FACILITATING DECISION MAKING, RE-USE AND COLLABORATION

A Knowledge Management Approach for System Self-Awareness

Shelley P. Gallup, Douglas J. MacKinnon, Ying Zhao, John Robey and Chris Odell
Distributed Information Systems Experimentation (DISE) Group
Naval Postgraduate School, Monterey, CA 93943, U.S.A.

Keywords: Program self-awareness, Decision making, Knowledge management, Data mining, Text mining, Cluster, Association, Visualization, Search.

Abstract: Decades of reform have been largely ineffective at improving the efficiency of the DoD Acquisition System, due in part to the complex processes and stovepipe activities that result in duplication of effort, lack of re-use and limited collaboration on related development efforts. This research applies Knowledge Management (KM) concepts and methodologies to the DoD acquisition enterprise to increase "Program Self Awareness" (Gallup and MacKinnon, 2008). This research supports the implementation of reform initiatives such as Capability Portfolio Management and Open Systems Architecture which share the common objectives of reducing duplication of effort, promoting collaboration and re-use of components. The DoD Maritime Domain Awareness (MDA) Program will be used as a test case to develop prototype data schemas and apply text and data mining tools to identify duplication and/or gaps in the features of select MDA technologies. This paper will also provide the foundation for future development of the System Self-awareness concept and KM tools to support decision making and collaboration in diversified commercial and military applications.

1 INTRODUCTION

1.1 Decision Making and System Self-awareness

As development and management of systems of systems (SoS) has progressed, these systems have increased in component, organizational, technical and management complexity. What has emerged is a set of increasingly challenging tasks for decision makers as they seek to know the "edges" of the systems acquisition efforts, development of technical components, funding lines associated with specific elements of effort, and other, often unknown dimensions. This effect is noted in both civil and military development and acquisition programs. This effect surfaces in a macro sense in the difficulty that decision makers express in obtaining constant awareness of what is going on in their domains of decision making because information that is needed is increasingly overwhelming. And, methods to sort information have remained largely undeveloped past use of flat-file databases, some simple search tools and visualizations using PowerPoint.

The interface between what is cognitive on the part of decision makers, and methods for understanding what is important across the span of SoS and attendant documentation may be expressed in a range of terms. We have borrowed from notions of "awareness" in this work, and are employing the term *Self-awareness* of a complex system as the collective and integrated understanding of system attributes and surrounding environment by decision makers. A related term, "situational awareness" is used in military operations, but carries with it a sense of immediacy, cognitive understanding of relationships in the moment. We seek understanding of past, present and future in our view. Here, system self-awareness allows decision makers to recognize relationships among attributes and seize collaboration and re-use opportunities to support cost effective management of a complex system. DoD acquisition is an extremely complex system, comprised of the myriad of stakeholders, processes, people, activities, and organizational structures involved, which navigate an array of procurement processes in an uncertain environment, to deliver useful military capability to users at the best possible value to the government. Acquisition reforms have

Gallup S., J. MacKinnon D., Zhao Y., Robey J. and Odell C. (2009).
FACILITATING DECISION MAKING, RE-USE AND COLLABORATION - A Knowledge Management Approach for System Self-Awareness.
In *Proceedings of the International Conference on Knowledge Management and Information Sharing*, pages 236-241
DOI: 10.5220/0002332002360241

been largely ineffective at improving the efficiency of the development and acquisition system, due in part to stovepipe activities that often result in duplication of effort, lack of re-use and collaboration on related development efforts. Achieving Program Self-awareness in the DoD program acquisition is a necessary goal, if savings are to be achieved across DoD, while improving capability.

This research intends to establish strategies and methods using advanced Knowledge Management (KM) concepts, methodologies, and apply them to various needs of DoD acquisition program managers. In general, we seek to determine how KM tools and methods may be employed to improve decision making, enable collaboration and re-use of components of a complex system. We believe that self-awareness, enabled by KM tools, will dramatically improve decision making and collaboration.

1.2 System Theory of Organizations

The Congruence Model (Nadler & Tushman, 1997; Mercer Delta, 1998) (Figure 1), is a useful framework from which to consider the implications of SoS projects. In an open system, the elements of the system have attributes in the form of qualities or properties that are mutually affecting all other elements, including the possibility of constraining each other. This mutual nature within the SoS also is affected by relationships to the environment. The framework provided in this model includes the internal mechanisms of people, technology, formal and informal systems, embedded within the effects of strategy and measured in some way, through metrics. As system improvement improves, so does it's "fit" between resources, strategy, work of people within the project, and metrics showing progress or improvement. It is here that the difficulties arise for management of the specifics of a SoS project. That is, as the number of individual elements of the SoS increase, so also will the need for definition of what each of those elements means, in context with each other, and to the program overall. Managers often lack the means to construct situational understanding of individual elements and their relationships to both the whole of the program, and to the other individual elements. What we advocate here is a framework to document these relationships, dynamically through the normal documentation that emerges in project development, and create tools to enable SoS managers to refine investment requirements, limit redundancy throughout the project, and enable reuse of system elements. Our research suggests that KM tools can be used to discover and monitor such system interdependencies from dynamic and real-time data and form a sort of "glue" among components, therefore ultimately improve the overall "fit" of the complex system, thereby improving output efficiency and facilitate implementation of policy objectives such as Capability Portfolio Management (CPM) and Open Architecture (OA) in the DOD acquisition context (Section 3).

1.3 Analytics

Data mining is a class of information analytic methods that looks for hidden patterns in a collection of data, typically in structured data which are stored in relational databases, Excel and XML files. Patterns can be used to predict future behavior (Turban, Shardra, Aronson, & King, 2008). Text mining is the application of data mining to non-structured or less structured text files, for example, word, pdf, PowerPoint documents, memos and emails. Much of the data in the world remains unstructured despite rapid development of database and data management technologies. Every organization must analyze a large amount of unstructured data to create management summaries and other decision aids. In this analysis of text rendered information, one task is usually to separate meaningful and important keywords from the remaining words used in the document, to create themes or categories for all that follows. This is very similar to an ethnographic coding methodology (Schensul et al.., 1999). As an example, when an unstructured data set is used to describe an object, one often wants to extract the features of the object, i.e. a set of keywords representing important properties of the object. An object can be a DoD system or an entity of interest. Text mining is very important for developing new meanings and relationships from unstructured data to support decision making.

The set of KM analytics used in this research and their contextual definitions includes the following:

<u>Cluster</u>: Objects can be grouped together based on keywords or attributes that describe their properties.

<u>Association</u>: Objects share properties and can therefore be linked together, or associated.

<u>Social network</u>: Behavior of objects in an interconnected network.

These analytic tools may be applied to both structured and unstructured data to confirm previously determined patterns, or to surface patterns that as yet are unknown.

1.4 Data Warehouses and Data Marts

Data mining techniques generally require that a set of data (data warehouse or data mart; Turban et al, 2008).be made available, and it is on this data set that various data mining algorithms can be applied and subsequent analysis can be performed.

The development of data warehouses into the structured form required to support data mining is not a trivial process. The data warehouse needs to be developed to support the functional area being supported and have the following fundamental characteristics: Subject oriented, integrated, time-variant, and nonvolatile. The data warehouse may also be developed to include the following capabilities: web-based, relational/multi-dimensional, client/server, and include metadata (data about data) (Turban et al, 2008).

Unstructured data or text documents often reside in directories or folders. Such repository style data warehouse or data marts are typical in real world. These repositories do not require the same conditioning of the information found in relational databases, and if it is possible to properly analyse these data files, this will represent a great savings in time and effort.

1.5 Visualization and Search

The KM tools used in this research are used to highlight relationships among object “features” to support decision making. For the purposes of this research, a "feature" is a marketable behavior or property of an object (see Figure 2, the range of features inherent to the Maritime Domain Awareness effort). In this research, the use of KM tools is applied to the notion of technical features, using the following KM tools:

Visualization: Use clusters and associations in a visualization of the data to help decision makers to see the “big picture” and understand results. Displaying the links of the objects in a network format can help visually validate links among objects, and also identify key objects in an interconnect environment.

Search: Clusters and associations need to be resolved from the unstructured data, noted and mapped to support. This effort becomes critical when analyzing unstructured data for two reasons, specifically:

1) Searching for features, often represented as keywords in multiple text documents, can be overwhelming. A search concept called anomaly search, which separates unique and interesting features of the programs from other features, can be helpful.

2) Searching also provides for mapping newly discovered keyword associations back in the original documents for validation.

2 DOD PROGRAM ACQUISITION

The Department of Defense (DoD) fiscal year 2009 budget for Research, Development, Test and Evaluation (RDT&E) and procurement exceeds $180B (Gates, 2009). Given such huge budget outlays and the increasing pressures of shrinking discretionary budgets and fragile economy, the DoD Acquisition System is the subject of intense scrutiny from government oversight activities, industry, and the general public. This scrutiny has been amplified by highly publicized acquisition program failures, continued cost and schedule overruns and lengthy development cycles.

DoD acquisition has endured an environment of seemingly perpetual reform to arrest this chronically poor performance, resulting in complex acquisition process models, increased executive oversight, and incremental policy changes. The effectiveness of acquisition reforms has yet to be evidenced in the overall performance of the DoD Acquisition System. Other models for improvement have not had much effect. Independent and government chartered studies and reports have repeatedly highlighted the need for improved systems engineering and business processes to incorporate best practices from the commercial sector.

The DoD has embraced several recommendations from these critical reports and moved to adopt several commercial best practices and process initiatives. Two such policy initiatives relevant to this research are the adoption of Capability Portfolio Management (CPM) and Open Architecture (OA) approaches.

In 2006, the Deputy Secretary of Defense released a memorandum to introduce the Capability Portfolio Management (CPM) approach to DoD Acquisition. The intent of exploring the CPM approach was “to manage groups of like capabilities across the (DoD) enterprise to improve interoperability, minimize capability redundancies and gaps, and maximize capabilities effectiveness.

Joint capability portfolios will allow the Department to shift to an output-focused model that enables progress to be measured from strategy to outcomes. Delivering needed capabilities to the joint warfighter more rapidly and efficiently is the ultimate criterion for the success of this effort." (Deputy Secretary of Defense, 2006). Open Architecture (OA) is critical in the design of software intensive systems has been the focus of the Navy PEO-IWS Software Hardware Asset Reuse Enterprise (SHARE) Repository, which serves as a searchable library of ship combat systems software and related assets available for re-use by eligible contractors (Johnson & Blais, 2008). CPM and OA are relatively early in their implementation and address different levels of the acquisition process, but reflect the overarching DoD goals of improving decision making regarding systems of systems (SoS) acquisitions to avoid duplication, identify gaps, and decrease costs and development times.

The tools and processes used by acquisition decision makers to support implementation of CPM and OA are not well defined. A fundamental requirement of both CPM and OA approaches is that acquisition managers develop an awareness of related efforts and activities across an enterprise and/or community of interest (COI) to identify duplication of effort, capability gaps, re-use and collaboration opportunities. It is the premise of this paper that development of improved "Program Self-awareness" is fundamental to the success of the CPM and OA reform initiatives.

3 A CASE STUDY: MARITIME DOMAIN AWARENESS

The DoD Maritime Domain Awareness (MDA) Program was used as a case study for this research. Application of KM decision support tools provided normalized "views" of program elements and attributes, termed "features," to support informed program decision making. The premise of this research is that application of KM tools will improve Program Self Awareness and support the informed decision making required to realize the full potential of CPM and OA initiatives.

Figure 2 also represents what program self-awareness embodies in the MDA COI, supported by collaboration and use of KM tools to enable improved decision making (Gallup and MacKinnon, 2008).

The National Plan to Achieve Maritime Domain Awareness (MDA) from October 2005 defines the Maritime Domain as "all areas and things of, on, under, relating to, adjacent to, or bordering on a sea, ocean, or other navigable waterway, including all maritime-related activities, infrastructure, people, cargo, and vessels and other conveyances." Additionally, it defines MDA as "the effective understanding of anything associated with the maritime domain that could impact the security, safety, economy, or environment of the United States." The stakeholders in this enterprise make up the Global Maritime Community of Interest (GMCOI), which includes "federal, state, and local departments and agencies with responsibilities in the maritime domain. Because certain risks and interests are common to government, business, and citizen alike, community membership also includes public, private and commercial stakeholders, as well as foreign governments and international stakeholders." (Department of Homeland Security, 2005)

The problem set that faces the Navy, as a key member of the GMCOI, is that "commanders lack access to, and the ability to process and disseminate, the broad spectrum of information and intelligence that enables cooperative analysis necessary to understand maritime activity in their area of responsibility, and requisite to early threat identification and effective response against these threats; and when appropriate, to enable partners to respond" (U.S. Chief of Naval Operations, 2009). Navy MDA is key to addressing this problem set because it will "enable the warfighter to sustain decision superiority to successfully execute its missions. MDA is fundamental to decision making superiority at all levels of command" (U.S. Chief of Naval Operations, 2009). The Navy plans to improve the following capabilities to achieve MDA; "focused data collection; technological enhancements; greater cooperative information sharing; supporting enduring and emerging maritime security partnerships; and the professional development of navy personnel within the maritime operations.

We began at NPS by using knowledge management tools from Quantum Intelligence, Inc. such as Collaborative Learning Agents (CLA) (Quantum Intelligence, 2008) and expanded to other tools, including AutoMap (Carnegie Mellon University, 2008)

3.1 Apply to Structured Data

Each year, the Distributed Information Systems Experimentation (DISE) group at the Naval Postgraduate School (NPS) provides standard methodologies for defining metrics, collecting data and performing analysis used in large-scale experimentations that assess and evaluate new systems and technologies for Navy acquisition.

Figure 3 shows a sample of 30 MDA processes (i.e. the vertical list in the matrix) defining the workflow – the activities that constitute Maritime Domain Awareness – and metrics for evaluating the insertion of solutions measured against 11 MDA Spiral-1 technologies (i.e. the horizontal list in the matrix) assessed in FY2008. System features are marked with an x or * if it was helpful in an MDA process. This is an example of structured data emerging from associations with MDA technologies.

Cluster analysis was then employed, using the unstructured data found in project documents, as an alternative. We first clustered the 11 technologies into 5 clusters and associating with all 30 intended program features. More weight was placed on less common features, e.g. features that appear in less than five systems. The colors show the five clusters (Figure 5), where three clusters (blue, yellow, green) are grouped by the less common. The purple cluster contains only one system which has unique features only to itself. The red cluster contains three technologies that share ten common features, i.e. the features appear in more than five technologies. The clustering results would facilitate decision maker's investment of resources in some areas, and scaling back in others. For example, a system including a unique feature may be considered for additional or sustaining resources because it is fulfilling a program requirement and is not found anywhere else within the group of technologies being analyzed. The technologies that share common features could be merged, etc.

We then applied an association algorithm to look into details of how these technologies are related. In Figure 4, MDA associations illustrate how many features two systems share (e.g. CMA vs. Global Trader). In Figure 5, the associations are shown in a social network overlaid with the clusters from Figure 3. This allows highlight more meaningful links among technologies.

3.2 Apply to Unstructured Data

In order to look into more detailed inter-connections among MDA technologies, we took a few sets of unstructured documents that are generated from experimentation, for example, documents belonging to programs such as: CMA, TAANDEM, and PANDA ranging from initial requirements, to designs, architectures, testing, and fielding reports. We applied text mining to each individual set of documents representing these technologies and extracted initial feature-like word pairs, then applied an anomaly search algorithm to separate the interesting, key features from the rest. We used a network visualizer in AutoMap to visualize the relationships of three technologies based on the final selected features as shown in Figure 6.

In Figure 6, three clusters of connected keywords centered around the technologies, CMA, TAANDEM, and PANDA. Keywords describing unique features of three systems are separated and pushed away from the center and colored in green, orange, and yellow. Shared keywords among systems are in different colors in the middle of the figure. Different colors indicate different clusters of centralization among word groups. They are produced using a social network analysis method (Girvin and Newman, 2002) and are connected as if they are in a social community.

4 CONCLUSIONS

Using the DOD Acquisition and Maritime Domain Awareness as examples, we have demonstrated in this paper a set of powerful knowledge management tools applied to both structured and unstructured data to develop system self-awareness for a complex system, in effort to facilitate decision making and collaboration in diversified commercial and military applications. We look to continue refining our methods to further improve self-awareness among multiple systems and search for other applications in which this methodology may be useful.

REFERENCES

Chief of Naval Operations, 2009. *Establishment of the Navy Maritime Domain Awareness Office*. Administrative Message 181837Z MAR 09. Washington, DC.

Carnegie Mellon University, 2008 http://www.casos.cs.cmu.edu/projects/automap/

Deputy Secretary of Defense, 2006. *Capability Portfolio Management Test Case Roles, Responsibilities, Authorities, and Approaches*, Memorandum. Washington D.C.

Department of Homeland Security, 2005. *National Plan to Achieve Maritime Domain Awareness for The National Strategy for Maritime Security*. Washington, DC.

Gates, R., 2009, A Balanced U.S. Military Strategy, *Foreign Affairs*, January/February 2009.

Gallup, S & MacKinnon, D., 2008, Status Assessment of Maritime Domain Awareness Capability Development. Monterey, CA: Naval Postgraduate School.

Girvan, M. and Newman, M. E. J., 2002. Community structure in social and biological networks. http://www.pnas.org/content/99/12/7821.full.pdf.
Johnson, J., & Blais, C., 2008. Software Hardware Asset Reuse Enterprise (SHARE) Repository Framework Final Report: Component Specification and Ontology, Monterey, CA: Naval Postgraduate School.
Mercer Delta, 1998. The Congruence Model: A roadmap for understanding *organizational change*. Boston, MA.
Nadler, D., Tushman M. and Nadler, M. B., 1997. Competing by Design: The Power of Organizational Architecture, Oxford University Press.
Quantum Intelligence, 2008. http://www.quantumii.com
Schensul, Jean J., Schensul, Stephen L., and LeCompte, Margaret D. 1999. Essential ethnographic methods: observations, interviews and questionnaires, Rowman Altamira.
Turban, E., Shardra R., Aronson, J.E. ,& King, D. , 2008. Business Intelligence A Managerial Approach. Upper Saddle River, New Jersey, Pearson Education Inc.

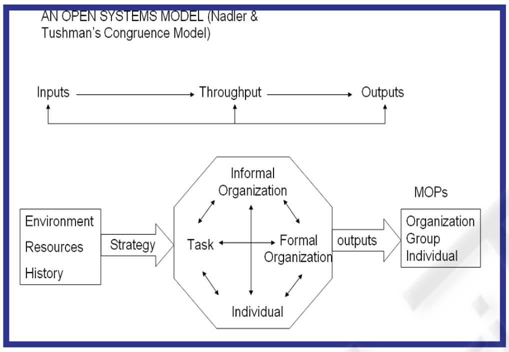

Figure 1: The Congruence Model (Nadler &Tushman, 1997).

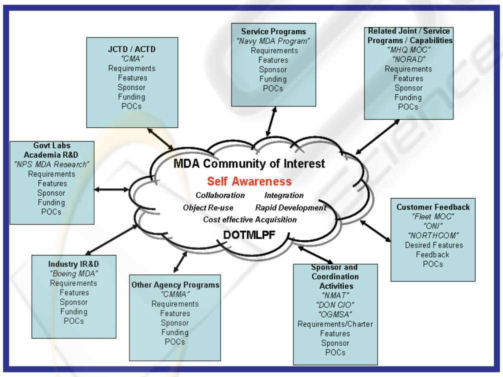

Figure 2: MDA Program Self-awareness feature space.

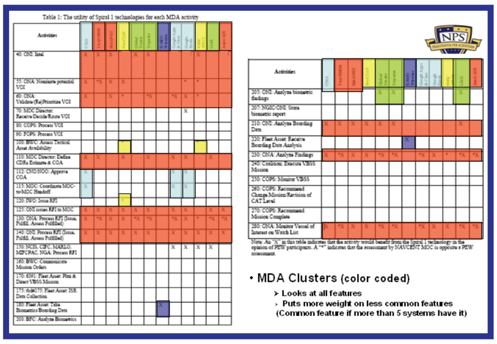

Figure 3: Cluster MDA Spiral-1 technologies based on business processes.

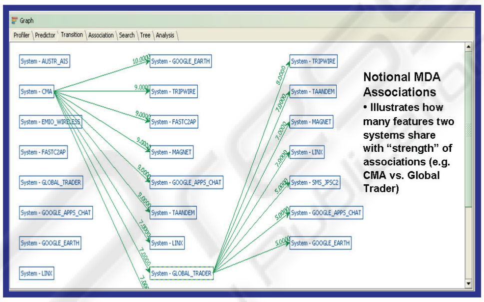

Figure 4: Associations among MDA programs.

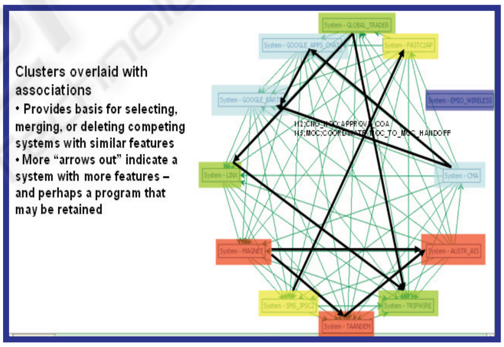

Figure 5: MDA program in a social network. Program clusters from Figure 3 overlaid with associations allow highlight more meaningful links among programs.

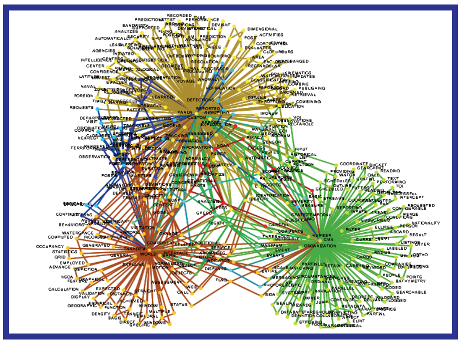

Figure 6: Visualization of MDA program inter-relationships discovered from the shared keywords in their documentation.