Big Data & Analytics
White Paper

Predictive Asset Optimization

Optimize the production time and life expectancy of your assets

IBM

Introduction

This paper discusses an IBM vision, point of view, business use cases and a solution framework that collectively showcases IBM's thought leadership, capabilities and differentiators in this emerging discipline of predictive asset optimization (PAO).

Although the paper focuses on applying PAO across multiple industries in the industrial sector, the vision, point of view are applicable to a wide variety of domains and business use cases.

IBM vision

Imagine a world where we have made our machines so smart that we can help them predict when they are falling "sick" and arm them with a 911 button – to call for help.

Imagine a world in which, during a normal and routine scheduled maintenance cycle, we could predict which component parts are likely to fail in the near future and subsequently be proactive in fixing or replacing the component, thereby avoiding costly unscheduled downtime in the near future.

Imagine that we can actually detect patterns in operator and equipment performances that lead to production inefficiencies and losses; publish alerts and suggest mitigation and remedial actions in real-time.

In a world where machine downtime and production losses are being accepted as normal, PAO holds the promise to pave the way for predictive and prescriptive operations and take us into a future of self-healing autonomous machines.

What if the actual maintenance of heavy equipment and oil platforms is performed only when they are needed? What if our cars gave us an alert on the dashboard saying that they need maintenance and immediately follows that up with suggested certified workshops and appointment timeslots? What if aircrafts can send enough information about operational issues to the ground so that the plane, upon landing, receives an analysis on what is required to address an issue and whether it can be completed and available to meet scheduled departure?

These are important machines that influence and define our lifestyle. They bring fuel to our cars, bring electricity to our households, extract natural resources from the earth, and promise a safe and enjoyable flying experience. It is important that they are properly maintained and their life expectancy optimized with better operating conditions.

Take a step back and now imagine a world where machine downtime, related operator inefficiencies and product losses are non-existent.

This is a game changer. In a world where machine downtime and production losses are being accepted as normal, this holds the promise to pave the way for predictive and prescriptive operations and take us into not too distant a future of self-healing autonomous machines.

This is the art of the possible. Technology is at the forefront helping us make this a reality today.

An IBM point of view

Currently the industrial sector is facing a set of imminent realities:

- A significant amount of catastrophic failures, be it machine or operator related, will continue to be random in nature unless we make a conscious decision to act upon it. It is important to point out that most of these random failures can be predicted with a high degree of confidence, if a commensurate amount of investment is made in addressing the problem – we are just choosing not to do so.
- The total cost to the company, of unscheduled downtime, ranges and can continue to range, anywhere between three to ten times that of a typical cost profile for scheduled maintenance.
- Enterprises continue to prefer and adopt a more conservative approach to maintenance scheduling and planning in an effort to avoid the randomness of failures.
- Companies staying ahead of the game plan on investing into PAO capabilities and infuse it into base business operations.

The companies that will have a competitive advantage in the marketplace will be the ones that will break away from the traditional approach of human intuition and expertise based sense and respond mode of business operations to one in which the next generation of manufacturing and production and in-service efficiencies will be achieved by providing precise, contextual analysis at the point of impact, thereby adopting a real-time, fact-driven, predict and act modus operandi, as depicted in Figure 1. This fundamental shift can only be made possible through a serious investment in analytics as a business strategy.

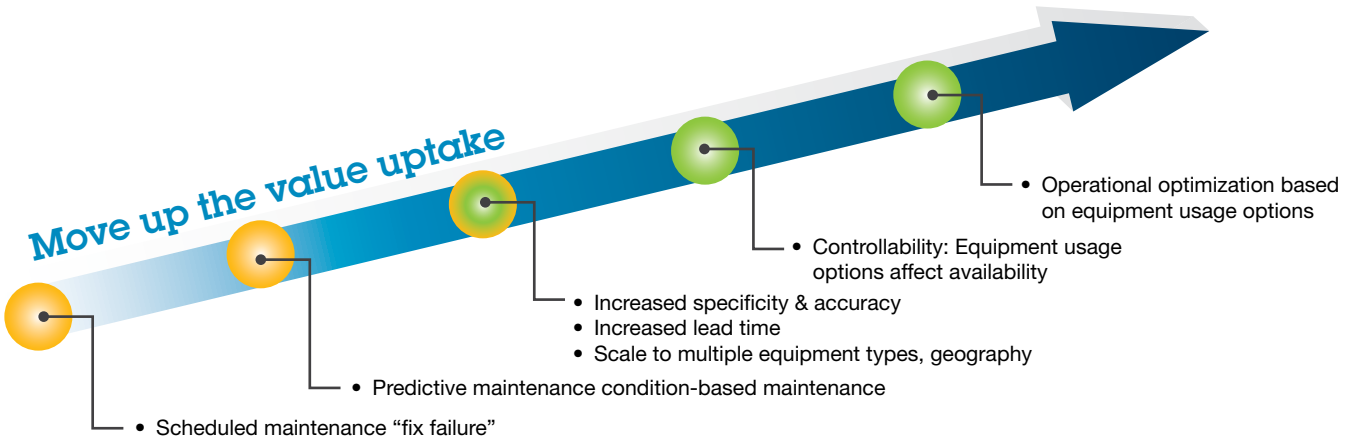

Figure 1: **Gaining competitive advantage requires breaking away from the current norm.**

The investment in analytics needs to be supplemented with an innovative solution approach that is built on a strong foundation of various complimentary advanced analytics techniques that collectively provide a 360 degree view of equipment health, along with operator and machine production parameters, in real-time.

Enterprises need to realize that the most advanced analytical solution may not be practically realizable at the onset. However, there should be a strong, focused and committed intent to embrace a continuous investment in predictive asset monitoring, management and optimization through iterative improvements in prediction accuracies.

As an example, applying PAO to equipment health monitoring, the advantages of achieving higher prediction accuracies is incremental in the mean time between predicted failure and the actual failure. Referring to Figure 2, this implies moving farther left and top in the performance-failure (P-F) curve.

Early issue detection and resolution remains the primary driver for increasing equipment/asset uptime and reducing operator and machine production inefficiencies.

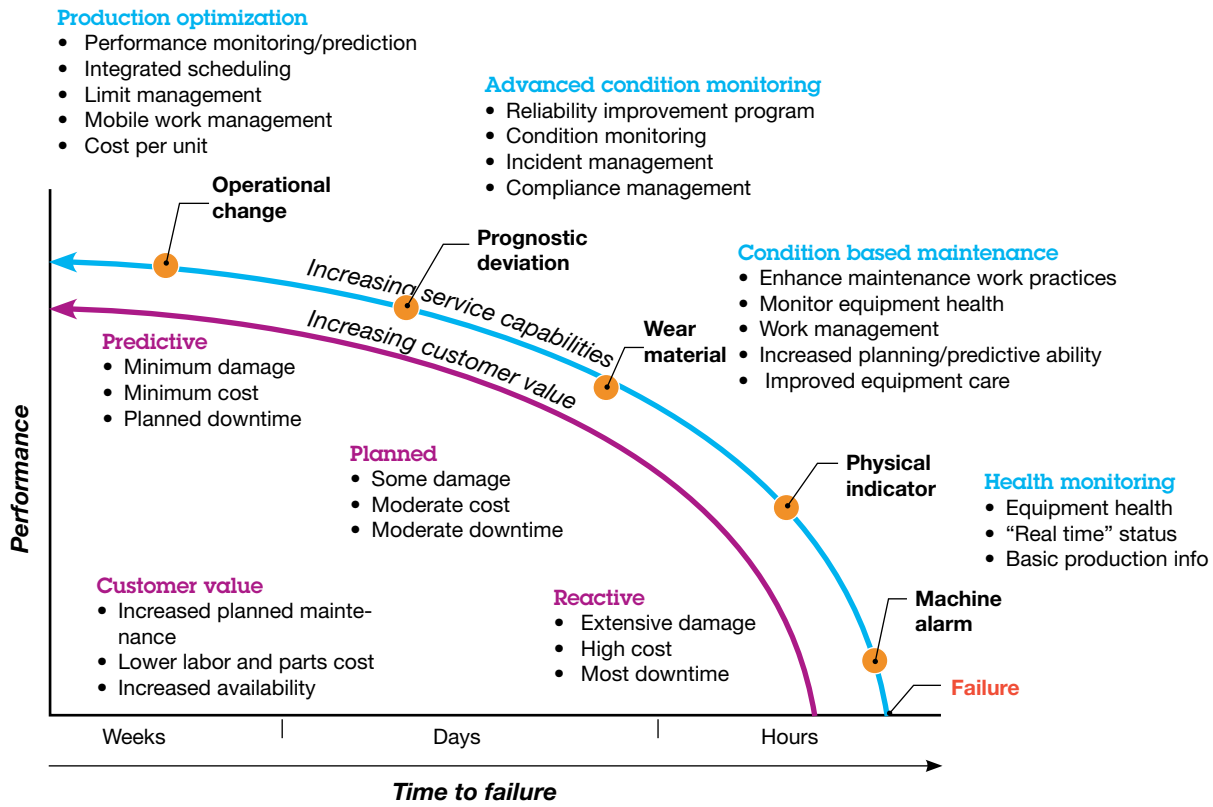

Figure 2: **An illustrative P-F curve.**

Applications of PAO

A typical industrial value chain has five stages or phases, namely:

1. Product development
2. Manufacturing
3. Distribution
4. Sales and marketing
5. After market or in service

Although the solution framework for PAO is well positioned to support and implement use cases across the value chain, we have seen the most common usage of PAO in the following areas:

- Predicting failure of critical equipment components
- Predicting which components are likely to fail together
- Optimizing recommendations to address the problem through continuous recalibration based on maintenance history
- Condition monitoring of rotating equipment in oil production platforms
- Manufacturing plant optimization
- Warranty and service prognostics
- Product quality monitoring
- In service fleet reliability analysis
- Unscheduled maintenance reduction

The PAO solution framework is capable of creating value across a wide variety of business use cases.

The analytics play

We recognize and acknowledge that to "listen to what machines are saying" it is critical that the data from the machines, along with other related information (for example, maintenance reports, warranty, and so on), needs to be gathered, transformed, cleansed and made available for any analytical introspection or processing to be initiated.

One of the key activities after the data is harnessed in an analytically usable form is to determine two things:

1. What type of analysis process is needed?
2. Which type of modeling is more appropriate?

Data mining focuses primarily on addressing these two questions. It is about automating the process of searching for patterns in the data.

It is imperative to follow a systematic process in an effort to maximize the efficiency and productivity of the task at hand.

Analysis process

We recommend a well-proven, industry standard data mining method called Cross-Industry Standard Process for Data Mining, which is more popularly known as CRISP-DM, its abbreviated form.

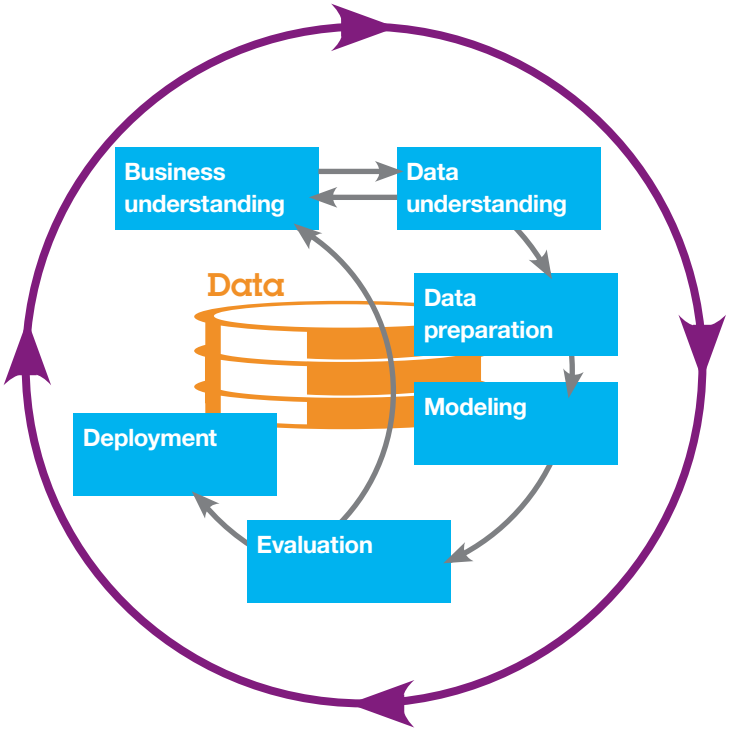

Figure 3: **CRISDP-DM methodology.**

CRISP-DM in its essence is a process model comprised of six phases in its lifecycle.

The six phases of CRISP-DM are the following:

1. ***Business understanding*** – Defining the project objectives and a thorough understanding of the requirements; helps in defining the data mining problem.
2. ***Data understanding*** – Collecting initial data, familiarizing and ensuring that proper data quality checks are applied.
3. ***Data preparation*** – Formalizing the data set (such as data tables, record set, attributes, and so on) and performing data cleansing and any required transformations.
4. ***Modeling*** – Identifying the specific modeling techniques to be applied and their customization or parameter calibration specific to the problem domain.
5. ***Evaluation*** – Analyzing and documenting the project outcome and assessing against the objectives.
6. ***Deployment*** – Deploying of the model(s) which are outcomes of the data mining process.

Further details about CRISP-DM are available in its official website (See ref [1]).

Predictive modeling

After the data is acquired, cleansed, transformed and provisioned in an analytics sandbox, the effort shifts to mathematical modeling to detect trends, patterns in the data and extract insight by leveraging stochastic and probabilistic modeling techniques.

Although each of the six steps is equally important, we see an emphasis on the data understanding and modeling aspects: the focus area of the latest and hottest IT role: the Data Scientist. We have also seen that most analytical projects start off with an effort to demonstrate prototypical value from employing analytical algorithms to predict future behavior.

Modeling and data mining focuses on machine learning techniques to develop predictive models capable of predicting some future events by analyzing historic data that characterize the underlying mechanism that generated the data. Machine learning can be broadly categorized into two major types: supervised and unsupervised.

PAO employs both supervised and unsupervised techniques to generate predictive models that characterize the phenomena hidden in the data.

An IBM solution approach

PAO is a framework to predict, analyze predictions, and suggest an optimized set of actions to mitigate the losses that may result from the predictions. At the core of PAO is advanced predictive and prescriptive analytics; its ability to detect trends and patterns in equipment or operator performances that may result in inefficiencies in operations and imminent production losses.

Predicting and optimally fixing equipment part issues before they fail, not only avoids costly unscheduled maintenance, but also avoids inefficient frequent scheduled maintenance.

The IBM solution approach is to build a framework that is based on a simple five step approach:

- Predictive analytics
- Standardizing prediction results
- Analyzing predictive outcomes
- Deterministic action identification
- Iterative resolution

Executing on the approach follows a prescriptive well defined sequence of analyzing data, predicting failures, analyzing possible root causes, and generating mitigating actions for it. This framework approach is the core implementation of IBM's PAO solution.

Figure 4 illustrates the schematic that captures the essence of IBM's PAO solution framework.

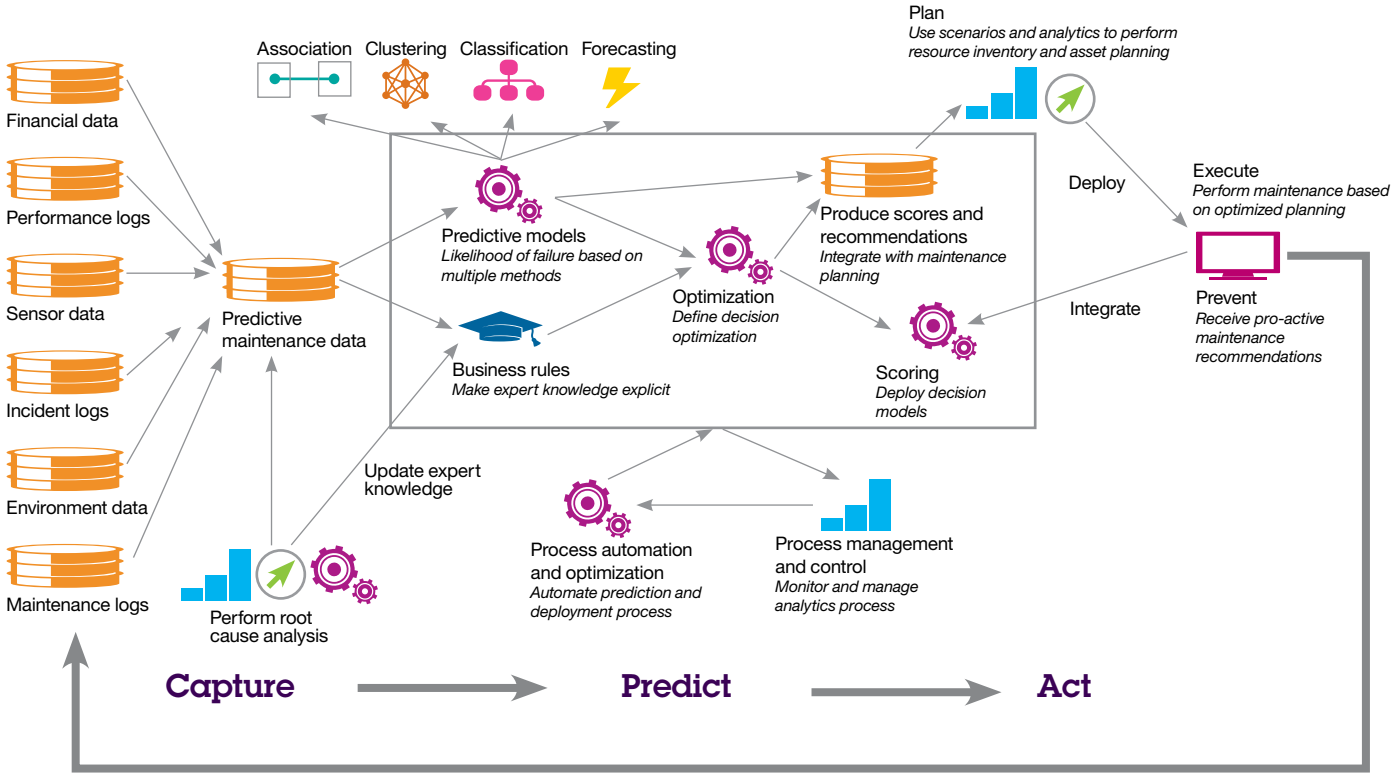

Figure 4: **IBM's PAO solution framework.**

IBM's PAO solution framework objective is to suggest actions based on current predictions. IBM's solution recognizes that newer, more risk-prone, scenarios may surface after the resolution of the first failure prediction. In recognition of the same, PAO employs the principle of optimization through an iterative increment in its prediction accuracies. In each iterative step the framework provides the capability to incorporate more data sets, additional analysis routines and newer business rule sets. This foundational principle enables IBM's PAO solution to provide convergent solutions: resolutions which completely and satisfactorily resolve the predicted issues.

We recognize that in the real world it is not realistic to assume that PAO can completely eliminate the human intuition and expertise factor especially in mission critical asset manufacturing and operations. Subject matter experts, field engineers, experienced operators may well assess the predictions from PAO and determine that not all of the predicted issues are critical enough to warrant attention right away. IBM's solution framework of PAO helps the predicted events to remain active until a proper maintenance window is identified for the recommendations to be implemented.

The rest of the section describes the key aspects of the IBM solution framework for PAO.

Predictive analytics – from multiple analytic engines

IBM's PAO solution recognizes that there are multiple analytics engines in the marketplace offered by different vendors who bring valued niche analytic capabilities. Leveraging the best of breed and specialized analytics engines increases the robustness and accuracy of predictions.

The IBM solution approach for PAO involves using the analytics engines in predicting, identifying and resolving issues. While any one particular analytic engine can predict a single class of failures, an overall approach requires PAO to identify and predict as many modes in which failures can occur. In using multiple analytical engines, the objective is to predict outcomes independently. Additionally, if multiple engines predict the same outcome, the confidence of such prediction increases, which in turn increases the recommended priority indicator to fix the issue. However, different analytics engines can also come up with exactly opposite predictions as well. In such circumstances, IBM's PAO solution analyzes the results in a different correlation, which will be addressed later.

Standardizing prediction results – through a common interface

The key to the success of such a multi-analytic-engine based approach is its ability to define and support a common interface through which an anomaly or a failure prediction is conveyed to the PAO solution framework. As an example, if one is implementing a clustering technique and can identify that data points that fall outside a cluster are a cause for concern, then this cause for concern needs to be translated to a standardized "prediction" such that it can then be compared with prediction outcomes from other analytics engines. IBM's PAO solution supports this flexibility by implementing adapter architecture: an analytics adapter for each of the analytics engines such that their output is presented in a standardized data format to the subsequent modules of PAO.

Analyzing outcomes – using hierarchical predictions

As identified earlier, different analytics engines can generate different failure predictions. And sometimes two different analysis engines can also provide similar prediction with varying confidence levels.

The IBM solution approach for PAO involves using the analytics engines in predicting, identifying and resolving issues.

IBM's PAO solution employs a simple but innovative technique to formulate how multiple predictions may be related and identifies which of these failure predictions are actually relevant enough to warrant an action to be taken. The IBM solution implements a categorization scheme that enables it to classify predictions hierarchically whose natural ordering are interpretive of their relative priorities through criticality of actions. The higher priority actions are more "actionable" than others.

The categorization scheme, to formulate hierarchies, may involve various algorithmic techniques. One such technique is to correlate the confidence scores of the predictions from different analytic engines and map them to a unique set of actionable predictions along with their confidence scores.

An implementation of IBM's PAO solution can plug and play with multiple analytic engines and thereby bring the best of breed analytical capabilities to realize the distinctive capabilities that organizations are seeking to harness.

Determining actions – using financial prudence

Each actionable prediction can be associated with one or more actions on one or more equipment component parts. Identifying appropriate actions requires a few things to occur simultaneously:

- The understanding of the association between equipment parts and their type of failures;
- The mode in which the failure is predicted;
- The engineering design of the component part;
- Component assemblies which are related to and interact with the component.

IBM's PAO solution determines and configures a mapping between the different components and the possible actions which may be performed on them. The solution also leverages an ontology based semantic model to represent the asset being monitored. However, it is commonplace that the products are highly customized when they are built. So, it is important that these relationships are identified based on the "as built" model of the product as opposed to its "as-designed" version.

Once such an actionable list is created, a financial judgement can be made for each actionable prediction by performing a cost-benefit analysis to determine the timeliness of acting on the predictions. This financial judgment is leveraged to prioritize what actions need to undertaken and when. As an example, some actions may warrant an immediate action while some others may be scheduled for a later time e.g. when the asset component undergoes its next planned scheduled maintenance.

It is important to note that IBM's PAO solution framework provides the foundation for organizations to consider adopting an optimized maintenance plan as opposed to their standard and often times conservative maintenance schedules.

Improving outcomes – with iterative feedback

The actions associated with the actionable predictions are typically performed based on its applicable time as identified in the previous step. After each of these actions is successfully performed, the expectation is that, subsequently thereafter, the failure predictions which caused the recommended actions in the first place, would no longer show up. However, other related failures may surface. It is therefore essential that the situations are evaluated repeatedly.

If the failure prediction that was assumed to have been solved shows up again, it is evident that the action that resulted from the prediction did not solve the problem. Such conditions imply that the above process of predicting, correlating predictions, identifying actionable predictions and determining associated actions needs to be performed iteratively such that the prediction accuracies keep increasing to the point that the recommended actions converge to a singularity. Singularity here implies the most precise recommended action with the highest confidence level to fix the problem; it is the essence of optimization.

IBM's PAO solution incorporates maintenance logs as an input source in each step of the execution chain of predicting, correlating predictions, identifying actionable predictions, and determining associated actions.

The future of PAO

The sustenance of PAO lies in its resiliency and ability to evolve in a modular fashion to incorporate newer and innovative capabilities.

The future outlook of IBM's PAO solution reflects its capability to harness the potential of text analytics and of augmented reality.

Integration to text analytics

A key area of improvement will occur when unstructured text analysis (including translation of text) is introduced as an analytics engine. This would not only allow the analytics engines to incorporate review comments and observations of technicians to better identify failures, but also to better identify actions associated with the predictions.

However, these engines need to understand the language of the technicians, because more often than not, depending on the industry, or even the enterprise for which the technicians work, the acronyms, choice of words, grammar and domain specific terminology usage have different semantic interpretations in the context of PAO.

The future value of PAO will be greatly enhanced by efficient text analytics features. Text analytics require well built dictionaries that capture the semantic relationships relevant to the enterprise. Building custom Dictionaries for every PAO implementation to capture these semantic interpretations can become cost prohibitive. Developing industry specific dictionaries and semantic models greatly reduce the cost, and increase the future value propositions for implementations of IBM's PAO solution.

Augmented task book

While describing identification of actions, it was apparent that the corrective actions performed on a single part can lead to the need for corrective actions to be performed on other related parts as well. Performing the corrective action can require the technician to perform multiple steps to reach the defective part to perform the corrective action. This could involve disassembly and/or reassembly of certain component parts.

When one has to deal with complex assemblies, especially those with extensive customization, visually aiding the technician to perform these tasks increases the probability of correctly performing the recommended action the first time. The key is to dynamically create the visual aid based on the recommended action. This is the basis of augmented task book, and will be an important feature for future implementations of IBM's PAO solution.

How to get started with PAO

In our experience we have seen a significant number of industrial enterprises that realize that they have a significant amount of data, especially machine generated, but they are not taking full advantage of the analytical insight possible.

It is imperative to align an organization around the key information which is required to maximize the business throughput of an enterprise.

IBM has developed an information agenda (IA), a proven industry-specific approach built from deep industry knowledge and applied experience gained through client experience. It is a proven formula for aligning information with the enterprise business objectives.

At the core, IBM's IA establishes a strategy for information so executives, decision makers and processes are supported with the best information to execute business strategy optimally.

The first facet of IBM's IA is a strategy that determines what information is needed to drive the key business priorities. The second area of focus is to develop a roadmap to identify the specific projects that will return the greatest value in line with business priorities and timescales. IBM leverages templates for over 16 industries that can be refined or customized to the client-specific needs.

IBM's IA also assists clients to define and govern their information supply chain that ensures information remains trusted and protected throughout its lifecycle. IA also develops a fit for purpose information infrastructure through which enabling technologies and services are assessed to ensure that the underlying platform will allow the realization of the desired transformation.

IBM's IA has demonstrated, with more than 600 engagements and 100 client references so far, to be one of the most successful initiatives which IBM has deployed lately.

To summarize, IBM's IA is a systematic approach to:

- ***Plan an information agenda*** – to align with the enterprise business strategy and plan for the future; its applicability to PAO is natural.
- ***Master the information*** – to ensure that the information is accurate and governed.
- ***Turn information into insight*** – to see, predict and shape business outcomes.
- ***Turn insight into action*** – to transform for breakaway results.

Summary

Machine parts fail; it is the nature of mechanical and electrical devices. The challenge for organizations is to plan, monitor, manage and mitigate such failures proactively enough to maximize the productive time of such costly equipment and increase its life expectancy. IBM's PAO solution can provide you with a unique and innovative approach through which the value of predictive analytics can be harnessed to solve your real world asset monitoring,maintenance and optimization problems by adopting a real-time, fact-driven, predict and act mode of business operation.

It is important to realize that the fundamental solution component building blocks of PAO are available from industry leading solution providers that ought to be assessed and potentially leveraged before home grown solutions are planned to be developed. Where a PAO solution in one industry will differ in another is the specific business context, the business use cases and the machine specific predictive models.

IBM's PAO solution is yet another example of IBM's pioneering and advocacy toward building a Smarter Planet.

Acknowledgment

The authors would like to acknowledge Gerald Kurtz in IBM Business Analytics and Optimization (BAO) for his leadership, sponsorship and continued support to focus on PAO as a key IBM solution and offering. We would like to thank Evelyn Rosengarten for her insights and feedback on the mathematical models.

Authors

Tilak Mitra is the Chief Architect of the Industrial sector in the Business Analytics & Optimization (BAO) group within IBM Global Business Services (GBS). Tilak can be reached at tmitra@us.ibm.com

Karthikeyan Seetharaman is a Solutions Architect in IBM's smarter cities team. Karthikeyan can be reached at karthik@us.ibm.com

Reference

[1] CRISP-DM Methodology – http://www.crisp-dm.org/

IBM Global Services
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
July 2013

Please Recycle

GBW03217-USEN-01