Implementing Content Technologies on an Enterprise Scale

A Case Study of the
Implementation of SGML, XML and Content Technologies
within the Canadian Department of National Defence
during the period 1992 to 2002

Joe Gollner, M.Phil.

29 February 2008

Ottawa, Canada

ABSTRACT

Starting in 1992, the Canadian Department of National Defence (DND) undertook to design, develop and deploy an enterprise solution for the management, interchange and publishing of complex bilingual technical documentation. In being undertaken in the aftermath of, not one, but two dismal failures in this very area, the initiative faced almost insurmountable obstacles.

Under the auspices of the Continuous Acquisition and Lifecycle Support (CALS) initiative, DND took on this challenge and led the design, development, demonstration, documentation and deployment of an enterprise standard for Digital Technical Documentation. DND then coordinated the development of a shared application environment that could be distributed throughout the supply chain associated with a wide range of equipment systems. A well thought out standard, coupled with a shared application environment, would constitute a solution to the long-standing challenges associated with managing, cost-effectively, complex technical information holdings.

In developing this shared application environment, the initiative deployed a flexible approach to editing, managing and publishing SGML and XML content and successfully implemented, for the first time, the arduous formatting and control specifications for side-by-side bilingual technical documents which represented a high water mark for complexity in the world of loose leaf publishing. The environment also included mechanisms for publishing XML-based Web IETMs (Interactive Electronic Technical Manuals) that could be distributed across strictly governed Intranets to deliver advanced functionality at a very reasonable cost.

The case study opens with a review of early attempts to modernize the enterprise technical documentation process, starting in 1985, and it traces events as they unfolded through to 2002. A comparison is made between a key business case analysis, conducted in 1992, and the actual implementation experience that then accumulated over a decade of *hard-fought* deployments. The original objective was to reduce technical information lifecycle costs by 50% and it has been determined that these savings have indeed been achieved. The paper concludes with a survey of the lessons learned from the experience and a review of what connections can be made between these lessons and the more recent interest surrounding such emergent standards as **S1000D** and **DITA**.

BACKGROUND

This paper records the life of a long-standing undertaking within DND to modernize the way in which technical documentation was created, managed and published. The author possesses a unique perspective on this effort, having participated in the following capacities between 1989 and 2002:

- Project Officer in the DND Directorate of Computer Systems Engineering and Maintenance (1989 – 1991)
- Lead Analyst and Author for the DND Documentation Modernization Business Case (1991 – 1992)
- Information Architect and Project Manager for the DND Documentation Modernization Initiative (1994 – 1998)
- Solution Architect and Senior Project Manager for the DND CALS Shared Application Initiative for Technical Documentation (1998 – 2002)

INTRODUCTION

This paper could be classed as something of an *über*-case study. It recounts the efforts of a reasonably large organization to come to grips with the magnitude of its information holdings and the complexity of the processes that annually churn through hundreds of millions of dollars. It recounts how a series of efforts were undertaken to apply new ideas to mastering the problems in information management. There were failures to be sure. In some cases, there were disasters of almost biblical proportions. There were successes as well and, happily, this case study can ultimately be classed as a success story in that it proves that the new approaches to managing technical information using open standards and emerging techniques for enabling efficient information reuse *do* in fact work.

For the many people who were involved over the years, this adventure provided an invaluable stock of lessons and a bank of solution components that continue to find new applications even today. The basic notions that launched this undertaking were, at the time, novel, even heretical, and it now seems incredible that these same ideas have become universally accepted and even commonplace.

Reviews of initiatives on this scale and over this time scale are not common and, with this effort, it has become clear that much more time could be productively spent in this direction - mining lessons for those that will follow. And while those that follow may be emboldened by the variety of tools at their disposal, the really hard challenges, those chronicled here and those that have little or nothing to do with syntax or scripts, remain, lurking in the weeds, waiting for the next generation of implementers. It seems only fair to pass on some of what has been learned so at least some of the problems overcome in the past can be avoided altogether in the future.

It would fortunate if someone was in the position to look back over this many years and reflect on the birth, the ups and downs, and finally the outcome of an enterprise-wide solution initiative in the relatively new field of *structured information standards* and their application to the management and processing of information. It would be almost too much to expect that one person, albeit unwittingly, would have had the opportunity to play a pivotal role in almost every phase of this journey. And yet, this case study attempts to summarize just such a history and to do so from just such a privileged perspective.

CONTEXT

Organization. The Canadian Department of National Defence (DND) is the chief protagonist in this case study. Somewhat like "Mini-Me" in the *Austin Powers* movies, DND is small when compared to some military institutions but despite this fact it attempts to acquire at least one example of every conceivable piece of equipment. As can be quickly guessed this leads to serious sustainability problems and these problems have grown at an exponential rate with the introduction of a new generation of high-tech equipment systems. These sustainability problems are then compounded by the erratic funding levels on one hand and a steady increase in operational commitments on the other. As has been demonstrated in some of the current hot-spots, the troops of the Canadian Forces have been exceedingly effective at what they do and they have done so frequently without the levels of logistical support that other militaries sometimes enjoy.

Metrics. In terms of metrics, this enterprise represents approximately 150,000 people, military and civilian, with an annual budget that hovers around $10 billion (all figures are in Canadian dollars). At any one time, the infrastructure is endeavouring to support troops operating in around 20 theatres, including the Gulf Sea, Afghanistan, the former Yugoslavia and some less-than-charming locations. So while not directly comparable to its major allies in terms of size, DND still represents a significant enterprise by any normal measure.

Information Management. In terms of information management, the department's active technical information holdings currently run into the tens of millions of pages. Included within this number are sets of highly volatile information that must adapt in near real-time to the constant adjustment and re-purposing of the main equipment assets – multi-role fighter aircraft, surveillance equipment, long-range transports, armoured vehicles and advanced frigates. While it is a trait shared across the military community, the Canadian military seems pride itself in modifying the configurations of its equipment items to precisely fit a wide range of operational roles and thereby convert already expensive equipment systems into outrageously costly investments.

Bilingualism. As Canada is officially a bilingual country, an additional wrinkle is introduced into the information management matrix. All applications, interfaces and documentation must be prepared and presented in both official languages: French and English. Over several years of experience in addressing this requirement a number of proven practices have evolved for the handling of bilingual material. Perhaps the most significant of these, from an information management perspective, is the need to display content in the two languages so that users can move quickly between the two renditions. As a significant number of users are in fact bilingual it has been found that users routinely cross-check the content between the languages. This is in part a mechanism that ensures that users can work around the inevitable translation hurdles. It also goes without saying the translation represents a huge annual expense and that it introduces a significant task into all content creation and publishing processes.

Publishing Standards. One consequence of bilingualism has been the establishment of arduous standards that apply to how bilingual content is to be displayed in paper and online. In essence, the standards are designed to ensure that content is presented in an equitable fashion, meaning that neither language is given a privileged position, and in a manner that facilitates use by personnel that may be bilingual or unilingual in either of the languages. In print, the primary presentation specification requires the presentation of content in a dual-column, side-by-side aligned format. When these bilingual requirements are added to the inherent challenges of technical documentation, the result is among the most complex of loose-leaf publishing challenges. Online, the Canadian Government Common Look and Feel (CLF) guidelines mandate that all content be presented in a manner that permits toggling between the respective languages. This toggling requirement also applies to all application screens and to help documentation.

IN THE BEGINNING

1985. In and around the time the CALS (then known as Computer Aided Logistics Support) initiative was first coming to life, DND found itself confronted with several million pages of active technical documentation and with a number of acquisition programs underway to deploy new fleets of ships and aircraft. Well in advance of most organizations, DND looked to modernize the infrastructure on which technical information management depended.

ONCE BITTEN

Major Experiment Fails. Before the ink was dry on ISO 8879, DND had launched its first "SGML" (Standard Generalized Markup Language) project. The scope was breathtaking: an integrated content management and publishing system that provided an advanced collaborative working environment for a departmental publishing group. Even when compared to modern content management systems, this application offered an impressive array of features. The initiative, as would be expected, encompassed a major data conversion effort, an effort that ultimately became the birthplace of the OmniMark language (which interestingly remains the leading *content processing* tool in the marketplace). The resulting mainframe-based environment may well be the first "content management system" ever constructed. Despite its leading-edge qualities, or perhaps because of them, this initiative ultimately ground to a halt. The reasons for its failure are difficult to determine precisely, after so much time, but the following doubtless contributed to its downfall:

- system complexity made maintenance increasingly expensive;
- the alignment of the initiative with a centralized publishing bureau biased the requirements towards facilitating production efficiencies and this introduced political resistance elsewhere;
- the implementation focused on introducing a new "typesetting" system which would yield only limited benefits because it remained focusing on facilitating print publishing;
- the introduction of a sophisticated content management system made process modifications unduly difficult to implement and prevented the adaptation of the system to stakeholder needs.

Outcome. Several years later, the system was *decommissioned* and, quietly, the rapidly aging mainframe hardware infrastructure was dismantled and its storage units destroyed. A $17 million dollar initial investment and a phenomenally large volume of converted data were *lost*.

Recollection. Returning to DND Headquarters in 1989, the author of this paper entered the informatics division during the “persecution of the guilty and the celebration of the uninvolved” phase of this project. An entire generation of technology managers within the department had, at a stroke, become declared enemies of SGML and content technologies in general. “SGML” literally became “the acronym that must not be spoken”.

A BUSINESS REVIEW

A Time of Reckoning. At the time the repercussions of this initial disaster were being digested, the department was prompted to look at the problem of technical information management again and this time in the context of a large and high-profile weapon system. After this particular equipment fleet had been forced to convert over one million pages of technical documentation twice in two years, questions started to be raised. In this particular case, an initial conversion effort in the late 1980s had to be almost immediately repeated when the vendor for the initially targeted publishing system declared bankruptcy. The format required for this system, and to which this content had been converted, was suddenly obsolete. At a higher level, a major review of the entire equipment program revealed, to the horror of the General Staff, that documentation maintenance had become one of the single largest line items in the annual budget for sustaining this fleet. Hard questions began to be asked. Surely documentation could not be *that* expensive.

A BUSINESS CASE FOR CHANGE

Detailed Review Undertaken. Responding to the growing concerns over the annual cost of information maintenance, the Department launched a new initiative but this time the focus fell on gathering concrete information upon which management decisions could be based. This review would start with a massive metrics gathering effort to determine how much content was really being managed and how much this was really costing.

The Numbers. For the first time in the Department’s history, and perhaps one of the first times anywhere on this scale, a project team was assembled specifically for the purpose of compiling detailed metrics on enterprise information holdings and the change processes operating upon them. This project involved junior staff foraging through massive documentation warehouses, ones reminiscent of closing scene of *Raiders of the Lost Ark*, pulling randomly selected technical manuals and documenting page counts, distribution lists, change histories, and general content composition. Based on these concrete facts, extrapolations were made to determine an enterprise-wide profile. So cautious was the team that the sampling approach and extrapolations were independently reviewed by a team of academic statisticians. These results, presented in Table 1, were to be presented to the highest levels of the organization and, given the bitter memories that still lingered about past investments in content technologies, there was no room for error or for questions of credibility.

DND Active Technical Documentation Holdings (1992)	
Total number of publications	51,559
Average number of pages	176
Total master pages	9,074,384
Average distribution	239
Total page distribution	2,168,777,776
Average annual rate of change	9%
Total number of master page changes (annual)	816,695
Average cost per master page change (including translation)	$350 / Page
Cost of master page creation (per annum)	$285,843,250
Total change page distribution	195,190,000
Average distribution cost per page	$.50 / Page
Total distribution cost (per annum)	$97,595,000
Total Annual Cost of Publication Maintenance	**$383,438,250**

Table 1
Analysis Results from the
1992 Documentation Modernization Business Case

Observations. The costing factors used in this study underwent an unusual level of scrutiny. The cost associated with change page processing was the subject of an independent time and motion study. Parallel research was undertaken in other government departments which directly corroborated the findings. The high-level results of this research are presented in Table 2. Each change page entailed effort in the identification and verification of the change requirement, research activities involving frequently inaccessible sources and occasionally direct recourse to the equipment item and the drafting of the change (an activity that frequently entailed collaboration amongst contributors). All this was then followed by a tortuous process of review, validation, translation, re-validation, formatting, dissemination and administration. It should be noted that the formatting of change page packages in accordance with the bilingual publishing format constraints significantly increased the costs of what should have been a more straightforward process of *formatting*.

Page Modification Metrics	
	Modification Cost per Page
Problem identification and verification	$50.00
Research / cross-referencing to sources	$50.00
Revision drafting	$70.00
Revision review and approval	$30.00
Translation	$50.00
Change page formatting & packaging	$50.00
Transmittal, insertion and administration	$50.00
Total Cost per Change Page	**$350.00**

Table 2
Time and Motion Study Results (1992)

Analysis Exclusions. Due to the magnitude of the numbers being unearthed, the project team elected not to attempt to incorporate collateral research that could be used to further characterize the status quo. Over the course of this study interesting metrics were established for the rate at which change pages were mis-filed during the distribution process (12%). It was determined that there were over 100,000 filing cabinets deployed within those units of headquarters that dealt with technical information. Meanwhile, parallel studies within the US Military were identifying the high cost of false replacements on equipment components due to inaccurate or out-of-date documentation, on the costs associated with unnecessary re-testing of components by equipment manufacturers due to errors or uncertainties in the maintenance process, and on the overall equipment availability profile given the time requirements for routine maintenance and minor problem resolution. Perhaps most striking of all were the metrics gathered on turn-around times for publication changes with the average being determined to be *18 months*. Publications, in effect, were never up-to-date.

Qualitative Benefits Set Aside. All of this information, although highly relevant, and especially so to the operational units, was deemed to be too *intangible* for the purposes of a strict calculation of quantifiable payback. Fortunately, the projected payback on modernization investments for the management and publishing of technical information were so large that the operational improvements were allowed to stand as purely qualitative benefits. Once these qualitative benefits were re-introduced into the considerations as supplemental to the quantifiable savings that were achievable, the case for taking action was effectively sealed.

Outcome. This business case, which was presented to and accepted by the highest levels of departmental management, identified a concrete opportunity to realize a financial savings well in excess of *$100 Million a year*. The investment required at the time, however, was daunting. The study team documented, and established the costs for, the requirements for an upgraded departmental network infrastructure, the acquisition and deployment of computing hardware to maintenance and operational units, and the development of a departmental protocol for technical information management, publishing and exchange. Taken together, the projected investments were very substantial and, somewhat surprisingly and on a distinctly incremental basis, these investments were in fact undertaken. As many of the recommended expenditures were associated with such rudimentary considerations as providing computing platforms and enhanced connectivity to military bases, these investments were clearly taken for many reasons in addition to the benefits claimed in this early research. Once these investments were made and they were added to those specifically associated with introducing state-of-the-art content technologies, the department indeed began to realize very substantial savings as well as operational benefits.

TWICE BITTEN

A Second Experiment Fails. Emboldened by the results of this quantitative research, the headquarters unit nominally responsible for managing and publishing technical documentation, set out to implement *the solution to rule all solutions*. And once again, the initial investments were directed towards the identification, acquisition and implementation of a *sophisticated content management environment*. The combination of ornate technology and a "monopolistic" service-bureau mentality could only have been entertained with reference to a payback model that offered some quick returns - the proverbial low-hanging fruit. This initiative sought to capture the savings associated with the automation of *formatting processes* and the potential for streamlining in the administration of change pages. But as could have been easily forecast, the combination of premature investments in a complex content management system with a *centrist service model* (and one that ultimately sought to constrain the practices of operational fleets in order to simplify the administrative work) led to a *second major failure*. The equipment management teams simply withdrew their support from the initiative and, in a number of high-profile cases, elected to outsource the publications management activities to industry rather than rely upon the headquarters unit. The actions of these major equipment programs turned out to be highly prudent as the headquarters publishing unit was not able to maintain, or even operate, the sophisticated content management, workflow automation and print-publishing system that had been acquired. Once again the investment quickly became an embarrassing legacy system the retirement of which was quietly performed.

Outlook in 1993. Given the downfall of two major initiatives that sought to bring under control the now well documented costs of the department's technical information holdings, the prospects for any new initiative in this area did not look promising. Undeterred, however, the department mobilized for a third assault.

THE LONG ROAD TO SUCCESS

A New Initiative. In 1993, the department formed, and for the first time properly funded, the DND CALS Office and set as its mandate "to make paperless materiel management a reality by 2002". Its scope was enterprise-wide and it was specifically tasked with aligning departmental efforts with those of its allies and its primary industrial supplier base. In fulfilling this mandate, this office would start by reviewing the events that had transpired in the past and the research that had been conducted.

THE BIG PICTURE

A Review of the Numbers. A detailed review was undertaken of the research conducted in the 1991-1992 timeframe. The cost and potential benefits calculations were given particular attention and new findings, largely imported from the experiences in the first generation of US Department of Defense (DOD) CALS implementations, were added into the mix. The results of this review were made interesting by virtue of the fact that the CALS mandate was such that the scope of analysis was broadened to include all aspects of materiel management including logistics planning, engineering management and training (in addition to publications management). One outcome of the review saw the costs that had been attributed to publications management come into serious question – they were in fact too *low*. A second outcome of the review was significant in that it illustrated two important facts:

- the cost of publication maintenance represented less than half (50%) of the total annual cost of the *information set* associated with an equipment item (including the maintenance of the logistics, engineering and training content as well as that portion appearing in publications);
- the actual annual expenditures on technical information management were in fact heavily weighted towards a specific group of *main flagship equipment fleets* upon which high volumes of engineering changes were being actively implemented.

The departmental understanding of the technical information management problem needed to undergo a root-and-branch adjustment. The annual cost of maintaining the technical information holdings was in reality much higher than had been previously acknowledged and the centers of those costs were now clearly visible. In essence, annual equipment information costs for the department were driven by the following:

- the originally identified costs for managing the departmental technical information holdings
- the significant concentration of the publication management costs around the main *flagship operational equipment fleets* which were undergoing a high level of engineering changes which in turn caused significantly higher rates of publication change (e.g., +20%)
- the effort associated with maintaining the complete technical information set upon which the publications were based, notably the logistics support analysis records, parts information, source engineering data, and training materials,
- the training materials represents a particularly noticeable area of expenditure because the early adoption of new technologies (eLearning) had in fact encouraged a further isolation of training materials from the technical information set (including publications) upon which they were based, with this isolation causing exceedingly high levels of information redundancy (+80%).

What's in a Page? A number of considerations arose during this review. Among the most obvious were the limitations of the concept of a "page" as measure of information volume. It was found however that despite its limitations, the page could not be easily replaced as a measurement unit. Instead, the page came to be understood as a physical instantiation of a complex matrix of underlying information sources. Consequently, the change page became a distillation or representation of a much more elaborate process and therefore came to be the embodiment of a greater set of costs. Whatever weaknesses this approach may have in specific contexts, it remained a useful measure when applied at the equipment program and at the departmental level. It remained a tangible control item that could be counted and around which costs could be crystallized.

A REAL SOLUTION EMERGES

The Real Problem Identified. It had become clear, by this point, that the real problem with how technical information had historically been managed lay in the fundamental disconnect in the various ways that constituent parts of this information was created and managed. On one side, equipment engineering and logistics information was typically created and maintained within the context of a number of *equipment breakdown structures*. As was usually the case, this was complicated by the fact that a constellation of suppliers was responsible for the creation of this information for different equipment components. The scope of this information was typically restricted to engineering and logistics planning data which was to be maintained in a mixture of databases and engineering tools. On the other side, operation, maintenance and training information was produced and maintained as *documentation items*, and these documentation items were produced and maintained in accordance with structures and procedures that served to disconnect it from the structure of the logistics and engineering information that in many ways served as the source. On a grand scale, the problem was recognized as one of massive *content silos*. The problem is illustrated in Figure 1.

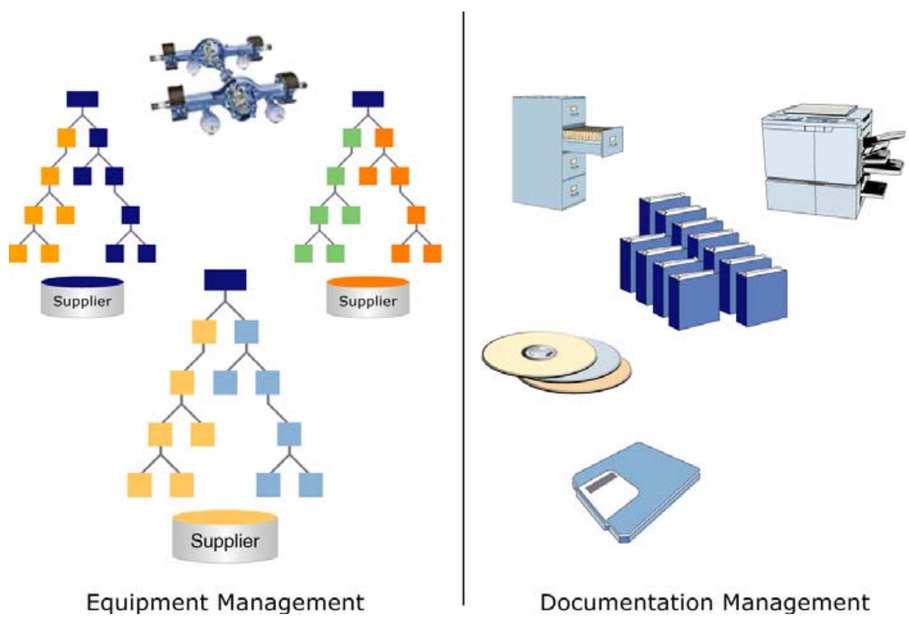

Figure 1: The Real Challenge
Content Silos resulting in Massive Information Redundancy

Any solution to this problem would need to introduce a method for managing documentation items in such a fashion as would enable the integration of all content into *a single management framework* that could be intimately aligned with the *governing equipment breakdown structures* and upon which content modification, automated publishing and output dissemination processes would run.

The Real Solution. The missing ingredient upon which any real solution would rely was a mechanism for establishing an *integrated information architecture* as a core component of an integrated solution that, in turn, would encompass, as its scope, *the full lifecycle of all technical information assets* associated with an equipment fleet.

A governing information architecture needed to conceptually subsume the two domains, both equipment and documentation structures, and then physically enable an instantiation that would allow all information resources to exist, if only for a time, as *a single integrated and interchangeable information set*. This information set would need to be portable amongst potential suppliers and it would need to permit the orderly decomposition of the content into separately managed libraries that could be distributed across the supply chain. On one level, the solution would demand a high degree of abstraction so that one solution framework could be applied in an infinite number of different circumstances. Prior experiences had shown that the tendency to create hyper-specialized content technology solutions for individual types of equipment was not affordable and, even if it were, not sustainable given the natural tendency of complex systems to continuously change. The chosen approach would also need to avoid the proliferation of different solutions to the same basic problem, a circumstance which has seriously undermined the CALS initiatives within several NATO countries. On another level, the solution would have to provide a simple, relatively low-tech, approach to implementing the essential components of the capability so that expense and application sophistication would not thwart progress *as it had in the past*.

The approach that was ultimately taken is illustrated in Figure 2. The solution depended upon the introduction of SGML (Standard Generalized Markup Language) as the modeling and implementation mechanism.

The modeling flexibility of SGML allowed DND to specify a modular information architecture that provided a framework of content structures for accommodating the wide range of individual information items used to document various aspects of modern equipment systems. The information architecture also provided a mechanism for associating these artifacts with the *multitude* of equipment breakdown structures used within the equipment engineering and logistics domains to describe, produce, manage and sustain complex equipment systems. Another key objective of this approach was to enable the exchange of content amongst suppliers, with the foreknowledge that each supplier would be using different content creation and management systems. An important consideration in enabling effective content interchange was the establishment of comprehensive content validation capabilities that could be incorporated into each exchange event.

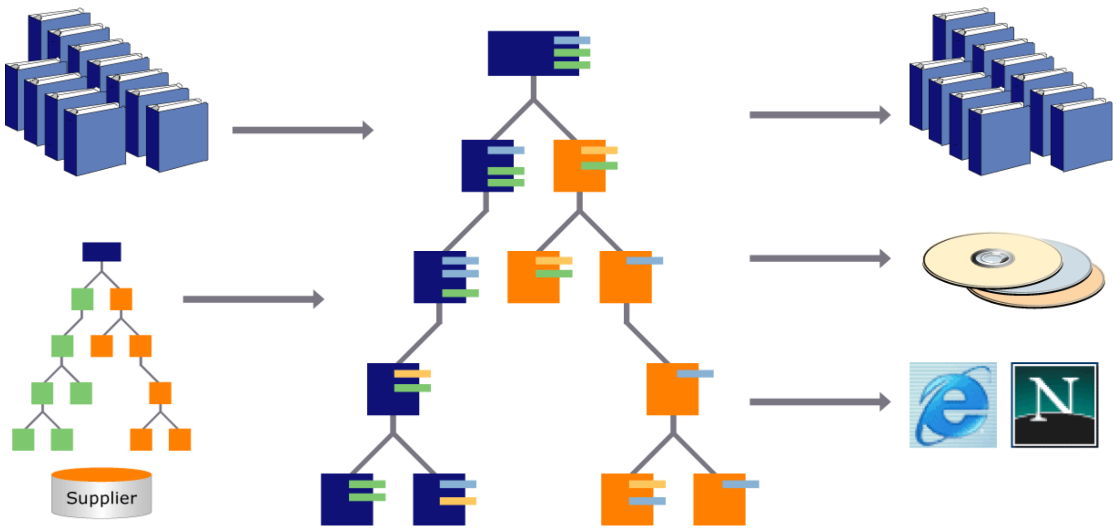

Figure 2: The Real Solution
Standards-based Content Integration

The Path Taken. A number of things are interesting about how this solution came into being. Firstly a collaborative environment was established using a custom Lotus Notes database within which a variety of solution stakeholders could pool their experiences, perspectives and ideas. Past efforts had been dominated by a limited subset of the stakeholder perspectives with this being a major contributor to repeated failures. The facilitation of a free-wheeling but also very intensive discussion about objectives, constraints and opportunities was a critical feature of this new effort. A second step taken was a comprehensive analysis of past investments in this area. Specifically the early implementation experiences associated with the US Military CALS initiatives provided a critical source of input. The international CALS community provided an invaluable venue for networking with other militaries on their experiences, including those from across the US DoD. Numerous standards were collected and subjected to detailed reviews, with these reviews typically entailing contact with the responsible agencies and various implementation projects. Included within the analysis scope were the early releases of S1000D, the Society of Automotive Engineers (SAE) J2008, the Swedish CALS Office (the Grund Document Type Definition), the US Navy Interactive

Electronic Technical Manual (IETM) family of standards and tools, and several others. Also included in the analysis were a number of *foundation* standards, specifically ISO10744 HyTime, and a number of initiatives in domains remote from equipment information but still instructive resources (in particular the Text Encoding Initiative). Both the collaborative mechanisms used to facilitate the analysis across a community of stakeholders and the range and quality of resources made available to the team could be seen as auspicious omens for the future success of the initiative.

ASSEMBLY INFORMATION ARCHITECTURE

One Structure to Rule Them All. The result of this effort was a core *Assembly Information Architecture* that could be used to describe all information resources associated with an equipment item. In finalizing this model, the team had to find ways around many of the weaknesses being seen at the time with *first generation* SGML implementations: loosely defined data structures, out-of-control semantic complexity, impractical application processing scenarios, pervasive biases towards print publishing, grossly inadequate schema documentation, unrealistic illustrative instances and, most seriously of all, the presumed availability of highly-refined, even idiosyncratic, technology components that would be needed to make implementations possible. Each of these historical weaknesses made its presence felt and occasionally caused missteps although the team was able to find effective response measures, the most important of which was the *validation project*.

An aggressive program of validation projects (i.e., pilot projects), each seeking to field an operational information product instead of performing yet another technical prototyping experiment, effectively expelled these ghosts. Projects included implementations in each service element (Army, Navy, Air Force and Communications) and specific efforts were directed towards testing the information model against a wide array of different equipment types. One aspect of this validation program was expressly aimed at winning back a wary stakeholder community by demonstrating regular successes and showcasing the fact that a completely different approach was being taken than had been seen in previous efforts.

The combination of the collaborative analysis and design effort with the program of validation projects saw the assembly information architecture emerge together with a growing constituency of supportive stakeholders. It is easy now that so much time has passed to forget just how difficult this process was to undertake. Many people sacrificed a great deal to make this process of *re-engagement* successful. And these sacrifices were made quite consciously by the people involved with the goal of saving precious time and money in the maintenance of the equipment systems upon which the operational troops regularly relied.

Figure 3 illustrates the general structure of the *Assembly Information Architecture* which stood as the backbone of DND's enterprise approach to planning, creating, managing, publishing and exchanging all forms of technical information.

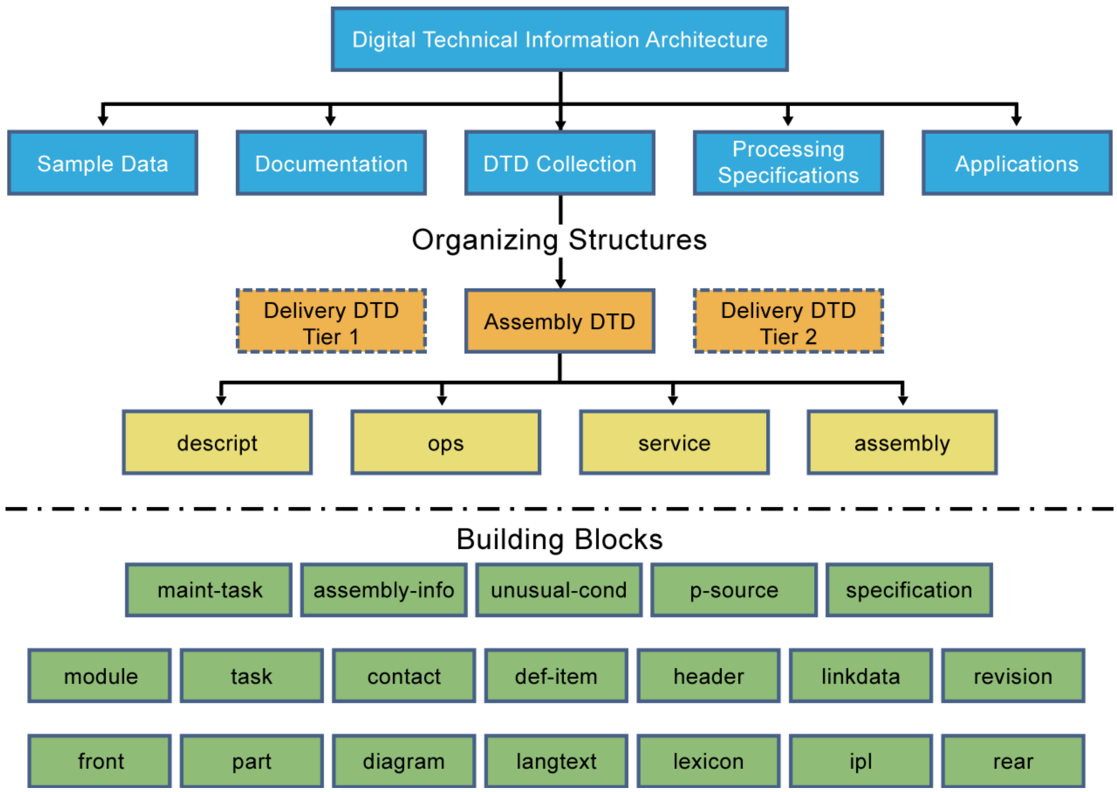

Figure 3: Assembly Information Architecture
A Simple Mechanism for Encoding Large Volumes of Technical Information

Architectural Scope. A number of features are worth pointing out in this architecture. Firstly, the highest level of the structure declares that the information architecture encompasses more than the information model itself. Specifically, the information architecture includes extensive documentation as must accompany any information model for it to be genuinely useful. Accordingly, the information architecture developed under the DND CALS Office was supported by over 1,000 pages of documentation including demonstration scenarios that illustrated exactly how the model was to be tailored to meet specific requirements. A comprehensive set of *functionally-realistic* sample instances was supplied which, in being integrated with the demonstration scenarios, ensured that the documentation could be more than just read - it could be explored. The amount of energy that came to be invested in these functionally-realistic instances was somewhat surprising but these investments did pay strong dividends in the long run. The challenge lay in making these illustrative instances reflective of the types of content that equipment teams actually supported while at the same time ensuring that the content remained *unclassified* so it could be broadly shared for use in the demonstration and testing of deployed solutions. On a marketing level, if it can be called that, the content within the demonstration instances had to be selected so as to be compelling to any equipment team (Army, Navy, Air Force, or Communications) and for this reason electronic components were chosen as representing the most universally applicable.

The solution architecture also included detailed definitions of the main processing scenarios that would be associated with implementing a working solution for any particular equipment fleet. Continuing with the theme of making the information model something that was *tangible*, these processing scenarios were supported by *application components* that could be used to run complete *end-to-end demonstrations* and that would actually produce the types of information products needed. Going still further, these application components were far from simple tools for use in demonstrations. They in fact provided a substantial amount of *production-grade functionality* so that equipment programs could tailor the applications to meet their requirements rather than develop functionality from scratch. The quality of these components was such that the temptation appeared to see them as a central part of the standard and as something that should be centrally maintained on behalf of all stakeholders. This temptation had to be resisted however as it pointed in a direction that would undermine sound strategies for managing the equipment information supply chain.

Information Supply Chain Management. Within contractual scenarios, a firm distinction had to be made between the standard which included all mandatory instructions as well as the information model and supporting documentation, and the shared application components that were provided as *reference resources* that the suppliers could use to accelerate their respective implementations. This was an important distinction to make because, in parallel with these modernization efforts, DND was adopting an aggressive posture towards *outsourcing* and this demanded that individual suppliers remained completely responsible for implementing production solutions for the creation, management, publication and delivery of equipment information. The standard, and the verifications that would be made with reference to the standard, had to remain the key determinate of supplier compliance and thus needed to be differentiated clearly from the application components suppliers may, or may not, use. It was an indication of the success of the standard that numerous suppliers were able to successfully implement content management and publishing environments that conformed to the requirements established by the standard but did so using completely different technologies than those leveraged within the shared application suite of components. The reference application components were found to be useful as resources even in these circumstances, again, because they made the objectives behind the standard, as well as specific requirements, clear.

A Solution Template. It had been learned, largely within the DoD CALS initiative, that the proclamation of a standard effectively solves nothing. Conversely, as DND had so gloriously proven, large scale technology investments in all-encompassing solutions prove to be just as ineffective and frequently far more disruptive. The middle ground was found to be occupied by something that might be described as a *solution template*. The Assembly Information Architecture established by DND represented just such a solution template as it set forth a suite of protocols governing information creation, management and exchange, established the fundamental rules that ensured that the information provided would support the full array of downstream requirements, and provided a set of application components for use in kick-starting solution implementations without inhibiting the innovation, or responsibility, of individual suppliers. Perhaps most importantly of all, the Assembly solution template enabled equipment teams to effectively and inexpensively adapt the provided capabilities in order to solve the many practical problems that invariably appear within any real equipment program. It was this adaptability that made the Assembly Information Architecture something closer to being a solution template than simply being either a mandated standard or a one-size-fits-all system.

Adaptability as a Core Feature. If there was a defining feature of the Assembly Information Architecture it was adaptability. Accordingly the architecture made critical distinctions between the structures in which information was to be created, the contexts within which it would be managed and the organizing structures to be used to plan and execute processing scenarios. All organizing structures, including the governing Assembly structure, were supported by a library of information building blocks which could be used to create new organizing structures should these become necessary. It was not imagined that the initial analysis and information modeling effort, however exhaustively it was pursued, would ever achieve perfection. The ability to construct new organizing structures, sometimes referred to as “hub structures”, “book plans”, “views” and “maps”, was assumed to be a natural part of implementing a working solution. As one example, early in the life of the standard, a specialized organizing structure had to be constructed to support a fighter aircraft fleet with undeniably unique requirements.

By a similar token, the building block library could also support the modification of individual building blocks to produce equipment-specific variants of the building blocks where necessary. It could even support, as was openly encouraged, the development of new building blocks to support specialized needs. It was identified that where such efforts were undertaken, the resulting building blocks would become a candidate for inclusion within the library of building blocks made available to all equipment fleets under the auspices of the overarching architecture. Interestingly, in each of the circumstances where the possibility was identified for the creation of a new building block, the active participation of the architecture support team determined that the requirement, once fully understood, could be effectively supported, sometimes through minor adaptation, using already available building blocks. So although adaptability was a design feature of the information architecture, the team in place to support the standard worked energetically to ensure that any adaptations that were made were indeed legitimate and not merely a result of poorly understood requirements or unfamiliarity with the architecture and its goals.

As was illustrated in the example where a new organizing structure was needed by an aircraft fleet, legitimate adaptation was something that was genuinely encouraged and for the time between 1997 and 2002 a respectable budget was in place to ensure that the standard was supported by a capable team that could drive this process of *responsive adaptation*. This ensured that equipment fleets saw that their requirements could be addressed quickly by working *with* a headquarters agency and this in turn would encourage collaboration as a preferable alternative to isolated and redundant investment. The success of this model, for the time period in which it was actively supported, provides an instructive lesson for all organizations wishing to realize benefits from the coordinated adoption and adaptation of content standards. A standardization effort is best understood as a *service* that must, at the end of the day, consistently deliver solutions and help their stakeholders save money and time in achieving their goals. Only then will the enterprise goals of improved standardization and investment consolidation be realized and sustained.

The intention throughout was to avoid the circumstance in which many organizations found themselves wherein slight differences in requirements gave rise to multiple, often competing, solutions. Instead, it was argued, all equipment teams would be enabled to work within a shared framework and thereby to leverage what was common without sacrificing their ability to cost-effectively address genuinely unique requirements. In practice, the experience of implementation efforts between 1997 and 2002 validated that this goal was both achievable and worthwhile.

Information Model. Given the scope set out for the architecture, as that of supporting the definition, creation, management, publication and exchange of all technical information associated with any equipment system, it comes as something of a surprise to learn that the Assembly Information Model, as a DTD (Document Type Definition), contained fewer *elements* than the Hypertext Markup Language (HTML) DTD. This conciseness stems from the fact that the concept behind the information model was fundamentally simple.

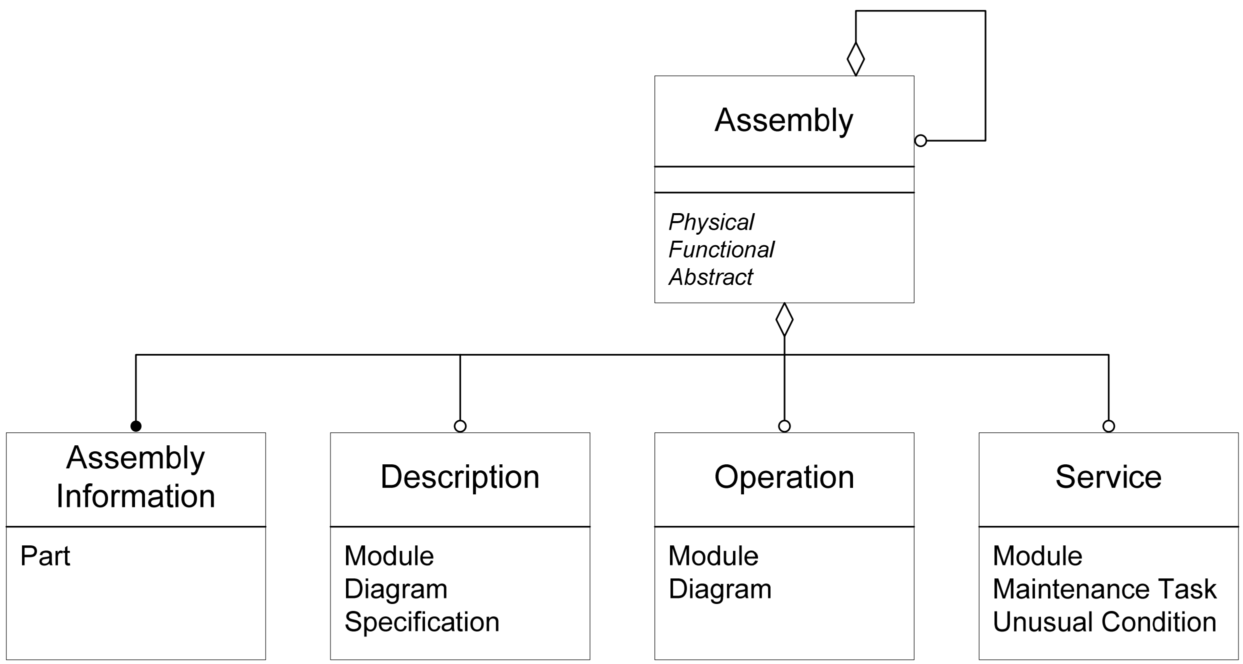

Figure 4: Recursive Assembly Structure
Reflecting the Equipment Breakdown Structure and Associating Information Objects with Assemblies

Assemblies represented the universal building block for modeling the structure of any equipment system, including its various physical, functional and abstract breakdown schemes. Information objects, addressing specific content requirements, could then be associated with the description, operation, or servicing of any assembly. Using these simple recursive and associative structures, the assembly hierarchy could be used to document equipment systems of *massive* complexity.

Publishing Information Products. The availability of delivery structures then allowed the assembly structure to act as the content source for an unlimited range of information products. As part of the overall solution template, information products included Interactive Electronic Technical Manuals (IETMs), online renditions enabling web access, as well as high-quality printed manuals that conformed to a number of different output specifications. The intent of the overall architecture, from the perspective of publishing, was to streamline the publishing process by allowing formatting stylesheets to be designed to process a base level construct, the Module, from which more semantically informative models were derived as specializations. The publishing framework would organize reusable information objects into the delivery structures needed (Tier 1) and then streamline all markup to align with the Module (Tier 2) against which final rendition processes could be run. This particular strategy was in fact critical to making this type of solution scalable across a wide range of potential implementation scenarios. Past experiences had also demonstrated, graphically, how complex and expensive publishing stylesheets could become if they were expected to handle significant ranges of content sophistication. Using this publishing architecture, DND was able to replace legacy print publishing solutions that depended on as many as 20 FOSI (Format Output Specification Instances) stylesheets with an environment that could support over a dozen discrete publication types with a single FOSI. The quality of the resulting print publications, it should be noted, was in fact improved in conjunction with this consolidation.

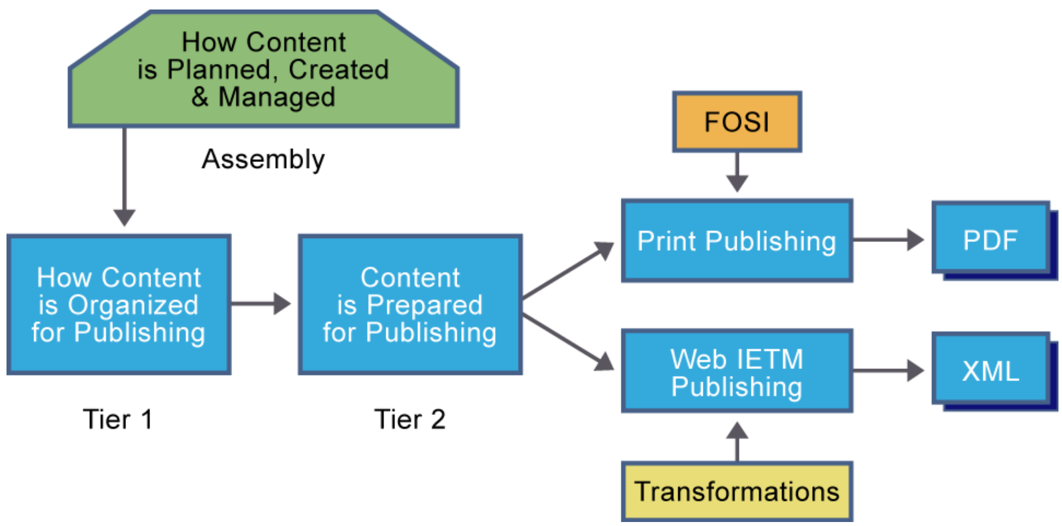

Figure 5: General Deployment Framework
Multi-Tier Processing Reduced Implementation Costs Significantly

In terms of capturing content for different equipment systems and implementing graceful support for a wide range of output information products, the assembly model proven to be remarkably robust. In the almost ten years since the underlying models were finalized, fewer than ten (10) substantive modifications have been made even though, at this same time, dozens of implementations have been successfully completed. When it is recalled that there was no prohibition against change requests, the effectiveness of this approach should be recognized as noteworthy.

Validation Projects. From the period of 1995 through to 1997, validation projects were undertaken as part of the finalization of the overall Assembly Information Architecture. In several cases, specific technology integrations were undertaken to refine the information model and the associated application components within the context of interacting with mainstream information management tools such as Lotus Notes, PC DOCs, Documentum and various Product Data Management (PDM) systems. In other cases, information migrations to SGML were undertaken for different equipment systems so as to determine whether or not different equipment types would give rise to new requirements for the information model. Under this guise, implementations of integrated information management systems were completed within the context of tactical command and control systems, aircraft fleets, naval air defence systems, transport and combat vehicles, protective clothing, computing platforms, surveillance and electronic warfare environments, and software applications.

The most interesting validation project was undertaken on a NATO joint weapon system design and development program. Among the reasons for the Canadian CALS assembly architecture being selected for deployment on this program, which had been formally chartered as a flagship implementation of the CALS vision of concurrent engineering and integrated technical information services, was the fact that it intrinsically supported multilingual content, it could be adapted to support a wide array of interchange scenarios, and it provided an integrated strategy for facilitating both print and online information products. Within the context of this deployment, enhancements had to be introduced into the architecture that would address extremely onerous security requirements.

It should be noted that within these validation projects, the objective in each case was to implement an effective working solution for the participating equipment fleet. The validation and refinement of the assembly information architecture and its associated components became a secondary objective and its fulfilment was enhanced by being subordinated to practical implementation successes.

Shared Applications. Once the validation period was concluded and a final baseline was established for the Assembly Information Architecture in 1997, a period of sustained investment commenced on both large-scale implementations and the evolution of a suite of shared applications needed to support them. Tellingly, the investments in shared applications were not sponsored by the DND CALS Office at all, but rather by a community of equipment programs that had determined that their common interest lay in coordinating their requirements and jointly funding the development of a suite of advanced management and publishing tools. These tools effectively elevated the demonstration capability, associated with the finalized Assembly Information Architecture, to production status.

The publishing tools encompassed the implementation of full support for bilingual loose-leaf printed technical manuals, a number of IETM products, and the ability to publish content to intranet and extranet servers in accordance with the government online standards. As should be expected, the success of the initiative depended very much on the ability of the stakeholders to articulate and integrate their requirements so that the sharing of application components remained a feasible goal.

Beyond the publishing components, a fully functional web-enabled Content Management System (CMS) was designed, developed and deployed that could be run in a federated mode across an equipment supply chain. The resulting CMS could be described as *minimalist* in that it was designed to provide a basic operating capability for smaller suppliers and a staging platform for suppliers exporting content from larger internal content repositories. Again, this environment needed to facilitate a wide range of security considerations while continuing to facilitate efficient working relationships through continuous digital information interchange. This environment incorporated process control aspects that allowed suppliers to manage their publishing processes and to enforce increasingly stringent quality control measures over the content assets being created and exchanged.

Migration to XML. Once XML entered the stage starting in 1996, a series of investments were undertaken so that the possibilities emerging with the appearance of commercially mainstream support for open markup could be leveraged. As part of the effort to enable web-based IETMs starting in 1995, DND was already aware of, and eager to further exploit, opportunities to include intelligence within the published HTML content and the gradual appearance of XML-support in browsers encouraged further investments in browser-enabled IETMs.

Very early on, an XML-compliant version of the assembly information model was established. Interestingly, the sound design of the original SGML schema underlying the assembly architecture, and its careful avoidance of features that would later be deprecated under XML, meant that making an XML adaptation required almost no extra effort. The one exception to this emerged around circumstances when SGML inclusions and exclusions had been used to enable widely-used *overlay markup*, such as is used to explicitly manage changes in the content and the workflow processes associated with them. The prohibition against these SGML mechanisms seen in XML has been one reason why many implementations have continued to leverage SGML as the master representation of their content while exploiting XML as a deployment venue and an integration mechanism with external data sources. This is a circumstance that remains commonplace across the aerospace and defence sector and it has proven to be a very effective compromise.

IMPLEMENTATION RESULTS

2002. The efforts that had commenced with the business case analysis of 1992 and the formation of the DND CALS Office in 1993 had, by 2002, introduced into production a modernized approach to creating, managing and publishing technical documentation for a wide range of DND equipment systems. These implementations had been successful in delivering both substantial savings and tangible operational benefits. Although not performed with the same rigour of the initial analytical efforts, a review was conducted to assess the progress that had been made over the 10 years.

BENEFITS REALIZATION

Implementation Scope. The Assembly Information Architecture and the associated solution components were implemented for a significant number of equipment systems with a priority being placed on larger and more complex fleets where the rate of change was highest. Based on the savings associated with rationalizing the content holdings, streamlining the publishing process and decreasing the reliance on printed manuals, the net result of implementing this modernized approach to creating, managing and publishing technical information was the reduction of the lifecycle ownership cost of the technical information assets by more than 50%. This is a very significant amount.

Information Rationalization. By far the most important consequence of the implementation of the Assembly Information Architecture, from the perspective of benefits realization, was the elimination of redundant information across publication sets and between the different content silos (logistics, configuration management, training, engineering, and documentation) that had historically existed. On average, the elimination of information redundancy, which was sometimes termed "information rationalization" or "content refactoring", reduced the volume of managed information holdings associated with an equipment system by well over 50%.

When it is considered that the content associated with a major equipment system will be measured in terms of several hundred thousand pages of technical information, and on occasion much more, the magnitude of this impact becomes clear. The effect of implementing the Assembly Information Architecture was the reduction, by over half, of the total amount of information routinely being modified, managed, translated, reviewed, and published in support of a typical equipment system.

It is almost equally unbelievable that, prior to this information rationalization, these equipment programs were literally expending funds modifying and translating the same content often dozens of times a year. It was noteworthy that while the equipment management teams knew that there was redundancy within their information collections, in each case, these stakeholders were surprised by how extensive the redundancy was. It was conceded by these stakeholders that apart from the costs of repeating tasks needlessly, there were serious risks associated with information redundancy. If important changes were made and if they had no systematic way of knowing where the modifications would need to be made across the information set, then it was almost certain that the change would only be partially reflected in the documentation set being used by equipment operators.

Streamlined Change Processing. Through the introduction of enhanced automation into the publishing process and the activation of improved information management practices, the Assembly Information Architecture reduced the cost of processing changes significantly. Quite apart from the efficiencies associated with rationalizing the content, it was determined that the time and effort associated with modifying content, translating the changes, and formatting the change packages was significantly streamlined. It is conservatively estimated that this streamlined environment reduced the cost of each *change page* by over $100. Whereas moving to a new work environment is never easy and rarely welcomed enthusiastically, it was found that no one was in the least bit unhappy about abandoning the traditional tasks associated with creating change packages. The joke amongst the solution implementers was that they never expected to hear, and indeed never did hear, anyone say they missed having to modify the List of Effective Pages associated with a changed publication.

Increased Information Product Digitization. The availability of an effective electronic publishing capability that could produce high-functioning web-based Interactive Electronic Technical Manuals (IETMs) meant that programs could aggressively trim the number of hardcopy manuals being maintained in the field. Although minor, this was seen as contributing immediate savings and, over time, as becoming the source of increased savings as more and more documentation would be predominantly used online. Interestingly, there were equipment programs that were so aggressive that they sought to eliminate hardcopy reference materials altogether. This had been discouraged because the challenging deployment scenarios undertaken by the military made hardcopy manuals an excellent option to have available at times. In one vividly remembered case, an emergency publishing effort was undertaken to produce sanctioned printed manuals so that an equipment system could be deployed to Afghanistan on short notice. It turned out that one thing that is compelling about printed materials is that they do not need to be plugged in.

LESSONS LEARNED

Stepping back from all of the details, a participant can become an observer, and from the perspective afforded, a number of lessons become visible:

- Premature investment in complex Content Management Systems can see precious resources being expended implementing functionality that, while eventually necessary, does not deliver visible business improvements and thereby runs the risk of losing crucial stakeholder support.
- Technology initiatives, even when stocked with significant financial resources and the best technology available, frequently collapse under the weight of their own *sophistication*.
- Technical *sophistication* alone can obstruct implementation within any significantly distributed and heterogeneous community of stakeholders.
- Continuous technical innovation, unless carefully managed, can destroy vital implementation momentum and in so doing become self-defeating.
- Among the chief threats to any enterprise initiative will be the allure of apparently magical answers (silver bullets) whether these are new standards emerging into the limelight or fiercely promoted commercial products. Many a project failure can be linked to a combination of management naiveté, that wants to believe in magic, and the harsh fact that magic generally evaporates in the cold light of project implementations.
- Solution architectures must balance portability, adaptability and universality with *practicality*.
- Solution architectures should leverage content processing in such a manner as to ensure that content exists in the form that is optimal for its various lifecycle stages. Content should be authored and managed in the way that is optimized for ensuring its quality and currency. Similarly, the content should be presented to consuming applications a form that is optimized for each circumstance.
- Refactoring content for reuse (information rationalization) is the most important single factor in realizing concrete savings within an information management modernization initiative.
- Refactoring content provides additional benefits associated with the optimization of change processes and improvements in information quality.
- Standards and protocols alone do not constitute solutions.
- The efforts of existing, or historical, standardization initiatives represent an invaluable resource and as much as possible from these precedents should be reused.

- Modernization initiatives that seek to improve the levels of standardization being achieved within an enterprise must be funded sufficiently to establish *genuinely useful* standards.
- Modernization initiatives, to be successful, must align business considerations, political realities, and legacy practices with the benefits of delivering on new opportunities.
- Modernization initiatives should be undertaken with reference to some form of business case that establishes clear and quantifiable objectives, in addition to any qualitative goals.
- Any enterprise modernization initiative, once launched, must be actively supported so that it can adopt a *service model* and thereby effectively engage stakeholders with the purpose of making them successful in conjunction with participating in the enterprise initiative.
- The cornerstone of success for any information management modernization initiative will be the specification and validation of a robust and adaptable information architecture that can provide the *common reference point* for various development and implementation efforts.
- And perhaps most importantly of all, in the design of standards, development of solutions and management of change initiatives, it is critical to remember that *simplicity wins*.

COMPARISONS WITH S1000D AND DITA

The approach taken within this initiative in many ways resembles, indeed prefigures, the general architectural strategies seen in such initiatives as:

- S1000D, the AIA specification for producing and publishing technical documentation;
- Darwin Information Typing Architecture (DITA) for a reuse-oriented approach to topic-based documentation.

On the cautionary side, the experience of the DND CALS Office, both with triumphs and challenges, indicates that some of the architectural features that are seen in DITA, but absent from S1000D, are in fact critical to success. Specifically important are the mechanisms seen in DITA for conceptual abstraction with an information typing architecture, which can be used to manage complexity, and a practical extensibility framework, which can be used to adapt solutions, both gracefully and cost-effectively, in response to the inevitable and inescapable forces of variability and change.

On the positive side, the experience of the DND CALS Office categorically validates the core claims being made by both DITA and S1000D – that a modular approach to creating and managing technical content is feasible and it generates significant benefits. In the more specific case of S1000D, which like the CALS initiative focuses on complex military and aerospace systems, the experience of the DND CALS Office demonstrates that further benefits can be realized when modularized content is managed within the larger context of the equipment system of which it is intrinsically a part.

EPILOGUE

This story does not end on an entirely happy note. For a number of reasons, most of them dubious, there has been a marked decline in the organizational support provided to the standard and shared solution described in this study. As a consequence, many of the benefits associated with enterprise-wide collaboration have come under siege and by this time have probably been reversed. At both the equipment fleet and departmental levels, the absence of strong support for a collaborative approach is leading, once again, to fragmentation and will see a new generation repeat mistakes that have, in this story, been proven to be unnecessary.

