A DISCUSSION ON AUTOMATED DIAGNOSTICS

August 2022

Table of Contents

Introduction

Power Generation is undergoing many changes, including a digital transformation to improve or adapt to a dynamic energy landscape. It is time to leverage data and digital tools and realize measurable benefits from expanded data and analysis for many reasons. Common drivers include; perceived reduced maintenance cost, improved asset reliability, improved safety, etc. A digital transformation is a targeted journey that begins with existing data, software, and digital tools. For power generation, this includes existing sensors and instrumentation used in the plant, along with other inspections or measurements to estimate the health of a piece of equipment. One of the more powerful and near-term data uses is to diagnose a fault or degradation of a component or piece of equipment. Diagnostics is rapidly expanding but its foundation centers around expanded data sets aligned with algorithms and computations. For equipment health, diagnostics is viewed as an essential capability that leads to prognostics. Figure 1 is an abbreviated attempt to show the association between monitoring, diagnostics, and prognostics.

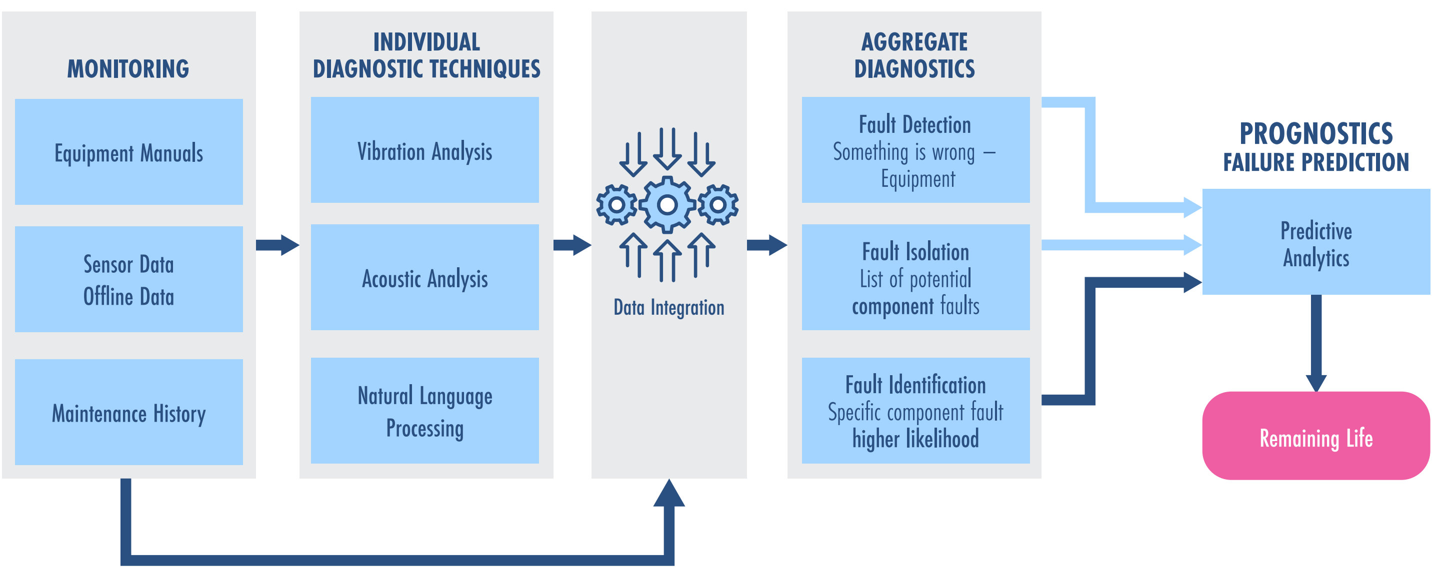


Figure 1. Association of Monitoring, Diagnostics, and Prognostics (abbreviated)

This paper aims to examine current diagnostic practices, R&D themes for automated diagnostics, and EPRI's role. The paper intends to create engagement and dialogue within the power generation industry and technology developers with the overarching goal of accelerating diagnostic technique development and application.

It would be prudent to define some of the nomenclatures before further discussion of the topic:

What is monitoring?

Monitoring can be thought of as surveillance tools and technologies to track an asset's behavior. Continuous online monitoring sensors, offline measurements, technical walk down, etc., can be considered surveillance tools to monitor power plant assets. In the context of diagnostics, additional data sources such as drawings, maintenance reports, and asset manuals are also considered monitored data sources.

What is a fault?

The fault of the power plant asset can be thought of as a "condition or state of an asset that is a precursor to the failure" (1). Failure of an asset can be considered a partially or fully disabled condition or state of an asset to perform its intended function.

What is fault diagnostic?

Diagnostics can be thought of as assessing an asset's health, condition, or state, based on evidence gathered through monitoring and analysis of monitoring data—the fault diagnostic aims to identify the fault(s) before an asset fails to perform its intended function.

Diagnostics of a power plant asset is not a new concept; utility subject matter experts (SMEs) routinely gather information from online and offline sources to estimate equipment anomalies. However, the current process is manual and sometimes takes days to assess data manually.

What is automated fault diagnostic?

Automated or continuous online diagnostics can help drastically reduce the time required for data analysis and improve safety, reliability, and cost-effective maintenance. Automated diagnostics can be thought of as assessing the health of an asset using automated, computerized, or robotic techniques and tools based on evidence gathered through continuous online monitoring means. Automated diagnostics can provide correct information at the right time to the right personnel for making the right decisions that can improve safety and cost-effective maintenance and reliability strategies.

As shown in Figure 2, automated diagnostics can be divided into individual and aggregate diagnostic techniques.

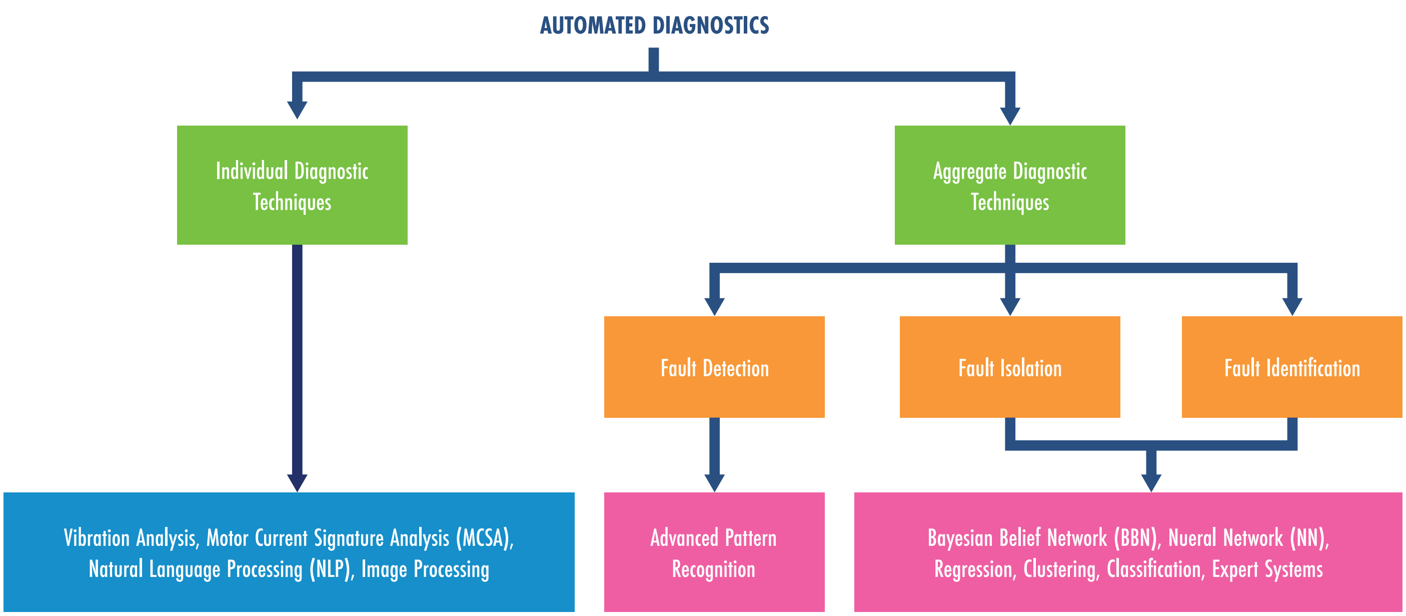


Figure 2. Automated Diagnostic types

Individual diagnostic techniques are signal processing or analysis techniques that use dynamic raw data to calculate key performance indicators (KPIs), features, symptoms, or virtual sensing parameters. Examples of conventional individual diagnostic techniques can include vibration analysis, motor current signature analysis (MCSA), partial discharge analysis (PDA), acoustic analysis, etc.; some of the newer individual diagnostic techniques can include image processing, natural language processing (NLP), mathematical models or equations for virtual sensing, etc. Some of the faults can be detected by only one individual technique. However, any equipment or system SME would generally agree that it is never just one technique that can provide diagnostics with high confidence.

The aggregate diagnostic technique can take into account specific KPIs, symptoms, or features from individual diagnostic techniques and particular static signals (examples include; temperature, pressure, flow, and so on) for detecting, isolating, and identifying specific faults with higher confidence.

As mentioned in (2), aggregate fault diagnostics can be in three forms based on available tools and techniques. It can be:

- Fault detection which is detecting abnormal behavior of the system or equipment;
- Fault isolation can be thought of as determining a component of the system or equipment experiencing a faulty condition.
- Fault identification can be thought of as estimating the type of the fault associated with the component and an estimated severity of that fault.

According to the I4Gen® advanced M&D maturity model in (3), the highest maturity of the M&D center or centralized analysis center is developing and deploying prognostics models for power plant assets. Developing a prognostics model based on specific goals can help further improve safety, reliability, and maintenance cost. The fault detection, isolation, and identification diagnostic techniques have uncertainties. These uncertainties can propagate in the uncertainty calculation of prognostics. As shown in Figure 1, fault detection and fault isolation can provide some basis for some versions of prognostics. However, fault identification with estimated fault and severity can provide grounds for higher confidence in prognostics calculations. One of the goals of automated diagnostics is to reach a fault identification state. Some of the possible benefits of automated diagnostics are listed in Figure 3.

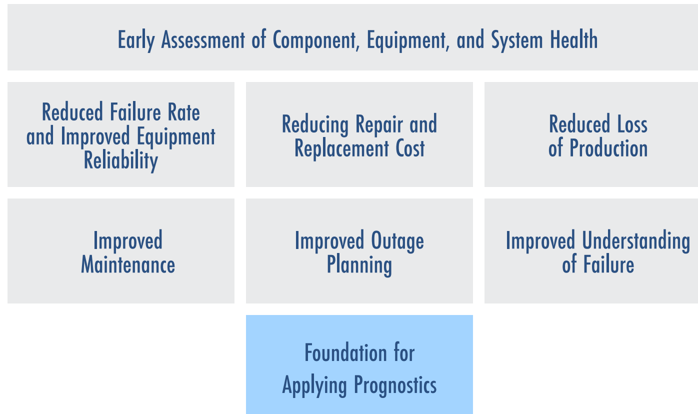


Figure 3. Possible benefits of automated diagnostics

This paper focuses on aggregate automated diagnostics and will not discuss prognostics in detail.

Two other factors are significant in setting up the infrastructure required for aggregate automated diagnostics – cyber security and data architecture. Cyber security plays a major role in any digital transformation project. It is essential to include cyber security and IT team members at the initiating stage of any digital transformation project.

Much of the static data and calculated KPIs, features, or symptoms live in siloed databases at the power plant facility. This architecture arrangement hinders the implementation of effective aggregate fault diagnostic techniques (for example, data analytics software). The diagnostic algorithm must have access to all these data sources to work successfully. Data integration platforms can help achieve this goal. In addition, it will also provide the opportunity to clean the data. Some basic information regarding data integration platforms is available in (4) and (5).

Current Diagnostic Practices

The following two main types can characterize current diagnostic practices in the utility:

1. Manual Preventive Maintenace (PM) Task-based and
2. Hybrid with the support of centralized monitoring and diagnostic (M&D) center

There are multiple varying practices in each category; however, generalized types are mentioned in this discussion.

Manual Preventive Maintenance (PM) Task-based Diagnostic

Utilities without data historians and a centralized M&D center or Centralized Analysis Center perform a diagnostic task using information gathered from periodic manual PM tasks. EPRI's Preventive Maintenance Basis Database (PMBD) (7) provides information regarding various PM tasks for many of the power plant assests. Figure 4 shows an example of such diagnostic practice. This figure shows some manual PM tasks performed on a medium voltage induction motor. Depending on utility preference, these PM tasks are performed by plant personnel, and in other cases, it is outsourced to a conditioned-based maintenance (CBM) or predictive maintenance (PdM) vendor. Several maintenance technology SMEs may carry out these manual PM tasks; for example, SME for vibration analysis might differ from SME for thermography. During these periodic manual PM tasks, if there are measurements out of normal bounds, the SME responsible for that asset considers related measurements from other PM tasks and SME knowledge and experience with the asset to estimate potential fault or degradation.

Diagnostic with Centralized M&D Center Support

Many utilities have deployed centralized M&D centers in the last two decades. Figure 5 and 6 show generalized diagnostic approaches with a centralized M&D center. Figure 5 shows a method where utilities with centralized M&D centers perform diagnostic; based on existing sensors (for example, temperature, pressure, flow) and significantly depend on manual PM tasks. In another case, as shown in Figure 6, some of the utilities have started deploying continuous online monitoring (COLM) sensors based on EPRI developed COLM quick guides (QGs) available in volume 7 of (6) or failure mode effect analysis (FMEA). This installation of additional sensors provides substantial time savings for the SMEs from a reduced number of PM tasks. The added sensors can provide more visibility into the asset's health.

Figure 5 and Figure 6 show generalized data transfer mechanisms. In the case of Figure 5, existing sensor data are stored at the plant historian. In the case of Figure 6, static data and KPIs, symptoms, or features calculated from additional COLM sensors' raw data are sent to plant historians. Advanced Pattern Recognition (APR) software compares data available in plant historian with model data. During

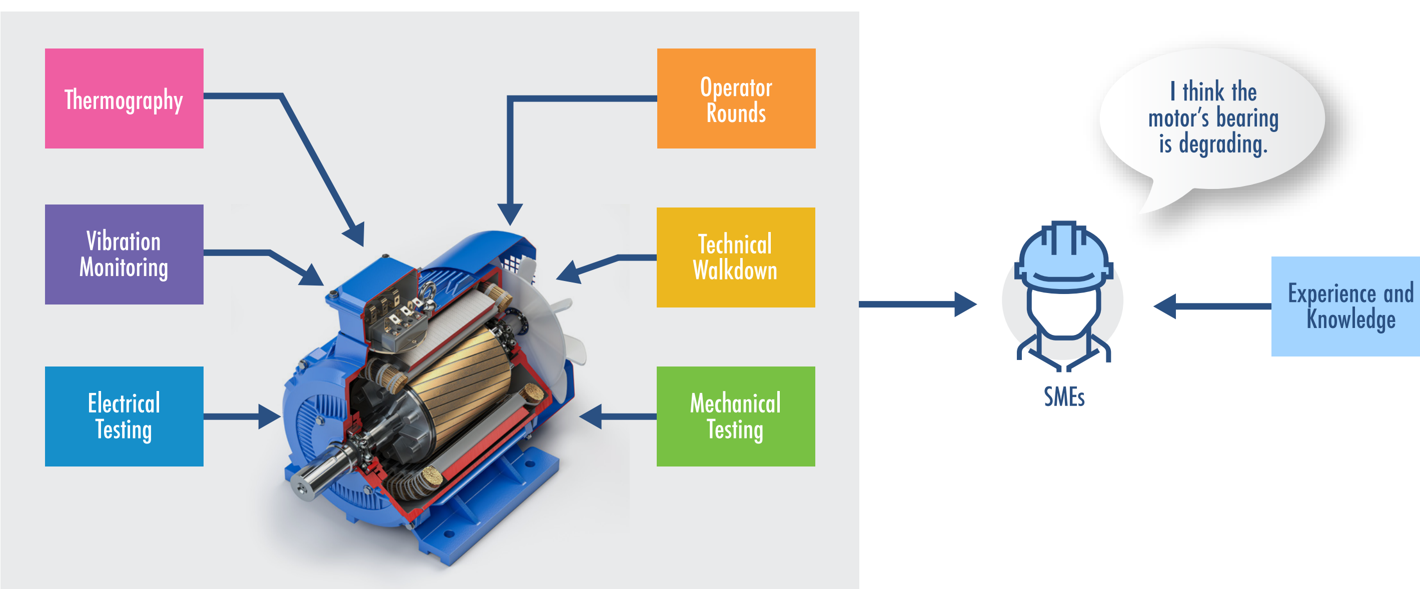


Figure 4. Generalized practice for PM task-based diagnostic

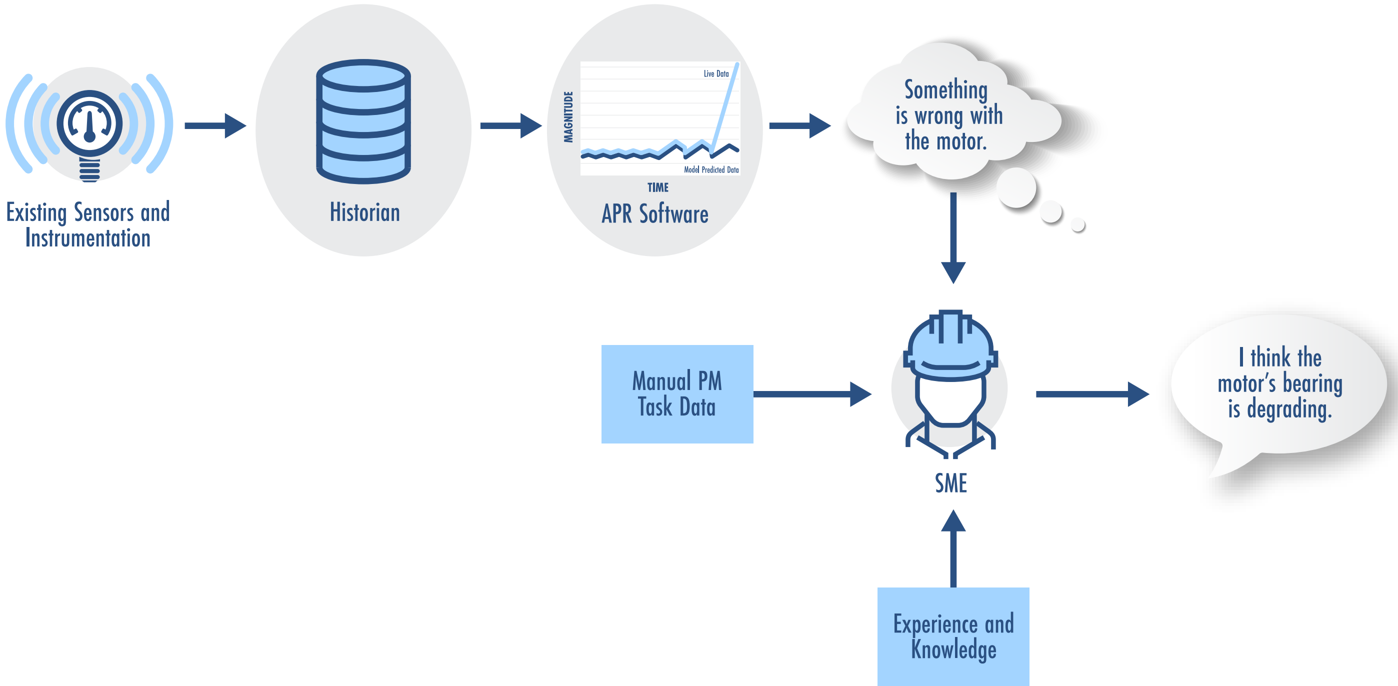


Figure 5. Generalized practice for diagnostic using existing sensor/instrumentation and centralized M&D center support

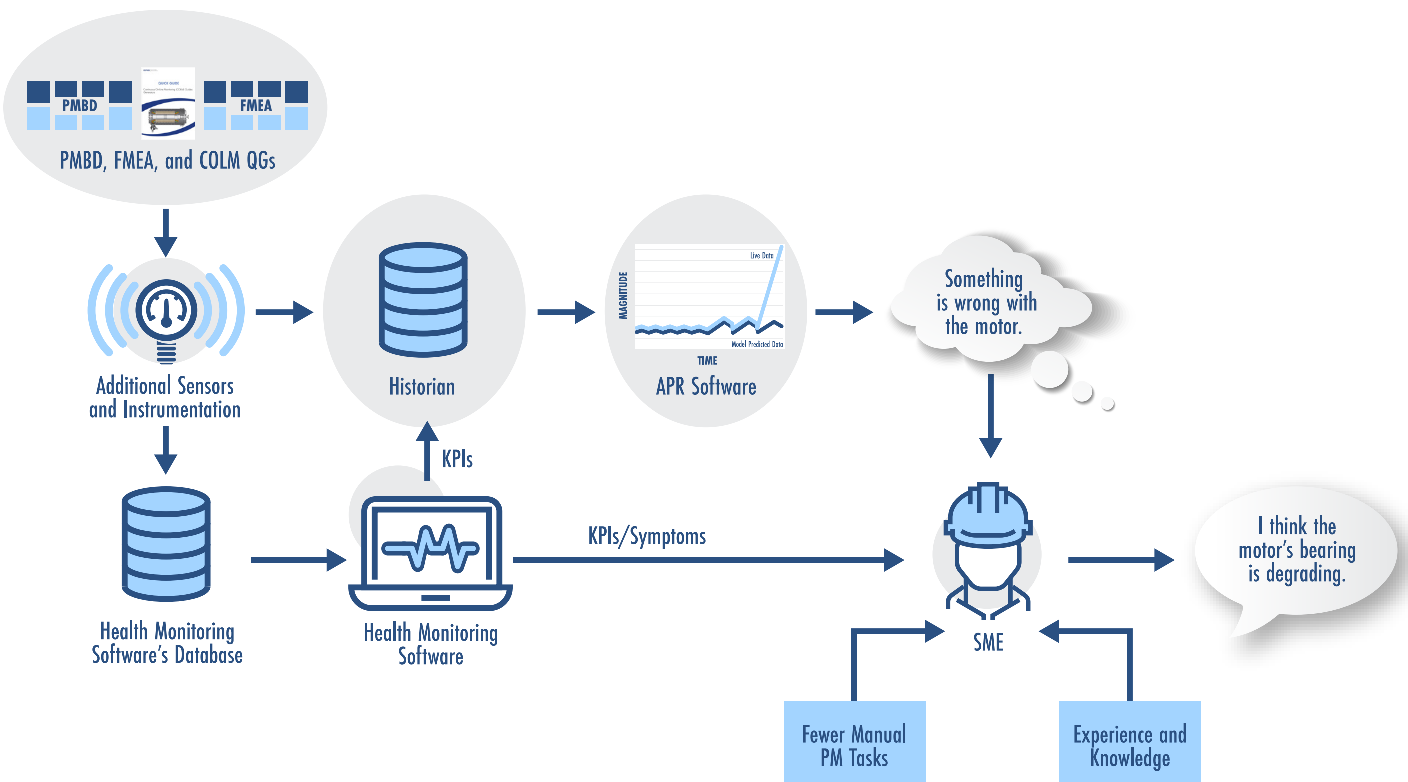


Figure 6. Generalized practice for diagnostics using additional sensors/instrumentation and centralized M&D center support

the APR software implementation phase, plant personnel creates models based on past data, including seasonal data. During plant operation APR software tracks whether live plant data matches the model data, as shown in Figure 7. An alarm is sent to centralized M&D personnel if the difference between live plant data and model data exceeds a preset limit or threshold. APR software may indicate something wrong with the equipment; however, it can rarely identify the specific component anomaly. M&D center personnel can request plant preventive or conditioned-based maintenance personnel to investigate for further information. SMEs determine potential faults based on all manual and continuous online data from the centralized M&D center. The Figure 5 method has a higher dependence on manual PM tasks than the method shown in Figure 6. Manual PM tasks are time-based and can cause delays in assessing the condition of a power plant asset. The automated diagnostics with permanently installed sensors providing continuous measurements can significantly reduce this delay.

Some utilities have deployed in-house centralized M&D centers, and others have outsourced centralized M&D centers to external organizations. Information regarding a centralized M&D center business case, setup, staffing requirement, and related supporting topics are discussed in volumes 2 to 5 of (6).

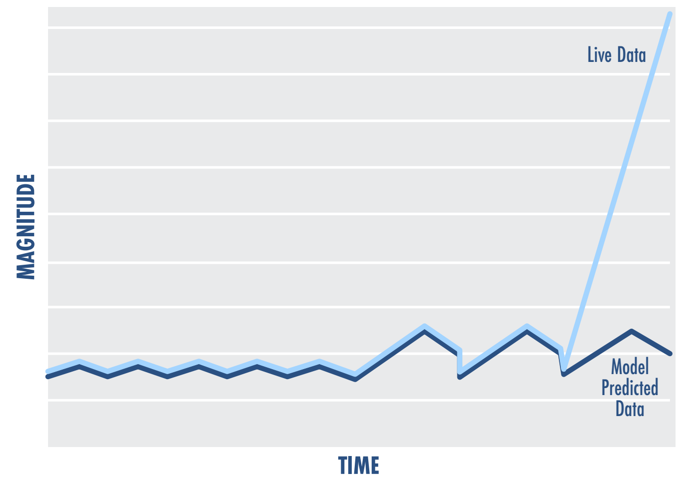


Figure 7. Example of APR comparison of live data and model predicted data

The Targeted Future State of Automated Diagnostics

Many utilities are exploring or piloting various advanced diagnostic algorithms for fault isolation and identification. Many technology providers provide statistical, physics-based, hybrid, rule-based, or neural network type advanced algorithms or engines for automated diagnostic. Some of APR software vendors have also updated APR software to include a diagnostics module. Figure 8 shows a gener-

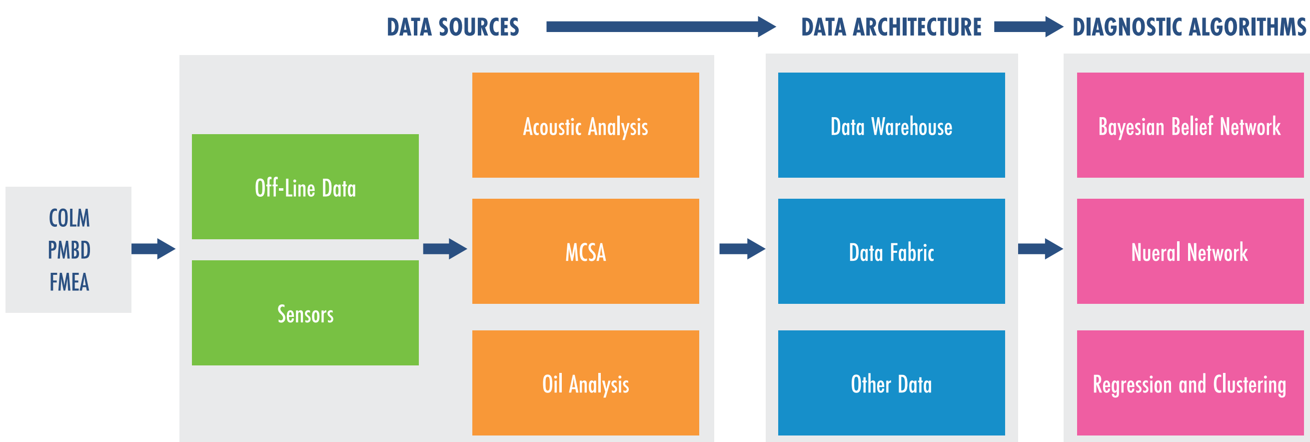


Figure 8. Generalized data flow to the diagnostic engines

alized data flow to the diagnostic algorithms. Various diagnostic algorithms have their advantages and disadvantages depending on their application. These diagnostic algorithms may be perfect for a diagnostic application. However, there are some implementation challenges associated with each of them. The following section discusses the elementary information regarding Bayesian Belief Network (BBN), clustering and classification type of algorithm, and associated implementation challenges.

Bayesian Belief Network

A Bayesian Belief Network (BBN) is a probabilistic causal network. Figure 9 shows an example of a simple BBN, a directed acyclic graph for a diagnostic approach. There may be multiple faults associated with exact causes, effects, and related symptoms in a realistic environment. The boxes next to each occurrence are either individual or conditional probabilities based on prior experience. BBN uses these probabilities and current observations to diagnose the fault.

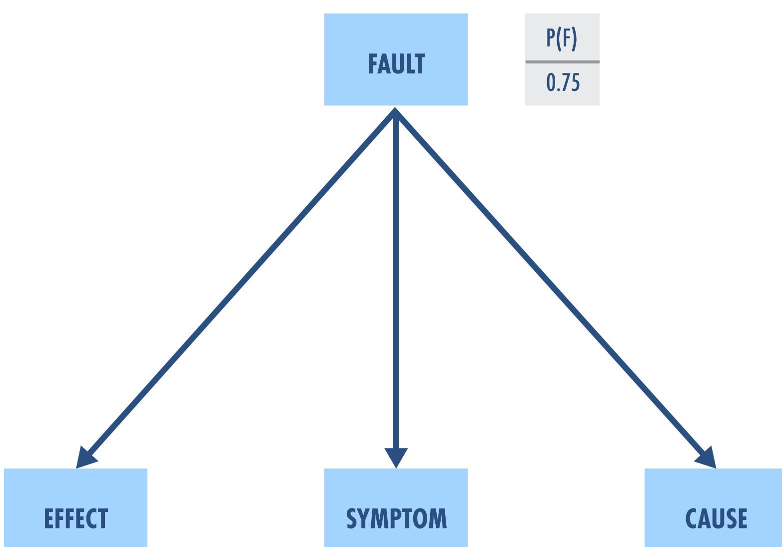
A couple of fundamental challenges of setting up a BBN-directed acyclic graph is knowing each fault and its causes, effects, and related symptoms. Another challenge is to possess the required knowledge or data for assigning individual and conditional probabilities.

Clustering and Classification

Figure 10 shows an example output of clustering and classification type of algorithms. Examples of clustering algorithms include K-means, Hierarchical, Normal mixture, etc. These algorithms use unsupervised learning. Examples of classification include logistic regression, Neural network, decision tree, etc. In either of these algorithms' outputs, it is easy to see that one data set looks different from another. However, it is difficult to determine the reason(s) why. Is it because of other operational conditions, seasonal changes, or a genuine asset fault? Based on past site data, algorithms can be trained on seasonal or operational changes. Fault diagnostic, however, will still require knowledge or data of the associated causes, effects, and related symptoms.

What are the Gaps?

As utilities and technology providers are moving beyond the use of APR software for diagnostic, the use of diagnostic approaches with data analytics (DA), machine learning (ML), and artificial intelligence (AI) algorithms similar to the above examples will increase. An example data flow for an automated diagnostic with DA, ML, or AI algorithm is shown in Figure 11. It is desired that the output of DA, ML, or AI algorithms be able to inform plant personnel of potential fault or degradation. However, as seen in the above examples, commercially available algorithms can recognize when one data set differs from another. Still, it can not determine the fault or degradation without prior knowledge of faults, causes, and related symptoms.



P(F)
0.75

E	S	C	P(F1 | E, S, C)
T	T	T	0.95
T	F	F	0.95
F	T	F	0.95
F	F	T	0.03

Figure 9. Example of a Simple Bayesian Belief Network for a diagnostic approach

In conclusion, any DA, ML, or AI type of algorithm will require knowledge, either from data or SMEs, as shown in Figure 12. As shown in this figure, it is hypothesized that at least during the first few years of implementation, the output of these algorithms may need input and verification from SMEs. Please note that this SME knowledge-based or past data-driven fault database requirement is not unique to the advanced algorithms discussed in this paper; it will be a requirement for any advanced algorithms.

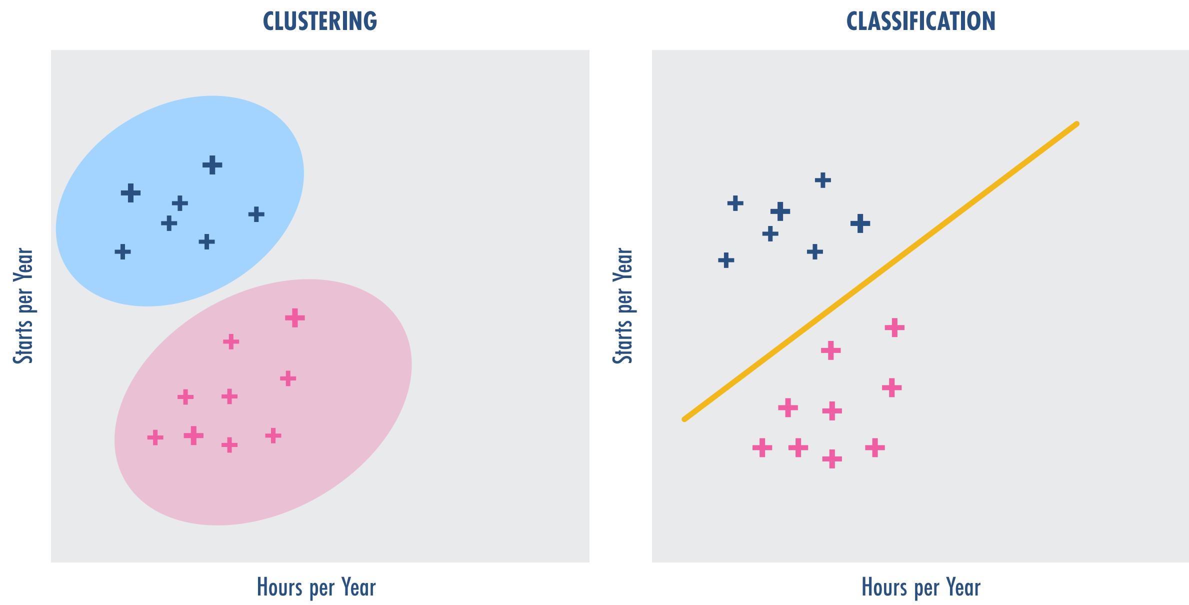


Figure 10. Example output of clustering and classification type of algorithms

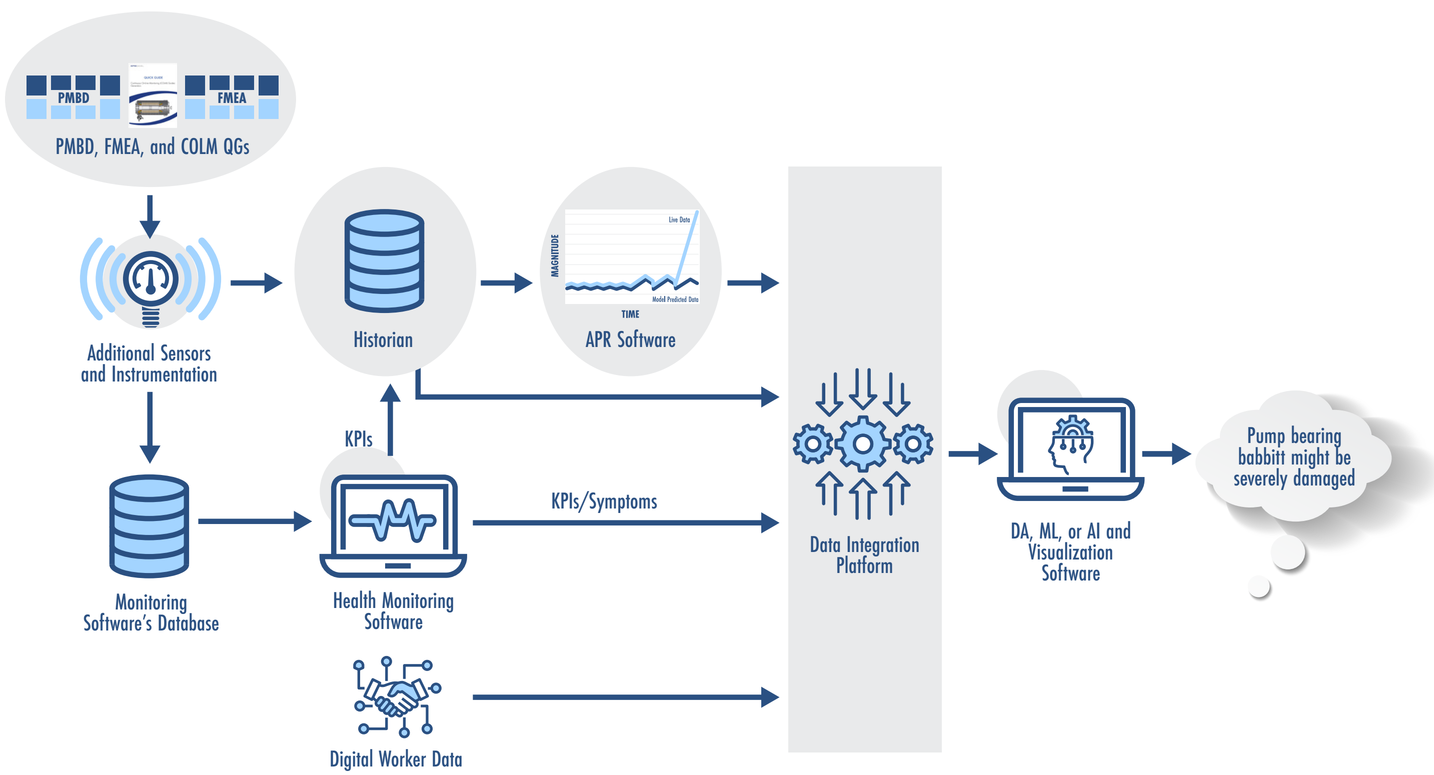


Figure 11. Example data flow for an automated diagnostic using DA, ML, or AI algorithms

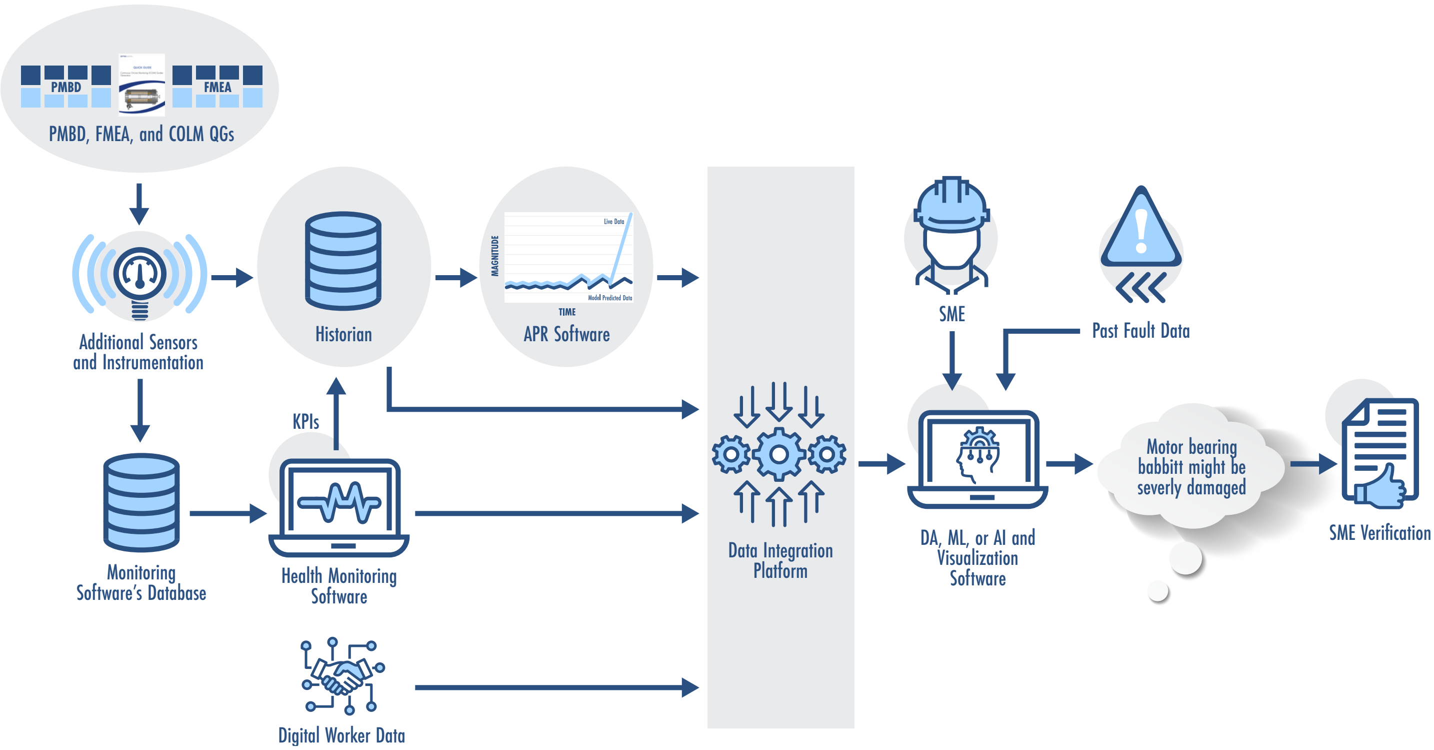


Figure 12. Example data flow for DA, ML, or AI-based automated diagnostic using SME knowledge and past fault data

EPRI's Role

A comprehensive fault database will need to be created for power-generating utilities to use these new advanced algorithms. Envisioned role of EPRI:

1. EPRI is uniquely positioned to collaboratively lead the development of fault databases in collaboration with utility members, technology providers, original equipment manufacturers (OEMs), national laboratories, and academia.
2. In collaboration with utility members and diagnostic algorithm technology providers, participate in fault signature database implementation and demonstration.

EPRI has developed a vast knowledge base in the form of EPRI deliverables that can accelerate the development of these types of fault databases.

As shown in Figure 13, EPRI envisions diagnostic technology provider agnostic fault signature databases:

1. Knowledge-based fault signatures
2. Data-driven fault signatures
   a. Plant data-based fault signatures
   b. Simulated or Modeling based fault signatures
   c. Emulated or laboratory experiment generated fault signatures

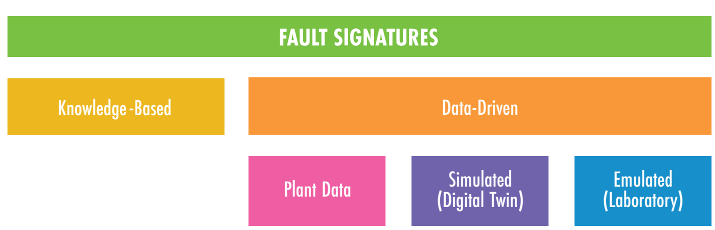


Figure 13. EPRI envisioned types of fault signature databases

Knowledge-Based Fault Signature Database

Knowledge-based fault signature database is structured fault information gathered through SME knowledge and experience with the equipment design, operation, maintenance, and controls. This fault information may include fault location, degradation mechanism, degradation influence, etc.

Over the years, EPRI has developed a Preventive Maintenance Basis Database (PMBD) (7) to provide maintenance guidance for various power plant assets. This database has a vast amount of fault information based on SME knowledge and experience. The PMBD has over 280 power plant component templates, including cables, gearbox, motor, transformer, turbine, and generator. EPRI also develops power plant component-specific continuous online monitoring (COLM) quick guides (QGs). These COLM QGs include information regarding sensors that can help replace or extend the interval for some of the periodic manual PM tasks. There are over 30 COLM QGs, and that number is increasing every year. Similar to the PMBD, COLM QG information is also gathered through SME knowledge and experience with various COLM sensors and individual diagnostic techniques.

EPRI products, PMBD, and COLM QGs provide the basis for monitoring power plant assets. Specifically, the PMBD provides information regarding failure modes associated with the power plant assets, and the COLM QGs provide information regarding value sensors that can detect those failure modes.

Capturing EPRI and utility SMEs' extensive experience and knowledge with power plant asset diagnostics into a structured format can help implement an automated diagnostic algorithm. After PMBD and COLM QGs, the next natural step for EPRI is to develop knowledge-based fault signatures, thus expanding the available information in PMBD and COLM QGs.

Data-Driven Fault Signature Database

Data-driven signatures are structured information derived from the time-series sensor data and processed raw data with the support of maintenance records, OEM manuals, and drawings. Data-driven fault signatures can be further divided into three categories –

a. Plant data-based fault signatures

b. Simulated or Modeling based fault signatures

c. Emulated or laboratory experiment generated fault signatures

Plant Data-based Fault Signature

Power plants are currently performing manual diagnostics with the support of a centralized M&D center and periodic manual PM tasks. Over the years, power plants may have gathered some of the fault data using online and offline monitoring techniques. Power plants are also installing new sensors with the help of EPRI COLM QGs or failure mode effect analysis (FMEA). All dynamic and static data from sensors (COLM, route-based, and offline), maintenance records associated with fault data, drawings, as well as centralized M&D center observation records can help develop data-driven fault signatures from plant data.

EPRI plans to request this type of fault data from participating utilities to develop data-driven fault signatures. Gathering plant data from multiple utilities provide the required variations for data-driven diagnostic approaches as sensor installations, PM practices, maintenance practices, ambient conditions, and operation modes will differ. Plant data can also support benchmarking and fundamental statistical analysis, for example, the mean time between failure (MTBF) for a component, independent and the conditional probability of failure of a component within one system or equipment, etc.

Simulated or Modeling-based Fault Signatures

Researchers have developed mathematical models based on physics and available data to simulate physical systems or processes. Modifying parameters within these models can mimic specific faulty conditions. Acquiring data using this type of simulation can help with the data-driven fault signature database.

EPRI has been working on developing digital twins for many of the critical assets. These mathematical models can support efforts to build a data-driven fault signature database. Academia can also help with the mathematical model development to create fault signatures.

Developing, testing, and verifying such mathematical models may require significant resources. EPRI envisions the development of such models based on the value proposition and member needs.

Emulated Fault Signatures

Emulation of fault on a physical asset in a controlled laboratory environment can also help generate fault data. On a case-by-case basis, EPRI envisions collaborating with various laboratories and universities to support the development of data-driven fault signatures.

Conclusion

Automating the analytics for asset fault isolation and identification has been an increasing industry need. Analytics algorithms' underlying mathematics was developed a long time ago. However, there is a radical change in processing power and data storage; for example, current smartphone processing power and memory are considerably better than an average laptop or desktop a decade ago. This drastic change in processing and data storage has made it possible for years-old magnificent mathematics to analyze vast amounts of data in near real-time. Some analytical algorithms based on mathematics are excellent at disaggregating one data set from another. However, it can not explain why one data set looks different from another. These algorithms will require help from power plant SME knowledge or similar past data to answer that question for power plant assets. A fault signature database can help structure SME knowledge and past fault data that can be used with analytics algorithms.

It can be agonizing for an individual utility to develop a fault signature database. EPRI is uniquely positioned to lead collaboration with utility members, technology providers, OEMs, academia, and national laboratories to build knowledge-based and data-driven fault signature databases. As a result of this collaboration, it can be possible to create a comprehensive database.

EPRI has developed PMBD and COLM based on industry SME knowledge and experience. Knowledge-based fault signatures are the expansion of information generated in PMBD and COLM. Consequently, knowledge-based signature database development may not take as long as data-driven fault signature. However, data-driven signatures can provide better insight into the faults.

EPRI envisions the development of knowledge-based and data-driven fault signature databases. Periodic reviews of the analytic algorithm landscape and its requirements will guide the path forward for these databases.

Frequently asked Questions

1. How will utilities without a centralized M&D center utilize these fault signatures?
   a. EPRI is planning to publish knowledge-based fault signatures in an Excel workbook format. Utilities without a centralized M&D center can manually use filtering functions of these workbooks to isolate faults based on observed symptoms.
2. Will these fault signatures include threshold values (unit and magnitude) for each fault signature symptom?
   a. No, threshold values for a symptom can be unique to each plant; however, the change in direction is not. Consequently, knowledge-based fault signatures will include only the direction of magnitude change in the observed KPI, symptom or feature. Data-driven fault signatures may be able to provide threshold values. However, that remains a hypothesis at this time. This hypothesis can be tested as more data-driven fault signatures are gathered from various utilities.
3. Will these fault signatures help implement fully automated diagnostic in a power plant environment?
   a. It will be closer to fully automatic diagnostics but not a comprehensively fully automated process. As seen in figure 1, currently, there are no sensors or virtual sensors for every fault of power plant assets. A type of inspection method still monitors some of the faults. Consequently, there remains a gap in achieving a fully automated diagnostic state. This gap also provides the opportunity to invest in value-based research for sensors and individual diagnostic techniques.
4. Will these signatures be accurate for my plant?
   a. EPRI envisions the development of fault signature databases as accurately as possible. However, it won't be perfect for all plant. Improving and editing these databases will be a collaborative effort between participating members and EPRI.

      From time to time, end-users and SMEs will find some errors, including missed symptoms, missed faults, and false diagnoses. It is envisioned that as utilities find these misses, those misses will be reported back to EPRI for review, verification, and potentially update the EPRI fault signature databases for broader use.

      It will be faster to make these databases accurate with more users of these databases.
5. If I can access the format for developing knowledge-based fault signatures, wouldn't it be better if I create fault signatures for my plant?

a. EPRI is planning to publish the process of developing knowledge-based fault signatures. This future document can help the individual utility create their own fault signature database.

However, the development of fault signatures for power plant assets will require tremendous personnel time and funding resources. In addition, this individually developed fault signature database might not be as comprehensive as fault signatures collected from multiple sources.

EPRI plans to lead the collaboration with the utility members, technology providers, OEMs, national laboratories, and academia to develop an extensive database.

6. Will the EPRI fault signature database include all possible fault signatures?
   a. EPRI envisions developing a comprehensive fault signature database. However, some unknown faults might still be missing from this database. As the industry experiences these unknown faults, EPRI envisions including these faults in the database.
7. Will these signatures be compatible with any technology providers' diagnostic software offering? Or will this be plug-and-play?
   a. It is envisioned that a considerable part of the generalized or agnostic fault signature developed by EPRI will be suitable for any technology provider's algorithm. However, each technology provider's algorithm will have some unique information requirements that may be missing from the generalized fault signature. EPRI can help fill this gap and add some of the missing information unique to the algorithm of utility choice.

      In addition, including fault signature database information in the technology provider's diagnostic software will be a manual process for the time being. As these databases become popular enough, in collaboration with a diagnostic software provider, an API or other mechanisms can be developed to automate the inclusion of fault signature database information in a technology provider's diagnostic software.
8. How can EPRI help?
   a. It is envisioned that EPRI will develop technology providers diagnostic software or algorithm agnostic fault signature database that can help utility members achieve an automated diagnostic state.

      EPRI can provide the following support for the implementation and use of a diagnostic algorithm:
      - Help identify and include specific fault signature fields required for the technology provider's algorithm that may be missing from the agnostic fault signature.
      - Help determine unique threshold values for each fault signature's symptoms specific to your plant.
9. Will EPRI collaborate with diagnostic technology providers?
   a. Collaborating with various stakeholders is a key to the success of this project. EPRI envisions demonstration projects with utility and diagnostic software providers. In addition, OEM participation is welcomed for fault signature development.

Reference

1. *Guide to Develop Asset Fault Signature Using Fleetwide Prognostic and Health Management Suite Software.* EPRI, Palo Alto, CA: 2017. 3002011218.
2. *Remaining Useful Life Approaches for Rotating Equipment.* EPRI, Palo Alto, CA: 2019. 3002015411.
3. *I4Gen® Integration of Intelligent Generation White Paper: Maturity Level for Key Application Areas within EPRI's Intelligent Plant Concept.* EPRI, Palo Alto, CA:2018. 3002013663.
4. *Introduction to Data Integration for Power Generation.* EPRI, Palo Alto, CA:2021. 3003022113.
5. *I4Gen® Intelligent Generation: Fundamentals of a Digital and Analytics Platform for Data Integration, Analysis, and Actionable Information.* EPRI, Palo Alto, CA: 2018. 3002013664.
6. *Continuous Online Monitoring Guidebook.* EPRI, Palo Alto, CA: 2021. 3002019924.
7. https://pmbd.epri.com.

EPRI RESOURCES

Prayag Parikh, *Senior Technical Leader*
704.595.2502, pparikh@epri.com

Enabling Technologies and Digital Transformation

About EPRI

Founded in 1972, EPRI is the world's preeminent independent, non-profit energy research and development organization, with offices around the world. EPRI's trusted experts collaborate with more than 450 companies in 45 countries, driving innovation to ensure the public has clean, safe, reliable, affordable, and equitable access to electricity across the globe. Together, we are shaping the future of energy.

3002025266 September 2022

EPRI
3420 Hillview Avenue, Palo Alto, California 94304-1338 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com