ROLTA

INSIGHTS. INNOVATION. IMPACT

The Way forward to Digital Business

Consumerising Data Science and Predictive Analytics for Superior Business Outcomes

Featuring research from

Gartner

Welcome

We are pleased to welcome you again through this Newsletter series from Rolta featuring research from Gartner. In earlier newsletters, we showcased a solution approach to turn big data into business outcome, with a tailored focus on the various asset intensive industries. The four key strategic objectives of Big Data namely, operational excellence, customer intimacy, new business innovations and risk management are best achieved by exploiting the power of Big Data. More importantly, the power of analytics must be made available to business users, not limiting it to niche data scientists and analytics professionals.

The way to exploit the power of data and bring out over 100x its value is a three-step process. First, the real-time data needs to be processed and correlated with historical and empirical data. Second, Data Science models need to be applied to predict what will happen. Third, the models need to be able to prescribe the action that needs to be taken. Understanding enterprise and non-enterprise data (e.g. from the internet) is a challenge that has been addressed by big data processing engines like Spark, Hadoop and other such solutions. The focus, therefore, shifts to a second complex challenge – that of generating real-time, predictive and prescriptive analytics and thereby, insights from that underlying data.

Data Science originated from academic areas like statistics, operations research, artificial intelligence, machine learning, etc. Enterprises have been struggling with Data Science and the approaches to ensure industry class solutions.

In this newsletter, we shall focus on what constitutes Data Science as a discipline and what is required to make it a part of corporate practice. Most importantly, we will discuss how an enterprise can derive immense value out of Data Science. This is done by bringing together business experts, data scientists, and solutioning experts to consumerize the solutions for business users. We strongly believe that an integrated solution approach, with Data Science as a key enabler, has immense potential to transform any organization.

The lifecycle of Data Science from ideas through business understanding, prototyping, refinement and firefighting, requires serious consideration by CXO's of any enterprise. The earlier a solutioning approach is adopted, the better it is to achieve the intended benefits of Data Science. Rolta is a leading provider of innovative IP-Led, Data Science and Big Data Analytics solutions. Leveraging its three decades of expertise in Engineering, Geospatial, IT and experience of working with Global Oil & Gas, Chemical, Utilities, and BFSI leaders, it offers an innovative solution built on best-of-breed platforms and models. We believe that this newsletter will help you in your journey towards consumerizing Data Science for business users to achieve greater business outcomes.

Rajesh Ramachandran
Chief Technology Officer & President GPTS, Rolta

Dr. N R Srinivasa Raghavan
SVP, Big Data Analytics & Data Science COE, Rolta

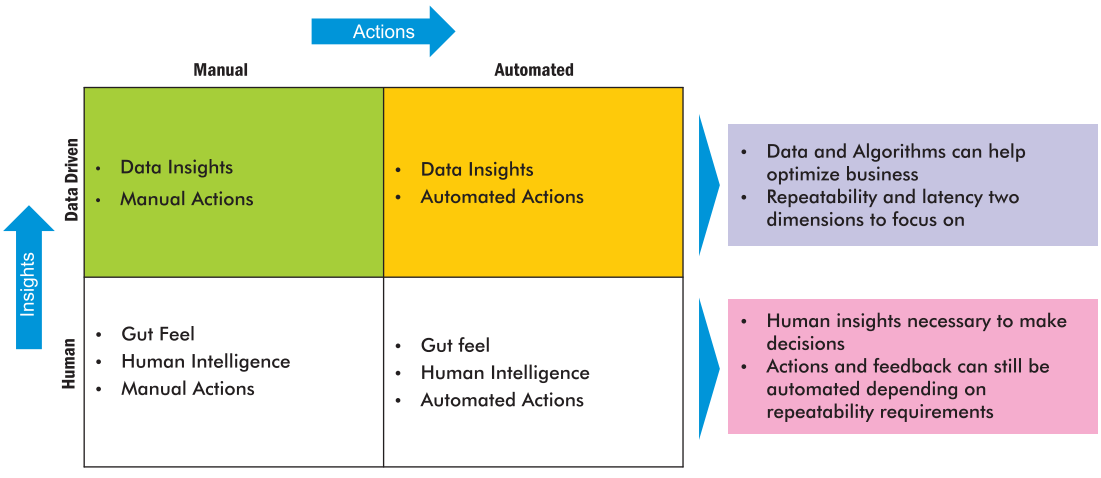
Rolta Insight

In the quest for improvement, organizations have increasingly started to streamline the 'data supply chain' from creation to business use. The increasing availability of systems and the establishment of processes to curate both structured and unstructured data have turned the spotlight on Data Science. The impact of Data Science programs in enterprises is typically gauged by the impact on the bottom line of the business, the ease of use of solutions and the level of engagement of the business stakeholders from definition to deployment. The transformation from a gut feel, judgment driven and manual action based organization to a data insights driven and automated action based organization is pivoted on the concerted action of domain experts, data scientists, and solutioning engineers. The availability of statistical, machine learning and text mining packages has made it possible to craft bespoke solutions to a plethora of operational and tactical problems that the organizations intend to address via Data Science.

However, the value of Data Science is best realized when it is used along with the domain context of where it is being applied to move from Data Science to decision science. Accomplishing this as a bespoke solution often results in longer cycles and iterations with uncertainly of the outcome and desired business impact. Considering that the time to market is often a golden touchstone, we have seen the pre-built industry specific analytics solution with integrated Data Science work bench, big data integration, domain specific models and consumerization of analytics can yield quick and rich benefits. This helps to expand the power of Data Science not just to the data scientist and functional specialists, but the decision science to business users as well. Consumerization is the key to the success.

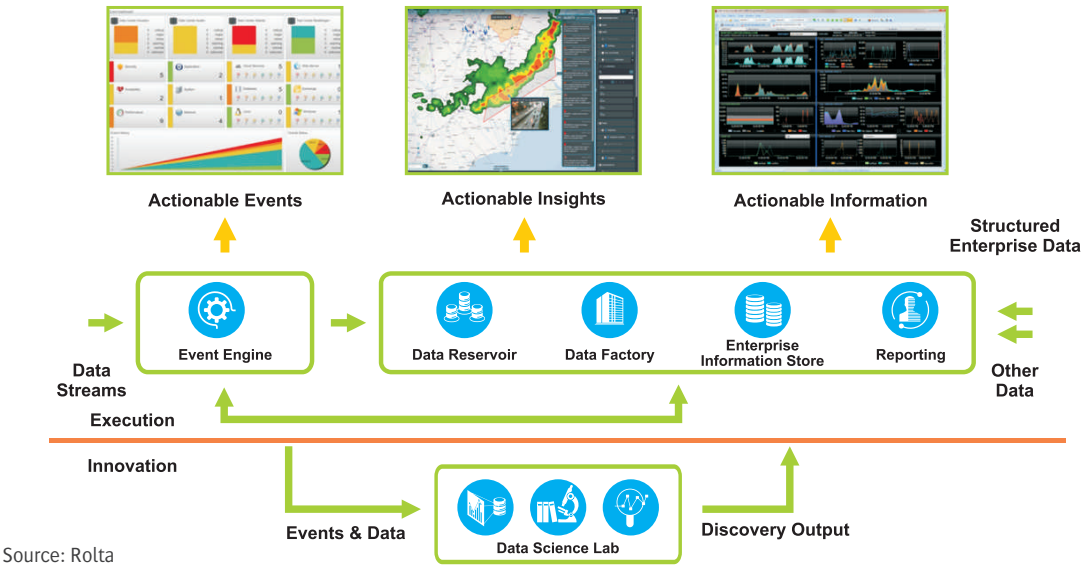
Data Science has been evangelized by all research and consulting majors as a key competence for not only digital and internet companies, but for all other enterprises as well. The flurry of activity in academics and in the software industry to embrace this discipline is incredible. It has even carved a niche for itself as a self-sustaining industry. While the proper use of Data Science results in high ROI for enterprises, it is often a rigorous and profound task to create unique IP based approaches in this area. The success of the diagnostic, predictive, and prescriptive analytics that Data Science can support is not only based on innovations in disciplines like statistics, computer science and operations research, but also in how well the business problem is understood, and in how well a solution is architected based on real data assets from within and outside the enterprise.

Figure 1 : Closing the insights to action loop through automation



Source: Rolta

Figure 2 : Architecture for embedding data science innovation into BI



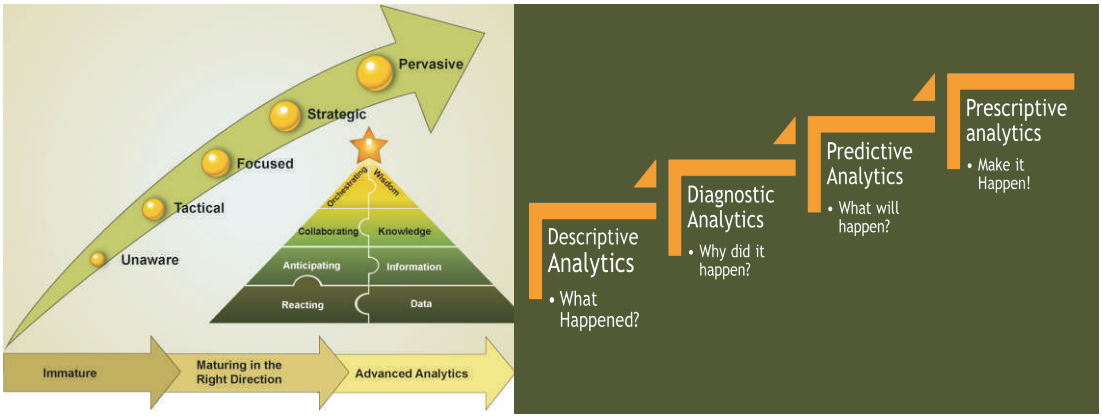
Source: Rolta

Demystifying Data Science for Business

Data Science is becoming vital for business in the digital era. The data landscape has significantly widened. Leading companies need to exploit the power of big data to excel and not just rely on enterprise internal data to control performance. Organizations need to move from reactive to proactive. The analytics maturity journey is no more left as a choice.

Data Science brings value in every step of the analytics journey. The questions, enablers and outcomes are different at each step. Further, value increases 100x moving in the direction of descriptive to prescriptive.

Figure 3 : The maturity map for advanced analytics



Source: Rolta

Figure 4 : The enablers and outcomes for different analytics requirements

	Descriptive	Diagnostic	Predictive	Prescriptive
Questions	What happened? What is happening?	Why did it happen? What went wrong?	What will happen? Why will it happen?	What should I do? Why should I do it? (What If)
Enablers	• Business Reporting • Dashboard and Scorecards • Data Warehousing	• Statistical modeling • Machine Learning • AI	• Data & Text mining • Web/Media mining • Forecasting	• Optimization & Simulation Decision Modeling • Machine Learning
Outcomes	Well defined business problems and opportunities	Identification of root causes of events; field quality analytics	Accurate projections of the future states and conditions	Best possible business decisions and transactions

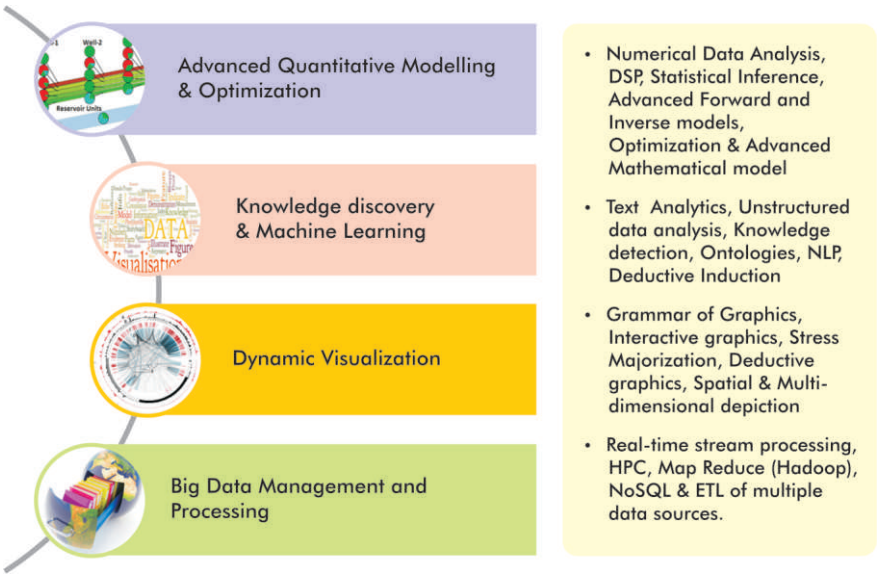
Source: Rolta

Data Science: The critical components

Some of the core components in internalizing Data Science, as illustrated above, are techniques from advanced statistical modeling and optimization, techniques to enable identification of hidden patterns in large data sets with possibly unknown ontologies, techniques to help visualize the data including the geo-spatial dimension and techniques for mugging petabytes of data in quick time.

Several licensed and public domain technologies help in addressing this vast landscape of requirements. Large organizations with deep pockets have started to create their vision around a center of excellence and possibly a laboratory to test these technologies. Frameworks to develop and protect new IP in this area are critical as well. Equally important is the ability to hire and retain a world class talent pool.

Figure 5 : Data Science: The Crital Components



Source: Rolta

Rolta's Integrated Approach for Consumerizing Data Science and Analytics – Rolta OneView™

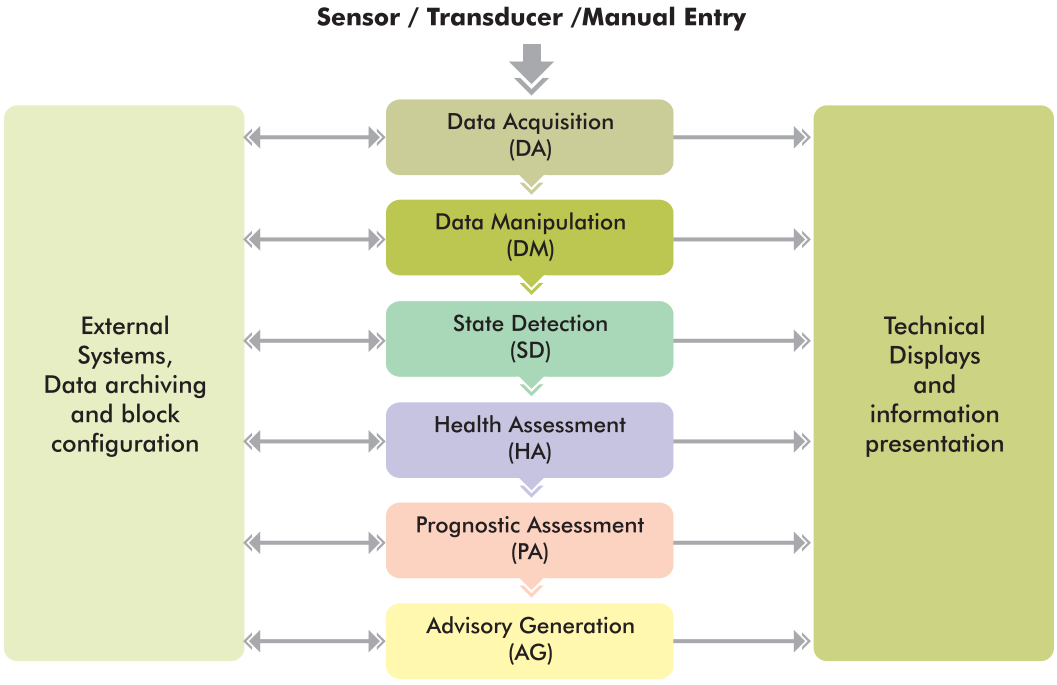
With the availability of the industrial internet of things, ever increasing big data from the internet and Data Science to look for unseen patterns in these data streams, it is extremely critical to look at an end-to-end solutioning perspective for business users. Several standards bodies define the overall framework, which if adopted by data scientists, can enhance the utility of the solutions that they develop.

Rolta OneView™ is an industry proven, pre-built Big Data Analytics Solution, which brings the unique value of OT-IT integration, 1100+ domain specific business process content, comprehensive Data Science platform and 3000+ vertical specific pre-built analytics and business value scenarios. In the field of real-time diagnostic, predictive, and prescriptive analytics in industrial systems that are highly automated for data collection, standards such as the ISO provide useful frameworks to adopt. Rolta OneView™ adopts such standards as part of the Data Science workbench enabled solutions for example for asset health monitoring. It is imperative for the data scientists to work in close collaboration with the solutioning team to ensure all the necessary stages are captured by appropriate applications.

Rolta OneView™ – Integrated solution brings the power of Data Science to create business outcomes

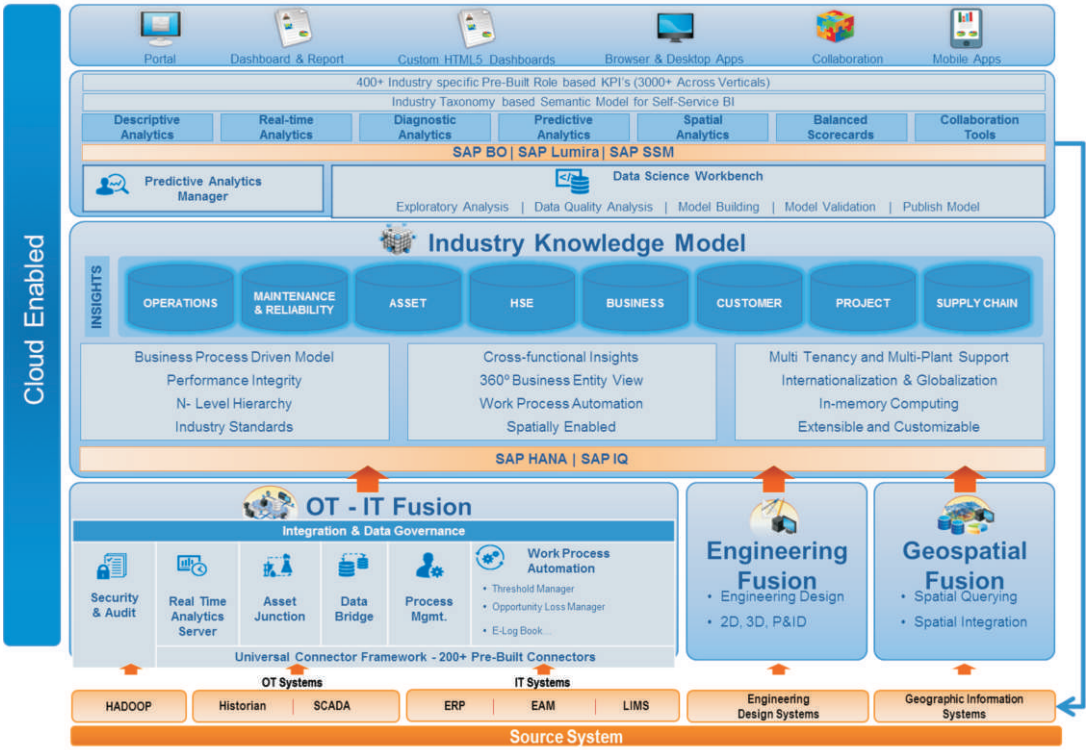
Usability being a fundamental touchstone in the industry, it is necessary to incorporate interfaces for data acquisition, data exploration, visual diagnostic analytics led by the user in a self-service mode, and (where possible) automated advisories based on paid databases on say, failure of complex rotating equipment. With Data Science being embedded in several organizations, users also expect interfaces for model building, model maintenance, and variety of configurable dashboards to support decision making.

Figure 6 : Rolta's standards based on model approach for different use cases / scenarios - ISO 13374



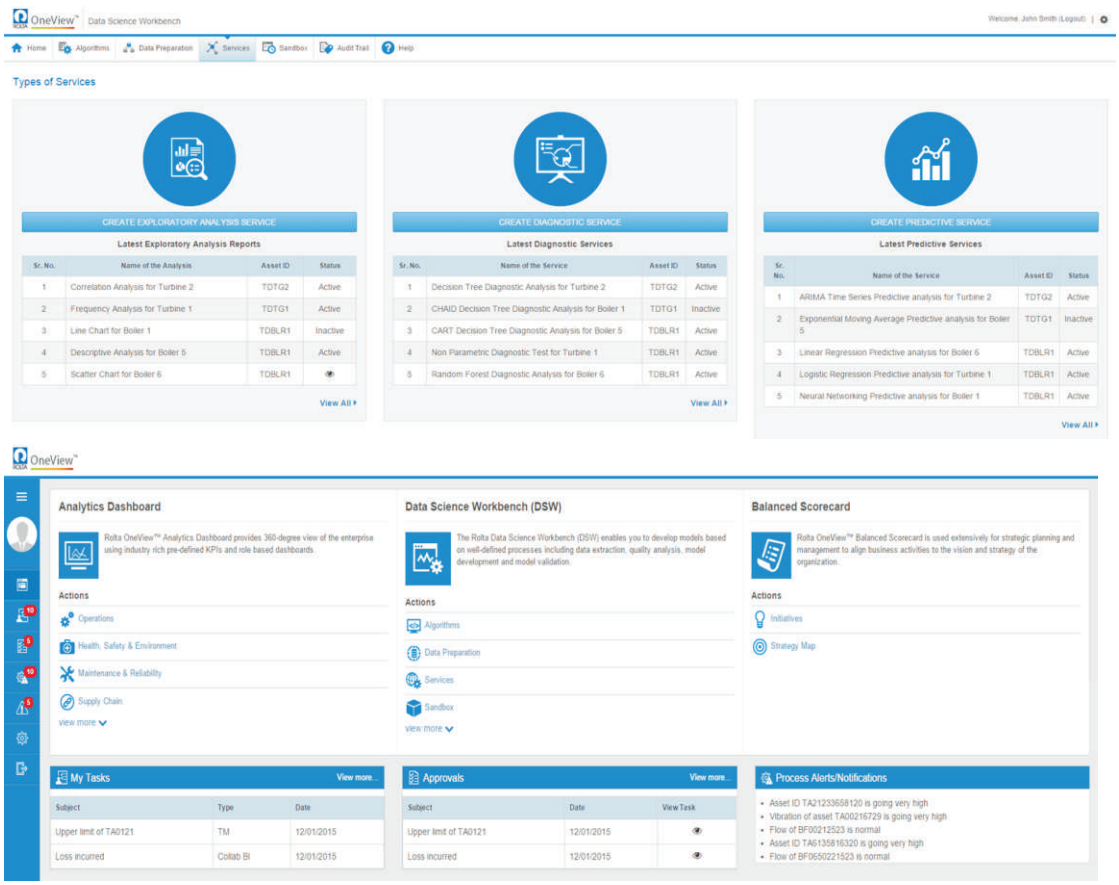
Source: Rolta

Figure 7 : Cloud Enabled



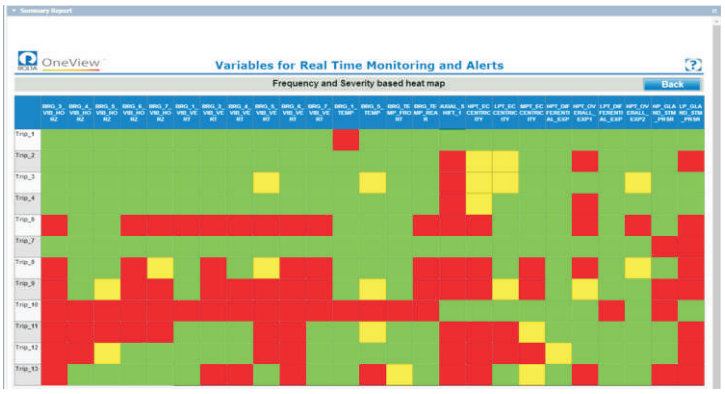
Source: Rolta

Figure 8 : Data Science Workbench with highest usability certified standards



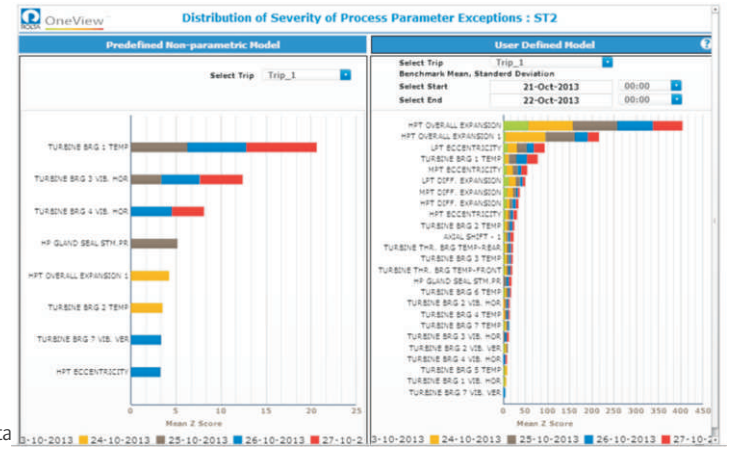
Source: Rolta

Figure 9 : A sample diagnostics analytics enabled by Rolta's Data Science Workbench™



Source: Rolta

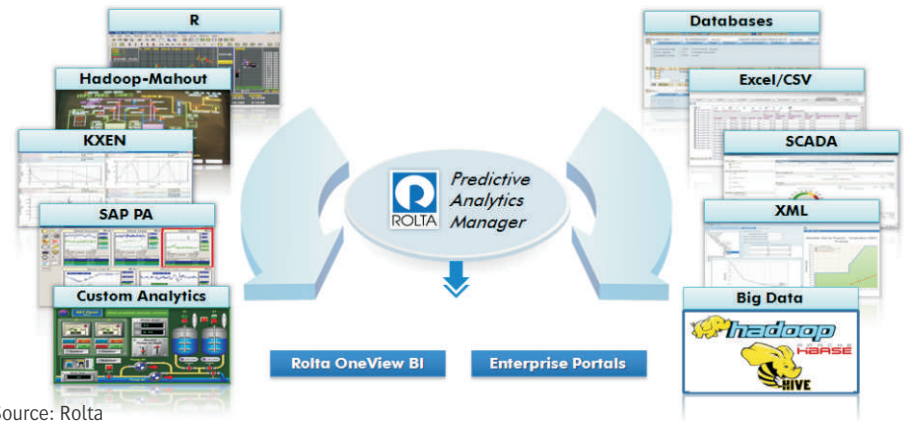
Figure 10 : Distribution of Severity of Process Parameter Exceptions : ST2



Source: Rolta

Figure 11 : Technical Realization of Prediction

- Configurable model deployment option
- Pick and Choose from multiple predictive algorithms (R, Apache Mahout, MATLAB and other predictive libraries)
- Ability to schedule and run processes with zero coding



Source: Rolta

As much important of creating the models through the Data Science Workbench through the standard based approaches, it is equally important to be able to manage and consume in the Analytical dashboards and applications in a seamless way for business users. Again Rolta OneView™'s integrated approach addresses this through its Predictive Analytics Manager that uses Rolta's patented technologies.

Source: Rolta

Case Study 1: Condition based Asset health monitoring for Oil & Gas, Chemicals, Petrochemicals, and Utilities

Asset intensive industries should drive operational excellence through actionable insights with fact based and data centric decision making – combining the power of data science, engineering and operations domain.
The organizations must optimize asset performance to increase availability, minimize costs and reduce operational risks.

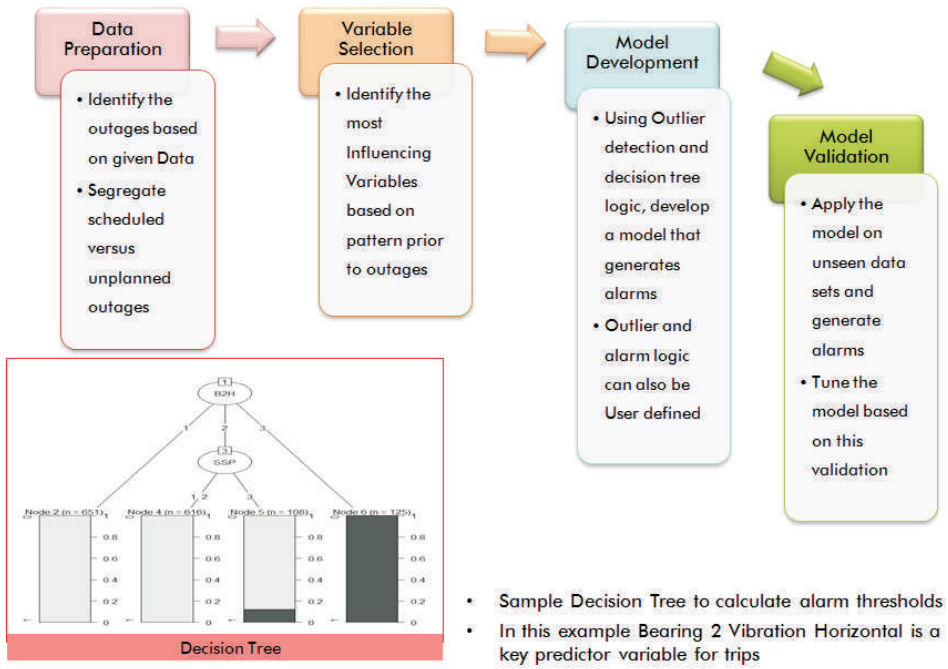
Business Challenges:

- Huge investments in assets necessitates focus on maximizing ROI via asset health monitoring. The business networks and operations networks are not integrated
- Asset data is stored in IT networks and operations data is stored in OT networks
- Meaningful analytics is difficult to be built by aggregating data from disparate network systems
- Lack of comprehensive 360 degree view of operational insights, business insights and asset insights
- Delay in timely decision making and prompt actions to be taken

Solution Features:

- Real-time predictive diagnostic and analytic advisory system that can detect a broad array of incipient abnormalities and provide diagnostic information prior to degradation or failure
- Ability to communicate in real-time with servers located at remote locations such as PI servers, OPC compliant data historians, web services, standard online data sources like SQL, ODBC, XML etc, and occasional CSV/ text files.
- Facility to upload plant/ unit/ equipment specific data sheets, relevant files and drawings as part of knowledge capture.
- Configurable GUI that displays information, diagnostics, analytics, playback, advisory management, workflow support, and knowledge capture with graphics, plots and trends, dashboards, analytics and an user friendly interface

Figure 12 : Diagnosis and Predictive Modeling Approach



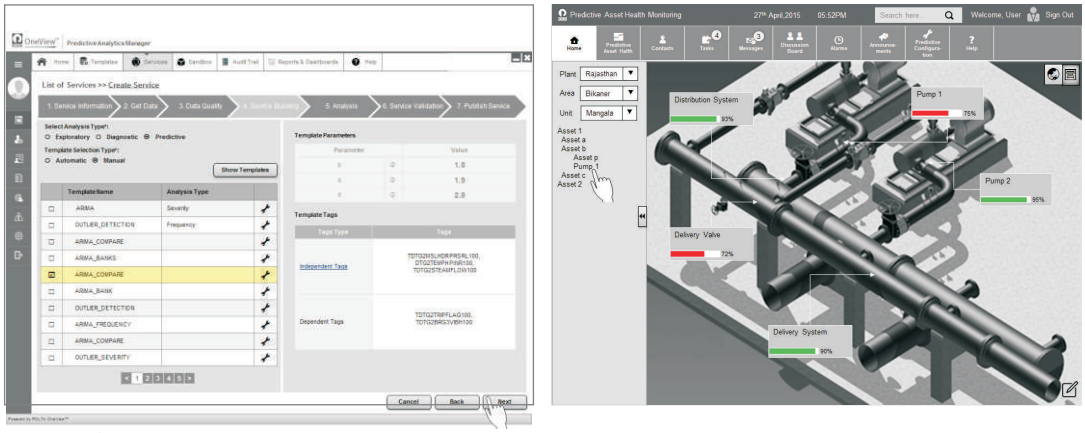
Source: Rolta

Business Benefits:

- Rolta OneView ™ is embedded with diagnostic and predictive modeling tools that enable organizations leverage investments in enterprise and operational technologies
- Phase-wise and methodical approach to enable citizen data scientists within asset intensive industries
- Early warning alerts are generated by patterns of historical data as well as expertise by several standards organizations
- The solution is based on a robust and scalable architecture
- Rolta Predictive Analytics Manager orchestrates real-time data ingestion components with an analytical and view layer based on configurable SAP Design Studio/Lumira
- Options exist to plug in to other visualization frameworks
- The solution is extremely customizable
- Pre-built library of predictive models
- Ability to integrate engineering drawings and geo-spatial data on assets
- Contextual visualizations that can help both plant managers and senior management of the company

Figure 13 : Asset – Predictive Model Integration

- Provision for selecting a prediction model at Unit/Area/Asset level
- Intuitive Mapping of tags with prediction services.



Source: Rolta

Figure 14 : Real Time Condition Based Monitoring Of Steam Turbines

Source: Rolta

Case Study 2: Early Warning Solution for BFSI

Business Challenges:

Banks are typically beset with the problem of having to identify non-performing assets in their corporate lending portfolio. Enough advance notice, if given to the field managers, can help banks to initiate appropriate actions in order to counter the P&L implications. Enormous transactional information pertaining to clients is available with the banks. However, a variety of unstructured text information is also available in various reports, as well as in internet based sources. Leveraging these information assets to arrive at an early warning signal for each customer is thus essential for operations.

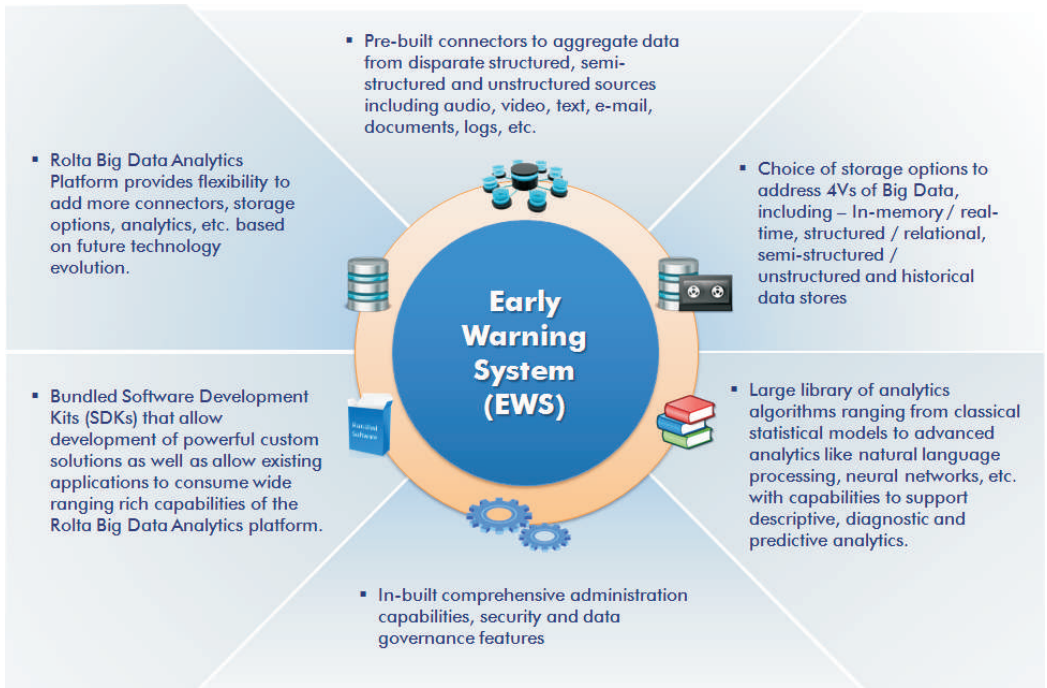
Solution Features:

The Early Warning System is an automated system to identify assets at the incipient stage of stress, categorize the assets/customers into various risk levels and raise alerts to the relevant business owners for appropriate action.

The system gathers and integrates data from both, banks' internal systems as well as a variety of external sources - like web, social media sites, external credit rating agencies, regulatory, legal bodies, industry associations, news sites, etc. It uses this data to derive useful information to monitor and analyze customer behavior, near real-time. Triggers are defined within the system (e.g. return of cheques, delay in submission of stock statement, FFRs, delay in implementation/commencement of project, etc. Different user customizable weightages are assigned to these triggers as per different customer segments using statistics on past data of the bank as well as business experience and knowledge.

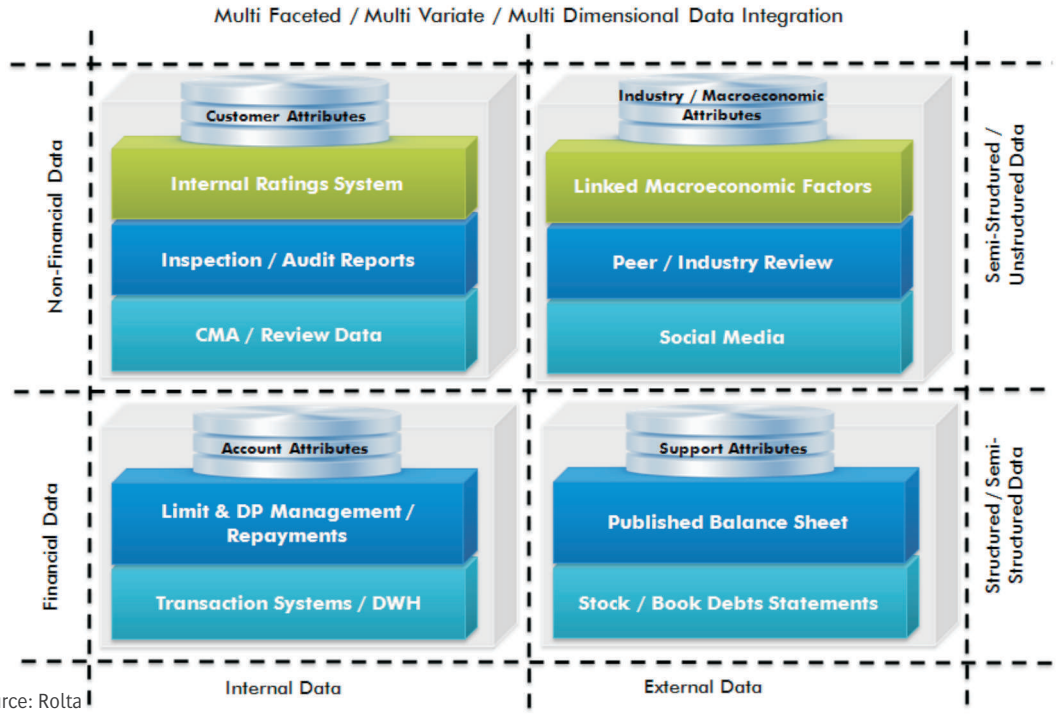
For each customer, a composite score is arrived at based on number of triggers hit, criticality of triggers and type of triggers using advanced analytics algorithms. The customers are categorized into low risk, medium risk, high risk, very high risk based on such calculated scores. Alerts are sent to the relevant business owners as soon as a customer score crosses a predefined threshold level, for appropriate action.

Figure 15 : Technical Features of the Early Warning System



Source: Rolta

Figure 16 : Rolta Early Warning Solutions for BFSI

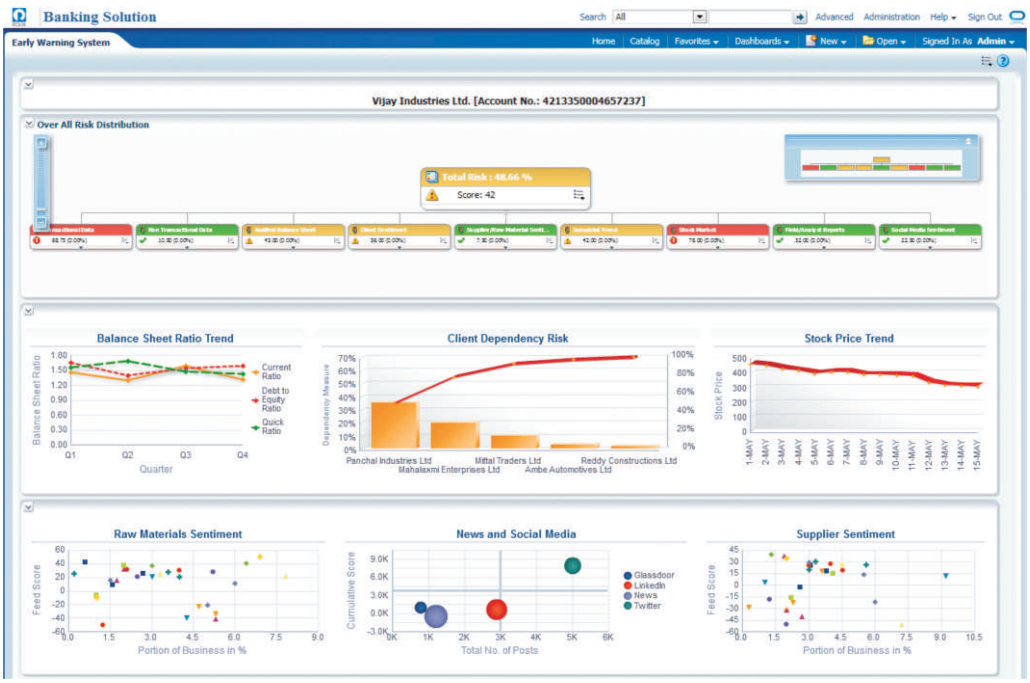


Source: Rolta

Figure 17 : Sample Dashboards

Source: Rolta

Figure 18 : Overall Risk Score



Source: Rolta

Business Benefits

The benefits of Rolta Early Warning Solution are:

- Taking appropriate action early-on, would result in higher cash recoveries of stressed assets and lower additions of NPAs
- Control NPAs, restrict exposure, thereby directly improving profitability as a result of lower provisioning and write-offs requirements
- Facilitate informed decision making and proactive action
- Caters to regulatory, management reporting requirements and facilitates effective, actionable RDA (Risk Data Aggregation) requirements
- A powerful risk data repository to assist / support credit policy formulation and policy back testing

Summary

As Data Science is increasingly being prioritized as a key source of competitive advantage in the digital economy, it is essential to garner as much enterprise wide adoption as possible for it to yield significant benefits. In that direction, companies need to mature in the data supply chain in terms of the level of digitalization, and embedding Data Science into company operations via enterprise applications. Building the capability both internally and externally through trusted partners is a key enabler for Data Science implementations. An integrated approach of bringing the data integration, domain business processes, Data Science driven modeling, managing and consuming the model, seamlessly integrating it with the analytical dashboards to see the insights contextually would consumerize the Data Science to business users and reveal significant business outcomes. Rolta OneView™ is a proven solution that has been architected grounds up to bring the value of this integrated approach and turns data into business outcomes in weeks to provide faster and higher ROI with lowest TCO.

Source: Rolta

Research from Gartner:

How Data Science Projects Deliver Business Impacts

Analytics leaders need to understand the different business impacts that data science projects can deliver on business. We group them here into five: innovation, business understanding, prototyping, refinement and firefighting.

Impact

- Data science projects can be characterized by their business impact, helping analytics leaders to understand the business benefits they deliver, from innovation and business understanding to prototyping, process refinement and firefighting.

Recommendations

Analytics leaders:

- Ensure that senior data scientists are part of innovation projects — only then can you be sure not to miss out on innovations that can be framed as data science projects.
- Use your data science team to support production teams for continuously improving enterprisewide model management and performance monitoring.
- Create a portfolio of analytical scenarios, including those your organization is already executing or planning, to better rationalize funding decisions for data science projects.

Analysis

As data is already everywhere, so data science problems are becoming increasingly prevalent (see Note 1 for data science definition). Some enterprises have counted 50 to 100 data science use cases. To better cope with the sheer mass of projects, some leading organizations are starting data science teams whose general mission is to become a shared resources across the organization (see Note 2 for data science team definition, and "Staffing Data Science Teams").

Figure 1 summarizes the impact of and top recommendations for data science projects.

At the macrolevel, data science projects can be utilized to deliver the following high-level business impacts, which we discuss throughout the note in more detail:

(a) **Innovation** — Foster new thinking based on data science

(b) **Business Understanding** — Explore yet unknown patterns in data

(c) **Prototyping** — Challenge the status quo with radical new solutions

(d) **Refinement** — Continuously improve existing in-production solutions

(e) **Firefighting** — Identify the drivers of certain upcoming situations

Figure 1 : Impact and Top Recommendations for Analytics Leaders

Impact	Top Recommendations
Data science projects can be characterized by their business impacts, helping analytics leaders to understand the business benefits they deliver	• Ensure that senior data scientists are part of innovation projects so that innovations that can be framed as data science projects. • Use your data science team to support production teams for continuously improving enterprisewide model management and performance monitoring. • Create a portfolio of analytical scenarios to better rationalize funding decisions for data science projects.

Source: Gartner (June 2015)

At the microlevel, of course, data science projects and teams can have many more contributions:

- Coaching citizen data scientists and validating their work (see "Smart Data Discovery Will Enable a New Class of Citizen Data Scientist").
- Educating the entire organization to become more analytics-driven and moderate the data discussion.
- Fostering networking across an organization.
- Driving discussions on data sources, standards, governance and interesting case studies.
- Communicating the state of the art in data science to all data scientists.
- Cataloging internal data sources and relevant external ones.
- Suggesting data development initiatives for acquiring more data, whether internal or external.
- Evaluating tools and service providers.
- Creating a portfolio of analytical scenarios, including those your organization is already executing or planning, to better rationalize funding decisions for data science projects.
- Studying what your industry peers and adjacent industries are doing, and including these activities in your portfolio of analytical scenarios.

Impacts and Recommendations

Data science projects can be characterized by their business impact, helping analytics leaders to understand the business benefits they deliver, from innovation and business understanding to prototyping, process refinement and firefighting

Innovation — Foster New Thinking Based on Data Science

Without data scientists and their knowledge, many issues surrounding the digital business age will remain unresolved — possibly even untouched. Data scientists frame complex business problems as machine-learning or operations research problems. Data scientists know which new information sources should be collected or acquired from external sources, to solve old burning business issues in radically new ways. (Remember Moneyball, the 2003 book and 2011 movie, where sabermetrics was popularized to completely question the old method of evaluating the performance of individuals and teams in baseball.1)

There are many more examples of disruptive projects and new "business moments" (see Note 3) made possible through data:

- In the 1990s, Google achieved its incredible success by using a previously untapped data source: the hyperlinks encoded in Web pages.
- Also in the mid-1990s, Amazon started one of the earliest recommendation services ("here are four other items that customers buying this product also bought"), which became one of the most prominent and lucrative data science projects in history. Rumor has it that 15% to 20% of Amazon's retail business is due to this simple product recommendation. In fact, it became a desirable feature, with customers wanting to explore related items for any given product.
- UPS On-Road Integrated Optimization and Navigation (ORION) revamped route optimization using many new data sources. It has enabled UPS to significantly improve its routing schedules, saving hundreds of millions of dollars per year while improving customer service.
- IBM Watson's Jeopardy-winning natural-language system was based on crowdsourced data and cutting-edge assembly of different machine-learning and natural-language approaches.

Companies also use data and the corresponding analytics in novel ways. For example, Progressive was one of the first to create an insurance product that used GPS-based location intelligence to keep it better informed about the actual risks against which it is insuring

Many online companies have been masters of data-driven innovation. The likes of Amazon, Google, Airbnb, Uber and Facebook constantly introduce new systems to collect better information. This enables them to create better or new services.

Recommendations for analytics leaders:

- Use your data science team to frame complex business problems not yet sufficiently solved as data science problems.

- Find inspiration for data-driven innovation from three sources:
- Your own thinking — You are your most important source of inspiration. Constantly think about your own business model, industry and understanding of new types of customer or equipment interaction points.
- Technology screening — Learn what you can from successful case studies from your own industry or other industries. But be cautious, because many publicly available case studies may not fully reflect exactly what happened.
- Induction from data — Examine how data expeditions can support your thinking process and how they can uncover novel and insightful patterns that teach you more about the underlying business mechanics.

Business Understanding — Explore Yet Unknown Patterns in Data

Data scientists must engage with big data expeditions, especially when there is no clear objective other than to explore the data for insights and tidbits. Such expeditions are a form of inductive thinking or inductive reasoning (see Note 4) — an example of "letting the data speak." The process can be tactical and ad hoc. Alternatively, it can be part of a more systematic practice in which you give the data science team a data dump for diving into and exploring. The lab then looks for anomalies, seeking something new. The most basic techniques are:

- Looking at the basic shape of the data by examining univariate and bivariate distributions, correlation coefficients and all sorts of visualizations.
- Drilling deeper into the shape of the data using more advanced techniques. These might include cluster and factor analysis, anomaly detection, regression, decision trees, Monte Carlo simulation and link analysis.

The objective is always to:

- Discover which events are drivers — or inhibitors — of other events, or of good or bad outcomes (such as reducing equipment failure and increasing customer satisfaction).
- Gain an understanding of events that could be new customer touchpoints or engagement points. Such information could be used to foster data-based innovation.

These kinds of projects can be a bit like fishing trips. The available data may give hints about what you may gain from the process, but there is no guarantee. Sometimes the output may "only" be:

- A better understanding of business mechanics
- Validation that the data is clean or you need to answer some questions about it

Recommendations for analytics leaders:

- Use your data science team to spot anomalies in data before you notice any problems, not after a crisis happens. View it as a form of prevention or a means of solving problems early, as with police doing regular patrols or people going for routine medical checks.
- Ask your data science team to look at the data again when new information sources appear or when you gain new understanding. This can prove very worthwhile.

Prototyping — Challenge the Status Quo With Radical New Solutions

Data science and especially machine learning excel in solving complex, data-rich business problems where traditional approaches, such as human judgment and exact solutions, either increasingly fail or deliver inferior solutions (see "Machine Learning Drives Digital Business"). Data science methods have been proven to often deliver superior results, when the space of critical variables is highly dimensional and very noisy.

Hundreds of new business problems exist that data science teams could tackle. Companies are already using data science teams for tasks such as:

- Improving product categorization. Many large online retailers realize that their product classification may have errors or not fit the way customers think about products or want to access them. Data science teams are seeking to improve product categorization by using all available features. These include: look, shape, purpose, codes (such as European Article Numbering and North American Industry Classification System codes), product text descriptions and user-generated tags.

- Predicting more accurately which passengers who buy airline tickets will fail to arrive for their flights. More accurate predictions enable airlines to oversell their planes. This minimizes potential lost revenue from empty seats while reducing the risk of passengers arriving to find that there is no seat available for them.
- Inventing devices to provide early detection of diseases — for example, the competition with the highest reward on Kaggle in April 2015 concerned the detection of diabetic retinopathy. The competition challenged data scientists to design a model that would result in an enhanced automated detection system for this disease.

Recommendations for analytics leaders:

- Assess whether it is best to design a radical new solution or buy or outsource one. Sometimes it is better for the business to have an unexceptional solution soon than a potentially great solution in a year. It is also possible that you could use both.
- Be very cautious when your data science team uses particular data for the first time. Some data was never intended for serious advanced analytics, so scrutinizing data lineage and making the data make sense are paramount.
- Allow your solution to go through a "model factory," which involves a metasearch that tweaks a set of acceptable solutions to gain a few extra percent in lift, classification or estimation accuracy.

Refinement — Continuously Improve Existing In-Production Solutions

- Most data scientists in the industry work in the production part of the business. In such areas, established models are already "in production." For example:
- Banks, retailers, telcos and insurance companies are constantly refining their existing customer segmentation, in order to gain a better understanding of customer profitability and customer behavior and engagement optimization.
- Retailers keep recalibrating propensity-to-buy models while online retailers are constantly improving and updating price elasticity prediction, in order to optimize their dynamic pricing.
- Financial services providers are continuously working to improve their risk models — the more accurate their assessment of risk, the better their chances of profitability.

In all these use cases, organizations must constantly improve advanced analytics because:

- New data keeps arriving, simply due to new products being created all the time.
- Customer behavior changes not only by season, but also year after year (through competition, the zeitgeist and an ever-changing marketplace).
- Customer touchpoints change each year, and sometimes even more quickly.
- New customer segmentation strategies can lead to better results — and require many existing models to be adjusted.

Recommendations for analytics leaders:

- Use your data science team to support production teams in creating and improving enterprisewide model management and performance monitoring.
- Use your data science team to help production teams create a more homogeneous and cutting-edge compute architecture in terms of hardware, cloud and software stack.
- Ensure your data science and production teams jointly explore the external data landscape and deploy cutting-edge algorithms (for example, ensemble techniques).

Firefighting — Identify the Drivers of Certain Upcoming Situations

Sometimes it may be almost impossible to avoid a crisis because insights into issues that may cause problems can be so well hidden. In such cases, use your data science team to help resolve the crisis. This use is a variation of the big data expedition use of data science teams. Many analytics projects are triggered by crises. When you ask a data science team in this way, you already know the "symptom" of the crisis. For example:

- Customer complaints have suddenly risen
- Customer retention has fallen dramatically
- Quality defects have increased
- Profitability has dropped

This means that the data science team has to identify "only" the cause, which narrows the datasets it must scrutinize.

Everything else in this use scenario is very similar to the work the lab does in big data expeditions. As in big data expeditions, the lab does not know at the outset whether it can identify the cause of the problem. Indeed, it is possible that the lab may never be able to identify the cause.

Basic data discovery/self-service business intelligence can often help. However, a deeper dive by a data science team can extract more from the data about what is really happening. For example:

Manufacturers worldwide are looking into the causes of quality fluctuations by combining "what if" analysis with sensitivity analysis or inversion of predictive models.

Technical support operations are trying to understand the drivers of maintenance costs. It is known that certain customer segments are more difficult to deal with than others. Factoring these risks into pricing can be crucial and is a well-established practice in the insurance industry.

Online retailers are investigating the reasons why customers return purchased goods.

Recommendations for analytics leaders:

Apply the Occam's Razor principle: Data scientists shall establish trust by applying the "simplest" methods that still deliver the key insight.

Leverage firefighting projects to expand the data science team's corporate network whenever possible.

Evidence

[1] M. Lewis, "Moneyball: The Art of Winning an Unfair Game," W. W. Norton & Co., 2003.

[2] The Diabetic Retinopathy Detection competition on Kaggle started on 17 February 2015 and is due to finish on 27 July 2015. The California Healthcare Foundation sponsored it with a reward of $100,000.

Note 1

Data Science

Data science is the discipline of extracting nontrivial knowledge from all kinds of data, to improve decision making. It involves a variety of steps, ranging from business understanding and data preparation to building and deploying analytic models. It is, to some extent, a replacement term for data mining, but is also much more: data science is the unification of several quantitative disciplines (statistics, machine learning, operations research, computational linguistics, and others). For the first time, people trained in these different disciplines are all willing to unite behind the banner of data science — which is a very profound development.

During the past year, this notion of data science has become more widely used and many more academic institutions continue to offer data science courses and degrees. In addition, organizations hiring data scientists and building data science teams and data science labs are on the rise. Gartner expects that within a few years, the term data science will gain widespread recognition as an umbrella term for many forms of sophisticated analytics.

Organizations that want to increase the maturity of their analytics and extend their portfolio of analytics capabilities need to improve their data science skills. They need to leverage new data sources and demonstrate business value using predictive and prescriptive (and often diagnostic) capabilities. However, organizations must recognize that data scientists are in very short supply — recruiting them internally may be difficult, but not impossible. They must also leverage their "citizen data scientists" in their lines of business to increase the reach and impact of analytics.

Data science drives a vast array of use cases across all industries, for example, customer relationship management, supply chain management, optimization and automation of diverse production processes, drug research, quality and risk management, smart cities, smart systems and many more.

Note 2

Data Science Team or Lab

A data science team, or lab, is a team disconnected from — but close to — the business intelligence competency center. Its individual members usually have different skills. For example, these might be in:

- Advanced statistics
- Business process engineering
- Programming of distributed processing
- Information architecture
- Management

The team becomes a "lab" when you provide it with resources (for example, server and storage sandboxes or relief from its daily workload). It often has a ratio of solutions to "dead end" efforts in the region of 1:10.

Note 3

Business Moments

Gartner defines a business moment as a transient opportunity that is exploited dynamically. It is very short in duration — perhaps only seconds — depending on the nature of the opportunity. This catalyst sets in motion a series of events involving people, businesses and "things" that span multiple industries and multiple ecosystems.

Note 4

Inductive Reasoning

Inductive reasoning aims at creating broader generalization from observations. Even though the facts that produce the generalization can be true, the generalization itself might not always be accurate. For example, if it has been sunny each time you have visited Dusseldorf in Germany, you might conclude — falsely — that it is always sunny in Dusseldorf.

Source: Gartner Research, G00274703, Alexander Linden, 30 June 2015

About Rolta

Rolta is a leading provider of innovative IT solutions for many vertical segments, including Utilities, Oil & Gas, Manufacturing, Federal and State Governments, Defense, Homeland Security, Financial Services, Retail, and Healthcare. By uniquely combining its expertise in the IT, Engineering and Geospatial domains Rolta develops exceptional solutions for these segments. The Company leverages its industry-specific know-how, rich repository of field-proven intellectual property that spans photogrammetry, image processing, geospatial applications, Business Intelligence, Big Data analytics and Cloud computing for providing sophisticated enterprise-level integrated solutions. Rolta OneView™ Enterprise Suite is one such innovative Big Data Analytics solution for asset-intensive industries with field-proven benefits. It brings unique value through its IT-OT integration and Predictive Analytics capabilities using Rolta's patented technology. Rolta is a multinational organization headquartered in India with revenues exceeding US$ 600M and over 3500 people operating from 40 locations worldwide and has successfully executed projects in over 45 countries.

Corporate and Global Headquarters:

Rolta India Limited
Rolta Tower A,
Rolta Technology Park, MIDC,
Andheri (East), Mumbai 400 093. India
Tel : +91 (22) 2926 6666, 3087 6543
Fax : +91 (22) 2836 5992
Email : indsales@rolta.com

India Regional Offices

Delhi NCR
Rolta Technology Park, Plot #187, Phase I,
Udyog Vihar, Gurgaon - 122 016.
Tel : +91 (124) 439 7000

Bangalore
Mittal Towers, 'C' Wing, 8th Floor, 47 / 6,
M. G. Road, Bangalore - 560 001.
Tel : +91 (80) 2558 1614/1623

Kolkata
501, Lords, 5th Floor, 7/1 Lord Sinha Road,
Kolkata - 700 071.
Tel : +91 (33) 2282 5756/7092

International Offices

Rolta International Inc.
(North America Headquarters)
Rolta Center, 5865 North Point Parkway
Alpharetta, GA 30022, USA.
Tel : +1 (678) 942 5000
Email : info.na@rolta.com

Rolta International, Inc. -- Chicago
333 E. Butterfield Road, Suite 900,
Lombard, IL 60148, USA.
Tel : +1 (630) 960 2909
Tel: +(905) 754 8100

Rolta UK Ltd.
(European Headquarters)
100 Longwater Avenue, Green Park,
Reading RG2 6GP, United Kingdom
Tel +44 (0)118 945 0450
Email : info.eu@rolta.com

Rolta Canada Ltd.
140 Allstate Parkway, Suite 503 Markham,
Ontario L3R 5Y8, Canada.
Tel : +1 (905) 754 8100

Rolta Benelux BV
Jupiterstraat 96, Building Pluspoint Nr. 2,
2132 HE, Hoofddorp, Postbus 190,
2130 AD Hoofddorp, Netherlands.
Tel : +31 (23) 557 1916

Rolta Middle East FZ-LLC
(Middle-East Headquarters)
Office No. 209-211, Building No. 9,
P.O. Box 500106 Dubai Internet City,
Dubai, United Arab Emirates
Tel : +971 (4) 391 5212
Email : info.me@rolta.com