# DIDS

*Haphazard reality of working ideas in progress*

DRAFT

# DIDS

## TABLE OF CONTENTS

# DIDS

Distributed Data-Informed mobile *ad hoc* Decision Support System-as-a-Service

*These are old ideas which may not have been feasible for pragmatic implementation and adoption at the dawn of the 21st century. Many principles and concepts are aggregated under the DIDS umbrella. Perhaps few (for example, FIDSE) will gain momentum in view of the urgency with which we must re-visit tools and technologies to address the growing demand in domains related to FEWSH (food, energy, water, sanitation, and healthcare). It is extremely difficult to make a difference in FEWSH at the grass-roots level. The latter is one reason why we, academics, are taking cover under the "digital" umbrella and trying to contribute without the trials and tribulations of change on the ground. Hence our attempt to be creative in the digital space with the hope that digital transformation in the networked physical world may become an useful system and digital FEWSH (dFEWSH) may sufficiently diffuse, globally, where needed, to enable the practice of science, engineering and technology to serve society as a catalyst for progress and purveyor of civilization.*

# *Part I ◆ In Principle*

## Atoms to Bits Connectivity in Decision (ABCD) Support: Embracing the IoT (internet of things) metaphor in a cyberphysical system (CPS) design for public goods, FEWS (food, energy, water, sanitation) and public health

***Green papers, potential projects and edge applications: Concepts & tools to make sense of data from DAMS/DIDS to serve society***

**DAMS -** Digital Architecture & Mapping Sensor Ecosystems - https://dspace.mit.edu/handle/1721.1/123984
Big picture: DWSN (distributed wireless sensor network) plus MANET (mobile *ad hoc* networking). Data acquisition, visualization (mobile), analytics (decisions) in **sensor ecosystem** test-beds? Consider potential integration / inclusion of Michael Stonebreaker's ingres/postgres type DB extensions (eg Amazon Redshift).

**Starting Point:**
Sensor data acquisition selected by features based on theoretical understanding (science and engineering) and pragmatic applications. Think feature selection, feature extraction and automated feature engineering, see https://people.eecs.berkeley.edu/~dawnsong/papers/icdm-2016.pdf and https://ballet.github.io/

**Events / Data:**
Involves humans, animals and plants (CDC "OneHealth"). Enable visualization of real-time data (if/when possible) using mobile apps (tablets, laptops). Contextual data curation, data fusion and non-obvious relationship analysis (NORA) followed by innovative analytics. Feed output (information) at various levels and feedback to re-optimize processes, data collection, analytical tools/techniques (ML, knowledge graphs, digital twins). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.4446&rep=rep1&type=pdf

**Converge:**
PEAS as a guide to create connected architecture (linked data nodes) and design/determine software and emerging tools (802.11bf). PEAS in "P3" book chapter (see below in 'review'). Include concept of MANET (mobile *ad hoc* networks, US Army) using sensors embedded with tinyDB and tinyOS features (David Culler, Sanjay Sarma) which can wake-up to transmit when secure/authorized upload connection is available via mobile gateway on drones (think remote, DARPA dust, LEO satellites; review Intel motes and mesh).

**Broad spectrum of field applications**
Water, irrigation, agriculture, food supply chain, public health (routine ops as well as pandemics/epidemics). Also think: collection of bat droppings from remote areas; samples from Ebola River/Forest, Congo; open sewers in more than half of the world; phosphorous in algal blooms; soil content of nitrates and phosphates leaching into the water table or polluting fresh water resources, monitoring environmental pollution, etc.

**Why Pursue**
Applied research in systems engineering, sensor ecosystems and data analytics tool for users and industry (ag, meat industry). Monitoring to reduce waste through controlling spoilage (biogenic amine sensors, for example). Tracking/tracing Salmonella, ammonia in poultry, animal feed, slaughter-houses, meat markets and sewers. Follow CDC's "onehealth" perspective. CADS platform for fruits/crops, plant viruses and infectious agents. Consumer applications in affluent nations: digital use-by date, packaging digitally communicates shelf life, safe handling instructions, cooking temperatures, recipes via "proximity" handshake within Bluetooth or IR range from packaged goods/produce in grocery aisle/shelf, mobile spoilage alert from embedded sensors.

**Review:**
Ongoing - SENSEE - in PDF docs marked 00 and 01 - here - https://dspace.mit.edu/handle/1721.1/123984
Context for DAMS is outlined in "HIP" - find "HIP" pdf here - https://dspace.mit.edu/handle/1721.1/123984
Concept of PEAS in "P3" - please find "P3" pdf here - https://dspace.mit.edu/handle/1721.1/123984
Fig 1 and Appendix 1 in "ADD" as examples - "ADD" is here https://dspace.mit.edu/handle/1721.1/128017
Science as a Service to Society – please find **"PrEP"** pdf here - https://dspace.mit.edu/handle/1721.1/123984

**ADD**
Cybersecurity infrastructure as an integral part of architecture (not as an after-thought in an operating layer).

**HELP NEEDED WITH TECHNOLOGY in the CONTEXT OF DIDS ARCHITECTURE and SENSOR ECOSYSTEM**
Straw sketches for "DAMS" serves only as an idea guide – see https://dspace.mit.edu/handle/1721.1/123984

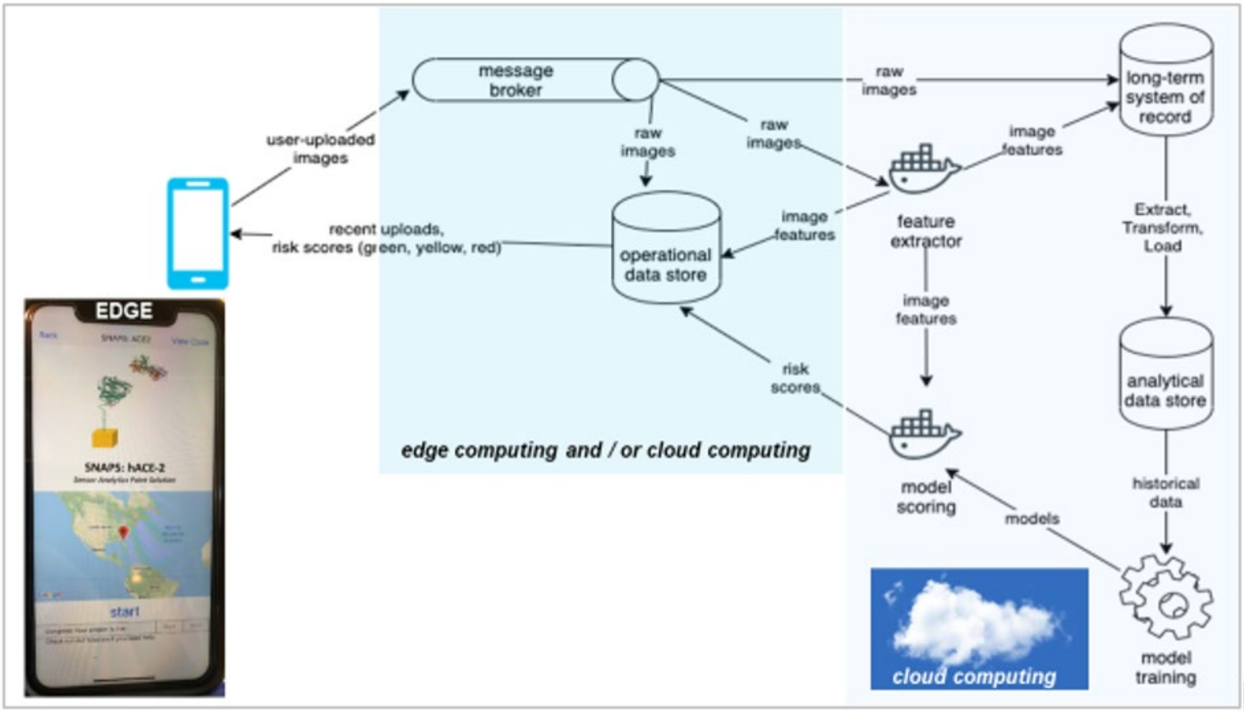
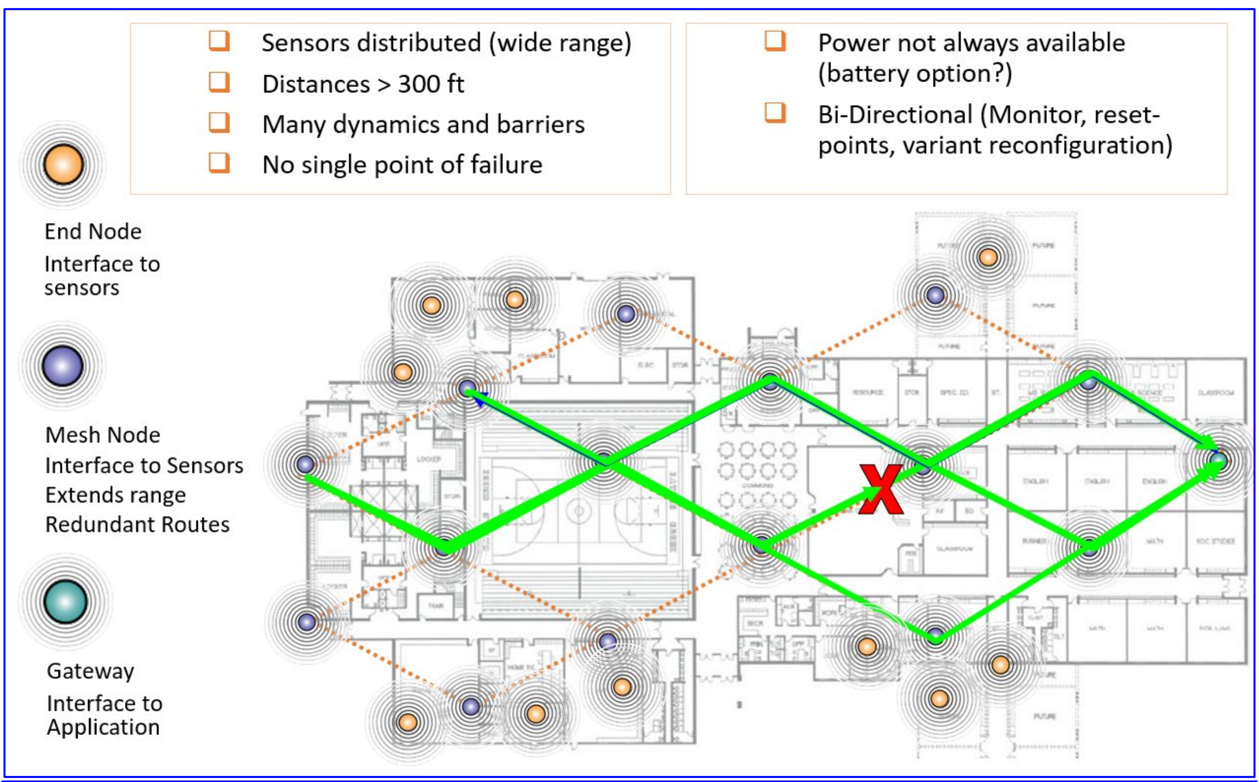
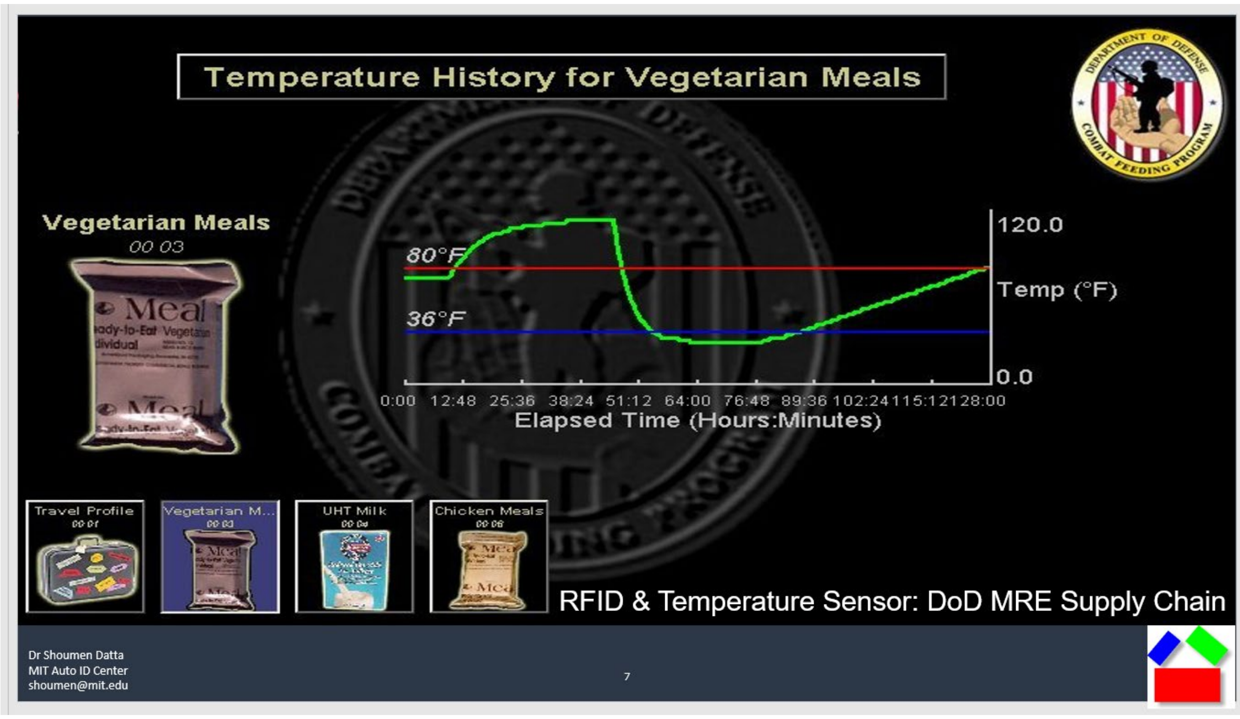
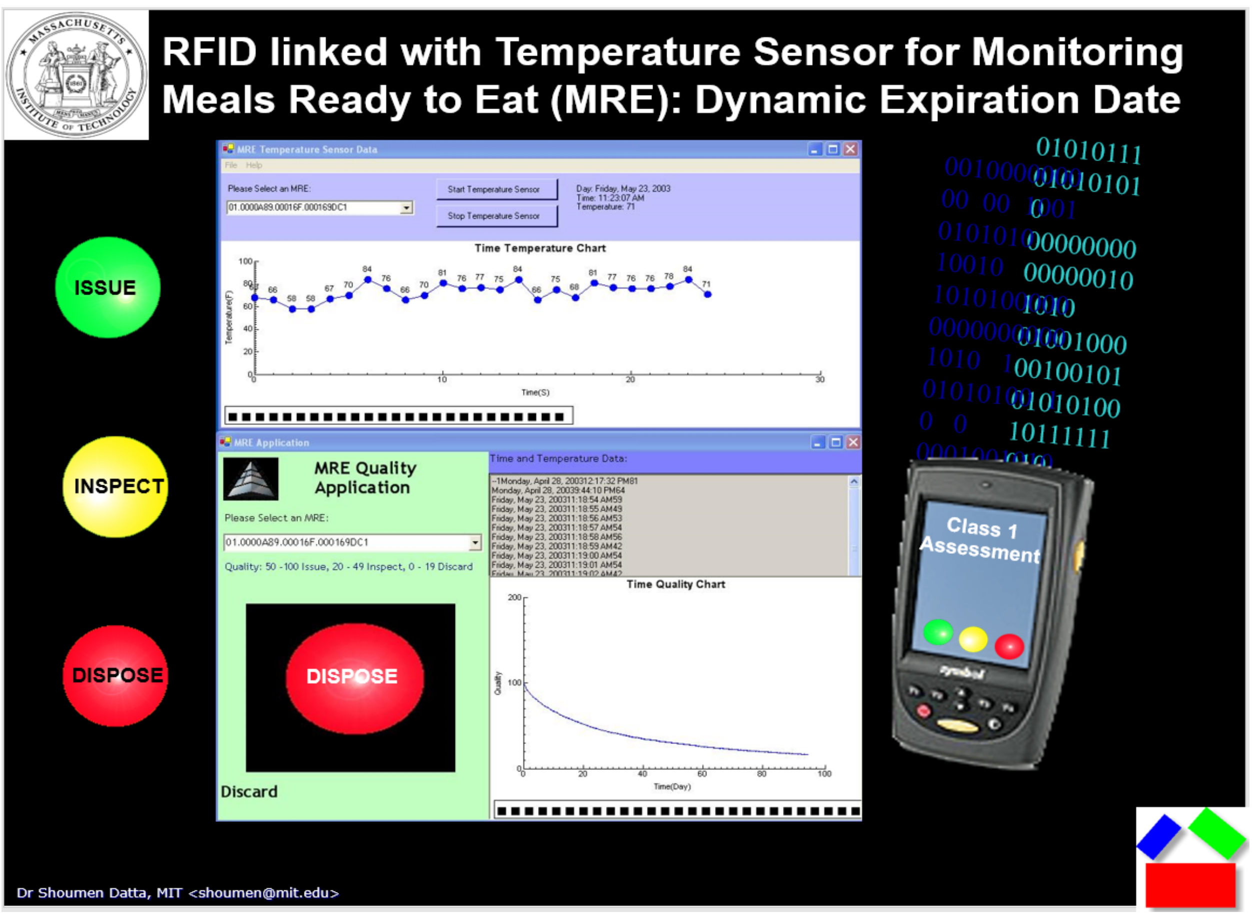



Figure 1 (also see Appendix 1) from "ADD" is my amateur and cherubic sketch of "software infrastructure"
Source: "ADD" PDF may be downloaded from MIT Library https://dspace.mit.edu/handle/1721.1/128017

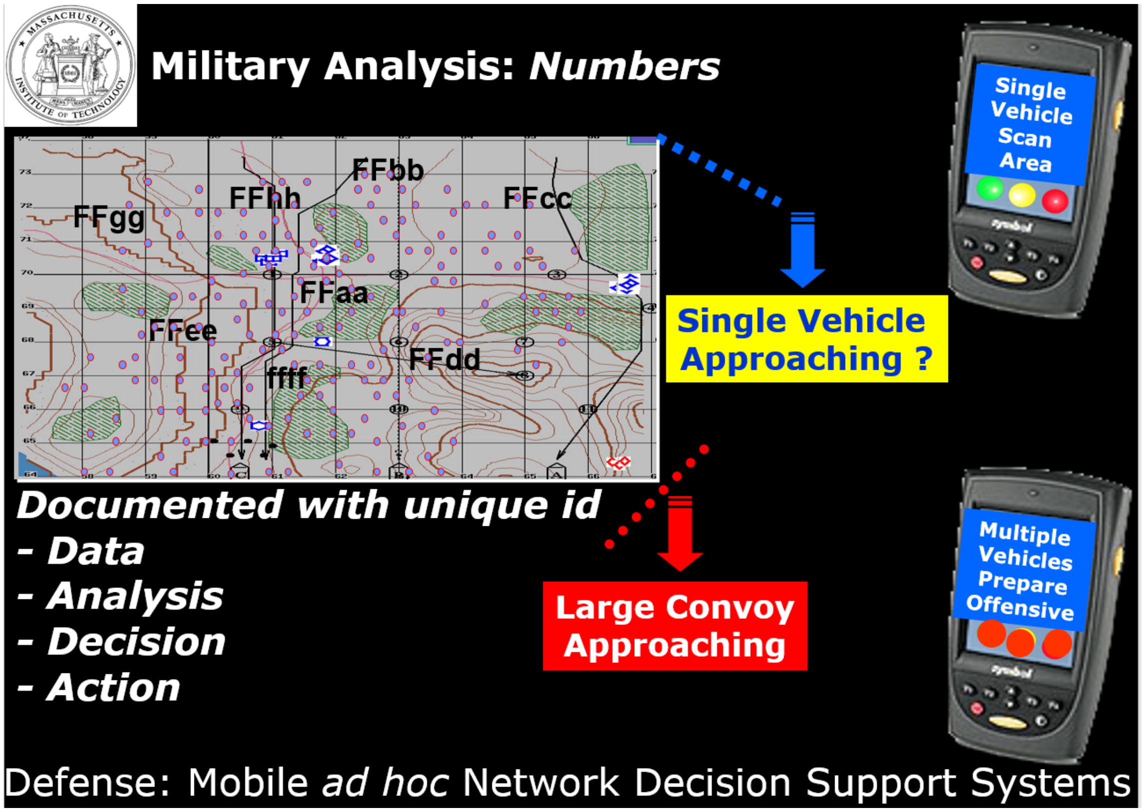


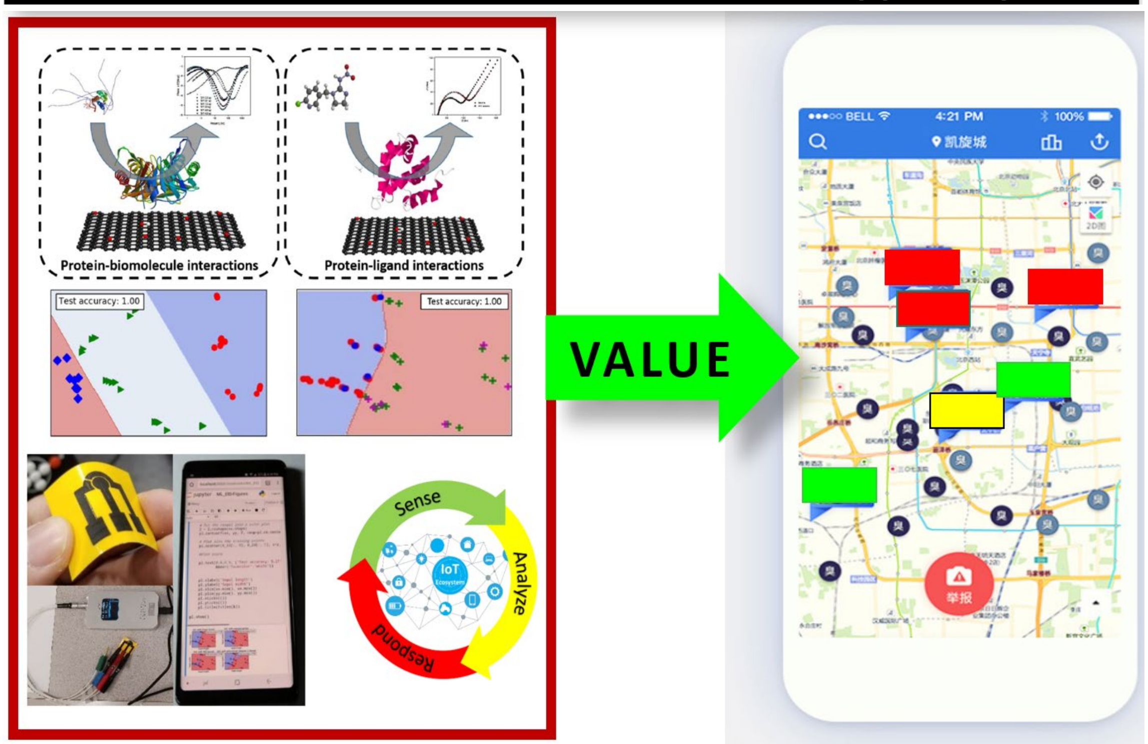

Non-expert (amateur) cartoon representing a distributed sensor network in hospitals, farms, meat packaging facilities, food manufacturing factories (locations where remote monitoring is essential for data acquisition). (Source: Shoumen Datta, MIT Auto-ID Center, MIT Forum for Supply Chain Innovation and teaching material: MIT Sloan School of Management. https://web.mit.edu/12.000/www/m2008/teams/awickert/motes.html)



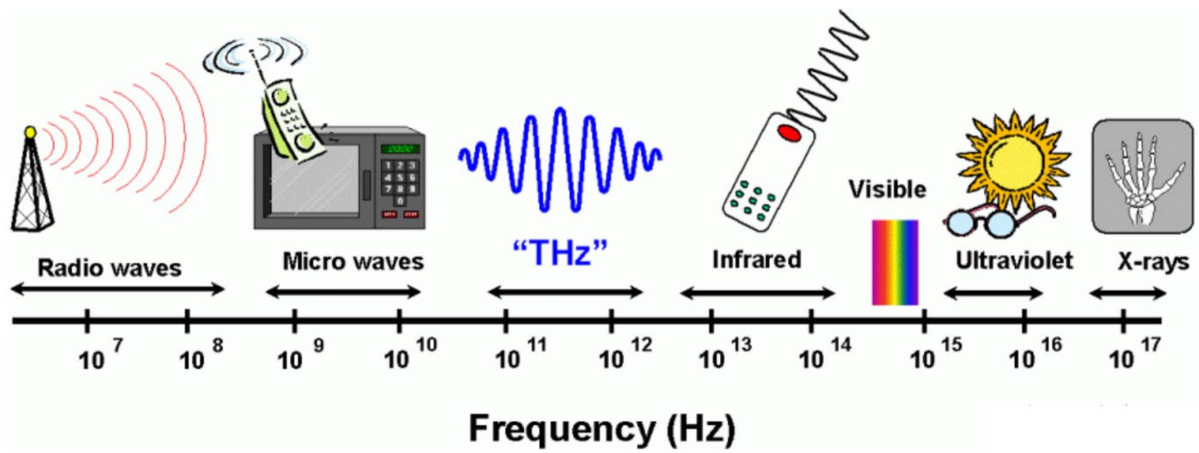

Example of data and analytics outcome (numbers) is visualized by non-experts on mobile devices in the form of very simple information: issue, inspect, dispose (answers). (Source: Shoumen Datta, MIT Auto-ID Center, MIT Forum for Supply Chain Innovation and teaching material used at the MIT Sloan School of Management.)

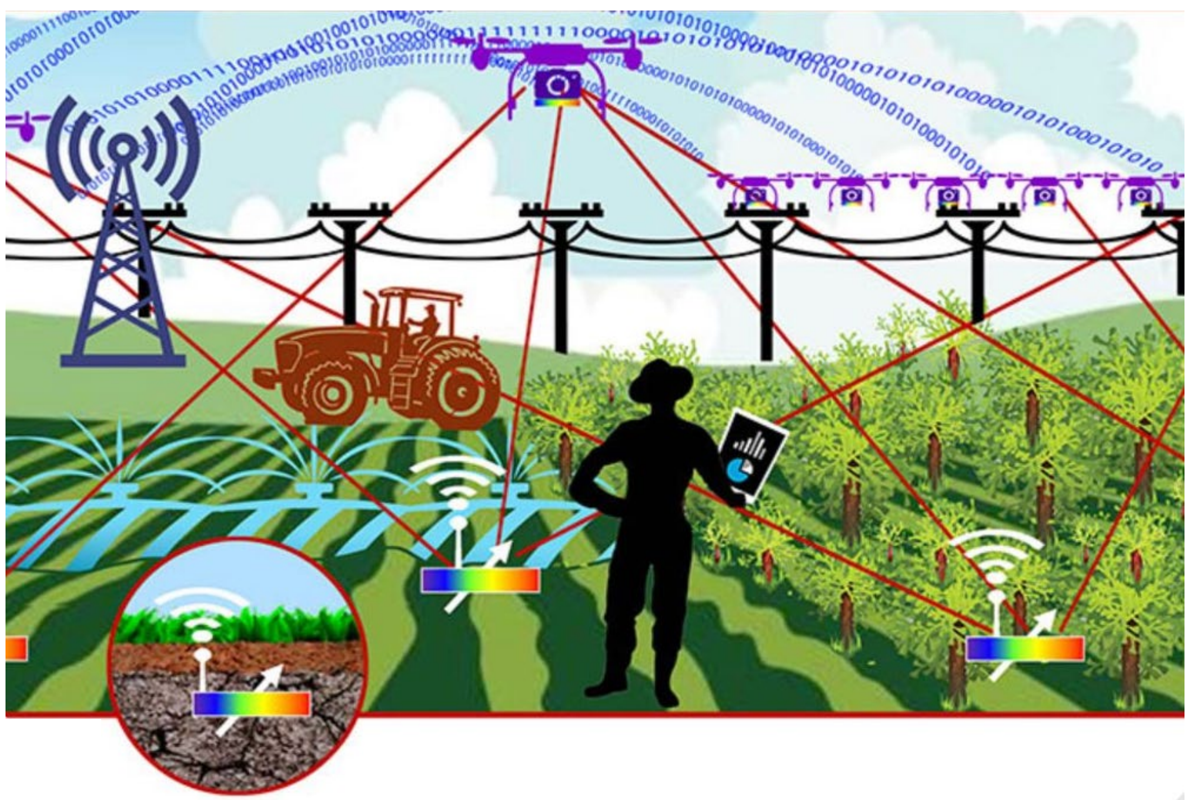


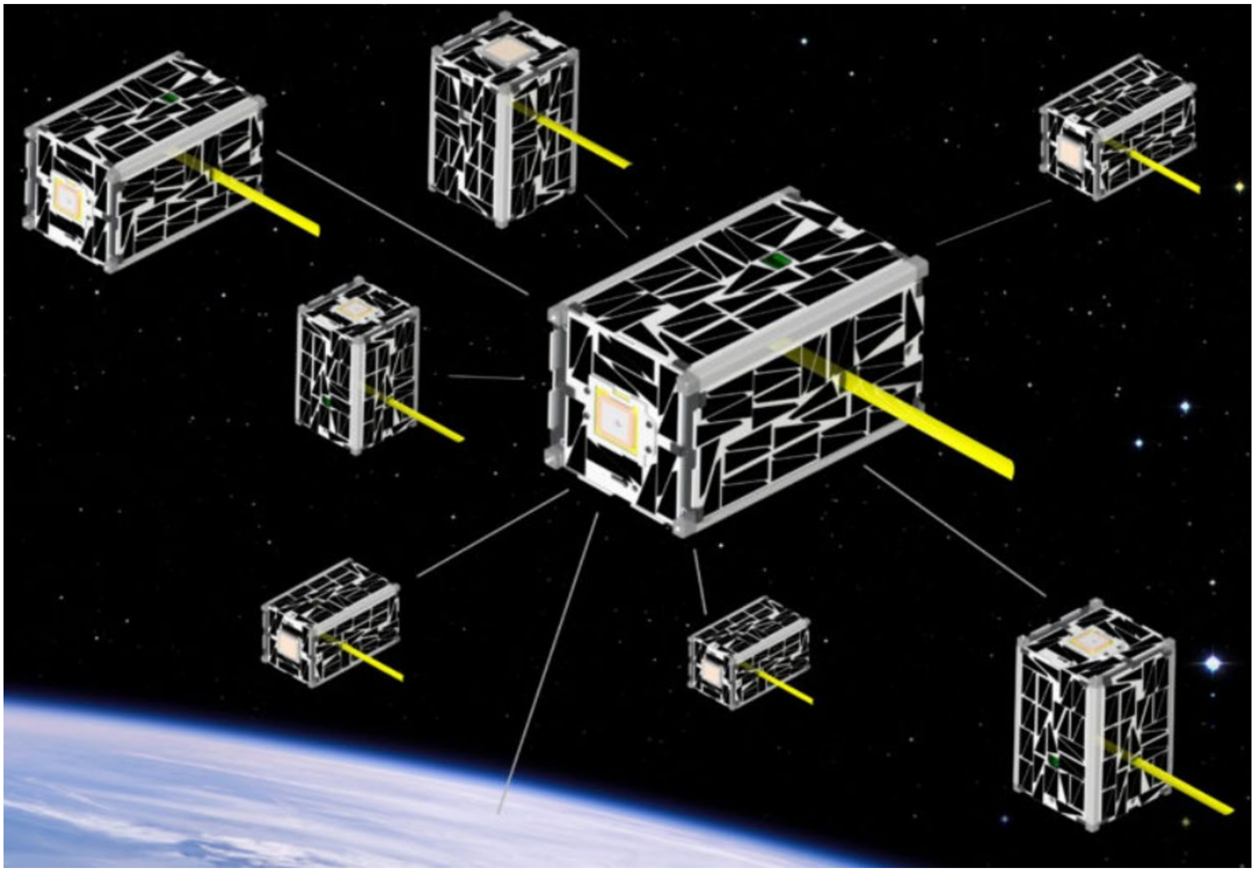

**Pay-Per-Answer.** End-users at the edge only consumes actionable information. Cartoon (top) shows military MANET acquiring data but transmitting information to handheld mobile devices to update the warfighter in the field. Cartoon (bottom) shows sensors, data and analytics which are converged (but concealed) to provide "health safety" answers on smartphones for users (awareness of the level of chemical pollution in the river water ecosystem). Source: Shoumen Datta (top). Eric McLamore and Shoumen Datta (bottom cartoon).

Terahertz (THz) spectrum, occupying frequency range between 0.3 and 3 THz, has potential for transformative applications in communication, sensing, spectroscopy, and imaging due to its desirable properties such as non-ionizing photon energy, penetration capability through optically opaque materials, unique spectral signatures for macro-molecules and chemicals. https://www.nature.com/articles/s41467-019-09868-6

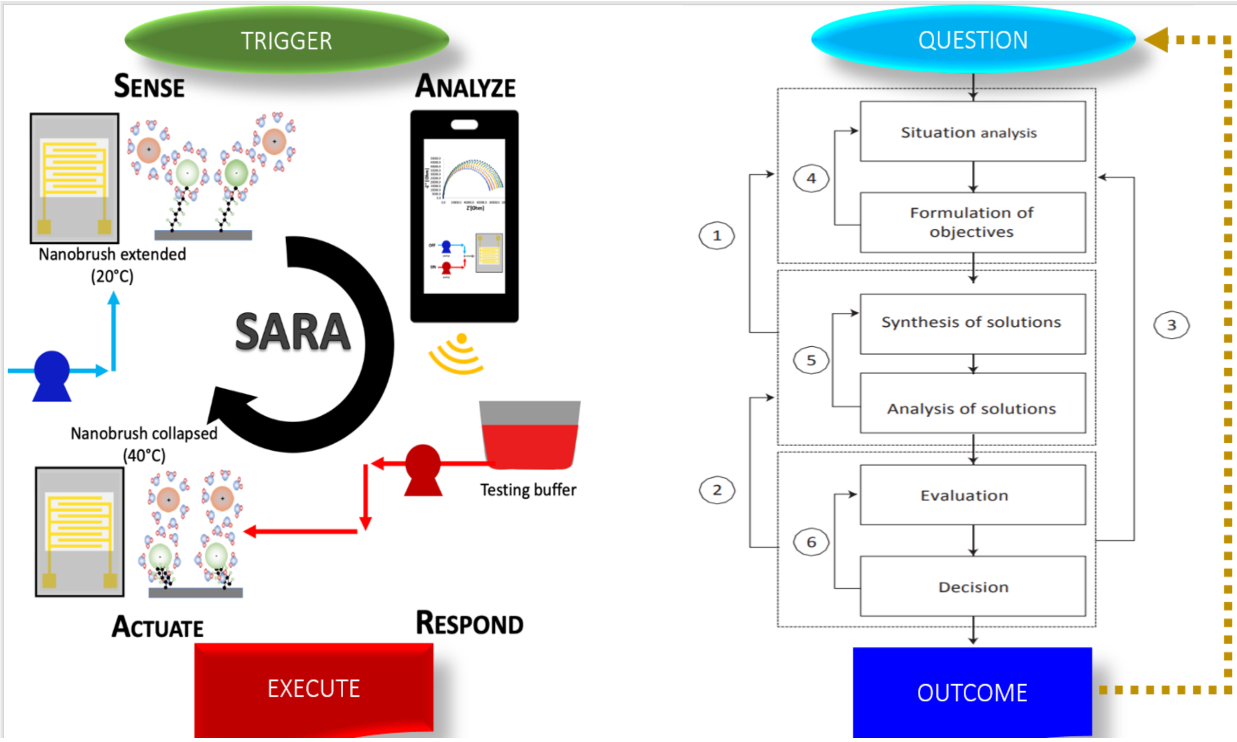


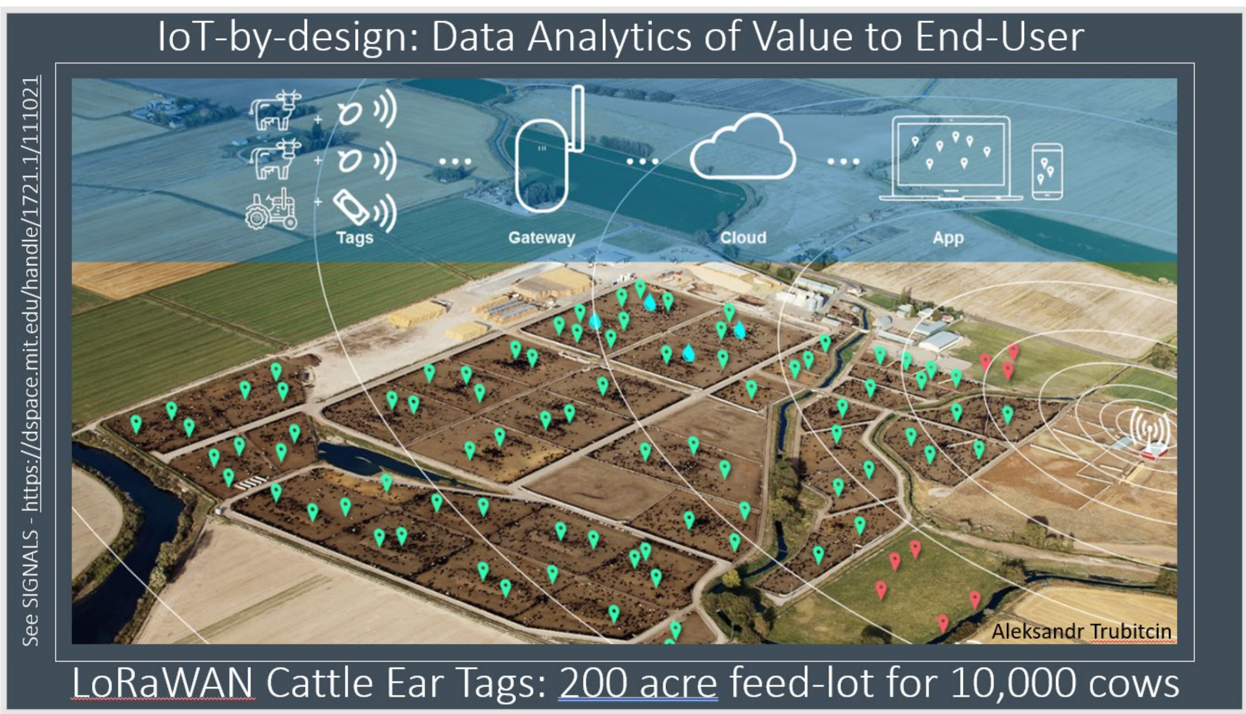



Various types of frequencies may be used in different type of sensors or tools to acquire signals (from the soil). Data acquisition by drones or swarm of drones improves access (applicable to inhabitable locations). Technologies (top) catalyze applications (bottom) for end-users who are non-experts in quest of answers. Cartoon (bottom) created by University of California at Merced and copied/used without any permissions.

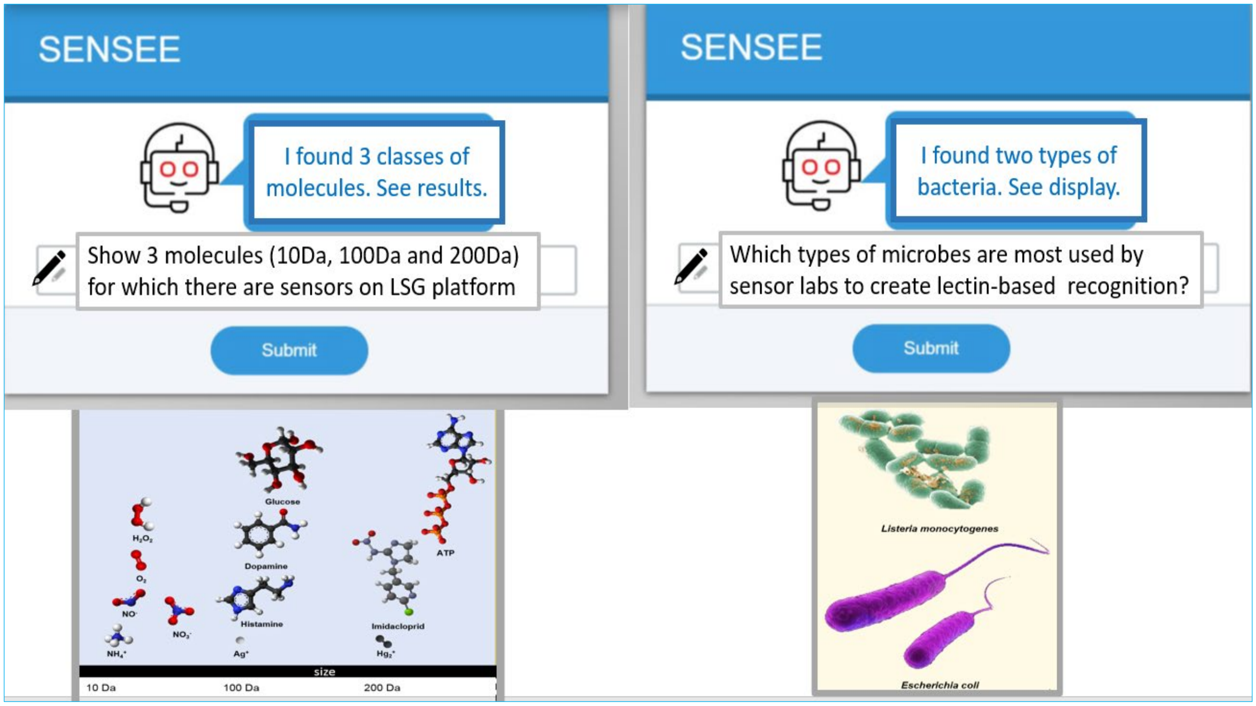


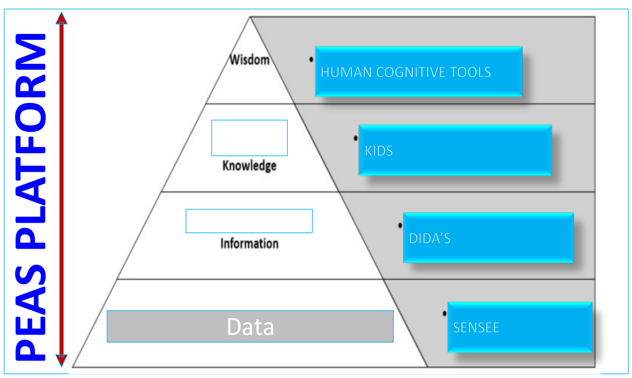

Convergence of technologies (swarm of satellites and radioactivity sensors on drones) providing dynamic mobile information solutions to keep citizens out of harms way after the Fukushima nuclear plant explosion. Technology in the context of architecture (soft, hard) and ecosystem. (Eduardo Castello, MIT Media Lab)

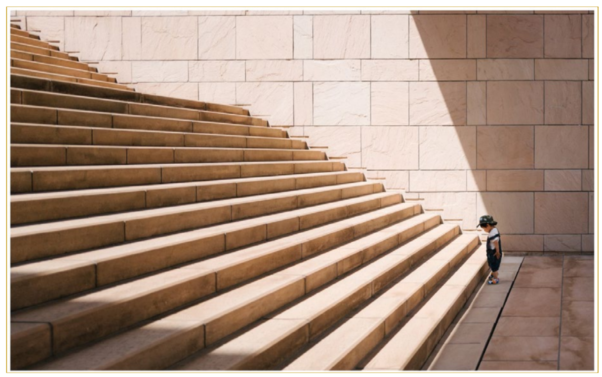


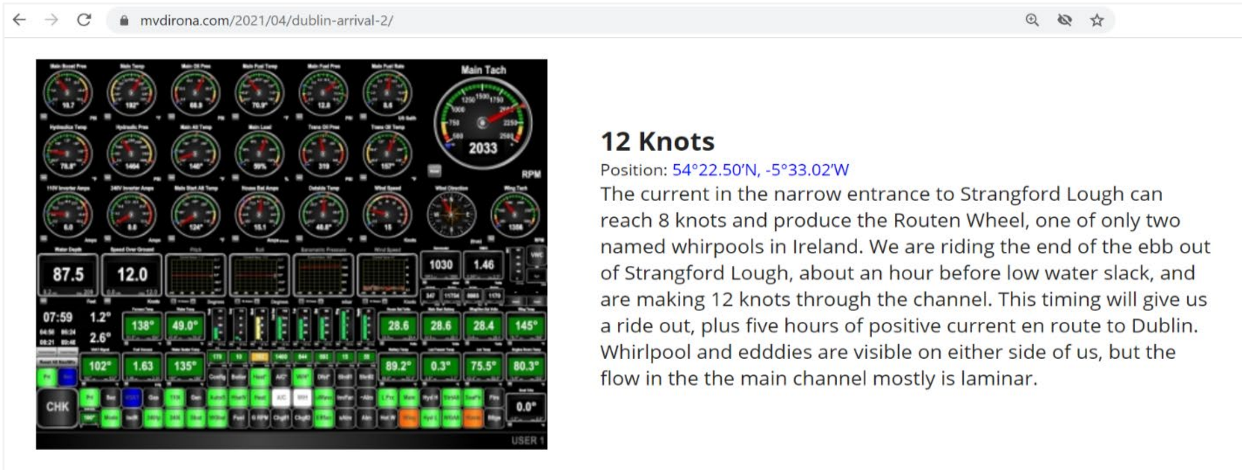

SARA (sense, analyze, respond, actuate) paradigm in using distributed wireless sensor network on dairy farm. In addition to location (this LoRaWAN application), animals with ingested sensors can transmit real-time metabolomic data which could be used to optimize productivity, for example, the biochemistry of the bolus in dairy cows is directly related to the quality (content) and volume (amount) of milk production. The value of data granularity convergence from many domains: crate of fish or meat or tomatoes (sensor ecosystem, food).

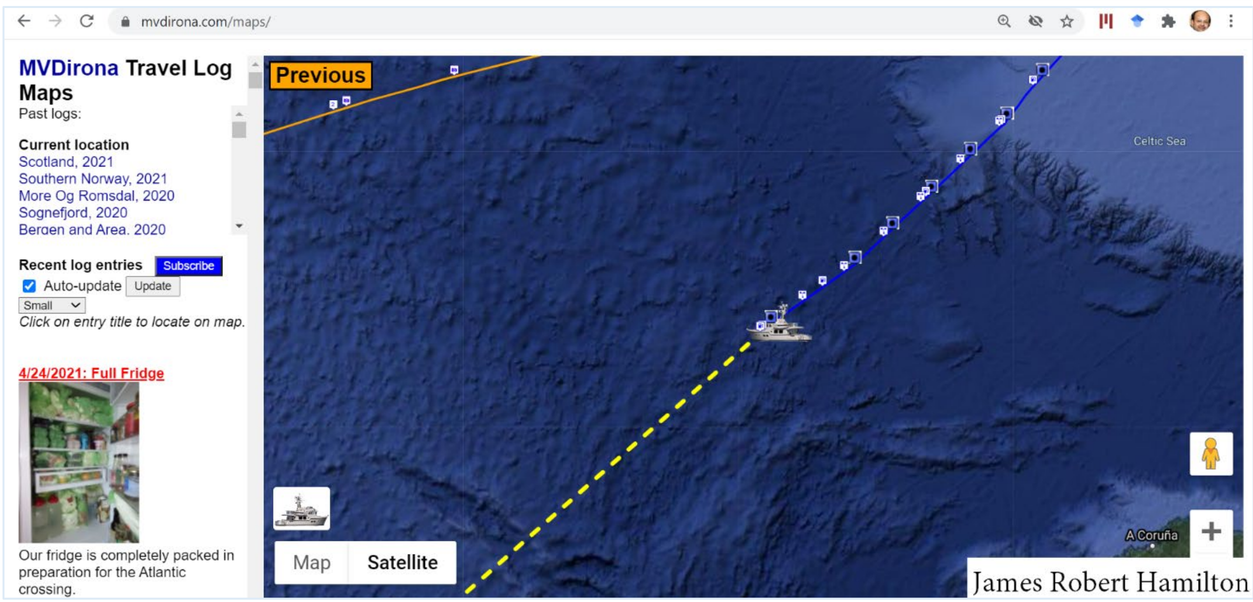


ABCD (Atoms to Bits Connectivity in Decision Support) is an overarching umbrella which can be applicable to almost any domain and application. SENSEE (SmartPath) outlines an attempt to create an information base to enable non-expert end-users (farmers, produce growers) to use natural language (NL) queries via an app on a mobile smartphone to obtain relevant information about sensors based on the user's context or problem. It is one component of a multi-component system which must be dynamic, composable and context-aware. The system starts with data acquisition and in this case (SENSEE/DAMS) sensor data is supposed to be the basal layer. But, how does the non-expert end-user (farmers, produce growers) determine what to sense and how to sense? The conventional wisdom is to contract a consultant or a company to outline a system to sense the attributes of interest and then a logic layer that can (hopefully) make sense of the data to provide the user with actionable information. SENSEE is a supporting software tool if users want to explore the information base. Asking the correct question eventually determines if the outcome of that exploration is information or frustration. If the end user (farmer, grower) is able to ask the correct question, ***then*** SENSEE may be helpful. The SENSEE system may need other linked information which may be a pre-requisite for the user in order to arrive at the correct question about the type of sensor (at this time these other functions are not available). Synthesizing contextual data and context-aware data fusion to optimize actionable information to extract value from investments in decision support tools is a herculean task mired in the quagmire of epistemology, syntax and semantics, as indicated by the PEAS (percepts, environments, actuation, sensors) DIKW pyramid. DAMS may be just one tiny step in our attempt to make sense of data and retrieve information, if there is any information in data. Our idea is to contextualize tools and technology with architecture and ecosystem.




**12 Knots** (top) is the non-expert's readily usable/actionable information which may influence decisions with respect to navigation. The rest of the interfaces on the panel (top) are data for the process and of value to the system and experts (eg James Hamilton) but not the "edge" information for ordinary users ("speed" of boat). James is now (bottom) crossing the Atlantic (location: mvdirona.com). About James Robert Hamilton: please see https://mvdirona.com/jrh/Resume/Resume.pdf & www.wired.com/2013/02/james-hamilton-amazon/ as well as: www.businessinsider.com/amazon-engineer-lives-on-a-boat-and-works-from-hawaii-2013-2

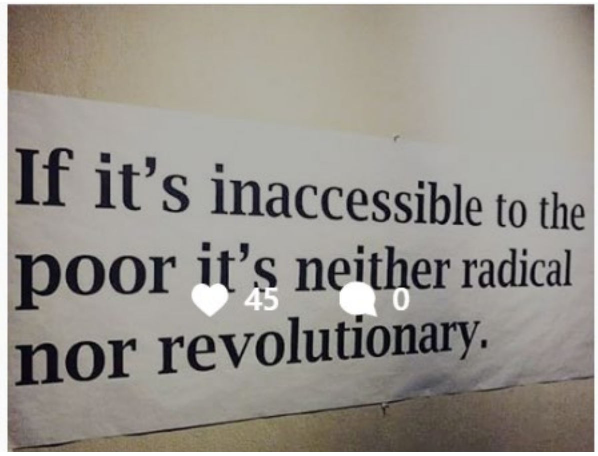
**From:** James Hamilton <jamesroberthamilton@gmail.com>
**Sent:** Friday, April 30, 2021 4:26 PM
**To:** Shoumen Pa Datta
**Cc:** jrh@mvdirona.com
**Subject:** Re: Request for zoom call - AWS academic issue

I like the track and trace application of the stolen trailer. It's getting increasingly affordable to put full tracking onto even medium priced objects. And it makes sense to instrument even very low value payloads to detect shipping problems (temp, water, ...) or potential supply chain of custody lapses.

I have 6 Raspberry Pis on the boat tracking 100s of data points every 5 seconds going back nearly a decade. And, of course, there are a lot of solutions less expensive than a Pi. What's becoming affordable right now is truly interesting.

On Fri, Apr 30, 2021 at 9:08 PM Shoumen Pa Datta <shoumen@mit.edu> wrote:

> This is an unique email James and glad to hear from the N. Atlantic via satellite. Let's talk when you are back in terra firma and you have access to usual modes of telecommunications (cell phone). Your insights and pointers will be invaluable. Your MV Dirona tracking triggered nostalgia - see attached PPT. BR, Shoumen

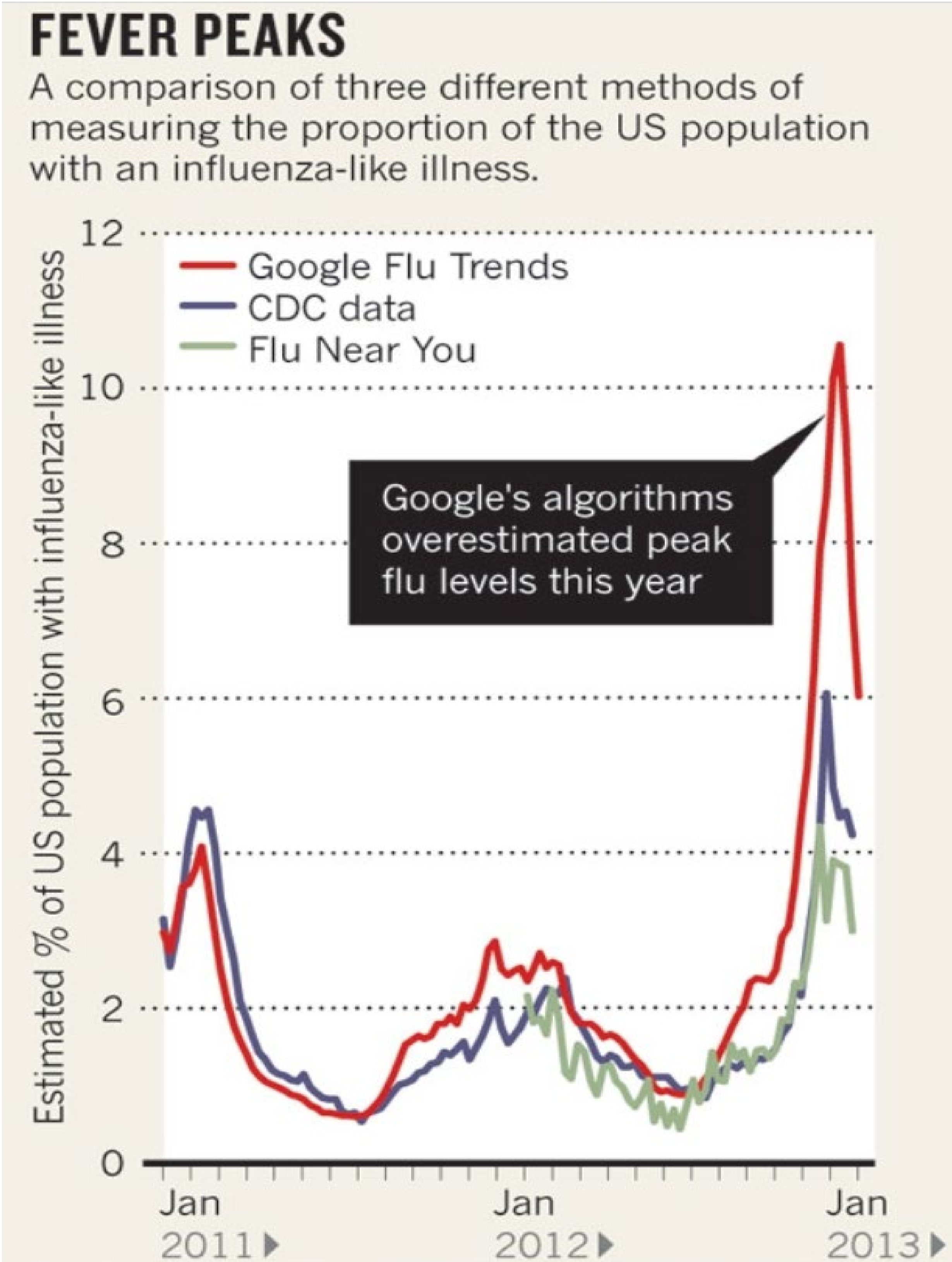

technologyreview.com/magazines/the-food-issue/

MIT Technology Review

Humans and technology

# How technology might finally start telling farmers things they didn't already know

In the Salinas Valley, America's "Salad Bowl," startups selling machine learning and remote sensing are finding customers.




https://www.nature.com/news/when-google-got-flu-wrong-1.12413

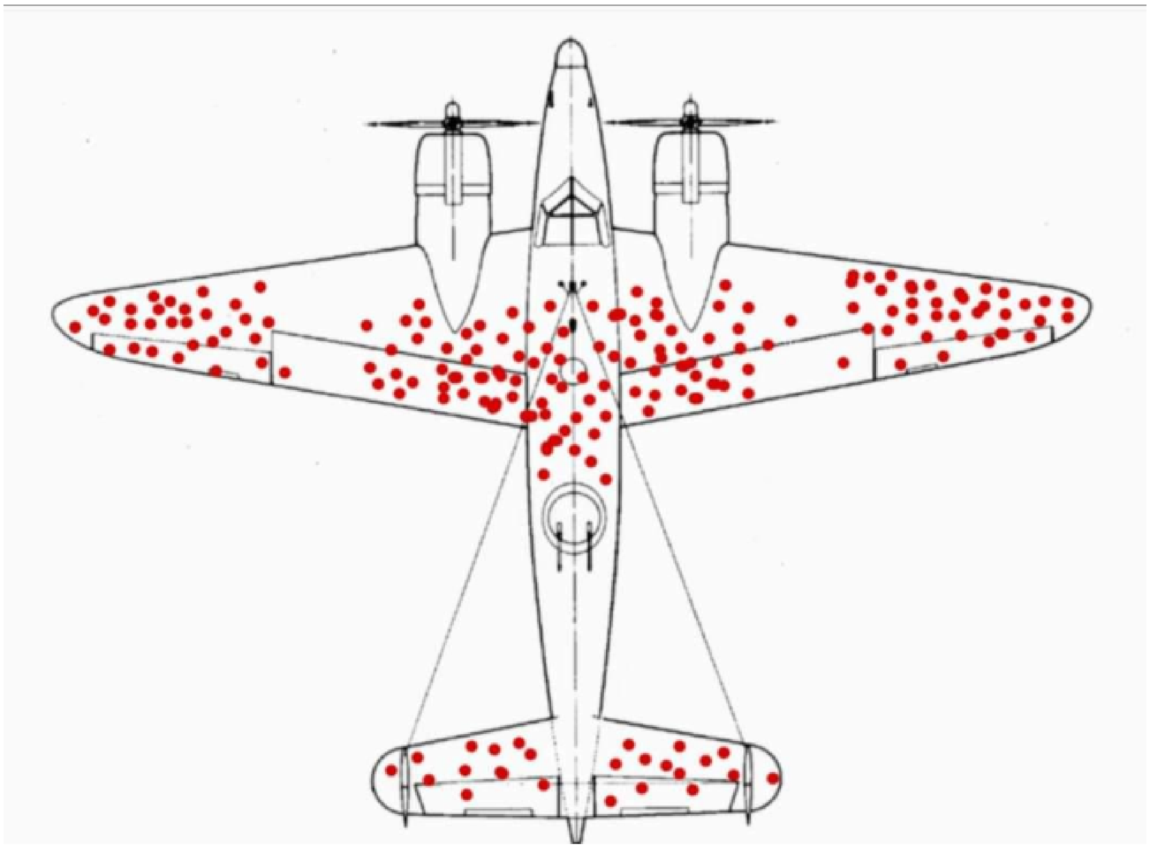
In systems thinking a common analogy is that of six blind men from Indostan (India) touching various parts of an elephant and claiming that "elephant is a tree" (man who touched the leg), "it is like a rope" said the blind man who touched the tail. "Like a snake" (man who touched the trunk). "It is like a big hand fan" (man who touched the ear). "Like a huge wall" said the man who touched the belly of the elephant. "Like a spear" said the blind man who touched the tusk. They were misguided in describing the part (**architecture**) they touched and as a consequence their interpretation is incorrect in the **context** of the whole animal (**ecosystem**).

During WWII, the Navy tried to determine where they needed to armor their aircraft to ensure they came back home. They ran an analysis of where planes had been shot up, and came up with this.

Obviously the places that needed to be up-armored are the wingtips, the central body, and the elevators. That's where the planes were all getting shot up.

Abraham Wald, a statistician, disagreed. He thought they should better armor the nose area, engines, and mid-body. Which was crazy, of course. That's not where the planes were getting shot.

Except Mr. Wald realized what the others didn't. The planes were getting shot there too, but they weren't making it home. What the Navy thought it had done was analyze where aircraft were suffering the most damage. What they had actually done was analyze where aircraft could suffer the most damage without catastrophic failure. All of the places that weren't hit? Those planes had been shot there and crashed. They weren't looking at the whole sample set, only the survivors.



Abraham Wald 1945 - https://www.jstor.org/stable/2235829
Abraham Wald 1946 - https://www.jstor.org/stable/2236089


The Helicopter on Mars: https://mars.nasa.gov/resources/20143/dare-mighty-things/
https://www.wired.com/story/nasa-lands-ingenuity-the-first-ever-mars-helicopter/

chicagotribune.com/news/ct-xpm-1992-01-01-9201010041-story.html

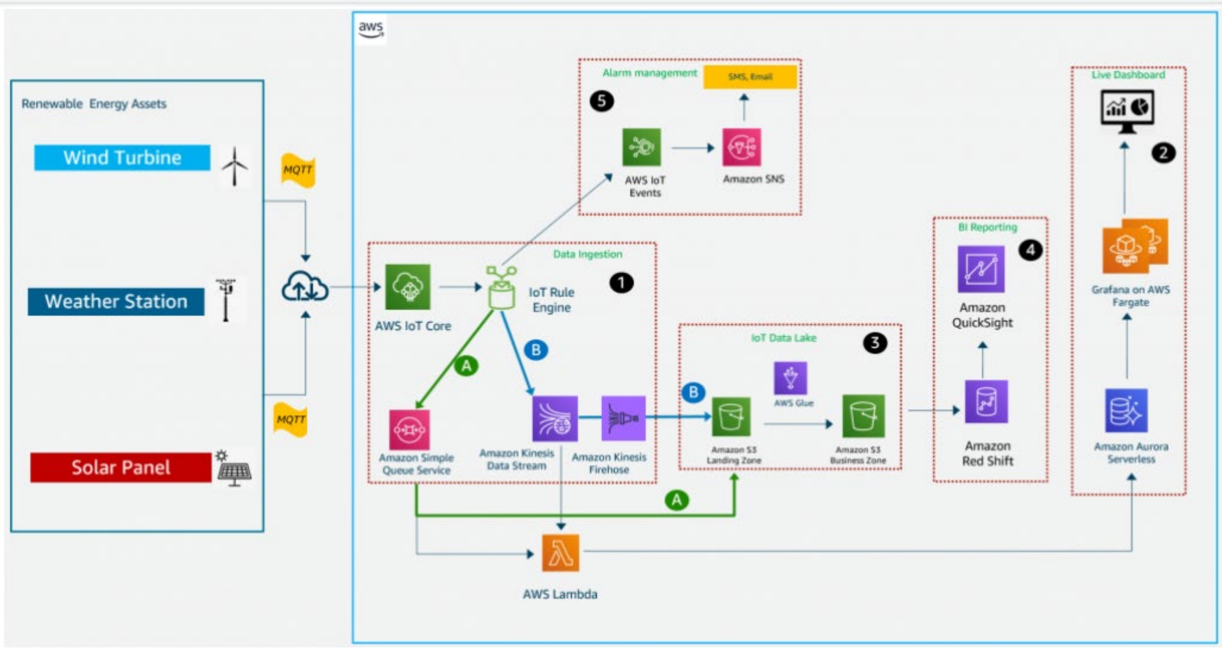
You can almost see Daniel Burnham, filled with dignity and vision, standing before a turn-of-the-century Chicago audience-the City Council, perhaps-exhorting his listeners to live up to his dream:

"Make no little plans; they have no magic to stir men`s blood and probably themselves will not be realized. Make big plans; aim high in hope and work, remembering that a noble, logical diagram once recorded will never die, but long after we are gone will be a living thing, asserting itself with ever- growing insistency. Remember that our sons and grandsons are going to do things that would stagger us. Let your watchword be order and your beacon beauty."

Daniel Burnham (September 4, 1846 -  June 1, 1912, buried in Graceland Cemetery, Chicago, IL) was the architect and overseer of the 1893 World`s Columbian Exposition and co-authored the very far-sighted 1909 Plan of the City of Chicago that resulted in the city`s park-bordered lakefront with outer-edge ring of forest preserves within a future-proof yet elegant urban architecture.

# Where is this leading to?

*One potential path – F I D S E – is a collection of old ideas viewed with an updated perspective which is illustrated as an architecture cartoon (below) outlining real-time monitoring of energy assets. FIDSE (discussed below) is a similar conceptual idea applied to sensor ecosystem, which is a superset of the application/architecture shown below (because sensors are ubiquitous across many verticals and domains serving a plethora of applications). The outcome from real-monitoring architecture is shown on the next page and represents the data visualization on mobile devices, expected for potential applications discussed in FIDSE.*

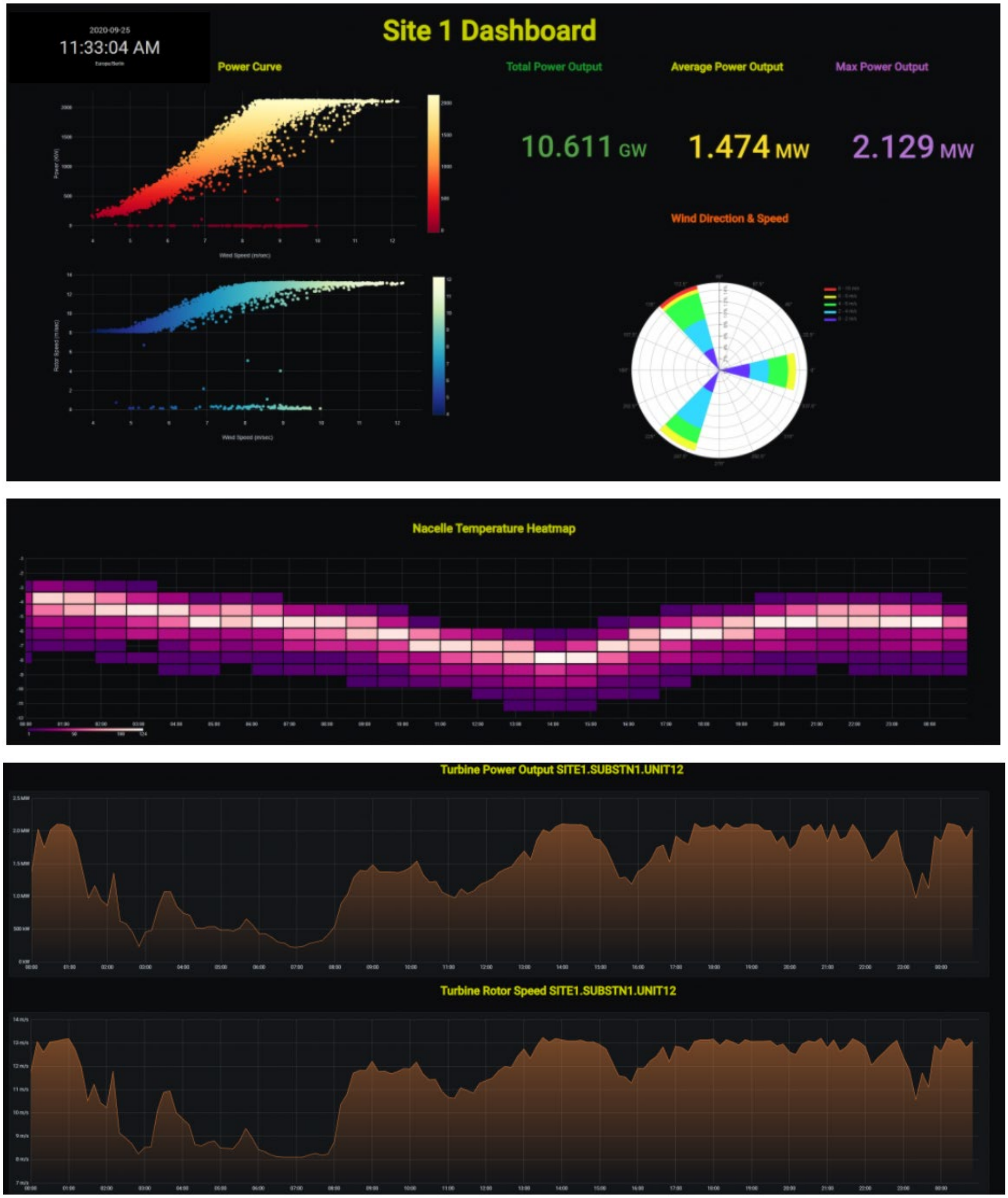



In this scenario, we have wind turbines, weather stations, and solar panels generating telemetry data and sending it to AWS IoT Core. This reference architecture achieves the following objectives:

- Scalable and secure high velocity data ingestion
- Scalable and real-time situational awareness serverless dashboard
- Hydration and curation of IoT data lake
- BI reporting
- Alarm management

Source: https://aws.amazon.com/blogs/industries/real-time-operational-monitoring-of-renewable-energy-assets-with-aws-iot/

# *Data visualization on mobile devices, expected for potential applications discussed in FIDSE.*




Source: https://aws.amazon.com/blogs/industries/real-time-operational-monitoring-of-renewable-energy-assets-with-aws-iot/

Data acquisition from electrochemical sensors are an important part of the DIDS data granularity but instrumentation to capture and analyze the raw waveform data is generally absent because devices transmit only numeric data after "software" processing where "model" fitting is the staple tool. There are very few provisions to include system architecture to store raw waveform data (except perhaps in healthcare) on devices by including tinyDB by design, based on the principles of tinyOS

*Physiol. Meas.* **41** (2020) 035008 A J Goodwin *et al*

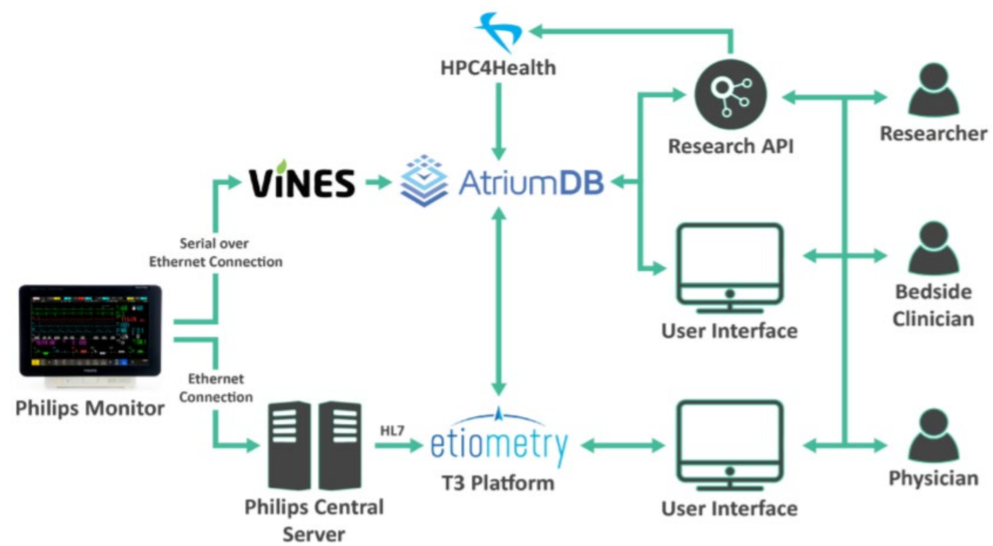
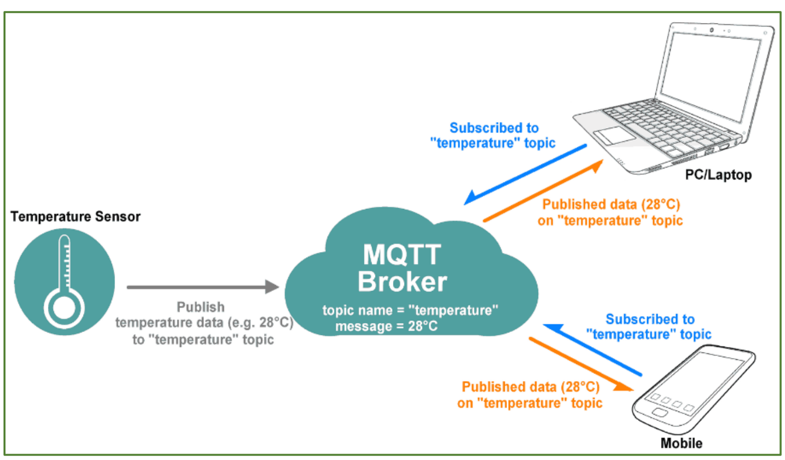



**Figure 1.** System architecture. Physiological waveforms and 1 Hz numeric variables are collected by ViNES device integration software and delivered to AtriumDB where the signals are compressed and stored on disk. Physiological variables are emitted from the Philips central server at 0.2 Hz and aggregated in T3 database. Data can be programatically retrieved from either storage system via an application programming interface (API) .

https://stackoverflow.com/questions/1050331/storing-waveforms-in-oracle

Lightweight publish-subscribe MQTT protocol (Message Queuing Telemetry Transport protocol) may transfer data from devices to cloud storage. https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.pdf Documentation of open source MQTT standard is preferred over citizen sourcing (Arduino, Raspberry Pi). Centralizing data in secure MQTT server may offer advantages for edge operation using mobile devices for interaction with users in the field. MQTT thin client on the device uses MQTT broker to connect to other devices as well as MQTT server (data store). Masked authentication messaging (MAM) offers data security.



**Fluid-like Infrastructure of Data from Sensor Ecosystems (FIDSE): A perspective of data as a public utility analogous to water, energy, electricity, gas, sewers and roads.**

Thought log:

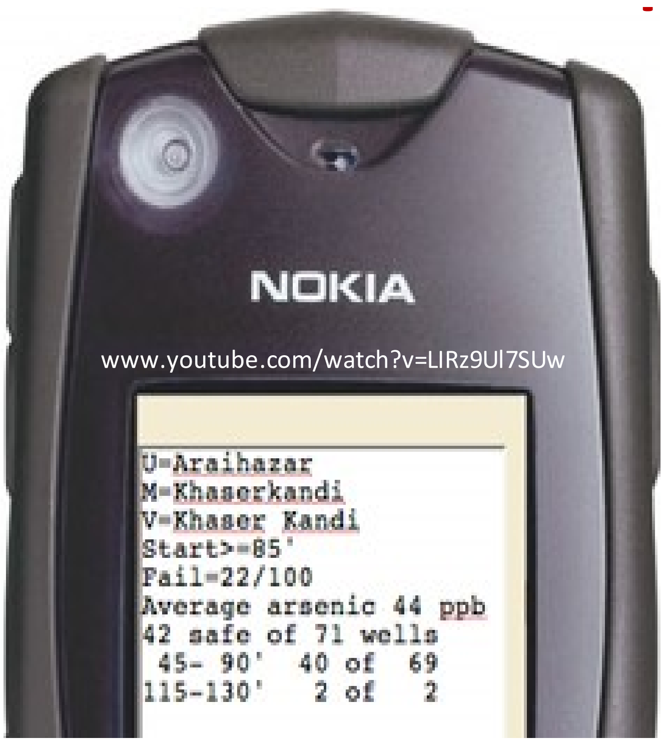
In most affluent nations, centralized water distribution systems may be viewed as pipelines. A source is linked to a water purification station (think data curation) which leads to trunks and tributaries (connects to the last mile, as in, homes, offices, schools, stores, etc). This is the analogy suggested in the title (FIDSE) where data from objects (physical things, IoT), processes and decisions (not objects but events) "flows" in a "pipe" which is open yet secure, within reason. The data is tagged to be identifiable (think id, for objects it may be RFID). The data "flowing in the pipe" is accessible on demand analogous to your faucet at home or office or farm, which, when turned on, gives you access to drinking water or water for irrigation. The data in the pipe is not only identifiable (by source) but it can be sorted, categorized and analyzed by selecting "features" associated with the data in the pipeline/platform. By accessing the "pipe" or platform, hospital can download your routine health data (avoiding the quagmire of syntax and semantics) or if you are in an accident in a different location then the emergency room can search for your data in the "pipe" using "features" linked to you or your specific health condition which may be critical at the point of emergency treatment (for example, if you are a diabetic then the treatment should exclude using steroids or you may be comatosed). Security by design is essential in data flow pipeline concepts (health/healthcare are difficult for open implementation).

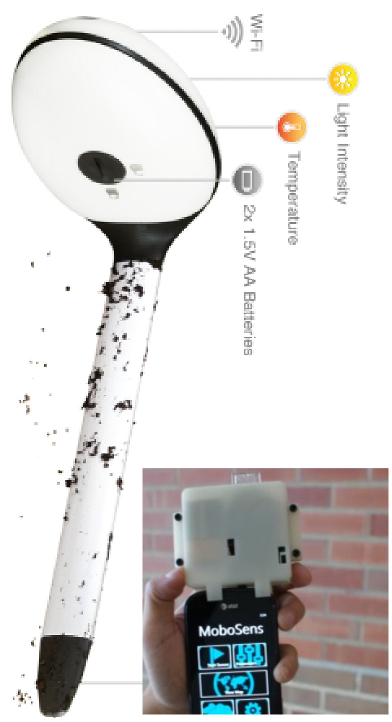
DIDS digital thread is the foundation of ABCD (DAMS, DWSN, MANET) where domain specific data pipes converge and selective data fusion may improve the value of data to extract obvious as well as non-obvious information. FIDSE is an infrastructure perspective which may be complicated to construct due to lack of trust between humans in the loop (similar to creating standards). But, the skeleton (DIDS open platform) for FIDSE may be worth exploring. If entities feel comfortable they can "plug in" to the platform (standards compliant APIs) to send/analyze data or others decision support tools/analytics/visualization/mobility.

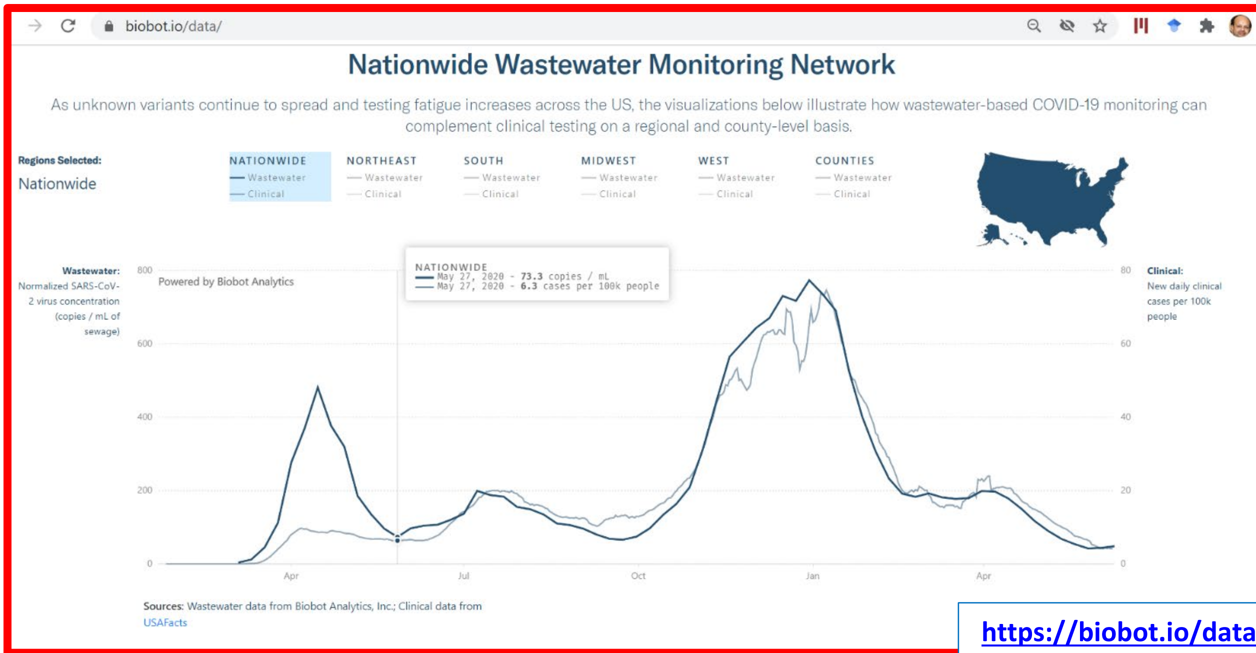
In the eye of the pandemic, the purpose of FIDSE must emphasize essential public goods and FEWS (food, environment, water, sanitation). Remote environmental monitoring is key but what do we wish to sense? Using the MANET component of FIDSE may be one route but the choice of sensors remain an open question. From an epidemiologic perspective, signals indicating emergence/presence of infectious biological agents will provide the most useful data from remote locations, environments and sewer systems (unfortunately, half of the world lacks sewer systems and managed sanitation services).

Water, wastewater and other forms of water quality management are important domains for affluent nations (developing nations are far removed from centralized water systems infrastructure) which must embrace monitoring sewer systems. The combined/aggregate data output from these activities could be central to public health risk mitigation strategies if "data pipes" were accessible (democratized) for remote analytics. The scenario in Flint, MI (Pb) and Tampa, FL (P) is not far (at least in principle) from Arsenic in Bangladesh. Without detection (sensor data) it is impossible to estimate the presence of problems and search for any solution. The potential for cyberthreat in water treatment plants and effuse regulation presents potential danger to the public if disenfranchised individuals (malicious hackers) may regulate valves and override safety systems to release very low levels of harmful contaminated water to mix with drinking water and farm water used for fresh produce (lettuce in the field in CA or hydroponic lettuce growing indoors in MD or FL). It is of little use to keep data local or regional in stand-alone utilities databases without the power of analysis which can be applied to the data from tools operated by state or national agencies (for example, CDC). The convergence of "data pipes" around the nation may offer greater transparency through the power of analytics.

## Reality Check ☑ Arsenic in Water, Bangladesh






[https://biobot.io/data](https://biobot.io/data)


Ubiquity of water as a central medium in the context of FEWSH (food, energy, water, sanitation, healthcare) makes it a critical asset for **monitoring** from the perspective of security (drinking water), agriculture (food, irrigation, use/re-use of brown-water for farms/fields), the nexus of food-energy-water as well as the meteoric importance of wastewater in the current climate of a raging pandemic (shown in cartoons, above) which has thrown healthcare into a pandemonium. Real-time data from testing water quality – by monitoring heavy metals in water for irrigation or viruses in wastewater or quality of drinking water – may be one of the most important tools we have at our disposal to catalyze prevention. The umbrella of DIDS offers tools and the technology is becoming affordable (in affluent nations) but implementation of digital transformation for monitoring water-related domains are few and far between. It is time to view "digital" water with a lens which can bring the different fields into focus. Solutions and services must shed the myopia and address systems level synergy and integration with a renewed sense of urgency. We must invest in the system to obtain reliable data and analyze data with rigorous techniques to offer actionable information for users.

Ag, ag-related services (machinery) and the meat industry may benefit from platform approach of DAMS and DWSN where businesses can create private / public "views" of data and contribute/connect data to analytics. The bigger picture is the food industry, as a whole, where preventing waste alone ( shelf life, spoilage) may save $285 billion each year because $285 billion worth of food is thrown away in the US. Most food is produced with scarce resources (water, soil, slaughtered animals) but will not be on the plates of people who go hungry every day because 54 million tons of food are trashed in the US, annually. Optimization of resource utilization is unlikely to happen if data dies in silos. Data sharing and democratic collaboration may be one option to reduce waste. Exposure of business data can be secured but the cost of inactivity is in trillions.

FIDSE will be difficult to build and even more onerous to maintain but it may not be a philanthropic project for charitable purposes or solely for affluent nations. Distributed use in everyday essential aspects of life and living makes it imperative for markets of billions to interact with digital tools, if it provides ***value*** through FIDSE. By lowering the transaction cost through digital services it may be possible to enter global markets even with very low purchasing power parity (PPP based on per capita GDP) by reducing traditional barriers to market entry through the low-cost ubiquity of mobile platforms. It is not a panacea but a model for the less fortunate world where access to public goods, FEWS and healthcare are still only for the privileged few. Data cannot feed people and FIDSE cannot provide their daily bread but data informed service may be an economic muse. By incorporating ethical profitability, services related to essential goods may be even profitable. Social businesses can sustainably earn micro-revenue by adopting *pay-a-penny-per-use* type of micro-financial tools. The important criteria in any data catalyzed system is the delivery of ***value*** which can improve quality of life.

The disconnect between science and human values erupts from the grave discord between advances in usable technologies vs the economics of technology in the context of applications and sustainable transaction costs. The democratization of data enabled by the explosion of software defined services over the past fifty years is a shining example of convergence of science and engineering which has reached mass users, albeit unequally. The lack of understanding of the economic impact of engineering tools at the hands of non-expert edge-users suffers from the collusion between big tech and exclusive cartels such as the world economic forum (WEF). Peddling prosperity for the affluent few (>20% of the global population) excludes thinking and designing potential benefits from technology to help farmers, sewer workers and laborers who aren't invited in Davos.

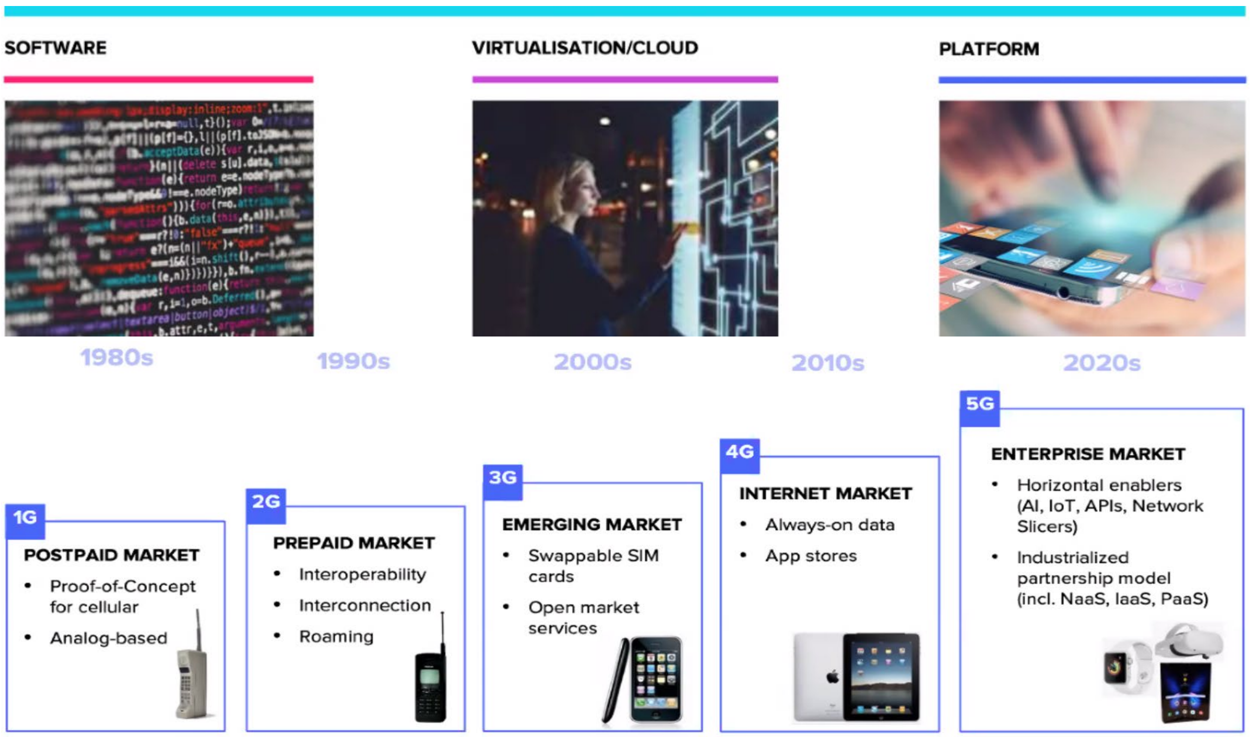
The tools for democratization of data may not be in short supply but the supply chain of value delivery to the less affluent working population suffers from a laissez faire approach which has prevented building a purpose driven ecosystem. The ecosystem of sensor networks and the mundane granular application environments in this discussion is aimed at essential goods and services we need rather than non-essential luxuries we want.

This discussion (FIDSE) is incomplete and useless if we cannot translate the economic value of these tools. If non-expert edge-users (farmers, growers, truck drivers, public health workers) can integrate and synergize the outcomes and functions of these tools with their daily routine and/or domain specific operations, only then we can extract tangible value. We must aim to improve our ability to articulate this potential value by building bridges between technology and capturing its value to the economy (workforce development, job creation, direct value to users and indirect but positive influence within the domain or ecosystem). The granularity of sensors and data from sensors must make analytical sense to generate actionable information. The value of information must be measured in terms of key performance indicators and economic growth.

Connecting distributed domains to converge data and information is the Holy Grail. Source: https://aws.amazon.com/blogs/aws/introducing-amazon-memorydb-for-redis-a-redis-compatible-durable-in-memory-database-service/?

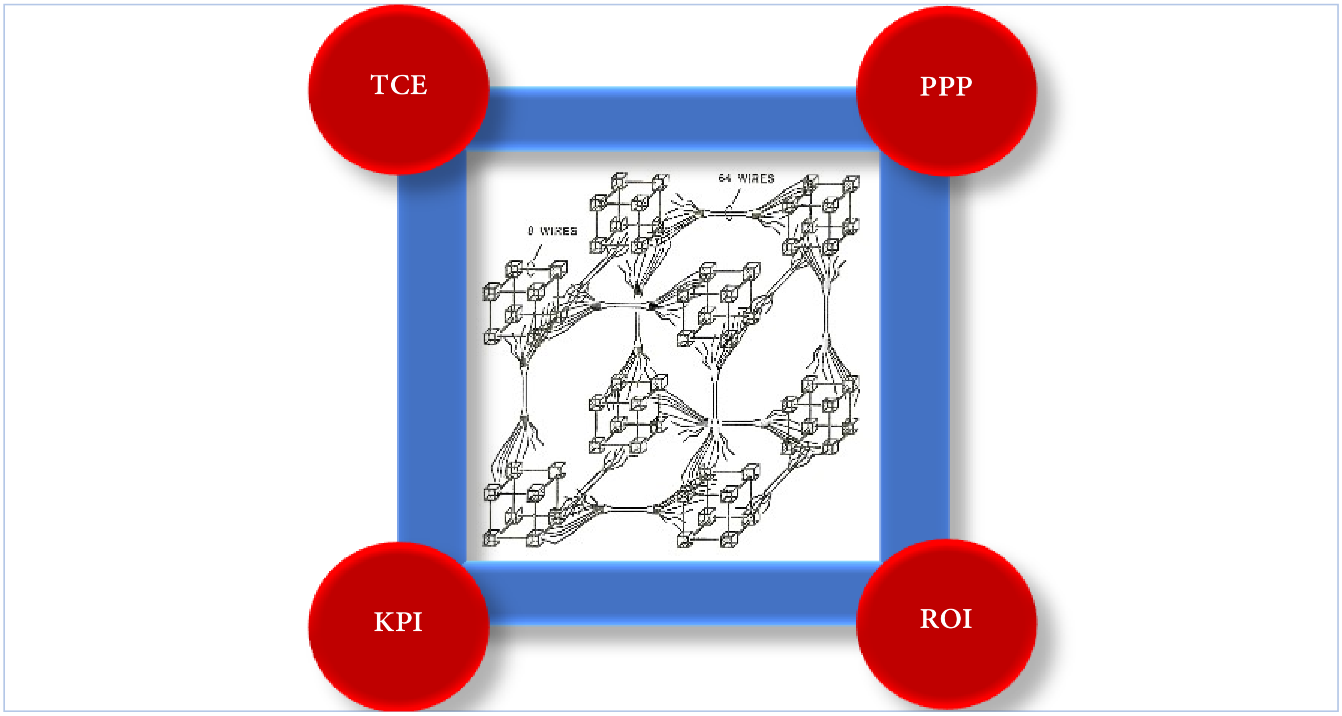
## Pay-A-Penny-Per-Use (PAPPU): CONNECTING MARKETS for PAPPU CONSUMERS

Economies of scale is pivotal to profit from the paradigm of PAPPU-based financial tools. Lowering transaction cost in each step and in every instrument in the transaction process will make or break this business model. Telecommunications is the information backbone of this system and software is the medium of connectivity. Bells and whistles are optional.




**Democratization of data is made possible by lowering transaction costs through a convergence of software engineering, device engineering and telecommunications. The issue of time assurance is largely overlooked in software infrastructure except for discussions in the 5G domain where time sensitivity of cyberphysical systems may be difficult to ignore. See [https://www.youtube.com/watch?v=Dm92rsguzzA](https://www.youtube.com/watch?v=Dm92rsguzzA)**

The thorniest question for FIDSE ("tempest in a tea cup") is its value and valuation at the hands of the next billion edge users (PAPPU consumers). Without consumption these ideas will die from paralysis by analysis. Short term commercial growth within the realms of ethical profitability is a context-dependent and/or case-specific push-pull in social business innovation at the nexus of four pillars: PPP, TCE, KPI and ROI (purchasing power parity, transaction cost economics, key performance indicators and return on investment) with an emphasis on the relationship between transaction cost and performance (value to users).

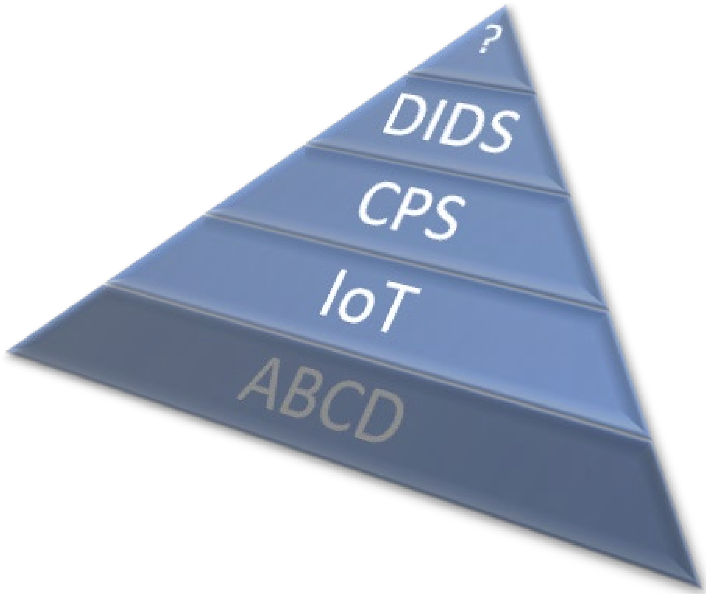



Short-term market growth may depend on the strength of the pillars (TCE, PPP, KPI, ROI) but interdependencies between subsets and case-specific interrelationships will lead to a myriad of permutations and combinations (cartoon of agent based "cubes-on-cubes" from Marvin Minsky's *Society of the Mind*). Determining the value of what the end-user values and what cost the system (community, PPP) can bear will influence digital service design.

## *Long-term global market creation: Expanding ethical consumerism*

Long term **value** creation isn't a business tool. Appreciation of value is rooted in education. *An educated consumer is the best customer.* A future section of this thought log may discuss socio-economics of education as a catalyst for freedom, development, economic growth and quality of life, by ***focusing*** on "mathematics for mothers" (**MATHHERS** *pronounce matters*). We espouse the idea that if mothers are mathematically proficient (to a certain extent) then the children will be mathematically inclined (https://bit.ly/MATH-HERS). If the female children are mathematically inclined, she may create a future family proficient in math. Logical analysis, rationality of thought and data-informed decision making depends on grasp of philosophy and mathematics. Maternal aptitude/education shapes children, their life and future. Please see fig 4 & 5 https://nces.ed.gov/programs/coe/indicator_tbe.asp *Does the apple fall far from the tree? Explore https://dspace.mit.edu/handle/1721.1/131129*

SUMMARY IN PRINCIPLE – **dFEWSH** – digital: food, energy, water, sanitation, healthcare




The central theme in a plethora of **dFEWSH** endeavors may depend on the diffusion of data which can inform decisions in a manner that non-expert end-users are able to extract/access actionable information which provides value (example: profit). The alphabet soup at the bottom of the pyramid starts with the granular concept of "atoms to bits" which imagines that every "atom" (physical thing or process layer) may contain information or "bits" which must be connected to harvest what improves or catalyzes decision making (ABCD – Atoms to Bits Connectivity in Decision). In more traditional descriptions, this ABCD foundation is the "data" acquisition layer where sensors, tags, manual input and other "sources" contribute "raw" data. The internet of things (IoT) is the canonical data layer for objects and things which feeds on the granularity of the common denominator (ABCD). Cyberphysical systems (CPS) combine data from "atoms" and physical "objects" with instructions (bits without atoms) from commands, processes and other cyber procedures which directs/determines actions/reactions in the networked physical world. The user may benefit from the synthesis of these layers (with very blurred boundaries) through the umbrella referred to as DIDS where the distributed data may be curated for quality control, analyzed with one or more mathematical and/or statistical techniques, perhaps fused with other internal or external data. The outcome of this data processing may be information which "informs" the user (human in the loop) how to make better decisions (hence, this is a service). The "edge" interaction may happen in an immobile context (control tower, office, factory) or more likely to be dynamic interactions between the outcome (data, information) and the user via a mobile platform (smartphone, tablets) where users can access the decision support system at will (*ad hoc*), anywhere, anytime. The mark of interrogation at the tip of the pyramid offers further room for imagination, invention, innovation and interpretation of the collective path which data may take to arrive at information which users can synergize/integrate to improve or profit from their decisions/actions. This ability in some tiny part of the [ ? ] segment on top of the pyramid is suggestive of an "always on" real-time digital proxy of operational systems (farm, factory, flying saucers). The latter is increasingly and erroneously referred to as "digital twins" to enhance the marketing panache of what may be simply a vanilla variety of a dumb digital duplicate, which can, in appropriate circumstances, provide useful information and collective system status.

# *Part II ◆ In Practice*

**Diffusion of DIDS in applications of dFEWSH: From Fundamental Principles to Practice**

Because we are complex systems and generally surrounded by or immersed in complex system of systems, it may be difficult to recognize or it may take a long time to arrive at the realization, that most rational actions are governed by a set of immutable scientific principles based on a few (simple?) natural laws (physics, chemistry, biology). The latter, taken separately, is sterile, like a seed, which grows to form a complex system, if planted. Perhaps also analogous to music, silent unless performed. In other words, immutable non-individual molecules, when orchestrated within rules of development, makes organisms and individuals with individuality (who may break the rules).

There is a latent framework in this journey from individual molecule to individuality. It is often referred to as Nature. Scientists are aware of the "framework" and institutions that sponsor science and research value those who can contribute highly detailed research with very specific granular outcome which may elucidate or shed some light on the basic principles which contributes to the development of complex systems. The academic and bureaucratic penchant for details is quintessential and indeed necessary but the incentive for the reductionist approach (for example, the Nobel Prize) far outweighs the inclination to view the whole. The sum of the parts is far greater than what may be gleaned from gluing together fractured and partial studies of details from complex non-linear systems, for example, human behavior.

The over-arching importance of understanding, creating and implementing frameworks may be exemplified by the theory of quantum mechanics. The latter is less of a theory but rather serves as a foundational framework where all (?) physical theory fits or must be fitted (?). A lesser known but another highly relevant framework is the theory of quantum chromodynamics (QCD) based on the discovery of quarks and gluons. But, unlike the theory of quantum mechanics where physical theories "fit" (nicely?) in case of QCD the "fit" appears to be more complicated due to lack of mathematical tools.

The underlying mathematical infrastructure is at the heart of frameworks, which is in turn the common conceptual feature in complex adaptive systems. Conceptual framework or "schema" is a common term. The granularity of such schemas makes it highly modular mathematical expressions (from example, equations for rates/flows) which can be "mixed and matched" (from a repertoire of model schemas) to generate different frameworks depending on the desired outcome (that is, the complex adaptive system it is expected to create/serve).

The modular mathematical expressions that constitute frameworks or schemas may be viewed as a scaffold for data, for example, equations for rates/flows. An equation for rate or flow in a complex adaptive system is only alive or of value when it acquires data or is involved in information arbitrage (for example, data or information from its environment or interaction with/within the environment of the system which can be referred to as the "percept" from the "environment"). The source of data (data acquisition system) may be human input or data harvesting tools, such as sensors. This (sensor) data represents actions in the real world (or system) which will feed the schemas (rates, flows, weights) and the feedback (in a closed loop system) will influence the schemas to change/modify their actions (actuate) to improve/optimize system performance.

Multiple arrays of schemas may have multiple data feeds (may switch between data feeds depending on the "whole" task or final outcome) but their individual outcomes must be synthesized and/or synergized as a whole for the most relevant semantic interpretation which is of value for the whole or the "performance" based on the framework made up of schemas for percepts, environments, actuators and sensors (PEAS).

DIDS offers clues as to how frameworks, performance and PEAS may influence decision support. Unbeknownst to solution providers in the real world, they are attempting to create systemic solution strategies for systemic problems by using this rubric (translating these concepts from ideas to pragmatic implementation). Data, dependencies, relationships and ratios are granularities which influences and/or informs weights, rates and flows to enable titration/optimization of system outcome/performance. If there is a "desired" outcome/performance level that is pre-ascertained by the user then application of a retrosynthetic[1] approach is justified if the system attributes are identifiable (parts, components, characteristics) and numeric values, based on rigorous metrics, are available.

In this retrosynthetic approach, we choose a couple use cases where we have identified parts (fragments/synthons) that create the whole. Could we create mathematical frameworks or schemas for individual parts of the system? Each schema is then subjected to the PEAS treatment and then the components (think individual gears) are brought back together (inside a clock) to deliver the "whole" performance. Complexity of the complex adaptive system may limit the scope of retrosynthesis, in principle, because the ability to deconstruct the end point performance may be influenced by too many "synthons" (in a manner what may happen when state space explodes in an optimization protocol with too many dependencies which influences the final outcome).

There is a bigger question is this process which may or may not be addressed. Since we have claimed that these schema or mathematical expressions are modular entities, then can we create a "library" or a repertoire of these expressions (weights, rates, flows) which can be re-used in other explorations of system performance? The idea may be similar to simulated (virtual) 3D concurrent engineering workbenches where functions and components may be sourced from a repertoire (drag and drop from a menu of choices) to modify/re-construct an engineering object (engine, gear box, pump).

If we want to reach non-expert users then we must create a system where every attribute/characteristic is a feature (explained in natural language) with which the non-expert user is intuitively familiar. We must further simplify the Lego Mindstorm approach which will be fool-proof to serve users in any country as long as they can access/download it a mobile device (smartphone, tablet). Using schemas and modifying the framework will be as simple as dragging and dropping an icon which is profitable for their purpose.

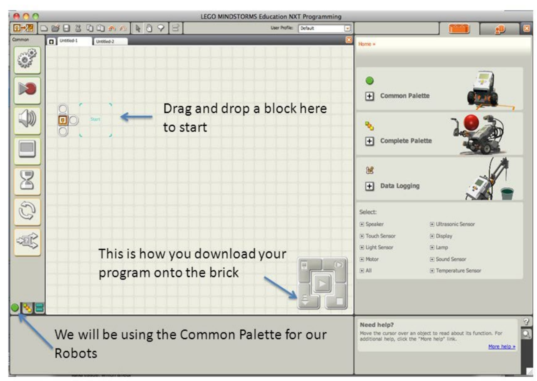
---

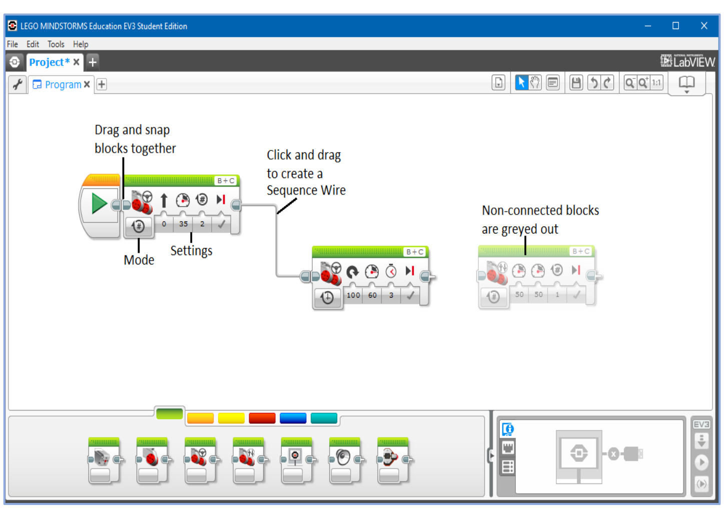
[1] Retrosynthesis is a method of chemical synthesis which involves "deconstructing" a target molecule into its readily available, simple starting materials in order to assess the best synthetic route. This is achieved by breaking the bonds of the target molecular structure into constituent fragments, known as synthons, and by conversion of functional groups into others, known as functional group interconversions. The concept of retrosynthesis was framed and formalized by Elias James Corey, for which he won the Nobel Prize for Chemistry in 1990.

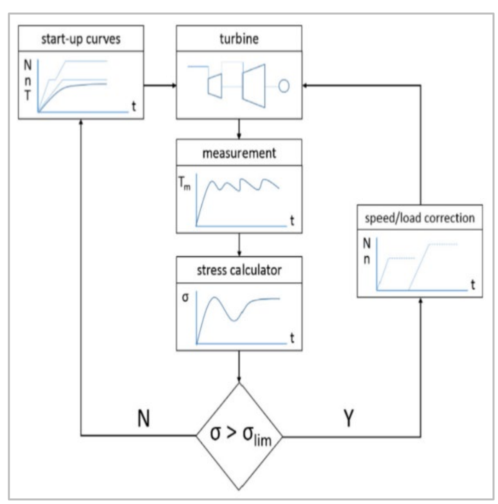
Cartoons representing embedded programming (programmable) as well as drag and drop features in mobile applications with the scope to test "what-if" scenarios by end users. This is a higher than desired level of utilities in the proposed tool for dFEWSH non-expert users. Several tiers of design thinking with end users may be necessary to configure ease of use.


*Ordinary tools/elements which can fit many scenarios need not be "programmed" each time for each application (if using a common platform/workbench). Animal calls and colors are examples shown in the cartoon. Without any knowledge of programming or any code embedding, an user may create a robo-dog using an user interface (UI) and make it look like a black and white spotted Dalmatian but insert "cat purr" as the animal call instead of dog sound (whine, bark1, bark2, growl, etc). Non-technical operators (eg: farmers) will be able to choose/drag and drop criteria that fits their purpose or optimizes for a specific crop.*

*Conceptual user interface for real applications may further simplify the Lego workbench (right) and the series of steps (bottom left). Workbench (UI) and palette of tools for each application (eg: [turbine energy optimization](), bottom right) may display the features / data ranges using simpler icons that users are allowed to modify. Follow-up dashboard displays a simulated outcome, before actual start, to allow users observe the "what if" results using their input values (eg: "what if" the nitrate is too low / how does it impact outcome).*

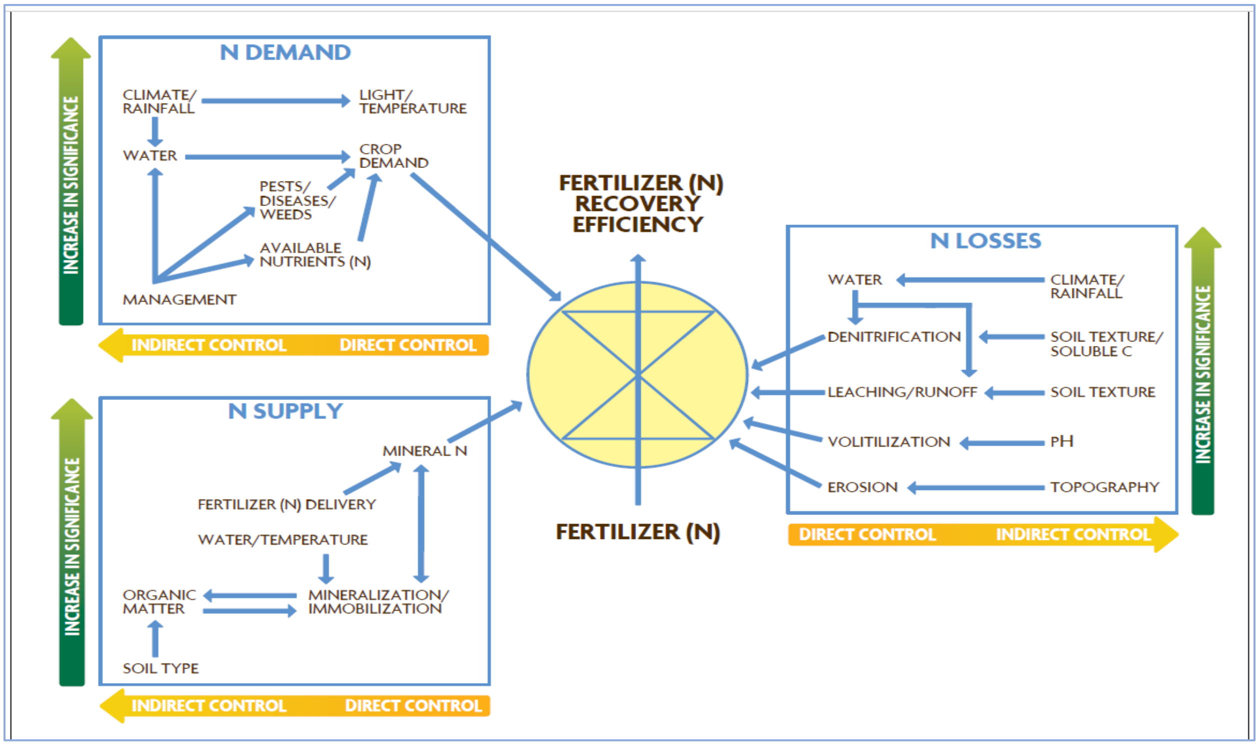


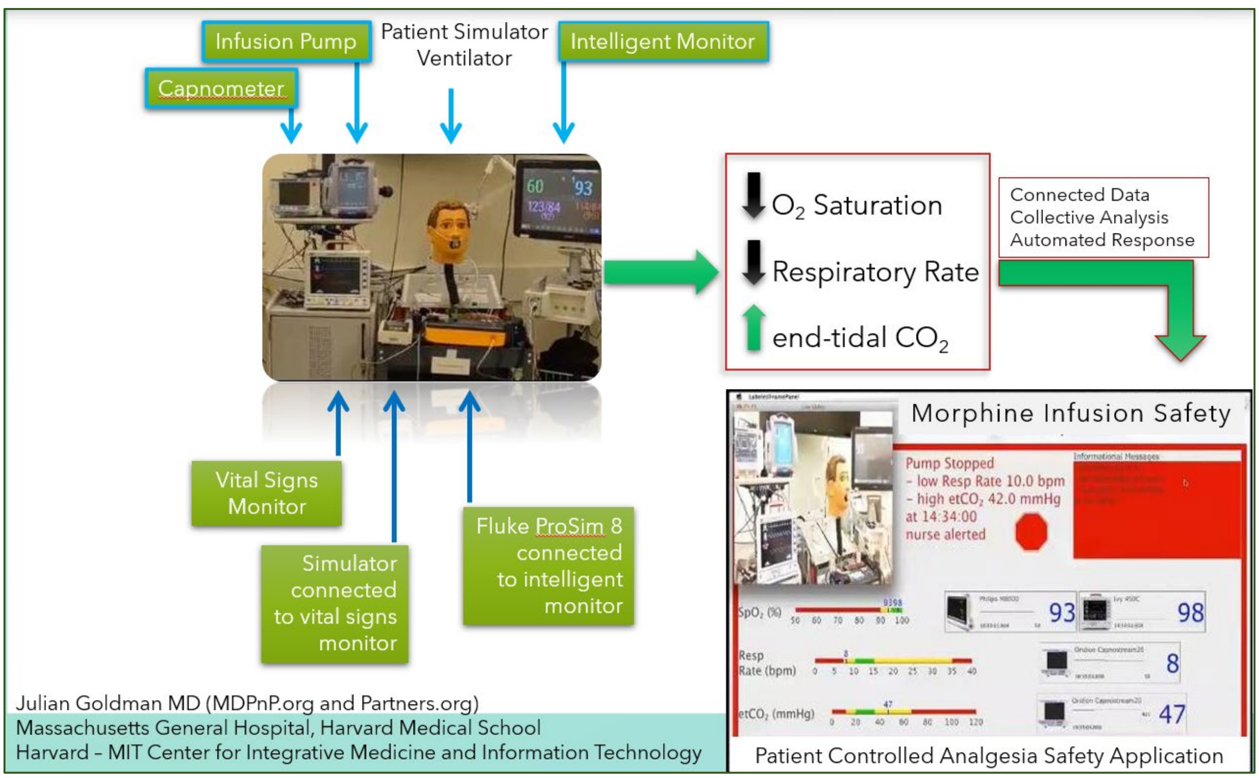



Applying retrosynthetic model building to complex adaptive systems: Can we practice what we preached?

In a sharp turn to reality, let us leap from quantum mechanics and quantum chromodynamics to explore the practicality of our idea in the context of a complex system (nitrogen use in agriculture) and a tiny slice of a post-surgical scenario known as patient-controlled analgesia (PCA) in healthcare (bottom panel).



In these use cases (on the previous page: agriculture and healthcare), can we transform our idea of creating mathematical models or schema to represent the two scenarios? The model may serve as an "engine" operating in the background of an application where data and users are interacting through UIs.

Using the retrosynthetic approach to analyze the factors that affect efficiency of nitrogen use in ag, it is clear, almost instantly, that the cartoon on the previous age is far too complex, as a whole, to arrive at the "efficiency" metric or outcome (circle in the center of the cartoon, upper panel, previous page). Dissociating the problem into three blocks, shown in the cartoon, reveals that "water" and "temperature" are common elements in all three sub-domains.

This common thread may form a good example to highlight that "water" should become an entity which may be stored in the "library" and "water models" may be searchable using discovery engines (such models are expected to be complex depending on a plethora of features associated with water in agriculture). Discovery of the repertoire increases re-usability when other teams/groups working on related agricultural models/topics (eg carbon cycle, phosphorous cycle) may search, find and use the "water" model created in this instance. Models may be imported (inserted/integrated) directly using common APIs (application programming interfaces) or modified after import to adapt to the task at hand, if the pre-coded model language is interoperable with the language of the current task. This "new" modified model may be uploaded to the library and contributed as "water model version x.y" which will enrich the searchable repertoire. This "crowd sourcing" approach will increase the granularity of models and add to the richness of the library.

Thus, each factor in each sub-domain presents many attributes, dependencies, relationships and ratios which creates the need for sub-component models and further drilling down to their data, weights, rates and flows. For example, in the sub-domain "N Demand" (top, left) we need for sub-component models of "light" and "temperature" which are key elements used in multiple applications and may not be restricted to agriculture only but become part of higher level libraries (once the models are created), similar to (page 31) examples of sound and color described elsewhere with respect to Lego Mindstorm systems.

The model for "light" and "temperature" suggests the need for sensor data to supply data feeds (of course, manual input is also possible). The source description of sensor data should be kept "open" with user exists to add different sources based on specific needs of the application. The latter can happen if the sensor data source and the receptacle/database in the model can synchronize data using common (standard?) APIs.

The model for "crop demand" may be quite different compared to near real-time sensor data source. This may be an instance where the data source may be linked to a "publish/subscribe" data/information feed which may be less dynamic (not real-time) but updates a storage log (time series database) with information which the sub-component model may search, discover, retrieve and use, when and if necessary, to improve the quality and accuracy of the predictive analytics contributed by the sub-domain "N Demand" toward the final goal (in this case, efficiency).

Thus, the mathematical framework appears to become an increasingly formidable task even for small sub-segments of any operation. One must exercise caution in their enthusiasm for model building because discrete mathematical functions may often fail to capture the intrinsic non-linearity of complex adaptive systems which are rather continuous than discrete. Values, weights, data and information may operate within ranges and create a "push-pull" dynamic scenario which leads to the outcome, perhaps not as optimal or as precise as mathematicians desire but a rational "approximation" in a "zone" where >80% of the problems may be addressed/solved rather than trying for that perfect fit for 95% or 99% of the cases. Outcome of mathematical models, the frameworks they create and the schemas which will act as scaffolds for data, must act in concert to harmonize dynamic re-optimization for continuous complex systems, in flux, almost always. More complex scenarios (next page) require further deconstruction to use retrosynthetic analysis as a tool.

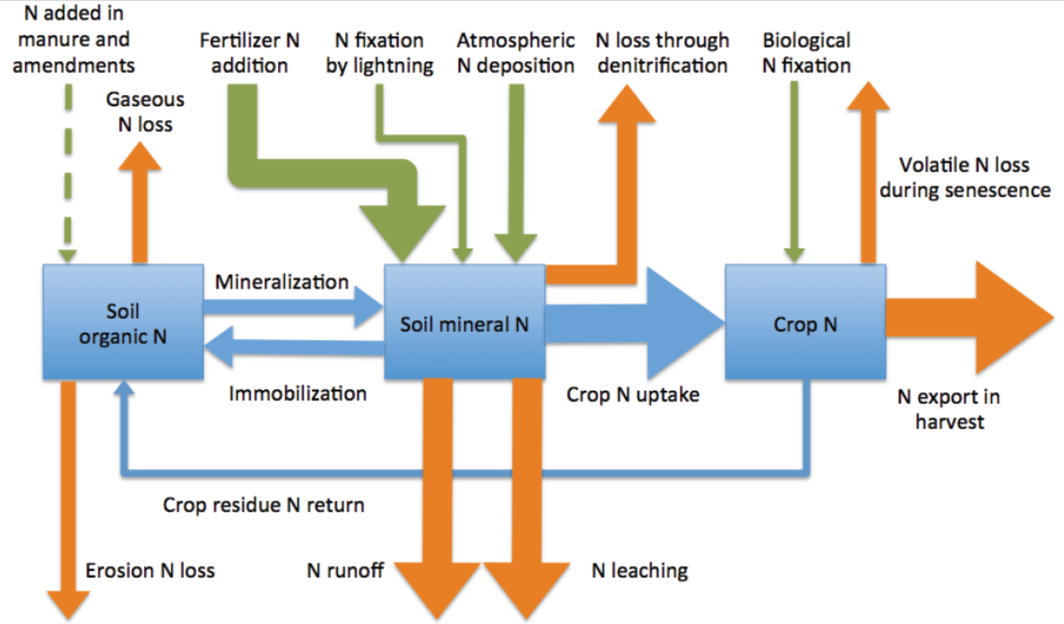



FIGURE 7-A-4 Hypothetical N stocks and flows for a cropping system using mainly mineral N inputs. Boxes representing N stocks and arrows representing N flows are not drawn to scale.

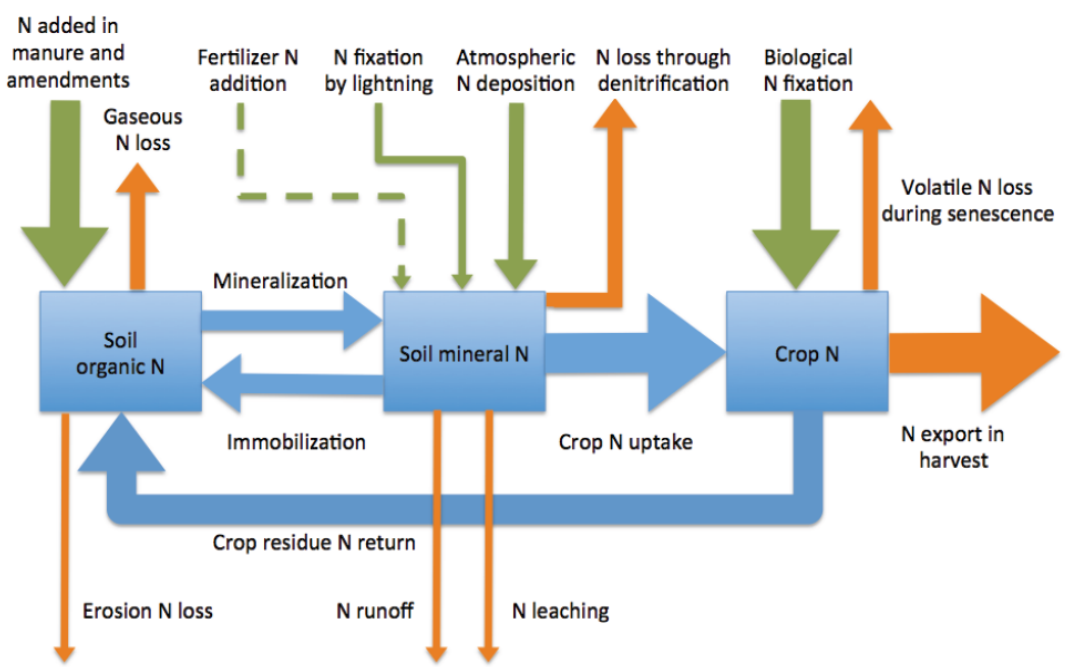



FIGURE 7-A-5 Hypothetical N stocks and flows for a cropping with low reliance on mineral N fertilizer, but with emphasis on biological N fixation, manure and organic matter amendments, cover crops, and perennial crops. Boxes representing N stocks and arrows representing N flows are not drawn to scale.

*Framework for Assessing Effects of the Food System* (2015). National Academies Press. http://nap.edu/18846

Having encouraged the creation of mathematical models it will be remiss not to point out that rigid frameworks and complex adaptive systems are not synonymous. The outcome from mathematical models must be linked to data sources and models within sub-components (of components which are part of the sub-system and sub-domain) at the granular level contribute, often in a hierarchical fashion, as a small part of the system performance. The "systems" level performance of complex systems (Nitrogen in agriculture or PCA in hospitals) are not "points" in terms of performance but a fabric with flexibility, ranges and tolerances which represents a continuum of adaptation, the hallmark of complex adaptive systems.

This need for certain systems to accommodate volatility is not a trait common to mathematical models yet these frameworks, at every level, must allow its operation to deal with levels of "fault tolerance" in order for such a schema to be a relevant/worthy contributor to systems performance.

Data, information, value and action (DIVA) are all connected and related but weighted differently because their semantic distances are subject to change based on the context and/or specificity of the domain, application and needs of the end user. Mathematical models from the very granular up to the systems level must be viewed through the lens of DIVA and this perspective allows an old wrinkle to raise its astute head: are equation based models (EBM) of mathematical frameworks deemed copasetic as a schema for adaptive system performance? The critics of EBM will point out that software agent based models (ABM) are one solution. In reality, it may be a push-pull system where EBM and ABM must co-exist in a dynamic balance to support the "continuum" of adaptations that complex systems must achieve in order to survive, successfully.

For example, in a specific instance of glucose metabolism, the equation based model may be the only model required to account for the outcome from hydrolysis of disaccharides, sucrose and lactose. The kinetics of enzymatic hydrolysis of each molecule of sucrose will produce exactly (mathematically) one molecule of glucose and one molecule of fructose. With the same mathematical precision, one molecule of lactose, if hydrolyzed, will generate one molecule each of galactose and glucose. In this scenario EBM = 1 and ABM = 0. In another instance of glucose metabolism, the concentration of glucose in the blood (humans) can vary from 60mg/dl to 600mg/dl. The extreme values, will present a series of (fabric of) pathophysiological dysfunctions but as a "whole" at the human level (systems performance) the outcome is unlikely to be death, at least, not immediately. In this scenario the inflexibility of mathematical frameworks makes them almost useless for any decision support system (not a place for DSS but humans in the loop). In this scenario, EBM = 0 and ABM = 1.

Taken together, mathematical models are not a panacea but an essential element in our elusive quest for systems-level optimization. In fact, by definition, optimization may be an incorrect descriptor because unequal coalition of many sub-parts may lead to a sub-optimal level of compromise which addresses >80% of the system needs/attributes (rather than attempting 95% optimization of one system at the cost of another).

Linking data sources and sinks to substantiate models is the "devil in the detail" in this endeavor. Which data and how many attributes/dependencies should/must be included in the model is a task where we need human resources with very deep granular knowledge of sub-domain specific processes and the myriad of interrelationships between processes (internal/external) which can influence the outcome. At the same time, the model builder and the process expert must, first, select the rate limiting processes key to the model before fine tuning the system. The team must have the wisdom to stop before they start *gilding the lilly*.

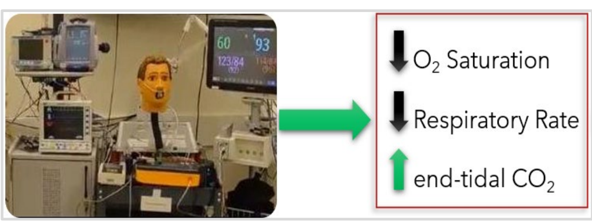
Orchestrating the sub-component models at the component and sub-system level is another massive task where the value of the outcome from the ***synergy*** will far outweigh engineering excellence of integration. Feeding sub-system data/information/outcome to the system level synergistic integration engine will be of value if the predictive analytics from the system is useful in the real world or closely mimics observed values. Titration of the system parameters (fine tuning models) based on observed versus predicted values may offer some "learning" experiences (machine learning?) and over time may improve system performance / value.

The value of these models, irrespective of their forms, figures or limitations, with respect to the mobile tools, workbenches and dashboards, are immense, if we must find ways to empower the non-expert user to adopt and benefit from data-informed decision support systems (DIDS). The daunting complexity of even basic systems (for example, efficient use of nitrogen for optimization of crop production in agriculture) questions whether these ideas are just pseudo-intellectual froth on the virtual cappuccino or tempest in a teacup or may be limited to pilot implementations either because of the unknowns when integrating with the "environment" (part of the PEAS paradigm, page 27) or due to introduction of bias, inherent in PEAS/DIDS.

In our next example, the post-surgical scenario (PCA on page 32, bottom) is an infinitesimally tiny healthcare slice with remediable options to mitigate mortalities due to overdosing of analgesics (eg: use of morphine) to reduce post-operative or post-obstetric pain, especially in patients with low pain threshold.

This example of PCA highlights the need for interoperability between and ***portability*** of models from one system or application to another (open APIs). For biomedical experts, "mobility" of models (frameworks and schemas created by humans) may be described by digressing to discuss biomimicry due to transposons or mobile genetic elements capable of shuttling between genomes, as an example of portability of models. These bio-vehicles can deliver tool kits which may be used for genetic modifications and/or transference or conferring immunity, even between species. One interesting instance of portability of tools in the biological kingdom is the "shuttling of defense cargo" where plants deliberately (by design) send small RNA (sRNA) to silence the virulence in a pathogen, an innovative disease control model worth exploring for response to pandemics, if we can delineate the factors involved in mobile molecular genetics of virulence.

In the PCA cartoon, only three critical values are indicated (below): blood oxygen saturation (SpO2), respiratory rate (RR) and end-tidal carbon dioxide (etCO2). These three are the tip of proverbial iceberg. A plethora of other factors and sub-factors (underlying comorbidities) can be critical in determining the suitability of of PCA and calls for "whole" patient status monitoring using ***integrated information platforms.***

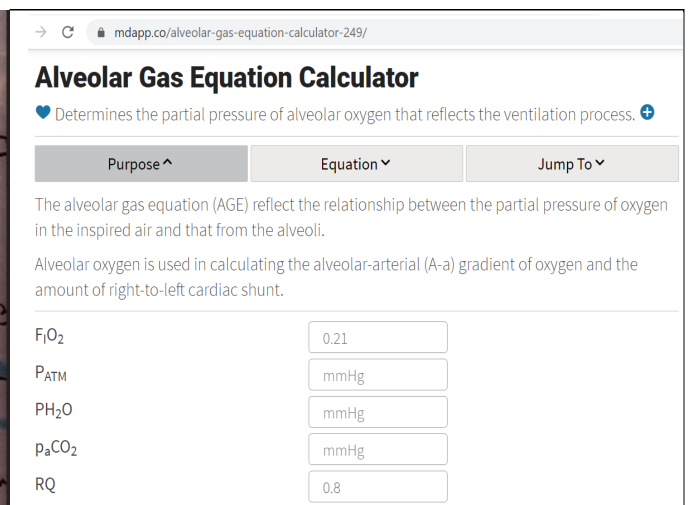



US opioid crisis (~100,000 deaths in 2020 may be attributed to opioids including heroin, codeine, hydrocodone, oxycodone and morphine) suggests the need to re-evaluate administration-approved patient safety protocols and medical device operations/alerts to reduce the risks from PCA.

Opioid-induced episodes of bradypnea (abnormally slow breathing rate: adult respiratory rate (RR) is 12-20 breaths per minute) and desaturation (lowering of blood oxygen saturation) can escalate to respiratory depression (RD). Low RR (RD) may require rescue (ROSC, return of spontaneous circulation) but in-hospital cardiopulmonary resuscitation is successful in fewer than one in five patients. Hence, careful planning to administer PCA based on patient history and stringent synergistic integration of medical devices and device data are necessary for real-time patient monitoring. Reducing risks due to false alarms from add-on/*ad hoc* medical devices may reduce mortalities/morbidities (deaths) contributed by the lack of medical device interoperability in the healthcare domain). **Different medical devices** involved in this (tiny) PCA scenario include: [a] CPOX or continuous pulse oximeters (safety) for oxygen saturation [b] respiratory rate measurements have evolved from the pneumotachometer to spirometry for standard pulmonary function testing in a clinical setting to continuous measurements which may be reliable (in future) using sensors and [c] capnography instrumentation to measure end-tidal carbon dioxide (pressure in mm Hg). Instrument manufacturers sell devices and often ignore how the data must be integrated for patient safety platforms to avoid death and anoxic brain injury from unrecognized postoperative respiratory depression (PORD).

DIDS-based outcomes are essentially about reconstruction of the synthons of the retrosynthetic approach which provides value and specificity of actionable information. The deconstruction of synthons to create mathematical models and frameworks must be a rigorous endeavor yet must leave room to adapt. The use of instrumentation and devices in these processes are essential. Great deal of effort is invested by device manufacturers to "streamline" the process by which the end-user inputs (or input data) are readily converted into charts and graphs in the form of predictive analytics. Not always but in general the problem is with the embedded middleware (software devices) which process the input and generates an output. Middleware is the "back box" which holds a number of so-called workflows which are linked to "model fitting functions" in the software that processes the data and "shapes" the output based on its "fit" with the embedded model.

Limitations of "model fitting" arising from, for example, the dependence on limited number of training sets may lead to generic optimizations/extrapolations of "how to fit" and "what to fit" which may introduce errors, corrupt the input data and allow bias of interpretation. One or more of these and other sources of discrepancy makes it unreliable to trust the scientific authenticity of the output and find ample confidence in the significance of the result but these minutiae only matters if the investigators are sufficiently astute. In most cases, the output from the instrument/device is almost viewed as sacrosanct, which we know is not the case specially if one dares to deconstruct the data (for example, exploring raw waveform data). It appears new tools and ideas are leaving calculations to algorithms and "apps" allowing room for artefacts.


MDApp offers a growing collection of medical (apps) algorithms, scores & calculators grouped by specialty.

The unending "app" inventory has added value for commerce and its contribution to healthcare is not without merit, depending on the case. The extrapolation from relationships between data from individuals to model fitting of individual data to generic models based on aggregated training sets, leaves room for ample doubt. In healthcare, depending on the acuity, automated decision support systems or prediction analytics based on "learning" models can easily go awry and even prove to be fatal, unless processes are supervised by human medical professionals and experts in the loop are actively engaged in the healthcare information arbitrage. The source and quality of input data in stand-alone apps may be error prone and lack systems integration.

The reductionist approach to make things "bite" size for "app" stores is a profitable *modus operandi* for grocery store design teams aiming to please the mobile smartphone user in quest of rapid retail therapy. In the real world of therapeutics and health, the risk of apps pandering to the lowest common denominator may not be without consequences. If unrecognized by low-skilled untrained workers, there may be fatalities.

Equations are essential in mathematical models which are at the heart of the "fitting" engines that run in the background of many apps which ingests data to provide data-driven outcome. Often, if not always, the data-driven outcome is of a poorer quality and less dependable if compared to data-informed information. But the "cheaper and quicker" app-economy is also "dirtier" and perhaps that is one reason which contributes to the not-so-subtle distinction between data-driven outcome versus data-informed information, which not intended to be a semantic jugglery of buzz words in the context of prediction analytics. It is this distinction which becomes far too important when lives are at risk if interpretations and extrapolations are incorrect.

For our discussion of PCA, what happens in the alveoli of the lungs due to administration of the opioid (morphine) is of specific concern with respect to respiratory depression (RD) and potential for fatal consequences as a result of PORD. Thus, accurate composition and partial pressures of alveolar air are super critical. It is pivotal to recognize with unambiguous clarity that the alveolar air is most likely not in a state of equilibrium in a post-operative or post-obstetric patient. Calculations and standards relating to diffusion and exchange of inhaled/exhaled breathing gases (air) at the alveolar-capillary unit may not ***fit*** the physiological status predominant in patients who are under PCA. This distinction which becomes excruciatingly important when PCA patient data are evaluated using frameworks, models, apps based on data (mass balance, steady-state equilibrium, cartoon below) from normal individuals where values (normal range) of parameters may be drawn from equilibrium phase of gaseous exchange. Hence, the cheaper and quicker point-of-care easy-to-use app may soon transmogrify to become dirty and deadly due incorrect interpretation and extrapolation.

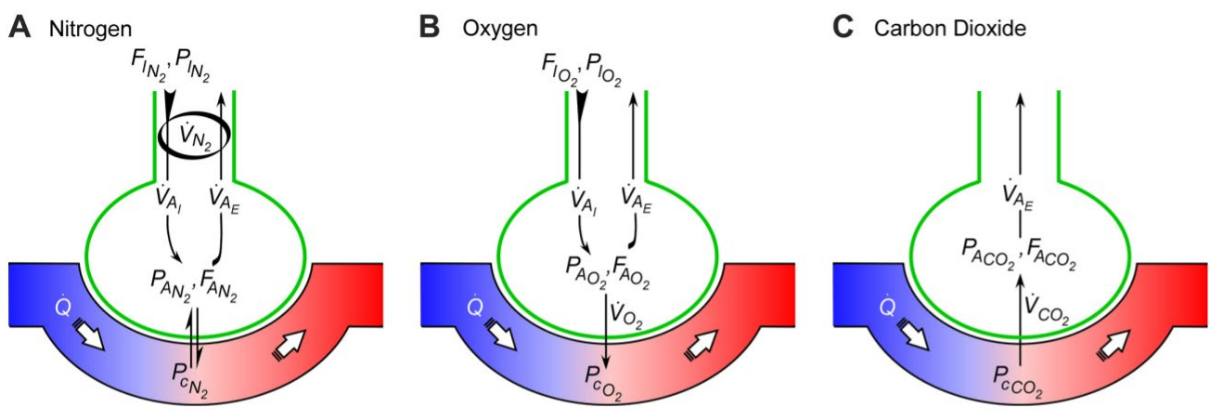
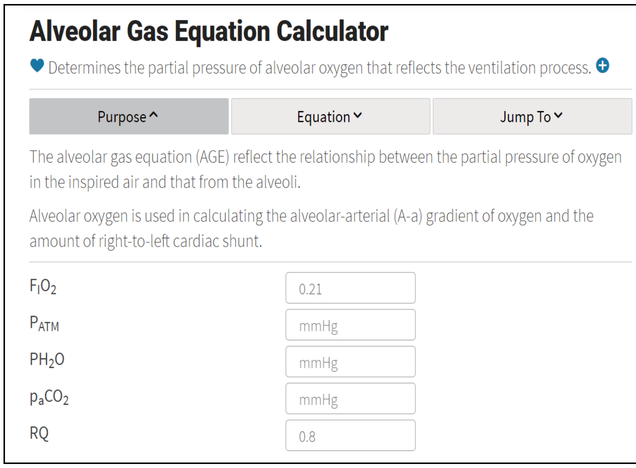


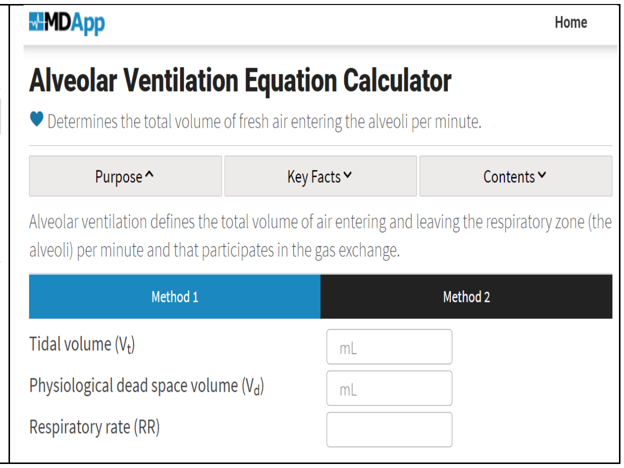

Fig. 1. Mass balance diagrams for $N_2$, $CO_2$, and $O_2$ in a representative alveolar-capillary unit. The principle of mass balance implies that, at steady state, the rate at which each gas enters the alveolus is equal to the rate at which each gas exits the alveolus. *A*: the rate $N_2$ is inspired equals the rate it is expired, denoted by $\dot{V}_{N_2}$. There is no net movement of $N_2$ between the alveolus (green outline) and the associated capillary (*c*). *B*: $O_2$ enters the alveolus by inspiration and leaves by two routes: *1*) net diffusion into the capillary, denoted by $\dot{V}O_2$, and *2*) expiration. *C*: $CO_2$ present in inspired gas is negligible. There is net diffusion of $CO_2$ from the capillary into the alveolus, a quantity denoted by $\dot{V}CO_2$, which equals the rate of $CO_2$ expiration. See Table 1 for definitions of terms used.

| Ventilations and Rates | | Other Terms | |
|---|---|---|---|
| Term | Definition | Term | Definition |
| $\dot{V}A$ | Alveolar ventilation (L/min) | VT | Tidal volume (Liters) |
| $\dot{V}A_I$ | Inspired alveolar ventilation (L/min) | VD | Dead space volume (Liters) |
| $\dot{V}A_E$ | Expired alveolar ventilation (L/min) | RR | Respiratory rate (1/min) |
| $\dot{V}E$ | Minute ventilation (L/min) | RQ | Respiratory quotient |
| $\dot{V}D$ | Dead space ventilation (L/min) | $FI_X$ | Inspired fraction of gas *X* |
| $\dot{V}_{N_2}$ | $N_2$ ventilation (L/min) | $FA_X$ | Expired (alveolar) fraction of gas *X* |
| $\dot{V}O_2$ | Rate of $O_2$ consumption (L/min) | $PI_X$ | Inspired pressure of gas *X* (mmHg) |
| $\dot{V}CO_2$ | Rate of $CO_2$ production (L/min) | $PA_X$ | Alveolar pressure of gas *X* (mmHg) |
| $\dot{Q}$ | Perfusion or cardiac output (L/min) | $Pa_X$ | Arterial pressure of gas *X* (mmHg) |
| | | PB | Barometric (atmospheric) pressure (mmHg) |

Millennials growing up on a diet of short-cuts, videos and apps tend to gravitate to user-friendly user interfaces. Such habits accelerate through undergraduate years and spillover into medical school, residency and fellowships. Using apps in the medical profession has benefits as long as the medical professional is well versed in the foundation and has the wisdom when not to use the app or seek additional data/confirmation.

Pulmonary physiology at the heart of PCA depends on the alveolar gas equation (oxygen equation) and the alveolar ventilation equation (carbon dioxide equation). Medical students may not be versed in the derivation of these equations. The apps shown below use standard equations and published normal ranges. The principle of mass balance which is relevant to physiology of multiple organ systems, is also applicable to pulmonary physiology. Relative rates with which substances (chemicals, gases, water) enter and exit these systems are subject to multiple conditions which can vary between individuals. Steady-state equilibrium occurs when the rates of entry and exit are equal under "normal" conditions. If existing tools are computed based on mass balance then one must question whether steady state equilibrium is applicable to PCA.




**Ideal Alveolar Equation**

$$\frac{V_D}{V_T} = \left(1 - \frac{863 \times \dot{V}\,CO_2}{\dot{V}_E \times Pa\,CO_2}\right);$$

**$PaCO_2$ prediction models**

$$PaCO_2 = 5.2 + 0.82 \times P_{ET}CO_2$$

$$PaCO_2 = 5.5 + 0.90 \times P_{ET}CO_2 - 0.0021 \times V_T$$

**$PaCO_2$ estimated from predicted $V_D$**

$$\frac{V_D}{V_T} = \frac{Pa\,CO_2 - PE\,CO_2}{Pa\,CO_2}$$

**$V_D$ prediction models**

$$V_D = 64.56 \times V_T + 138.73$$

$$V_D = 0.077 \times V_T \div 138.4$$

$$V_D = 0.049 \times V_T + 1.54 \times \text{weight}$$

$$V_D = 0.285 \times V_T - 64$$

The holy grail of physiology is the maintenance of homeostasis. For pulmonary physiology it means the parameters for steady state equilibrium must optimize/maintain gaseous exchange within ranges suitable for normal functions and states of exercise (activity). Gas exchange in the alveoli is a sub-part of the recorded gas exchange in the lungs because conducting airways (connecting air passages) do not have gas exchange potential and referred to as "dead" space volume ("wasted" breath) denoted by $V_{D.}$

In the table (previous page), $V_D$ was calculated via 'ideal' alveolar equations, whereas $PaCO_2$ or $V_D$ models were based on end-tidal $CO_2$ tension ($P_{ET}CO_2$), tidal volume ($V_T$), and/or weight. Breathing faster or deeper can enhance gas exchange while rapid shallow breathing tends to be less efficient at gas exchange. This is important in PCA where standard "models" of breathing may not be applicable. For PCA patients with pre-existing conditions like COPD destruction of alveolar walls can result in the coalescing of multiple alveoli, giving rise to enlarged air spaces that are relatively poorly perfused (e.g., emphysema). In this instance the physiological dead space volume in PCA patients with COPD are a sum of $V_D$ due to connecting air passages (dead space without gas exchange potential) and $V_D$ due to alveolar air that no longer participates in normal alveolar respiratory gas exchange. Thus, the true value of $V_D$ must be taken into account but is it likely to happen in a static app? Similar scenarios may be presented due to pulmonary embolism which may block perfusion to entire alveolar capillary units, which will significantly alter the alveolar gas equations.

The ability to adapt mathematical models and frameworks for specific scenarios is a key part of the inclination to embrace precision medicine. In practice, we are now practicing imprecision medicine if we are at the mercy of software or apps with middleware models incapable of changing the "fit" to fit the patient.

However, fitting the equation to the patient is highly recommended and desirable but certainly not an easy task. The "fitting" to patient will require patient-specific data for multiple parameters (please see the table at the bottom of page 39) which may be difficult (or impossible) to obtain for the post-operative patient. Theoretically, if we had workbenches or user interfaces with each parameter provided in a menu of choices, then an expert may be able to partially adjust/adapt the values but even then it may not be enough for the purpose of improving point-of-care services. To practice pragmatism, we must combine the gains from data and information with respect to the value it may provide in carrying out actions (DIVA) that may improve patient safety and quality of care.

DIDS in principle is a task worth pursuing. DIDS in practice will depend on the choice of applications to determine whether transforming some of these ideas into reality delivers the anticipated real-world value. DIDS in the context of complex adaptive systems needs new thinking, new tools and new paradigms which I have been unable to provide in the context of the examples presented. DIDS in principle and in practice poses many open questions for the next generation of thinkers who may probe with new perspectives/creativity.

The robust mathematical richness of Ansys-esque workbench and the user-friendly Lego Mindstorm, if converged to better extract information from data may even help the Dairy Farmers of America, a real-world application which is too complex to address. The biochemistry of the bolus holds clues to quality and quantity of milk production. By modeling the end-to-end bolus biochemistry (feed to waste) workers can input known attributes/characteristics of their feed by simply using natural language keywords in dialogue boxes and adding feed names/chemistry from a drop down menu. Perhaps the most difficult part is the data the model must ingest with respect to the biochemistry of the feed components, the kinetics of the formation of intermediary products and the rate of release of volatile fatty acids (as a consequence of digestion in the lumen of the cow). Ingested sensor data may be essential and may be a limiting step in data acquisition. If the data is acquired and if there is information in the data that the model can use, then the model may churn out a series of predictions which users can choose to modify by re-titrating the system with quality/quantity of food/feed to optimize the outcome of value, that is, milk production. Is this just a pipe dream or a challenge?

## *Part III ◆ Comments*

Mathematical models, frameworks and schema are the core building blocks for any synthon in a retrosynthetic approach to problem deconstruction and solution reconstruction, even for very challenging problems such as the biochemistry of the bolus.

The complexity of adaptive systems and vastness of the DIDS landscape should not deter us from attempting to create mathematical models in domains of interest, especially those constructs which may provide value to 80% of the global population who are not a part of < 1 billion people in affluent nations.

Perhaps simple point solutions could benefit from these models. Progress may be incremental but can a series/sequence of improvements trigger allosteric advance? Can we explore building models as seeds, planted globally to trigger thinking (differently) with DIDS. Seeds are sterile unless planted (and nurtured).

The cleavage between the immutability of mathematical constructs and the adaptive complexity inherent in biological flexibility indicates that there are "transforming principles" which are in practice in Nature (but may be unbeknownst to humans) which makes it possible for nearly-indivisible molecules (atoms) confer individuality (humans) when synthesized, coalesced, or aggregated in a series of steps guided by fundamental laws of physics, chemistry and biological evolution, naturally.

The fact that the mathematics and physics of temperature (heat and cold) are biologically interpreted by the nervous system is evident when we perceive heat (ion channel TRPV1[1]) or cold (ion channel TRPM8[2]). This seminal research[3] encourages us to build mathematical models because we recognize that bridges exist to ferry between mathematically rigorous laws of physics and (laissez-faire) biological systems. Uncovering the mechanisms to enable suitable application-specific balance between immutability and flexibility may be one way to build ramps to access that hypothetical bridge of knowledge. The bridges may be simple for point solutions or may start out as haphazard reality if we attempt to use DIDS in the context of digital swarms[4].

The principles and practice of DIDS is a call to harness the power of mathematical rigidity in the service of society through case-specific interpretations. Any branch of decision science, for example DIDS, which must involve humans, must also find tools and technologies to dissect, curate and synthesize the convergence of rational with the occasional irrational[5]. Even while immersed in the quagmire of rules and tools, one must remain creatively cognizant of the fact that the inevitable irrational (non-deterministic?) accompanies the rational (deterministic?) not only by chance but also by choice, the irrational[6] choices[7], made by human design.

REFERENCES

[1] Caterina MJ, Schumacher MA, Tominaga M, Rosen TA, Levine JD, Julius D. *The capsaicin receptor: a heat-activated ion channel in the pain pathway.* Nature. 1997 October 23; 389(6653):816-824. Doi: 10.1038/39807 https://pubmed.ncbi.nlm.nih.gov/9349813/

[2] McKemy DD, Neuhausser WM, Julius D. *Identification of a cold receptor reveals a general role for TRP channels in thermosensation.* Nature. 2002 March 7; 416(6876):52-58. DOI: 10.1038/nature719
Peier, Andrea M., et al. "A TRP Channel That Senses Cold Stimuli and Menthol." *Cell* 108 (5) March 2002, pp. 705–715. https://doi.org/10.1016/S0092-8674(02)00652-9

[3] https://www.nobelprize.org/prizes/medicine/2021/summary/

[4] Bonabeau, Eric, et al. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, 1999. ISBN-13 978-0195131598 http://docshare04.docshare.tips/files/20663/206639475.pdf

[5] Kahneman, Daniel. *Thinking, Fast and Slow*. Doubleday Canada, 2011. http://dspace.vnbrims.org:13000/jspui/bitstream/123456789/2224/1/Daniel-Kahneman-Thinking-Fast-and-Slow-.pdf

[6] Ali, Mohsin and Manoranjan, Branavan (2013) *On Vaccines and Irrationality*. McMaster University, Canada. https://journals.mcmaster.ca/meducator/article/view/827/794

[7] Pollitt, Katha. *The Age of Irrationality*. September 2021. *www.thenation.com* https://www.thenation.com/article/society/covid-denial-irrational/