Mathew, Avin D. and Ma, Lin and Hargreaves, Douglas J. (2008) A conceptual data modelling methodology for asset management data warehousing. In Gao, Jinji and Lee, Jay and Ni, Jun and Ma, Lin and Mathew, Joseph, Eds. *Proceedings World Congress for Engineering Asset Management*, pages pp. 1086-1095, Beijing, China.

A CONCEPTUAL DATA MODELLING METHODOLOGY FOR ASSET MANAGEMENT DATA WAREHOUSING

Avin Mathew, Lin Ma, Doug Hargreaves

Cooperative Research Centre for Integrated Engineering Asset Management, Queensland University of Technology, Brisbane, Australia.

In an attempt to turn data management into a profitable enterprise, many businesses are seeking to integrate and centralise their data through data warehousing. A data warehouse then allows businesses to turn data into knowledge, and turn knowledge into tangible profits. One key success factor of a data warehouse lies in its ability to integrate data from multiple sources through a unified data model. Within asset management, several such integrated data models have been proposed, however these individually only cover a limited number of areas within asset management data and are not designed with data warehousing in mind. This paper presents the development process of a novel conceptual data warehousing data model that holistically integrates numerous asset management data areas. The comprehensive ethnographic modelling methodology involves a diverse set of inputs (including data model patterns, standards, information system data models, and business process models) that describes asset management data. The outputs of the process were verified by more than 20 experts in asset management and validated against four case studies.

Key Words: Asset management, Data warehousing, Conceptual modelling, Data modelling, Integration

1 INTRODUCTION

The last 30 years have led to an explosion in the amount of data available within organisations. Technology enhancements in information and data acquisition systems have allowed for huge volumes of data to be collected and retained within an organisation, while innovations such as the Internet have produced new sources of data such as e-mail systems. Data management has become a critical function for IT (information technology) departments as they attempt to gather, maintain, and analyse the overwhelming amounts of organisational data. To aid their data management, many organisations are now seeking to integrate and centralise their data through data warehousing. A data warehouse serves as an information management platform that integrates information across domains, organisations, and applications. To provide a conduit of accurate and valuable information, such integration must occur throughout the different layers in a data warehouse, from the underlying data models to the graphical presentation of information to users.

Data warehouses critically depend on data models to integrate data from different information systems and databases in an organisation. As data flows into the data warehouse from sources that have different underlying data models, the data warehouse's data model must consolidate several disparate data models into a unified structure. This issue is notably evident within asset management, where data often resides in multiple information systems and databases. As such, the uptake of data warehousing in asset management has been slower than other business areas. While there have been a few initiatives by standards bodies and information system vendors to create data models that integrate asset management data from various systems, the resultant models are not complete in their coverage of data areas. Additionally, these data models are not designed for data warehouses (instead, for databases) and subsequently do not address data warehouse-specific concepts such as temporal support, aggregate values, and hierarchies.

To address the lack of asset management data models for data warehousing, this research embarked on a data modelling venture that would fulfil the requirements of an asset management data warehouse and subsequent decision support system. The process follows the three-staged modelling paradigm of beginning with conceptual modelling, progressing to logical modelling, and ending with physical modelling. This paper presents the first stage of conceptual data modelling and details the comprehensive ethnographic data modelling methodology undertaken as well as the case study evaluation of the methodology's outputs. The process and experience detailed within this paper serves as a guide and reference for both data modellers and system architects seeking to develop not just an asset management data warehouse, but any type of asset management information platform.

2 EXISTING WORK

While there are no models that purport to be a conceptual data model for asset management data warehousing, broadening the scope by excluding data warehousing reveals areas of relevant work. These conceptual data models for asset management are found within standards presented by industry bodies and standards organisations, as well as in the data models from which information systems are built.

There are three relevant standards – ISO 15926, MIMOSA OSA-EAI, and ISA-95 – that are intended for the exchange of asset management data between systems. All originate from different backgrounds and needs, and their influences are apparent in the design of the models. While the primary objective of each standard is in communicating data, the relevance to data modelling stems from the fact that one step in integrating data is identifying the content of the data (which can be expressed as a model).

ISO 15926 is entitled, “Industrial automation systems and integration—Integration of life-cycle data for process plants including oil and gas production facilities”. Its roots are planted in ISO 10303, more commonly known as STEP (Standard for the Exchange of Product model data) whose focus was on providing a taxonomy of equipment related data. Although initially beginning with the process industry, the coverage of ISO 15926 has increased as it has become more generic and less specific to a particular industry. It is now being proposed as an upper level ontology by its primary maintainers [1], and has subsequently received criticisms levelled against its ontological applicability [2]. Ontologies and data models both deal with conceptualisation, however, ontologies are typically generic and task-independent while data models are task-specific and implementation-oriented [3]. ISO 15926 is more the former, with its entities more focused on modelling the universe rather than specifically focusing on asset management. Abstract and generic models require longer periods for implementation, as a greater number of design choices are required. Coupled with the huge nature of the model and lack of supporting technology [4] no complete implementations of ISO 15926 exist at present.

MIMOSA OSA-EAI is entitled, “Open Systems Architecture for Enterprise Application Integration”. It is poised as a standard interface for operations and maintenance data, but its vibration data exchange heritage is apparent with its comprehensive support of XML (Extensible Markup Language) Schema for the transmission of raw measurement data. Apart from the area of measurements, the OSA-EAI also supports the areas of equipment, agents (personnel and organisations), work management, events, equipment health and diagnosis, alarms, and reliability.

ISA-95, which is entitled, “Enterprise-Control System Integration”, was born out of the need for a standard interface between enterprise and process control systems. Thus the areas it covers involve equipment, personnel, materials, capabilities, schedules, and performance. While it is clear that the standard does not try to define the entire gamut of asset management unlike the previous two, it provides unique concepts not touched upon by the others with its capability and performance models.

There are also a multitude of information systems within the EAM (Enterprise Asset Management) category (as well as Enterprise Resource Planning and Computerised Maintenance Management System categories) that are based on asset management data models. As public publishing of these data models is limited in order to maintain a competitive advantage, it is difficult to ascertain the scope of the data models on which the information systems are based. Regardless of their unavailability, their scope is questioned as there are no information systems that support the complete functionality of an asset management conceptual data model. While it is a false inference to assert that this is because no asset management conceptual data models exist, it does strengthen the probability. An organisation would create a complete asset management data model either because (1) it plans to expand their current solutions into other areas, or (2) it plans to integrate their solution with other products. With (1), even if a company develops a data model, it would suffer from publishing restrictions as previously stated. With (2), as the lack of integration between different classes of information systems is the primary motivation behind the above integration standards, the information systems-based data models must be lacking in their scope.

3 CONCEPTUAL DATA MODELLING METHODOLOGY

The main considerations of a data modelling methodology revolve around the process as well as the process inputs, both of which are discussed below. The outputs of the methodology as well as their validation are discussed in subsequent sections.

3.1 Modelling Process

The foundation of most data model theory is derived from the ANSI/SPARC three schema architecture proposed by Tsichritzis and Klug [5]. The architecture separates the conceptual, logical (external or application view), and physical (internal view) levels. The purpose of a conceptual data model is to explore high level domain concepts; the purpose of a logical data model is to define the entities, attributes, and relationships for an enterprise project; and the purpose of a physical data model is to design the schema of a database [6]. Moving from the conceptual model to physical model entails an increase of structured information at each level. Thus a conceptual model is required before either a logical or physical model can be developed.

The two important groups of conceptual data modelling methodologies are the ER model and Object Role Modelling (ORM) [7]. The methodology in this paper has its roots based in the ER model due to the simplicity and pervasiveness of the technique. Despite the application of object-oriented techniques to data modelling "show[ing] exactly the same concepts as ER models" [8], UML object techniques are used to enhance the data modelling process. Literate modelling [9] is also used to enhance the ER and UML methodologies by providing a narrative of each model.

It is important to note that object modelling itself was not conducted. The intent of object modelling is to model a software system and how it operates, rather than how data are stored. Instead, only the notation elements of object modelling are used, rather than its philosophical purpose.

3.2 Modelling Inputs

A comprehensive data model requires a comprehensive modelling process and a thorough knowledge of the domain. In order to understand the field of asset management and the data requirements, seven points of investigation formed the inputs to the modelling process. These were the examination of data model patterns literature, standards, information systems, business process models, interviews, analysis procedures, and business documents.

3.2.1 Data Model Patterns

There are several definitions attributed to the word 'pattern' [10]. A pattern is a template from which something can be derived. A pattern is also reoccurring characteristics of multiple objects. Patterns literature within computer science revolves around providing exemplars of good practice. As one of the aims of this work is to provide a reference model, the characteristics of patterns are significant to note.

There are numerous branches of patterns literature ranging from object models, integration models, data models, to metadata models. While the area is broad, the focus for this work is on data model patterns in which there are three seminal works. These are books by Hay [11], Fowler [12], and Silverston [13]. Both Hay [11] and Fowler [12] provide more conceptual models where the focus is on entities, their relationships, and cardinalities. Silverston [13] concentrates on providing logical models that include attributes, and are one step removed from physical models.

The difference in the approaches to the patterns contributes to the differences in the covered business areas as shown in Table 1. The table lists the patterns relevant to asset management and the authors that describe the patterns. Silverston [13] tends to be more implementation driven, and goes into domain specific patterns such as ordering and invoicing, and shipment of goods, while the other two authors abstract the detail with the broader categories of contracts and activities, respectively.

There are several works in enterprise object model patterns [9, 14, 15] that are useful to conceptual data modelling. The qualifier 'enterprise' is used to differentiate from creational, structural, and behavioural object patterns such as those by the Gang of Four [16]. Despite the impedance mismatch between object models and data models [17], a data model can be derived as object modelling forms a superset of relational modelling. As the aim is to provide a conceptual data model, the object model patterns need to be abstracted through techniques such as class categorisation to discover the underlying theory upon which they are based.

In a similar vein to patterns, albeit from a commercial perspective, ADRM (Applied Data Resource Management) provide enterprise, business area, data warehouse, and data mart models for various industries [18]. Starting from a common enterprise model, contextual elements are added for various organisations. Thus portfolios are added for financial service companies, rate plans are added for telecommunications companies, policies are added for insurance companies, and metering is added for utilities. There are some patterns that can be used for asset management, but such support is varied.

Table 1
Comparison of data model patterns literature

Area/Sub-area	Hay	Fowler	Silverston
People and organisations	✓	✓	✓
Assets and objects	✓	✓	✓
Documents	✓		
Contracts	✓	✓	
Ordering and invoicing			✓
Procedures and activities	✓	✓	
Shipment			✓
Work	✓	✓	✓
Accounting	✓	✓	✓
Measurement	✓	✓	
Units	✓	✓	
Ranges		✓	

3.2.2 Standards

Standards are an agreed upon set of rules that are established by an authority [10]. However, the proliferation of standards-issuing bodies signifies the lack of a universal authority. This is true within asset management, where there are a multitude of standards from different organisations that span the same areas [19]. As Grace Hopper said, "the wonderful thing about standards is that there are so many of them to choose from" [20].

Despite the plethora of standards, their importance on conceptual data modelling lies within the endorsement by authorities. These are experts that have significant domain knowledge, and who may have conflicting views but have come to a compromise to ratify a standard. Thus the established concepts, models, and nomenclature in standards can be applied in the modelling process.

For the patterns they describe, only Arlow and Neustadt [9] reference and comply with standards, primarily those by the International Standards Organization (ISO). Thus money is phrased in terms of ISO 4217, countries in terms of ISO 3166, and books in terms of ISBNs (ISO 2108). The standards referenced only deal with categorical data, and thus their influence on the patterns is only reflected in the model attribute types. For example, as ISO 4217 describes currencies as a three letter code (e.g. "USD"), the Currency class has an alphabeticCode attribute of String type.

There are hundreds of standards that are related to asset management, and it would be an impossible task to examine each. There are several summaries of existing and future standards and how they fit into particular areas of asset management [19, 21-23], but these summaries are not exhaustive. Even when examining data related asset management standards, not all are applicable in conceptual data modelling. For example, IEEE 1451 provides a standard transducer interface, while and IEC 60870 provides a standard controller interface. Despite their importance to asset management, the mechanism used by systems to transmit data is not important (although the data contained within are important).

Table 2
Asset management related standards

Area	Standards
Specifications	PAS-55
Assets / Activities	MIMOSA OSA-EAI ISO 15926 / ISO 10303 – STEP ISA-95 / B2MML ISO 14224 OASIS PPS KKS
Personnel / Organisations	MIMOSA OSA-EAI ISO 15926 ISA-95 / B2MML
Documentation	OMG ManTIS
Life cycle costing	AS 4536
Measurements / Diagnosis / Prognosis	MIMOSA OSA-CBM MIMOSA OSA-EAI ISO 13374 IEEE 1232 – AI-ESTATE
Units	IEEE SI 10-1997
Reliability	ISO 14224 AIAG FMEA-3 SAE J1739 MIL-STD-882
Risk	AS 4360

Table 2 shows a list of asset management standards that were used for asset management conceptual data modelling. It is not a complete list of all applicable standards, but it covers a large majority of asset management data areas relevant for conceptual data modelling. Standards such as MIMOSA OSA-EAI, ISO 15926 and ISA-95 cover multiple asset management areas and provide either object or data models. General standards that are domain independent are used for less traditional asset management activities such as risk management or life cycle costing.

3.2.3 Information Systems

As one of the goals of conceptual data modelling is to provide an integrated model from which a data warehouse can be based, it is important to look at information systems. An information system is "the system of persons, data records and activities that process the data and information in a given organization, including manual processes or automated processes" [24]. Information systems are based upon a data model, which form constituent areas of a total asset management model. Information systems also supply the data that are extracted, transformed, and loaded into a data warehouse.

As with standards, there is an overwhelming number of existing information systems with the key differentiator being their functionality. The areas covered by each information system are similar for their class. Thus most EAM systems will cover asset registries, financial management, materials management, maintenance work management, etc. while reliability systems will cover FMECA (Failure Mode, Effects, and Criticality Analysis), reliability block diagrams, root cause analysis, etc. The similarity between systems in the same class benefits the modelling process by narrowing the systems to be examined to a sample representing the population.

Examination of information systems can be undertaken in two ways. The first is a direct physical inspection of the front end functionality

Table 3
Asset management information systems

Category	Information System
Enterprise Resource Planning (ERP)	SAP
Enterprise Asset Management (EAM)	Maximo, Mainpac, Hansen
Documentation	Comos
Geographic Information Systems (GIS)	ArcView
Reliability	Relex
Process and control	CitectSCADA
Condition monitoring	Emonitor Odyssey, IMS Watchdog

and/or back end database. The database does not directly need to be analysed as the data model can be derived by looking at the functions the system supports. Particularly with large and older systems that have a huge amount of tables with unintelligible names, examining the front end is more productive, as asset management areas can be studied in manageable portions. The second means to examine information systems is indirectly via a list of supported functions. This can be either gathered from documentation, advertising materials (such as vendor websites), or word of mouth.

Convenience sampling was used in selecting the information systems to be analysed. As many of the systems require intricate deployment setups that can cost thousands and sometimes millions of dollars, the systems of CIEAM and IMS associated organisations were investigated. The list of information systems examined is presented in Table 3. The systems listed encompass a wide range of asset management data, particularly as ERP and EAM systems are expansive in their coverage. SAP, Maximo, ArcView, and Emonitor Odyssey form part of the world market leaders in their respective categories [25-27], while the Citect is the Oceania market leader [28] for process control systems. Although popularity does not indicate technical superiority, it does indicate the significance and relevance of functionality required by organisations.

3.2.4 Business Process Models

In order to understand, redesign, and optimise existing business processes, companies are undertaking business process modelling to capture their business processes. These models show the relationship between activities, data, entities, resources, and goals. The connection between activities forms a flow, and branches can be defined using Boolean logic to indicate where decisions are required. As the main goal of data warehousing is to provide a foundation for decision support, business process modelling has an impact on conceptual data modelling by dictating the types of data required to be supported in the model.

Business process models of asset management processes at a water utility organisation in Australia were analysed. However, the models did not completely cover all of the asset management business functions within the organisation. Using the ARIS (Architecture of Integrated Information Systems) methodology of business modelling, the models covered the business value added chain, a high level asset management process, and the processes involved in information acquisition and analysis, risk analysis, strategic and operational planning, scheduling, and maintenance work. The models were described through process, data, organisation, and function views, with supplemented descriptions (i.e. literate modelling).

3.2.5 Analysis Methods and Functions

While a large percentage of data warehouse decision support systems solely use OLAP in their analysis, there are domain specific analysis techniques that are more intricate in design and data requirements. These can be in the form of methodologies, or at a lower level, algorithms.

Reliability and condition monitoring are two areas of asset management that use complex analysis methods. The reliability area presents simple reliability measures (e.g. mean time before failure), FMECA and root cause analysis, to reliability prediction using historical failures [29]. There are generally two ultimate goals of condition monitoring: diagnosis and prognosis. Diagnostic methodologies focus on using intelligent classifiers to compare healthy and non-healthy equipment states, while prognostic methodologies focus on using regression to predict the health of an asset. These two areas are also covered by the conceptual data modelling process.

3.2.6 Interviews

One of the metrics by which a decision support system is measured is its ability to satisfy end users. Thus eliciting requirements and knowledge from people within the field who may use an asset management data warehouse is important. While standards do present the consensual views of people and groups, there may be elements missing from standards. This could be due to issues where it becomes too difficult to reach a consensus, issues that do not have enough resources for standardisation, or issues that do not require standardisation. Thus interviewing experts can fill gaps not discussed within standards. Interviews also have the advantage of exploratory analysis, as topics can evolve.

People from three types of organisations (industry organisations, universities, and committees) were interviewed. The interviews included structured and unstructured questions, presentations given by the interviewer or the interviewee, and field demonstrations. The interviews also were conducted individually as well as in groups for both the interviewer/s and interviewee/s.

3.2.7 Business Documents

Although data model patterns, standards, information system data models and functionality, business process models, and analysis methods can be embodied as documents, there is other information that this medium provides. Documents include business vision and strategies, policies, manuals, log books, presentations, diagrams, and charts. While the information content in the aforementioned documents could be stored within business process models or information systems, their existence indicates a need for the document format (e.g. technological limitations within business process models/information systems, or a greater utility with a document format).

With the Internet increasing the ease of information dissemination, many business documents, such as annual reports, can be found online. While confidential documents are usually locked away on the corporate Intranet, some documents are made publicly available – particularly those by government organisations. The documents used in this research were a combination of business-supplied as well as Internet-harvested documents. While the majority of their content areas were uncovered via the other six modelling inputs, the documents provided a validation for these areas, as well as insight into business operations around the globe.

4 ASSET MANAGEMENT CONCEPTUAL DATA MODEL

As the main thrust of this paper is to detail the methodological development of the conceptual data model, the model itself is only briefly summarised in this section. Terminology primarily follows the MIMOSA OSA-EAI 3.1 Terminology Dictionary, although some modifications and additions are made for clarification. Graphical representation follows a combination of ER and UML syntactical conventions, and the approach by Fowler [12] is used for positional conventions in that models are split into knowledge and operating levels. Microsoft Visio was chosen as the modelling tool for its abilities in quickly codifying thoughts to paper and ease in graphically manipulating schema objects.

Table 4 shows a list of the areas covered by the asset management conceptual data model. The model is sectioned into seven major areas to improve navigation through the model; however, many common elements exist within these major areas. Resources are basic objects within an organisation which can have various types and that can have different attributes and associations between them. For example, an asset can be of an electric motor type, have a maximum RPM capability/attribute of 1725 RPM, and have associations with other assets (e.g. bearing, shaft) in its containing pump structure. Attribute and association patterns were extensively used throughout the modelling process to facilitate reuse. The area of motivation describes objectives/goals, and the means to achieve those objectives/goals. It also describes influential factors upon these objectives and means, as well as their potential risks and rewards. Activities are one type of means, and can consist of hypothetical, proposed, or actual activity types. Measurements are a type of activity often used within asset management, and require the modelling of units, measurement regions, and alarms. Events are distinguished from activities in that they occur at a point in time rather than over a duration. The cause and effect nature of events and activities is also captured through this area. Finances deal with monetary accounts and transactions involved when dealing with resources and activities. Contracts specify legal obligations between agents, and insurance and warranties are covered in depth for asset management. Documents provide a way to refer to any (instantiated) entity described by the model and their interactions with other entities.

5 EXPERIMENTAL TESTING

The conceptual data model was examined through both verification and validation processes; verification and validation are two stages within software testing to ensure quality. Verification ensures that a system meets the specified requirements while validation is to demonstrate that a system fulfils its intended use when placed in its intended environment [30]. As summarised by Balci [31], verification is about building the model right, while validation is about building the right model.

5.1 Verification

During the development of the data model, the model would be periodically reviewed by experts within the field of asset management. As asset management is an expansive discipline, it is virtually impossible to find an expert that is knowledgeable in all areas, and thus a variety of experts were engaged. The people who provided input into the model via interviews (see Section 3.2.6) were also asked to analyse both the modelling process and the model itself, enabling a cyclical feedback loop. More than 20 people from these organisations reviewed the model for accuracy and coverage.

These reviews were conducted by presenting firstly a justification of asset management conceptual data modelling; secondly a background on data modelling; thirdly the methodology used in this research; and finally the models themselves. While Microsoft PowerPoint was used to present the first three items, the modelling tool, Microsoft Visio, was used to present the model. The latter was used as (1) it was difficult presenting readable models that fit on a

Table 4
Areas within the asset management conceptual data model

Topic area	Coverage
Resources	Assets, Models, Segments, Geographic Locations, Agents, Specifications, Capabilities
Motivation	Ends and Means, Business Rules, Influencers, Risks / Rewards
Activities	Hypothetical and Actual Activities, Measurement, Units, Regions, Alarms
Events	Hypothetical and Actual Events, Cause and Effects
Finances	Accounts, Transactions
Contracts	Contracts, Insurance and Warranties
Documents	Indexes

single presentation slide and (2) changes to the model could be made in real-time.

As the conceptual data model was presented at the same time as being reviewed, a concern was that each expert would not have enough time to examine the model. Due to confidentiality concerns at the time of review, the model could not be given out to reviewers, and all reviews needed to be conducted in the presence of researchers. These sessions lasted one to two hours, with the variation resulting from the amount and type of questions asked.

Section 2 showed that whilst there were no equivalent existing models, there were three standards as well as certain information systems that were relevant for asset management conceptual data modelling. These items were used in the verification process by comparing the corresponding areas of this research with the other models. The objects within the conceptual data model were examined to see if they were ideologically compatible with the objects within the compared model. The term *ideologically compatible* is used because a simple one-to-one comparison would produce erroneous results, and thus the intent of each object needed to be compared.

5.2 Validation

As pointed out by Tichy [32], "experimentation is central to the scientific process [as] only experiments test theories". Thus if possible, a system should be subjected to empirical evaluation to prove its scientific merit. Validation was conducted in the form of both practical and theoretical usability experiments. Four case studies were conducted, each at differing levels for software development. Case Study 1 provides an implementation case of a section of the model; Case Studies 2 and 3 provide coverage comparisons against two detailed asset management software specifications; and Case Study 3 provides a comparison against a high-level asset management framework.

5.2.1 Case Study 1

Due to the overwhelming nature of the number of asset management areas covered by the conceptual data model, it could not be implemented in its entirety for validation. A complete implementation would require an unattainable amount of resources, and hence a different approach needed to be taken. A partial implementation was conducted in conjunction with a CIEAM research project.

The research project entitled "Integrated Decision Support System for Asset Management in the Water Utility Industry", looked at developing an asset management health system called BUDS (Bottom Up Decision Support). The goal of BUDS was using condition monitoring and reliability data in order to aid in asset renewal decision making. Thus sensor data from condition monitoring and SCADA systems were stored in the database from which diagnoses and prognoses were calculated. The system also stored failure modes, RDBs, and fault trees for reliability prediction calculations. Maintenance cost data was also stored and used with the condition monitoring and reliability predictions to translate the figures into a format that managers could understand – costs and dates.

The software was developed in Microsoft Visual Studio .NET 2003, using C# as the programming language. A relational DBMS (Microsoft SQL Server 2000) and flat files were both used as the storage mechanism as it has been shown that each have their own merits [33]. The conceptual data model was translated to a logical and physical data model for the sections stored in the database. For the objects stored in the relational database, a Structured Query Language (SQL) program was created that would create tables and insert data.

As data was gathered from a variety of systems (ERP/CMMS, condition monitoring, SCADA, Microsoft Access database, and Microsoft Excel spreadsheets), a semi-automated ETL (extraction, transformation, and loading) process was developed for each source. In addition to extracting the data from the sources for loading into the data warehouse, a data cleansing process was added for selected sources.

As the ETL process integrated data from several source systems, analysis could be conducted not only with the aforementioned program, but with any tool that could read from a database. Thus tools such as Microsoft Excel and MATLAB were used for report generation in areas that BUDS did not have the required functionality.

5.2.2 Case Study 2

A subsequent CIEAM research project, looked at extending the functionality of BUDS to cover additional functionality for asset management. As the project is currently in progress, only the software requirements were validated against the conceptual data model.

The requirement specifications in the subsequent project enveloped the past project to include functionality in the areas of asset operation, inventory, safety procedures, project management, contracts, and asset performance. The specifications itemised specific data requirements and described the required functionality for each area.

A comparison between the specifications and the conceptual data model showed that the model adequately covers each area in the specification. As is the nature of research, the specifications are in constant flux and some sections are incomplete, and for these areas, a less detailed comparison can only be made.

5.2.3 Case Study 3

The specifications for an asset management system under development by World In One Technology Sdn Bhd in Malaysia were also compared to the conceptual data model. The system forms the base module for a future system implementation of change management, maintenance management, and fleet management. It currently covers areas in asset registry, configuration, failure, movement tracking, value and depreciation, warranty tracking, measurements and reporting, and performance and reliability measures.

A comparison against this case study showed that the conceptual data model covers each area in the specification. While metadata was not modelled in the data model, it was addressed by this research, and this is used with the specification area of user permissions. Tohe areas cover similar areas to Case Study 2, with less emphasis on condition monitoring measurements, and more emphasis on metering. Warranties are addressed through the contract model, and currency conversion is addressed through the units of measurement model.

Table 5
Comparison to the CIEAM Asset Management Framework

Framework area	Relation to conceptual data model
Strategic Planning	Motivation
Asset Ownership	Documents, Motivation
Risk Management	Motivation, Risk
Budgeting & Costing	Finances
Data Management	Assets, Segments
Condition Monitoring	Measurements
Tactical Planning	Activities
Human Resources	Agents
Assets Usage Life Cycle	Activities
Performance Measure	Measurement
Information Systems	N/A
Financial Management	Finances

5.2.4 Case Study 4

The comparison process was also applied to the CIEAM Asset Management Framework (the results of which are shown in Table 5). While not a software specification, the framework provides guidance on the broader functions within asset management, and through an identification of the data needs of these functions and sub-functions, a comparison could be drawn.

The functions described by the framework are at a higher level than the two previous case studies, and hence only the top level areas in the conceptual data model are compared. In addition, as the framework areas are broad categories, only the primary conceptual data model areas are listed. For example, Documents and Motivation are the primary data areas for Asset Ownership, but in reality, data in the Assets, Segments, Agents, and Activities models would also be used for functions in this framework module. The Information Systems framework module is listed with a N/A as the sub-functions relate to the creation of information systems, rather than the use of information systems (as is the case with the other eleven areas).

5.2.5 Case Study Discussion

The case studies validate the models through different levels. Case Study 1 validates the conceptual data model using detailed implementation-level data models but for the very specific scope of condition monitoring and health assessment. Case Studies 2 and 3 validate a larger (but still limited) portion of the conceptual data model, but through a more abstract means. Similarly, Case Study 4 validates almost the entire conceptual data model through an even more abstract comparison.

As the end use of the conceptual data model is to provide an implementation (or physical) data model as in the first case study, the ideal validation of the model would require the implementation of the whole model. However, this is not feasible due to time and financial constraints as well as a lack of research in certain asset management areas to effectively use of the data by a computer system. Thus, it is left up to future research to provide full, implementation-level validation of the conceptual data model.

6 CONCLUSION

The importance of data modelling is evident in system development, as it has far reaching consequences on system design. Within a data warehousing context, data models influence low level factors such as storage space and performance speed, to higher level factors such as the types of analysis that can be conducted. With an increasing number of systems and data areas, asset management organisations are seeking to integrate these areas for advanced data warehouse-styled reporting. An understanding of asset management data needs to be in place before integration can occur, and this research has attempted to provide one stage in developing this understanding.

This paper addresses the development of a conceptual data model for asset management data warehousing and presents a unique modelling methodology and approach. Relational, object, and literate modelling were combined to harness the benefits of each modelling language, provide a systematic approach, and to present a more intuitive understanding of the data model. By examining data model patterns, standards, information systems, business process models, analysis methods, and conducting interviews, a comprehensive model was developed.

The types of data available within asset management systems are diverse, and the primary areas identified by this research include: Resources, Motivation, Activities, Events, Finances, Contacts, and Documents. Each of these areas have unique structural elements, and both associative and attributable patterns are clearly evident in their structure. The conceptual data model presents the integration between each area and clearly highlights the interrelationships within asset management data. The models were verified through expert reviews and validated against four case studies that ranged from a low-level asset management software implementation to a comparison against a high-level asset management framework.

The methodology presented provides a valuable resource for asset management system designers. Amongst the list of modelling inputs is an evaluation of these resources and their applicability for asset management data modelling. While the comprehensive methodology has been demonstrated on the area of asset management data, it does not preclude an application to other business domains.

7 REFERENCES

1 West M. (2006) ISO 15926 – Integration of lifecycle data. *Upper Ontology Summit*, Gaithersberg, MD. National Institute of Standards and Technology.

2 Smith B. (2006) Against idiosyncrasy in ontology development. Bennett B & Fellbaum C (Eds). *Formal Ontology and Information Systems*, Baltimore, MD. pp. 15-26. IOS Press.

3 Spyns P, Meersman R & Jarrar M. (2002) Data modelling versus ontology engineering. *SIGMOD Record,* 31(4), 12-17.

4 Teijgeler H. (2007) *The process industries and the ISO 15926 semantic web*. Netherlands: OntoConsult.

5 Tsichritzis D & Klug AC. (1978) The ANSI/X3/SPARC DBMS framework report of the study group on database management systems. *Information Systems,* 3(3), 173-191.

6 Simsion G & Witt G. (2004) Data modelling essentials (3rd ed.). San Francisco, CA: Morgan Kaufmann.

7 Venable JR & Grundy JC. (1995) Integrating and supporting entity relationship and object role models. *Object-Oriented and Entity-Relationship Modelling*, Gold Coast, Australia. Springer.

8 Hay DC. (1998) Making data models readable. *Information Systems Management,* 15(1), 21-33.

9 Arlow J & Neustadt I. (2003) Enterprise patterns and MDA: Building better software with archetype patterns and UML. Boston, MA: Addison-Wesley.

10 Merriam-Webster. (2003) Merriam-Webster's Collegiate Dictionary (11th ed.). Springfield, MA: Merriam-Webster.

11 Hay DC. (1996) Data model patterns: Conventions of thought. New York, NY: Dorset House.

12 Fowler M. (1996) Analysis patterns: Reusable object models. Reading, MA: Addison-Wesley.

13 Silverston L. (2001) The data model resource book: A library of universal data models for all enterprises (2nd ed.). New York, NY: John Wiley and Sons.

14 Coad P. (1997) Object models: Strategies, patterns, & applications (2nd ed.). Upper Saddle River, NJ: Prentice Hall PTR.

15 Nicola J, Mayfield M & Abney M. (2001) Streamlined object modeling: Patterns, rules, and implementation. Upper Saddle River, NJ: Prentice Hall PTR.

16 Gamma E, Helm R, Johnson R & Vlissides J. (1994) Design patterns: Elements of reusable object-oriented software. Reading, MA: Addison-Wesley.

17 Behm A, Geppert A & Dittrich KR. (1997) On the migration of relational schemas and data to object-oriented database systems. *International Conference on Re-Technologies for Information Systems*, Klagenfurt, Austria. pp. 13-33.

18 Applied Data Resource Management. (2008) Data Environments. Retrieved January 2, 2008, from http://www.adrm.com/7_products.htm

19 Koronios A, Nastasie D, Chanana V & Haider A. (2007) Integration through standards: An overview of international standards relevant to the integration of engineering asset management. *World Congress on Engineering Asset Management*, Harrogate, UK. pp. 1066-1088.

20 Garfinkel S, Weise D & Strassmann S. (1994) The UNIX-HATERS Handbook. San Mateo, CA: IDG Books.

21 Johnston A. (2006) *Enabling open standards-based interoperability for operations & maintenance people, processes and systems*. Tuscaloosa, AL: MIMOSA.

22 Mills S. (2002) *Condition monitoring and diagnostics of machines: Standards in progress*: ISO.

23 Wolfson Maintenance. (1998) *Guidelines for types and locations of measurements & ISO standards references*. Manchester, UK.

24 ISO. (2008) Information technology - Security techniques - Information security management systems - Overview and vocabulary, *ISO 27000*.

25 Cline ME. (2005) Mapping solutions under $500. *ONLINE,* 29, 27-30.

26 Liang H, Xue Y, Boulton WR & Byrd TA. (2004) Why Western vendors don't dominate China's ERP market. *Communications of the ACM,* 47, 69-72.

27 O'Hanlon T. (2005) *Computerized maintenance management and enterprise asset management best practices*. Fort Myers, FL: Reliabilityweb.com.

28 Wooldridge S. (2007) MyCitect News Mar 2007 Oceania. Retrieved November 16, 2007, from http://www.citect.com/index.php?option=com_content&task=view&id=422&Itemid=458
29 Sun Y, Ma L, Mathew J, Wang W & Zhang S. (2006) Mechanical systems hazard estimation using condition monitoring. *Mechanical Systems and Signal Processing,* 20(5), 1189-1201.
30 Software Engineering Institute. (2007) *Capability Maturity Model® Integration v1.2*. Pittsburgh, PA.
31 Balci O. (1997) Principles of simulation model validation, verification, and testing. *Transactions of the Society for Computer Simulation International,* 14(1), 3-12.
32 Tichy WF. (1998) Should computer scientists experiment more? *Computer,* 31(5), 32-40.
33 Qi K & He Z. (2005) Data management in monitoring and diagnosis system for electromechanical equipment based on LabVIEW. *Computer Measurement & Control,* 13(1), 24-30.