

# Towards Defence Strategic Data Planning

**Li Jiang, Nicholas Tay, Hossein Seif Zadeh and Gary Bulluss**

*Strategic Capability Analysis Branch, JOAD, DST Group  
Email: li.jiang@dsto.defence.gov.au*

**Abstract:** Defence generates and receives vast quantities of data in relation to the specification, development and intended operation of its capabilities and related decision processes. As an organisation, Defence has made good progress in the collection and availability of these types of data; however, the growing pace of technological advancement and the rapid evolution of the spectrum of potential adversaries mean that Defence must become even more effective and efficient in its data analytics to ensure agile decision making at all levels of the organisation. Achieving this objective requires focused strategic data planning, rigorous data design, speedy and accurate data collection, and effective data validation, processing, and management. Current Defence data management practices in Australia do not, for example, provide a unified asset identification number, nor a unified asset naming convention, leading to confusion and ambiguity in the provision of timely and accurate advice to decision makers. This is in part due to the lack of rigorous definitions, data structures and effective data management mechanisms and policies, making universal data integration difficult. This also increases the risk of inconsistencies between Defence's many data repositories. Based on reviewing data management practices of Defence organisation internationally, the authors argue that the Australian Defence Organisation needs an improved strategic data planning mechanism in order to develop a more holistic and effective approach to managing its capability-related data. We propose a framework based on review of the literature and best practices internationally, as well as our experience in analytical support of Force Design and capability-related decision processes, albeit being mindful of the evolving structure of Defence as the implementation of First Principles Review is taking shape. The value proposition of the framework is four-fold: (1) it proposes an iterative strategic data planning process that helps to improve the effectiveness and efficiency of data management in support of effective decision making, (2) it provides a set of guidelines for developing a strategic data plan, helping to facilitate and streamline data management across the organisation, (3) it provides guidelines for data design, collection, and management, as well as resource planning, and (4) it helps develop consistent capture of 'quality' information and other critical data required for Defence capability planning. Our framework is intended to be useful to strategic data planners and data managers working at different levels of Defence, and will also better support robust analytics for capability planning. This can, in turn, help reduce costs in data management but, more importantly, enhance Defence's capability development and trade-off decision processes.

**Keywords:** *Information management, data planning, strategic management, capability management*

## 1. INTRODUCTION

Managing and using data effectively are becoming increasingly critical issues that influence every part of our society: individuals, business, organisations, as well as the nation's economy and Defence. Like many other sectors, Defence generates vast quantities of data in relation to the specification, development, and intended operation of its capabilities and related decision processes.

As an organisation, Defence has made good progress in the collection and availability of all types of data. Many information systems have been developed over the last two decades at various levels of the organisation. For instance, the Defence Preparedness Management Information System (DPMIS) Suite (DPMIS 2015) comprises a suite of tools that capture and present data for preparedness reporting and analysis. Program Viewer (Bulluss, Tay et al. 2014) is another Defence decision support tool providing a rich visualization of Defence capability scheduling, costing, personnel, and project slippage. The Capability Development Management and Reporting Tool (CDMRT) (CDMRT 2015) is an information-collection tool in which data is entered by project managers and used by various stakeholders to aggregate decision support information. With the information provided by CDMRT, Capability Development Group (CDG) can make coordinated and complementary judgements about directions and choices in developing capability options before they are presented to Government for approval.

Despite availability of a plethora of data management systems and tools, many data management issues still persist in Defence. These include:

- Information systems are developed independently within each section of the organisation. Different data formats, interfaces and/or software tools are used which render data sharing and information integration very difficult.
- Data collected and information developed are consistent within a division and yet could be quite different from other divisions or branches. This creates data management issues across Defence as the responsibility for managing phases of Defence's capability life cycle are dispersed and shared.
- There are no unified global unique identifiers used across the organisation. This often causes confusion and renders data sharing and integration impossible. For example, Unapproved Major Capital Investment Program (UMCIP) and Approved Major Capital Investment Program (AMCIP) use project names (e.g. AIR 6000 Phase 2A/B) whilst Material Sustainment Agreements (MSA) uses force element names (e.g. Joint Strike Fighter) as the unique identifier. Major Capital Facilities Program (MCFP) uses a yet different identifier.
- Datasets are incompatible, even amongst common datasets such as costing, making the attribution of costs from one budgetary account to another very difficult. In some instances, similar datasets refer to quite different sets of data. For example, UMCIP uses Initial Operating Capability (IOC) as the date when a capability is available, whereas AMCIP uses Initial Material Release (IMR) as the date when a capability is complete.
- Lack of effective ways to trace changes to data, which often results in decisions being made based on information that is out of date. For instance, project slippage occurs when actual project expenditure deviates from the budget. Slippages between AMCIP and UMCIP budgets are so frequent that a deterministic slippage model is regularly used (Lo, Weir et al. 2015). This reduces confidence in the data.
- There is no enterprise (whole of Defence) level data standard or policy across Defence. No policy or naming convention is used for data definition, tracking changes of data, data structural maintenance, and data quality control across Defence. This often leads to lower quality data containing ambiguous, repetitive, and redundant information. On one hand, this causes utilization of inappropriate data or information for decision making, and on the other, often leads to higher cost in reconciling, integrating, and managing data.
- There are few contemporary principles and best practices that address strategic data quality and data management issues. It was estimated in 2006 that the combined annual cost of bad data in the US was over US\$30 billion (Lee 2007). While a cost estimate does not exist for Australian Defence, the impact of incomplete or inconsistent data on both capability and warfighter lives could potentially be just as significant.

These issues are mostly a consequence of the stove-piped nature of the data sets, which has contributed to not only inefficient and ineffective data management and decision support, but also to higher costs of managing and using the data. The growing pace of technological advances and the rapid evolution of the spectrum of potential adversaries require Defence to become more effective and efficient in its data management to ensure more agile decision making. This requires more strategic data planning, rigorous data design, speedy and accurate data collection, and effective data validation, processing, and management. To help improve current and meet future data management challenges and to support decision making in strategic planning and operations we argue that implementing a sound data management framework is vital for Defence (Simon 2006).

Analysis of existing data management issues at Defence suggests there is a need for a new Strategic Data Planning (SDP) framework in Defence. Early research in SDP (Goodhue, Kirsch et al. 1990) advocated a top-down and centralised database approach in which SDP is defined as “a formalized, top-down, data-centered planning approach that builds a data model of the enterprise.....”. However, our experience has shown this approach is unlikely to work as the complexity and diversity of Defence business inevitably leads to distributed and dynamic changes to data across different parts of the organisation.

To aid development of an effective data strategy, we propose a strategic data planning framework and a more holistic and effective approach based on a review of the literature and best practices internationally. This is supported by our experience in analytical support of Force Design and capability-related decision processes, bearing in mind the evolving structure of Defence as the implementation of First Principles Review is taking shape. In our framework, Strategic Data Planning (SDP) is defined as a formalized, data-centered, value-driven approach that identifies and implements an integrated set of high quality data management and data utilization systems to meet the needs of Defence business. This definition focuses on the following aspects of data management:

- Achieving information superiority by taking a value and utility driven approach. The approach focuses on how data can provide a competitive edge in supporting decision making and military operations
- Achieving data integration between all related divisions and branches across Defence throughout the data lifecycle (creating, processing, implementing change management, using, and retiring data) and the capability development lifecycle
- Addressing continuous management of data change
- Considering data value and quality across the data lifecycle

The volume, frequency and variety of data generated in Defence are rising significantly. Furthermore there is an ever increasing volume of data that is generated by other organisations for Defence, such as technical data (e.g. specifications, standards and engineering drawings) and operational data (e.g. operational environment, operational views, joint force operations and joint international operations). The focus of SDP is the data used in support of decision making, however, the proposed approach can be equally applied to other types and applications of data as well. Our framework is intended to be useful to strategic data planners and data managers working at different levels of Defence, and will also provide better support to robust analytics for capability planning. This could, in turn, help reduce costs in data management but more importantly enhances Defence’s capability development and trade-off decision processes.

## 2. REVIEW OF LITERATURE

Strategic data planning (SDP) is a methodology within the general field of information engineering that is focused on addressing two critical phases of information engineering: organisational analysis and strategy-to-requirements transformation (Martin 1982, Hackathorn and Karimi 1988, Lederer and Sethi 1988, Martin and Finkelstein 1988). As discussed in Section 1, Goodhue et al (1990) describes SDP as a formalized top-down, database centered and process-driven planning approach. However, early implementations of the SDP methodology were not without their problems. A rigorous top-down approach can be very time consuming and expensive, especially in large enterprises with data that is generally heterogeneous. Furthermore, a process-driven approach might not necessarily reveal all the data required. More importantly, it does not explicitly address the issues of data value, consistency, redundancy and fragmentation. While conducting case studies on data management approaches, Goodhue et al (Goodhue, Quillard et al. 1988) observed that “for many firms, the approach is too expensive, its benefits are too uncertain, and it is organisationally difficult to implement”. Whilst the success of SDP was uncertain, Goodhue et al (Goodhue, Quillard et al. 1988) observed that it did provide an architectural base for subsequent data management efforts. Goodhue et al (Goodhue, Quillard et al. 1988) also noted that “in most of the data management efforts studied, the planning and implementation process did not proceed as suggested by strategic data planning methods.” This indicates that a proper SDP implementation may not have been followed in most of the case studies.

Many early limitations of SDP are addressed in the work of Adelman et al. (Adelman, Moss et al. 2005), however, the work focuses more on the issues, and does not provide a systematic methodology for SDP implementation. Moreover, several limitations of SDP relevant to Defence business are yet to be addressed. For instance, the use of naming conventions and naming spaces for military operations are important issues (USJFCOM\_J87 2004) that seem to be lacking in the existing methodologies. Additionally, as the tempo of Defence decision making tends to be fast, the temporal dimension of data management also needs special attention. Achieving information superiority is yet another consideration in strategic data planning in the operational context. According to UK Joint Doctrine Note (DCDC 2013), information superiority is defined as the competitive advantage gained through the continuous, directed and adaptive employment of relevant information principles, capabilities and behaviors. There is no consensus and no endorsed doctrinal definition

across Defence on what information superiority specifically is and how it is achieved (DCDC 2013). Perry et al. (Perry, Signori et al. 2004) proposed a methodology that can help assess quality of information and its impact on shared awareness in military operations. However, this assessment focused exclusively on the measure of information quality, which left the majority of other quality measures untouched.

Data management standards are often developed for managing the quality of data. ANSI/GEIA-859 is an international data management standard intended for managing performance-based data. It also provides a set of best practices for the identification, management, and protection of intellectual property and competitive edge. Another well-known framework is DoD Architecture Framework (DoDAF) (DODAF 2010), which introduces a Data and Information Viewpoint incorporating conceptual, logical and physical levels of abstraction. However, these do not provide methodological support on how to implement SDP.

### 3. STRATEGIC DATA PLANNING FRAMEWORK

The full potential of the value of data can only be realised if Strategic Data Planning (SDP) is carried out at various levels in the organisation responsible for generating and using data; groups, divisions, branches, directorates etc. Based on experience in Defence and industry, and based on a review of relevant literature, a framework is proposed to help develop a sound strategic data plan with a focus on data integration and data utility evaluation. The framework contains a methodology and a set of distilled principles and practices for SDP and can be applied at various levels of the organisation. The detailed five-stage process is discussed in more detail below.

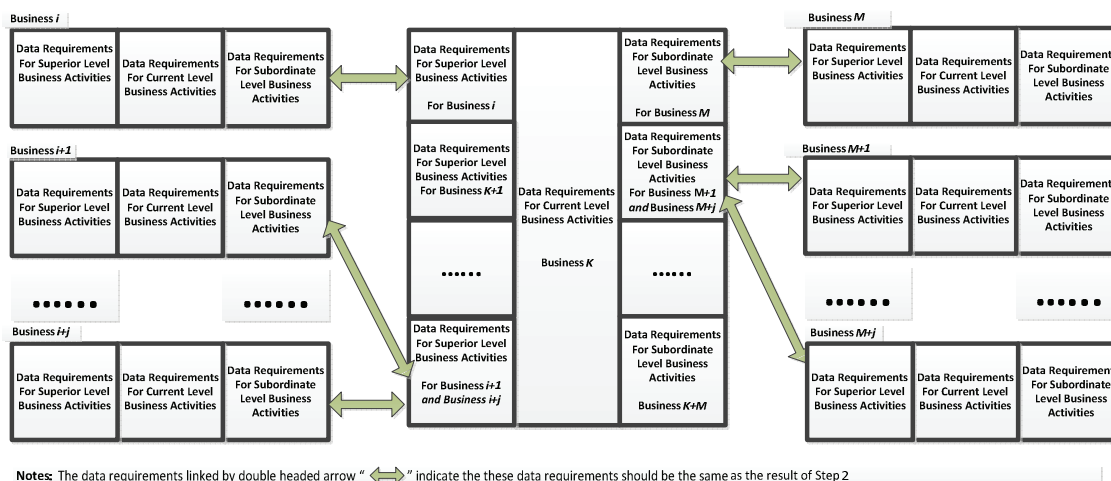
#### Step 1: Define Strategic Data and Concepts

Defining data is the first stage in strategic data planning. The aim of this stage is to develop a set of data requirements for strategic decision making at Defence. This includes the core data concepts; general data requirements, such as what data are required; as well as data security requirements, such as who has what type of access to what data resource. This will be discussed as follows:

- 1) Core data concepts. Defining core data concepts is very important for sharing common understanding of data items (such as Defence capabilities and/or assets) and consistently referring to them among all stakeholders across Defence. Therefore, a set of core concepts that are well-defined and shared across Defence must be developed. This will be discussed further in Step 2 of the paper.
- 2) For each level of business management, the general data requirements are categorized by a data input, processing and output model (Ballou and Pazer 1985). These data requirements include:
  - a) Strategic data requirements for any target level (Business level  $K$  in Figure 1) business activities; a set of well-defined data items (denoted as a set  $BR^D$ ), produced and used within the target level of management. These data are essential for supporting the function of the target level business activities.
  - b) Strategic data requirements for the parent level (Business level  $j$  in Figure 1) business units (with respect to the target level business) as illustrated in Figure 1. These are a set of well-defined data items (denoted as a set  $SR^D$ ) produced in the target level and required by the parent level business or management units. These data are essential for supporting parent level business or management activities. These data requirements are one part of the data integration requirements for the business unit.
  - c) Strategic data requirements from subordinate levels (Business level  $M$  in Figure 1) of business units (with respect to the target level business) as illustrated in Figure 1. These are a set of well-defined data items (denoted as a set  $IR^D$ ) required for the business activities of the target level, and obtained from and produced by the subordinate levels of business units. These data requirements are another part of the data integration requirements for the target business unit.
- 3) Data security requirements are related to data security, privacy, and access authorisation. These requirements should be defined before data is even generated. Details of these requirements are beyond the scope of this paper.

In the strategic data definition stage, ancillary issues such as data modelling, data structure, data categorisation, data definition and naming conventions and metadata definitions should also be considered. Naming conventions will be discussed briefly in the second stage of the strategic data planning process. Discussion of the remaining ancillary issues is beyond the scope of this paper.

At each level of the organisation, it is the responsibility of the data manager to establish a set of critical data items and mandate their use and collection. Each subordinate group or branch has the discretion to collect more data as they see fit, but they will need to at least collect the data mandated by their first-level line management. Other data and data structures they decide to collect or use must be subordinate to and compatible with the mandated data structures.



**Figure 1.** Definitions of Data Integration Requirements for Business level K.

## Step 2: Define Strategic Data Integration Requirements

Defining strategic data integration requirements is a stage specifically designed to address chronic data management failure experienced by many organisations. Addressing data integration issues at an early stage of strategic data planning is vital in reducing potential inconsistencies in data definition and data fragmentation across the organisation. Defining strategic data integration requirements needs to address:

- (1) general data integration issues. This includes defining a core set of data integration requirements across two or more divisions and their subordinates. As discussed in Section 1, a prevailing problem in Defence is the absence of effective data integration, data integrity and data consistency checking mechanisms, rendering much data unusable. Defining strategic data integration requirements aims at synchronising the definition of requirements for data integration across interconnected business units to ensure data consistency and seamless data integration. Activities of a 'data engineer' required to achieve this are:
  - a) effective communication with related business units across Defence. These lines of communication are shown as green arrows in Figure 1. For each business unit, the requirements for data might come from several higher level business units (refer to Figure 1). Similarly, data requirements of each business unit might influence more than one subordinate level. Therefore, an effective method to define data integration requirements across many stakeholders is to develop a community of interest COI (DoDI 2013) consisting of all stakeholders. The main advantage of a COI approach is the focus on the value and consistency of data among all stakeholders. COI can also be used for defining core data concepts discussed in Step 1.
  - b) Develop data quality standards across interconnected business units. Research has shown data quality has a lasting impact on the success of organisations (Wang, Storey et al. 1995) and the cost of using inaccurate data is very high (Paradice and Fuerst 1991, Strong, Lee et al. 1997). Therefore, developing data quality standards at all levels of the organisation is critical. The following specific issues in data quality standardization are critical in Defence:
    - Developing a unified policy on naming conventions. Defence has traditionally used many names, even though sometimes only slightly different, for the same data items in different divisions and groups. This frequently leads to confusion and inconsistencies.
    - Developing a data naming registry. A centralized database should be developed to serve as a central data name management mechanism where core data concepts, data syntax, semantics, meta data requirements, scope, storage location, and responsible personnel for managing the data are clearly defined and shared across Defence (with due consideration for possible access and security concerns). Structural and architectural aspects of data may also be addressed at this stage.
    - Developing and using a Defence asset unique identification (ID) number management system where a unique ID number is defined for each key Defence asset.

To help resolve these issues effectively, establishing a Whole-of-Defence Data Management and Analytics Centre will be helpful for managing the development of data quality policy and guidelines, and to provide assistance in Defence critical data analytics.

- (2) Data management system integration issues. Our research has found that Defence has little centralised enterprise level data management and data is distributed across a very large number of different database



systems. Database management systems (DBMS) in each group or branch are introduced and developed separately without proper data integration plans. This, in turn, results in isolated solutions which lead to fragmented and inconsistent data sets across Defence. It is possible to institutionalize policies to ensure a compatible set of DBMS being used in Defence, albeit we believe developing one centralized database management system covering all aspects of Defence data requirements might not be feasible. For the existing DBMS, data integration software should be developed to ensure critical data can be seamlessly accessed and integrated across different groups and organisations.

### **Step 3: Define Measures of Data Effectiveness**

There are many metrics to assess quality of data (Ballou and Pazer 1985, Redman and Blanton 1997, Ballou, Wang et al. 1998). Examples of these include:

- (1) Measures relating to data value in decision support
- (2) Measures relating to consistency and integrity of data
- (3) Measures relating to data redundancy
- (4) Measures relating to data structure and how well data is defined
- (5) Measures relating to data completeness

It is important to define appropriate and relevant data quality measurements aligned with Defence business while monitoring the ongoing effectiveness of the data. For each metric a relevant scale can be defined; for example, “The percentage of data used in decision support” can be the scale used in measuring the value of data, while “The percentage of inconsistent data found at a target level” can be used to measure the consistency and integrity of the data. Most of the measures defined above can be collected automatically without requiring significant extra manual effort.

### **Step 4: Assessment and Gap Identification**

Assessing the value and quality of existing data is essential for understanding data gaps in the operation of Defence. Technical quality issues are relatively easy to discover and correct, however assessing the value of data to a business unit both strategically and operationally is a much harder task. The main measure of assessing the value of data should be based on the utility of the data in decision making, as discussed in the previous step. The fundamental aim of this step is to identify data integration gaps in the organisation. An assessment of the value of data and the quality of data go hand-in-hand and should focus on:

- Verifying appropriate level of data value and quality,
- Identifying and removing redundant data or data with little value (i.e. does the benefit of having data outweigh the effort in acquiring or collecting the data),
- Identifying, correcting or removing inconsistent or ill-defined data and data structure,
- Identifying gaps in the data (and the associated quality) and identifying potential new data items to collect and the required level of data fidelity.

### **Step 5: Develop and Execute Strategic Data Action Plan**

Addressing gaps identified in the previous step may be achieved through an action plan at the following levels of abstraction:

- Policies and governance. To achieve best results, clearly define whole of Defence SDP policies and clearly define roles and responsibilities.
- Human resources. Train or recruit workers with knowledge and skills in data definition, modelling, integration, management, and analytics.
- Infrastructure. Invest in infrastructure for storing and analyzing data commensurate with the scale and scope of data in each business unit.

## **4. CONCLUDING REMARKS AND FUTURE WORK**

Empirical research confirms the advantages of data-driven decision making and the positive impact this can have on business performance (Brynjolfsson, Hitt et al. 2011, HMGovernment 2013). This equally applies to Defence business, as using data effectively creates further insight, positions new capabilities and strengthens existing capabilities. To achieve information superiority requires access to sufficient, consistent, high quality and highly integrated data. We believe it is time to develop and implement a new strategic data planning methodology (SDP) in Defence. This paper proposes a SDP framework which includes a methodology and a set of practices and principles to help strategic data planning in Defence. The proposed framework is based on our hands-on experience as well as international best practice. The significance of the proposed SDP is that it:

- Provides a framework for data planning and a foundation for advanced analytics, thereby, enabling more accurate and valuable support to decision making.

- Provides an integrated data management framework and principles that can be used to promote best practices for data management.
- Provides a set of recommendations for addressing SDP problems for Defence.
- Makes data problem identification (such as data interdependency issues) possible.
- Focuses on the value of data and therefore reduces overall organisation cost of data creation, storage, and processing
- Advocates higher level analysis and analytics in support of critical decision making

The five steps of a strategic data planning process discussed above should become a continuous and integrated business process. It is also worth noting that the five steps introduced here should not be separate activities, but rather be intertwined and iterative. Specifically, Steps 2, 3, 4, and 5 should be conducted in parallel with Step 1. For instance, it is necessary for a data engineer to acknowledge data integration requirements (Step 2), and the value of data (Steps 3 and 4) for data definition in Step 1 to be effective.

Applying and validating the proposed SDP framework in targeted areas is part of our future research direction and will be reported in due course.

## REFERENCES

- Adelman, S., L. T. Moss and M. Abai (2005). Data strategy. Addison-Wesley.
- Ballou, D., R. Wang, H. Pazer and G. K. Tayi (1998). "Modeling information manufacturing systems to determine information product quality." Management Science 44(4): 462-484.
- Ballou, D. P. and H. L. Pazer (1985). "Modeling data and process quality in multi-input, multi-output information systems." Management science 31(2): 150-162.
- Berner, M., E. Graupner and A. Maedche (2014). "The Information Panopticon in the Big Data Era." Journal of Organization Design 3(1): 14-19.
- Brynjolfsson, E., L. M. Hitt and H. H. Kim (2011). "Strength in numbers: How does data-driven decisionmaking affect firm performance?" Available at SSRN 1819486.
- Bulluss, G., N. Tay, K. O'Shea and P. Pong (2014). Innovations in Understanding the Whole of Australian Defence System of Systems. IEEE 9th International System of Systems Engineering (SoSE), Adelaide, Australia.
- CDMRT. (2015). "Capability Development Information Management", from <http://spintranet.defence.gov.au/CapabilityDevelopment/ICTSupport/Pages/default.aspx?showtabno=1>.
- DCDC (2013). Joint Doctrine Note 2/13, Information Superiority. M. o. Defence, The Development, Concepts and Doctrine Centre (DCDC), Ministry of Defence, UK.
- DODAF (2010). DoD Architecture Framework, version 2.02. Department of Defense, USA, Department of Defense Architecture Framework Working Group.
- DoDI (2013). INSTRUCTION: Sharing Data, Information, and Information Technology (IT) Services in the Department of Defense. D. o. D. DoD CIO, Department of Defense.
- DPMIS. (2015). "Defence Preparedness Management Information System (DPMIS) Suite." from <http://intranet.defence.gov.au/vcdf/sites/JCC/comweb.asp?page=71348&Title=DPMIS>.
- Goodhue, D. L., L. J. Kirsch, J. A. Quillard and M. D. Wybo (1990). "Strategic data planning: lessons from the field." MIS Quarterly: 11-34.
- Goodhue, D. L., J. A. Quillard and J. F. Rockart (1988). "Managing the data resource: a contingency perspective." MIS quarterly: 373-392.
- Hackathorn, R. D. and J. Karimi (1988). "A framework for comparing information engineering methods." MIS Quarterly: 203-220.
- HMGovernment (2013). Seizing the data opportunity: A strategy for UK data capability. D. f. B.-I. a. Skills, Her Majesty's Government
- Lederer, A. L. and V. Sethi (1988). "The implementation of strategic information systems planning methodologies." MIS quarterly: 445-461.
- Lee, S.-g. (2007). Challenges and Opportunities in Information Quality. E-Commerce Technology and the 4th IEEE International Conference on Enterprise Computing, E-Commerce, and E-Services, 2007. CEC/EEE 2007. IEEE.
- Lo, E., T. Weir, G. Bulluss and Y. Wang (2015). Improving Project Budgeting through Slippage Analysis in Program Viewer. Systems Engineering Test and Evaluation Conference 2015 Canberra, Australia.
- Martin, J. (1982). Strategic Data-Planning Methodologies. Englewood Cliffs, Prentice-Hall, NJ.
- Martin, J. and C. Finkelstein (1988). Information engineering. Savant.
- Paradice, D. B. and W. L. Fuerst (1991). "An MIS data quality methodology based on optimal error detection." Journal of Information Systems 5(1): 48-66.
- Perry, W. L., D. Signori and E. John Jr (2004). Exploring information superiority: a methodology for measuring the quality of information and its impact on shared awareness. Rand Corporation.
- Redman, T. C. and A. Blanton (1997). Data quality for the information age. Artech House, Inc.
- Simon, A. J. (2006). Overview of the department of defense net-centric data strategy, DTIC Document.
- Strong, D. M., Y. W. Lee and R. Y. Wang (1997). "Data quality in context." Communications of the ACM 40(5): 103-110.
- USJFCOM J87 (2004). Joint Coalition Data Strategy Status Brief, USJFCOM, J87, JBMC2 Integration.
- Wang, R. Y., V. C. Storey and C. P. Firth (1995). "A framework for analysis of data quality research." IEEE Transactions on Knowledge and Data Engineering 7(4): 623-640.