Digital in 4D

COMBINE

Telecommunications (COM) + Bio (B) Info (I) Nano (N) Ecology-Environment (E)

Shoumen Palit Austin Datta

Digital in 4D • Page 1 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u> This blog is dedicated to the memory of Mrs Sheila O'Sullivan. If diagnosis and treatment of her breast cancer, had access to these tools, there would be so much reason to celebrate her life, rather than her death anniversary, on 25th September.

Digital in 4D • Page 2 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u>

Detection, Data, Diagnosis, Decision

Digital in 4D • Transformation of Public Health, Environmental Health, Healthcare & Preventive Medicine

Dr Shoumen Palit Austin Datta, Massachusetts Institute of Technology (MIT) and Massachusetts General Hospital, Harvard Medical School (HMS)

PREFACE

The ideas and suggestions in this blog are not uncommon. This is neither an unique proposal nor a report about a new widget or novel technology. The strength, and purpose of this blog, is to seek convergence. Suggested scenarios, highlights the synthesis and confluence of tools, without any new inventions. The suggestions offer room for innovation, social business, economic progress, plethora of entrepreneurial opportunities, ways to improve quality of life and lift many boats.

Suggestions in this blog may not provide any final solution, it is a compass to create solutions, far beyond the scope of this discussion. The primary focus is on detection and orders of magnitude we may gain, in terms of sensitivity of detection, from micro-arrays vs nano-arrays. Unless we can detect, we can neither measure nor diagnose. Unless we detect long before the event precipitates, the actions and analytics necessary to trigger predictions for prevention, may be moot and sterile.

PROBLEM SPACE - SOLUTION PORTAL

US DHS CBP officer at Boston Logan Airport Global Entry kiosk hands me a tube, and instructs me to spit inside the tube, *after* I take it over to the attendant, in the holding area. I am returning back to the US, after giving a talk at the Internet Forum in Goma, DRC. During the event, I lodged at the Lake Kivu Serena Hotel in Gisenyi, Rwanda. To re-enter the US, I must be tested for Zika virus, first detected in 1947 in the Ziika forest, Entebbe, Uganda, near Lake Victoria. The forest is more than 500km east of the locations I visited. Goma and Gisenyi, are both located on the shores of Lac Kivu, but in different countries, Democratic Republic of Congo and Rwanda, respectively.

My sputum is diluted, and a miniscule drop is applied (using a micropipette), on a flash drive with a micro-USB. The attendant hands me the flash drive and suggests I insert it in my phone, *after* activating the DHS Global Entry app. The app display reveals a distinct red button. The dialog box prompts me to insert Foreign Agent Detection (FAD) flash-sensor drive. I insert the flash drive. In about 30 seconds, the red button turns blue. It reads: "SUCCESS! Foreign agent was not detected. Welcome to the United States of America." I return the drive. On inquiry, I learn the assay[¹] used gold nanostars conjugated with antibodies to detect Zika non-structural protein 1 (NS1). The sensitivity to the biomarker was improved using surface enhanced Raman spectroscopy (SERS). The assay did not detect NS1 in my sputum. Hence, I was free to walk out of the secure area. A DHS agent re-checks my app for the blue button. I proceeded to baggage claim to pick up my bag.

ABSTRACT

This blog analyzes tools and technologies, to collect data, and the convergence, to transform this scenario (above) to reality. This is one of many potential^[2] applications in healthcare. However, 'detection' is useful in agriculture, security, environment, manufacturing, etc.

INTRODUCTION

The public health scenario, in practice, may prevent epidemics and pandemics. It may help identify the point of origin, pattern of dissemination and spread of infection, due to a pathogen.

The sensor, in the scenario, can sense any material which can be engineered[³] and detected, within the limits of what can be sensed, if the limits of sensitivity, are useful for the purpose. Detection of air-borne pathogens[⁴] or toxins[⁵], arsenic[⁶] in water, pesticides[⁷] in soil[⁸], bacteria in food[⁹], methane[¹⁰] from animals, gas[¹¹] in the environment, nerve agents[¹²] in conflict zones[¹³], chemical[¹⁴] pollution and glucose[¹⁵] in blood[¹⁶], are a few, of the many use cases.

Transmission of the outcome (data resulting from action of the sensor) must generate a wireless signal^{[14}], which can be captured^{[11}] in an app^{[17}], is essential for connectivity and communication. It allows us to thread the "digital needle" in order to move the needle, and take advantage of the wave of digital transformation, and benefit from the imminent tsunami of digital economics^{[18}]. Various efforts target only part of this end-to-end scenario (detect, connect, transmit, decide).

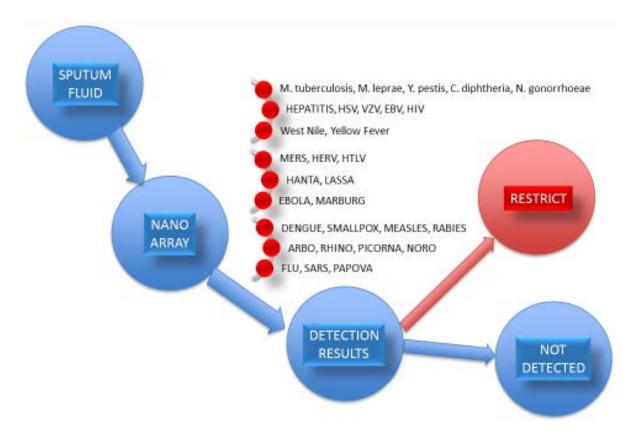
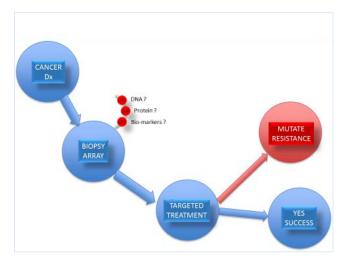


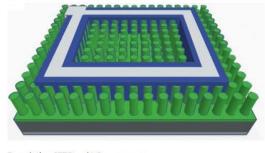
Figure 1: Simple illustration of public health application. The basic research, in principle, and the tools to connect the data, to an application, on a mobile device, are available. What are the key barriers to implementation? Growth in air travel and shipborne trade, are reasons for appearance of communicable disease pandemics, disease vector invasions, vector-borne disease movement[¹⁹] and infected people. A risk mitigation strategy, without emphasis on detection, is blasphemous.

Digital in 4D • Page 4 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/11021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u>

APPLICATIONS – INDIVIDUAL STUDIES



← → C 🔒 Secure | https://ftoor.lab.uiowa.edu/research



Silicon Nanowire based Optoelectronic Biosensors

Current cancer detection methods are insufficient for early detection; at present, cancer detection is performed using a combination of physical examination and CT scans or MRIs. Early detection both lowers the cost of treatment and increases survival rates. Nanowires (NWs) are effective sensing structures due to their large surface area to volume ratio. However, contacting the NW arrays is challenging. Our silicon (Si) NW optoelectronic biosensor is made by a bundle of vertically oriented NWs, allowing us to electrically contact millions of NWs per cm² simultaneously, compared to 10's of NWs in other state-of-the-art NW biosensors.

☆ 🖸 🖸 🚳 🖻

The project is in collaboration with a Massachusetts-based start-up, Advanced Silicon Group, and Professor Aliasger Salem's lab at UIowa, and funded by the National Science

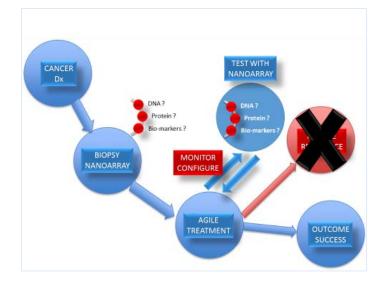


Figure 2: Monitoring bio-markers may help prevention, and treatment, in progress. Prediction of potential risks can reduce morbidity and mortality due to common diseases (stroke, congestive heart failure, COPD). What often happens during cancer treatment (upper panel) may be reduced by monitoring biomarkers with high sensitivity nano-arrays, to alleviate multi-factorial resistance.

Digital in 4D • Page 5 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u>

Foundation STTR and I-Corps programs.

APPLICATIONS – POPULATION STUDIES

If Twitter (Figure 3) can be used as a crowdsourcing app to map hate^[20], imagine the immense benefit to public health (and epidemiology) if we can crowdsource sensor data, about air-borne pathogens or arsenic in water (for example, Bangladesh). Using the appropriate sensor, attached to a mobile device (microUSB flash drives with swappable slots for analyte-specific sensors), one can generate global maps of analyte distribution, and track and trace, any molecule, anywhere, anytime, if the sensor and signal transmission is compatible with smartphone microUSB platform.

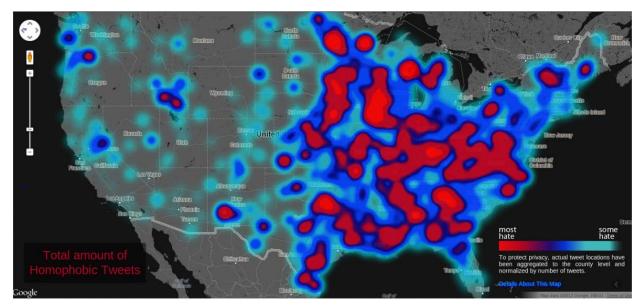


Figure 3: Potential of crowd-sourced sensor data, in convergence with GIS, to map harmful agents.



Figure 4: Converge: crowd-sourced data and swarm^{[21}] intelligence^{[22}] for data analytics?

Digital in 4D • Page 6 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u>

LOCAL DECISIONS USING GLOBAL PLATFORMS: SENSE-TRANSMIT-COMMUNICATE-RESPOND

Project Ara from NK Labs^[23] (Cambridge, MA), failed to ignite the spark, at the hands of Google^[24] but its potential, if viewed with 'new' eyes, may be innovative for sensor applications.

The swappable modular architecture may be the preferred "housing" infrastructure for the sensor and serve as a platform. The form factor of this *housing platform* is a standard (USB) connector based flash drive-like unit which can transmit, compute and communicate with smart devices, for example, phones, tablets, automobiles, cameras, lawn mowers, refrigerators, washing machines. For mobility applications, the device must be mobile, but for stationary monitoring, any widget, with a connector, may be adequate, as long as it is functional, within the end-to-end ecosystem.

Computational requirements within the housing platform may need tinyOS (operating system), tinyDB (database), tinyAPI (application programming interface). Seminal work by David Culler^[25] and Howard Shrobe^[26], among others, are available for innovators, to introduce these functions, within this form factor (housing platform). Security, identity, authorization, verification and data privacy are cybersecurity^[27] issues which depends on data de-identification.

Identification of the housing platform, and creating an unique identification of the test subject, in combination with the tool (drive), may use RFID EPC code plus the IPv6 address of the device or combine with biometric id. This may be relevant for travel and immigration.



Figure 5: Any analyte, or pathogen, can be sensed, anywhere. If the sensor, and smartphone, can capture the data, it may be transmitted, anywhere. One assumes the transmission infrastructure exists, is protocol-agnostic and connects or uploads between wired or wireless, nodes or gateways. Examples of molecular diagnostics, illustrated, are benefits from medical internet of things (IoT). It may help preventive medicine, and diffusion of healthcare tools, to remote places in the world.



Figure 6: Arsenic contamination of water^{[28}] leads to severe morbidity, for many, in Bangladesh.

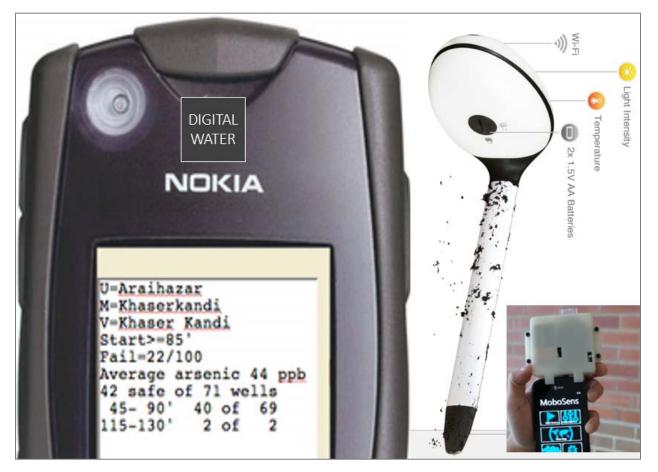


Figure 7: Pioneering work by Nokia. Connecting sensor data, for detection of Arsenic in water, with communications, may be rejuvenated in the era of IoT. Digital transformation, and detection of Arsenic, can help to improve the quality of life for inhabitants of Bangladesh. Millions of people may use a mobile phone, with attached sensor, to test any water, they encounter. The data can be uploaded to the cloud. Dynamic data analytics may provide precision maps of the areas to avoid. It may trigger measures for water safety^[29], calibrate purification^[30] and preventive healthcare.

Unless we can make sense of the data from sensors, in near real-time, the value of detection may be zero. Hence, the importance of telecommunications in this end-to-end ecosystem. Nokia's entry into "digital water" was not triggered by altruism but due to the business case, based on pay-per-use micro-payment business model (discussed in "Digital Enterprise" which is a part of "Digital Transformation" here <u>https://dspace.mit.edu/handle/1721.1/111021</u>). Telco standards and interoperability between protocols, may help or hinder adoption of solutions, in different areas.



Figure 8: Sense and response systems (in discussion) must transmit data to centers or to the edge.

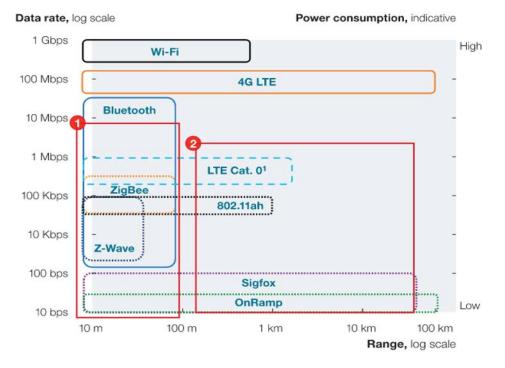


Figure 9: Competing standards and protocols in the short range (1) are an increasing problem. This is the telco segment where "detection" tools are likely to be used. Hence, compatibility and interoperability between protocols and standards, is a pre-requisite for function. The longer range (2) is used for other purposes, eg, smart grid, where there is less competition and more regulation.

Digital in 4D • Page 9 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u> Open standards and use of global standards, for example, Open Systems Interconnection (OSI) model, are important elements in the solutions approach. OSI (Figure 10) characterizes and standardizes the communication functions of a telecommunication or computing system. Its goal is the interoperability of diverse communication systems with standard protocols (Figure 11). The 7-layered OSI design is due to Charles William Bachman III (December 11, 1924 – July 13, 2017). Aspects of the OSI design are similar to ARPANET[³¹] and related networks, in other countries.

Application Layer	IT / ICT		ОТ
Presentation Layer	• XML • JSON	• XMPP • MQTT	 OPC UA MTConnect FieldBus (App Layer Messages)
Session Layer	• HTTP	AMQP	Common Industrial Protocol (CIP)
Transport Layer	• TCP/IP		D ZD Z
Network Layer	LAN / PAN Ethernet 	Ethernet • GPRS/3G/4G/5G • Wi-Fi PAN: Bluetooth BAN / WBAN • LPWAN (LoRA, NB-IoT)	• TSN
Data Link Layer	PAN: BluetoothBAN / WBAN		
Physical Layer	• Lo6PWAN	Satellite: Ka Band, VSAT	

Figure 10: OSI model. ICT WAN (information communication technologies wide area network) standards are bifurcating to high-power/high-bandwidth and low-power/low-bandwidth segments. Most ICT WAN protocols aid in sensing (not in actuating). Operational Technology (OT) is not directly relevant to this discussion but it is an integral part of the 7-layer OSI model.

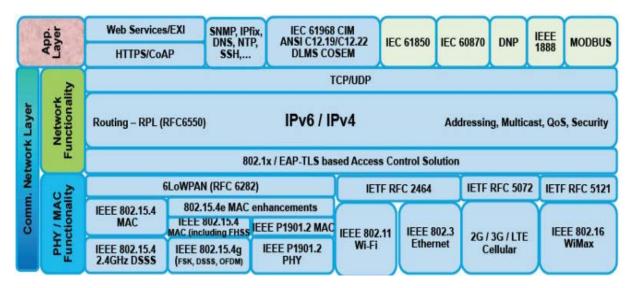


Figure 11: Standards relevant to OSI model. Open standards reference model from David Culler^[22]

Digital in 4D • Page 10 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u>

REVIEWING THE COMPONENTS OF THE SYSTEM – EXPLORING THE ROOM FOR INNOVATION

The scenario presented at the beginning of this blog starts with

- [1] fluid (sputum) applied to
- [2] sensor (FAD flash-sensor drive)
- [3] which communicates via an app (data).

The sample containing the analyte or pathogen or bio-marker is the natural product, to be tested. Innovation is not applicable to this phase. Selecting alternate fluids may not be an innovation, for example, if an invasive test, involving blood were to be used, instead of the non-invasive sputum.

Signal detection, data transmission, storage, transport and visualization through an app, may not be as direct as procuring the components, in the bill of materials (BOM), from an "off-the-shelf" catalogue. However, these steps are used in millions of processes. Information is available for configurable components (for example, tinyOS and tinyDB). Communication systems must use standards and protocols. Apps, APIs, GUIs are generic software products. Taken together, it must converge to generate the result – the data – which must be analyzed, "intelligently" if possible[³²].

Data analytics is of paramount importance. It is fertile grounds for innovation, to fuel decisions, following detection, and diagnosis. Unlocking actionable information from data is the Holy Grail. Data, analytics, AI, and platforms, are brimming with hype. Examples of how the cookie^[33] can crumble^[34] are cautionary tales. Even if you have a hammer, does not mean everything is a nail.

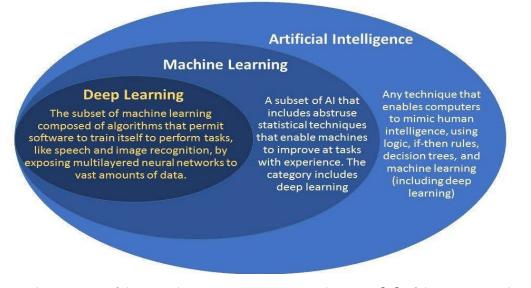


Figure 12: Making sense of data, with respect semantics and context^[35] of data, is crucial to obtaining useful information. Data must deliver value to justify the investment in the system. The application of AI (artificial intelligence) algorithms, and its various manifestations, are a part of the portfolio of approaches to data analytics. Choice of algorithms, the type of solvers and models which may be used in analytical engines, may evolve with the application, or the application may trigger functions based on context, an old idea^[36], reincarnated through SDN^[37]. The nature of the data, transmission speed, bandwidth and the latency the data can tolerate, (before the data or information perishes, in value), may dictate whether the analysis, and response, may occur via the cloud, the fog or mist or at the point of contact^[38], that is, at the edge (as in the DHS scenario).

Digital in 4D • Page 11 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u>

ROOM FOR INNOVATION ?

We are left with the actual sensor. This is a critical point of convergence and demands fusion of transdisciplinary domains. Generating viable analyte-responsive sensors require orchestration of medical parameters, biology of molecules, chemistry of molecules and material, material science, biomedical device engineering, and manufacturing processes necessary to produce sensors.

There is plenty of room for innovation but one must start with asking correct questions. What are we planning to detect? This is where the depth of science, and knowledge about the subject area, is critical. The sensor will be worthless, unless identification of the target is unambiguous, and is relevant, from the perspective of biology and medicine (if applicable), in terms of function.

The next question is the knowledge about the complementarity of the detection process. For example, if we wish to detect an antigen or an epitope of an antigen, then its specific antibody must be available and presented, as the immobile component, on the surface of the sensor.

How many immobile components will be available for interaction on the surface of the sensor? This is where the "game changes" as we move from micro-arrays^[39] to nano-arrays, which may owe its origin to the principle of nanowires^[40], but modified, over the past quarter century. The move to nano will increase the surface area but will that result in a concomitant increase in the sensitivity of detection? Detection ability of novel nanowire sensors are in the picogram per milliter (pg/ml) range and at a 100x reduction in cost compared to other manufacturing processes for silicon nanowire sensors (Marcie Black, Advanced Silicon Group, *personal communication*).

Sensitivity of detection, is one reason why nano-wire sensors (nanoarrays) could change the face of public health and preventive medicine. Coupled with digital transformation, and IoT, as a digital by design metaphor, the scope of digital health, and medical IoT are, truly immense.

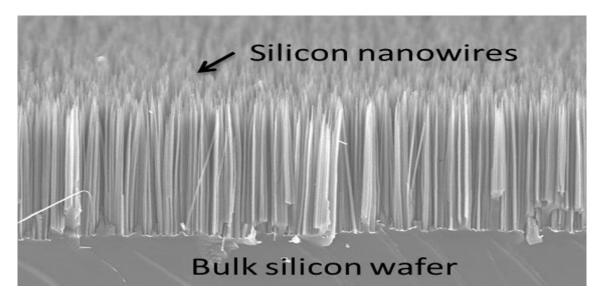


Figure 13: The revolution in detection due to nano-arrays using silicon nanowires. Compare 5-10 molecules on other forms of nanowire devices versus this configuration (above) which has the potential to be far more sensitive due to >1 million nanowires per device (six orders of magnitude increase, per sensor surface). Theoretically, each nanowire can serve as an anchor for an analyte.

Digital in 4D • Page 12 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u> The biochemical challenge is *retaining the functional viability* after attaching the analyte to the nanowire. An assay to measure the amount of attached analytes cannot be used to extrapolate the gain in *sensitivity* due to nanowires. To estimate the gain in sensitivity, we must reproducibly assay the percentage of the attached molecules which are functionally active to detect the target. What is the tolerance for variability in assay conditions with respect to functional sensitivity? Is the nano-sensor suitable for multiple uses or single use? Will the complex formation (necessary for detection?) auto-dissociate to rejuvenate the nano-sensor for another cycle? In the absence of auto-dissociation, can the nano-sensor return (treated to return?) to basal or ground zero state? What is the nature of false positives? Is error correction and/or data curation/cleansing required to obtain the result? What is the detection signal when the analyte detects the target? In what ways this signal can be modulated? How is the signal converted to data? Attempting to answer these and other related questions, provides ample room for innovation in biology/biochemistry.

Can we create standard operating processes (SOP) for nano-arrays, using a modular approach? In this vision, the components, except the target, are pre-fab "kits" from vendors (for example, New England BioLabs). Democratization of the nano-array tools, beyond specialized labs to any clinic or industry, may benefit other domains, for example, exploration of microbiomes[⁴¹].

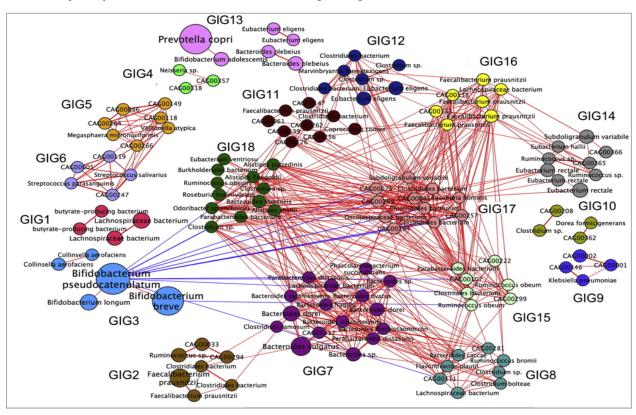


Figure 14: Network of interactions between gut microbia in obese children^[35]. With low cost, routine nano-metabolomic assays, out-patient or primary care clinics, may detect molecules of interest, while they are still below the threshold (picogram or femtogram per milliliter) of disease or dysfunction. Nano-arrays for prevention, in home test kits, can use non-invasive body fluids (sputum, urine, feces) to detect target molecules, in a manner similar to pregnancy tests or blood glucose strips, available from pharmacies (CVS, Walgreens, in USA). Nano-kits may be purchased from brick-and-mortar stores or ordered online (WalMart, Petsmart) or delivered-by-drone to your door-step or window-sill (Amazon, Alibaba), in the near future.

Digital in 4D ● Page 13 ● Dr Shoumen Palit Austin Datta ● <u>http://autoid.mit.edu</u> ● <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> ● Please email to <u>shoumen@mit.edu</u> ● <u>sdatta8@mgh.harvard.edu</u>

CONFLUENCE OF IDEAS

The vast room for innovation, in optimizing sensitivity of detection of analytes, will determine the ultimate power of, and benefit from, nano-arrays. Availability of home nano-array kits in gas stations and convenience stores, can reach far corners of the world, where availability of medical professionals may not need a miracle, but is as rare as a solar eclipse. What was not feasible with microarrays may be possible with nano-arrays in the detection of agents for flu, cholera, plague, TB, malaria, leprosy, meningitis, diabetes, myocardial infarction and so many other diseases, which affect the rich and the poor. If detected, before the threshold, millions of lives can be saved.

Digital transformation catalyzed data collection, enables quantitative measurements, establishing metrics, changing the rigidity of "the model mind set" and driving key performance indicators. In the data driven economy, KPIs must be guided by the *context of the problem* or the environment, rather than feeding models from antiquity, limited by the scope of their equation based structure. In the digital era of near real-time data, the dynamic composability of agent based models and its modular components (parameters, features), must be continuously constructed, de-constructed and reconstructed, to keep pace with the volatility of data. The data must not be made to *fit* the model. The model must fit the observed data, to be useful, at least per the Pareto principle. Data from detection is no longer the "Epic of Gilgamesh" available to the Queen. Digital transformation can make this data accessible to anyone, anywhere and we may access that <u>same</u> information (please see page 5 in "Digital Transformation" here <u>https://dspace.mit.edu/handle/1721.1/111021</u>).

Data repeatedly re-appears to be at the heart of decisions. Suggestions in this blog are "obvious" as well as data dependent. Mobile phones analyzing sensor data can detect tremors related to Parkinson's disease^[42] and display in apps. In the near future, the smartphone, which you plug in the car, to charge, while you are driving, can also inform you if the tire-pressure is low or the car needs wheel balancing – without the automobile's sensor. *Sensors in phones* (accelerometer, gyroscopic sensor) can "sense" the operation and use machine learning (ML) to predict status. Sensors can upload real-time data about emissions and record reduction of pollution by EVs.

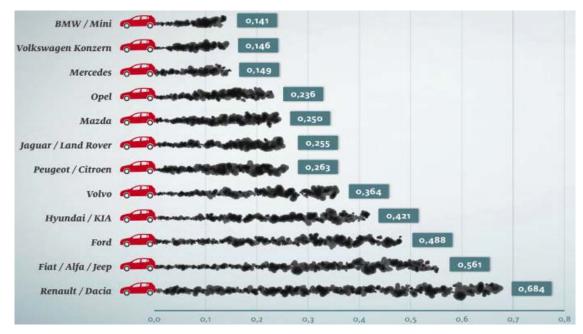


Figure 15: We do not need nano-array sensors to detect NOx (g/km). EU data from diesel cars[43]

Digital in 4D • Page 14 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u> The ability to connect the data to health monitoring or healthcare systems, unleashes multiple benefits, for the patient, individuals, families, local community, national governments and global organizations (PIH, WHO, MSF). Data and data analytics are essential elements in the closed loop abstraction, that is, cycle of data and analytics feeding the decision, and data from implementing the decision, feeding the next round, of the dynamic decision process. Data and contextual analytics, at the right time, may reduce [a] the 43 million cases[⁴⁴] per annum, worldwide, due to medical errors, [b] 99,000 deaths per year due to healthcare related infections and [c] 770,000 injuries and deaths due to problems with medication. The outcome, hospital readmission, costs[⁴⁵] an estimated US\$17 billion, each year. This is the immense value, from connected data.

How do we collect data, how do we communicate it, how we store it, how we transmit it, how to query it, and what to do with it, in order to extract value for the decision process, with or without humans, are questions that must be addressed, case by case, because it is not formulaic. We have come a long way from the stethoscope of 1816. In about 200 hundred years, we can use a mobile smartphone, which can connect to an USB flash drive, with a sensor. It is a sign of the times.

The type of sensors suggested in this blog, are able to track and trace analytes, related to health and environment, among other things, using nano-wires for ultra-high sensitivity nano-arrays. Anybody who doubts "if" it will happen, may wish to pay attention to recent developments in chemistry. Two groups from the same department at the University of Manchester, Chilton & Mills^[46] and Richard Layfield^[47], have reported a dysprosium molecule ^{[66} Dy] with switchable magnetic properties, single-molecule magnet (SMM), with the ability to store a single bit of data.

The stunning chemistry of the dysprosium molecule runs parallel to the thrilling potential for advancing information technology. Molecular data storage using SMMs could handle about 30 terabits of data per square centimeter, or more than 25,000 GB of data (>3 terabytes), on a device the size of a flash drive. That is about 10-fold more than the current (256GB) data storage capacity of Apple's latest iPhone 8, which relies on magnetic nanoparticles.

Cross-pollination of multi-disciplinary approaches, and innovation, may guide us to create new vistas, to analyze analytes, in an attempt to inform the processes, which may improve the health of animals, humans, plants, and the collective health of the interconnected environment.

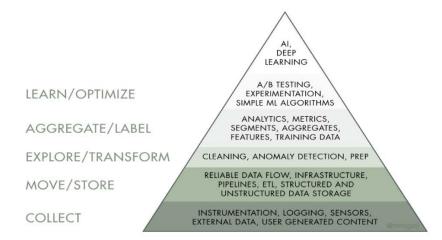


Figure 16: In this blog we have focused on the bottom of the pyramid. "Detection, Diagnosis and Decision" needs data. Nanosensors and mobile devices are tools we may use for collection of data.

Digital in 4D • Page 15 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u>

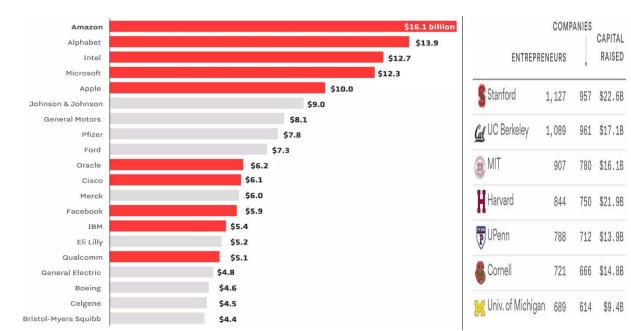


Figure 17: US corporate R&D budgets. Tech companies (red bars) in the Top 20 list, by R&D spending, invested almost US\$100 billion in research and occupy 5 out of the Top 5 positions[⁴⁸]

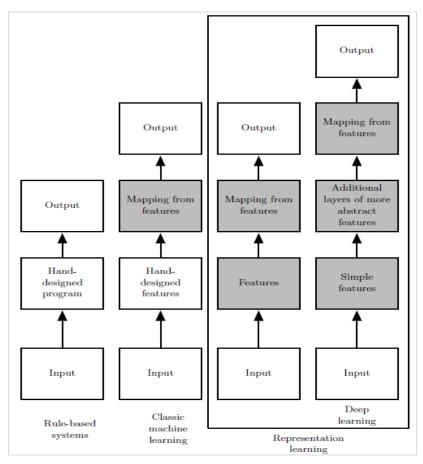


Figure 18: Feature engineering is **key** to AI. Insufficient, noisy, unstructured data, with respect to features, renders these tools impotent. From^[49] Ian Goodfellow in *Deep Learning*. MIT Press. 2016

Digital in 4D ● Page 16 ● Dr Shoumen Palit Austin Datta ● <u>http://autoid.mit.edu</u> ● <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> ● Please email to <u>shoumen@mit.edu</u> ● <u>sdatta8@mgh.harvard.edu</u>

CONCLUDING COMMENTARY

Science and engineering are at the heart of this blog. The identification of biomarkers and analytes, as molecules of interest, is fundamental. Scientists who can transform these ideas into reality, may not care much about the end-to-end qualitative scenario, outlined in this blog. In other words, majority of scientists will not consider reading a "blog" by a non-expert. Engineers who can fabricate the material, to create the nano-wires, will follow in the footsteps of scientists. The engineers in telecommunications, architecture and systems connectivity, will disregard the blog because it deals with "biology" or medicine. The computer scientists and the data scientists (we used to call them statisticians), will ignore the blog because there is "nothing in it about actual data" to discuss. Most business folks and investors will roll their eyes, because they are interested in ROI and business models, not in analytes or nano-wires or tools to save millions of lives. Foundations and "angels" will ignore this blog due to its lack of focus. Publishers of august journals will point out that there aren't any new advances, in this blog. Other publication outlets will ask who is the intended audience, what is the target market? There are so many, so justified, and so good reasons, why this blog was a waste of time and considered junk mail, in in-boxes.

Ian Goodfellow, answering a question from the audience, remarked that the reason why those who tend to delve into deep learning are not sufficiently informed in neurology, and those who are in neurology, are not engaged with AI or deep pursuits in computer science, due to the fact that mathematics is not a priority in biology and biology is not a priority for artificial intelligence. He went on to add, computer scientists abhor learning declarative facts and prefers developing processes from first principles. But, neurology or biology, emphasizes the learning of facts. The chasm between the groups, and the fact that we may be at the level of arthropods (insects, bees), in terms of understanding neural networks, makes the process of AI and representational learning almost entirely, and precariously, dependent on trial and error. The science of AI is still an art.



Figure 19: The panel on the left attempts to fit data and reveals, once again, that common sense is uncommon. Extending that idea to neural networks, the right panel, shows how artificial neural networks (ANNs) are highly vulnerable due to their linear nature. In this example^[50], applying a very small perturbation (+ 0.007) to the dataset (panda image recognition with 58% confidence), can dramatically change the output and emphasize an incorrect answer with very high confidence (image suggests gibbon with 99% confidence). This is due to the linearity of ANN, which is the principal tool for ML, DL and AI. Linearity is an artifact of ANN and not at an intrinsic property of neural networks. From the perspective of human neurology, humans are well equipped to "think back" and go "back" in thought and thinking, often in the presence of significant perturbation. Back propagation is difficult yet back propagation algorithms are used in training deep models.

Digital in 4D • Page 17 • Dr Shoumen Palit Austin Datta • <u>http://autoid.mit.edu</u> • <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> • Please email to <u>shoumen@mit.edu</u> • <u>sdatta8@mgh.harvard.edu</u> In the ANN model, data perturbation by an input of 0.007 completely changed the output, and sharply increased the confidence of the result. Imagine what could happen to medical diagnosis! If using an AI tool for diagnosis, minor perturbations in cerebral metabolic rate (CMR) or GABA (gamma amino butyric acid, a neuro transmitter) or tissue plasminogen activator (tPA), may lead to a diagnosis of idiopathic epilepsy^{[51}], when a normal person or animal, was suffering only from anxiety or stress due to duress. AI is useful but how can it be trusted without humans in the loop?

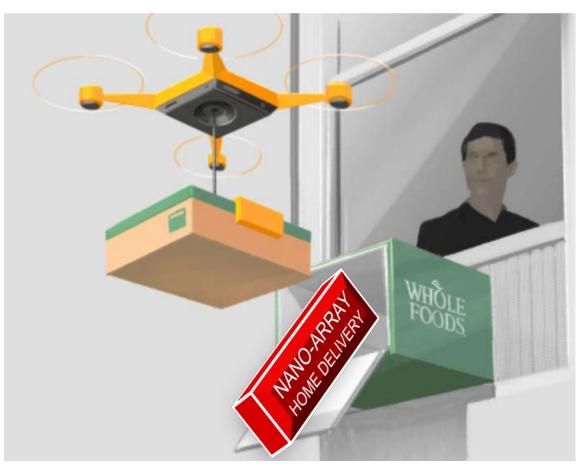


Figure 20: Home delivery-by-drone of nano-array kits. Arriving soon on a window-sill, near you? Measuring Arsenic in Bangladesh or Mercury^{[52}] in Colombia or CEA in cancer, can save millions.

One expected outcome (illustrated in Figure 20), is not impossible, because we do not have to break any rules of physics, to make it happen. The barriers are due to the rules of finance and economics. An investor commented that it will take a major catastrophe in a wealthy nation to mobilize the detection tools, drive the convergence, and for industries, to seek the end-to-end solution. The fact that eleven million children die, annually, due to mainly preventable causes, is not sufficient reason to catalyze the convergence. About 70% of the deaths are attributable to six causes (diarrhea, malaria, neonatal infection, pneumonia, preterm delivery, and lack of oxygen at birth). The fact these deaths are concentrated in 10 countries, mostly, in Sub-Saharan Africa, is the fly in the ointment. If that "fly" from Bamako could fly and trigger a public health havoc in Boston, then, try to prevent angels from investing or foundations from opening the funding gates.

Digital in 4D ● Page 18 ● Dr Shoumen Palit Austin Datta ● <u>http://autoid.mit.edu</u> ● <u>http://mdpnp.mgh.harvard.edu</u> <u>https://dspace.mit.edu/handle/1721.1/111021</u> ● Please email to <u>shoumen@mit.edu</u> ● <u>sdatta8@mgh.harvard.edu</u>

- ¹¹ <u>https://swagergroup.mit.edu/organocobalt%E2%80%93carbon-nanotube-chemiresistive-carbon-monoxide-detector</u>
- ¹² <u>http://onlinelibrary.wiley.com/doi/10.1002/anie.201604431/abstract</u>
- ¹³ https://swagergroup.mit.edu/chemiresistor-devices-chemical-warfare-agent-detection-based-polymer-wrapped-singlewalled-carbon

¹⁴ <u>https://swagergroup.mit.edu/integrated-gas-sensing-system-swcnt-and-cellulose-polymer-concentrator-benzene-toluene-and-xylenes</u>

- ¹⁵ <u>http://care.diabetesjournals.org/content/40/1/63</u>
- ¹⁶ <u>https://swagergroup.mit.edu/quaternized-polymer%E2%80%93single-walled-carbon-nanotube-scaffolds-chemiresistive-glucose-sensor</u>
- ¹⁷ https://ubicomplab.cs.washington.edu/pdfs/biliscreen.pdf
- ¹⁸ <u>https://dspace.mit.edu/handle/1721.1/111021</u>
- ¹⁹ https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3145127/pdf/ukmss-35894.pdf
- ²⁰ https://www.technologyreview.com/s/514816/researchers-create-hate-map-of-the-us-with-twitter-data/
- ²¹ <u>https://www.intechopen.com/books/swarm intelligence focus on ant and particle swarm optimization</u>
- ²² <u>https://users.info.uvt.ro/~dzaharie/ma2016/projects/techniques/FireflyAlgorithm/Yang_nature_book_part.pdf</u>
 ²³ https://www.nklabs.com/
- ²⁴ https://venturebeat.com/2016/09/02/a-brief-history-of-googles-ill-fated-modular-smartphone-project-ara/
- ²⁵ <u>https://people.eecs.berkeley.edu/~culler/</u>
- ²⁶ https://people.csail.mit.edu/hes/
- ²⁷ <u>https://dspace.mit.edu/handle/1721.1/107988</u>
- ²⁸ <u>https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4772235/pdf/ijerph-13-00215.pdf</u>
- ²⁹ <u>http://news.mit.edu/2016/faculty-profile-rohit-karnik-0901</u>
- ³⁰ <u>http://web.mit.edu/karnik/www/home.htm</u>
- ³¹ https://en.wikipedia.org/wiki/ARPANET
- 32 https://dspace.mit.edu/handle/1721.1/108000
- ³³ https://www.energymanagertoday.com/how-ge-plans-to-make-billions-with-predix-0127475/
- ³⁴ <u>https://ca.reuters.com/article/businessNews/idCAKCN1B80CB-OCABS</u>
- ³⁵ <u>https://www.nature.com/polopoly_fs/1.12413!/menu/main/topColumns/topLeftColumn/pdf/494155a.pdf</u>
- ³⁶ http://www.cs.columbia.edu/~feiner/courses/mobwear/resources/schilit-mcsa94.pdf
- ³⁷ https://www.cisco.com/c/en/us/solutions/software-defined-networking/overview.html
- 38 http://www.rle.mit.edu/eems/wp-content/uploads/2016/11/eyeriss_jssc_2017.pdf

³⁹ <u>http://www.pnas.org/content/94/6/2150.full.pdf</u>

⁴⁰ <u>http://science.sciencemag.org/content/261/5126/1316</u>

- ⁴¹ http://www.ebiomedicine.com/article/S2352-3964(15)30064-5/pdf
- ⁴² https://www.newscientist.com/article/2123459-super-fast-parkinsons-app-will-track-symptoms-more-closely/
- ⁴³ <u>https://www.adac.de/infotestrat/tests/eco-test/default.aspx</u>
- ⁴⁴ http://qualitysafety.bmj.com/content/qhc/22/10/798.full.pdf

45 http://www.nejm.org/doi/pdf/10.1056/NEJMsa0803563

⁴⁶ <u>http://www.nature.com/nature/journal/v548/n7668/full/nature23447.html</u>

⁴⁷ http://onlinelibrary.wiley.com/doi/10.1002/anie.201705426/full

⁴⁸ <u>https://www.recode.net/2017/9/1/16236506/tech-amazon-apple-gdp-spending-productivity</u>

⁴⁹ <u>http://www.deeplearningbook.org/</u>

50 https://arxiv.org/pdf/1412.6572.pdf

⁵¹ <u>https://www.ncbi.nlm.nih.gov/books/NBK28033/</u>

⁵² <u>http://mercury2017.com/program/technical-program/4c/</u> (WO-069) PARTICIPATORY MONITORING AND COMMUNITY PROTECTION FROM MERCURY POLLUTION GENERATED BY ARTISANAL GOLD MINING IN COLOMBIA (Dr Eric Mclamore, University of Florida, Gainesville, FL)

¹ <u>http://pubs.acs.org/doi/abs/10.1021/acsinfecdis.7b00110</u>

² www.usda.gov/media/blog/2016/11/30/nanostructured-biosensors-detect-pesticide-help-preserve-environment

³ https://swagergroup.mit.edu/nanowire-chemicalbiological-sensors-status-and-roadmap-future

⁴ <u>http://www.nuwavesensors.com/products/lumair-lu45</u>

⁵ <u>https://ll.mit.edu/publications/journal/pdf/vol17_no1/17_1_3Petrovick.pdf</u>

⁶ http://pubs.rsc.org/en/content/articlelanding/2016/ew/c5ew00276a#!divAbstract

⁷ https://www.elsevier.com/books/new-pesticides-and-soil-sensors/grumezescu/978-0-12-804299-1

⁸ <u>http://www.mdpi.com/journal/sensors/special_issues/sopsawf</u>

⁹ https://swagergroup.mit.edu/janus-emulsions-detection-bacteria

¹⁰ https://swagergroup.mit.edu/wireless-oxygen-sensors-enabled-feii-polymer-wrapped-carbon-nanotubes