Blog 10:Avoiding the Data Swamp in Asset Management

(Getting Ready for ISO 55000 – Part 10 of 10)

Insights from the "Asset Management for the 21st Century - Getting Ready for ISO 55000" Seminar, May 2013, Calgary: *This blog is based on a series of interviews with John Woodhouse from The Woodhouse Partnership (TWPL), who delivered this well-received seminar. John Woodhouse is CEO and Managing Director of asset management consulting firm TWPL, is a founder member of the Institute of Asset Management. He chaired the development of the PAS 55 standard and is UK Principal Expert in the development team for the ISO 55000 standard.*

In the not so distant past, most enterprise applications were based on a range of data that was heavily constrained by what was collectable. It was hard to acquire data and ensure it was of high quality (and it still is for some types of data!). Now in the modern world, with the Internet of Things and the ability to collect data from a variety of different sources, automatically in many cases, you can easily end up with overwhelming amounts of data. But Big Data can often result in Big Confusion.

Some of the data has a natural home, such as asset registers and technical records. Some can be distilled, analyzed and converted into useful management information. But a large amount of it falls under the category of 'it might be useful one day'—a large, often unstructured mix of activity records, asset performance and condition attributes, sometimes having localized or temporary usage but often collected just because it is now easily collectable.

The real challenges therefore are to understand what data is *worth collecting* in the first place, and why (how it would and should be used). Then we have to put it into organized repositories that are more like a library and less like a swamp. Here are some ideas on how you make sure that you can store more relevant data, with clearer understanding of why it is needed, without it becoming a messy liability that is neither used nor trusted.

First, think of data as part of a demand-driven supply chain in which justification for collection, retention and usage has to be made from the business risk or cost of *not* having it (to the appropriate standard at the right time). The apparently low cost of acquiring data and the motive that 'it might prove useful' are not enough to justify collection and retention. This bucks the trend of data provision being seen as an availability-driven process that triggers a search of usages. Demand-driven thinking requires greater understanding of how the data will be used, selective extractions from it, and what business value is achieved from using it.

The SALVO Project, a multi-industry R&D program to develop innovative approaches to asset management decision-making, has yielded good examples of this approach. Three of the necessary six steps in the SALVO decision-making process illuminate the demand-driven data specification. Step 1, "Identify problems and improvement opportunities" spells out the business impact criteria for which assets need what attention in the first place, and the desirable evidence to support this identification. This includes definition of asset health indices (relevant mix of performance and condition features) and criticality measures.

Step 2 is the drill-down into the identified problems or improvement opportunities to ask *why* they are problems (root cause analysis). This often reveals a mismatch between expectations and realities in the use of data to demonstrate patterns and correlations. The noise in the system, the inherent limitations of data samples, and the volatile business environments in which data is collected (including consistency of collection method) mean that pattern-finding or 'non-randomness' is rarely provable, irrespective of the clever data analytics that are applied. Except in very rare cases, the available data will normally be constrained and 'censored' in various directions, so the collectable evidence needs to be used with great care—and with a healthy dose of realism and 'tacit knowledge' from asset design, operations, and maintenance experts. Step 3 of SALVO covers the selection of potential actions or interventions, and these can be a far wider range of options than the technical tasks normally considered (such as inspection, maintenance or renewal). SALVO has identified 42 practical options that might be applicable to solve asset management problems.

Step 4 then covers the business value-for-money evaluation of the potential solutions, requiring assumptions and, if obtainable, evidence of costs and short-term and long-term consequences. This step combines observable facts (mostly helpful in quantifying the 'do nothing' implications) with external data needs and the tacit knowledge of the experts in forecasting and estimating the degrees of improvement that might be achievable. This is a stage where reliance on collectable hard data is fairly limited, but at least we can be clear about the questions that need to be asked (that is, what data is desirable to support the decisions). SALVO has mapped the information needs for all 42 common decision and intervention types — the information required to determine if the interventions are worthwhile and, if so, when. For example, there are 13 specific questions or data elements that must be considered in deciding whether to buy a critical spare part and how many to hold. These decision-specific checklists help to focus on the *relevant* and *useful* information within the background swamp of confusing evidence. They, and a 'what if?' approach within the evaluation process, reveal the role of the data to support decisions. They demonstrate the business value of collecting the right stuff, by quantifying the 'cost of uncertainty' when forced to rely on range estimates or assumptions.