Machine Intelligence Through Data Transformation

After several years of hype about big data and analytics, it's truly amazing how misunderstood and underestimated data management tools are in the development of the Internet of Things (IoT) and Smart Systems market. Aggregation, transformation and management of data from sensors, machines and equipment is the holy grail of machine learning and the IoT; a fundamental core enabler.

smart systems design | Harbor Research

The fact that a wide range of sensors, machines and equipment can transmit information about status, performance and usage, and can interact with people and other devices anywhere in real time points to the increasingly complex role of data in IoT systems. This only compounds when we consider the many billions or more of networked devices that many observers are forecasting will be deployed and the scale of data they will produce.

Gathering and analyzing machine data is not a role for human beings. The only way to achieve it is to have the product's own "machine intelligence" continually delivered back to its creator. This requires three things:

» Giving machines something worthwhile to say and the ability to say it. "Something to say" comes from sensors and embedded computing in the machine and the ability to "speak" comes from giving the machine wired or wireless connectivity to the Internet.

» Putting the machine in an "atmosphere" where it can be heard, and most importantly, be understood. This requires transforming diverse machine data types, including sensor signals and values, messages, and machine logs, as well as data feeds and streams into a "normalized" technology-neutral information architecture.

» Listening to and interpreting what the machine has told you. This means integrating and analyzing the device data to yield actionable business information.

We can now see that the Internet of Things and big data "story" has been both a blessing and a curse. It has evolved to become the fundamental organizing schema for all intelligent machines and devices to share data and information, but the dilemma lies in the fact that data transformation solutions today can analyze some, but not all data, and mostly just simple structured data. The ability to normalize different data types and fuse diverse data sets requires new tools and technologies. Without advanced data management and transformation tools, users are only scratching the surface of the full value of their machine data.

TABLE OF CONTENTS

EXHIBITS

THE AGE OF SMART SYSTEMS AND SERVICES

Peer-to-peer information, social networking and embedded computing are combining to create new modes of collaboration and decision making. People, information, and technology are becoming more connected, distributed and pervasive enabling the convergence of the physical and virtual worlds.

These forces are informing a new trend we call "Smart Systems." In its simplest form, Smart Systems is a new generation of computing systems and information architecture, that when combined with artificial intelligence, machine learning and Internet of Things technologies are breaking away from today's information, computing and telecom (ICT) paradigms to enable the data from machines, sensors, video streams, maps, people, news feeds and more to become an integral part of all information systems.

Our society is at the cusp of a "perfect storm" of network connectivity. This phenomenon is not just about the dichotomy between people communicating with people or machines communicating with machines: it also includes people communicating with machines (e.g. a networked ATM), and machines communicating with people (e.g. automated stock ticker alerts on your smart phone).

This new paradigm is driving all information systems and, more importantly, their interactions towards real-time, state-based, context-driven capabilities that integrate people, processes, physical equipment and knowledge to enable collective awareness and better decision making -- thus, Smart Systems.

SMART SYSTEMS ARE REALLY ABOUT THE DATA .... DUMMY!

As networks continue to invade the "physical" world, traditionally unique components and interfaces between and among electronic as well as electro-mechanical elements are becoming more and more standardized.

The implications of these trends are enormous. No product development organization or its suppliers will be able to ignore these forces -- product and service design will increasingly be influenced by common components and sub-systems. Vertically defined, stand-alone products and application markets will increasingly become a part of a larger "horizontal" set of standards for hardware, software and communications.

As it becomes easier to design and develop smart systems, competitive differentiation will shift away from unique product features towards how the product is actually used, how the product fosters interactions between and among users in a networked context and, most importantly, how the data from the product will inform these

new insights. Even though we have been steadily designing devices and products with more and more intelligence, this information has gone largely unleveraged and unharvested. This may seem surprising to some observers but it has only been in the last few years that the world has come to understand the value of device data.

Machine data can offer extraordinary business advantages to the companies that manufacture, deliver and service machines, especially in terms of customer relationships. The ability to detect patterns from large scale sensor and machine data is the "holy grail" of smart systems. Machine data analytics, often thought of as part of the evolving "big data" story, allows not only data patterns but a much higher order of intelligence to emerge from large collections of ordinary machine and device data.

Smart Systems technologies are combining with new innovations in data and information architectures to work together in unprecedented ways to solve more complex business problems than previous generations of computing.

But, are they?

DON'T WE ALREADY HAVE BIG DATA?

Before delving into the new thinking that makes all this possible, let's talk about why it's necessary at all. Don't we already have big data and analytics tools? Aren't these tools helping us to manage and analyze all this data we keep hearing about?

Almost everyone will answer with a resounding "Yes!" But consider this analogy from Buckminster Fuller: Suppose you are traveling on an ocean liner that suddenly begins to sink. If you rip the lid off the grand piano in the ballroom, throw it overboard, and jump on it, the floating piano lid may well save your life. But if, under normal circumstances, you set about to design the best possible life preserver, are you going to come up with the lid of a grand piano?

Today's so-called data management tools are like that piano lid. In a period of great change and tumult, it's been working—in the sense that it kept us afloat. But that does not make it the best possible design, or qualify it to be something that we should plan to live with forever.

Today, IT professionals speak a great deal about "data management" tools that can be made available anywhere, anytime, for any kind of data and information. However, the tools we are working with today to manage and analyze data coming from intelligent sensors and machines were not designed to handle the diversity of device data types and the massive volume of datapoints generated from real-time machine and equipment interactions. These challenges are diluting the ability of

WHAT IS MACHINE DATA?

Machine data includes all data generated by equipment, devices and sensors, including:

» Computer network, and other equipment logs

» Temperature, pressure and related sensor readings from pipelines, factories and the environment

» Satellite and telemetry data

» Location data such as RFID readings and GPS system outputs

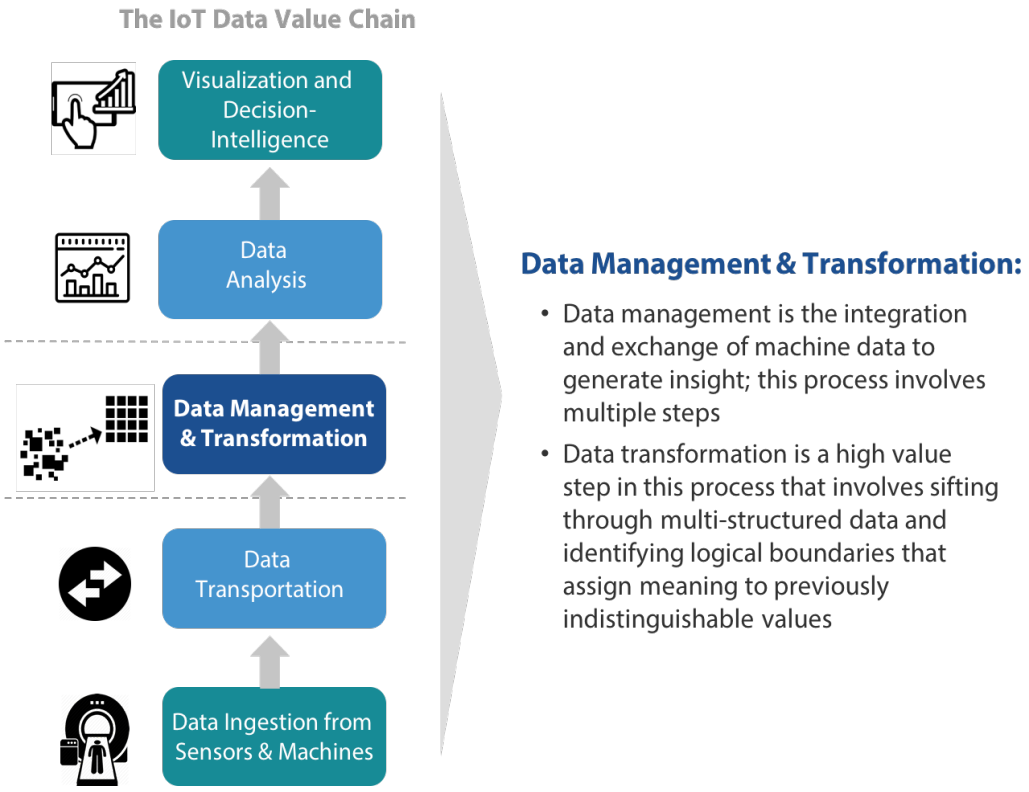
» Medical device readings for human health parameters

technical organizations to efficiently and effectively organize the data to model it and analyze it. The fragmented nature of software offerings available today to transform, model and analyze data make it extremely difficult, time consuming and costly to get results.

NOT ALL MACHINE DATA IS CREATED EQUAL OR IS EVEN THE SAME

The "Achilles Heel" of data tools for Smart Systems does not originate in its data collection or aggregation capabilities, or the analysis tools. Those inventions are not necessarily ideal, but they are useful enough today, and they can be replaced over time with better alternatives. Rather, the weakness lies with basic data management technologies—in particular, data transformation and modeling tools—and the restrictions they place upon organizing and utilizing device data to conduct analytics.

Exhibit 1: The Scope of Data Management and Transformation



Historically, computing systems have stored information in one of two basic ways: utterly unstructured, or completely structured. At the unstructured end of the spectrum are static Web pages, blog postings, emails, etc., which are free-form and

lack any fundamental identity. At the other end of the spectrum are very structured relational databases that are not at all flexible and make rigid assumptions about the meaning and context of the data they store.

Between these opposite extremes, intelligent machines on networks are now producing a vast array of semi-structured data types, including machine logs, data streams, sensor values, control signals and more. Sensor data and simple log data comprise the vast majority of data gathered from machines today. These simple data types comprise only a fraction of all potential data value and, on their own, cannot enable more advanced use cases, such as predictive maintenance on an MRI machine. Furthermore, this data cannot effectively be leveraged for advanced machine learning analytics technologies.

Many customers and users are surprised to learn there are multiple types of data produced by their machines and often underestimate the challenges involved in managing this data.

THE FUTURE OF MACHINE INTELLIGENCE

In the IoT arena today, most networked machine applications are limited to remote monitoring and maintenance services, including alerts, alarms, remote diagnostics as well as tracking and location services. This is due to several factors including technical complexities, business model challenges and a lack of significant embedded intelligence in machines. Existing technology has proven cumbersome and costly to apply with many conflicting protocols and incomplete component-based solutions. The challenges of gathering machine data and integrating diverse data types have been big adoption hurdles for customers wanting to analyze the data from machines and systems.

Return from simple applications, while extremely valuable, is limited to the manufacturer's service delivery efficiency. Contrary to what current market offerings depict, however, the value of connectivity does not have to end with simple applications focused on a single class of device. Moving from "Simple" to "Compound" applications involves multiple collaborating machines and systems with significant interactions between and among devices, systems and people. No longer is the focus solely on the machine builder's ability to deliver support for their product efficiently. Rather, value is brought to the customer through business process automation and machine optimization.

As technologies mature, particularly embedded computing and software tools, machines will continue to evolve to much higher levels of intelligence. As machines become more and more complex, so to will the challenge of extracting intelligence from the machine's data. Because more advanced intelligent machines produce

WHAT IS DATA MANAGEMENT and TRANSFORMATION?

Data transformation is a high value step in the data management process that involves sifting through multi-structured data and identifying logical boundaries that assign meaning to previously indistinguishable values. These processes can include the following:

DATA QUALITY PROCESSING
Cleansing
Filtering
De-duplication

SYNTAX TRANSFORMATION
Format normalization
Data normalization

SEMANTIC TRANSFORMATION
Semantics assignment
Context injection

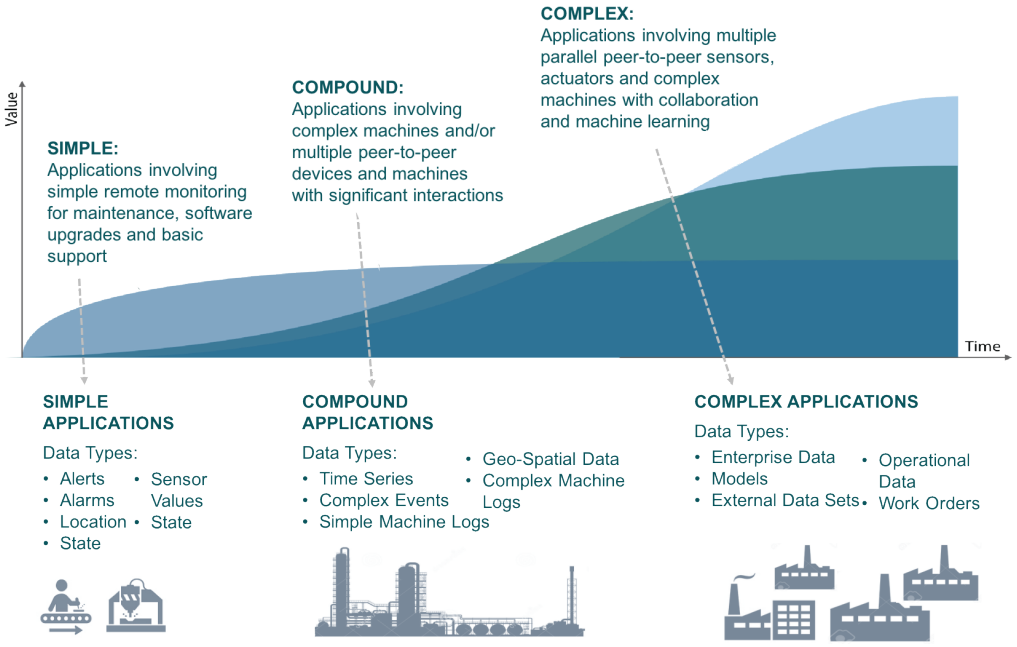
GOVERNANCE
Ownership assignment
Anonymization
Protection

PARSING
Syntax analysis
Data structure

EXTRACT, TRANSFORM, LOAD (ETL)
Data reading
Data converting
Data writing

a variety of more complex "machine logs" in a relatively predictable manner, it is an ideal "staging" area for designing, building and deploying a new generation of advanced data transformation, management and analytics tools.

Exhibit 2: Connectivity Produces Device Data Value Across Value Chains

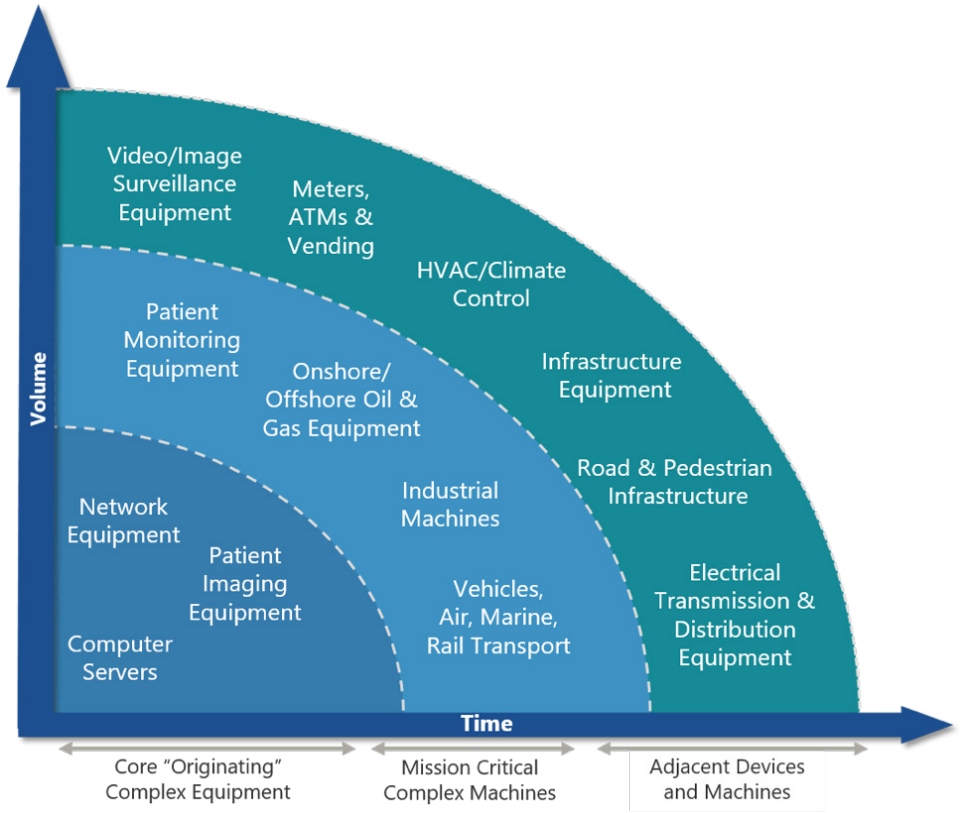


Analyzing why an asset has failed requires investigation of the patterns and hidden signs within machine log data. The bottom line is there's a huge difference between the world of asset monitoring, which is driven by sensor and simple log data, and the world of advanced analytics. This difference is dictated by data sources: sensor and simple log data can provide alerts that something has gone wrong, but only complex machine log data can be used to truly uncover and address the root cause of the failure. Furthermore, complex machine log data provides a much richer context than sensor data. For example, sensors cannot provide information about what applications within an asset's operating system are being underutilized, but machine log data can be used to understand these sorts of usage patterns and suggest user operational improvements.

Advanced forms of machine data will evolve beyond simple sensor and simple log data and will become far more robust. This opens up the opportunities for many diverse and valuable applications. These compound applications will involve more complex machines (such as medical imaging machines) as well as significant

interactions between and among many simple and complex machines and data sets (combining, for example, data from medical imaging, diagnostic monitoring and patient records) creating new collaborative business model opportunities that have the potential to drive much greater value for the customer.

Exhibit 3: Evolution of Complex Machine Intelligence



A major driver of the need for new data management tools is the diversity of data types users want to analyze. Because machine and sensor data cover a broad range of data types and structures - diverse formats, often analog, and high-velocity - there are major challenges that traditional data transformation tools and techniques do not handle well.

COMPLEX MACHINE DATA MANAGEMENT CHALLENGES ARE UNDERESTIMATED

Many factors contribute to the current simplistic state of applications, but one of the most underestimated and significant elements is data transformation. Unfortunately, this is a challenging, time consuming and costly step. Today, a growing number of

COMPOUND APPLICATIONS DATA and VALUE

A single large oil refinery produces more data in a day than all of the New York Stock Exchange and AMEX combined

In a 200 turbine wind farm, each turbine has 50 sensors with over 100 data points collected every 40 milliseconds, producing over 6,000 data points every second

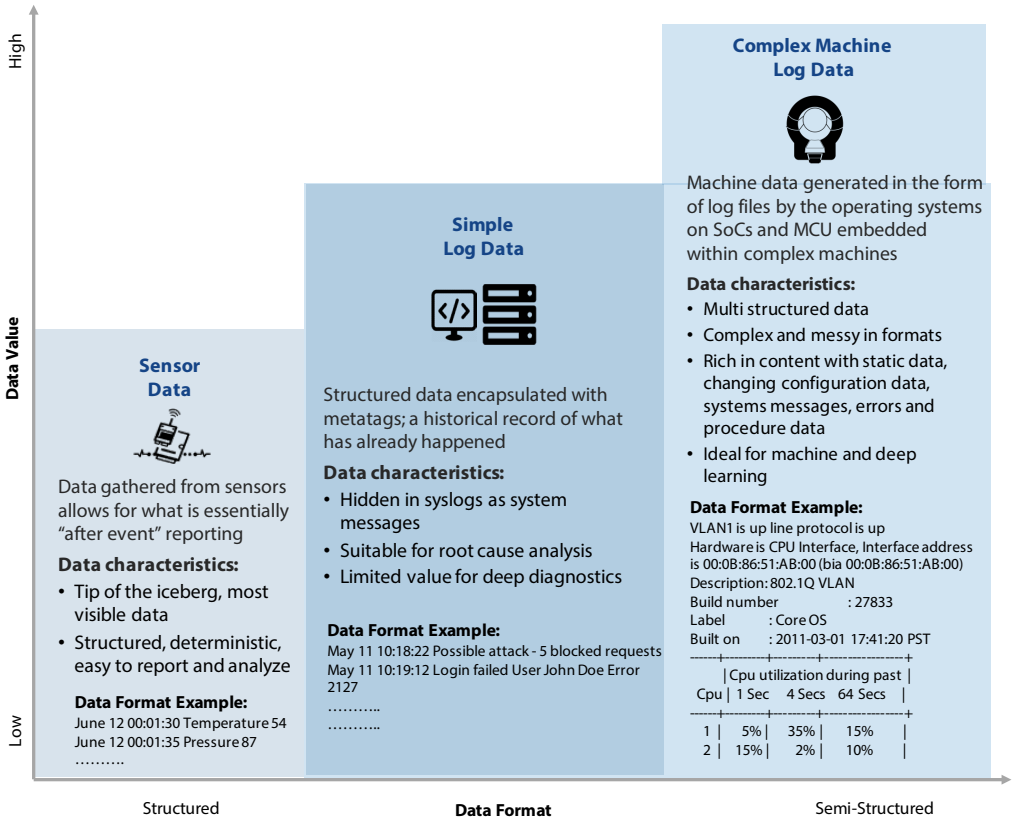
Estimates of data produced by Smart Grid and power generation applications could reach between 35 and 1000 petabytes per year

There are over 500,000 data centers in the world suffering an average of 2.5 outages per year with an average duration of 134 minutes. Globally that translates to 2.84 million hours of annual data center downtime at an estimated cost of $300,000 per hour of downtime, resulting in $426B a year in losses

high-end and complex machines have significant computing power within them, and as this population of machines grows over time, the need for advanced data management and transformation solutions will become critical. The vast majority of IoT data solutions currently on the market can only address sensor and simple log data; these solutions are not able to address the growing amount of complex, multi-structured log data produced by today's advanced machines.

The vast majority of today's IoT data solutions can only address sensor and simple log data; these solutions are not able to address the rapidly growing amount of complex, multi-structured log data produced by today's advanced machines

Exhibit 4: Not All Machine Data Is Created Equal



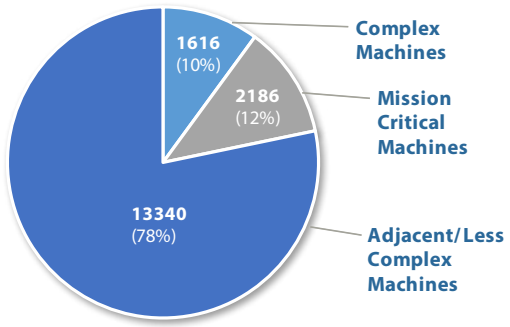
IoT Data Comes in Many Forms, Data Transformation is Critical to Analyzing Complex Data Produced by Advanced Machines

Data management and transformation is a critical step in the data value chain, but unfortunately, this is a topic that has been too confusing and misunderstood for way too long. As a baseline, many organizations know that data is extremely important to their organization, but they are frustrated because they are not getting enough value from the data they are gathering from their connected machines. A major driver of the need for new data management tools is the diversity of data types from an ever-growing number of devices being connected each year.

This value is primarily driven by complex data and the associated data transformation and analytics solutions that derive this value. However, very few machine OEMs or end-customers understand how challenging this dimension truly is. The critical step of data management and transformation accounts for approximately 70% of all data analytics cost and time.

Exhibit 5: High Growth Expected For Complex Intelligent Machines

Installed Base by Device Complexity, 2017 (Units, Millions)



Complex Machines Will Drive The Bulk of Data Transformation, Management and Machine Learning Innovation

In 2017 alone, the Harbor Research Smart Systems Market Model forecasts 4.8 billion new IoT devices will go on-line, driving over $30 billion in new value

Complex Machines Installed by Industry Segment, Worldwide (units, millions)

	2017	2018	2019	2020	2021	2022	2017-2022 CAGR
Servers-Storage-Net Hardware	1194.58	1394.02	1598.72	1795.58	2006.51	2225.93	13.3%
On Road Vehicles	261.97	339.51	433.73	541.60	661.97	794.38	24.8%
Off Road Vehicles	94.64	117.01	145.27	179.15	220.34	271.07	23.4%
Cranes and AGVs	36.86	43.76	51.96	61.34	72.04	84.68	18.1%
Turbines and Other Power Gen	26.40	33.86	43.87	56.47	72.63	93.33	28.7%
Medical Imaging	1.30	1.68	2.24	3.00	4.07	5.59	34.0%
Rail Systems	0.90	1.16	1.50	1.90	2.40	3.01	27.4%
Ships	0.15	0.20	0.27	0.35	0.44	0.56	30.5%
Aircraft Systems	0.08	0.11	0.15	0.19	0.25	0.33	32.8%
Total	**1616.88**	**1931.31**	**2277.71**	**2639.59**	**3040.65**	**3478.88**	**25.89%**

Most existing approaches to data management and transformation have several failings that lead to costly and time consuming data challenges. First, they cannot handle the variety, velocity and volume of data produced by today's increasingly complex assets. Second, most data management tools available to users today are, at best, cobbled together solutions that take an extremely long time to organize data. And third, these outdated solutions do not derive even a fraction of the full value of the machine data being produced.

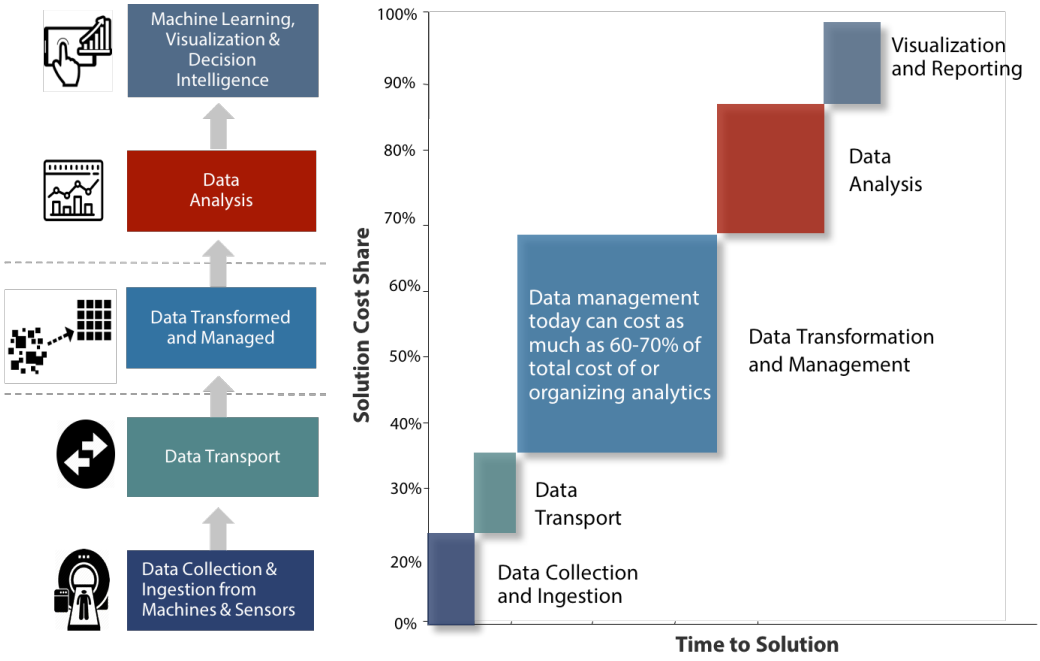
Because many end-users do not understand the challenges associated with complex machine data management, they attempt to develop in-house solutions. These end-users quickly get themselves in trouble by developing hard-to-scale "point" solutions that take months to years of internal development and does not deliver the desired level of value.

ENTER GLASSBEAM

As Smart Systems move beyond the first base of connectivity, the service delivery story becomes critical to deriving new levels of value from gathered data. The need to manage and transform complex data for analysis becomes paramount, and the complexity of building this sort of solution in-house becomes unrealistic. And customers who opt to go with OEM service contracts are spending huge dollars and getting little value in return. This is why an increasing number of companies who are looking for new levels of data value from their advanced machines are turning to Glassbeam; they realize that other offerings cannot address complex machine log data in the manner that Glassbeam is able to.

Exhibit 6: Data Transformation is Critical to Analyzing Complex Machine Data

Data transformation is the bridge between connectivity and high value analytics, and accounts for as much as 70% of all data analytics time and costs



Glassbeam's key differentiators include:

» Ability to ingest, parse and analyze multi-structured, complex log data from advanced assets; the solution goes beyond the simple value provided by analyzing sensor or historical data.

» Much faster time to deployment, which also leads to significantly reduced cost (rapid deployment and reduced man hours). Overall, Glassbeam can deliver 10x the functionality in 1/10th the time at half the cost of other solutions.

» Much more granular development of Rules & Alerts when compared to competitors. Complex log data management uncovers many more variables for analysis than traditional solutions can.

» Integration of multiple data types to enable increasingly complex applications.

» Enablement of machine learning and predictive analytics on complex log data.

» Ease of use for multiple user personas with data different needs.

Glassbeam's key capabilities that drive differentiation:

» SCALAR Platform and Semiotic Parsing Language (SPL): SCALAR is a purpose-built machine data management and analytics solution; the platform leverages Glassbeam's SPL to combine data parsing, ETL and Rules/Actions into a single processing element (ability to address both unstructured and structured data in one single development step).

» Rules & Alerts Engine: Complex event processing technologies that model and capture threshold and anomalous conditions, then send alerts when pre-defined conditions are met.

DATA TRANSFORMATION IS CRITICAL TO ENABLING NEW SERVICES

Glassbeam is a horizontal solution that is actively used in many industries today. The complex data challenges are numerous across all industries, but they can be most clearly seen within the healthcare vertical. All healthcare organizations strive to minimize asset downtime and expenses while also providing a better experience for patients. This can be accomplished by conducting analysis on data from the wide range of medical assets that produce complex machine log data, but almost all healthcare organizations are woefully unequipped to properly leverage this data for advanced use cases such as predictive and preventative maintenance.

Glassbeam is an end-to-end technology solution that works across multiple market segments that utilize compute-intensive machines that produce complex multi-structured log data. Healthcare is an extremely complicated market, and Glassbeam has been able to enter this market and overcome challenges that in-house solutions and OEM solutions simply cannot address.

Moving beyond an industry specific context, OEMs want to extend relationships with their end customers by provide services, but many OEMs have become complacent with the traditional "aftermarket" view of the services opportunity. Complex machine data will play a major role in driving new services for OEMs.

COMPLEX DATA IMPACTS HEALTHCARE SECTOR

Complex data challenges are numerous across all industries, but they can be most clearly seen within the healthcare vertical. All healthcare organizations strive to minimize asset downtime and expenses while also providing a better experience for patients. This can be accomplished by conducting analysis on data from the wide range of medical assets, such as medical imaging systems, that produce complex machine log data.

Enabling advanced healthcare use cases by leveraging complex machine data is a valuable opportunity, but few software suppliers have the technical ability to address this challenge.

Environmental Monitoring

Case: Proper room temperature and other environmental considerations are critical for MRI machines to function optimally. A hospital network's historical batched data gathering identified environmental problems in an untimely manner, leading to significant downtime.

Solution: Glassbeam's Rules & Alerts Engine analyzed environmental data and sent hospital engineers instant email alerts when room temperature was beginning to fluctuate out of required range.

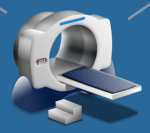
Impacts: Significant cost savings due to avoidance of equipment damage and downtime.

Asset Utilization

Case: A hospital network is unsure of utilization rates for MRI machines and cannot accurately determine if additional machines need to be purchased.

Solution: Analyzing complex machine logs with Glassbeam allowed the hospital to understand utilization rates of MRI machines across the whole hospital network.

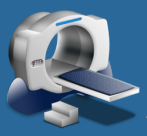
Impacts: This resulting utilization metric provided the hospital network with a clear understanding about current utilization rates, clearly informing if additional assets needed to be purchased.



Glassbeam's Semiotic Parsing Language (SPL) and Rules and Alerts Engine can transform and analyze complex machine data from healthcare imaging equipment with much more depth and efficiency than solutions developed in-house by OEMs or software solutions from competitors.

As healthcare organizations begin to transition away from sorely outdated software solutions, leveraging advanced data management and transformation solutions is a key building block for success. Glassbeam has played this role within many healthcare organizations that are looking to get real value from their asset data. The solution is able to quickly parse complex machine log data and provide advanced analytics that were not previously available. The Glassbeam solution can:

- Quickly transform and analyze complex log data from a wide variety of assets produced by any OEM;
- Enable a wide variety of use cases that were not previously possible, including real-time data visibility, predictive maintenance, and environmental monitoring;
- Allow hospital maintenance staff to reduce downtime by proactively predicting machine failures and alerting staff to order new parts much faster than is possible when using OEM monitoring solutions.



Operator Training

Case: Task completion time by ultrasound machine operators was inconsistent across a hospital network, and the hospital management wanted to identify which operators needed more training to complete tasks more efficiently.

Solution: Complex machine log data from ultrasound machines was used to identify how long each task takes for individual machine operators. Operators taking too long are identified for additional training.

Impacts: 15-20% increase in operator efficiency after training.

Remote Services & Maintenance

Case: Clinical Engineering team at a leading hospital had no centralized data analytics solution for MRI and CT Scanners in place, and had to manually analyze data on key operational parameters such as cold head temperature, magnet pressure, and tube arcs to react to potential machine downtimes in their network.

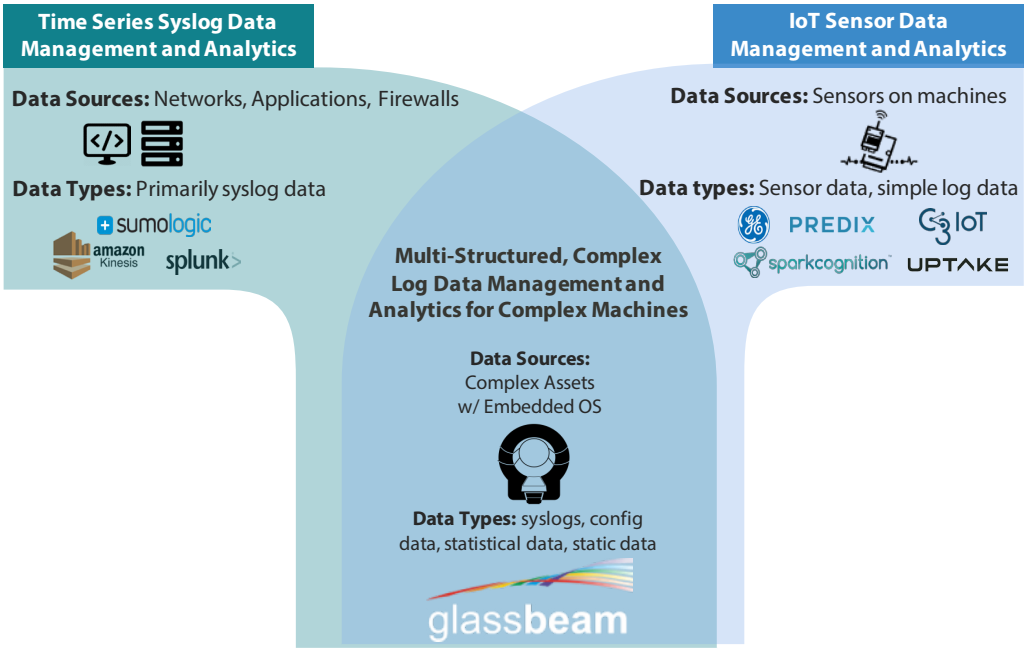
Solution: Transforming and analyzing asset data proactively through Glassbeam, the hospital was able to identify leading indicators of failure and resolve the problems by proactively dispatching maintenance techs or ordering appropriate parts well in advance.

Impacts: Machine availability improved to 99% uptime re-claiming $millions of dollars in lost revenues per year and enhancing patient safety and satisfaction ratings.

Harbor Research projects that leveraging complex machine data from healthcare imaging equipment will produce $11.1 billion of potential revenue value by 2022. Glassbeam's capabilities directly address the bulk of this opportunity.

For OEMs, new data transformation and analytics platforms will open many connected services and value adding opportunities to capitalize on and helps to grow the overall services opportunity because it appeals to customers who previously did not see the value in a services contract. Evolving services opportunities will drive new revenues by as much as 10-20% increase over traditional support sales for OEMs.

Exhibit 7: Glassbeam Addresses Complex Machine Log Data in Ways Others Cannot



Complex log data management enables advanced analytics

Examples of business and economic value for OEMs created by new general services offerings include:

» **Building Aggressive Customer Retention Capabilities:** Many OEMs further along the maturity curve of data analytics have come to recognize that the cost of acquiring new customers pales in comparison to the cost of retaining customers, meaning that ongoing machine data analytics helps accomplish better customer retention by providing customers with continued product intelligence.

» **Driving Recurring Revenue Opportunities:** Although it is still early in the life cycle of new services driven by advanced data management and analytics, many OEMs see the shift away from reactive, transactional services towards a more predictive and managed services stance (larger established OEMs are all approaching attach rates for services contracts in the 90+% range).

» **Capitalizing on Cross- & Up-Sell Opportunities for Equipment & Services:** OEMs are recognizing that as they increase their understanding of systems and support performance through better analytics, they have the ability to focus this data on service performance to enhance their ability to sell that performance as part of their offerings. Further, comprehensive discovery and inventory tools allow customers and OEM partners to see more clearly where systems require updated hardware and support and where equipment is either not supported or at the end of its support life.

» **Improving Business Mix and Profitability:** Many OEMs are coming to believe these systems will help grow new services revenues, drive mix and larger contribution margin. While it is still early from a quantification standpoint, the most aggressive partners are seeing huge performance impacts on their business as a result of adopting advanced data management and analytics solutions.

The rise of the IoT is transforming machine OEMs into value-added service companies. Manufacturers are learning that by putting products on networks they are essentially placing themselves into continuous contact with their customers, thereby enabling them to better understand their customers' needs and act appropriately. The intersection of these trends creates an opportunity for product OEMs to evolve their business model and drive competitive differentiation through new collaborative product and service offerings.

SMART SYSTEMS REQUIRE COMPLEX MACHINE DATA TRANSFORMATION

Implementation of an advanced data transformation and analytics solution from Glassbeam can drive competitive differentiation for a wide variety of customers, including OEMs, end customers and IoT platform providers. For all potential solution adopters, Glassbeam plays a key role in enabling an end-to-end data platform, spanning from data acquisition to data visualization, and is built to handle complex machine log data.

There is huge promise in the value of conducting data management and analytics for multi-structured machine log data. However, many customers underestimate the importance of data management and transformation, and do not understand that the full value of this complex log data can only be realized through advanced data solutions. Glassbeam is purpose-built to address these challenges, allowing end-users to realize new levels of value from data, while also achieving significant cost and time savings.

ABOUT HARBOR RESEARCH

Founded in 1984, Harbor Research Inc. has more than thirty years of experience in providing strategic consulting and research services that enable our clients to understand and capitalize on emergent and disruptive opportunities driven by information and communications technology. The firm has established a unique competence in developing business models and strategy for the convergence of pervasive computing, global networking and smart systems.