MARY KAY O'CONNOR PROCESS SAFETY CENTER
TEXAS A&M ENGINEERING EXPERIMENT STATION

18th Annual International Symposium
October 27-29, 2015 • College Station, Texas

Equipment Data Collection…Simplified

Keith Lapeyrouse
Process Reliability Solutions, LLC
kmlapeyrouse@ProcessReliabilitySolutions.com

John Harrison
SAP Labs, Canada
john.harrison@sap.com

Nils van Heijnsbergen
Intergraph Corporation - Part of Hexagon
nils.van.heijnsbergen@intergraph.com

Ana Pacheco
Intergraph Corporation - Part of Hexagon
ana.pacheco@intergraph.com

Richard "Rich" Anderson
Utopia Inc.
randerson@utopiainc.com

Michael "Mike" Jordan
Utopia Inc.
mjordan@utopiainc.com

Abstract

Lack of Equipment Data is a fundamental barrier to understanding asset performance inside of and outside of ERP solutions. Defining the solution is highly dependent on the organization’s existing IT infrastructure, the effectiveness of implementation of these solutions and continuing support to keep data accurate. The costs of the solution can vary by one or more orders of magnitude depending on the organizational goals and value propositions communicated to management.

This presentation will address this topic and describe three levels of solution:

- Non-Integrated/Unstructured Data – Field Data Collection Tools w/o Comparative Reporting in SAP
- Integrated/Structured Data– Intergraph / SAP/ Comparative Reporting
- Integrated/Structure Data + Data Governance For Master Data – Intergraph MDG Master Data Governance SAP Solution

We will discuss a Multi-generation approach and benefits associated with each generation. We will show a product in each space that can accomplish the above stated objectives.

Preface: Why do we care about data?

Data, both Master and Transaction (which is defined later in the paper), when properly collected, provides the raw information that can be transformed via metrics to monitor business process performance. These metrics are becoming an increasing topic of discussion in Industry organization and are embedded in practices that define "Recognized and Generally Accepted Good Engineering Practice (RAGAGEP). Though memberships in these industry organizations, we have a responsibility to improve our business practices.

- **Cost Avoidance in Mega Projects**

Vendors who provide data management solutions are quick to claim cost avoidance. A literature search to quantify these hard savings failed to provide specific examples. Intergraph reports one customer expects cost avoidance of up to 3.5% of a plant's total installed cost through: 1) Design optimization of material and labor, 2) Increased engineering productivity, 3) Reduced cycle time, 4) Duplicated success and reduced risks through reuse of proven designs, and 5) Increased operating efficiency and shorter shutdowns via effective control of as-built and asset information (Intergraph Corporation, 2005).

- **Knowledge Retention in Aging Work Forces**

The issue of knowledge loss associated with the aging workforce is a topic that has been well described. Quotes such as "In our process engineering group, more than 50% of the engineers are new to the company in the last three years because our experience has been walking out the door." are frequent in industrialize settings (David W. De Long - Accenture Institute for Strategic Change, 2002). Attracting and retaining new talent is difficult when the industry is not seen as a leader in deploying innovative, cutting edge technologies. As long term employees leave and work continues to transform to contract-based, mobile workforce, so too does the inherent knowledge.

- **Accident Avoidance**

Trevor Kletz writes in the "Introduction of Lessons from Disaster" in 1993 about Individual Knowledge loss associated with employees leaving organizations where these accidents occur. I believe the author is referring to the ability to contextualize information and to apply it to specific settings. He raises the question, "How can we persuade people to use knowledge that is already available?". He further states in the introduction that so many accidents are exact enough in nature for him to make the observation, "Don't bother to send me the accident investigation. I'll send you a copy from my files" (Kletz, 1993).

The issue of data contextualization and communication when it is needed, where it is needed, perhaps has the greatest opportunity in accident avoidance. Today, this is largely done through the concepts associated with Critical Procedure Use. Repeat accidents are still occurring today. Understanding how to apply lessons learned and internalizing this knowledge needs to be an area which is further documented in both academic and commercial literature.

- **Performance Monitoring has Increasing in Presence in Codes and Standards**

Recent changes in codes and standards are recognizing the need for business process performance management. Business process performance management describes the process of defining acceptable performance (hence unacceptable business performance is an excursion from the defined boundaries and monitoring the number of times these boundaries are exceeded).

Recommended practices and standards are being developed to support performance management. Examples of these are American National Standards Institute (ANSI)/American Petroleum Institute Standard (API), Recommend Practice (RP), Process Safety Performance Indicators for the Refining and Petrochemical Industries and ANSI/API RP585 Pressure Equipment Integrity Incident Investigation. Examples of practice Standards that are introducing performance measures include API Standard 510 and IEC 61511.

- **CCPS Vision 20/20**

The American Institute for Chemical Engineers (AIChE), Center for Process Safety (CCPS) published a bulletin in 2014 entitled “Vision 20/20”. “Vision 20/20” established five industrial tenants and four social themes to promote process safety. These themes and tenants include:

Industry Tenets

- Committed Culture
- Vibrant Management Systems
- Disciplined Adherence to Standards
- Intentional Competency Development
- Enhanced Application & Sharing of Lessons Learned

Societal Themes

- Enhanced Stakeholder Knowledge
- Responsible Collaboration
- Harmonization of Standards
- Meticulous Verification

We will examine Disciplined Adherence to Standards, Enhanced Application & Sharing of Lessons Learned and Meticulous Verification relative to both their definitions and the implications for data collection (American Institute for Chemical Engineers, 2014).

Disciplined Adherence to Standards means using recognized design, operations, and maintenance standards that should be followed all the time. Companies identify, document, and diligently follow RAGAGEP for new and existing designs. As industry standards evolve, companies codify significant new learnings in their standards for new and existing equipment there by continually improving their standards which establishes the minimum expectations for design, operations, and maintenance. Companies identify and manage process safety risks arising from gaps against these standards ensuring a constant standard of safety for assets regardless of age (American Institute for Chemical Engineers, 2014).

Inherent in this definition is the need to understand codes and standards in force at the time of construction along with the current edition. Further, it is necessary to analyze new standards to understand the changes so that a management plan can be developed to manage process safety gaps identified by the new standard. This will require that construction code data be added to existing equipment data model.

Enhanced Application & Sharing of Lessons Learned communicates critical knowledge in a focused manner that provides continual learning in the organization. It required learning from accidents, near misses, internal/industry benchmarking, and success stories. This is accomplished through identifying the learning and efficiently sharing the learnings without overwhelming the organization. The process should also determine if it is appropriate to update internal standards or practices. The final element is to follow the practices defined in Disciplined Adherence to standards (American Institute for Chemical Engineers, 2014).

Internally, management systems should be in place to collect accidents, near misses, industry benchmarking data, and success stories. Data governance systems become essential to provide quality data.

This definition has a strong focus on internal learnings with the exception of the reference of industry benchmarking. Industry benchmarking is highly dependent upon clear metric definitions, the underlying data cleanliness and its' quality to support the effort.

Harmonization of Standards for the safe design, operation, and maintenance of equipment streamlines practices, eliminates redundancy, and cooperatively addresses emerging issues. Various standards have the potential to conflict with each other, resulting in an increased risk of confusion. Harmonized standards provide consistent guidance for design, operations, and maintenance, to support effective understanding of process safety risk (American Institute for Chemical Engineers, 2014).

The lack of Harmonization adds complexity to the development of Business Metrics. One example is the group definition of a Hazardous Fluid. Fluid hazard definitions have been developed by the National Fire Protection Association (NFPA), American Petroleum Institute (API), and American Society of Mechanical Engineers (ASME) within the United States alone. The result of this is a recognition that data definitions and agreements are essential when collecting data.

Data 101:

In order to begin a discussion on Enterprise Resource Planning (ERP) Systems, Plant Information Management (PIM) Systems, Data Governance/Quality and a Multi-generational approach to these items, a common dictionary needs to be established.

- **Data Management**

According to the Data Management Association: Data Resource Management is the development and execution of architectures, policies, practices and procedures that properly manage the full data lifecycle needs of an enterprise." Some of the challenges associated with data management include: 1) Many data sources, 2) Data Duplicated across sources, 3) Different forms of data, 4) Data gets out of synchronization and 5) Many different uses of the same data.

- **The System of Record**

The system of record (SOR) or Source System of Record (SSOR) is the holder of the authoritative (the Golden Record) for any data in Information Technology (IT) solution which utilizes multiple systems where data may be replicated. There may be many systems that contribute data to a Master Data Management (MDM) solution relative to a particular functional location or equipment depending on the business systems.

Typically, the MDM acts as a hub for multiple systems, many of which could allow (be the source of truth for) updates to different aspects of information on a given entity. For example, the CRM system may be the "source of truth" for most aspects of the customer, and is updated by a call center operator. However, a customer may (for example) also update their address via a customer service web site, with a different back-end database from the CRM system. The MDM application receives updates from multiple sources, acts as a broker to determine which updates are to be regarded as authoritative (the Golden Record) and then syndicates this updated data to all subscribing systems.

- **Types of data**

System Configuration Data defines the computers, processes, and devices that compose the system including the specific definition of the elements that define and/or prescribe what a system is composed of in terms of data families, data fields, data values and value definitions, etc. (Wikimedia Foundation, Inc., 2015).

Master data: In general, this is relatively static data, the data values do not change frequently, and are used to describe an object and its' capabilities. E.g. A pump: master data could be the name of the manufacturer, the operating temperature range, its capacity, the type of fuel that is needed to operate the pump etc. Another characteristic of master data is that in general there is a need to control any changes to the master data since other processes depend upon having the data correct and current.

As the basis for what you are going to analysis, correct and complete master data is essential. Ensuring accurate, complete asset master data can also have benefits outside of correct analytics. It can have a direct impact on reducing environmental risks, health and safety risks, improving operational performance and maintaining regulatory compliance.

Since the volume of master data is overwhelming, prioritization is the key IDC, a market intelligence firm estimates that less than 1 % of data collected is analyzed, and with the advent of the Internet of Things, Connected Manufacturing, and Connected Logistics, there will be more and more data to potentially be analyzed. If we keep going the way we have been, even less of that data will be put to use. We have been collecting data on the basis that it exists and the answers will come, but this is not the “Field of Dreams”. We have to step back from the emphasis on just collecting data, and concentrate on collecting the right data, the data that can be used to develop insights, improve our lives, increase our safety, etc.

One strategy for this prioritization is to look first to what master data is needed, and by whom. Typical users of asset master data might be the plant engineering, plant maintenance, reliability, or operations organizations. If you cannot identify a specific use or potential 'owner of the master data, you can safely assume that the data is non-essential. Some analysts have suggested that less than 10% of the information generated during a Greenfield capital project is necessary or useful for ongoing asset operations and maintenance. The critical factor of course is identifying and capturing that essential 10%.

ISO 14224 classifies three types of Master Data for equipment reliability calculations. These data types consist of technical, operational and environmental data. The standard further defines the types into data parameters characterizing the design and use of an equipment item (ISO copyright office, 2006).

It should be noted that the ISO 14224 database was patterned after the Offshore Reliability Database (OREDA) and was not designed to collect new equipment data. Therefore, when collecting data for new projects, additional data definitions may be applicable. Examples of these are:

- **Template Data –** Data definitions designed by an organization as appropriate to a specific type of functional location or equipment. It may be documented in a data taxonomy and include a detailed definition of what is included and priorities for field.
- **Engineering Specifications –** Documents used to convey process data, equipment design, equipment components and/or equipment configuration. Significantly more detailed than the information typically required by operations and maintenance.
- **Operating / Performance Ranges –** Define the operating context for the equipment. While much equipment is specified for a single operating point, equipment is frequently specified with a variety of conditions which affect performance. These conditions may include maximum flow/pressure/temperature, minimum flow/pressure/temperature, normal flow/pressure/temperature and various chemicals/compositions.
- **Location Data (Geographical or Geospatial Information System) –** The advent of mobile devices and mapping software had made the use of location data more important organizationally.
- **Maintenance Process Data** (Inspection/Testing, Preventative Maintenance and Predictive Maintenance Task scheduling).
- **Other**

Transactional (operational) Data: In general, this is data that changes and is used to describe what is happening at a point in time. E.g. a pump: operational data could be the current operation temperature, or the amount of material that was pumped during a specific time period.

No matter what data is being gathered, it should go through a review / data governance process. This will help to further cleanse the data prior to being added to the various processes. Having multiple qualified people review the information results in a number of things, which include: 1) More than one person understands the data requirements, 2) Errors are more likely to be identified, 3) Misunderstandings about the data can be clarified, and if needed, an investigation can be undertaken to clarify the information.

Process Operational Data – Gathering temperature data alone, does not provide valuable information. Additional data will be necessary, such as: The water, the oil, inside temperature, outside temperature, temperature of equipment at 50% load, 70% load, etc. The value alone does not tell you when the temperature was recorded, nor does it tell you how many times the temperature was recorded. Let alone the circumstances around the value, e.g. what product was being produced, quality of the product, who was operating the equipment, how the value was taken, and who took the value. All of these parameters increase the understanding of the single value of "Temperature".

With enough parameters surrounding the value being recorded, you can gain an understanding of the context surrounding it. A single occurrence is not normally sufficient to add value. The data values need to be looked at (analyzed) over a period of time. How frequently the data values are recorded and are stored are important considerations in analyzing data.

Some companies are very good at recording data at a high frequency, but are limited by storage technology. Others are bad at data collection, but very good at data storage. The technology to do this varies from paper (slow data recording, good long term storage, hard to analyze) to digital (quick data recording, error free, costly, easy to analyze), and every variation in-between. As you start your journey down the road of data collection and analysis, remember it is not just the value you are recording, but the circumstances around it.

- **Data Governance/Data Cleansing**

Both Data Governance and Data Cleansing have the same objective, the creation of data intelligence. Data intelligence can be defined as a company's use of internal data to analyze their own operations or workforce to make better decisions in the future (https://www.techopedia.com/definition/28799/data-intelligence).

Data governance is a control process that ensures that the data entry meets precise standards for data quality by monitoring incoming production data and communicating errors in data back to operational team members, or to the technical support team, for corrective action. Data validation is a part of data governance and almost invariably means data is rejected from the system at entry and is performed at entry time, rather than on batches of data.

Data governance ensures that important data assets are formally managed throughout the enterprise so that data can be trusted. It puts people (data stewards and data custodians) to improve data quality, including the fixing of incorrect data and preventing issues with data (Wikipedia). Data governance differs from data cleansing in that it works to ensure that first pass quality is of the highest caliber (Wikimedia Foundation, Inc.).

Data cleansing, **data cleaning** or **data scrubbing** is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database. The term refers to identifying incomplete, incorrect, inaccurate, irrelevant, etc. parts of the data and then replacing, modifying, or deleting this dirty data or coarse data.   After cleansing, a data set will be consistent with other similar data sets in the system. The inconsistencies detected or removed may have been originally caused by user entry errors, by corruption in transmission or storage, or by different data dictionary definitions of similar entities in different stores. Some data cleansing projects will also include data enhancement, where data is made more complete by adding related information, is a common data cleansing practice. For example, appending addresses with phone numbers related to that address. Data cleansing may also involve activities like harmonization or standardization of data. Harmonization refers to the updating of records to a consistent single definition. Standardization of data refers to changing a data set to use new standard codes. After data cleansing, the data should have the following characteristics:

- **Clean:** the data does not contain spelling mistakes, is free of unnecessary, complex syntax and tables do not contain duplicate records.
- **Complete:** all of the required fields in each record have been populated.
- **Accurate:** all of the required fields in each record have been populated correctly.
- **Consistent:** the data is described in a standardized, structured manner and adheres to an agreed-upon naming convention.

- **Data Governance and Management of Change**

The last aspect of data governance we will address is the timeliness of data management. This issue deals with the point at which the data should be processed and its status changed to installed. It is generally recognized that this is part of the commissioning process.

However, a larger, more complex issue of MOC looms ahead of us. CCPS in its publications regarding “Vision 20/20” in Disciplined Adherence to Standards refers to the ability to assess the impact of change standards on the suitability of assets (American Institute for Chemical Engineers, 2014) This view has profound implications of change for the chemical industry. It brings us to the point where we must monitor standards for change, identify assets that are impacted by the changes in the standards and determine what actions must be taken to our existing assets. This change alone is a sufficient driver for us to recognize that our existing systems are incapable of this function and the changes described here after must occur.

The Multi-generation Vision

The multi-generational view of data management below is presented based on changing technologies and changing business needs. The expectations of performance and abilities to perform are constantly changing due to process technology changes, feedstock changes, regulatory changes, shareholder expectations and a host of other issues. Generally, three primary levers have been used to optimize performance. These levers were: costs of feed stocks, costs of energy, uptime and costs of labor and materials (primarily maintenance). We believe that by improving data management practices through a multi-generational perspective, a positive gain can be accomplished for both uptime and maintenance.

There are multiple generations or ways in which this information can be collected either manually, semi-automated, and fully automated. Unfortunately, the sources of this information are spread across multiple phases of a facility’s lifecycle and from many sources.

We are proposing three generations of data management practices, which are technology based. The goal of each generation is to improve data for more effective decision making.

Generation	Nature of Data	Business Results (Goal)	Theme
1	Non-Integrated / Unstructured Data	Improving the completeness and accuracy of ERP System Master & Transactional Data	The search for the truth
2	Integrated/Structured Data	Data supports maintenance effectiveness and drives Equipment Reliability Improvements	One Truth … validated but not evergreen
3	Integrated/Structured Data + Data Governance	Data supports Product Reliability Improvements	One Truth … validated & evergreen

Table 1, The three generations of data management

Multi-generational planning provides a mechanism where large complex changes are involved. The MGP process has three major steps: 1) defining the goal, 2) identifying the generations, and 3) identifying and categorizing the needed technologies. For MGP purposes, the needed technologies fall into three categories:

1. The technologies that are already understood and being used.
2. The technologies that are known, but that an organization needs to be much more capable at than they are today.
3. And finally, the technologies that are entirely new (McDonough).

We see this view of changing technology in our multi-generational view of data management. We see the outcome of implementation as safer, more efficient, and more cost effective facilities.

- **Identifying the Goal**

The most important part of the multi-generation project planning process is identifying the overall goal. An organization does not undertake an MGP for some trivial goal. The goal of the MGP must be something worthwhile, compelling and difficult (McDonough).

It is important to have time pressure and value pressures for these difficult projects. Projects of this nature tend to go on forever if a stake is not placed in the ground. A gated project approach along with recognition for organizations with success is essential for accomplishing wins and driving acceptance (McDonough).

- **Identifying the Generations**

Once the goal is visible, it is time to identify the generations. Each generation should be identifiable, something that everyone can understand and relate to. Each generation should represent a significant achievement in its own right, and yet be a step towards the larger goal. Each generation should embody some subset of the capabilities needed for the ultimate goal. The overall goal becomes more accessible as each of the generations proves a smaller number of technologies (McDonough).

- **Identifying the Technologies**

The generations help break up the combination of capabilities required into more manageable pieces. Introduction of new technologies represents change which much be recognized and managed.

Generation 1: Improving the completeness and accuracy of ERP System Master & Transactional Data - Non-Unified, Unstructured Data

You can't go wrong with paper. It is inexpensive, and everyone understands how to use it. All of the information required to execute a business process or maintenance activity can be easily printed out from the various maintenance and asset management systems. The work package given to the maintenance worker can have all information required to do the job; that is, if everything is kept up to date and all revisions to the documents are applied. This is a known and proven technology. As long as there is discipline and adherence to procedures, information can be collected around what and how that information should be collected, edited, and maintained.

Figure 1, Examples of Data Contextualization

Data can be collected by your team in the field using paper forms and existing documentation. In this process, it is quite easy for the engineer to collect information such as: equipment pictures, equipment model and serial number verification, actual condition of the equipment, location, configuration changes, etc.

However, there are a couple of problems with paper, transferring the corrections or additions on the paper document to the appropriate systems is an error prone process at best (after all, we were not hired for our penmanship). Transference of paper to digital data can be a very difficult and expensive process due to the need to analyze the documents for codification. There are also storage issues with paper use. In systems that lack sufficient resources, it can accumulate while waiting for resources to digitally codify. Occasionally, when the backlog becomes too big and additional resources are unavailable, the paper is discarded and with its discarding, the opportunity to capture this data is lost. Additionally, paper degrades over time and is easily lost.

Since the data is un-unified, it is possible for multiple sources of the truth to be present in multiple systems of record. This redundancy of data requires time-consuming effort to determine the document that is most reflective of the current functional location or equipment.

Generation 2: Data supports maintenance effectiveness and drives Equipment Reliability Improvements - Unified, Structured Information Readily Available

Engineering data and documentation are essential for effective and efficient operations and maintenance of an industrial facility. Every person involved in operating and maintaining a facility needs ready access to trustworthy information to perform their job effectively. Owner operators struggle to manage large volumes of unstructured, unintelligent information – documents, drawings, models, lists, and data sheets – which are often duplicated in various folders and databases throughout the organization. Even companies that have implemented a central document management system continually battle against unofficial storage locations that spring up over time. This lack of control leads to duplication, fragmentation, and degradation of information integrity, and sometimes, its outright loss.

Not addressing these challenges can lead to the inefficient execution of operation and maintenance tasks, and increased cost for modifications and turnarounds. Another often under evaluated challenge is knowledge drain. With an ageing plant, there is usually also an aging workforce that both personally and collectively may carry much valuable information about the plant, having worked there for years. If this knowledge is not kept in-house, it will be unavailable once these important employees retire. So it is essential to retain information by appropriately maintaining the engineering data and documentation and providing centralized access.

Analyzing and structuring information can be split into the following major steps:

- Rapidly capture, organize and link documents and information.
- Rapidly capture, organize, centralize and link documents from any source and link information.
- Identify document masters and eliminate access to duplicates and outdated revisions.
- Extract intelligence from documents and other sources.
- Create central tag register and load tag properties from various sources.
- Create central document register and identify Meta data.
- Automatically create links between tags and documents.
- View inconsistencies of tag attributes from various engineering data sources.
- Correct document relationships and identify missing reference files.
- Compare data and documents with 3D models and laser scans.
- Validate the facility’s “as-built” configuration at your desktop prior to field verification.
- Make information accessible to remote stakeholders.
- Visualize and navigate all facility information through an intuitive web portal.

The above provides a trusted environment of the As-built status of the plant asset. Local and remote users can quickly find the documents they're looking for. Below are several examples of how such viewing portals may look:

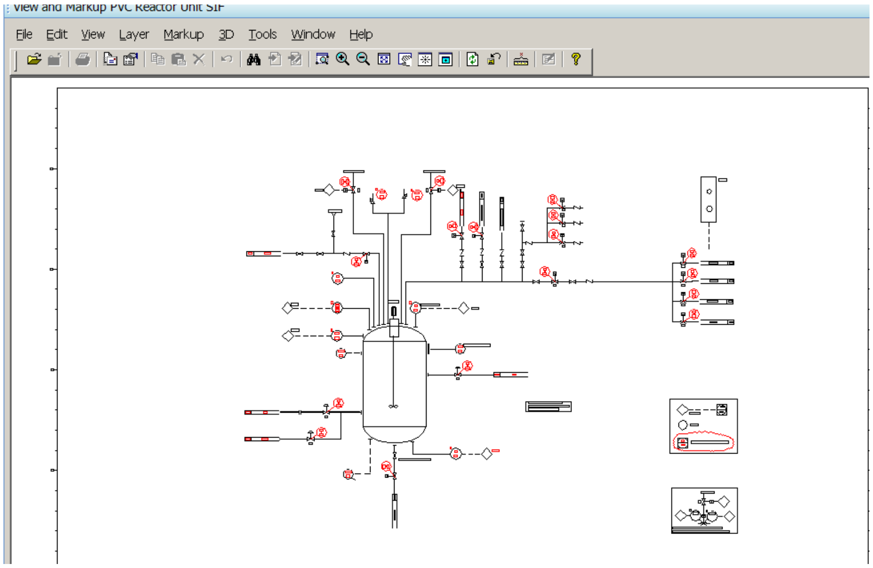

Figure 2 Start with a formerly dumb pdf of a PID

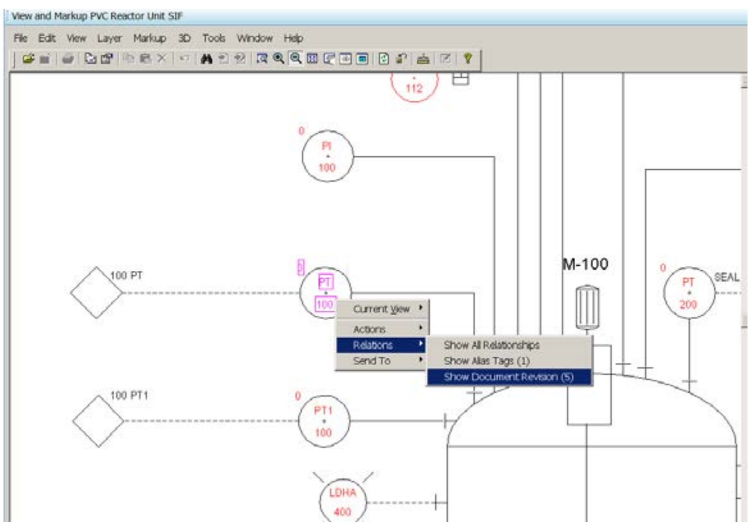

Figure 3 Navigate via intelligent hotspots into data details of a pressure transmitter

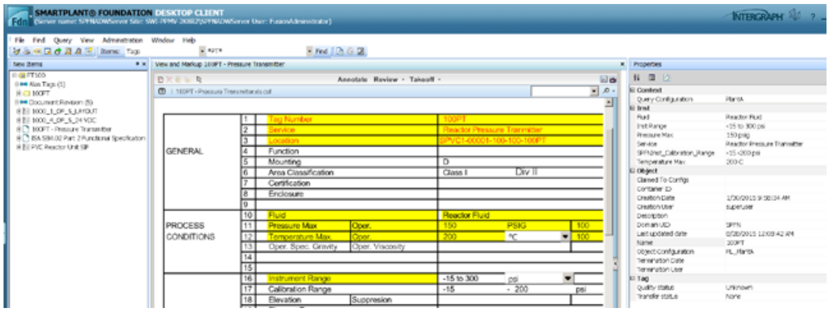

Figure 4 Have the Transmitter data sheet on hand easily and review tag properties

- Quickly find a certain equipment item everywhere it appears – specification sheets, service manuals, maintenance procedures, inspection certificates and records drawings, data sheets, P&IDs, purchase orders, invoices, construction plans, etc.
- Navigate using 3D models or Leica TruView laser scans to compare data and documents with the as-designed and as-is facility configuration.

While the above is an important milestone in achieving consistent and correct plant asset information, this data will have to be extended with information from various disciplines and sources around maintenance, operations, finance, procurement, health and safety and so on. This environment can serve mobile applications with required data and data interfaces to feed downstream processes.

Generation 3: Integrated/Structured Data + Data Governance

Even more so than going "mobile", automatic data collection and analysis requires good /correct master data. After all, once you have the "data pipe" connected from the equipment to your solution, whether ERP or analytics, the data just flows and it is very difficult to correct if the information is associated with the incorrect equipment / process.

One of the problems with automatic data collection, as strange as it might seem, is collecting the data. In some of the older plants, the sensors generate analog signals or are analog dials. For the devices that generate analog signals, it is possible to install an analog to digital converter that converts analog signals into digital ones that can be interpreted by a computer. For those sensors that are purely analog and do not generate any signals, you might have to replace these pieces of equipment (adding additional cost to the project or pilot).

Another consideration is what standard your equipment is communicating in. One very common standard is the Open Platform Communications (OPC) standard. This is an interoperability standard for the secure and reliable exchange of data in the industrial automation space. To connect your equipment to your business applications, you might need to invest in additional interface equipment such as an OPC server. This computer and the associated software converts the hardware communication protocol used by a Program Logical Controller (PLC) into the OPC protocol which then can be read by other devices. There are other protocols that you need to be aware of, for example, the HART Communications Protocol (Highway Addressable

Remote Transducer Protocol) is an early implementation of Fieldbus, a digital industrial automation protocol.

Once you have the communication (connection) problem resolved, consideration must be given to how much data you need to retain. Most sensors, gauges, and equipment have very limited data storage capabilities, and if your desire to analyze data over time including a data or process historian in your landscape becomes necessary. Data historians are special-purpose data repositories for environments in which vast quantities of time-series are acquired and stored.

Once the data is in the historian, then it is possible to automate the analysis of the store or just acquired data. This is where familiar techniques of SPC and trend analysis can be utilized. Control Charts can be generated and actions can be triggered by what an element is outside the control limits. There is no limit to the amount of variables that can be monitored and plotted.

- Data Analysis – Maintenance

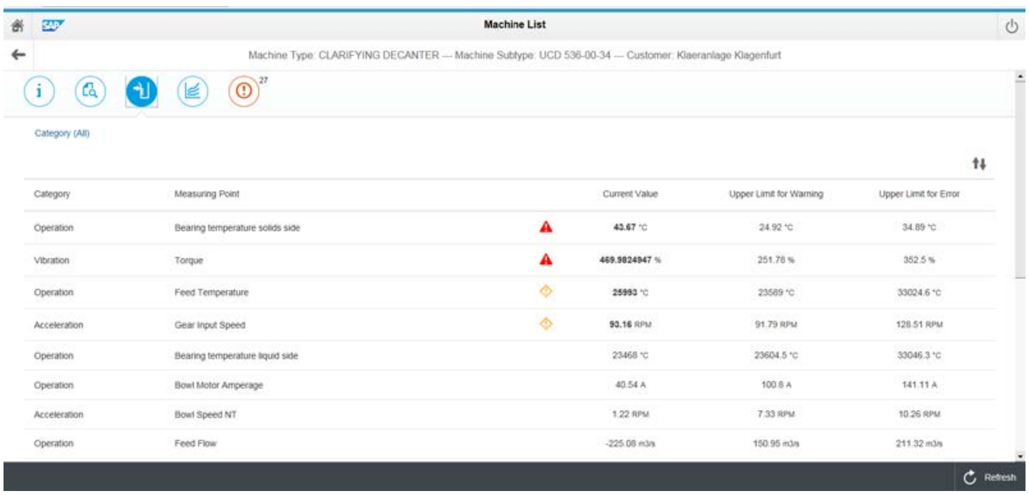

Figure 5: SAP Current Operational Data with limits

Within the maintenance world, data analysis leads to operational improvements and condition based maintenance. At its basic, when a variable goes out of control, maintenance is notified that there is a problem that needs to be addressed. However, when you start looking at variables that might be going out of control (e.g. trending out of control) a company is entering the world of predictive analytics. As companies transition from triggering an event based on a simple violation of a control to triggering the event based on trends or predictions, more data and variables become part of the analysis. Eventually, additional techniques are needed to address all these variables that are being analyzed.

As companies become more sophisticated in their analysis, they progress towards a more predictive modelling process using multi variant analysis. Multivariate Data Analysis refers to statistical techniques used to analyze data that arises from more than one variable. This allows one to essentially model reality where each situation, product, or decision involves multiple influencing variables. Building a model of the situation allows current operation data to be fed into the model so that predictions of failure, or control limit violations can be created and acted upon.

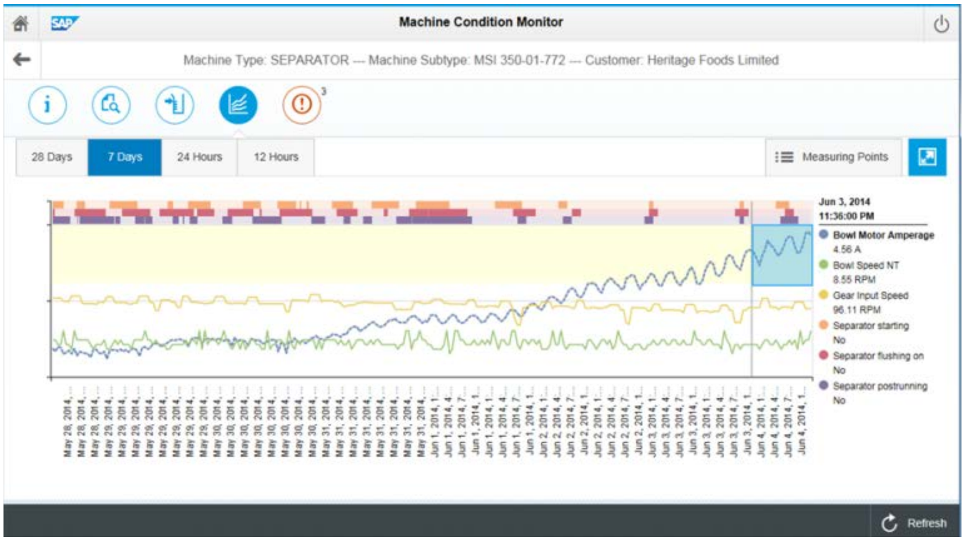

Figure 6: SAP Machine Health Monitoring

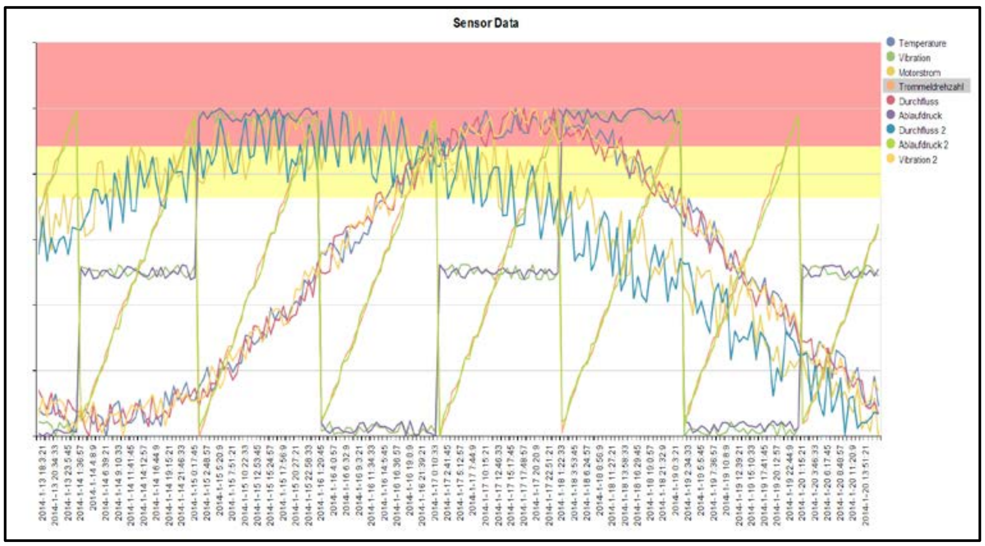

Figure 7: SAP Real Time Analysis

As with any analysis, care must be taken when selecting the variables and how much data needs to analyzed. Current tools allow people other than data scientists to identify potential causality between variables and events, apply sophisticated analysis tools, and to generate statistical models that can be embedded in monitoring applications. This enables the model to run with current operating data and to use it for predictions.

Generation X - Mobility

Many elements of master data can be validated and collected by the mobile workforce during the course of day-to-day activities. Plant maintenance workers are frequently affected by inaccurate or incomplete asset master data when maintenance work orders are based upon bad information. As a result, the maintenance workers arrive at a job with wrong work instructions or repair parts. Mobile automation tools will allow the maintenance worker to initiate a change request on-the-spot which will kick off an auditable process to remediate the problem and insure that the inaccurate information is corrected.

Mobile workforce tools will also enable incremental inspection/validation of structured and unstructured master data as field operations are performed. These field reviews might include digital redline mark-ups of engineering drawings and diagrams, or recommended improvements to maintenance or inspection procedures.

Perhaps the greatest advantage of mobility for improving data quality is the application of RFID's to Functional Locations or Equipment. Frequently, operations personnel can only describe the issue at a functional location. An example of this is, "The pump does not start". This could be an accurate assessment made at the field or it could be a misstatement of the truth "The pump is not supplying the required (or any) flow". The operator may or may not perform basic trouble-shooting depending on culture and workload. At this point, there are numerous opportunities for the work to be misdiagnosed. If the failure is, "The pump is not suppling the required flow," this could be caused by: power loss, stop/start switch failure, motor starter problem (including fuses), wiring issues, motor issues, coupling issues or pump problems. The problem could be mechanical or electrical. A best practice solution to this might consist of a work notification /order written to the functional location requesting that the coupling be broken to check the pump and motor for proper functioning as shown below:

WO 123456 P-256 Pump is not flowing ➔Functional Location Identified by Operations
EQUIP 1211487 Electric Motor Fuses Blown and Motor Bearing Bad

In this example, the work notification/order pair was assigned to an equipment item installed in the field with a malfunction. This was made simpler by the attachment of a RFID to the Electric motor and communicating that the electric motor included all electrical parts (wire, motor starter, stop/start station and motor) as a part of the motor.

Conclusion

This paper has simplified the spectrum of data collection and management systems into three technology based generations in the hopes of providing vision to organizations trying to improve their data collection to improve safety and reliability. It is the belief of the authors that in the last few years, technology has advanced sufficiently to make it possible for owner/operating companies to use these generations in their quest for the above goal. It is recognized that improving data quality is a journey of many years and that there are few articles in the public domain about the competitive advantage of data quality improvements.

We also see a tread in representative trade organizations to begin to increasingly stress the need for deconstructed data to meet business objective. We believe the changes described in the paper will occur in the coming years and that this paper will assist you on your data improvement journey.

Appendix A: Generation 1 – Improving the completeness and accuracy of ERP System Master & Transactional Data

Data Maturity/ Data Governance/ Themes	• Data unstructured or structured in multiple systems to a solution where the multiple systems can be viewed. • Management of change for engineering data and documents for as-built facility may or may not be completed (or completed in a timely manner) depending on the asset type (example ... upgrade transmitter material of construction) in the Engineering Data System and/or Enterprise Resource Management System. • Data Governance is a Manual Process - Data is cleansed and validated manually prior to load. • Training on the importance of data governance has begun. • The concepts of Functional Location and Equipment designation are being implemented into EAM.
Characteristic of the Generation	• The system that holds the master data to manage change is un-unified. • Change may be only recognized when the field equipment does not match the Maintenance System Master Data or non-integrated Test Forms. • Data transfer between Engineering and Maintenance is difficult. • Little reconciliation of Engineering Master Data with Maintenance Master Data. • Change is recognized when the equipment does not match the Maintenance System Master Data.
Engineering and Maintenance Master Data	• Appropriate Functional, Locational, and Equipment Data is in the ERP to support Reliability and Mechanical Integrity efforts. • Engineering Data: Implemented/implementing as-built. engineering data management system with data transfer to SAP. • Maintenance Data: Engaging the work force to validate ERP Master Data as part of work execution. • Extract engineering information from different sources (Shared drives, papers, and database).
Maintenance Master Definition	• Engineering data is both paper and electronic depending on age of system. • Manual processes exist for paper updates. • Engineering Master Data transfers utilizes manual extract, transform (including cleansing, quality control and validation processes) and loading.
Requirements	• Manual effort to collect data typically through paper based systems leveraging physical documents and brute force labor.

Challenges	• People, paper, and definition of data to be collected. • Challenges include validity and cleanliness of data. • Potential for incomplete data collection not recognized until team leaves the field. • Transcription of data into target environment, time consuming, but can use relatively unskilled labor.
Business Process Benefits	• Benefits include collection of missing information and possibility of leveraging existing operational resources to do data collection on an incremental basis while in the field.

Generation 2: Data Supports Maintenance Effectiveness and Drives Equipment Reliability Improvements

Data Maturity/ Data Governess/ Themes	• Low cost - Data Maturity +. • Electronic PID and instrumentation information controlled by processes that ensures quality and accuracy (Managed P&ID and instrumentation database). • Engineering and Maintenance Master Data are consolidated. Which system is the master or "Golden" system for which data is defined? A process exists for periodic comparison of fields for disparate data. • MOC is has engineering data integrated in a fashion than improves effectiveness from the basic approach. • Data Governance has been implemented into the work process. Master data is validated when work is performed. • A process existing to ensure work performed is assigned to the correct Functional Location and Equipment. Mobility & Simplification (Digital in the field) improve worker involvement and satisfaction with documentation process. • Quality Work Performed records are part of the basic job expectation for all employees. • RFID's ensure that work performed records are assigned to the correct location.
Characteristic of the Generation	• A unified view of data exists with the master system defined key information sets for Engineering and Maintenance Master Data. • Master data from the system of record. • Data transfer between Engineering and Maintenance is electronic. • Validation of Maintenance Master Data is part of the organization culture with follow-through to update Engineering Master Data.
Engineering and Maintenance Master Data	• Lowest cost, easiest to implement data+ • Electronic Engineering Drawings & Specifications are available for use in the field. • Engineering Data: Implementing detail engineering design to feed the as-built engineering data management system with

	reporting to SAP. • Maintenance Data: Maintenance Master Data is supplied from Engineering Master Data.
Maintenance Master Definition	• Work is in progress to develop unified picklists shared by both Engineering and Maintenance Systems. • Setup engineering tools and synchronize with engineering management system for as built facility. • Extensive engineering data gathering is integrated into Maintenance Master Data System. • Perform engineering design in specialized tools enforcing quality control and standards.
Requirements	• Utilizes mobile solution with pre-scripted data collection tasking as determined by upfront data profiling. • Data compiled in the field and sent back to central processing team. • Central processing team manually or systematically loads data into target environments. • Allows for activities like RFID tagging, digital redline markup, and geo-locating using mobile devices.
Challenges	• Data profiling tools, offline/online mobile devices, skilled resources.
Business Process Benefits	• Benefits include early identification of problem areas through profiling, pre-scripted system allows for prompt to technicians.

Generation 3: Data supports Product Reliability Improvements

Data Maturity/ Data Governance/ Themes	• Medium cost - Business Case +. • *Reduction in safety and reliability events. • *Automated data transfer between engineering, maintenance and inspection systems. • *Reduction in work process cycle time.
Characteristic of the Generation	• Mobility & Simplification. • Internet of things. • Predictive Analytics.
Engineering and Maintenance Master Data	• Lowest cost, easiest to implement data+. • Moderate cost with data transfer +. • MoC systems ensure Engineering Data is evergreen. • Master Data is sourced from Engineering Records. • MoC systems ensure maintenance master data is evergreen. • Engineering and Maintenance Master Data. • Master data management processes are in place and supported by the culture. Data management systems are fully integrated with maintenance and inspection tools and workflows.

Maintenance Master Definition	• Data analysis and mapping with full integration to maintenance and inspection systems.
Requirements	• Fully connected mobile solution. • Work lists automatically populated to the workers in the field. • Data collection by mobile device. Includes validation of data and governance workflow to kick off change requests in Engineering and Maintenance systems whereas built doesn't match "digital equipment data." • Includes GEO positioning, RFID tagging, digital redline markup or potentially "smart connected monitor" installation or ongoing maintenance activity.
Challenges	• Challenges - most costly, includes several interconnected systems, additional skilled resources..... • Skilled resources, connected mobile devices with speed to collect and transmit data, workflow driven Data Governance engine (MDG for EAM).
Business Process Benefits	• Benefits include governance at the point of collection and entry, automated workflows, real-time feedback and validation provided to technicians in the field and at command center, ongoing visibility to information from IOT sensors.

Bibliography

Wikimedia Foundation, Inc. (n.d.). https://en.wikipedia.org/wiki/Data_governance. San Francisco, CA, USA. Retrieved 09 13, 2015, from https://en.wikipedia.org/wiki/Data_governance

American Institute for Chemical Engineers. (2014). *Vision 20/20.* American Institute for Chemical Engineers.

David W. De Long - Accenture Institute for Strategic Change. (2002). *Confronting the Chemical Industry Brain Drain: A Strategic Framework for Organizational Knowledge Retention.* Cambridge, MA: Accenture Institute for Strategic Change.

https://www.techopedia.com/definition/28799/data-intelligence. (n.d.). Data Intelligence. Retrieved 09 14, 2015, from https://www.techopedia.com/definition/28799/data-intelligence

Intergraph Corporation. (2005). *Plant Information Management.* Huntsville, AL: Intergraph Corporation.

ISO copyright office. (2006). ISO 14224: Petroleum, petrochemical and natural gas industries — Collection and exchange of reliability and maintenance data for equipent. Geneva, Switzerland.

Kletz, T. (1993). *Lessions From Disasters: How organizations have no memory and accidents recur.* Houston, TX: Gulf Publishing Company.

McDonough, J. (n.d.). Multi-Generation Project Planning’s Fit with Six Sigma . Retrieved 09 13, 2015, from http://www.isixsigma.com/implementation/project-selection-tracking/multi-generation-project-plannings-fit-six-sigma/

Wikimedia Foundation, Inc. (2015, 09 13). System configuration. San Francisco, CA, USA. Retrieved from https://en.wikipedia.org/wiki/System_configuration