Cutter IT Journal

The Journal of Information Technology Management

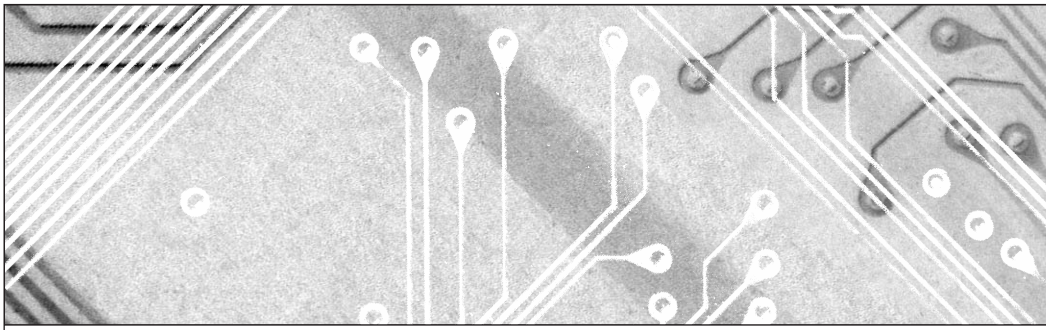
Vol. 29, No. 4
April 2016

"The sheer variety, volume, and velocity of data the IoT generates create a major obstacle to realizing its full potential. This calls for a strategic approach to data management."

— Bhuvan Unhelkar and San Murugesan, Guest Editors

IoT Data Management and Analytics:

Realizing Value from Connected Devices



Cutter IT Journal

About Cutter IT Journal

Part of Cutter Consortium's mission is to foster debate and dialogue on the business technology issues challenging enterprises today, helping organizations leverage IT for competitive advantage and business success. Cutter's philosophy is that most of the issues that managers face are complex enough to merit examination that goes beyond simple pronouncements. Founded in 1987 as *American Programmer* by Ed Yourdon, *Cutter IT Journal* is one of Cutter's key venues for debate.

The monthly *Cutter IT Journal* and its companion *Cutter IT Advisor* offer a variety of perspectives on the issues you're dealing with today. Armed with opinion, data, and advice, you'll be able to make the best decisions, employ the best practices, and choose the right strategies for your organization.

Unlike academic journals, *Cutter IT Journal* doesn't water down or delay its coverage of timely issues with lengthy peer reviews. Each month, our expert Guest Editor delivers articles by internationally known IT practitioners that include case studies, research findings, and experience-based opinion on the IT topics enterprises face today — not issues you were dealing with six months ago, or those that are so esoteric you might not ever need to learn from others' experiences. No other journal brings together so many cutting-edge thinkers or lets them speak so bluntly.

Cutter IT Journal subscribers consider the *Journal* a "consultancy in print" and liken each month's issue to the impassioned debates they participate in at the end of a day at a conference.

Every facet of IT — application integration, security, portfolio management, and testing, to name a few — plays a role in the success or failure of your organization's IT efforts. Only *Cutter IT Journal* and *Cutter IT Advisor* deliver a comprehensive treatment of these critical issues and help you make informed decisions about the strategies that can improve IT's performance.

Cutter IT Journal is unique in that it is written by IT professionals — people like you who face the same challenges and are under the same pressures to get the job done. *Cutter IT Journal* brings you frank, honest accounts of what works, what doesn't, and why.

Put your IT concerns in a business context. Discover the best ways to pitch new ideas to executive management. Ensure the success of your IT organization in an economy that encourages outsourcing and intense international competition. Avoid the common pitfalls and work smarter while under tighter constraints. You'll learn how to do all this and more when you subscribe to *Cutter IT Journal*.

Cutter IT Journal®

Cutter Business Technology Council: Rob Austin, Ron Blitstein, Tom DeMarco, Lynne Ellyn, Vince Kellen, Tim Lister, Lou Mazzucchelli, Ken Orr, and Robert D. Scott

Editor Emeritus: Ed Yourdon
Publisher: Karen Fine Coburn
Group Publisher: Chris Generali
Managing Editor: Karen Pasley
Production Editor: Linda M. Dias
Client Services: service@cutter.com

Cutter IT Journal® is published 12 times a year by Cutter Information LLC, 37 Broadway, Suite 1, Arlington, MA 02474-5552, USA (Tel: +1 781 648 8700; Fax: +1 781 648 8707; Email: citjeditorial@cutter.com; Website: www.cutter.com; Twitter: @cuttertweets; Facebook: Cutter Consortium). Print ISSN: 1522-7383; online/electronic ISSN: 1554-5946.

Opening Statement

by Bhuvan Unhelkar and San Murugesan, Guest Editors

The world of plentiful miniaturized computers in phones and clothes creates a human-technology mosaic that we find difficult to define, and so we call it *things*.

The Internet of Things (IoT) effectively comprises two keywords or concepts:

1. **Internet** — signifying connectivity and communication
2. **Things** — representing a variety of objects

The explosion in networking coverage and speed is felt in short-range and wide-area mobile connectivity, as well as wireless communications ranging from near-field communications (NFC) and Wi-Fi to satellite-based signaling, to name but a few. This has enabled hitherto unimaginable realities of road navigation, health monitoring, contactless purchasing, crime monitoring, and even groceries monitoring in the home fridge, as well as smart homes, smart cities, autonomous vehicles, and drones. Advances in wireless communications and networking complemented by the shrinking size and cost of microcomputers, sensors, and actuators are driving several IoT applications. Affordable miniaturized wearable devices (pens, watches, buttons, and even underclothes) are further driving the popularity and permeation of IoT in our daily lives.

However, the sheer variety, volume, and velocity of data the IoT generates create a major obstacle to realizing its full potential. This calls for a strategic approach to data management and usage that addresses business strategies, business processes, enterprise architecture, systems and applications, and security and privacy considerations. To glean maximum value from the data, we must decide how to capture, store, extract, and use it in ways that can result in new or improved products and services, better decision making, and enhanced policy development.

So, in this edition of *Cutter IT Journal*, we focus your attention on an important issue facing us in leveraging the potential of the IoT: data management and analytics. To do so, we bring to you five excellent discussions around the IoT. We plucked these articles from an overwhelming response to our call for papers, identifying those with a fine balance of the rigors of theory and research together with examples and case studies that discuss direct practical applications of IoT.

Advances in wireless communications and networking complemented by the shrinking size and cost of microcomputers, sensors, and actuators are driving several IoT applications.

IN THIS ISSUE

In the first article in this issue, "Solving the Jigsaw Puzzle: An Analytics Framework for Context Awareness in the Internet of Things," authors Anurag Agarwal, Ramakrishna Govindu, Sunita Lodwig, and Fawn Ngo highlight the value of context in IoT and propose an analytics framework powered by a *context awareness engine*. Their framework outlines an approach to handling multiple and heterogeneous sources of data to develop accurate models that provide 360-degree context awareness for IoT-based applications. This framework can help create a much richer suite of data points for IoT applications, thereby offering the potential for improved analytics and decision making. To further illustrate the potential for improving the operational effectiveness of an IoT-based application through context awareness, the authors offer a real-world example of a breath alcohol analyzer used to keep impaired drivers off the road. You will find some very practical advice on how to detail a context (through the four W's — Where, When, Who and What) and how this context can be applied in practice.

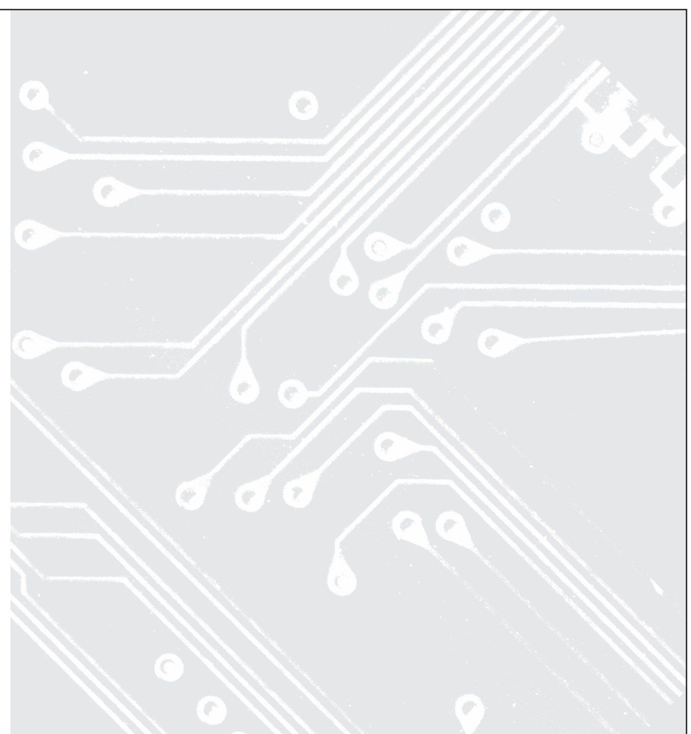
Next, Pranav Shah, Suman Datta, Rekha Vaidyanathan, Sudhakara Poojary, and Vidyut Navelkar dive deep to provide us with "IoT Data Management Challenges: A View from the Trenches" — which is indeed an excellent view from the trenches. Derived from their experience of building and deploying an enterprise-wide Remote Energy Management Solution, they share their insights on issues dealing with a diversity of data sources as well as quality aspects such as missing and noisy data. In telling us how they strategized to manage the high-velocity data, handled its storage, and ran analytics on the data in near real time, the authors provide invaluable practical lessons in developing IoT-based applications.

Needless to say, the data generated through IoT-based devices is indeed vast, unstructured, and changing rapidly — all the characteristics of big data.

In our third article, Donald Wynn and Renée M.E. Pratt answer Agarwal and his coauthors' "four W's" with "Exploratory, Embedded and Enabled: The Three E's of IoT Data Management Maturity." The exploratory, embedded, and enabled levels of IoT maturity they describe provide a helpful basis for IoT adoption in real-life organizations. Because the IoT devices at the exploratory level are standalone, their impact on financial and operational efficiency is not a major concern. The main value derived here is knowledge and experience. The embedded maturity level focuses on integration of devices and networks to result in mature workflows. At the enabled maturity level, organizations are able to establish guidelines, standards, and best practices for incorporating the technology into their strategic and operational plans. Wynn and Pratt further examine the three crucial phases of data management (data collection and storage, data integration, and data analysis) at each of the maturity levels.

Like Shah et al., Tushar Hazra offers us another real-world case study. In "Leveraging EA and IoT Synergy for Digital Transformation," Hazra details how a multi-million-dollar US healthcare IT service provider with more than 60,000 public sector employees used EA practices in its IoT implementation to effect digital transformation across the enterprise. He writes, "While EA and IoT are not similar disciplines, I believe there is an inherent synergy between them, stemming from the fact that EA and IoT each support digital strategies for disruptive business operations. They deliver a converged enterprise that is collaborative, connected, and often socially infused." Hazra's account will be of immense interest to readers interested in the crossover of EA and IoT.

In our final article, you will find the fascinating concept of "hex elementization" developed and applied to IoT. Mapping it to the "nature of things," authors Girish Nair and Yi-Chen Lan create a six-part (hex) structure of elements that can be used to define a data point. Their framework helps in breaking down a source of data — be it structured data, unstructured data, sensor data, or what have you — into such hex elements and then integrating them. They show how these elements



can be assembled to produce a rich context for a suite of IoT-based applications and processes. This framework further aims to update the elements with "enough attributes to help them automatically seek and match themselves with information seen in other unrelated pieces of hex elements." Such matching has the potential to create dynamically changing patterns and mosaics for future IoT applications.

We hope you will find these five articles on IoT interesting, challenging, and practical. We also hope they will encourage you to experiment with the ideas they present in your own approach to integrating IoT within your business systems and processes. Needless to say, the data generated through IoT-based devices is indeed vast, unstructured, and changing rapidly — all the characteristics of big data. We see this issue as a major contribution to the literature on IoT, with overlaps in identifying the context, incorporating IoT in applications, and merging it with EA. As always, we welcome your comments on these articles.

Bhuvan Unhelkar (BE, MDBA, MSc, PhD; FACS) is a Senior Consultant with Cutter Consortium's Business Technology & Digital Transformation Strategies practice. He has decades of strategic, as well as hands-on, professional experience in the information and communication technologies (ICT) industry. As an Associate Professor, he currently leads the IT department at the University of South Florida Sarasota-Manatee. As a Founder of MethodScience.com, he has demonstrated consulting and training expertise in big data (strategies), business analysis (use cases, BPMN), software engineering (object modeling, Agile processes, and quality), collaborative Web services, green IT (environment), and mobile business. His domain experience includes banking, financial services, insurance, government, and telecommunications. Dr. Unhelkar earned his doctorate in the area of object orientation from the University of Technology, Sydney, in 1997. Since then, he has authored or edited 17 books in the areas of collaborative business, globalization, mobile business, software quality, business analysis/processes, the UML, and green ICT and has extensively presented and published papers and case studies. Dr. Unhelkar is a sought-after orator, a Fellow of the Australian Computer Society, Life Member of the Computer Society of India, Rotarian (Paul Harris Fellow), Discovery Volunteer at New South Wales Parks and Wildlife, and a previous TiE Mentor. He can be reached at bunhelkar@cutter.com.

San Murugesan (BE [Hons], MTech, PhD; FACS) is a Senior Consultant with Cutter Consortium's Business Technology & Digital Transformation Strategies practice, Director of BRITE Professional Services, and an Adjunct Professor in the School of Computing and Mathematics at the Western Sydney University, Australia. He is Editor-in-Chief of the IEEE's IT Professional *and Editor of IEEE's* Computer. *Dr. Murugesan has considerable experience in both industry and academia, and his expertise and interests include the Internet of Everything (IoE), cloud computing, green computing, healthcare IT, advances in the Web, and IT in emerging markets. He offers training programs on IoE, cloud computing, and green IT.*

Dr. Murugesan is coeditor of the book Harnessing Green IT: Principles and Practices *and the upcoming* Encyclopedia of Cloud Computing. *He edits the "Cloud Cover" column in* Computer. *Dr. Murugesan is Standing Chair of the new COMPSAC Symposium on IT in Practice (ITiP). He is a fellow of the Australian Computer Society and served as a distinguished visitor of the IEEE Computer Society. Dr. Murugesan held various senior positions at Southern Cross University, Australia; Western Sydney University; and the Indian Space Research Organization, Bangalore, India. He also served as Senior Research Fellow of the US National Research Council at the NASA Ames Research Center. He can be reached at smurugesan@cutter.com; website: http://tinyurl.com/sanbio.*

REMEMBERING ED YOURDON

by Bhuvan Unhelkar

In the early 1980s, as a young COBOL programmer with Telco (Tata Engineering — nothing to do with the telecom industry), I discovered *Structured Design: Fundamentals of a Discipline of Computer Program and Systems Design* (Yourdon Press, 1979). Little did I know that almost three decades later, the lead author of that book — and many others — would pen a foreword to my own book *Mobile Enterprise Transitions* (CRC Press, 2009). I distinctly remember my dance of joy when Ed agreed to write the foreword for me. In introducing my book, he mentioned his own classics — *Rise and Resurrection of the American Programmer* and *Death March*. He also explained the key differentiators between successful and unsuccessful software development and deployment projects. Ed was a big asset to the software community internationally, including and especially in India and Australia, where he is quite revered. I know that up there, somewhere in the heavens, Ed Yourdon is espousing the cause of modeling and processes to improve the quality of whatever it is that gets done there. I deeply mourn his loss. Rest in peace and thank you, Ed.

Solving the Jigsaw Puzzle: An Analytics Framework for Context Awareness in the Internet of Things

by Anurag Agarwal, Ramakrishna Govindu, Sunita Lodwig, and Fawn T. Ngo

The Internet of Things (IoT) is a paradigm wherein ubiquitous, context-aware devices equipped to identify, sense, and process data communicate over the Internet to accomplish some intended objective(s). The Internet of Things Global Standards Initiative (IoT-GSI) defines IoT as:

> A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.[1]

Several technologies that have evolved independently over the last two decades (e.g., sensor networks, RFIDs, microchips, intelligent agents, the Internet, mobile computing) are now converging to enable the IoT paradigm. For example, it used to be the case that devices equipped with RFID chips could only be employed locally due to their limited near-field communication abilities. With advances in communication and Internet technologies, the same devices have now expanded their geographic reach globally. The development and adoption of these technologies have been so rapid that the number of active IoT devices themselves are estimated to reach 50 billion by 2020, up from an estimated 25 billion in 2015.[2] According to one estimate,[3] 40% of all data generated by 2020 will come from interconnected sensors and devices. Thanks to IoT and other related technologies, massive quantities of both structured and unstructured data are being generated on a continuous basis at a phenomenal rate, leading to a big data revolution, which in turn is providing opportunities for big data analytics.

The increasing business and social impact of the IoT paradigm is prompting researchers and practitioners to bring together and integrate more and more "things," resulting in a vast network of autonomous, self-organizing, and intelligent devices.[4] Consequently, the IoT holds the promise of creating a global network, supporting ubiquitous computing and context awareness among devices.[5] Ubiquitous computing and context awareness have become critical requirements of ambient intelligence, one of the key promises of the IoT.[6] The IoT helps embed technologies into everyday products/devices, such as audio/video receivers, wristwatches, smoke detectors, and home appliances, which not only enables them to communicate online, but also to receive and process data and information from other devices in a dynamic fashion, in real time. Thus, the real revolution of IoT goes beyond embedding a sensor and sending signals over the Internet to developing a 360-degree context awareness by analyzing data from multiple sensors or sources using complex advanced algorithms, in real time, for improved decision making.

For example, RFID technologies previously enabled organizations to track the location of products through the supply chain at various destination points, such as the warehouse or the retail outlet. With IoT, though, products can now be tracked while in motion and in real time, resulting in dynamic tracking and improved inventory management. As another example, with traditional RFID technologies, a machine stationed at a location could be monitored through sensors and a potential maintenance issue communicated to manufacturers to preempt failures (e.g., elevator manufacturers getting notification alerts on impending failures so that preventive maintenance can be scheduled, thus improving customer safety). With the evolution to IoT, this advanced monitoring feature can now be extended to machines that are mobile, such as engines in trucks, ships, and planes. Such intelligent behavior requires IoT devices to be aware of their context or surroundings.

CONTEXT AWARENESS

The IoT literature provides many definitions of context awareness.[7] For the purpose of this article, we define context awareness as the information necessary and sufficient to perform the intended function of the device effectively and efficiently. Typically, but not always, the context can be ascertained comprehensively by answers to some or all of what we like to call "the four W's": Where, When, Who, and What. A simple IoT device with limited functionality may only need to answer

Parameter	Description	Examples of Data
Where	Location of events	The latitude, longitude, and altitude of the IoT device
When	Time of events	The timestamp of all the signals/data received by the IoT device
Who	Person(s) or object(s) of interest associated with the events	Identification data for the object or person (e.g., biometric readings for a person or the IP address of another IoT device in network)
What	Various measures of interest besides location, time, and objects	Temperature, pressure, speed, level, weight, duration, sentiment, demographics, distance

Table 1 — Parameters defining context.

one or two W's, while a more complex IoT device may need answers to all four, and perhaps even to additional questions such as How, Why, Which, How Much, and so on. See Table 1 for some examples of contextual data needed for the parameters defining the context awareness of an IoT device.

AN ANALYTICS FRAMEWORK FOR CONTEXT AWARENESS

Figure 1 illustrates our proposed analytics framework for a typical IoT device. As shown in Figure 1a, within one's environment of interest, there is typically a network of IoT and non-IoT devices interacting with each other over the Internet to accomplish certain objective(s) for an entity (person or organization). Each IoT device has three main components: a set of sensors, a context awareness engine, and a solution engine (see Figure 1b).

Through the set of sensors, the necessary data is collected. There are typically many different types of sensors that generate data for a specific application, ranging from a simple thermometer for measuring the room temperature to a radio telescope for sensing radio waves from faraway galaxies. Sensors can be designed to capture data for different environmental characteristics, such as latitude and longitude, time, temperature and other weather-related characteristics, the presence of objects, motion detection, speed, and so on. These sensors provide data to answer questions for the four W's (see Table 1). The context awareness engine then analyzes the data to model the context. Taking into account the objective of the application, the solution engine uses the context information and determines the best possible solution, which could be some action. The action might simply be a notification to an appropriate person or device or it could be a corrective action that alters the state of some object of interest in the environment.

The context awareness engine of the IoT device has two main components — the data representation engine and the context modeling and analytics engine (see Figure 1c). The data representation engine acquires the sensor data from various heterogeneous sources and represents and stores that data in the appropriate formats. The context analytics and modeling engine combines data in heterogeneous formats and applies suitable algorithms to model the context. The output of the context awareness engine is the context, which is used by the solution engine. The context is also stored within the context awareness engine using the data representation engine, to enable dynamic update of the context and further refinement of context if needed.

Modeling a 360-degree view of the context is like solving a jigsaw puzzle, where each puzzle piece comes from a different source in the form of data.

As shown in Figure 1c, there is bidirectional communication between the data representation engine and the context analytics and modeling engine, because the way the data is represented influences the modeling algorithms and vice versa. Depending on the nature of the input data, it can be stored in a variety of different formats within the data representation engine, such as the standard two-dimensional relational table, as a class of objects with properties and methods, or as one of the NoSQL database formats (key-value pairs, column family database, graph database, etc.). For instance, for high-velocity, single-column data, the key-value pair format would be the most appropriate. For low-volume, multi-attribute, structured data with no built-in methods, the tabular format would be the most appropriate.

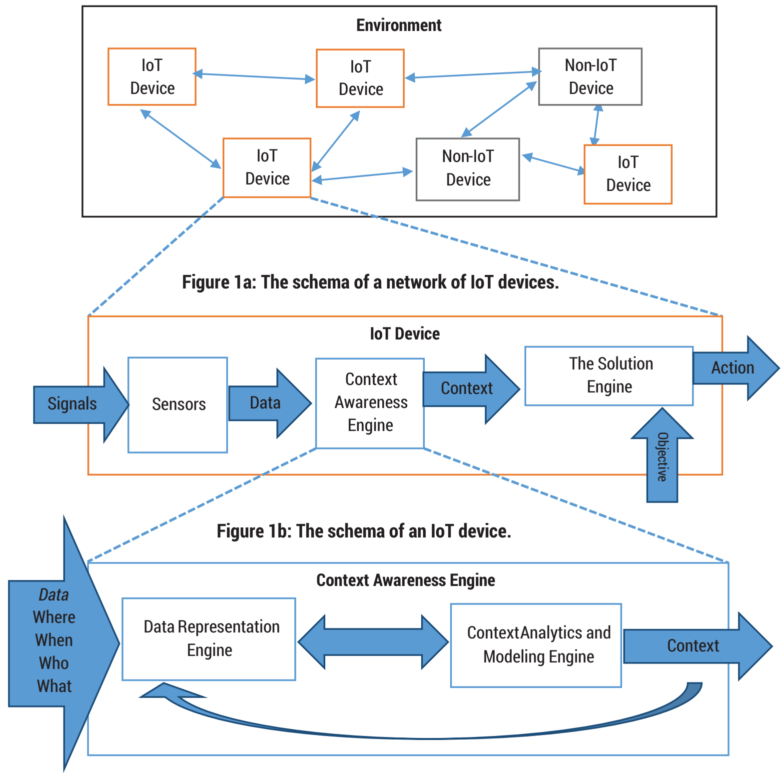


Figure 1a: The schema of a network of IoT devices.

Figure 1b: The schema of an IoT device.

Figure 1c: The schema of the context awareness engine.

Figure 1 — An analytics framework for context awareness of an IoT device.

For low-volume, multi-column, structured data with built-in methods, the object-oriented format would be the most appropriate.

Modeling a 360-degree view of the context is like solving a jigsaw puzzle, where each puzzle piece comes from a different source in the form of data. The completed puzzle may look completely different from any of the individual pieces. This puzzle solving happens within the context analytics and modeling engine by using the data collected to answer the four W's (in Table 1). Depending on the nature of the data and the application, it may employ a variety of analytics techniques, ranging from simple arithmetic operations to complex pattern recognition (facial, voice, image, or text). Other techniques might include data mining methods such as classification, clustering, association rule mining, and so forth. This context modeling is usually performed in real time, since the actions taken by an IoT device are typically time-sensitive.

To illustrate our proposed analytics framework for context awareness, let's look at a practical example of an IoT system that helps law enforcement agents keep the roads free of drunk drivers.

KEEPING IMPAIRED DRIVERS OFF THE ROAD: A PRACTICAL EXAMPLE

A breath-alcohol analyzer (BAA) is an IoT device used to monitor the breath-alcohol level of DUI offenders during their probationary period. DUI offenders are required to install a BAA device in their vehicle and provide their breath sample periodically while driving. The BAA analyzes the breath sample for its alcohol level, and if the level exceeds the legal limit, the device alerts the appropriate law enforcement agency. However, this approach is not foolproof. Drivers may potentially cheat the system by having a sober person provide the breath sample on their behalf, or they might adopt measures to manipulate their breath sample using some suppressants. Furthermore, DUI offenders may be risky drivers even while within the legal alcohol level.

To better assess the riskiness of an offender's driving, a more comprehensive context awareness is needed, one that builds from multiple sources of data. For example, in addition to breath-alcohol level, the smoothness of the drive itself can be assessed using appropriate sensors, such as a telematics device and the vehicle's curb camera. Thus, the functionality of the BAA as an IoT device can be enhanced by enabling it to gather data from multiple sources, thus developing a more comprehensive context by applying analytics on these multiple sources of data. Figure 2 shows the operational schema of an environment in which the BAA interacts with other IoT and non-IoT devices.

The BAA installed in the DUI offender's vehicle receives input from the driver (biometric readings, breath sample, voice sample), a telematics device (speed, timestamp), a curb camera (distance from curb, timestamp), satellite (longitude and latitude), and the law enforcement agency (voice sample and other biometric patterns of the offender). In addition to these external sources, the BAA itself may contain some internal sensors that provide data for the context awareness engine. Table 2 shows the details of the data collected for the BAA to answer the four W's.

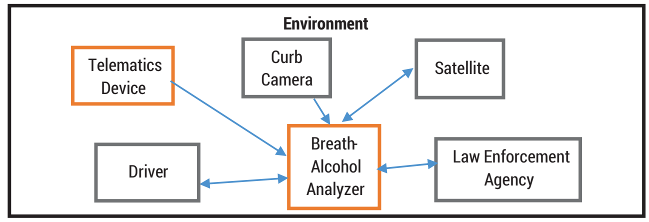


Figure 2 — The schema of the breath analyzer IoT device and its environment.

Data Representation for the BAA

Since the data sources for the BAA are quite heterogeneous, the appropriate data representation formats also differ, depending on the data characteristics. In our example, the telematics data and the curb camera data are considered high-velocity, since we are collecting the data every second. It is best to represent this data as key-value pairs. For the telematics device data, the key will be the timestamp, and the value will be the vehicle speed. For the curb camera data, the key will again be the timestamp, and the value will be distance from the curb. The offender data from the law enforcement agency is very low-volume, multi-attribute with no in-built methods, and therefore a tabular data format is the most appropriate for this data. The data for each trip and for each breath analysis is also low-volume, multi-attribute, structured data with in-built methods, and thus the most appropriate format would be an object-oriented class.

Figure 3 shows the data representation for each type of data discussed above. For the class diagram, we identify two classes: Trip and BreathTest. The Trip class captures data for each trip, such as the date, time, driver's biometric reading, and origin and destination coordinates. Its methods confirm the identity of the offender. The BreathTest class captures the results of the breath test multiple times for each trip. Its methods are designed to confirm the validity of tests and whether there was a violation. The data from the telematics device and the curb camera is shown as key-value pairs. This data will be used by the context modeler to determine if the offender is driving in a jerky manner or weaving within his or her lane.

Parameter	Description	Example of Data	Source
Where	Location	The longitude and latitude of the vehicle	Satellite
When	Time	The timestamp of the breath analysis Timestamp for speed Timestamp for curb distance Time of request for breath Time of breath input	BAA Telematics device Curb camera BAA BAA
Who	Person(s) or object(s) of interest	Offender's identity (through some biometric sensing)	Driver
What	Measures of interest	Alcohol level in the breath Speed of vehicle Curb distance Biometric data for the offender	BAA Telematics device Curb camera Law enforcement agency

Table 2 — Parameters defining context for a breath analyzer.

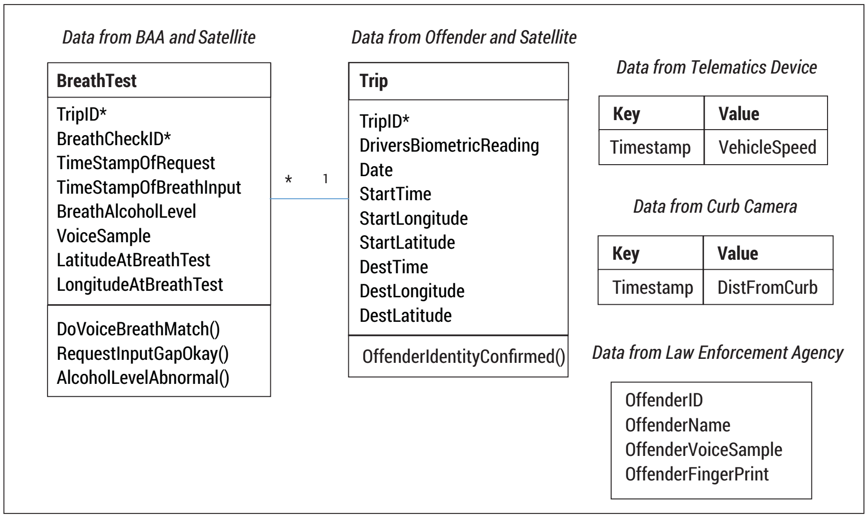


Figure 3 — Data representation for BAA data.

Context Analytics and Modeling for the BAA

The telematics data on vehicle speed, which is captured every second, is used to evaluate whether the drive is smooth or swervy. An algorithm calculates the acceleration and deceleration every second and determines if the variations are above normal in level and frequency. If they are, then the BAA device becomes aware that the drive is abnormal. Similarly, another algorithm calculates differences in the distance from the curb every second. If these differences are beyond a certain threshold and in both directions, and sustained over a period time, then the BAA device becomes aware that the offender is swerving.

Multidisciplinary skills will be needed in order to develop and deploy context-aware engines for IoT-driven applications.

When the BAA becomes context aware of an abnormal drive, it may send a notification to the law enforcement agency and a new request for a breath test may be triggered to the driver. The driver must provide a breath sample within a predetermined duration of time; if the gap between the request and the receipt of the breath sample exceeds that duration, the RequestInputGapOkay() method (see Figure 3) sends an alert to law enforcement. If the breath-alcohol level is above a pre-set threshold, the AlcoholLevelAbnormal() method sends another alert. While providing the breath sample, the offender is also required to make a sound. The sound characteristics are then compared with the offender's voice sample that the law enforcement agency has on file. If the voice does not match, the method DoVoiceBreathMatch() sends an alert to the agency. This check is used to ensure that the offender is not using another person's breath to circumvent the test. If he or she is, the law enforcement agency will receive this notification and charge the offender with the violation. If the voice matches, then no alert is sent.

POTENTIAL CHALLENGES

There are a few challenges associated with our framework. As our example shows, the need for a comprehensive context requires heterogeneous data collection and storage. Further, the analytics is being performed in real time on disparate data representations. Multidisciplinary skills will be needed in order to develop and deploy context-aware engines for IoT-driven applications.

CONCLUSION

The analytics framework we have proposed in this article is designed to handle multiple and heterogeneous sources of data to develop accurate models for 360-degree context awareness for IoT-based applications. By capturing and analyzing the details of the context, the IoT device enhances context awareness, which in turn leads to improvements in the operational effectiveness of an IoT-based application. For instance, the BAA example illustrates how law enforcement agencies could benefit as enhanced context awareness provides more accurate and comprehensive monitoring of DUI offenders on probation. This, in turn, reduces the time, cost, and effort spent to achieve the objective of keeping unsafe drivers off the road. We believe application of the proposed framework will result in a much richer suite of data points for many IoT devices, thereby providing more value to businesses and other organizations in their decision making.

ENDNOTES

[1]"Overview of the Internet of Things." ITU-T (www.itu.int/ITU-T/recommendations/rec.aspx?rec=y.2060).

[2]Evans, Dave. "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything." Cisco Internet Business Solutions Group (IBSG), April 2011 (www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf).

[3]Taslim, Allison. "12 Facts You Need to Know About the Internet of Things." Servicemax (blog), 17 July 2015 (http://blog.servicemax.com/12-facts-you-need-to-know-about-the-internet-of-things).

[4]Whitmore, Andrew, Anurag Agarwal, and Li Da Xu. "The Internet of Things — A Survey of Topics and Trends." *Information Systems Frontiers*, Vol. 17, No. 2, April 2015 (http://link.springer.com/article/10.1007%2Fs10796-014-9489-2).

[5]Hong, Jong-Yi, Eui-Ho Suh, and Sung-Jin Kim. "Context-Aware Systems: A Literature Review and Classification." *Expert Systems with Applications*, Vol. 36, No. 4, May 2009 (www.sciencedirect.com/science/article/pii/S0957417408007574).

[6]Bellavista, Paolo, et al. "A Survey of Context Data Distribution for Mobile Ubiquitous Systems." *ACM Computing Surveys*, Vol. 44, No. 4, August 2012 (http://dl.acm.org/citation.cfm?id=2333119).

[7]Perera, Charith, et al. "Context Aware Computing for The Internet of Things: A Survey." *IEEE Communications Surveys & Tutorials*, Vol. 16, No. 1, 2014 (http://arxiv.org/pdf/1305.0982).

Anurag Agarwal is a faculty member in the College of Business at the University of South Florida Sarasota-Manatee. He has a PhD in information systems from The Ohio State University and an MBA from the University of Wisconsin. His teaching interests lie broadly in the field of analytics, data mining, and information systems. Dr. Agarwal has taught courses on big data analytics, business analytics, business statistics, operations management, information systems in organizations, systems analysis and design, database systems, programming languages, and decision support systems. His primary research interests are in predictive and prescriptive analytics. He has published in journals such as INFORMS Journal on Computing, Naval Research Logistics, European Journal of Operational Research, Omega: The International Journal of Management Science, Computers and Operations Research, Annals of Operations Research, Information Systems Horizons, *and* Information Technology and Management. *Dr. Agarwal serves on the editorial board of* Information Technology and Management *and* Enterprise Information Systems. *He can be reached at agarwala@sar.usf.edu.*

Ramakrishna Govindu is a faculty member in the College of Business at the University of South Florida Sarasota-Manatee. He has a PhD in industrial engineering from Wayne State University and an MTech (Honors) in quality, reliability, and operations research from the Indian Statistical Institute, Calcutta, India. Dr. Govindu has taught courses in applied statistics, operations and supply chain management, Lean Six Sigma, and operations research and decision sciences. His primary research areas include predictive and prescriptive analytics, supply chain modeling and management, and multi-agent systems. Dr. Govindu has published in Computers & Industrial Engineering, International Journal of Modelling and Simulation, Wiley's Handbook on Technology Management, *and* American Journal of Criminal Justice. *He has about two decades of industry experience in the US and India that includes more than a decade of consulting work. Dr. Govindru has executed, led, and managed projects for organizations such as Caterpillar, Ford, Visteon, GE, and others. He has implemented solutions involving supply chain analytics, predictive and prescriptive analytics solutions for decision support, business reengineering, and Lean Six Sigma that saved millions of dollars for clients in the manufacturing, service, and healthcare sectors. Dr. Govindu is a member of several professional associations. He can be reached at rgovindu@sar.usf.edu.*

Sunita Lodwig is a faculty member in the College of Business at the University of South Florida Sarasota-Manatee. She has a PhD in theoretical plasma physics from the Indian Institute of Technology, Delhi, India. Dr. Lodwig has served as a faculty member at Northern Illinois University and as a Research Associate at Washington State University. She currently teaches a range of courses in information technology, such as UML, program design, programming languages, security, data structures, and the capstone course. Dr. Lodwig also served in management positions at AT&T Bell Labs and held various roles in Motorola's Global Software Group, where she was involved in technical marketing, project/product management, and globalization issues. She can be reached at slodwig2@sar.usf.edu.

Fawn T. Ngo is a faculty member in the College of Arts and Sciences at the University of South Florida Sarasota-Manatee. She has a PhD in criminology and criminal justice from the University of Maryland and an MS in criminal justice from the California State University, Long Beach. Dr. Ngo's teaching interests include crime analysis for problem solvers, applied statistics, research methods, and program evaluation in criminal justice. Her primary research interests include predictive analytic applications in criminology and criminal justice, survey methods, quantitative methods, and evaluative research. Dr. Ngo's most recent publication involves a comparative study of four statistical techniques — logistic regression, classification and regression tree, chi-squared automatic interaction detection, and neural networks — for their utility in predicting inmate misconduct. She can be reached at fawnngo@sar.usf.edu.

IoT Data Management Challenges: A View from the Trenches

by Pranav Shah, Suman Datta, Rekha Vaidyanathan, Sudhakara Poojary, and Vidyut Navelkar

All that has been said and written about the challenges associated with the Internet of Things (IoT) does not quite prepare you for the practical difficulties that crop up as you start implementing and deploying IoT solutions. Most of the publicly available knowledge about IoT challenges relates to high-level issues that are typically addressed through architecture and design decisions. One of our recent successful implementations, an enterprise-wide Remote Energy Management System (REMS), brought us face-to-face with an entirely new set of ground-level challenges, from data ingestion to data storage, to data processing, to data analytics and visualization.

> **Given the evolution of the organization and the diversity of its facilities (large campuses, single buildings, leased facilities, etc.), we expected the variety of data sources to some extent — though not fully, as we were to realize later.**

REMS, which was prompted by the organization's eco-sustainability goals, involved monitoring energy consumption across the organization's 100+ campuses spread all over India to find opportunities for energy saving. This meant instrumenting the infrastructure over all these campuses with over 10,000 meters and sensors to collect consumption data for various purposes (HVAC, lighting, computing, etc.) and correlating it with other relevant data such as occupancy, weather, and time. The system needed to provide a real-time integrated view of KPIs like energy consumption, consumption per occupant, and consumption per square foot at the enterprise and facility levels and across various load types and time frames. Now fully implemented, the system enables us to monitor consumption and identify abnormal consumption patterns to raise alerts in near real time, so as to trigger appropriate action. In the time that it has been operational, REMS has already delivered significant benefits in terms of reducing the organization's energy costs and carbon footprint.

In this article, we share our experiences related to data management in the course of implementing the system, both issues that we headed off at the pass and those that we discovered and addressed along the way.

MANAGING THE KNOWN: SYSTEM CHARACTERISTICS AND DESIGN CONSIDERATIONS

While the main focus of this article is to highlight the surprises that we had to deal with, below are a few observations and data design decisions that were especially relevant for achieving some of the desired system characteristics. We'd like to note that there were indeed data management challenges that we foresaw and preempted through appropriate design decisions (after all, we did carry out an elaborate architecting exercise!).

Scalability

Since we were dealing with a large, growing, geographically dispersed organization, it was clear that our solution would need to be scalable, with high performance and availability. This was achieved through the use of the TCS Connected Universe Platform, our cloud-based IoT platform. The underlying cloud infrastructure provided the levels of scalability, performance, and availability required to ingest and process data in near real time.

Storage and Processing

In addition to the performance enabled by the IoT platform and cloud infrastructure, other key decisions that impacted performance included those related to data storage and processing. We had to consciously consider how much historical data actually needed to be archived, and further, what processing could be carried out at the edge so as to reduce both storage and processing on the cloud. In our system, we found it very beneficial to pre-process and aggregate data at the edge before passing it to the cloud.

Variety of Data Sources

Given the evolution of the organization and the diversity of its facilities (large campuses, single buildings, leased facilities, etc.), we expected the variety of data sources to some extent — though not fully, as we were to realize later. This meant an exhaustive stock-taking of meters, metering topologies, and building management systems (BMSs) across locations, as well as developing connectors and agents as required. We knew that this would require talking to a lot of equipment vendors and manufacturers to obtain technical and operational details, and boy, were we right!

Varying Latency

Yet another aspect of variety was the varying latency of different data sources. While in the first phase of the system we had restricted ourselves to data from energy meters and BMSs, we later extended the system to include feeds from weather sources and employee feedback indicating whether they were uncomfortable with the temperature or lighting at their location at any point in time. Since the latency of these sources varied and there was always a lag between an action taken and its effect, we had to ensure that we took care of such variance when generating alerts.

Flexibility

Apart from basic scalability in terms of size and numbers, we also knew that REMS would need to be flexible enough to allow seamless addition of new data streams with minimal code change and minimal downtime. This again was largely aided by the use of our IoT platform, which helped in black-boxing the device layer from the other components of the system. Our subsequent experience of extending the system to cover newly added facilities and manage change in configurations — *all with zero code change and zero downtime* — proved the tremendous value that a good IoT platform brings to an IoT implementation.

Usability: The Operations Perspective

Based on our experience in software development, we ensured involvement of the right mix of people from operations and IT for understanding the requirements, especially from a usability perspective. While we got that bit right, we were grossly wrong about the enormity of the impact of this single decision. Involvement of operations had implications that went beyond requirements definition, deeply influencing the design and development of the system. Usability and user experience are not only about aesthetically designed screens and dashboards, nor only about the system's ease of use. They are ultimately about the ease of operations facilitated by the system and the value of the system as perceived by the user. This value went beyond decisions pertaining to the presentation layer and impacted other data-related design decisions such as:

- At what rate would data need to be ingested to be meaningful? (Is reading meter data every minute too much? Is reading meter data every day too little?)
- What transformations would be required on the data? (The meter data shows total energy consumed until the time of reading, and this would need to be transformed to energy consumed in a particular time slice.)
- What data would need to be stored? (At what granularity will the meter data be required for facilitating operations — by facility? by floor? by zone?)

In short, involving the operations team helped us address a number of data design issues early in the lifecycle that we may have overlooked or not fully appreciated otherwise. What further helped was that we had a customer who was not only highly technically evolved, but also very involved in the entire process of developing and implementing the system. (OK, we admit we got a bit lucky here!)

Involving the operations team helped us address a number of data design issues early in the lifecycle that we may have overlooked or not fully appreciated otherwise.

MANAGING THE UNKNOWN: "HEY, WE DIDN'T SEE THAT COMING!"

All planning, designing, and architecting aside, there were frankly quite a few situations that we had not anticipated — many of them related to data. The fact, as we now realize, is that these were probably not unique to our project and would surface to varying extents in some form or another across most IoT projects.

Missing Data

Infrastructure failure and misbehavior leading to missing data were things that we had not really accounted for — or at least foreseen to be an issue that could not be handled with some quick investigation and fixes. Like all good software engineers, our first reaction was "We need to get down to the root cause, and once we fix that, we should be good." While this approach did help us resolve some of the fundamental reasons for failures, we soon realized that in practice, the extent and variety of such points of failure (which could be at the sensor level, at the server level, or a result of communication failure) were simply far too great to be handled on a case-by-case basis. Accepting this reality meant defining a strategy to deal with missing data. The question was "What will we do when Murphy strikes"?

The answer lay in defining strategies for each data source, depending on the nature and usage of the data. Some missing data points could simply be ignored. For example, there were situations in which the missing data was due to the network's being down for days together, and it did not make any sense to try to impute the missing data points for such a long stretch of time. Then there were other missing data points that could not obviously be ignored, and we handled those by defining algorithms that helped us guesstimate the missing data points for data imputation. For example, in cases where we had missing data for a few hours due to failure of the network or proxy server, we decided to impute the missing values as soon as the next valid reading became available.

> **Consider all the possible combinations of space and time over 100+ campuses and try defining rules for that monster!**

Spiky Data

Spikes in data were another issue that we had to contend with. While these spikes could be interpreted as another manifestation of infrastructure failure and misbehavior, we realized that things were not always that straightforward. As we started analyzing the spikes, we found that while some of them reflected the right data values, there were others that were actually a result of infrastructure misbehavior. For example, switching on a UPS or a chiller led to a "good spike," since it represented an actual spike in consumption. On the other hand, there were "bad spikes" that occurred for no obvious reason and were clearly a result of erroneous behavior of the meter. The challenge, then, was to distinguish the good spikes from the bad spikes. This obviously meant building intelligence into the system to figure out this distinction, and that is what we did. As a very simplistic example of such intelligence, a spike of a few hundred KwH could be caused by devices being switched on and was interpreted as a good spike. On the other hand, a spike of thousands of KwH would probably indicate a bad spike.

Diversity of Data Sources

While we had expected that we would need to deal with a variety of meters, metering topologies, BMSs, and so forth, in hindsight we probably grossly underestimated the problems that such a diversity would pose. It was not only about developing connectors and agents to enable these components to talk to our platform using a variety of protocols (which we anticipated). What caused unanticipated data management challenges was the variety of data formats and measurement systems each of the underlying components used, and the varied data latency these components exhibited. This meant going deeper into understanding the behavior of each component type and programmatically handling such differences in data format and latency.

Interpreting Data

Given the diversity of data sources and the diversity of operations across campuses, we very quickly realized that interpreting the consumption data to distinguish between normal and abnormal energy consumption wasn't as simple as defining thresholds and rules that would apply across the organization. Normal and abnormal consumption, we discovered, were functions of space and time, and hence defining thresholds and rules manually was next to impossible. For example, what was normal consumption for a data center was not normal consumption for an employee office building. Similarly, the consumption pattern on a Monday differed from the consumption pattern on a Friday — and to add to the complexity, consumption patterns also varied through the hours of the day. Now consider all the possible combinations of space and time over 100+ campuses and try defining rules for that monster! The answer stared us in our faces — we had to build capabilities into the system that would allow it to learn from past data and define its own rules.

Data Traceability

Data collected from meters and other sensors in our system had to pass through a number of hardware and software components at the edge, on the platform, in the cloud, in the database, through the analytics platform, through the various apps, through the various reports and dashboards, and so on. Apart from handling the latency introduced by these multiple hops, the major challenge was ensuring that the data was traceable across all the components (from sensors to presentation layer) so that the component responsible for any data failure (missing or poor-quality data) could be identified early. This was done through the use of system-monitoring tools that provided reports on the availability and downtime of various systems and helped track the availability of various components. In addition to the system-monitoring tools, we also introduced customized reports that helped us monitor the availability of data and perform sanity checks on the data so as to highlight any missing or poor-quality data at the point of defect injection.

Collecting and Maintaining Master Data

Master data such as building details (area of different buildings, floors, zones, operating hours, geography, etc.) and metadata about meters (location, equipment type connected to, etc.) were collected manually by a dedicated team consisting of the administration and audit functions at each site. This entailed significant effort that we had not fully foreseen. Since this entire process was manual, there were also occasions where discrepancies crept in because of human error (incorrectly tagged locations, equipment connected to, etc.). This was further complicated by configuration changes that kept happening as a part of normal operations over time. Since all reporting and analytics were based on this master data/metadata coupled with the raw sensor data, it led to issues with reporting accuracy and quality. Once again, data availability and data quality reports came in handy in identifying such issues and getting them fixed through master data updates and rectifications, thus providing a feedback loop.

ABSTRACTING THE LEARNING: LESSONS IN IOT DATA MANAGEMENT

While, as a part of architecting an IoT system, you will obviously aim to build a system that is high on availability, performance, security, scalability, and the like, the complexity of the system is bound to cause surprises that you won't have anticipated. The large number of hardware and software components that come together in any reasonable IoT implementation result in a system that has multiple points of failure. Data management in IoT is no exception; our biggest lesson learned is that we need to be prepared for surprises. Regardless of how well you architect and design your system, there will be situations that you will not be able to preempt. The success of your system will depend on how it reacts to such situations. Specifically from a data management perspective:

- Be prepared for situations where (for whatever reason) gaps occur in the data stream. Determine the criticality of these gaps and implement a strategy within your system to deal with them.
- Spiky data may not always be a result of system misbehavior. You will need to figure out ways to distinguish between spikes resulting from natural behavior of the system and those resulting from misbehavior — and deal with the two differently.
- It is important to decide on the amount and format of data that needs to be stored on the cloud. How much history do you actually require and in what form? Do you really need every data point, or can you make do with an aggregate? Or are you only interested in every instance of a transition in value? This decision will depend purely on the demands of your application.
- In many systems, there will be opportunities to convert and pre-process data at the edge level before passing it to the cloud for further processing and analytics. Sometimes it may also be possible to do simple analytics at the edge. Every such opportunity could lead to significant impacts on storage, performance, and cost.
- Be prepared to deal with diversity, not least at the data level. Devices in the field, even if they are meant for the same purpose, may come from different vendors and be of different vintages. This could lead to a situation where different sensors, meters, and so on, use different units (e.g., WH vs. KwH) or report at different frequencies (e.g., every 10 minutes vs. every 15 minutes). You will need to deal with such diversity during ingestion, storage, and processing.
- In most cases, you will need to deal with legacy equipment and data sources. As part of handling the diversity, plan to establish connections with the original equipment vendors and work with them to get technical and other internal details. In some cases, you may also need the vendor to make certain changes.
- Traceability of data across components is critical. The dataflow from sensors to the presentation layer in an IoT application passes through various hardware components (sensors, devices, gateways, etc.) and software components (converters, code on the edge, code on the cloud, etc.), and the geographical spread of these components adds to the complexity. Depending on your needs, you may have to create data availability and quality reports to help trace the flow of data through the system and highlight or

raise alerts in case of possible failures. You may also need system-monitoring tools to keep tabs on the availability of at least a few of these components.

- Behavior of most real-life systems will vary with time and space. Hence in a system of any reasonable complexity and scale, it will be virtually impossible to manually define thresholds for distinguishing between normal and abnormal behavior. You will need to find ways by which the system can learn from historical data in order to define the thresholds and rules automatically.
- Unlike most IT systems, master data about *things* may not be easily available and may be hard to collect. Depending on your system, you might have to budget for significant effort for this activity. Further, if the real physical world that you are dealing with is prone to changes, you will need to work out ways to detect such changes and update master data accordingly.
- IoT is as much about operational technology (OT) as information technology (IT). Involvement of customer teams, especially operations, is critical for IoT success. Unlike other IT systems, this involvement will be required across the lifecycle and will have a significant impact on the usability and design of the system. Early and complete involvement of operations helps in making better design decisions, including those related to data.

While many of these lessons may seem obvious in hindsight, the fact is that we did not have these insights when we started building REMS. They came about as important learning for us through our experience of implementing the system. More importantly, we have benefited from this learning and experience in other IoT implementations that we have since been involved with. We hope that this attempt to abstract and share our lessons learned will likewise be of value to you in your own IoT journey.

Pranav Shah leads analytics in the IoT area at Tata Consultancy Services (TCS). He has extensive experience in business intelligence and analytics in the telecom, retail, utilities, and government verticals. He may be reached at pranav.shah@tcs.com.

Suman Datta has over 14 years of experience in various roles, such as Technical Architect, Project Manager, and Developer, in the areas of IoT, mortgage banking, and bioinformatics at TCS. He is currently working on IoT analytics and machine learning. He may be reached at suman.datta@tcs.com.

Rekha Vaidyanathan works in the area of IoT analytics at TCS with over eight years of industry experience in software development as Analyst and BI Application Consultant, and four years of research experience in text mining. An Associate Member of the Information Retrieval Society of India, she may be reached at rekha.vaidyanathan@tcs.com.

Sudhakara Poojary leads IoT delivery at TCS. He has over 19 years of experience of working on and leading large delivery projects for global customers across verticals. He may be reached at sudhakara.poojary@tcs.com.

Vidyut Navelkar is Principal Evangelist for IoT at TCS. He has vast experience in developing organizational capabilities and successfully leading various large programs. He may be reached at vidyut.navelkar@tcs.com.

Exploratory, Embedded, Enabled:
The Three E's of IoT Data Management Maturity

by Donald E. Wynn, Jr. and Renée M.E. Pratt

IOT IN TODAY'S WORLD

With the constant improvements in communications technology, the demands on connected things (ATMs, check-in systems, digital-operated locks, etc.) have increased, and the barriers to their adoption have greatly diminished. With these improvements, the added data collected from back-end systems, and the increased interest in data analyses, the Internet of Things (IoT) has emerged as a powerful paradigm among digital business strategies. IoT refers to the connection of a wide range of physical devices and objects to the Internet, including sensors, actuators, industrial equipment, consumer products, and more. These devices collect, store, analyze, and act upon data in accordance with some specified goal. These things will learn and gain information and knowledge directly from data received from other things. Ultimately, it is expected that things will become the main generators and consumers of data, instead of human beings.[1]

As IoT becomes more prevalent among organizations and individuals across the globe, the devices will generate a tremendous amount of data from many sources, in many different formats, and at a potentially high velocity. Although predictions regarding the level of adoption of connected things vary widely, Cisco estimates that as many as 50 billion devices will be connected by the year 2020.[2] Other predictions suggest that there could be as much as 500 zettabytes of Internet traffic attributable to these devices by 2019.[3]

These figures are not surprising when we consider the example of GE Aviation, which is developing the means to analyze data from its aircraft turbine engines while in flight. While flying, a typical engine can generate nearly 10 terabytes every 30 minutes![4] Eventually, this data may be analyzed *while the plane is in flight* to optimize flight plans and engine performance in order to increase flight safety and save fuel costs.[5] GE's rival, Rolls Royce, is also investing heavily in similar technologies.[6] In healthcare and other industries, personal IoT devices such as wearables and ingestibles promise to revolutionize the amount and quality of data that can be collected from an individual.

A number of organizations are already seeing significant cost savings and improved services from IoT deployments. For instance, the city of Chicago uses a network of sensors to measure such things as full trash cans, water conditions at city beaches, and traffic conditions.[7] Integrating this data from such a wide range of sources and formats into a cohesive platform requires a number of changes to traditional data management and analysis tools, including such aspects as common storage and analytical platforms and interoperability between multiple vendors.

In this article, we propose a maturity framework that addresses the data management challenges for organizations seeking to develop and deploy IoT platforms. We will highlight this framework through a data management lens, exposing the implications and requirements for data collection and storage, data integration, and data analysis. Finally, we will suggest a set of prescriptions for organizations wishing to improve their level of IoT maturity in order to capitalize on this phenomenon.

THE THREE E'S MATURITY FRAMEWORK

Organizations that successfully adopt enterprise-level technologies progress from the initial adoption of the technology to routinization among the organization and finally, infusion of the technology into governance, social systems, and work processes. In this final stage, the organization is able to use the new technology to its fullest potential to increase its performance and effectiveness.[8] Thus, the maturity of a given technology within an organization can be based on the degree to which it is successfully assimilated into the overall structure and processes. In much the same way, enterprise-wide IoT technologies will achieve effective assimilation through their management of generated data. Hence, the level of IoT maturity in organizations should be assessed using a framework that takes into account an organization's ability to develop and assimilate standards and practices that address each of the three data management processes. Our proposed framework features three levels of maturity: exploratory, embedded, and enabled. We will illustrate these levels with brief examples from the manufacturing and healthcare industries.

Exploratory

At the *exploratory* level, organizations introduce trial IoT projects into their information architectures. As is often the case when introducing emerging technologies, there are few concrete plans for how they will improve the financial and operational performance of the entire organization. Instead, the key focus is on automation of several functions or business processes. Such organizations acquire IoT devices in an attempt to capture early benefits and functionality as promised by case studies and vendor claims, but these devices are not fully integrated with their legacy systems and data, or even with other IoT devices in separate projects. Instead, they are configured in standalone systems that require intervention for aggregating and manipulating data across devices. Although initiated with the full intention of attaining a positive return, many projects will not be raging successes in terms of the stated financial or operational goals.

At this point, the biggest benefit will be the knowledge and experience gained throughout the projects' lifecycles. This knowledge of the organizations' ability to develop and leverage capabilities incorporating IoT technologies with their people and business processes will prove to be quite valuable as they attain higher levels of maturity. For example, many hospitals and their patients are currently experimenting with a number of wearable devices to record basic health behaviors and statistics, while manufacturers are deploying sensors that can send data to a central repository, where it can be analyzed to yield information about production efficiency and performance.

Embedded

Organizations at the *embedded* level of IoT maturity begin to integrate IoT devices and the resulting data into their existing technological architectures and workflows, instead of restricting the deployment to standalone projects. These organizations leverage their understanding of the people, processes, and technologies developed in the previous exploratory IoT projects. At this stage, IoT projects are more commonplace and routine. This leads to an increased understanding of the capabilities of the new technology, along with an increase in the knowledge available to the organization. As a result, these organizations are more likely to attain a positive ROI from IoT deployments. Based on these experiences, standards and policies regarding IoT implementations begin to evolve. Embedded platforms typically incorporate a wide range of IoT devices at multiple points throughout the organization. With the increased comprehensive analysis and reporting generated, the primary challenge is integrating these efforts and platforms to produce better insights and results.

Beyond merely gathering data for analysis elsewhere, healthcare providers at this maturity level are starting to integrate the data from IoT devices into their existing software platforms, and thus into their patient care initiatives. Manufacturing firms at the embedded level may deploy self-aware machines that, for instance, not only gather data about their performance, but also store job information to identify future performance improvements.[9]

Enabled

Eventually, organizations learn enough about the requisite technological and organizational characteristics of IoT implementation to reach the *enabled* level. These organizations are able to establish guidelines, standards, and best practices for incorporating the technology into their strategic and operational plans. IoT projects are thus more likely to be larger in scope and scale, more widely dispersed throughout the organization or specific business units, and more integrated with the overall IT architecture. In many cases, this enhanced integration even includes a reliance on IoT to revamp or augment specific operational processes or the overall business model itself. In essence, IoT becomes a significant component in the organization's competency base. Enabled platforms are fully interoperable with respect to the amount and sources of data they are able to incorporate from IoT devices. They are also able to prescribe actions to be taken by organizations and consumers alike.

Along with an analytical engine containing diagnostic and treatment suggestions, an enabled healthcare platform may suggest immediate lifestyle changes in an effort to mitigate any adverse clinical outcomes in advance, and even interact with other devices to change the patient's behavior.[10] Enabled manufacturing firms may use stored job data to alter their production capacity to yield such benefits as optimized efficiency, improved asset utilization, reduced lot sizes, and reduced downtime.[11]

Scaling the Three E's Ladder

Over time, successful implementations feature a movement to higher levels of maturity as organizations become increasingly reliant on the benefits of IoT technologies. In so doing, they may experience a number of significant changes in organizational structures, business processes, and competitive strategies. For example, GE's new approach to collecting and analyzing data on its aircraft engines has led to a completely new business model of leasing (rather than selling) these engines to customers. This approach would not have been possible without its IoT-enhanced capacity for managing the engine data. Likewise, the prolific use of IoT in the healthcare industry is introducing major changes in the way healthcare providers track costs and payments, diagnose patients, and assign treatment options. Some have even envisioned scenarios in which much of a patient's care is conducted in his or her own home, surrounded by a number of connected devices, which will lead to clinical changes as well.

These changes will also be reflected in the organization's technological architecture. Because of their highly distributed nature and the significant variety of generated data, IoT networks present several data management challenges that organizations will need to address in order to fully capture the value from their IoT investments.

IOT DATA MANAGEMENT

As with other data-intensive technologies, collecting and analyzing data from IoT deployments requires sound data management strategies and tactics. Collecting and analyzing this data can be generalized to include three distinct processes: (1) data collection and storage, (2) data integration, and (3) data analysis (see Figure 1). Let's examine each of these processes at each of the maturity levels.

Data Collection and Storage

The data collection and storage process includes tools to capture, store, and access data in centralized data stores, whether they utilize relational DBMS, NoSQL, or any other form of data storage technology. This process also

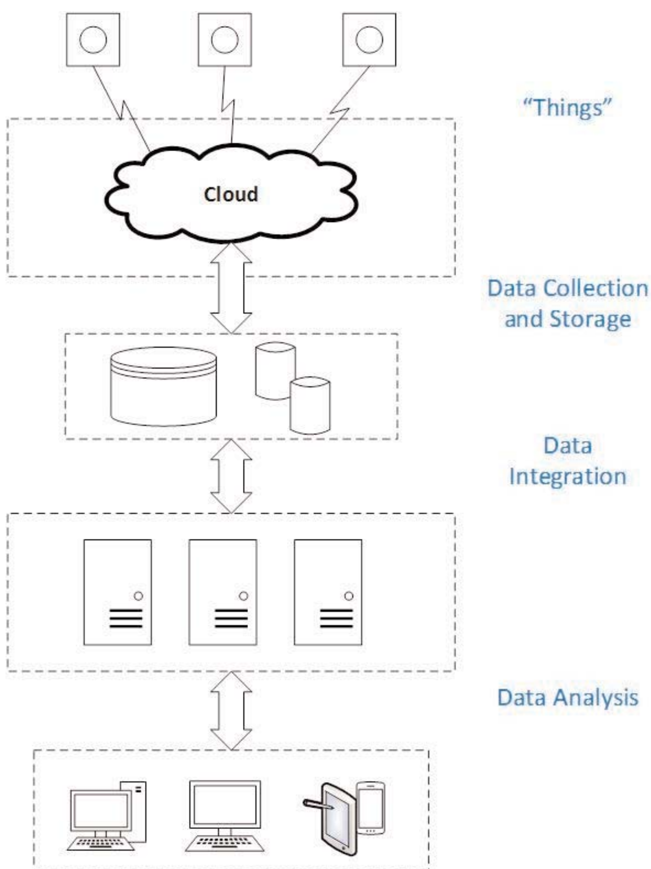


Figure 1 — IoT data management processes.

includes any storage and filtering inherent to any IoT devices that may have such capabilities. At a technical level, data collection and storage also consists of efforts to define both the physical and logical design of any database technologies to be used for IoT data.

In the simplest cases, IoT devices are little more than isolated data sources that stream data across the network. With a small number of these devices, the impact on the network is likely to be minimal. As noted in our discussion of the exploratory level, trial IoT projects and standalone systems are often connected to the network with very little attention to long-term strategy and placement. However, as more devices are added to the network, the IT staff needs to decide how the network can be segmented and distributed to ensure enough bandwidth to handle the increased traffic. This necessitates the capacity to monitor dataflow patterns in order to distribute the traffic effectively.

At the embedded level, organizations have much better data collection and storage capacities. The emergence of edge computing introduces the opportunity to handle some distributed storage and filtering of data before it is stored centrally. These organizations may also begin to develop experience with the design and development of distributed storage clusters. Rather than having to retrieve data from the data warehouse, they enable devices to access and utilize data locally. In addition, several of these organizations will deploy tools such as Hadoop, Pig, Spark, and Hive as part of their data management solutions.

At the enabled level, information architectures will likely include a hybrid combination of centralized data stores (e.g., a data warehouse or data lake) and distributed storage arrangements, with Hadoop and related tools to manage the efficient collection and storage of this data. IoT-enabled organizations will employ a wide range of sensors and controllers that are then integrated into the enterprise architecture, not merely as data sources but as fully capable components of the overall information systems platform. As a result, there is little need to fully transmit all streaming data from the device. Instead, the data is aggregated or filtered as required by other systems before being loaded into the data warehouse. These enhanced capacities allow the initial conversion, aggregation, and, in many cases, storage of IoT data to be handled remotely instead of transmitting everything centrally. This reduces some of the need to expand bandwidth on the back end.

Data Integration

In order for data to be made available for enterprise-wide analytics and decision making, the data from IoT devices needs to be combined with legacy data and other sources of data throughout the organization. Data integration consists of the tools and processes that allow IoT data to be combined with other data to generate opportunities to create meaningful information through analytics.

Although the data in exploratory-level organizations is generally stored in a centralized manner, it is likely not integrated with the organization's legacy data and systems. The tools for transforming and loading the data into the database are focused on speed and reliability. At the outset, organizations rely on traditional batch-oriented extract, transform, and load (ETL) processes to load the data into the database. However, the data generated by IoT devices is often streamed continuously, necessitating the adoption of real-time tools to manage the loading and integration processes. Exploratory organizations begin to develop the capacity to handle the integration of this data using these tools.

Thanks to a number of tools and techniques, as levels of maturity increase, the IoT data begins to integrate with remaining organizational data into a single application platform. One technique for making this a reality is master data management. It allows for a trusted and reliable reference for the IoT data, making it easier to share and integrate with other data across the organization. Another effective data management practice is the adoption of devices using established data standards, which would allow analytical engines and other reporting software to access data from IoT devices more easily. A third technique is the adoption of APIs and middleware tools. This aims to make it easy for these organizations to integrate data across various standards and protocols.

While embedded organizations do not necessarily have a full set of tools at their disposal, they have experimented with combinations of these tools as components of their data management platform by this stage, achieving an increased degree of success over time. Enabled organizations have deployed fully combined technologies, practices, and standards, leading to the establishment and maintenance of a fully integrated data platform. This platform includes a combination of tools and technologies that promote the integration of IoT and legacy data sources. Organizations at the enabled level are flexible enough to quickly and effectively utilize the data from each device as needed for analysis and reporting. These organizations are also willing to experiment with these tools to develop new combinations to allow for subsequent comprehensive data analysis.

Data Analysis

Data analysis tools and technologies are used to convert the raw data into more usable information. For IoT data, this analysis — whether ad hoc, descriptive, predictive, or prescriptive — is complicated by the attributes of the data itself. The tools and techniques employed at this stage must be capable of handling a continuous stream of structured and unstructured data, frequently in real time. In addition, the results of this analysis must often in turn be stored in a way that allows them to be integrated in future analyses along with the raw data itself.

Most of the data analysis at the exploratory level is descriptive and aggregative. The focus at this stage is largely explanatory, using the data to monitor and report on existing processes and events. Owing to the isolated silos in which IoT projects are implemented and the presumed experimental nature of early projects, there is little emphasis on combining IoT data with other sources of data to generate more complex analytical outputs.

Organizations at the embedded level have an increased focus on predictive analytics. This allows them to model the future performance of a given product or process. Organizations at the enabled level of maturity are not only able to generate descriptive, predictive, and prescriptive results using a centralized platform, but also to utilize the analytics and intelligence capacities built into IoT devices at the edge of the organization. These devices can even generate actions for themselves without communicating with a centralized data processing environment. For instance, one classic example is a traffic signal that can make decisions at a given intersection based on the number of cars it senses at that moment, based on intelligence built into the signal. The resulting bandwidth is more efficiently utilized and results are generated faster than with other solutions. In addition, self-governing devices have the potential to greatly change the organization's strategic and operational execution. Recall the GE Aviation example, in which the manufacturer was able to change its competitive strategy by establishing a fundamentally different relationship with the airlines that purchased its engines. This may ultimately lead to a completely different profit model and industry strategy as GE competes with Rolls Royce.

Table 1 summarizes the three levels of adoption, along with each data management attribute. This is not meant to imply that an organization is simultaneously at the same level across each of the three data management processes. Rather, it is more likely that organizations

	EXPLORATORY: Early Pilot Trials and Initial Forays	EMBEDDED: Modified Workflows and Integrated Data	ENABLED: Fully Integrated Data and Processes
Strategic and Operational Foci	Automation of simple manual processes	Integration of data and business processes; metrics beginning to emerge	Business model or operational changes
Data Collection and Storage	Simple configurations for IoT devices broadcasting to central storage	Emergence of edge computing, intelligent devices, and distributed storage architecture	Hybrid architecture consisting of centralized and distributed storage
Data Integration	Centralized, but rarely integrated with legacy data and systems	Some distributed storage and filtering; orderly integration with legacy systems (e.g., master data management)	Distributed storage and filtering of data
Data Analysis	Mostly descriptive analysis, albeit in isolated silos	Predictive data analysis beginning to reflect the integration of data	Prescriptive analysis and self-governing devices that act on the analysis without human intervention

Table 1 — Three levels of IoT maturity.

will exhibit different levels of maturity at any given point in time. For example, a firm may currently be at the embedded level in terms of storage and analysis and merely exploratory in terms of integration. In time, the same firm might reach the embedded level across all three processes as it completes implementation projects and upgrades its capabilities.

As IoT devices and standards continue to evolve, so too will new data storage, integration, and analytics tools.

RECOMMENDATIONS FOR ACHIEVING TOMORROW'S IOT MATURITY

Organizations at the highest levels of maturity with a particular technology are able to fully leverage its capabilities toward the achievement of their objectives. Organizations adopting IoT platforms will be no exception. By aligning their business and IT strategies and processes, these organizations may be able to create opportunities to develop meaningful competitive advantages. At its core, IoT focuses on the creation and use of an increased volume and velocity of data. Consistent with the framework described above, we offer four recommendations for organizations seeking to move toward IoT maturity:

1. **Implement flexible architectures that can accommodate enhancements in each of the three data management processes.** As IoT devices and standards continue to evolve, so too will new data storage, integration, and analytics tools. In addition, new hybrid forms of data management architectures will emerge to take advantage of such technologies as edge computing and distributed storage. Organizations should find ways to evaluate these new tools and strategies with respect to incorporating them in a roadmap to achieve their business and technological goals.
2. **Embrace and encourage the growth of internal knowledge with respect to IoT components.** As the IoT becomes more widespread, personnel with skills and knowledge in its underlying technologies will become critical to developing IoT applications, policies, guidelines, and best practices. Organizations will need to hire qualified personnel and/or fund training in order to fill any gaps that may exist between desired and current levels of expertise.
3. **Establish and constantly monitor the storage, transmission, and use of IoT data.** With the deployment of IoT devices, it is necessary to avoid the potential chaos that could result from the continuous streaming of unstructured and structured data. By tracking such measures as segment bandwidth, available data storage, and percentage of data used, organizations can make sound decisions regarding their current and future architecture to maximize its efficiency.
4. **Seek ways to incorporate the advantages of IoT implementations into the strategic and operational processes of the organization.** This includes finding new uses for the data generated across the organization to improve existing business processes, lower organizational costs, monetize internal data, and create competitive strategies and opportunities.

By heeding these recommendations, organizations should be able to manage their progression toward organizational and technological maturity.

CONCLUSION

IoT promises to revolutionize the way organizations deliver value to their stakeholders in the near future. It also increases the opportunities for organizations and consumers alike to reduce costs, save time, improve efficiency, and take advantage of new capabilities.

But the outlook is not all rosy. Security is a major concern for IoT devices — which are always connected and always on — and their shared data. As a result, there is an increased need for risk assessment and management to protect the organization as well as its linkages with other organizations. Interoperability is another concern. The fast pace of innovation by multiple vendors and providers results in a plethora of interoperability issues among various IoT devices, platforms, and organizations. Until we can arrive at a set of commonly accepted standards for connectivity, security, and data format, this concern may continue to fester.

Organizations can use the three-level IoT maturity framework we propose in this article to assess their current maturity and set desired goals with respect to the benefits they hope to derive from their IoT program. They can also capitalize on the recommendations we offer above to position themselves to achieve a positive return from their IoT investments.

ENDNOTES

[1]Qin, Yongrui, et al. "When Things Matter: A Survey on Data-Centric Internet of Things." *Journal of Network and Computer Applications*, Vol. 64, April 2016 (www.sciencedirect.com/science/article/pii/S1084804516000606).

[2]Evans, Dave. "The Internet of Things: How the Next Evolution of the Internet Is Changing Everything." Cisco Internet Business Solutions Group (IBSG), April 2011 (www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf).

[3]"Cisco Global Cloud Index: Forecast and Methodology, 2014–2019." Cisco, 2015 (www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.pdf).

[4]Palmer, Danny. "The Future Is Here Today: How GE Is Using the Internet of Things, Big Data and Robotics to Power Its Business." *Computing*, 12 March 2015 (www.computing.co.uk/ctg/feature/2399216/the-future-is-here-today-how-ge-is-using-the-internet-of-things-big-data-and-robotics-to-power-its-business).

[5]Robb, Drew. "Will the Internet of Things Revolutionize the Aircraft Industry?" *Forbes*, 23 June 2014 (www.forbes.com/sites/ptc/2014/06/23/will-the-internet-of-things-revolutionize-the-aircraft-industry).

[6]"Rolls Royce: Internet of Things in Aviation." HBS Open Knowledge, 22 November 2015 (https://digit.hbs.org/submission/rolls-royce-internet-of-things-in-aviation).

[7]Engebretson, Joan. "Chicago CIO Shares Internet of Things Success Stories." Telecompetitor, 17 October 2014 (www.telecompetitor.com/chicago-cio-shares-internet-things-success-stories).

[8]Cooper, Randolph B., and Robert W. Zmud. "Information Technology Implementation Research: A Technological Diffusion Approach." *Management Science*, Vol. 36, No. 2, February 1990 (http://pubsonline.informs.org/doi/abs/10.1287/mnsc.36.2.123).

[9]Teresko, John. "Preparing for the Factory of the Future." *IndustryWeek*, 15 November 2007 (www.industryweek.com/articles/preparing_for_the_factory_of_the_future_15330.aspx).

[10]Kvedar, Joseph C. *The Internet of Healthy Things*. Partners HealthCare Connected Health, 2015.

[11]"Leveraging the Industrial Internet of Things (IIoT)." *IndustryWeek*, 31 March 2016 (www.industryweek.com/cloud-computing/leveraging-industrial-internet-things-iiot).

Donald E. Wynn, Jr. is an Associate Professor in the School of Business Administration at the University of Dayton. He holds a PhD in business administration from the University of Georgia. His research appears in journals such as MIS Quarterly, MIS Quarterly Executive, Information Systems Journal, Cutter IT Journal, Journal of Organizational and End User Computing, Communications of the AIS, *and* Journal of the Academy of Marketing Science. *Dr. Wynn has published research articles in a number of areas, including open source software, big data and analytics, electronic health records (EHR) software, technological ecosystems, and research methodologies. He can be reached at dwynn1@udayton.edu.*

Renée M.E. Pratt is an Assistant Professor at the University of Massachusetts Amherst in the Isenberg School of Management in the Operations and Information Management Department. She received her doctorate from Florida State University in management information systems. Dr. Pratt completed research in 2013 on the implementation and post-adoption processes of electronic medical record software under a Fulbright Scholar Grant in Germany. Her research interests include post-adoption diffusion behaviors, traditional and clinical enterprise systems, healthcare IT, and the psychological contract. Dr. Pratt's work has appeared in numerous publications, including MIS Quarterly Executive, Cutter IT Journal, Information Systems Journal, Journal of Information Technology, *and* Information Systems Journal of Education, *among others. She can be reached at rpratt@isenberg.umass.edu.*

Leveraging EA and IoT Synergy for Digital Transformation

by Tushar K. Hazra

Over the past few years, the business world across the globe has been witnessing a new kind of change. It is unique, it is disruptive, and more importantly it is constant. This change is about the transformation of an entire enterprise and its ecosystem. It encompasses the convergence of all the digital assets, devices, systems, and applications as well as communication technologies and networks any enterprise uses. We commonly refer to this change as "digital transformation."

EA AND IOT: BETTER TOGETHER

Most business and IT leaders today are embracing the Internet of Things (IoT) to catalyze the digital transformation across their enterprises. The IoT is empowering businesses and consumers to connect and converge their world of physical objects with computer systems, applications, and services, and — most importantly — to handle the deluge of data they collect, consume, and analyze. In this context, the role of enterprise architecture (EA) is highly significant. While EA and IoT are not similar disciplines, I believe there is an inherent synergy between them, stemming from the fact that EA and IoT each support digital strategies for disruptive business operations. They deliver a converged enterprise that is collaborative, connected, and often socially infused.

Traditionally, EA has focused on delivering a set of guiding principles, frameworks, reference models, blueprints, and roadmaps to support operational excellence as well as strategic business and IT alignment goals. Today this focus is shifting toward leveraging collaborative, agile, disruptive, and innovative approaches to executing EA practices for digital transformation. Like many thought leaders, I believe that in order to successfully implement a new change, a new EA must be proactive and customer-oriented. This new EA must be innovative enough to deliver tangible business results more consistently and more frequently to capitalize on the IoT opportunity. EA and IoT together help enterprises to leverage their capabilities (people, process, and technology) while establishing mechanisms or conduits for the digital transformation.

Here are just some of the ways that EA and IoT complement each other:

- EA can decide on a set of technology standards, platforms, and network protocols (e.g., Wi-Fi, TCP/IP, RFID) that solution developers implementing the IoT devices can use. The IoT devices and sensors can employ these standards and protocols to connect with the enterprise and thereby enhance the quality of service and experience for the device users.
- The EA team can provide a set of proofs of concept or prototypes that solution developers can leverage to expedite IoT implementation.
- The primary focus for EA and IoT today is data analytics and how data needs to be ingested, processed, analyzed, disseminated, and subsequently used by data scientists or decision makers. Emerging technology trends such as social computing, cloud computing, mobility, and big data add key benefits to the EA/IoT combination. For example, cloud can facilitate the connection of IoT devices, while big data analytics can help professionals observe the behavior of customers using those devices.

Many companies in the healthcare, transportation, retail, manufacturing, and pharmaceutical industries are already embracing the concepts of EA and IoT together in their digital transformation initiatives. One of them is the company I worked for, which I will call "Health-USA" (a fictitious name to maintain anonymity). Over the past three to four years, the company has been trying to expand its operations with a new set of services and solutions for patient-centric care delivery, care coordination, and clinical decision support. Recently, the company decided to implement IoT devices and sensors as part of its strategic IT plan. In the remainder of this article, I present a case study of Health-USA's experience leveraging the combination of EA and IoT in its digital transformation effort and share a set of effective practices gleaned from that effort.

CASE STUDY: DIGITAL TRANSFORMATION AT HEALTH-USA

Health-USA is a multimillion-dollar healthcare IT service provider with more than 60,000 public sector employees; it has been serving US citizens for more than a decade. In 2013, Health-USA acquired two technology firms and a training provider, all involved in healthcare delivery services. Once the newly acquired organizations were integrated within the enterprise, senior leadership decided to incorporate medical technology devices such as life management (measuring and monitoring) gadgets and wearables into the company's enterprise solutions and services portfolio.

To this end, Health-USA formalized its partnership with a recognized manufacturer of wearable IoT devices. The hardware component of the chosen device consists of two sensors: one to wear as a contact lens, and the other as a patch that is capable of taking the patient's blood sample (preferably from an arm), temperature, and blood pressure. The software component of the device is a mobile app that Health-USA wants to develop inhouse and deploy using iOS, Android, and Windows platforms for versatility. With these components in place, Health-USA aims to provide services to more than a million diabetes and hypertension patients across the US within a year. Initially, the company plans to perform a clinical trial with 1,000 patients with Type 2 diabetes in the mid-Atlantic region.

The Vision

The CIO, in collaboration with senior leadership, led a set of strategy sessions to formalize and kick off the enterprise-wide digital transformation. For Health-USA, the primary business goals of the digital transformation included enhancing the customer experience and improving the quality and efficiency of service provided to patients and their families. It decided to get into the wearable IoT device–based services business in order to:

- Serve new and existing customers (diabetes patients with hypertension)
- Strengthen Health-USA's current market position
- Gain a competitive edge over other service providers in the same field

Once he socialized the strategic vision and mission among the rest of the executives, the CIO asked the chief architect and the EA team (of which I was a member) to define an actionable strategic plan so the EA team could help the entire enterprise manage its digital transformation effort while facilitating the IoT implementation.

The Plan

The EA team began its engagement with two business units as early adopters of the digital transformation and IoT implementation strategies: Clinical Decision Support (CDS) and Health Informatics (HI). It then planned to get the Public Health Affairs (PHA) and Enterprise Health Solutions (EHS) business units involved within the next three months.[1] The Health-USA leadership team directed both CDS and HI business units to work together with the EA team as a unified team to collaborate in developing the strategy, a roadmap, and an implementation plan for integrating the wearable IoT devices into the enterprise services portfolio and digital transformation initiative. The EA team devised an architecture strategy to:

1. Collect the data from the wearable devices using mobile apps and a Wi-Fi network
2. Ingest the raw data to an application in the CDS business unit
3. Store the raw data in secure cloud data storage

To harness the power of big data and data analytics, we selected a commercially available analytics tool to determine the quality, relevance, validity, and authenticity of the data.

The EA team made this decision based on current industry trends and research on commercially available tools and techniques for emerging technologies such as big data, cloud, data analytics, and IoT. We recognized that the use of private and secure clouds would enable us to control access to and protect the data collected (and shared) from the medical devices as well as the sensors. To harness the power of big data and data analytics, we selected a commercially available analytics tool to determine the quality, relevance, validity, and authenticity of the data. This subsequently allowed us to reduce the volume and velocity of the ingested de-identified data from the wearable devices before storing the data in a private cloud.

The CDS and HI business units plan to acquire and share a cloud-based data management and integration system to validate, process, cleanse, analyze, and disseminate the data to internal and external users for further analysis and informed decision making. Key external users include clinical research organizations, physicians, patients, and care delivery professionals. Internal users include data scientists, medical research scientists, and analysts from CDS and HI, as well as the PHA and EHS business units.

Identified Needs

During the initial strategy meetings the CIO had with his peers (from CDS and HI) and direct reports (including the EA team), my colleagues and I discerned the following needs:

- **Facilitating the adoption of the corporate vision of digital transformation across the enterprise,** specifically for the CDS and HI business units and associated technology organizations identified as early participants in the digital transformation. This was new territory, as the team members had little or no prior experience with the IoT or the use of mobile medical devices.
- **Recognizing the gaps in business and technical capabilities** while a number of capabilities related to leveraging the IoT were still under consideration. Key challenges with regard to leveraging the IoT included the privacy and security of the data (specifically for protected health information [PHI]) and the interoperability of the devices with existing applications, just to name a few.
- **Reestablishing the EA program and its governance** (charter, principles, review bodies, and processes) while adjusting its way of engaging with any enterprise-level initiative "on the fly" in order to meet the business needs and goals. The EA team concentrated on promoting the adoption of IoT (strategy, roadmap, and plan) by:
  - Offering technology briefings, proofs of concept, and vendor demos
  - Cultivating EA strategies for data protection and data access control
  - Facilitating sessions to make associates aware of issues and risks related to wearable medical devices and sensors and to familiarize them with the processes for resolving patient safety and security concerns
- **Engaging and empowering the right experts and resources to deliver business results and operational excellence successfully.** The most critical element here was not just ingesting or collecting the data from the IoT devices, but also making the data relevant and available for analytics by the data scientists and appropriate decision makers. The EA team guided CDS and HI business teams, application architects, application developers, and technology partners to develop and test a number of IoT, big data, cloud, data analytics, and mobile application development–related technology pilots, proofs of concept, and prototypes.

The EA team made a few additional observations while using IoT-based medical devices in conjunction with health informatics, mobility, and cloud computing. Since Health-USA already deals with PHI-based clinical and health informatics as well as biometric data, the management and analytics of the data discovered and captured from the IoT devices have been critical. In order to leverage cloud-based storage, we needed to secure the data in transit and at rest. Data loss prevention was another important issue for us. We considered multi-factor authentication for the data access control and NIST-recommended virtual disk encryption technologies to protect the devices. Requirements for compliance with government regulations, healthcare policies, and mandates intensified further since we were supporting US public sector clients; this added a new layer of security and privacy constraints for us. The lack of standards for connecting the IoT devices continues to pose challenges for the initiative today.

Essential Questions in Four Phases

In order to establish the credibility of the EA discipline, we cultivated a simple way to abstract the major elements for our digital transformation. As shown in Figure 1, we prepared a set of essential questions under the labels of Who, Why, What, Where, How, and When that we would ask ourselves (the EA team), business and IT sponsors, stakeholders, and business and IT teams over the four phases of the initiative: strategize, organize, socialize, and institutionalize. (See Table 1 for some sample questions.) We used the answers derived from the questions to develop different views and models of the transformation. The views and models also included the information pertinent to using the IoT, cloud computing, and big data analytics. The aforementioned four phases enabled us to develop a holistic approach to unite the technology, processes, interoperability, and security for the enterprise:

1. **Strategize.** During this phase, the EA team engaged senior leadership (sponsors and stakeholders) to get a clear understanding of the strategic vision. We identified the key business drivers for and desired benefits

from our digital transformation journey. We started this phase in parallel with the CIO's strategy meetings mentioned earlier. We frequently assessed our readiness for initial digital transformation projects and evaluated our ability to use the IoT (devices and connectivity), cloud, and data analytics effectively. We focused on the Who, Why, and What questions during this phase. The EA team's primary intention in conducting this exercise was to recognize the business vision and identify the capabilities needed to leverage new and existing digital assets.

2. **Organize.** In this phase, we started planning, blueprinting, and developing roadmaps for a set of projects that were prioritized according to their impacts on (and possible disruptions to) business operations. Although we primarily focused on Who, What, and Where questions, we also recognized certain specifics of How and Where. This exercise allowed us to evaluate our existing processes, technology, and regulatory compliance requirements. During this phase, we also laid out the initiative's communication plan, while setting the rules of engagement for the EA team to perform architecture planning, capability assessment, and architecture reviews. We established our relationship with the medical device, cloud, and data analytics tools vendors through technology briefings, demos, and evaluation and selection processes. The primary intention here was to pinpoint the gaps in business and IT capabilities and secure the new resources needed for digital transformation.

3. **Socialize.** We actually launched this phase to instill the concepts of the digital transformation in the organization. We cultivated the idea of the patient's use of wearable medical devices (as part of the IoT) and associated security- and privacy-related requirements. We also discussed the options for using cloud and data analytics across the enterprise early on. We refocused on answering a few more Who, Where, and How questions to explore data capture, protection, and storage needs. During this phase, we (the EA team members) also established a strong relationship with the individual project managers and the enterprise program management office (PMO), working together with them to build and/or revise the business cases to move forward with the selected initiatives and secure potential funding.

4. **Institutionalize.** This phase has been vital for formalizing risk management, governance, and compliance policies and processes across the enterprise. We shifted our focus to establishing review processes and procedures to manage and monitor adherence to EA principles while complying with government mandates and industry regulations related to the usage of wearable medical devices. We also faced challenges due to the lack of standards in the use of the IoT devices and, more specifically, stringent clinical restrictions associated with wearables. We refocused on Who, How, and When questions to establish the governance of the EA standards, principles, and our corporate governance policies.

Figure 1 presents the four phases in sequence. However, in reality, we have mostly worked on two or more phases in parallel to expedite the involvement of EA as a discipline in the enterprise-level digital transformation

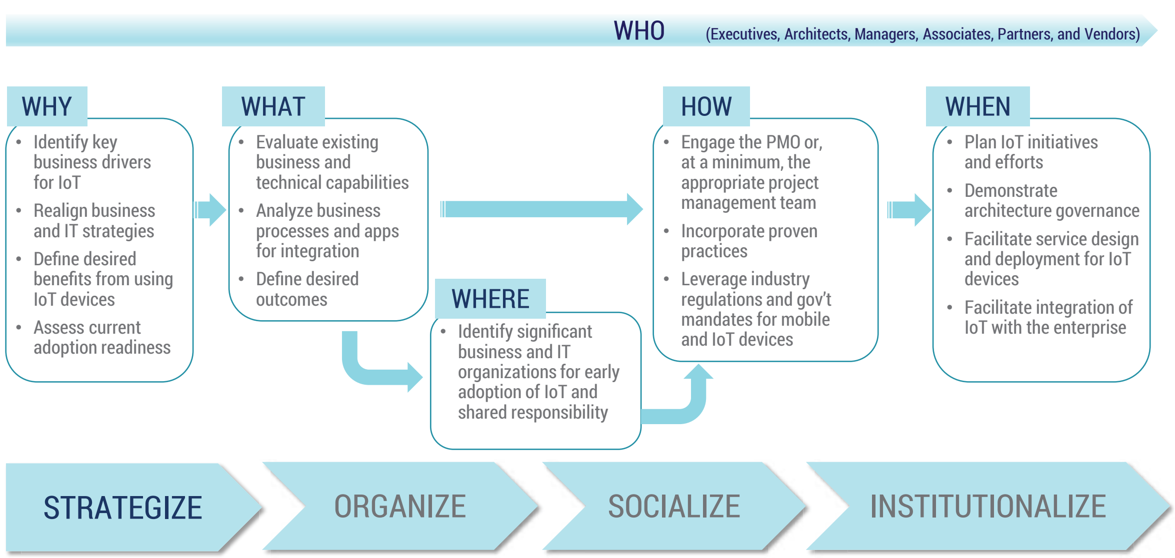


Figure 1 — A holistic approach to driving enterprise convergence.

Who	• Who are our target customers? Internal and/or external? • Who are the key decision makers? Business sponsors? IT leaders? • Who are our primary digital asset users? Data scientists? Research scientists? • Who are our digital technology or business partners and vendors? Software and hardware? • Who are our team members? Business and IT? Application development? Operations support? Infrastructure? Security and access control? • Who has prior digital transformation experience? • Who is responsible for interacting with the IoT vendors? For managing contracts and associated relationships?
Why	• Why do we need digital transformation? • Why do we need the IoT wearable devices? • Why do we need data analytics? • Why do we need big data? • Why do we need mobile applications? • Why do we need to consider new standards, new technologies, and/or a new development approach? • Why do we need to consider new integration technologies or network connectivity for the IoT devices?
What	• What are our key new business and technology requirements? • What are our existing business and IT capabilities for digital transformation? • What are our existing digital assets? What new ones would be beneficial? • What are our existing business processes? What new ones should we implement? What needs to change? • What other ongoing projects may be impacted (or interfaced)? • What are our reusable technology standards? • What impacts do new technologies and work products have on our enterprise, financially and operationally? • What savings do these technologies and work products hold for the CDS and HI units and subsequently for the entire enterprise? • What types of agreements/contracts do we need to have in place with telecom and/or IoT device manufacturers?
Where	• Where do we make the most business impact? Organizationally? Financially? • Where do we have most of our resources to start digital transformation? • Where we do we have the most architecturally significant components? • Where do we have most of the data needs? • Where do we currently store our PHI and PII data? Where do we store and archive our data in the future?
How	• How do we manage the initial digital transformation project(s)? • How do we ingest the data from the mobile devices and the sensors? • How do we engage our IoT business and technology partners? • How do we store the ingested data? How do we separate our PHI and PII data from the rest of the ingested data? • How do we incorporate our existing EA best practices? How about new ones? • How do we perform governance activities (e.g., architecture review)? • How do we incorporate industry regulations, standards, and government mandates for IoT or mobile device usage? • How do we manage or maintain interoperability issues? • How do we manage availability, scalability, and security issues related to the IoT devices? • How do we manage or monitor the power consumption of the IoT devices/sensors? • How do we integrate the IoT implementation with the entire enterprise?
When	• When are we ready to start the digital transformation? The IoT implementation? • When do we start to design, develop, and deploy the service for the IoT devices? • When do we start user acceptance testing of the IoT implementation?

Table 1 — Sample questions for digital transformation (including the IoT implementation).

initiative. We also empowered the project teams to leverage time and resources efficiently and consistently in order to make measurable progress with the digital transformation.

WHAT WE LEARNED: A SET OF EFFECTIVE PRACTICES

While we have had our own share of struggles in getting the organization ready, we have had a successful journey so far in establishing a collaborative, connected, converged, and socially responsible enterprise. The EA team has facilitated many focus group sessions with new and existing patients, clinicians, research scientists, and physicians and has collected various medical device–related requirements. We have also presented our findings from social meeting sessions (group discussions) to various care delivery organizations and the Health-USA leadership team. As part of this initiative, the EA team is still successfully utilizing the core guiding principles of EA, basic tenets of IoT strategies and blueprints, and a set of innovative digital health (including medical technology) concepts to build an effective digital enterprise, e-Health-USA. Here are a few of the foundational practices that made our digital transformation effective:

- **Recognizing and managing the expectations of the leadership.** Although this may sound like a common practice for any enterprise-level project, it plays an important part in digital transformation. It is important to keep the leadership well informed and up-to-date on the challenges, risks, issues, and impacts related to the ongoing projects while dealing with emerging technologies such as cloud, big data analytics, and IoT. Each of these technology trends requires the leadership to set a clear vision. However, the EA team must subsequently establish the right expectations and obtain leadership buy-in for alternative options as appropriate. For Health-USA, the leadership of the CIO was the most important impetus for the disruptive digital transformation journey.

- **Recognizing organizational readiness and the learning curve.** It has been an important practice to recognize the capabilities of the enterprise in embracing the IoT implementation. We have instilled a cultural change in the mindset of the team involved through formal training, workshops, seminars, technology briefings, and vendor demos. We have encouraged and empowered the project team members to take ownership of new techniques, approaches, or ideas that can expedite the project schedule and success. That said, the company had to make a conscious initial investment to help the team scale the learning curve associated with leveraging cloud computing, big data, and data analytics. For Health-USA, this initial investment demonstrated the leadership's commitment, which motivated the team to perform well in all phases of the initiative.

- **Adapting to the lack of IoT industry standards in formulating interoperability.** This has been a key practice for us. While many healthcare IT industry leaders have already embraced IoT, there is still a lot of work needed to establish standards for providing accessibility, interoperability, and protection for the data retrieved from IoT devices. We had to shift our focus onto the hardware aspects of the wearable devices in order to manage the aforementioned architectural elements. For example, we have worked closely with the IoT sensor manufacturer to encrypt the data captured by the sensor before allowing it to be received by a mobile device.

- **Addressing the privacy, security, and regulatory compliance issues.** This has been another major practice for us, and it requires attention throughout the digital transformation phases when leveraging cloud, big data analytics, and IoT. In most cases, it also involves securing the data while it is moving across multiple devices and the cloud or simply being streamed from a device to the enterprise system. We have established a federated or distributed approach to monitoring, maintaining, and managing the security and privacy aspects of the data while incorporating regulatory compliance factors under the EA governance principles. At the same time, the federated approach helped us connect our "silo" organizations and establish a mindset for shared responsibility across the enterprise.

- **Adopting emerging technologies with an open mindset.** This practice has been important for us in developing an adaptive EA and at the same time establishing a flexible culture across the enterprise. Leadership support was tremendously valuable in enabling us to evaluate emerging technologies and select technology options carefully and effectively. We have also had the opportunity to assess multiple open source–based products and tools while preparing multiple teams for the digital transformation. Our EA team — with the help of expert consultants, formal training, appropriate prototyping, and vendor demos — learned various ways of preventing PII and PHI data loss and leveraging real-time data analytics to minimize the risks of using IoT devices and big data.

I strongly believe that these practices can make any EA team effective in an IoT implementation while facilitating digital transformation. These practices have allowed us to integrate emerging technologies such as cloud and big data analytics with our enterprise-level initiative.

CONCLUSION

The initiative presented in this article is still a work in progress. However, I would like to offer four suggestions based on my experience in this digital transformation initiative and other real-world engagements:

- **Create value.** It is important to solidify the creation of value by leveraging EA practices for the IoT implementation early in digital transformation. I suggest that the professionals involved clearly define a set of desired business value(s) during the strategize phase, and EA professionals must help to obtain leadership buy-in for the created value proposition (i.e., the ROI existing and new digital assets will provide).
- **Develop value.** It is essential for EA as a discipline to take the lead in developing the value IoT, cloud, and big data will yield in the digital transformation initiative. In practice, value development requires collaboration between various teams. The EA team worked closely with the solution architecture, application development, and project management teams to develop various pilots, prototypes, and proofs of concept to demonstrate the value of these technologies to the enterprise.
- **Deploy value.** It is vital to deploy the value developed early to achieve the trust and confidence of the entire organization. We used the pilots and prototypes to demonstrate the value developed. We also used these pilots to train new team members. This value deployment becomes even more crucial due to the fact that most associates, employees, and executives do not have expertise or experience in dealing with the complexities of using the IoT along with cloud computing and big data analytics. In our case, we used the socialize phase to extend the value deployment across the entire enterprise while holding more review, inspection, and walkthrough sessions for the work products. These sessions allowed us to evaluate and test the work products using different use cases and what-if scenarios.
- **Realize value.** This may be the most difficult yet most rewarding part of the digital transformation journey. We have achieved a small but significant success in integrating the wearable IoT devices into the enterprise and have already managed to collect the data ingested by the IoT devices and support a team of data scientists to analyze it. We have also established a secure private cloud to store and protect a small segment of PII and PHI data. Most importantly, we have managed to leverage a set of industry-recognized standards for preventing data loss from the IoT devices. We have leveraged the institutionalize phase to achieve our goals here. The EA team is currently working with all four business units of Health-USA to roll out services for 1,000 wearable devices as a pilot project for Type 2 diabetes patients in a selected region of the US. The data captured from the wearable devices will be processed, protected, and used in five major US diabetes research and care delivery organizations. It will also help the Health-USA data scientists and research scientists study the patterns of a patient's self-management (using the pilot wearable devices) and subsequently offer relevant services to a wider range of the patient population.

I hope that the experiences I've recounted will resonate with *Cutter IT Journal* readers and that the suggestions and recommendations offered here will help you leverage EA/IoT synergy in your own digital transformation initiatives.

ENDNOTE

[1]Hazra, Tushar. "Establishing the Business Architecture Practice: A Case Study." Cutter Consortium Business & Enterprise Architecture *Executive Report*, Vol. 15, No. 1, 2012 (www.cutter.com/article/establishing-business-architecture-practice-case-study-389176).

Tushar Hazra is an Executive Enterprise Architect with over 22 years of experience in various areas of architecture development, implementation, governance, risk management, and compliance. Dr. Hazra is a recognized thought leader and expert in delivering enterprise-level business solutions, strategy, blueprints, and roadmaps; strategic planning; and implementing effective enterprise architecture for digital transformation. He has efficiently leveraged emerging technologies such as social media, cloud computing, big data analytics, and mobile computing in large and complex healthcare, finance, and insurance services initiatives. He has a proven track record of aligning IT with business goals through planning, prioritization, and implementation.

Dr. Hazra has been actively involved in designing, developing, and delivering mission-critical, innovative, and cost-effective IT solutions across multiple US federal, state, and local government agencies as well as in private healthcare (payers and providers), financial services, and insurance businesses. He has demonstrated C-level executive partnership, technical thought leadership, program management, P&L responsibility, and system engineering domain expertise. Dr. Hazra is an industry-recognized communicator, author, and speaker with 150+ publications and 100+ conference appearances. He can be reached at tkhazra@epitomione.com.

A Common Thread: Applying Hex Elementalization in IoT Data Analytics

by Girish Nair and Yi-Chen Lan

From the Stone Age to the Renaissance, from the Industrial Revolution to the Digital Age, there have always been upheavals and paradigm shifts in the way we humans have conducted ourselves in any field or phase of society. These revolutions helped us break through the boundaries we set up from our prior set of knowledge. The Internet of Things (IoT) is the latest focus of society's participants, consumers and producers alike. What is exciting is the vast array of IoT applications. IoT can be extended to the machine-to-machine (M2M), transportation, healthcare, nanorobotics, and utility domains. It can be used for vertical and horizontal IoT solutions. This disruptive technology has major implications for e-commerce and the industrial Internet.[1] IoT development will also accelerate the adoption of big data. In fact, the big data flow from IoT will explode to the extent that managing it smartly will be paramount.

IoT at its core is all about integration and marriage of "things," which can encompass anything from objects and machines to software and human beings. The question is how IoT will exist among a myriad of data types, information exchange systems, applications, and hardware and communication systems. Integration is not just confined to different kinds of devices and their mediums, but must also be undertaken as a business strategy.[2] The Digital Age seems to challenge long-held economic principles; abundance of data, not scarcity, will become the norm.[3] As IoT becomes ubiquitous on a lot of things (LoT), it will entail A Lot of Un-Interpreted Data (ALOUD). The challenge is to ensure that IoT is interconnected in an increasingly meaningful way. However, with so many companies working on so many devices, which will communicate in so many different ways and via so many different protocols, how are we going to find a common thread among them?

"Hex elementalization" is a concept conceived to deal with this. Hex elementalization involves breaking a source of information, irrespective of its form, into "hex elements" with six pieces of information and then integrating them. The source of information here can be a set of unstructured data, structured data, microwave or digital signals, sensor data, machine language, a lingua used by robots to communicate with each other, or any other future way of communication. The intention is to break the information sources down into the simplest form while embedding within them enough attributes to help them automatically seek and match themselves with information seen in other unrelated pieces of hex elements. Consider, for example, the hundreds of x-ray photos, CT scans, ultrasound scans, blood reports, and other pieces of data sitting in cupboards and digital storage in a hospital. A holistic, architecture-based approach can enable them to find meaningful information. An architecture into which you can throw anything and get unexpected and interesting relationships and information snippets is what we are theorizing in hex elementalization.

HEX ELEMENTALIZATION EXPLAINED

Hex elementalization aims to create a platform — an architecture, an environment — encompassing end-to-end integration. The traditional way of thinking about data in 1s and 0s (bits) is unsuitable if we want to create a common "playground" for IoT. No matter how large or diverse the data is, it needs to be broken down into smaller chunks that will enable and ease interaction. These smaller chunks (i.e., hex elements) can be combined and related into meaningful information.

Why "hex" elements? Because the hexagon as a shape can be found in the most basic of structures.[4] It is found widely in nature — for example, honeycombs, ice crystals, and even the recently discovered graphene.[5] In other words, the hexagon is a shape that occurs naturally and is trusted by nature. Nature finds this structure helpful, effective, and efficient. Figure 1 shows how data from disparate sources (IoT devices) can be broken down into six-sided hex elements.

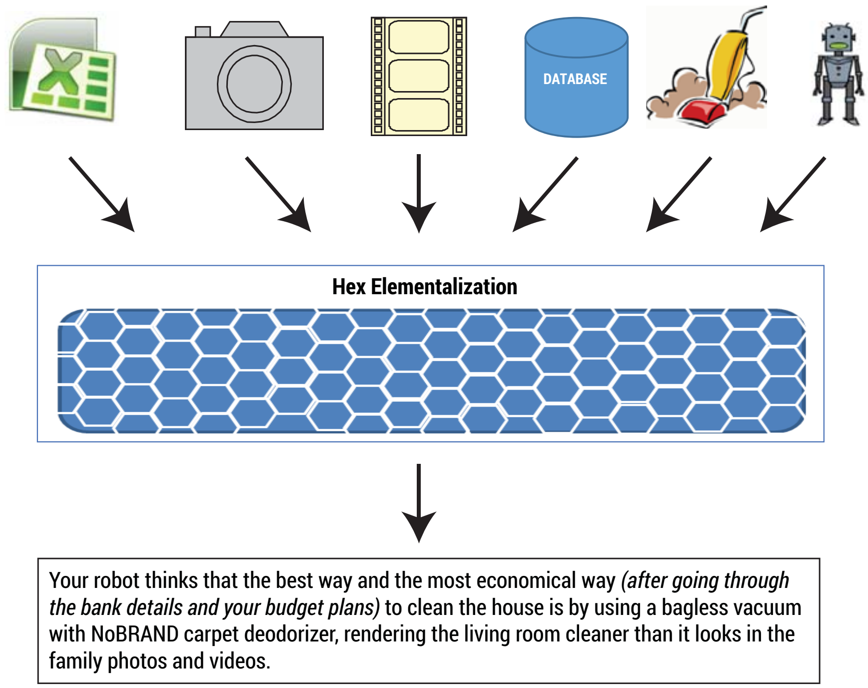


Figure 1 — The hex-elementalization process.

HOW HEX ELEMENTALIZATION WILL WORK

The dataflow from a given IoT device gets broken down into hex elements with a set of six properties (see Figure 2). There is no limit to how many IoT devices can be integrated via hex elementalization. Each set of hex elements from each IoT device tries to interconnect by seeking common properties (see Figure 3). This will enrich each flow of information as it gathers more data from each new stream of hex elements emanating from each IoT device. Each flow will gather the related information it seeks to create a new informational flow (see Figure 4). For example, hex elements from a smartwatch can get interconnected with the hex elements from your car and smart meters on the street, providing you with a unified snippet of information: "Your restaurant is a five-minute (250m) walk from this parking spot, and you have 45 minutes before the time-based parking restrictions will come into force." Thus, the information stream or "hex stream" will be made up of a chain of hex elements from disparate sources that are grouped and unified by a common property to create a chain of rich and interconnected information. The process is analogous to a creek getting bigger and turning into a stream and then into a river as it merges with smaller creeks and amasses more waterflow.

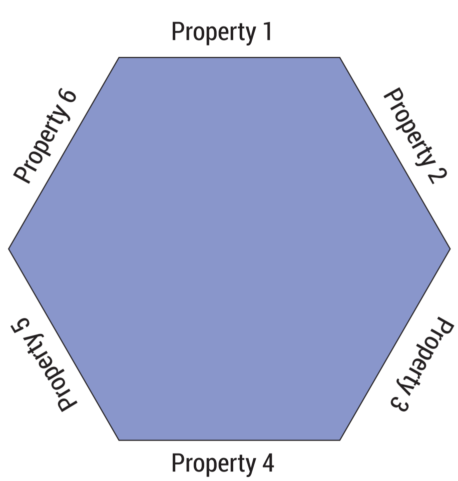


Figure 2 — A hex element and its six properties.

HEX ELEMENTALIZATION IN VARIOUS DOMAINS

IoT and Big Data

With the adoption of IoT, the propensity to gather data will increase exponentially.[6] But for IoT to become synonymous with pervasive or ubiquitous computing, its real potential lies in real-time data analytics. Business intelligence, business planning, product planning, and product design are some of the activities that can benefit from real-time data analytics. For example, real-time smartwatch usage information can help both product designers improve their product and software developers fine-tune user interaction. IoT, with its ability to drag a big trove of data (big data) with it, has the

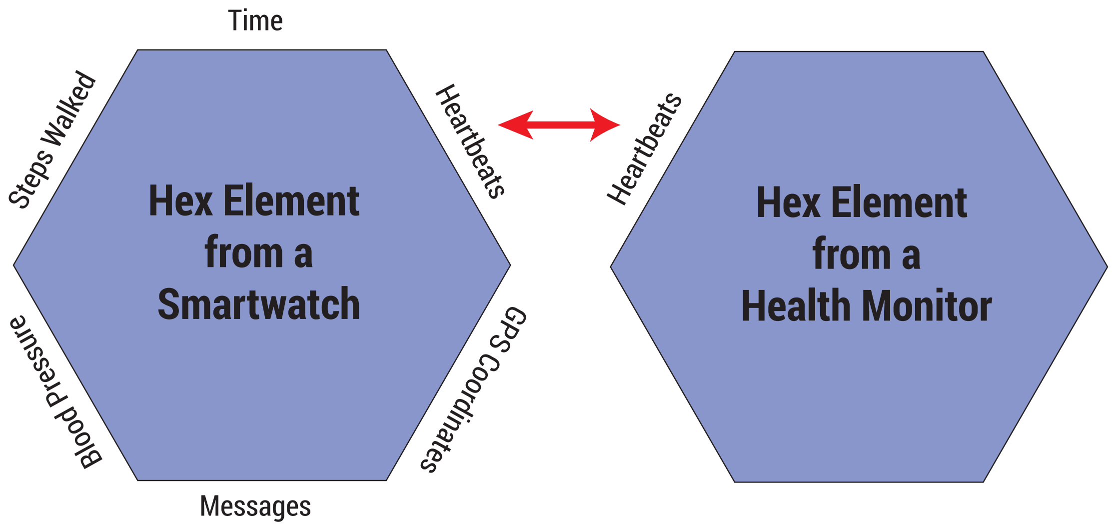


Figure 3 — An example hex element from a smartwatch and how it finds a common thread with a hex element from a health monitor.

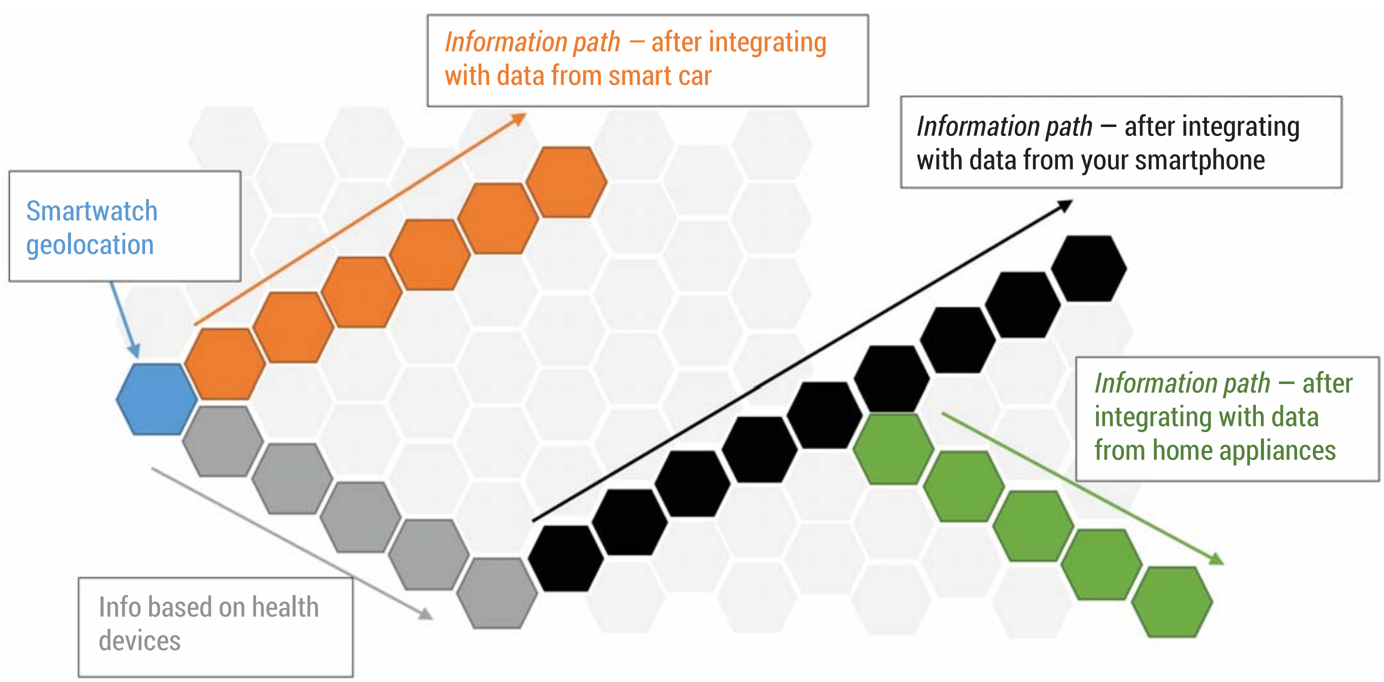


Figure 4 — Decisional path of hex elements extrapolated from Figure 1.

potential to improve the productivity of various industries, including healthcare, transportation, energy, and the obvious IT industry.

There are three challenges facing IoT and big data going forward:

1. Big data precedes IoT in form and adoption, but with IoT becoming mainstream, it will accelerate the adoption of big data, more in the form of cloud-based and real-time big data processing solutions.[7]
2. Hardware is not dumb anymore, thanks to IoT and the increasing push from companies to cram as much processing into the tiny chip embedded in IoT devices. But how much processing power will companies continue to push onto devices, sacrificing the ergonomics and aesthetics of their devices (e.g., thicker phones)?
3. IoT will help accelerate the adoption of business intelligence and analytics on a real-time basis to make real-time decisions, but the question is how? The traditional business intelligence approach is time-consuming and modular and linear in nature.[8] This makes it powerful from a management and oversight perspective, but not from an Agile approach.

IoT will help accelerate the adoption of business intelligence and analytics on a real-time basis to make real-time decisions, but the question is how?

Currently IoT is designed for linear flow of information (after processing big data) from an IoT device to the user and vice versa. However, IoT should encompass the availability of intelligence and data analytics across the entire value chain of the product in use. An example is the soon-to-arrive autonomous car, which can send real-time information to tire manufacturers on what is causing damage to their tires, to engine control unit (ECU) producers on what the optimal processing would be, to traffic departments on how numerous cars are avoiding certain roads, to municipalities on how certain streets are becoming parking lots, to smartphone manufacturers on which apps the user had problems with when using them inside the car, and many more. Thus, IoT and its subsequent marriage with big data entails sharing intelligence not only between the provider and the receiver, but across the entire IoT value chain.

Hex elementalization breaks down the barriers between the three stages of data analytics: data collection (big data), data analysis, and response. Once data from disparate IoT devices is broken into hex elements, the hex-elementalization process will find information paths (as shown in Figure 4). These information paths can provide snippets of unified information to the IoT user or real-time insights to a business that has incorporated IoT at various stages of its workflow. By automatically finding relationships based on the common properties between hex elements, hex elementalization aims to make big data into *quick* data. Marketing and sales campaigns, product designs, and customer response are more effective based on real-time metrics. In the fast-moving consumer goods (FMCG) sector, tracking fleet and transportation/distribution and inventory levels on a real-time basis will help FMCG companies respond to adverse macroeconomic conditions without any lags.

Vertical and Horizontal IoT Solutions

The proliferation of IoT products has forced current IoT producers to adopt either a vertical or horizontal business model. Vertical business models are constrained within the same product line and can operate vertically within the product's value chain.[9] The issue is that within this widely adopted model, users can demand different systems, which are designed to achieve predefined tasks with their own array of support services. Horizontal models, on the other hand, work across different value chains and across different product lines.[10] The goal here is to nurture innovation in the IoT industry by allowing multiple players to share, contribute, and adopt open functionality. The objective should be to make IoT devices and data exchange more easily available and distributed across vertical and horizontal business models. A hybrid business model will be ideal where a complex information system enabling free and easy sharing of data between IoT devices is facilitated.

Hex elementalization can offer this framework to enable the creation of hybrid IoT solutions, breaking down the barriers inherent in vertical and horizontal IoT solutions. For example, a technology company can utilize a hex-elementalization framework to enhance its software services and products; it can also learn about the user's interactions with the software in order to give input to its hardware division so it can improve its product design. For instance, Microsoft could use its knowledge of how its users interact with MS Office applications to better design Microsoft hardware (e.g., Surface, keyboards, mice, gaming controllers), or it could pass on the learnings from its hardware users to make its software more intuitive and easy to use — all on a real-time basis.

Consumer Services

The impact of IoT will be greatest on the consumer. From health-tracking bands, to tablets, to smart glasses, to self-driving cars, it's hard to imagine any aspect of a daily life that will not be impacted by the advances made by IoT. As more IoT devices enter our lives, they will broadcast a vast array of usage data. This usage data might emanate from a single device, a multitude of devices, or a multitude of ways of communicating with the same device, forcing us to act on individual chains of information from each IoT channel. For example, we interact with our phones through motion, voice, touch, and position to applications or real-world people. In an ideal world, we want an IoT ecosystem in which all the disparately built, differently programmed, and varied interactions across the IoT devices can communicate on the same level playing field.

Hex elementalization aims to facilitate this communication by laying down a carpet of ubiquitous communication channels. An ecosystem built on hex elementalization will not only let different devices (from different producers) communicate with each other, but will help in ubiquitous and real-time processing of data to make our day-to-day life (and decision making) easier. We need IoT to expose the data and information through the IoT devices, and we should get real-time help in our everyday activities through a combined and concentrated effort across all our IoT devices.

Hex elementalization will find the best-suited paths and decision trees while processing through the relationship set between hex elements. Figure 4 shows how the information processing and information flow will occur after getting the inputs from various devices linked to the hex elementalization framework. Based on the unique but related characteristics, hex elementalization will provide a platform for self-sustained and directional decision/information branches to provide real-time analytics to the end user. Having ubiquitous processing and computations is pivotal for the success of IoT. Finding correlations in real time will enable IoT to take center stage in the next data revolution.

Robotics and Automation

According to a study done by Oxford University researchers, 45% of American manufacturing will be done by robots in the next two decades.[11] Setting aside the philosophical side of the issue, robots will take IoT and its spread to a higher level. As robots enter our daily work and personal lives, their integration becomes important. They will need to have a common ground for communication as adoption of IoT and big data becomes mainstream. This will help businesses in making both strategic decisions and tactical decisions.[12] As IoT augments the adoption of big data, AI will augment the use of robots. AI, in other words, will be responsible for how robots and automation are managed and introduced in our lives.

In a scene from the movie *Transformers*, we see big robots emanating from small metal balls. The scene shows how, in the future, smaller, AI-driven, self-sustaining granular and intelligent robots (nanorobots) will come together to create a larger and unified intelligent robot (or an object). Hex elementalization can lend its basic principles to this physical "transformation" function, as well as laying down the principles of most tight-fit integration among the future world of nanorobots. It is difficult for small metal balls (circles in 2D) to combine to create a larger structure, but it is easier for six-sided hex elements to do so (see Figure 5).

From assisting this physical transformation in robotics to driving the development of AI, hex elementalization concepts can be extended in a multitude of directions. The basics of AI are about machine learning — the ability to find and learn patterns for decision making and intelligence gathering. Hex elementalization can help in:

- The establishment of integration and data interchange protocols between robots
- The AI driving these robots
- The integration of IoT data and devices to assist with IoT analytics

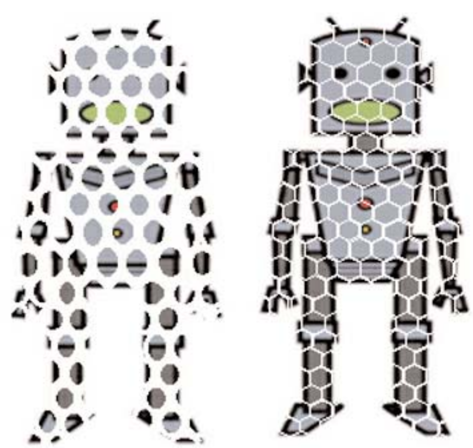


Figure 5 — A robot made of circles vs. a robot made of hexagons.

For example, if a smart dirt sensor learns from your smartwatch that you are allergic to the dust it has detected, it can instruct your robotic vacuum cleaner to clean your floors by the time you get home — which it determines after getting your estimated arrival time from your smart car.

HEX ELEMENTALIZATION IN PRACTICE VIS-À-VIS CONTEXT

Businesses need to navigate through the noise when it comes to finance[13] or business intelligence, but they also need to absorb and navigate through the noise produced by IoT and its increased use of big data. Big data has enabled companies to accumulate and accommodate large troves of data in sophisticated repositories. This data must be converted into information and from information to *intelligent contextual information* to be of value to business. Content itself is not sufficient — what is needed is the contextual underpinnings of big data associated with IoT. In the example above, the context provided around the information flow is important to its success. You don't want the dirt sensor to provide you with information on the amount of allergen-containing dust on your floor just to satisfy your idle curiosity; you want it to use that information to solve the problem before you get home. For IoT to be meaningful, the big data emanating from these devices needs to have contextual boundaries to provide smart info. For example, if the context is "wellness," then content in terms of metadata (which can be "amount of dust [in cc]," "allergen levels per cc of dust," "allergen levels triggering an allergic response in homeowner," etc.) can be derived and passed on to create boundaries for this wellness-related IoT application.

Hex elementalization can provide this context through semantics embedded in the hex elements within the same value chain (i.e., hex elements derived from the same vertical chain of IoT products). This can be achieved through the use of AI overlaid on hex elementalization, which can help limit flow of information and decision-enabling data through contextualization. For example, a phone manufacturer does not just need to know how much sales fell after a new product launch, but why they fell and what macro- or microenvironmental factors affected its customers' buying patterns. With this information in hand, the company might decide to stock its retail shelves not with an equal number of each model, but to tailor the stock according to the age and other demographical attributes of each store's location (e.g., targeting feature-filled phones to younger markets vs. basic, functional phones to an elder community).

CHALLENGES IN IMPLEMENTING HEX ELEMENTALIZATION IN IOT EFFORTS

New Form of Unstructured Data

Big data and the IoT ecosystem can collate, store, sift, aggregate, visualize, and analyze complex data sets in relational, structured, unstructured, sequential, and geospatial form. Hex elementalization as an architecture would need to create a platform to understand, decode, disintegrate, relate, encode, process, compute, and create directional decisional paths before reaching the end user. Text analytics, which is how consumer interactions are captured (via text, phone calls, emails, call center longs, globs, surveys, and social networks), is difficult to perform due to the element of language.[14] The same challenges apply to continuously evolving unstructured data sets (e.g., Vines or Live Photos).

Hex elementalization will need to create and simplify the complex structure of this new kind of data to harness its untapped power. As we described above, the process of hex elementalization starts with creating meaningful and identifiable hex elements, each of which will hold six distinct properties. The challenge is to break unstructured data (like voice) into identifiable and self-sustained granular elements. The processing of unstructured data and the data exchanges, when compared to structured data sets, are less precise or complete.[15] The hex elementalization process will be augmented once AI enters the realms of text and voice expression to help us understand the complex set of data around us in simple terms (think Jarvis in the *Ironman* movies).

IoT Node Interface and Communication Protocols and Channels

The uptake of IoT, which is evolving all the time, is going to explode in the coming years across every conceivable part of our society. Although hex elementalization can provide a platform to ease the communication, processing, and data interchange between IoT devices, the question will be whether each IoT data pipe will need an individual gateway that will have to be intelligent enough to break the outgoing data into hex elements. Given that current IoT devices have limited processing done at the host, this requirement will demand that IoT hosts become smarter and faster.

Heterogeneity of IoT

Heterogeneity of existing IoT operating systems or ecosystems will complicate the hex elementalization framework. Ideally, each IoT stream will entail a

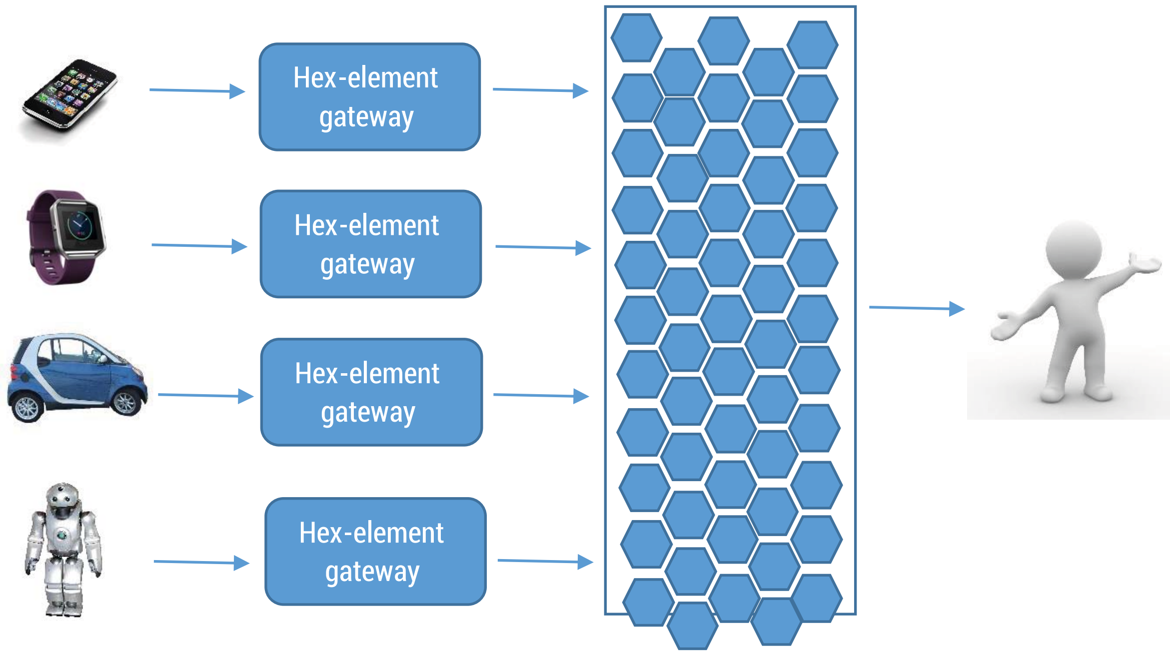


Figure 6 — Information gateway and hex elementalization.

process in which it will decrypt its own data set before these elements enter the hex-elementalization phase (see Figure 6), where correlations and decision trees start formulating. Each gateway will have to adhere to certain common goals and protocols for hex elementalization to work seamlessly.

Pull or Push

Another protocol that needs to be discussed at length and decided upon is the question of whether IoT will push or pull the data for intelligent processing. Going back to the example of autonomous cars, the smart meters operated by city governments need to push data out of their domain so that IoT devices (and hex elements) can access this data to find and build relationships. The autonomous car will only be really useful if it makes decisions autonomously across various functions (parking, distance to the destination, traffic diversion, etc.), not just when it comes to driving. The most desirable aspect of the hex-elementalization framework is having real-time intelligence based on easily accessible and related data sets. The protocols built on pull data over a security layer would push the need for smarter gateways for each IoT usage.

CONCLUSION

Connecting Wi-Fi chips and sensors to dumb hardware does not make it smart. Even placing a Wi-Fi chip and a sensor on a toaster to monitor its heat level and then notify us when the toasting is done does not make the dumb toaster smart. What *will* make it smart is its ability to perform functions beyond those for which it was primarily designed, such as informing us (by updating our grocery list on our smartphone) that we need to buy more bread (as we just ran out of it) and also recommending that we get some butter. Hex elementalization thus can provide a framework, an architecture, on which IoT can grow into an intelligent system.

ENDNOTES

[1]Evans, Peter C., and Marco Annunziata. "Industrial Internet: Pushing the Boundaries of Minds and Machines." General Electric, 26 November 2012 (www.ge.com/docs/chapters/Industrial_Internet.pdf).

[2]Lan, Yi-Chen, and Bhuvan Unhelkar. *Global Enterprise Transition: Managing the Process*. IGI Global, 2005.

[3]Shapiro, Carl, and Hal R. Varian. *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business School Press, 1998.

[4]Arthur, W. Brian. *The Nature of Technology: What It Is and How It Evolves*. Free Press, 2009.

[5]Allen, Matthew J., Vincent C. Tung, and Richard B. Kaner. "Honeycomb Carbon: A Review of Graphene." *Chemical Reviews*, Vol. 110, No. 1, 2009 (http://pubs.acs.org/doi/abs/10.1021/cr900070d?journalCode=chreay).

[6]Chen, Min, et al. *Big Data: Related Technologies, Challenges and Future Prospects*. Springer, 2014.

[7]Mayer-Schonberger, Viktor, and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. John Murray, 2013.

[8]Moss, Larissa T., and Shaku Atre. *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*. Addison-Wesley Professional, 2003.

[9]Bassi, Alessandro, et al. (eds.) *Enabling Things to Talk: Designing IoT Solutions with the IoT Architectural Reference Model*. Springer, 2013.

[10]Severi, Stefano, et al. "M2M Technologies: Enablers for a Pervasive Internet of Things." *Proceedings of the 2014 European Conference on Networks and Communications (EuCNC)*. IEEE, 2014 (http://goo.gl/B3VUbK).

[11]Rutkin, Aviva Hope. "Report Suggests Nearly Half of US Jobs Are Vulnerable to Computerization." *MIT Technology Review*, 12 September 2013 (www.technologyreview.com/s/519241/report-suggests-nearly-half-of-us-jobs-are-vulnerable-to-computerization).

[12]Unhelkar, Bhuvan. "Collaborative Business and Enterprise Agility." Cutter Consortium Business Technology & Digital Transformation Strategies *Executive Report*, Vol. 13, No. 9, 2010 (www.cutter.com/article/collaborative-business-and-enterprise-agility-379846).

[13]Bernstein, Richard. *Navigate the Noise: Investing in the New Age of Media and Hype*. John Wiley & Sons, 2005.

[14]Dey, Ian. *Qualitative Data Analysis: A User-Friendly Guide for Social Scientists*. Routledge, 2003.

[15]Feldman, Ronen, and James Sanger. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press, 2006.

Girish Nair is a member of the global quantitative strategy group within the research department of a major US investment bank. He is involved in in-depth analysis and research using quantitative methods (mathematical, statistical, and computer-aided models) to test efficacy of various fundamental, technical, and economic factors affecting equity markets. His team's work is published in professional research reports distributed to various fund managers globally. Mr. Nair has an MBA from the University of Technology, Sydney. He can be reached at girish124@gmail.com.

Yi-Chen Lan is the Deputy Pro Vice-Chancellor (International) at Western Sydney University. Dr. Lan holds a bachelor of commerce (honours) degree and a PhD from the University of Western Sydney. He is responsible for the implementation of the university's international plan; the development, management, and facilitation of all international relationships across the university relating to transnational programs; the student mobility program and international research partnerships; recruitment of onshore international students; communications to all international stakeholders; and regulatory compliance for international operations. He can be reached at Y.Lan@westernsydney.edu.au.

About Cutter Consortium

Cutter Consortium is a truly unique IT advisory firm, comprising a group of more than 100 internationally recognized experts who have come together to offer content, consulting, and training to our clients. These experts are committed to delivering top-level, critical, and objective advice. They have done, and are doing, groundbreaking work in organizations worldwide, helping companies deal with issues in the core areas of software development and Agile project management, enterprise architecture, business technology trends and strategies, enterprise risk management, metrics, and sourcing.

Cutter offers a different value proposition than other IT research firms: We give you Access to the Experts. You get practitioners' points of view, derived from hands-on experience with the same critical issues you are facing, not the perspective of a desk-bound analyst who can only make predictions and observations on what's happening in the marketplace. With Cutter Consortium, you get the best practices and lessons learned from the world's leading experts, experts who are implementing these techniques at companies like yours right now.

Cutter's clients are able to tap into its expertise in a variety of formats, including content via online advisory services and journals, mentoring, workshops, training, and consulting. And by customizing our information products and training/consulting services, you get the solutions you need, while staying within your budget.

Cutter Consortium's philosophy is that there is no single right solution for all enterprises, or all departments within one enterprise, or even all projects within a department. Cutter believes that the complexity of the business technology issues confronting corporations today demands multiple detailed perspectives from which a company can view its opportunities and risks in order to make the right strategic and tactical decisions. The simplistic pronouncements other analyst firms make do not take into account the unique situation of each organization. This is another reason to present the several sides to each issue: to enable clients to determine the course of action that best fits their unique situation.

For more information, contact Cutter Consortium at +1 781 648 8700 or sales@cutter.com.

The Cutter Business Technology Council

The Cutter Business Technology Council was established by Cutter Consortium to help spot emerging trends in IT, digital technology, and the marketplace. Its members are IT specialists whose ideas have become important building blocks of today's wide-band, digitally connected, global economy. This brain trust includes:

- Rob Austin
- Ron Blitstein
- Tom DeMarco
- Lynne Ellyn
- Vince Kellen
- Tim Lister
- Lou Mazzucchelli
- Ken Orr
- Robert D. Scott