

EQM

Dr Shoumen Palit Austin Datta

The Elusive Quest for Monetization

Pay a finder's fee for the notional needle in the humdrum haystack

This is an idea. It is not a proposal (not yet). Suggestion combines different topics with big data science.

SUMMARY

Synthesizing analytical engines as catalysts to unleash contextual value from combinatorial data and non-obvious relationship extraction from big data and very high volume data analytics.

GOAL (not limited to but including)

Increase the precision of intelligent decisions (data driven predictive analytics eg medical diagnostics) in the context of past (lagged) events, temporally weighted (trends v spikes, time-series) and in real-time.

Integrate these concepts at the software systems level which will offer time guarantees or operate within bounded latency to be useful in control, feedback and automation related to cyberphysical systems (CPS).

APPROACH (unbaked, amorphous, distributed)

Structuring data interoperability to extract information [a] by combining data related by context [b] from highly unstructured broad data [c] and explore cryptic contextual clues in statistically uncorrelated data.

Filtration of signal from noise in hellabytes of streaming data from trillions of devices, sensors and other sources which may be connected to and communicating with diverse global platforms, such as, the IoT.

Tapestry of emerging non-traditional concepts, bio-inspired ideas from complex systems and other advanced but traditional tools (statistics, time series, fractal theories) to reconfigure old approaches (non-obvious relationship analysis) to gain information (may be actionable or possess transactional value)

ELEMENTS (dots – to be / may be – connected)

[a] Space-time-node engines

[b] Stigmergic computation

[c] Cognitive matrices

[d] Dynamic networks

[e] Semantics of time

[f] Temporally integrated software systems

[g] Artificial retina pattern recognition algorithm

[h] Neural Image Caption (NIC) Generator

POTENTIAL AND POSSIBLE OUTCOME – The cybernetic treatment of information entropy

Connect these and other dots in the context of the goals (outlined above) using the suggested approach (page 1) to converge these diverse areas to create the skeleton of an intelligent predictive tool (engine) with a modular approach (all approaches or elements may not be needed all the time for all data sets).

Create bridges to integrate this engine with legacy ideas, older notions of equation-based models (EBM, PDE) and agent-based models (ABM) as well as semantic ontologies, econometrics, operations research (dynamic optimization, Markov chains) and control theory (feedback loops, recursive operations).

Convert this conceptual flow into programmable or intuitive and intelligent software tool which can be used as an “engine” and integrated with (think embedded multicore platforms) or distributed to edge devices to push the analytical operation to the point of origin of the data (operating in real-time and run-time modes as well feeding back the "lesson" to the machine learning core). Explore programming temporally integrated distributed embedded systems with this “engine” in the context of time criticality.

Make these “engines” relevant to data scientists as well as politicians by creating “app” like connections to rapidly define the value of the data, information and knowledge which may be relevant to operations (retail moment of truth, run-time intruder detection, just in time delivery) as well as policy development based on data (rather than ephemeral political whim or desire). For example, the uncharted path for rules, regulations and policies to guide autonomous vehicle activity on public roads or healthcare privacy boundaries that will test HIPAA when health data monitoring catalyzes n-directional flow of information.

AN UNCERTAIN PATH TO NON-OBVIOUS ANALYSIS – *where inventions or insights may be hiding*

Convergence of dissimilar tools and unconventional approaches to data analysis (sterilizing, stitching, synthesizing) may unleash otherwise cryptic patterns or non-obvious relationships as a by-product of conventional search for established parameters or KPIs. The vast flow of data requires improved pattern recognition algorithms ([artificial retina](#)) where speed (1 tera bit per second or higher) and diversity of data demands new approaches to uncover new information (eg elimination of congestive heart failure).

It is no longer sufficient to practice "aim, shoot" but the counter-intuitive modus operandi "shoot, aim" may be increasingly a sign of the times. We may use a dynamic or matrix approach (Alice Guionnet, MIT) where the chosen tools are from a repertoire of analytical engines each driven by a specific domain. The application / user or the use case chooses which engines to select and combines them for a particular type of analysis. This is the "obvious" approach and may be analogous to an automobile repair shop where large red cases of drawers store multiple little widgets. Depending on the repair job the mechanic selects one or more widgets from those tiny boxes. For traditional data another analogy may be the pendulum which is an IFTTT approach. When the amplitude of the swing is small, the period of vibration of a pendulum is directly proportional to the square root of its length. If a pendulum 1 meter long has a period of 1 second, a pendulum 4 meters long will have a period of 2 seconds.

What if you carry out the "obvious" approach for a fee but then re-visit the raw data and subject it to a matrix of analytical engines in various combinations. Some results may be questionable (customer buys 0.13 eggs every 13 years). But, computation of data through a matrix of engines may reveal what we

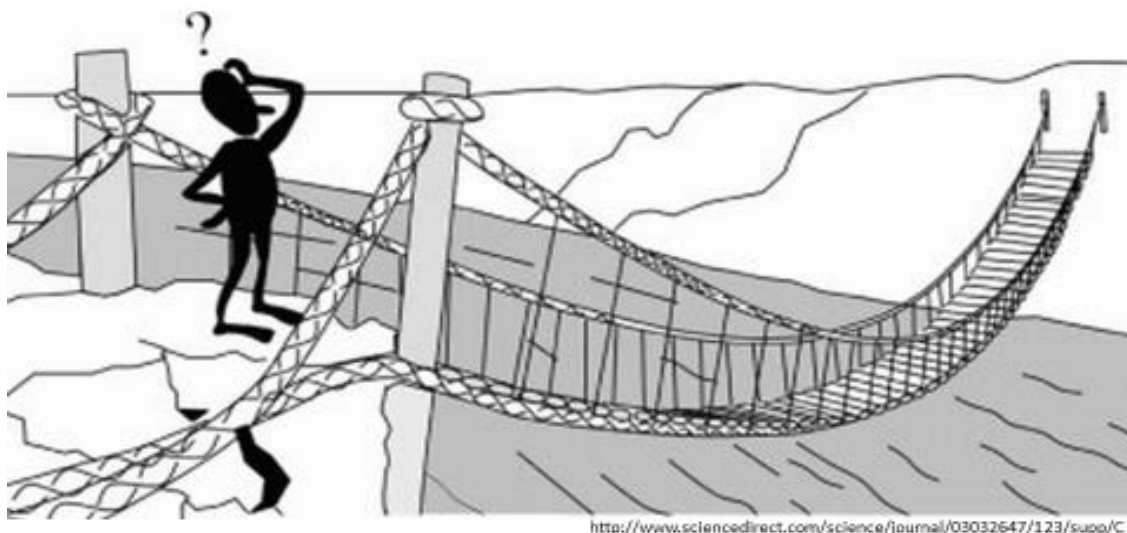
didn't know or could not have known. This may be similar to a RDBMS/SQL type table where the questions are in the boxes and data (eg high volume sensor data) is flowing through each box. If the question finds an answer in the data stream, it pings the system. It is also analogous to molecular screening for active site binding by synthetic organic molecules.

But the fact remains, if we do not know what we are looking for or looking at then how would we recognize the "positive" outcome when we see the results? Wouldn't we discard it because it didn't fit our sense of known knowledge? How do we know that we have found a non-obvious relationship?

This is where artificial intelligence may meet with natural stupidity, ie, humans. This is why breakthroughs may be catalyzed by the serendipitous prepared mind and ultimately human-aided systems may be required to hand-pick that needle from the haystack. The AI avatar may have narrowed that zone to a very tiny bounded area in order to accelerate the process. In other words, we deploy decision support system but not a complete execution of the decision, depending on exception management or the nature of the task. The complete execution of a decision system complete with controls and feedback loops may be equivalent to automation.

One problem in automation is the current inability of an automated object (eg vehicles) to differentiate between the threat potential of a pedestrian crossing the road, a dog, a piece of shredded rubber tyre flying in front of the vehicle or a plastic bag floating past the windshield. The ability of neuron image caption generator to [translate images into natural language descriptions](#) may refine collision avoidance. Please review *Future IoT* which may be downloaded from <http://dspace.mit.edu/handle/1721.1/86935>

Handrails – Constraints or Opportunities to Cross the River



<http://www.sciencedirect.com/science/journal/03032647/123/supp/C>

Questions and models:

1. How do we convert industrial data streams to real revenue streams?
2. Which applications in IoT and IIoT offers practical value which resonates with industrial customers?

a) Remote Monitoring

Connected buildings initiative - 11,000 government buildings to be risk optimized. Process billions of events per day from 96,000+ crucial industrial assets like boilers, chillers, pumps in real time with an intra-event interval of 3 seconds to the central command center. Pricing model - price per asset per month for business models involving remote asset maintenance. (www.flutura.com)

b) Predictive Action

Proactively recommend which electrical assets needs recalibration. Assets are generally installed with default settings and field conditions may necessitate recalibration.

c) Personalization

Energy audits for consumption of gas, electricity and water (irrigation) as a peer to peer comparison in municipalities to optimize public goods. Insurance related data (mileage, braking, turns, acceleration) to obtain discounts based on driving habits related to individual risk profile linked to dynamic pricing model (unique to the individual as opposed to being a part of a generic age profile). Retail applications of price comparison and dynamic pricing are abundant in operations research and supply chain management.

d) Machine Data

To improve machine performance and preventive maintenance by sampling data from physical systems which are yet to be connected to deliver data. For example, deep machine sensors in turbines, jet engines and other equipment (the emergence of true "industrial" internet of things).

Note

The process of converting these suggestions into a proposal is work in progress in collaboration with James Lamb. It is not our intent to pursue "money" models exclusively rather advance the theme of making "sense of data" and indicate, where appropriate, how the data or information may reduce transaction cost or add value for business or enhance public goods. I have collected some of the references (papers) that may substantiate these suggestions about big and small data science. To review the collection of papers, please email (shoumen@mit.edu). It is a folder >250MB (download from Google Drive). You are invited to comment, contribute or co-author this flight of ideas. External views of monetization are peppered all over industrial IoT forums. One example is copied above and it exemplifies the common adage "*the price of light is less than the cost of darkness*" (A C Nielsen).