# PROVIDING PLANT DATA ANALYTICS THROUGH A SEAMLESS DIGITAL ENVIRONMENT

**Aaron Bly and Johanna Oxstrand**
Idaho National Laboratory
2525 Fremont Ave, Idaho Falls, ID 83415
Aaron.Bly@inl.gov; Johanna.Oxstrand@inl.gov

## ABSTRACT

As technology continues to evolve and become more integrated into a worker's daily routine in the commercial nuclear power industry the need for easy access to data becomes a priority. Not only does the need for data increase but the amount of data collected increases. In most cases the data is collected and stored in various software applications, many of which are legacy systems, which do not offer any other option to access the data except through the application's user interface. Furthermore the data get grouped in "silos" according to work function and not necessarily by subject. Hence, in order to access all the information needed for a particular task or analysis one may have to access multiple applications to gather all the data needed.

The industry and the research community have identified the need for a digital architecture and more importantly the need for a Seamless Digital Environment (SDE). An SDE provides a means to access multiple applications, gather the data points needed, conduct the analysis requested, and present the result to the user with minimal or no effort by the user.

The study researches the potential approaches to building an analytics solution for equipment reliability, on a small scale, focusing on either a single piece of equipment or a single system. The analytics solution will consist of a data integration layer, predictive and machine learning layer and the user interface layer that will display the output of the analysis in a straight forward, easy to consume manner.

*Key Words*: seamless digital environment, data analytics, use cases

## 1 INTRODUCTION

As technology continues to evolve and become more integrated into a worker's daily routine in the Nuclear Power industry the need for easy access to data becomes a priority. Not only does the need for data increase but the amount of data collected increases. In most cases the data is collected and stored in various software applications, many of which are legacy systems, which do not offer any other option to access the data except through the application's user interface. Furthermore the data gets grouped in "silos" according to work function and not necessarily by subject. Hence, in order to access all the information needed for a particular task or analysis one may have to access multiple applications to gather all the data needed.

The U.S Department of Energy Light Water Reactor Sustainability (LWRS) Program initiated research in to what is needed in order to provide a model for nuclear power plants to reference when building an architecture that can support the growing data supply and demand flowing through their networks. The LWRS Digital Architecture for an Automated Plant effort published the report Digital Architecture Planning Model [1], which describes items to consider when designing an architecture intended to support the increasing needs and demands of data throughout the plant.

A well-designed architecture will be able to support the data demands. However, in order to ensure the data is adequately utilized to improve and support the plant operations there also needs to be an easy, quick and reliable method to access the data. A common method is to create a “one stop shop” application that a user can go to get all the data they need. A key to this approach is a method to integrate the data stored in multiple applications (e.g., work management system and plant information databases). In other words, there is a need for a Seamless Digital Environment (SDE). Without any effort by the user, the SDE will access all applications, gather the data points requested, conduct the analysis requested, and present the result to the user.

Research conducted in both the LWRS Computer-Based Procedures for Field Workers and the LWRS Automated Work Packages efforts indicates an increased interest by the industry to implement electronic work packages (eWPs) and computer-based procedures (CBPs) to improve system efficiency and reliability as well as increase human performance related to activities conducted in the plant. The Nuclear Electronic Work Package – Enterprise Requirements initiative, which was facilitated by Idaho National Laboratory (INL) in 2016, investigated how eWPs will enable immediate paper-related cost savings in work management and provide a path to future labor efficiency gains through enhanced integration and process improvement in support of the Nuclear Promise [2]. The deployment of eWPs and CBPs will create a new source and demand of data that needs to be incorporated into the SDE supported by the plant. The addition of eWPs and CBPs adds more near real-time data which can be used to make important plant decisions.

A study conducted by Oxstrand et al. in 2015 in the LWRS Automated Work Packages effort demonstrated means for automatic and wireless acquisition of plant process and components status information into the work order on a mobile device [3]. To enable this automatic acquisition of data, a prototype platform for data exchange between the field instruments and the mobile devices was designed. The researchers aimed to develop an architecture design that is prompt, robust, and interoperable with any technology.

To enable well-informed plant decisions the data need to be both accurate and relevant. Both quantitative and qualitative techniques need to be utilized to analyze available data points. The approach to use these types of techniques is usually referred to as data analytics. The ability to conduct data analytics is an important process needed to be incorporated in the SDE. Data analytics can help the utilities improve operational efficiency by being able to respond quickly to trends shown in the data. The trends analyzed, for example, could point to equipment failure detection and issues that could be causing delays in work completion. The entire data collection needed depends on what the desired subject to be observed is. In order to verify that you have collected the desired data, an analytical model is developed and used on the gathered data to test its accuracy. The model, and possibly the data gathered, is revised and tested again in a process known as “training” the model until it functions as intended and proven to provide the desired analysis.

During the 2016 annual Nuclear Information Technology Strategic Leadership (NITSL) group meeting the nuclear utilities identified the need for research focused on data analytics. It was suggested that the effort would develop and evaluate use cases for data mining and analytics for employing information from plant sensors and databases for use in developing improved business analytics.

The goal of the specific research effort, which is described in this paper, is to support the adoption of the SDE concept in the nuclear industry. To provide proof of concept to the industry, a use case study will be conducted in 2017. The use cases will help demonstrate both benefits of data analytics and how to present aggregated or analyzed data to the user in a meaningful way. In addition, the researchers will benchmark the use of data analytics both within the nuclear industry and other industries to gather lessons learned.

## 2 USE CASE STUDY

The use case study will research potential approaches to building an analytics solution, which will conduct analytics on data from multiple locations and provide the result to the user in a way that enhances the user’s productivity and/or ability to make well-informed decisions. Although many commercial software products exist that may be able to provide such an analysis, the use case is intended to be able to be referenced as a more cost effective and nuclear industry tailored approach.

The analytics solution will consist of a data integration layer, predictive and machine learning layer and the user interface layer that will display the output of the analysis in a straight forward, easy to consume manner. A middleware platform is used for the data integration layer between different plant applications, such as the work management system and the plant information database, e.g., NextAxiom *hyper*Service platform. The researchers will also use a statistical programming language, such as R or SAS, in the predictive and machine learning layer. The effectiveness of various user interface types will be studied to determine the most appropriate manner in which to present the output to the end user.

The use case study will be hosted by Arizona Public Service Palo Verde Nuclear Generating Station (PVNGS) in the second half of 2017. The use case study is therefore a collaborating effort between the researchers at Idaho National Laboratory, NextAxiom, and PVNGS.

## 2.1 Methodology

A large set of potential use cases was identified in the early stage of the effort. For example, the team considered targeting the tracking of design implementation work orders in order to get real time status updates related to specific plant design modifications. The team also identified the potential benefits of tracking work and equipment status by accessing data logged during operator rounds in the plant as well as in the main control room.

### 2.1.1 Use Cases

Two use cases were selected to be used in the study; an equipment anomaly detection use case and an engineering work management use case.

The equipment anomaly detection (EAD) use case describes the implementation of a system for detecting, evaluating and dispositioning potential equipment anomalies. These anomalies are deviations from the standard operating parameters of the equipment and can indicate the beginning of an equipment failure. Successfully implemented, this system would help identify early indications of equipment failures and improve equipment reliability by tracking the anomalies, allowing for preventative action to be taken before the failure could cause an issue that could lead to worker injury or shutdown of the reactor.

The engineering work management use case aims to provide a portal that would gather data from several systems and utilize the data to more accurately help assign work based on availability. The portal would then update the source systems about the work scheduled. This use case will only be studied in detail if time permits.

### 2.1.2 Develop the System for the Use Case Study

The research team decided to focus on the EAD use case and study the engineering work management use case if time permits.

This system in the EAD use case would consist of 3 parts:

1. A system for detecting “equipment anomalies”
2. A dashboard for engineering and operations to review and disposition these equipment anomalies
3. An integration layer for both retrieving contextual data and updating data according the anomaly disposition

While systems for detecting equipment anomalies based on time-series process data have existed for some time, these systems are expensive and labor-intensive to maintain. The researchers hope that by using new machine learning techniques much of this manual effort can be automated and scale anomaly detection more efficiently. Such techniques provide a method of data analysis that automates analytical model building. Using algorithms that iteratively learn from data, the machine learning techniques allows computers to find hidden insights without being explicitly programmed where to look. As the algorithms are tuned to find the correct data through an iterative process it is expected to see high false-positive rates, where many anomalies that are detected are mistakenly found and are inconsequential. By using an integrated dashboard, there is hope to improve the efficiency of screening anomalies as well as improve the ability to act on anomalies by efficiently generating condition reports and other work mechanisms. An integrated dashboard consists of information from various disparate data sources presented in one graphically intuitive user interface. Interactive organization and drill down capabilities allow quick access from the dashboard oversight down to an equipment's anomalous detail. This brings all the information into one location for a user to be able to make more informed decisions by having the correct information at the right time.

A system has been developed implementing these machine learning algorithms, and shows strong historical sensitivity to equipment anomalies. Data fed to the system is not real-time and access to the system is difficult. Unfortunately, the current system relies on data retrieved via the PI Datalink Excel Add-in, which is not a reliable nor sufficiently flexible tool for developing a near real-time proof of concept, where users have immediate access to the data. A near real-time system will be able to provide an easier access point to any department that could benefit from the data. Systems that are not near real-time usually are not updated as frequently and could have issues with providing access to all potential users. NextAxiom's *hyper*Service Platform will be used to gather the data from the various source systems.

Figure 1 illustrates the relationships between the different components involved in the use case.

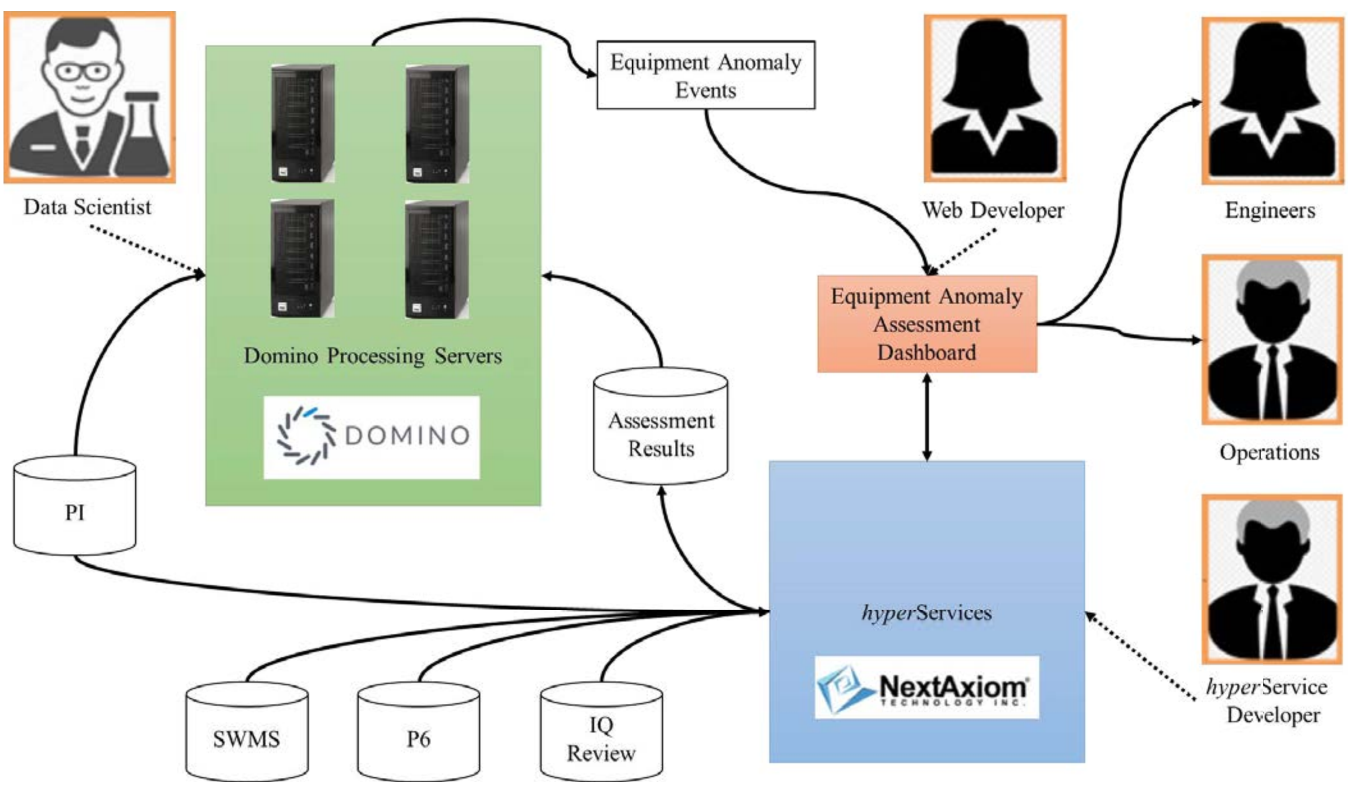


**Figure 1. The relationships between the different components involved in the equipment reliability use case.**

### 2.1.3 Develop the User Interface of the System

The users (in this case engineers and operators) use an integrated dashboard to access information important to their job functions. The information presented on the dashboard comes from the PI database, the scheduling tool (P6), and the work management system (SWMS). The *hyper*Service Platform is used to access the data to determine what to present to the user and how to do so. Also the *hyper*Service Platform will be used to log what specific information is accessed most by the users. This log will be used to adjust the data gathered in order to provide predefined data sets to the users.

As the engineers use the dashboard they will be able to disposition the anomaly events into three categories: Acknowledge, Watch, and Condition Report. Acknowledging the anomaly is used by the engineer to state that he has seen the anomaly and doesn't believe there is a cause for further investigation. Setting the anomaly in the Watch state informs other engineers that there is a possibility that there should be a concern and to continue watching the event. The Condition Report (CR) state allows the engineer to create a CR automatically from the information gathered by the EAD system with just a click of a button. This frees up the engineer to continue investigating other anomalies. The states are fed back to the Domino processing servers where the data is incorporated back into the data analysis in order to train the system on how to detect if future anomalies are actually real or may be a false positive event. The research team will collect and analyze data from the development and use of the architecture described in Figure 1 to report on the resulting lessons learned.

### 2.1.4 Evaluate

For the proof of concept, all valid PI points for several systems will be accessed from all of PVNGS' three reactor units. After an initial seed load, this data would be pulled daily at a 15-minute granularity. At ~250 PI points, for two systems and three units, this would be 1,500 PI points for 144,000 points a day. The shorter data feed interval will allow more fine-tuned trending of any anomalies and increase the speed at which equipment failure can be predicted.

The research team will assist PVNGS with the development and implementation of the equipment reliability study. In addition, the research team will present the use case study at the Data Analytics Initiative Special Interest Group (SIG) workshop during the NITSL Conference in July 2017. A report will be provided on the data analytic prototype's development process, data gathered, and outcome.

## 3 PRELIMINARY FINDINGS

The researchers designed the functional aspects of the user interface for the EAD use case by creating wire frame mockups. Mockups were demonstrated to potential users in order to receive feedback on what the potential users think is needed in the dashboard in order to accurately determine the course of action on the events reported. The team updated the designs after feedback from several focus groups each reviewed the mockups. This design process allowed the users to give their opinion on how the dashboard would function and what data they need in order to perform their roles.

Figure 2 shows the mockup of the dashboard's home screen. As shown on the top section of the dashboard, the users have the ability to filter and refine the list of anomalies they are seeing. This allows the engineers to focus on the systems they are responsible for. The engineer can select an anomaly event from the list in order to view more details regarding the event.

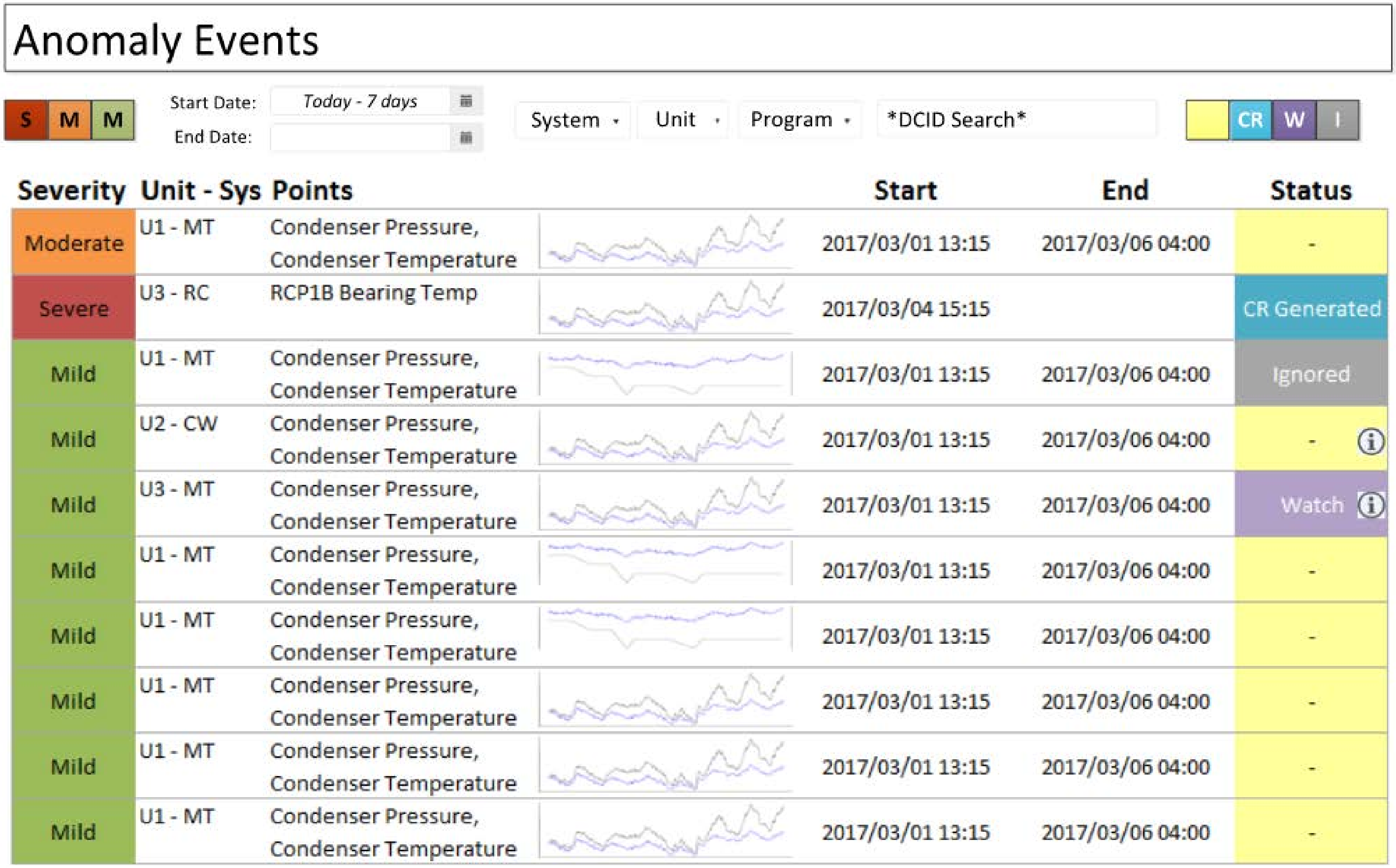


**Figure 2. Equipment Anomaly Detection Home Screen Mockup.**

Figure 3 shows the details page mockup of the event. The details screen allows the user access to the data located in various silo systems seamlessly. The source systems that provide the details are captured in the in the upper right hand corner of the screen. The initial systems include the PI database, the scheduling tool (P6), the work management system (SWMS), the IQ Review system, engineering modifications repository, and a file repository where reference material can be found, e.g. design documents. Each button on the details screen displays the number of records found regarding the equipment the anomaly is occurring on. The underlying data in the system can be accessed by clicking the respective buttons.

When the user has reviewed the information a decision can be made whether to acknowledge the event, put the event on a watch list, or generate a CR.

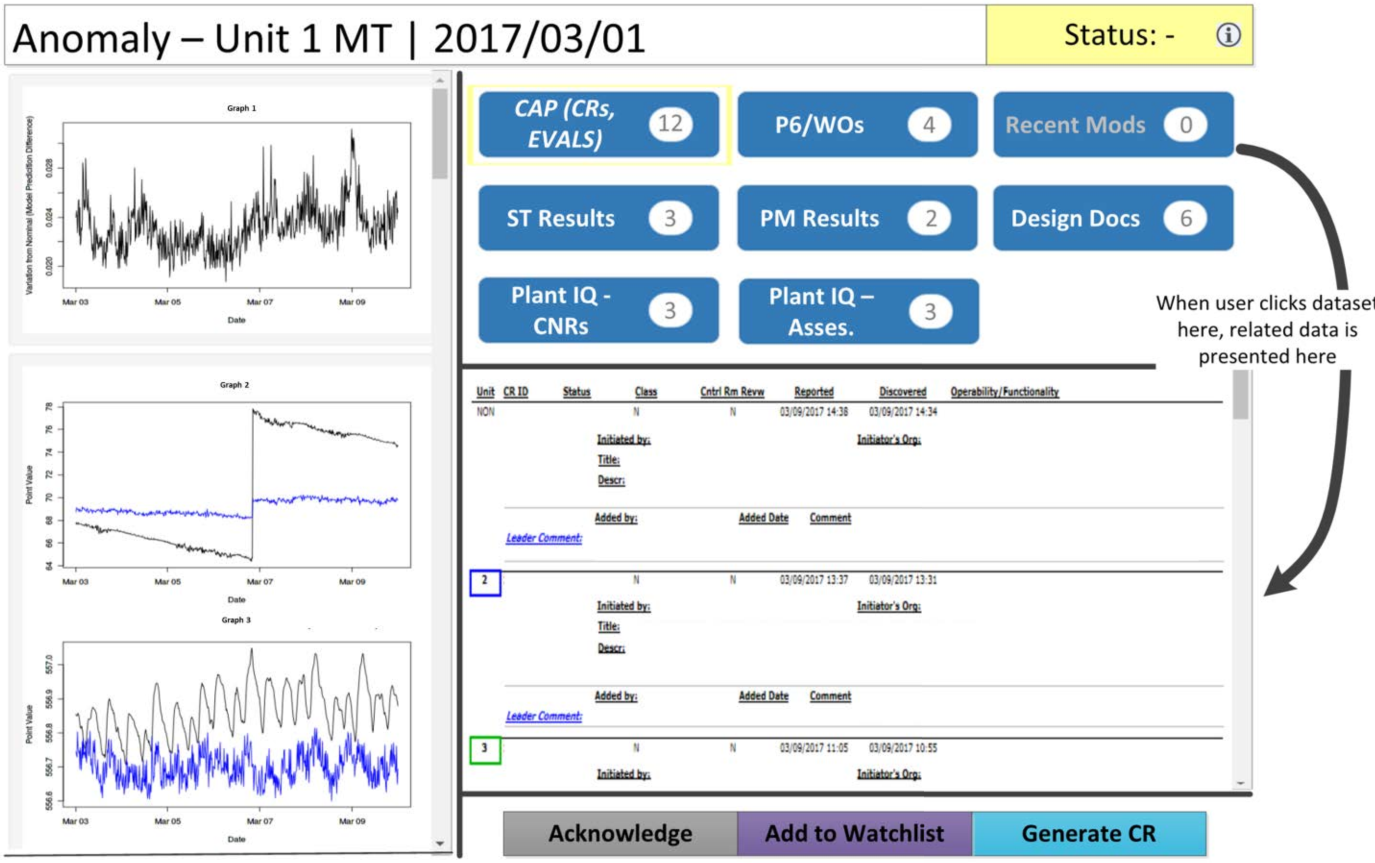


**Figure 3. Equipment Anomaly Detection Detail Screen Mockup.**

The team has created the services that connect to the PI database, the scheduling tool (P6), the work management system (SWMS), and the IQ Review system as shown in Figure 3.

When the users choose to generate a CR a popup is displayed to the user as seen on the left of Figure 4. The data presented to the user is a very small summary of the CR that will be generated: Title, Description, and optional Suggested Disposition. The title and description are auto-generated with all the information needed but the system allows the user to modify the fields if desired. Once the user clicks "Generate CR", the system automatically fills out a condition report and submits it. The right side of Figure 4 displays some of the data that is automatically filled out with the data already gathered in the background. The user never sees the right side of Figure 4. As the users look at the detail data and categorize the events, their actions are recorded in order to train the analytic model to have the ability to report accurate anomaly events.

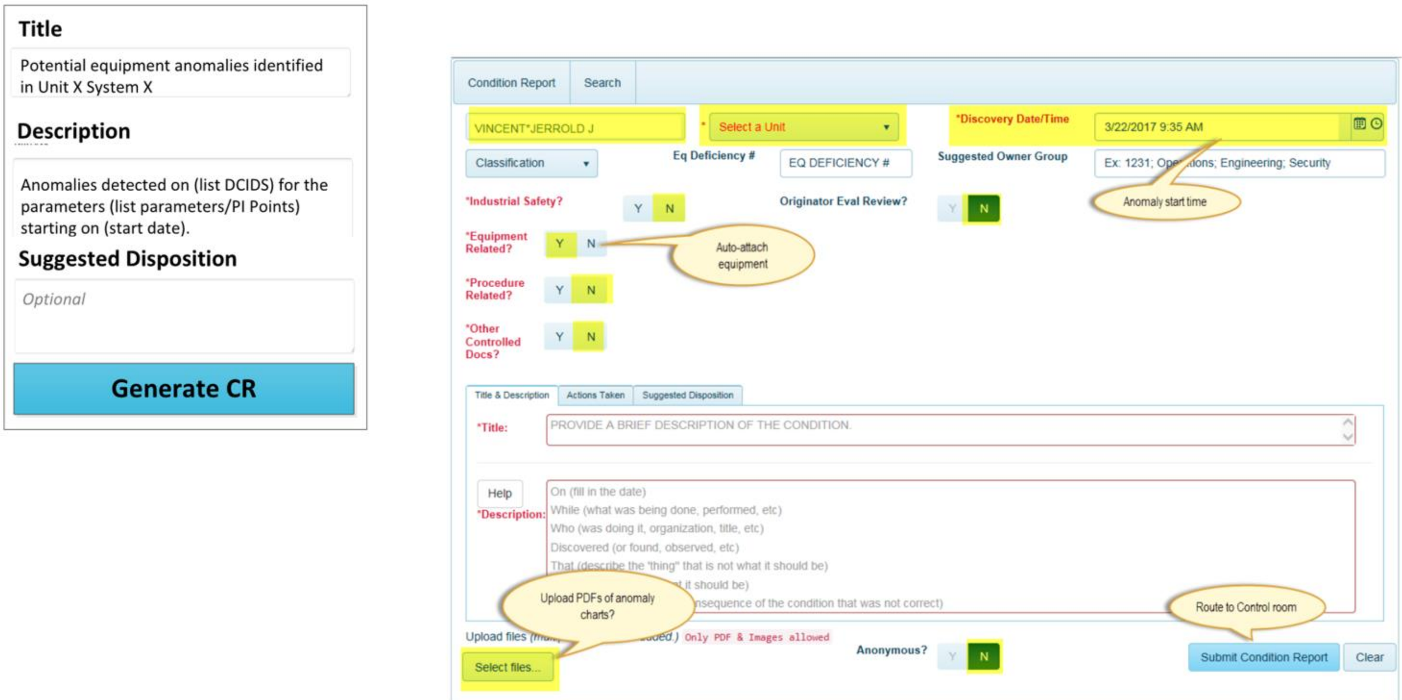


**Figure 4. Condition Report One-Click Creation Mockup.**

The design and data gathered is expected to change as the use case continues. The hope is to fine tune the features in the dashboard that improve work efficiency and tune the data feedback to the analytics model to improve anomaly detection. Also user feedback on the use of the dashboard will steer the any design updates.

## 4 SPECIAL INTEREST GROUP

In addition to the use case study, the researchers will benchmark the use of data analytics both within the nuclear industry and other industries to gather lessons learned. This input will be useful when utilities are moving forward with advanced data analytics. In the nuclear power industry, once a better work practice has been proven, there is a general expectation that the rest of the industry will adopt it. The benchmark is conducted through a SIG.

The SIG focuses on broader questions related to data analytics and how/when it should be used to support the Nuclear Promise. The purpose of the SIG is for members to share insights and lessons learned from related activities in their organization and learn from others' experiences. The SIG provides feedback on the use cases study and feeds the results from the study back to their organizations.

As of February 2017, the SIG currently consists of 30 members. Fifteen of the members represent four U.S. nuclear utilities (Arizona Public Service, Southern Nuclear Company, Dominion, and Duke Energy) and one international utility (EDF Energy). The other 15 members represent Idaho National Laboratory, Sandia National Laboratory, the Institute of Nuclear Power Operations, and six vendors.

## 5 PATH FORWARD

The research team will refine the design of the EAD system over the course of the study based on feedback from users. The lessons learned and development process will be presented in a report at the end of the study through the LWRS program. It is planned that the SIG will continue after the study concludes in order to promote the continuation of sharing the knowledge gained as more activities are completed by its members and their organizations.

## 6 ACKNOWLEDGMENTS

The authors would like to express special gratitude to the following people for their collaboration and support of this research effort: Ann Orr, Mark Johnson, Bruce Gordon, Veenu Ravi, Marko Mitrovic, Jerrold Vincent, and Bradley Fox at Palo Verde Nuclear Generating Station, for supporting the Seamless Digital Environment effort. Sandy Zylka, Richard McLaurine and Eric Rich at NextAxiom Technology Inc, for their support and input in defining a viable demonstration that would show the concept of this research.

This paper was prepared as an account of work sponsored by an agency of the U.S. Government under Contract DE-AC07-051D14517. The views and opinions of the author expressed herein do not necessarily state or reflect those of the U.S. government or any agency thereof.

## 7 REFERENCES

1. J. Oxstrand, A. Al Rashdan, A.Bly, B. Rice, K. Fitzgerald, & K. Wilson, *Digital Architecture Planning Model (INL/EXT-16-38200),* Idaho Falls, USA (2015).
2. J. Oxstrand & A. Bly, *Computer-Based Procedures for Field Workers - FY16 Research Activities (INL/EXT-16-39984),* Idaho Falls, USA (2016).
3. J. Oxstrand, A. Al Rashdan, K. Le Blanc, A. Bly, & V. Agarwal, *Light Water Reactor Sustainability Program Automated Work Package Prototype: Initial Design, Development, and Evaluatio*n (INL/EXT-15-35825)*,* Idaho Falls, USA (2015).