Murphy, Glen D. and Chang, Artemis and Barlow, M. (2008) Leveraging Engineering Asset Data: Strategic Priorities, Data Types and Informational Outcomes. In Jinji, Goa and Lee, Jay and Ni, Jun and Ma, Lin and Mathew, Joseph, Eds. *Proceedings 3rd World Congress on Engineering Asset Management and Intelligent Maintenance Systems Conference (WCEAM-IMS 2008): Engineering Asset Management – A Foundation for Sustainable Development*, pages pp. 1136-1365, Beijing, China.

Proceedings of the 3rd World Congress on Engineering Asset Management
and Intelligent Maintenance Systems (WCEAM-IMS 2008)
Beijing, China
27-30 October, 2008

LEVERAGING ENGINEERING ASSET DATA: STRATEGIC PRIORITIES, DATA TYPES AND INFORMATIONAL OUTCOMES

Murphy GD[a], Chang A[b] and Barlow M[c]

[a] *CRC for Integrated Engineering Asset Management (CIEAM) 2 George Street Queensland University of Technology, 4000, Australia*
[b] *Faculty of Business, 2 George Street Queensland University of Technology, 4000, Australia*
[c] *Australian Nuclear Science and Technology Organisation* (ANSTO), *PMB 1 Menai NSW 2234 Australia*

A common complaint heard within the engineering asset community is that while the capacity for data storage increases, the quality of ever increasing amounts of data remains poor. We propose a new model of engineering asset data management that helps explain why data collected by organizations frequently fails to assist in effective engineering asset management. The model situates a four component typology of engineering data between institutional drivers (e.g. organizational culture; organizational strategy; organizational life-cycle; consequence of asset failure) and asset management outcomes. We argue these outcomes (regulatory compliance; time-based maintenance; condition-based asset management; capacity development) are functions not only of the data collected by an organization, but its capacity to leverage that data. We develop a model suggesting that institutional drivers dictate the data requirements of engineering asset intensive firms, typically at the cost of data requirements for different phases in the asset's life-cycle. This paper will assist practitioners to re-conceptualize the manner in which they view their data, the manner in which it is utilized, and provide a better understanding of data and its intended outcomes. This will allow a better prioritization of data collection activities and offer an improved insight into ways in which engineering data may be better transformed into informational and knowledge outcomes.

Key words: Data Quality; Organisational Lifecycle; Data typology

A brief review of the practitioner and academic literature reveals extensive discussion based around the quality and hence usability of information required by organizations. For example, numerous publications lament that despite an exponential increase in the capacity for data storage, data quality remains poor [1]. Potential antecedents of data quality include systems integration/migration problems, age and redundancy, relevance and the temptation to collect data '*just because we can and we might need it someday*" [2,3]. Others suggest issues with an organization's capacity to source technical personnel with the requisite skills [4]. These however tend to focus their level of analysis at an operational level and consequently are unable to address some of the broader organizational and strategic level considerations that may have a direct impact on the type of data collected and the extent to which it is leveraged into informational outcomes.

In this paper we suggest that attention shift towards the manner in which organisational and contextual elements determine organizational data acquisition priorities, and ultimately their informational needs. We argue a significant step forward in the area of engineering asset information quality is the development of a framework emphasizing the value of adopting a strategic approach to the prioritization, acquisition and transformation of data into informational outcomes. Central to our argument is the idea that data often fails to result in usable information due to a clash in intentions from a managerial and engineering perspective. We suggest that engineering standards, OEM requirements and accepted technical practice dictate data acquisition (type, amount and frequency) but strategic priorities and resource limitations dictate data use (regulatory compliance, time-based maintenance, condition-based maintenance). A strategic model to facilitate alignment of the data acquisition processes with data use requirements and priorities will therefore improve quality of data and ensure effective utilization of data.

1 DATA MATURITY MODEL

We propose a new typology of engineering asset data in which we suggest that drivers such as asset value, asset life-cycle and the consequences associated with asset failure dictate the resultant or desired informational outcomes. Figure 1 illustrates the proposed theoretical model titled the "Data Utilization Maturity Framework". The model has three distinct components.

Located at the front end of the model are the "contextual" ***data drivers*** (e.g. organizational and asset life cycle, asset value, asset complexity) that may determine the priorities how much, how often and what type of data are collected and used by an organization. The centre of the model is concerned with ***data acquisition***, articulating a generic framework of engineering asset data types of concern to any engineering asset intensive organization. Self evidently the extent to which informational outcomes are achieved is dependent on the type of data collected by organizations. We assert that engineering asset data comprises of four main categorizations of data relating to a) *Configuration and baseline dat*a, b) *Asset condition*, c) *Event or incident data* and d) *Process data*. It is suggested that each of these data types can potentially relate to a number of data outcomes including a) *regulatory compliance,* b) *time based asset management,* c) *condition based asset management* and d) *capability development*. This final component of the model therefore is concerned with ***data utilization*** and the extent to which an organization is capable of leveraging the data it has collected into usable informational outcomes. A key contribution of this model is the manner in which it highlights that these outcomes are functions not of the data types collected by an organization, but its data utilization maturity - its capacity to use data at their disposal in an increasingly sophisticated manner. This is reflected in Figure 1 below.

In an important step forward in the data quality debate the model explicitly realizes the critical role played by organizational context, in particular organizational life-cycle, the nature of the assets under consideration and the contribution of the asset to organizational function and success. We suggest that for example an organization currently engaged in the design/build phase of an asset would typically devote adequate resources to collect data aimed at ensuring regulatory or contractual compliance. However the extent to which additional resources are allocated that allows the exploitation of data to achieve condition based maintenance during the operate/maintain phase is dependent on a broad understanding of the vital role data plays in determining asset performance. This argument is discussed further in the next section, dealing with data drivers.

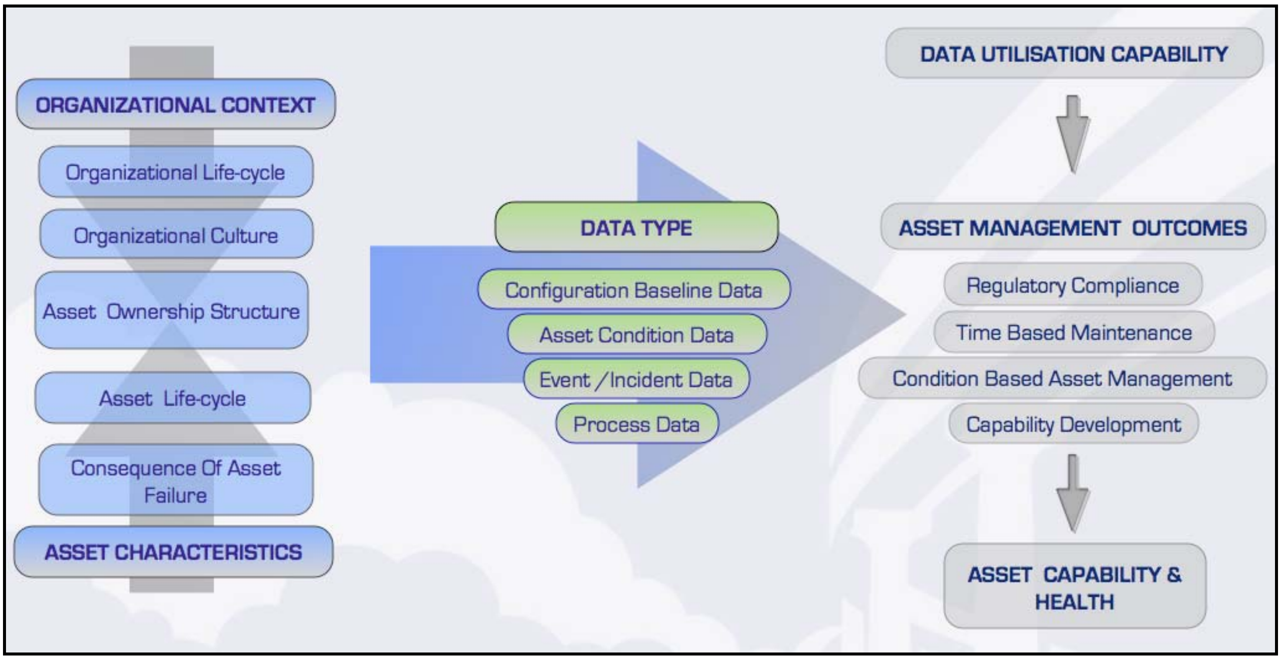

Figure 1.0 Data Utilization Maturity Framework

2 DATA DRIVERS

Within any one organization there are likely to be a range of operational, political and resource based constraints/enablers dictating data acquisition activities and informational outcomes. In broad terms however we argue there are two primary categories of drivers that dictate the resources and priorities devoted to the collection and management of engineering data. The first relates to ***organizational context*** - aspects such as organizational life-cycle; culture and strategy. The second relates to ***asset characteristics,*** suggesting that aspects such as asset life-cycle and consequence of asset failure dictate the level of resources allocated to the collection, storage, analysis and utilization of engineering data.

2.1 Organisational Context

Refers to the operational environment within which asset maintenance management activities are conducted. Factors in this environment critical to the type of data include (but are not limited to) industry and business strategies, organizational culture and organizational life-cycle. Two of these are discussed briefly below.

Organizational Life Cycle: Organizations evolve in a steady and predictable manner from the initial start-up, expansion, to maturity and decline [5]. An organization's life cycle stage can influence what resources and infrastructures are available to collect asset data, but can also determine the priority placed on different asset outcomes. We argue that regulatory compliance is a key focus during the start up stage, during which limited configuration and asset condition data is available and is the target of data acquisition efforts. As an organization transitions into expansion and maturity phases, more resources begin to be allocated to asset outcomes such as condition based maintenance and capability development. A simple diagram of the impact organizational life-cycle can have on data priorities is presented in Figure 2 below. The emphasis placed on each outcome, and therefore data type collected depends on an organization's business strategy and its capacity to utilise the data collected (discussed in section 3 below).

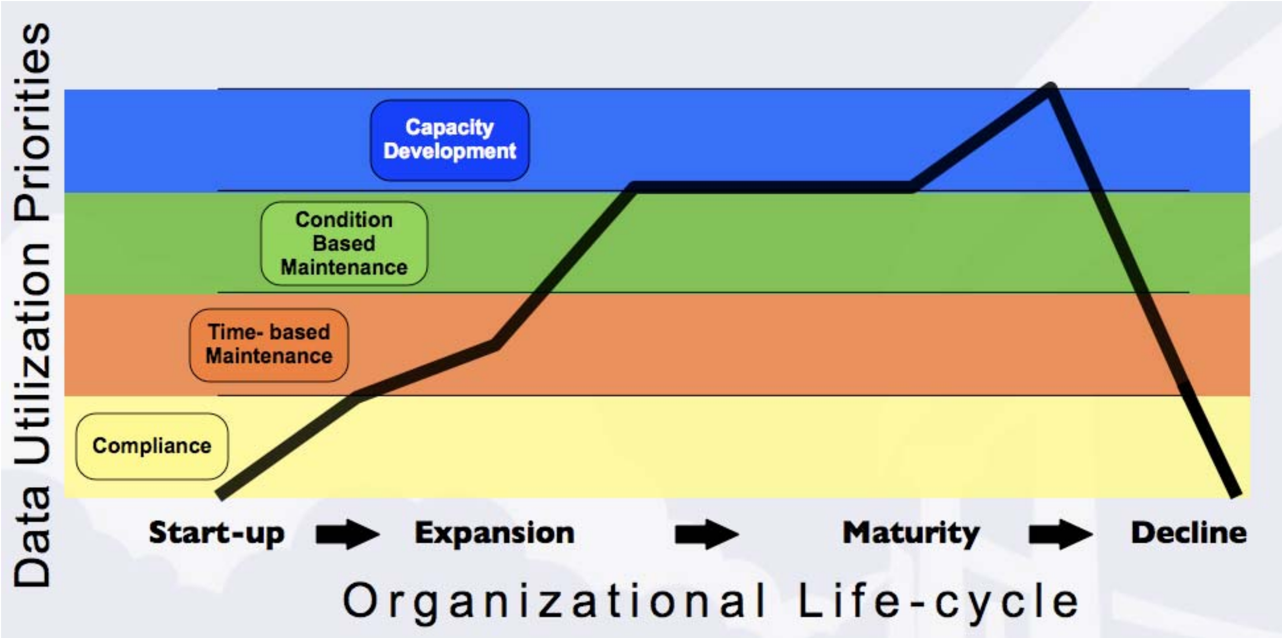

Figure 2 Organizational Life-cycle effects on data utilization priorities

Organizational culture: Organizational culture is typically defined as the set of basic assumptions about the functioning of an organization shared by the majority of employees that drive their perceptions, attitudes, feelings and behaviors [6]. The presence of strong embedded cultures within organizations can have a significant impact on the way in which the issue of engineering data is treated. Certain cultures may place a higher premium on the value of data and its importance to the success of the organization. For example Detert et al. [7] observe that Total Quality Approaches to management "embrace an approach to truth and rationality represented by the scientific method and the use of data for decision making "(p853). Quality and Six Sigma cultures by their nature therefore tend to be data driven and as a consequence are more likely to devote far more resources to the acquisition, storage and utilization of engineering data.

2.2 Asset Characteristics

At an operational level we argue that there are three primary drivers of information and data maturity within the majority of engineering asset intensive organizations - Asset value; Asset life-cycle; and the consequences of asset failure[1].

Asset life-cycle: As to be expected, each phase of an asset's life-cycle is likely to determine a different set of data and informational requirements. However our model suggests that those organizations failing to take a strategic approach to their data acquisition and information generation activities are likely to limit their data collection activities to specific short term goals at the expense of future requirements - this is most likely in the early "build and design" phases of an asset's life where efforts are focused on delivering the asset ready for service and complying to any contractual or regulatory requirements. We suggest that a failure to recognize the importance of historical data captured and stored in a usable fashion limits the ability of an organization to later engage in more mature data usage activities such as condition based asset management or organizational capacity building.

Asset Ownership Structure: Refers to the business structure in which the assets are operated and the effect this has on asset utilisation, maintenance and improvement. The ownership structure affects the management of assets, which in tern determines the type of data collected and the expected outcomes. For example, there is substantial difference between assets owned by business and the community in terms of data type collected and expected outcomes. The authors have observed that as public participation increases, data acquisition efforts focus on shorter term issues, with a consequent loss of possibly more significant data which would support strategic decisions and outcomes. This situation has been observed in government and public companies. In cases where the community is the asset owner, with the local council as steward, the focus can be on all three aspects of asset safety, reliability and performance, for the primary intention of being seen to be a good asset steward and not attracting criticism. In addition, the data itself may seem totally unrelated to strategic management. Parks and reserves are good examples, needing to be safe for children and at night, and to be suitable for the local football game

[1] We limit our discussion here to a singular "asset", acknowledging that "not all engineering assets are created equal", and that the nature of some assets will attract different forms of attention. However, it is possible that these drivers relate to a "class" of assets or apply universally to all engineering assets within an organization and reflect an overarching attitude to the acquisition, storage and usage of engineering data.

every weekend. Therefore data collection will have to provide answers to comments such as "We need more lights because the crime rate is too high", or "The grass is too long. When did the council last mow the field?".

Consequence of Asset Failure: This driver has particular relevance in an engineering asset context given the typical costly and potentially significant safety consequences of engineering asset failure. We suggest that the more significant the consequence of asset failure - whether that be in the case of asset repair or replacement costs, loss of revenue due to reduced service delivery, or health and safety risk, the greater likelihood that organizations will place a greater emphasis on greater utilization of engineering data - particularly in relation to performance based and capacity improvement information. Two case examples of "failure consequence" are provide below to further illustrate the point.

Case Example 1: Fault Analysis System. A project at a telecommunications company involved the development and introduction of a fault analysis system applied to a large diverse fleet of vehicles, trucks and earth-moving plant. The system required the repair assessors to detail the operations and times involved in repair and maintenance, with the main objectives being the capture of recurrent faults for later analysis. These "estimate sheets" included codes to identify the type of operation, e.g., "remove and replace", and the component, e.g., "radiator", as well as the time allowed for the operation. Significant technical effort was invested in the design of the database, equipment was purchased and the assessors and data entry staff trained. In this instance "*consequence of asset failure*" was a significant driver in this initiative. The company used the fleet, amongst other things, to install long stretches of communications cable, and failure of the asset item resulted in stalling of the job, with consequent loss of production and revenue. The key data type in this case was "event and incident" information (discussed below), with particular relevance to preventing future failures of the same type.

Case example 2: Open Space Assets. At a large western Sydney council, a requirement was identified for a database to record and report upon the open space assets which included parks, reserves and playing fields. The purpose of the database was to provide ready access to mainly static information to respond to residents' enquiries. An unusual aspect was the tendency for residents to make general enquiries, then change the enquiry, so that an important requirement was the ability to present information, then re-query the database according to the new enquiry. In this case the data driver could be classed as "Asset Value". As well as being valuable in terms of property prices, the community value is probably inestimable due to the intangible aspect of community involvement and use, lifestyle, quiet enjoyment and "green space". The implementation and continued use of the database was a success for these reasons:

1. The requestor had a very clear view and a complete specification of the structure and outputs of the database.
2. The data was drawn from the corporate database, which was accurate but had limited reporting facilities.
3. The requestor was to be a main user of the system.

3 DATA ACQUISITION

While the specific kinds of data that might be collected by any one organization for any range of assets might number in their thousands each of these can be classified based on the function they are intended to serve. When considering the form and function of engineering data we argue there are four key types of data that engineering organizations are interested in acquiring. This section explains each of these four data types in more detail. It should be noted however that acquiring and even storing the correct data does not ensure it is utilized effectively, this is dealt with in more detail in section 4 of the paper.

Configuration data: Refers to data that describes the ideal state of the asset as it was designed to comply with the end-user's requirements and conform to regulatory requirements. Typically this data is provided by those responsible for the design and /or manufacture of the asset. While configuration data may typically begin life as OEM related information, periodic enhancements, upgrades and other events such as hazard assessments may inform the configuration data over the life-cycle of the asset. The close links between this data type and regulatory compliance requirements ensure that configuration data is often given a high priority throughout the asset's life-cycle. However anecdotal evidence suggests that not all relevant configuration data is always available or provided to an adequate standard prior to the asset's operational deployment.

- Typically originates from OEM related asset data
- Informed by periodic enhancements & upgrades
- Hazard assessments requiring configuration changes
- Used to provide benchmark comparisons with condition data

Condition data: Relates to data concerned with the measurement, recording or documenting the existing condition of the engineering asset (e.g. calibration data, asset health related data). This data type typically involves simple, routine data collection activities carried out as part of planned maintenance routines. However asset condition data may be captured at different times by different populations within the engineering asset community (e.g. operators, maintainers, engineering staff). This data is recognized as highly relevant and useful by those tasked with maintaining and operating the asset.

- Used to confirm compliance with regulatory requirements
- Used to ascertain asset health
- May identify the need for reactive (unplanned) maintenance
- Can be used for the trending of asset health

It appears that in order to be useful this data type requires accurate and complete historical records for sophisticated trending applications. We suggest that the issues observed with the collection of high quality data may lie in the different priorities/value placed on the data by the groups collecting and consuming the data. Given the important role of condition data and the diverse set of capabilities represented by the various populations tasked with capturing the data it appears that ease of use and efficiency of data capture are important elements to consider.

Event & Incident Data: Principally refers to data captured following asset failure and documents the nature of the incident, the underlying causes (i.e. root cause analysis) and actions taken to rectify the asset failure. By its nature event data can be highly variable in the quality and quantity of data provided, particularly if the data is acquired manually. It is suggested that Event & Incident data quality is directly influenced by the data maturity of the organization and the priority it places on data required for predictive asset management.

- Used to identify appropriate actions to reinstate the asset back to its ideal state / operational state (component focus)
- Used to identify appropriate long-term strategies to prevent future asset failures of this type (system focus)
- Used to inform predictive asset health systems
- Used to improve future design enhancements

Process Data: Refers to a diverse collection of data documenting resources used and required during maintenance activities or data relating to work instructions or business process improvements. Typically this information is recorded as part of a maintenance work-instruction or work-package issued in response to a planned maintenance request or asset failure. Examples of the data collected by operators, maintainers of technical personnel include: time taken to complete job; tools required; safety issues not already documented; observations noted during inspection; additional or redundant steps required/not required; feedback on the quality and efficiency of the work instruction; and anecdotal comments regarding the asset. Like “Event & Incident” data the quality and quantity of process data recorded by organizations is largely determined by the data maturity of the organization. High quality process data is more likely to be observed in those organizations with well developed predictive asset management regimes and a focus on organizational capability improvement.

- Used to accurately determine asset management requirements
- Used for scheduling, workforce planning and material management
- Used to revise work instruction and safety hazard documentation
- Used to drive business process improvements
- Used to capture tacit knowledge

4 DATA UTILIZATION

Engineering asset organizations collect asset management data from multiple sources for multiple reasons of which four are applicable across all industry types. We represent these four data outcomes as a continuum of data use maturity from the most basic, superficial data collection and use, to the maximum exploitation of data for multiple uses, not just for current assets but the design, management and utilization of those in the future. This typology of outcomes reflects the broader concerns of engineering asset management while incorporating the four typical maintenance methodologies described by Tsang [8] evident within engineering asset intensive organizations.

Regulatory / contractual compliance: In many instances the consequences of engineering asset failure dictate a level of regulatory compliance for most engineering assets. If not governed by regulatory compliance asset condition can also be driven by contractual obligations either by the OEM or the asset owner as the service provider. We suggest that data collected in order to satisfy regulatory or contractual obligations is often an organization’s first priority and the data most likely to be used, especially during the early phase of the asset’s life-cycle. The outcome is heavily dependent on the quality of the configuration / baseline data provided by the OEM or collected by the organization.

Time Based Asset Management: Refers to institutionalized, reactive or planned maintenance where data collected is only used to maintain the current condition of the asset. Tsang [8] refers to preventative maintenance, where items are replaced or returned to good condition before failure may occur. Time based or preventative maintenance is typically driven by OEM maintenance stipulations or established planned maintenance routines. Organizations engaging in time based AM are either unable (e.g. lack the resources, personnel, or structural capability) or unwilling (e.g. competitive strategies may require the asset be “run to failure“ and be disposed) to engage in condition based asset management. Data used during sustainment maintenance may be sourced from configuration data, asset condition or event data.

Condition Based Asset Management: This asset management outcome largely relates to maintenance regimes such as Condition Based Maintenance (CBM) that rely on sophisticated predictive modeling to determine maintenance schedules. This capability represents a high degree of data maturity as this type of predictive modeling requires accurate, timely, reliable longitudinal data to not only be collected, but used. We suggest that asset age, asset replacement cost and availability are likely determinants of whether an organization is willing to expend the resources to engage in performance based asset management.

Capability Development: Refers to the use of data to improve the design, development and manufacture of future engineering assets or ancillary processes (maintenance routines, safety procedures). While not an outcome desired by all organizations the use of data to inform future developments is considered a valuable use of engineering asset data and reflects a high degree of data utilization maturity.

A key factor in the extent to which an organization realizes the benefits afforded to it by increasingly mature data utilization is the extent to which the organization is capable of managing the data required to achieve increasingly sophisticated outcomes such as condition based monitoring and capacity development. The elements determining the capability of an organization to utilize its engineering data depends on a range of factors, the discussion of which lie outside the scope of this paper. Examples of those elements include; the level of IT systems investment, the level of systems and functional integration, the quality of the data being used, and the level of training and technical competence of those responsible for data acquisition. It should be noted that not all organizations will desire or value the achievement of mature data outcomes such as condition-based maintenance or capacity development. For example, a hire-car company which replaces its fleet every 24-36 months is more likely to be focused on regulatory compliance and time-based outcomes (e.g. basic sustainment) than investing in long-term condition based maintenance regimes.

5 CONCLUSIONS

Anecdotally the authors have observed that three main reasons are often cited as to why organizations fail to achieve the informational outcomes they desire from their data acquisition efforts. The first and most often reason is that the data is "poor", that it lacks the requisite integrity in terms of completeness, accuracy or timeliness to be of any use. The second most commonly cited reason is that the data needed is not collected, for any number of reasons such as lack of infrastructure, resources or analytical capability. The final reason is that, ironically that there is "too much" data, that increased and more cost-effective storage capacity has resulted in an indiscriminate capture of any and all data. This appears to have the unintended consequence hindering the analysis of critical data, as poor data storage for example prevents access to the data and increases the need for costly activities such as data cleansing and refinement. The key theme within this paper is that much of these issues can be traced back to the lack of strategic data management within organizations. We argue that organizations wishing to improve the informational outcomes they seek from thier data will have a greater appreciation of the strategic and contextual drivers of what data should be collected, how it will be collected, stored and used. In summary this paper achieves three outcomes:

One, It provides a universally applicable *typology of engineering asset data types*. The model presented in Figure 1 has its origins in a model developed to begin the conversation about engineering asset data across organizations and industries without worrying about the technical aspects of data specific to each organization. Having a common language and universal reference points for the types of data and information outcomes expected enables cross-functional and cross-organizational comparisons with greater clarity and an improved understanding of what engineering data organizations are interested in.

Two, it provides a universally applicable *typology of engineering asset informational outcomes* that any one organization may wish to achieve - this allows engineering data researchers the ability to compare across organization and industries how data is used and exploited by organizations. Developing a well informed typology of data types and informational outcomes should also assist in the development of generic guidelines aimed at improving overall engineering data management.

Three, it provides a tool for engineering asset managers to articulate both up-stream and down-stream the type of data needing to be collected and why, providing a clear linkage between engineering decision making and desired organizational outcomes. As discussed we suggest that organizations begin to adopt a more strategic, longitudinal approach to their data acquisition, storage and usage priorities. Most acknowledge that the typical approach to "collect anything and everything" is an inadequate form of data management and rarely results in desired informational outcomes [6]. Instead we advocate efforts to better understand the aims of data acquisition and the uses of that data in the long and short term.

The framework presented here represents the beginnings of an attempt to investigate the strategic, organizational and contextual drivers of engineering data and its outcomes. A number of opportunities are open to data quality and engineering asset management researchers to take this work forward. One area of obvious interest is to investigate in more detail the potential drivers of data management priorities within organizations. In this paper we highlighted five that are considered to be of considerable impact (organizational culture; organizational life-cycle; asset life-cycle; consequence of asset failure; and asset ownership). Further work is required to validate the role of the five discussed here and to investigate the presence of others with equal or greater impact on data priorities within an organization. Another potentially fruitful research opportunity lies in investigating further the resource allocation processes within organizations relating to the acquisition, storage and utilization of data. In simple terms, who decides how much resources should be devoted to the data management process and what criteria do they use? For example, do managerial decision making schemas tend to centre around return on investment and risk based analysis in contrast to engineering and technical personnel that may place value on data relating to asset reliability, safety and utilization? If so then this may give some insight into the disconnect between the critical nature and value of engineering data to the firm, and the typical lack of organizational resources devoted to its acquisition, storage and utilization.

REFERENCES

1 Lin, S. Gao, J. & Koronios, A. (2006) The need for a data quality framework in asset management. Proceedings of the 1st Australasian Workshop on Information Quality (AUSIQ), June 22-23, 2006. Adelaide.

2 Koronios, A. (2006) Validating the Data Quality Framework in Engineering Asset Management. Proceedings of the 17th Australasian Conference on Information Systems, December 6 - 8, Adelaide, South Australia.

3 Orr, K. (1998) Data quality and systems theory. Communications of the ACM, 41(2), 66-71.

4 Sandtorv, H., Hokstad, P. & Thompson, D. (1996) Practical experiences with a data collection project: The OREDA project. Reliability Engineering and System Safety, 51, 159-167.

5 Adizes, Isaac. (1979) Organizational passages: Diagnosing and treating life cycle problems in organizations. Organizational Dynamics, 8, 3-24.

6 Schein, E. (1992) Organizational culture and leadership (2^{nd} ed.). San Francisco, Jossey-Bass.

7 Detert, J. Schroeder, R. & Mauriel J. (2000) A framework for linking culture and improvement initiatives in organizations. Academy of Management Review, 25(4) 850-863.

8 Tsang, A. (2002) Strategic dimensions of maintenance management. Journal of Quality in Maintenance Engineering, 8(1), 7-39.