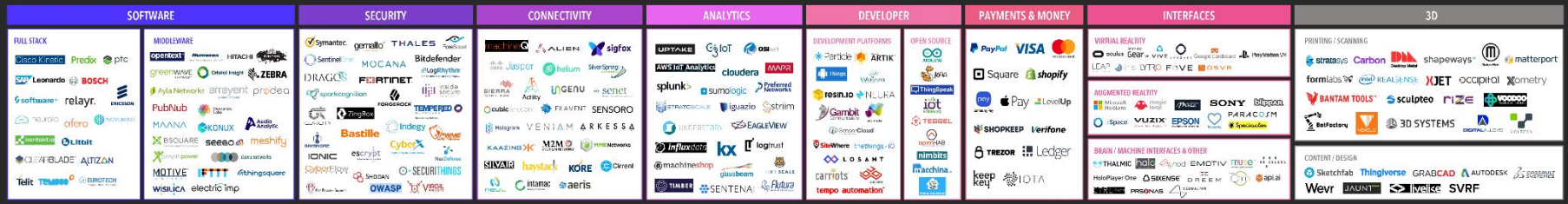
https://community.freescale.com/community/the-embedded-beat/blog/2013/03/19/how-do-you-promote-innovations-to-accelerate-the-roll-out-of-the-internet-of-things-iot

Internet of Things Landscape 2018

Applications (Verticals)



Platforms (Horizontals)



Building Blocks

Final version, revised and updated as of February 7, 2018

FIRSTMARK
EARLY STAGE VENTURE CAPITAL

IoT Applications

IoT Applications

And Now

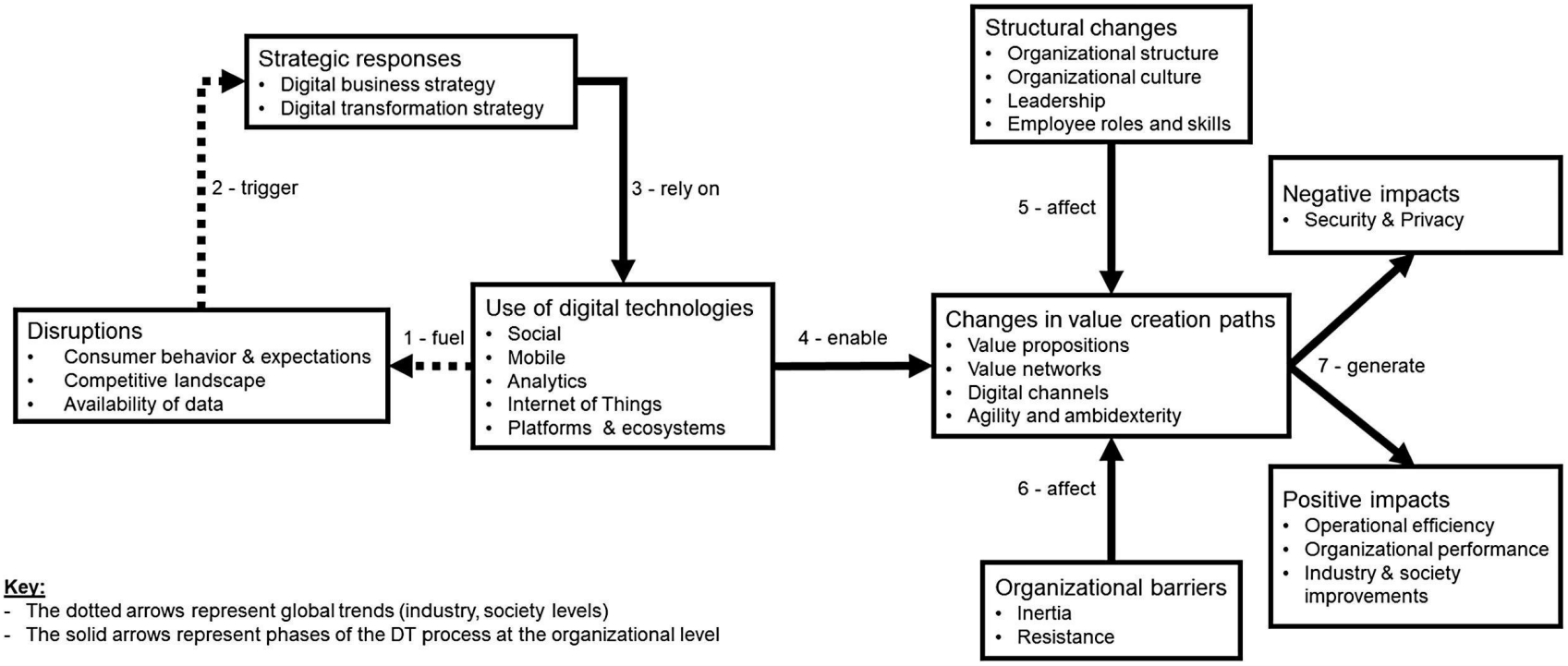
1. Devices in all form factors
2. Sensors everywhere
3. Rich variety of inputs and outputs
4. All wirelessly connected
5. Data-driven intelligence

Computation, communication, and sensing
all integrated into physical world

- Digital transformation & cyber physical systems

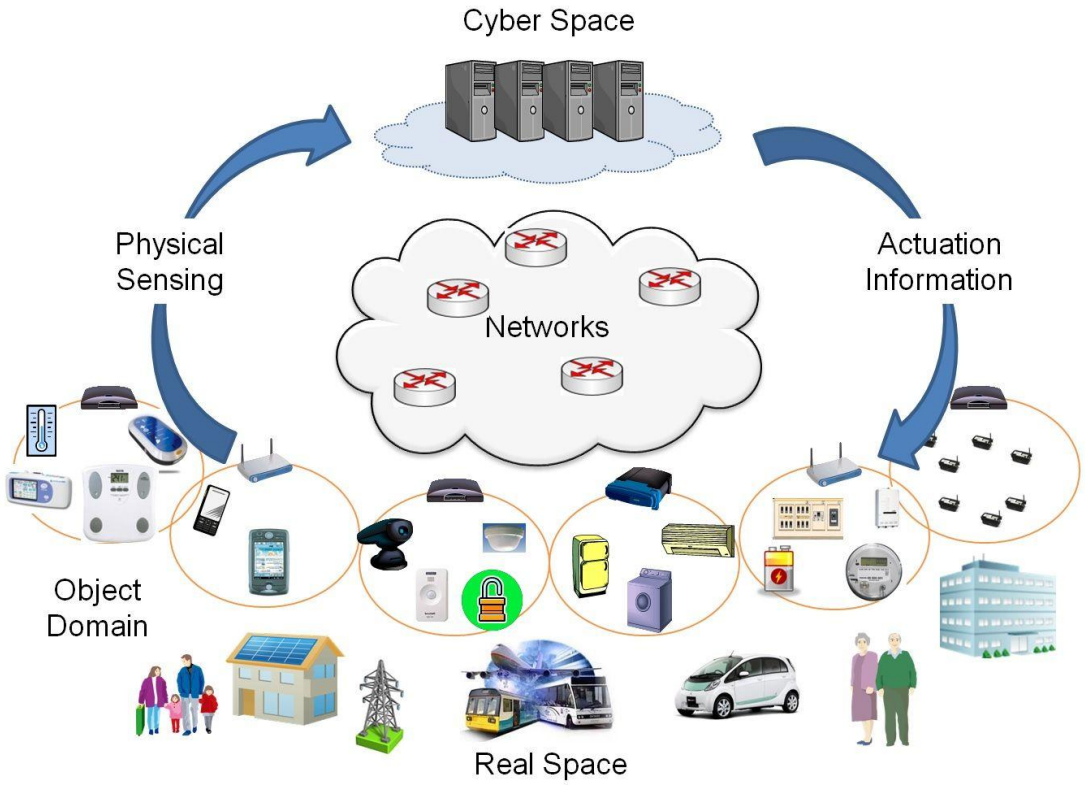
Digital Transformation

"a process that aims to improve an entity by triggering significant changes to its properties through combinations of information, computing, communication, and connectivity technologies" (Vial 2019)



Vial 2019: Understanding digital transformation: A review and a research agenda https://doi.org/10.1016/j.jsis.2019.01.003

Cyber Physical Systems



1. Devices in all form factors
2. Sensors everywhere
3. Rich variety of inputs and outputs
4. All wirelessly connected
5. Data-driven intelligence

Transformative technologies for managing interconnected systems between its physical assets and computational capabilities

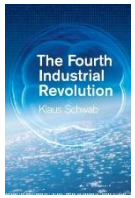
A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems, Jay Lee, Behrad Bagheri, Hung-An Kao Manufacturing Letters 3 (2015) 18–23

4th Industrial Revolution Led by **Cyber-physical-systems**

Navigating the next industrial revolution

WORLD ECONOMIC FORUM
COMMITTED TO IMPROVING THE STATE OF THE WORLD

Revolution	Year	Information
1	1784	Steam, water, mechanical production equipment
2	1870	Division of labour, electricity, mass production
3	1969	Electronics, IT, automated production
4	?	Cyber-physical systems



The Fourth Industrial Revolution, Klaus Schwab, Founder and Executive Chairman World Economic Forum 2016
The Fourth Industrial Revolution: what it means, how to respond, Klaus Schwab, January 14, 2016

Velocity, scope, systems impacts are great!

- Velocity: the speed of current breakthroughs has no historical precedent; exponential growth
- Scope: it is disrupting almost every industry in every country
- Systems impact: the breadth and depth of these changes herald the transformation of entire systems of production, management, and governance

The Fourth Industrial Revolution: what it means, how to respond, Klaus Schwab, Founder and Executive Chairman, World Economic Forum, January 14, 2016

Hype Cycle for Emerging Technologies 2020

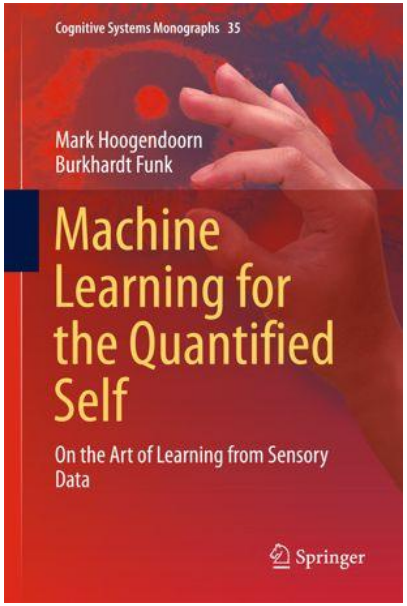
Sensor Data Science: Logistics

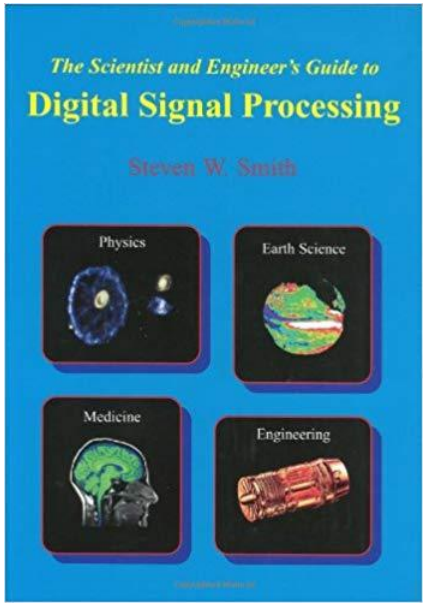
Syllabus URL: http://tiny.cc/y3wouz

Goals of this course

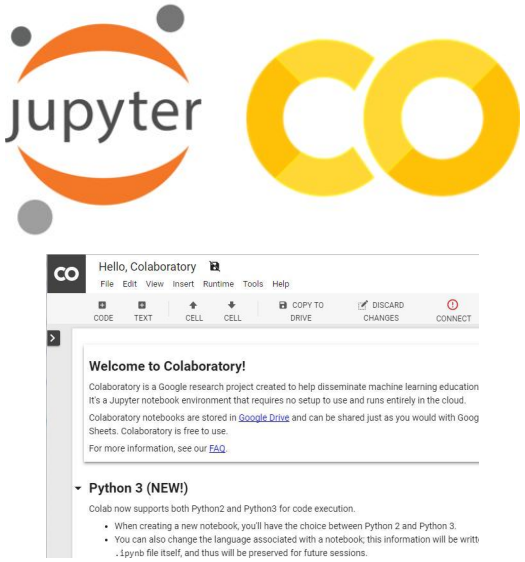
- Understand the key concepts of mobile sensor data science that constitute the recent innovations
- Improve complex problem solving skills by learning technical tools and applying those tools to solve real-world sensor data science problems
- Help advance the state-of-the-art sensor data science technologies

Topics in this course





Learning



In-class Practice

Basics

- Introduction & Overview (W1)
- Sensor Data Analysis Process (W1)
- Contextual & Persuasive Computing (W2)
- Sensor Data Visualization (W2)

Sensor Data Science Pipeline

Data Collection & Datasets (W3)	Data Pre-processing (W4)	Feature Engineering (W6, 7)	Supervised Learning (W9, 10, 11)
K-EmoCon K-EmoPhone	Noise, missing values Anomaly, reduction	Digital signal proc. Feature extraction	Traditional learning Deep models Evaluation techniques

Advanced Topics

- Active & Interactive ML (W12)
- Interpretable ML (W13)
- Causal Analysis (W14)

Applications

- Emotion & Personality (W5)
- Activity & Machine Monitoring (W14)
- Interruptibility & Productivity (W15)

Communications Tools

LMS @ KENTECH

- Homework
- Project

Piazza
(sign-up link)

- Q&A
- Critiques

Q&A must be done through Piazza!
(please, refrain from using emails)

Teaching Assistants

- Kevin (Head TA)

Logistics

- Lecture materials:
  - Google Doc Syllabus: http://tiny.cc/y3wouz
- In-class discussion:
  - Discussing the topics of the week (e.g., based on summaries/critiques)
  - Group activities
  - Programming/prototyping exercises

Logistics

- Grading policy
  - Participation: 10%
  - Mid-term: 30%
  - Homework: 30%
  - Project/Final exam: 30%

Logistics: Participation (10%)

- Class Participation (in-class & piazza)
  - Online class attendance
  - Ask or answer questions
  - Help your classmates (e.g., answering questions or tutoring)

  - Lecture note annotation & improvement
    - Provide additional explanation & leave comments to the slides
    - Goal is to improve current lecture materials for 2022 Class!
    - Starting week 2, TA(s) will share assignments

Logistics: Homework (30%)

- Homework
  - Mostly based on in-class co-lab practices
  - Homework is given to make sure that you run the code and can modify the code (help to check whether you really understood what you have learned)
  - You must do your homework to do the term project because you will use the same code for your term project
  - Homework will take less than a few hours (goal is to check that you correctly understood what you learned/practice)
  - Must submit within a week
  - There will be 10 assignments (3% per assignment)

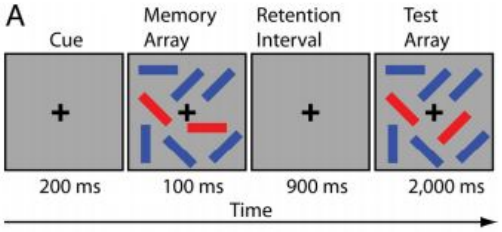
Logistics: Term-Project (30%)

- Must use K-EmoCon or K-EmoPhone dataset (or both)
- Possible term-project topics will be released by September, but you can propose your own topics
- Possible directions:
  - Replicating existing research papers by using K-EmoCon or K-EmoPhone
  - Improving existing research papers by applying recent techniques
  - Solving new problems by using K-EmoCon or K-EmoPhone
  - Deepening our understanding of K-EmoCon and K-EmoPhone datasets

Multitasking and Learning

- Distraction & divided attention: texting while driving?

- Multitaskers are inferior in cognitive control: filtering irrelevant info and task switching (Ophir et al., 2009)



- Lower scores (w/ laptop use) of the user and nearby users (secondhand smoking) (Cepeda et al. 2013)

Personal Technology Use Guidelines "Appropriate Use"

- When students bring personal devices such as laptops, tablets, and smartphones to the classroom or class lab, they should **appropriately** use them
- Communication devices will need to be on **silent** during the lecture
- Using devices for interacting on sites such as Facebook and Instagram, or playing online games is **not an appropriate in-class use** of technology
- Sending or receiving text messages, instant messages, or making or receiving phone calls in class can cause significant distractions to the teacher and fellow students
- Personal devices in the classroom are to be used for **class purposes only**
- Your **in-class participation scores** will be deducted if you use your devices inappropriately

Original statements from Bayless, Clipson, Wilson (2013). Faculty Perceptions and Policies of Students' Use of Personal Technology in the Classroom, Business Communication and Legal Studies