Visions of IoT
II. Machines

Fundamentals of Internet of Things

Andrea Omicini
andrea.omicini@unibo.it

Bologna Business School, Università di Bologna

Master in Digital Technology Management 2020/2021

Main Source

- most of the ideas here are extracted from [Chou, 2016]
- however, any mistake should be attributed to the author of these slides

Next in Line. . .

Not Just Humans

Focus on processes

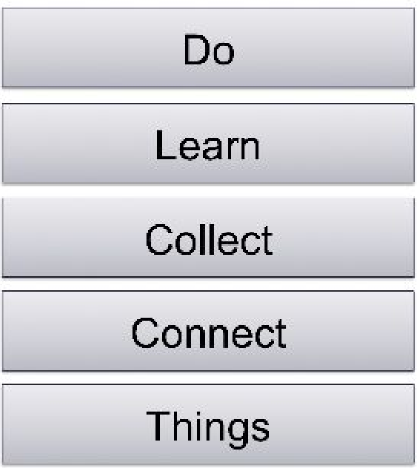
- IoT is not just about human using enchanted objects
- actually, it has the potential to change *manufacturing*, *agriculture*, and *services*
- changing *existing processes*
- making it possible to devise *new processes*

Precision

Exploiting knowledge for processes

- one of the key point here is *precision*
- we have much more knowledge on the processes
- in the *time* and *place* where we need them
- we have the potential to use all the knowledge about the process to vastly improve them—make them cheaper, more effective, faster, ...
- and, mostly, to make new processes possible

Five Layers of IoT



Five level for precision machines [Chou, 2016]

Next in Line. . .

Precision: Thing

- adding *sensors* to things make it possible to make them *aware* of the local surrounding environment—in real time
- *computing* and *storing* capabilities make it possible to take local instant decision about thing behaviour
- implemented via *actuators*
- thing behaviour can be *precisely* adapted to the local conditions in real time
    - e.g., no waste, no tries, as less mistakes as possible

Machine vs. Nomic Data

The process, the thing, the user

- *nomic data* tells us about the process in the specific application domain
    - in user centred processes, it may all us about the user(s)
- *machine data*, instead, give us understanding of the status of the thing itself
    - to make its behaviour the most effective and efficient, to prevent failures, . . .

Precision: Connect

- local information can be improved with data coming from nearby things (with sensors)
- local environment condition and change can be monitored in principle with much precision needed, for any parameter required
- by all things connected
    - e.g., getting information on blood pressure, heart rate, and body temperature altogether, so as to get the physical condition of a patient in real time, and possibly react accordingly
- machine data can be connected to provide for new features
    - e.g., to achieve fault tolerance, to increase accuracy and precision, . . .

Precision: Collect

- all local data from the same process (a production line, a patient monitoring, a racing car) are collected together and stored
- along with all other "locality" involved in a possibly articulated process
- to have not just an overall view in real time
- but also a view over time—*evolution* over time
  - e.g., 24h of a patient life, 500 miles of a car racing Indianapolis 500, one month of an assembly line in a manufacturing process, . . .
- machine data can be collected to *monitor* the evolution of the machinery controlling a process
  - so as e.g., to increase efficacy and efficiency, to save power, . . .

Precision: Learn

- all stored data from the same process are made available for analytics, machine learning, inference
- *new knowledge* can be deduced or induced—about the process, the machines, the users: their status, their evolution
    - e.g., data from a patient can be used to predict critical conditions, data from a production line can be combined with market forecast and warehouse status to govern production, ...
- machine data can be used to *predict* the evolution of the machinery controlling a process
    - so as e.g., to predict potential failures, to order new spare parts, to make technicians available in due time, ...

Precision: Do

- we can achieve real-time fine-grained control over the machines, the processes, the users
- recording the past, monitoring the present, predicting the future
- agricolture, manufacturing, services—all existing processes can be transformed by the *precision* of control
  - e.g., we can use IoT to deliver relevant information in real-time to workers in the course of their activity, or, even during their decision processes
- new processes can be invented
  - e.g., the SAM car
  - machine data can be used to provide assistance service on the machinery controlling a process
  - so as to improve availability, reliability, safety, maintainability, . . .

Next in Line. . .

Focus on. . .

Rio Tinto's Mine of the Future Initiative

History was made on June 14 2019. Rio Tinto completed the transition to entirely automated operation of its 1500km railway in the Pilbara region of Western Australia, becoming the first heavy-haul railway in the world to operate an automated network.

Described by Rio Tinto as the world's largest robot, AutoHaul is operating up to 50 automated and unmanned trains at any one time. Each 240-wagon, 2.4km-long consist, requires two to three locomotives, which haul 28,000 tonnes of iron-ore from the company's 16 mines to the ports of Dampier and Cape Lambert on an average 800km, 40-hour journey.

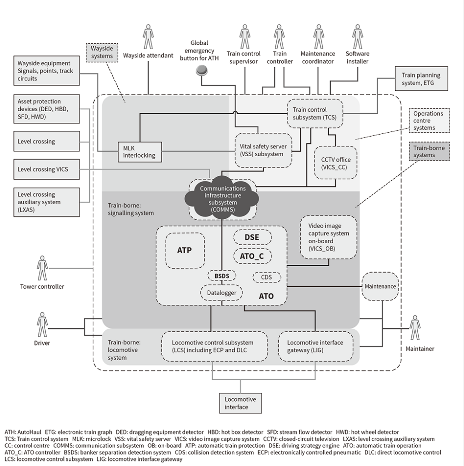
https://www.railjournal.com/freight/rio-tinto-completes-automation-of-pilbara-rail-network/

AutoHaul I

Overall view https://www.miningmonthly.com/tag/autohaul

- the achievement required international collaboration at the largest level
- Hitachi Rail STS (formerly Ansaldo STS) main developer of AutoHaul
- some core technology by NYAB

AutoHaul II



System structure of AutoHaul

https://www.hitachi.com/rev/archive/2020/r2020_06/

NYAB View

New York Air Brake (NYAB)

- NYAB is a division of Knorr-Bremse, a multi-billion-dollar firm based in Germany
  - a leading supplier to the railroad industry of braking systems and components, training simulators and train control systems for 125 years
- NYAB has been increasingly using information from Things to improve operational efficiency

AutoHaul: Thing

First thing for things: sensors

- data from the already-existing sensors made available
- data from new sensors that are added ad hoc
- sensors mostly sampled once per second, some five times, e.g.
  - brake cylinder PSI (pound per square inch) — the amount of pressure applied to the locomotive's brake cylinders
  - brake pipe PSI — the current PSI of the locomotive's brake pipe; compressed air is transmitted along the train to brake each of the cars
  - equalising reservoir PSI — a small reservoir used on locomotives to regulate the brake pipe
  - brake pipe airflow — a flow meter provides air mass flow rate into the brake pipe
  - current and voltage to traction motors
  - speed
  - engine RPMs
  - location

AutoHaul: Connect

Focus on locomotive

- most locomotives have a Communication Management Unit (CMU) that manages and optimises routes for data traffic
- the CMU contains a variety of data modems including WiFi, cellular and satellite.
  - it may even contain multiple network providers for cellular and satellite
- the CMU looks at the priority of the data and the cost of transmission and determines the best physical/data-link layer method to use
- the CMU creates a VPN connection with the railroad's back-office-network system so all traffic goes through private WAN
- in addition to communications, the CMU contains an integrated GPS provisioned for Differential Global Positioning System
- data is pulled every five minutes by NYAB

AutoHaul: Collect

- train data is transmitted to NYAB's cloud where they use the same detailed physics models running onboard the locomotive to expand the compressed data to a time-series database
  - roughly 40TB per month per locomotive
- to process the data, NYAB uses Splunk, which provides real-time and historical search, as well as reports and statistical analysis
  - Splunk's scope includes data searching, filtering, modification, manipulation, insertion and deletion
- NYAB can run a query to retrieve all of the data for a particular trip or set of trips for a set of trains or a single train
  - Splunk maintains the archive
  - when they want to analyse the data, it is decompressed on demand and a large CSV file is generated, which is uploaded to bulk storage on Amazon's S3

AutoHaul: Learn

Off-line

- comparing the actual train operation to the idealised computer-generated plan
  - e.g., to monitor driver compliance and assist the train operator in how to reduce in-train forces, reduce fuel spend and improve time-to-destination
- monitoring the health of the locomotive
  - e.g., to monitor degradation of the horsepower rating of the locomotive, before it hits the *negative impact zone*
- NYAB uses both supervised and unsupervised machine learning (ML), e.g.
  - unsupervised ML is used to implement train-handling risk mitigation
  - supervised ML is used for fuel-efficiency analysis
- NYAB also uses data to implement predictive and prescriptive analytical analysis of brake shoe wear

AutoHaul: Do

Benefits

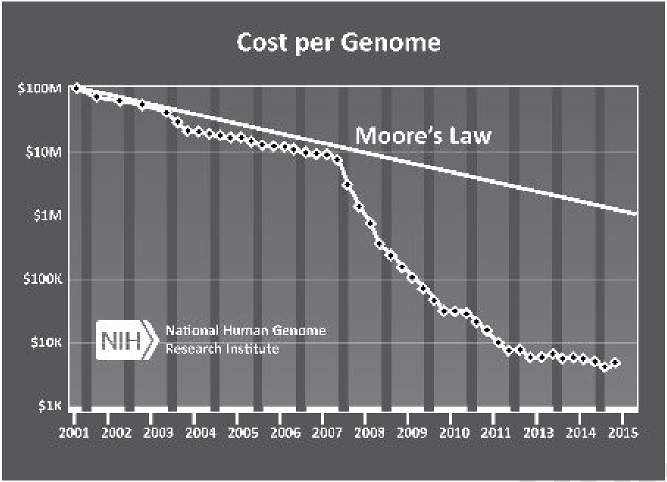
- fuel saving
- better and cheaper train maintenance
- saving on human labour—$300000 per year at Rio Tinto in 2010
  - unsupervised ML is used to implement train-handling risk mitigation
  - supervised ML is used for fuel-efficiency analysis
- better safety—e.g., human operators hit the alert button even when fully asleep

Focus on. . .

Gene Sequencers I

- *gene* or *DNA sequencer* is an instrument used to determine the sequence or order of the four bases from any plant or animal cell
  - bases are G (guanine), C (cytosine), A (adenine) and T (thymine)
- sequencers works by having a chemical reagent called fluorochromes mixed with the sample
- using lasers, light reflected from the samples yields the sequence information
- there are a variety of techniques used to create higher-volume and lower-cost sequencing
- in the last twenty years, the cost of a whole human sequence has dropped down by order of magnitudes
  - from $100M (2000) to $10M (2007), $1M (2008), $100K (2009), $10K (2011), ...
  - $100 per sequence would made cancer treatment affordable for many more people

Gene Sequencers II



Gene sequencing costs [Chou, 2016]

Gene Sequencers III

Use cases

- prenatal DNA testing
- cancer identification
- devising out genetic causes of rare diseases
- up to *personalised health care*

More on this at NIH "DNA Sequencing Fact Sheet"

https://www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Fact-Sheet

The Case of Illumina's MiniSeq I

- prenatal DNA testing
- amniocentesis is commonly used, yet it has non-negligible risks of complications
- ! the baby's genetic material can be found in the mother's blood sample
  - which can be tested with no risks for both baby and mother
- this is the case study of the Illumina MiniSeq

  https://www.illumina.com/systems/sequencing-platforms/miniseq.html

The Case of Illumina's MiniSeq II

The Illumina's MiniSeq System

https://www.illumina.com/systems/sequencing-platforms/miniseq.html

MiniSeq: Things I

Sensors

- built with many subcomponents that often have embedded sensors, so that the actual number of sensors is unknown
- nevertheless, the 40 sensors Illumina does record include
  - 1 CCD camera
  - 2 RFID readers, 2 sensors each
  - 1 photodiode
  - 5 motor encoders
  - 1 valve position sensor
  - 6 fan-speed sensors
  - 6 temperature probes
  - 1 fluidic pressure sensor
  - 15 digital-state sensors

MiniSeq: Things II

Data

- key-value pairs with time stamp
  - either periodical data collection or sequencing data
- more than two hundreds discrete metrics logged

MiniSeq: Connect

Network & security

- Illumina's gene sequencers were engineered to connect to the Internet and the Illumina cloud with Ethernet and a conventional TCP/IP stack running on the Intel single-board computer
- data is transferred with RESTful services over HTTP from the machines
  - customers can use their own intranet for genomic data transfers in order to isolate the machine from the network
- AES (Advanced Encryption Standard) used to *encrypt* both machine and genomic data as it is written to local storage
- data to the cloud is secured with OAuth v1.0 authentication and classic SSL transport encryption

  https://tools.ietf.org/html/rfc5849

MiniSeq: Collect

Cloud & warehouse

- flat files on the sequencer
- stored in the cloud on Amazon's S3 storage service
- once in the cloud, machine data can be processed by Illumina with ETL (extraction, transformation, loading)
- machine data are loaded into Amazon RedShift data warehouse for predictive maintenance (and the like) of the instrument
- separate software tools by Illumina deal with (ge)nomic data, possibly locally
  - dealing with issues such as patient privacy and medical-data security

MiniSeq: Learn

Quality score

- looking for *sequencing quality score*—that is, the right base is called (in)correctly
- analysis targets potential underperformance by the equipment
- machine learning and analytic engines are part of the learning stage

MiniSeq: Do

Lab services

- mostly, services to the lab activities
  - onsite service people
  - preventive maintenance and part substitution
  - availability of reagents

Focus on. . .

The Story of Sam Schmidt

- Sam Schmidt started as a successful business man
- started racing, was good—RoY 1995, rising star, win from the pole in Las Vegas 1999
- ! terrible crash in 2000, left him *tetraplegic*
- no way he could drive a car again
- yet,
  "Sam Schmidt Will Get The First Semi-Autonomous Driver's License" *(Autoweek, 26 September 2016)*
  `http://autoweek.com/article/car-news/sam-schmidt-will-get-first-semi-autonomous-drivers-license`
- ? what happened?

Long Story Short

- in 2013, Schmidt met with a group of engineers at Arrow Electronics Inc.
- together they decided to try and enable Sam to *drive again*
- by 2014, Schmidt was driving all on his own on the racetrack at more than 100 mph

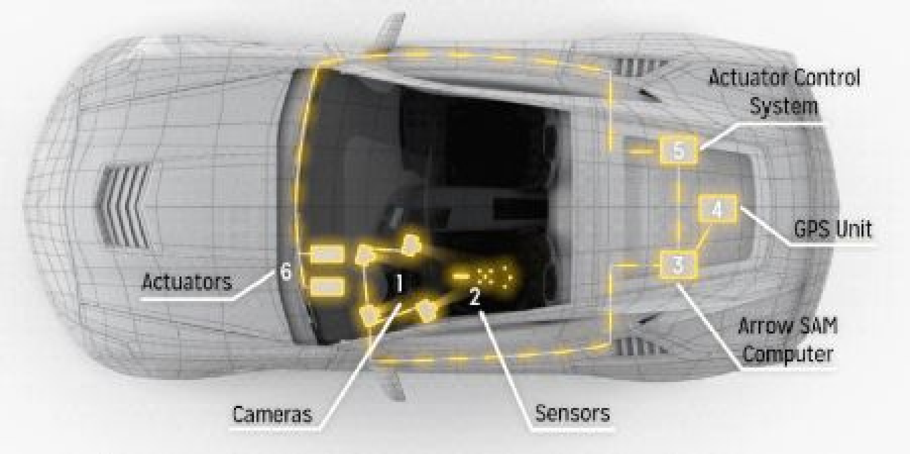
Before We Proceed

Questions

Thus is somehow a sort of a more familiar application scenario, in some sense: so we can start by trying to answer to some preliminary questions

- what scenario do we envision?
- which are the issues we can identify?
- what is the starting point for the solution?
- which *things* can constitute the solution?
- at which level(s) do we expect they should work?
- ...

Sam Car: Things I



SAM car architecture [Chou, 2016]

Sam Car: Things II

SAM car

- a number of cameras and sensors
- GPS unit for car's physical location
- actuator control system
- a computer as a gateway

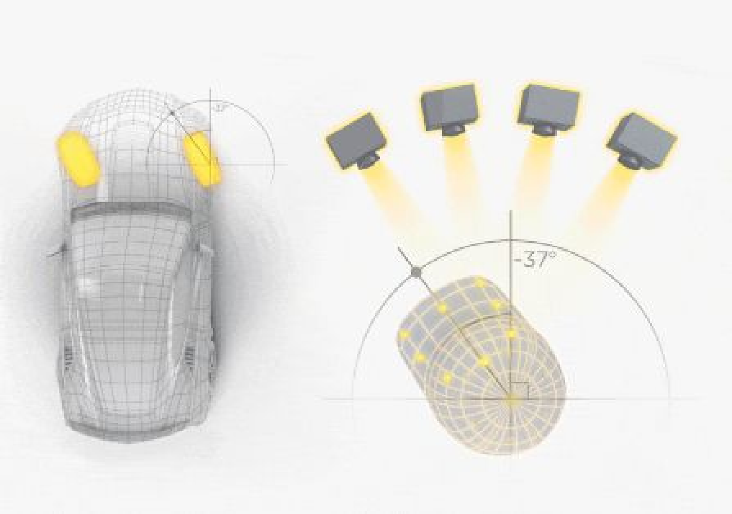
Car sensors

- data from both the car and the driver

car
: as a moving object (performance, speed, handling, ...), as an environment (brightness, temperature, humidity, ...)

driver
: heart rate, body temperature, ...

Sam Car: Things III



Driver's helmet [Chou, 2016]

Sam Car: Things IV

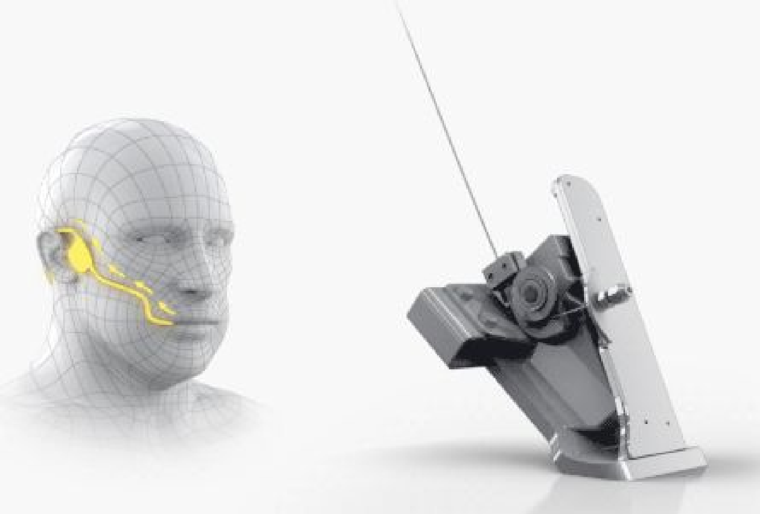
Helmet sensors

- eight reflective, infrared sensors fitted into the driver's helmet
- four infrared cameras face the driver

! to steer, the driver has just to look in the direction he/she means to go

- the driver's head angle and movements are traced in real time—via the SAM computer

Sam Car: Things V



Driver's mouthpiece [Chou, 2016]

Sam Car: Things VI

Accelerating and braking via mouthpiece

- *acceleration* via breath in the driver's mouthpiece
  - sensed and measured in *intensity* via a mouth-pressure sensor
  - the stronger the driver's breath, the stronger the acceleration
- *braking* via sipping through the same mouth-pressure sensor (negative pressure)
- a microprocessor transmit signals in nanoseconds to brake and gas pedal, and to a rotary actuator for the steering wheel

Sam Car: Connect

- through a GPS unit in the car, SAM computer connects over a 4G LTE modem to the Arrow Connect Cloud
- transmits telemetry, car's sensor and video data
- sensor data in real time, video when fast connection available
- e.g., telemetry for car safety, warns driver in case of danger

Sam Car: Collect

- Microsoft Azure Cloud service
- MongoDB as the database
- millions of records

Thing of Things, System of Things

- what is the car, here?
- what does the helmet represent?
- what about the overall system?
- including the GPS, and the cloud?

Sam Car: Learn

- not really deep here
- API to visualise, integrate, and analyse data from the SAM car
- available to third parties

Sam Car: Do I

- it happened
- in 2016, practice laps at the Indianapolis 500
- up to 152 mph

Sam Car: Do II

Sam Schmidt and the SAM car [Chou, 2016]

Next in Line. . .

The Point—Again

IoT changes the processes

- by providing more data and more knowledge on the processes, the machines, the users
- existing processes can be made more precise—more efficient, effective, . . .
- new processes can be devised out
- in every sort of economy—agriculture, manufacturing, services
- suitably exploiting both nomic and machine data

Visions of IoT
II. Machines

Fundamentals of Internet of Things

Andrea Omicini
andrea.omicini@unibo.it

Bologna Business School, Università di Bologna

Master in Digital Technology Management 2020/2021

References

[Chou, 2016] Chou, T. (2016).
Precision: Principles, Practices and Solutions for the Internet of Things.
CrowdStory Publishing.