



CarnegieMellon
Software Engineering Institute

Pittsburgh, PA 15213-3890

**SEI Monographs on the Use of
Commercial Software in Government
Systems**

Case Study: Evaluating COTS Products for DoD Information Systems

Carol Sledge
David Carney

June 1998

About this Series

Government policies on the acquisition of software-intensive systems have recently undergone a significant shift in emphasis toward the use of existing commercial products. Some Requests for Proposals (RFPs) now include a mandate concerning the amount of COTS (commercial off-the-shelf) products that must be included. This interest in COTS products is based on a number of factors, not least of which is the spiraling cost of software. Given the current state of shrinking budgets and growing need, it is obvious that appropriate use of commercially available products is one of the remedies that might enable the government to acquire needed capabilities in a cost-effective manner. In systems where the use of existing commercial components is both possible and feasible, it is no longer acceptable for the government to specify, build, and maintain a large array of comparable proprietary products.

However, like any solution to any problem, there are drawbacks and benefits: significant tradeoffs exist when embracing a commercial basis for the government's software systems. Thus, the policies that favor COTS use must be implemented with an understanding of the complex set of impacts that stem from use of commercial products. Those implementing COTS products must also recognize the associated issues—system distribution, interface standards, legacy system reengineering, and so forth—with which a COTS-based approach must be integrated and balanced.

In response to this need, a set of monographs is being prepared that addresses the use of COTS software in government systems. Each monograph will focus on a particular topic, for example: the types of systems that will most benefit from a COTS approach; guidelines about the hard tradeoffs made when incorporating COTS products into systems; recommended processes and procedures for integrating multiple commercial products; upgrade strategies for multiple vendors' systems; recommendations about when not to use a commercial approach. Since these issues have an impact on a broad community in DoD and other government agencies, and range from high-level policy questions to detailed technical questions, we have chosen this modular approach; an individual monograph can be brief and focused, yet still provide sufficient detail to be valuable.

About this Monograph

This monograph reports on a DoD program that undertook a detailed evaluation effort that examined several commercial products as candidates for a large information system. While the evaluation effort focussed on technical questions, the essential issues faced by the program and the lessons that were learned were more in the sphere of management and programmatic issues, and this monograph is written from that perspective.

Case Study: Evaluating COTS Products for DoD Information Systems

1 Introduction

This case study reports on the experience of a Department of Defense (DoD) organization that evaluated several commercial off-the-shelf (COTS) products as candidates for a military information system. This organization is implementing a major DoD program to modernize a number of legacy systems; they are to be integrated into a standard single, field-level system supporting one branch of the armed services. The goal is a system designed to interface with several legacy databases, and one that will collect, process, and distribute transactions both to DoD and to corporate-level systems. Since the program is still underway, we shall refer to it only as “the project.”

This paper reports on activities carried out during the project’s initial phases, and particularly on a succession of COTS product evaluations that were performed. In preparing this report we examined several in-house documents, including system descriptions and product evaluations. We also carried out interviews with management and technical personnel. The experiences described in this study provide ample evidence of the essential variability in COTS evaluation: two separate evaluations of the same set of products produced markedly different assessments, and led to different product choices for the system. As we describe below, these decisions were based only partially on technical issues.

In Section 2, we briefly discuss the background of the project. In Section 3 we describe the evaluations that were carried out. In Section 4, we provide an analysis of these evaluations and in Section 5 we list some of the key lessons learned. Section 6 is a summary of the case study.

2 Background

The system in question is a large and complex information system in the domain of human resources and personnel management. There are numerous specific requirements for the modernization of the system; these include that the new system be multi-layered and distributed, and that it run on Microsoft® Windows NT™ servers. In addition to the set of technical requirements, there are three other major factors that have affected this project: two are external to the program, and one is intrinsic to it.

These factors are

- the current DoD interest in using commercial products
- a stated desire to have, for this class of application, a single, joint system for the entire DoD
- an aggressive schedule for design, evaluation, and deployment of this system

The desire for increased use of commercial products is government-wide. Therefore, it was logical that when developing this system (which is an information system of a fairly common

type) management would make an effort to determine as precisely as possible the potential costs and benefits of a commercial solution, and to pursue a commercial strategy where feasible.

The second of these factors partially results from the DoD's Corporate Information Management (CIM) initiative, which recommended reduction to fewer systems in each functional area. However, in this application area, the ultimate goal for the DoD is to have a common, jointly-used system shared by all of the services. Therefore, attributes such as integrability and interoperability are important requirements for the new system, as is the need for compliance with the DoD standards for information systems development.

The third impacting factor, and the one that is particular to this project, was that it was given a very aggressive schedule. Milestone 0 approval (i.e., to begin concept exploration) was achieved in July 1995; at that time, Milestone 1 approval (i.e., permission to begin the new program) was scheduled to be achieved within eighteen months. This schedule encountered delays, which we describe herein. However, the current schedule is still highly aggressive, with IOC (initial operating capability) now scheduled for November 1998.

3 Product Evaluations

While a commercial solution was obviously a desirable possibility, any candidate products would need careful evaluation before a selection was made. There were, for instance, several significant military-unique aspects of the system that insured that no COTS business application could be used without at least some tailoring and development; the extent and difficulty of such tailoring would therefore be a key question. Also, since several other acquisitions of similar systems were underway, the potential to achieve uniformity (e.g., by selecting the same product) would influence the evaluation process. And finally, the aggressive schedule strongly suggested that a product that was immediately available was preferable to one that would not be available for several months.

A market survey revealed three candidate commercial products (products "X," "Y," and "Z") thought to have appropriate functionality for the proposed new system. The project carried out several evaluations. An initial evaluation was made of all three, analyzing features, business factors, and platforms. This narrowed the list to a single candidate product. This product was then the subject of a more detailed evaluation "fit" analysis to determine its precise capability to perform the requisite DoD functions. A second "fit" evaluation was performed later, and is described below.

3.1 Initial Evaluation

The initial evaluation indicated that neither product Y nor product Z was acceptable. Product Y had a very good user interface, and also made use of a very flexible architecture. However, the new system's requirements specified a multi-layered, distributed environment with a substantial number of sites; this kind of system was incompatible with product Y's design assumptions, which were geared toward a centralized system with few sites. Product Y's vendor also had a licensing strategy that did not easily accommodate the DoD purchasing requirements, with the result that the estimated licensing cost for product Y would apparently be astronomical. In the case of product Z, the product did not run in a Windows NT environment, and the vendor

indicated no intention to produce such a product. Both of these products were therefore removed from consideration.

The initial evaluation of product X—although its interface and architecture were not considered to be as good as product Y’s—showed no immediate problems relating to a distributed system with many sites. Further, there was an immediate attraction toward using product X, since it was already in place as part of a bundled set of products previously purchased by the organization. In addition, in another closely-related acquisition effort (one being run by a different branch of the services), product X had already been chosen for use in a very similar information system.

Individuals from the other acquisition were contacted to obtain information about their system requirements and about their methods of evaluation and analysis of product X. Unfortunately, this information was not immediately available, so personnel from the project undertook their own detailed evaluation of the product.

3.2 The “Fit” Evaluation

This detailed evaluation was performed by making a “fit” analysis of the capabilities of product X with respect to system requirements. The analysis was done in terms of “core fit” and “custom fit.” The former involved estimating the degree to which product X would immediately satisfy the system’s functional requirements; the latter was done by estimating the level of effort required to tailor product X to implement the new system. Additionally, the evaluation assessed the compliance of product X with DoD standards.

The goal of the evaluation was to determine the gap that existed between the product and the specified DoD system, and to expose any other issues that might indicate whether product X should be used to implement the new system. The program’s schedule constraints dictated that this fit analysis could only be done over a portion of the product’s functionality. The program’s technical developers therefore nominated a representative sample of the most complex system functions to assess against product X’s capabilities. This set embodied over 50% of the system’s data requirements and was indicative of the entire system’s functionality. Each function was weighted according to the expected degree of difficulty to implement. Then, a summary of each function’s fit produced a set of overall fitness statistics.

It is important to note that the product’s vendor was invited to participate in both the definition and execution of the evaluation. This participation was deemed valuable both because the project did not have expertise in hands-on use of the product, and because there was a long-range intention to engage consultants (during the development of the new system) from the vendor of whichever product was selected.

The evaluation was constructed as a pilot project. A skeletal system was defined that used the subset of functions and that ran on a subset of data from the legacy system. Functional personnel familiar with the legacy system then provided detailed expertise on business processes to be used and the rules that would apply. These personnel also provided insight into the extent of customization of product X that might be required. This pilot program culminated with a week-long review of the skeletal system by the functional representatives, program personnel, and consultants from the product’s vendor.

3.3 Outcome of the Evaluation

The fit analysis revealed a fairly low level of fit between product X's core capabilities and the new system's requirements. Even with the use of some advanced data mapping features of the product, the fit of core functionality was less than 50%. The analysis of the "custom" fit was inconclusive: It was also not known precisely how much tailoring and customization would be needed to be done to product X, nor the extent to which this would have to be redone for future upgrades of product X and its associated products. (Some amount of parameterization would certainly be needed, but the extent of customized data fields or screen revision was not precisely known.) While inconclusive, the analysis suggested that using product X would result in a productivity gain of 9% over building a custom system in-house. This was determined by considering the "out-of-box" capability, estimating the extent of customization needed, and comparing these against probable development costs.

The pilot also identified a number of other concerns and issues, technical and programmatic. First, there were four main technical limitations:

- The Windows NT version of product X was not yet available.
- The product was incompatible with the current version of an additional product needed by the system (and supplied by the same vendor).
- The product required a higher screen resolution than that available at several field sites, necessitating the upgrade of those older monitors (a short term fiscal impact).
- The product's graphical user interface (GUI) did not provide a consistent interface to the user.

There were also several issues concerning standardization and programmatic matters:

- The product was not certified as compliant with the Defense Information Infrastructure Common Operating Environment (DII/COE). (Some of the vendor's other products were so certified, but not product X.)
- Using product X would require the program (and, subsequently, other DoD organizations) to adapt their business rules and processes to fit the product. The alternative was to extend the product to accommodate whichever of these rules and processes were deemed essential in their current form.
- There were few people in the organization with expertise in the setup and use of the product.

Finally, there was concern about the future direction of the product; this included technical strategy, but also such factors as plans to gain DII/COE certification and ongoing conformance with DoD standards (e.g., the JTA).

The evaluation report was delivered to management within ten days of completion of the pilot. In spite of the concerns and issues described above, the report indicated that if the vendor could provide satisfactory responses to these issues and concerns, product X could be tailored to meet the new system's requirement to an acceptable degree. As a result, product X was selected. This decision was based on the perception that whether or not product X was functionally an ideal

solution, it was acceptable and, more important, immediately available. The implication of this decision is that the driver for this organization was not a perfect functional ‘fit,’ but rather schedule. Another implication is that by using this product, the organization was in conformance with the DoD’s stated desire to use COTS. In addition, the choice was perceived to be highly cost-effective, since the product was virtually free (as noted previously, this product was already in place in the organization), and would be harmonious with another acquisition that was committed to product X.

3.4 Events After the Decision

A major constraint on this decision was that the vendor should provide answers to address the above concerns and issues. For several of these issues, therefore, product X’s vendor was asked for further information and clarifications. In particular, the program requested assurance that the necessary expertise and resources would be available and committed to the project (which included immediately sending product experts to assist with the customization effort). The vendor was also requested to provide a statement of direction for product X that included all planned enhancements and planned release dates.

The vendor made no formal response to these requests. During the four months between the fit analysis report and the targeted Milestone 1 approval date—and despite numerous requests from program personnel—the vendor of product X failed to provide any requested official statements of intent for the product, its conformance to DoD standards, its planned enhancements, and other information deemed critical to the successful use of the product.

In addition, a severe technical dilemma had arisen: despite the presence of some on-site consultants from the vendor, the product could not be customized beyond its installed capability. There was a widespread perception that the vendor was sending poorly qualified consultants to assist in the technical work. For instance, of the consultants sent by the vendor to work with the program, it seemed that only one of the vendor’s consultants could even make the product run. The government personnel became increasingly frustrated, since it was clear that the vendor either did not have the resources, or was unwilling to commit those resources to support the project.

The program therefore ceased the attempt to customize product X. An outside contractor, one not associated with the project so far, was given a 90-day task order to assess the situation, perform further product evaluations, and make recommendations for producing a working system in the remaining time frame.

3.5 The Second “Fit” Evaluation

The new contractor chose to do a closed-door evaluation: most of the personnel who had worked on the program previously were excluded from the new evaluation. The rationale was that a second opinion was needed from a group that did not have any vested interest in the product already chosen.

Although almost two years had passed since the initial market survey, the same three products (“X,” “Y,” and “Z”) were still the primary candidates. In a preliminary screening, product X was

eliminated because of the problems of the previous several months; it was perceived that this decision needed no further justification.

In the case of product Z, its vendor had changed the previous position, and product Z would now be available on Windows NT. However, this version had not yet been shipped. In addition, at the time of the reevaluation, the product's user base was primarily European, with little US market share. For these reasons, product Z was again eliminated in the initial evaluation.

In the case of product Y, one of the major factors that had earlier ruled against the product had now changed: the vendor had redesigned its licensing and pricing agreements to be more in line with industry standard pricing in this application area, significantly lowering its cost. Product Y therefore appeared as the most likely candidate, and a more detailed fit evaluation was made of the product.

Together with product Y's vendor, the contractor carried out a 60-day prototype (i.e., another pilot evaluation) that analyzed product Y with respect to the requirements of the new system. This analysis used the previous fit analysis as a model, though the extent of the functional fit analyses was considerably smaller than in the first evaluation. The fit analysis report (written by the vendor of product Y) showed a higher fit than the first fit analysis had shown with product X. However, even using product Y's tailoring and customization facilities (which were extensive), the fit only approached 50% of the desired capabilities.

As had been true in the previous evaluation, the low fit was not a decisive factor:¹ the contractor recommended choosing product Y as the basis for the new system. Schedule was still the driving factor in making this choice. Some of the accompanying rationale for this choice included database independence, seamless integration of the associated development environment, perceived ease of tailoring and customization, and flexibility of the accompanying tool set. Product Y operated with both of the relational databases currently fielded by the program (whereas product X only worked with one of them). Product Y also had a substantial installed user base and a positive reputation within the community.

Implementation of the new system is proceeding using product Y. The total delay caused by the reevaluation was approximately four months, and Milestone 1 approval was achieved in May 1997. Initial operational capability for the new system is now scheduled for November 1998, still eighteen months after Milestone 1.

4 Analysis

In this section we provide a brief analysis of these evaluation experiences. We focus on two major items:

- the evaluations themselves: their accuracy, their benefits, and how they compared to each other
- the characteristics of the program itself: its schedule, the use of COTS, and differing programmatic expectations

¹ The question of "fit" being a decisive factor will be discussed in Section 4.

4.1 Analysis of the Evaluations

4.1.1 Accuracy and Completeness

Given the compressed schedule available to perform the fit evaluations, neither evaluation was complete, since neither involved a total analysis. Further, these fit analyses centered primarily on functional fit, data mapping, data display (e.g., production of screens), and product conformance to selected DoD infrastructure standards. Omitted from both evaluations were such factors as performance modeling or stress testing; the only indications about acceptable performance have been the vendors' assurances that the products would perform adequately. (This is a potential area of concern, since some doubts were expressed during our interviews about the expected performance characteristics of the new system.) However, while necessarily incomplete, both evaluations appear to have been unbiased. The presence of the products' vendors did not apparently produce any undue influence toward the individual products.

4.1.2 Benefits of the Pilots

Use of pilots as an evaluation mechanism was found to be extremely helpful. Construction of the pilot systems allowed the implementors to test their understanding of the features and functionality of the COTS products. The pilot programs also provided insight into the ease or degree of difficulty of installation and the various tailoring, customization, and enhancement capabilities. The pilots also allowed the functional users to contribute their expertise and knowledge of the legacy system and their expectations about the new system's functionality. The functional users became familiar with the look and feel of the proposed system. They subsequently worked with the system implementers to modify some of the less stringent system requirements, and to reduce some nonessential tailoring and customization of product Y.

4.1.3 Comparison of the Evaluations

The two detailed product evaluations were very similar, at least externally. Both were "fit" evaluations, both were implemented as pilot prototypes of the proposed new system, both were necessarily incomplete, and both were done in a relatively short period of time. The outcomes of the two evaluations also had certain similarities: neither product had a high level of fit using the core capabilities (i.e., the product as it is installed straight "out of the box"), and neither had a large degree of custom fit using their tailoring and customizing features (e.g., even with tailoring, the maximum fit for product Y only approached 50%).

However, the two evaluations were not truly equivalent. The first analysis was considerably more comprehensive than the second. While the second analysis had a goal to replicate the first analysis, it did not provide as much information nor as great a degree of functional detail as the first analysis. Another difference between the evaluations is that while both evaluation teams included representatives of the vendor of the product being evaluated, the second evaluation team, since it excluded anyone who had participated in the first evaluation, had a lesser degree of participation from personnel with intimate functional knowledge of the existing systems and of the requirements for the new system.

4.2 Characteristics of the Program

4.2.1 Schedule Constraints

Schedule constraints provided the major driver for decisions about the new system. There were multiple factors that caused this. One was an operational requirement to solve the “Year 2000” problem by replacing certain legacy systems; this in turn dictated the need for an Initial Operational Capability (IOC) prior to 2000. Maintenance costs of the legacy systems also contributed to the schedule pressure: the sooner the new system was operational, the sooner these legacy systems could be retired and their maintenance costs diminished. Finally, there was a general belief by program management that the productivity benefits of the new system will compensate for the significant downsizing of personnel that is currently underway throughout the organization.

In the aggregate, all of these constraints produced a project schedule that was highly aggressive, notwithstanding some other programmatic constraints (e.g., unfamiliarity with using COTS business applications) that are in conflict with such a schedule.

4.2.2 Use of COTS

The stated directions to the project concerning the use of COTS were that commercial applications were to be used *whenever feasible*. However, many individuals in the program perceive that this was actually a mandate to use COTS *in all circumstances*. One factor that led to this perception is that although some technical personnel had produced an estimate to build the new system that was significantly lower than the overall cost of using product X, management still favored the commercial solution. And when the failure by its vendor led to the rejection of product X, there appears to have been no genuine consideration of whether building the system was an option; instead, a new evaluation led to the choice of product Y, a product that had earlier been rejected. (This perception overlooks the fact that a key element in the rejection of product Y, its pricing and licensing structure, had been extensively revised in the interim.)

4.2.3 Differing Expectations

Expectations of management, technical personnel, and functional users were very different throughout this program. For instance, the perception of what constituted “a COTS solution” differed between technical and management personnel. The technical personnel considered that a COTS solution could consist of their using lower-level commercial products (e.g., operating system, COTS development tools) to develop a custom application. By contrast, management perceived that DoD practice indicates a declining role of organic development and mandates the use of commercial products at the business application level. Thus, for this program, the technical case for a custom application would be very difficult to make, absent some compelling argument in its favor.

A different divergence was found between the expectations of functional users and management. The functional users expected the new system to support their existing business processes; they approached the pilots accordingly. Management, realizing that COTS applications would not have a perfect fit with the existing practices, directed changes to the business processes. After the reevaluation and choice of product Y, for instance, there were a number of business changes that

stemmed from the inherent characteristics of the product. By directing that these changes be implemented, it became possible to reduce the customization needed.

Given the unavoidable downsizing that the DoD is facing, productivity increases and reduced costs are imperative. One of the main approaches to achieving these ends is through the use of COTS rather than in-house development. This approach will be taken even if commercial products do not provide a perfect match for existing business practices. Where misfit occurs, changing business practices is a more viable approach than creating custom applications.

5 Lessons Learned

Whether covered by an explicit mandate or not, commercial products will form the basis of DoD information systems for the foreseeable future.

We have already noted that the goal of using a COTS solution in this system was the principal driver of many of the decisions made in this project. As with many government projects now underway, there is a strong imperative to use a commercial product. This represents a significant change at all levels of an organization, and this change will require understanding by all participants in any project. In this project, for instance, personnel that had previously been systems developers only gradually became aware of the fact that a COTS solution was not merely an option, it was closer to a preordained decision. (While this may be a strong statement, it is certainly true that a COTS solution was a very strong management goal, one that would need a very powerful argument to displace.) The understanding that must become widespread, therefore, is that COTS solutions are now favored: the burden is now to show *why not* to use COTS, rather than showing why to use COTS.

The benefits of a COTS approach will take time to materialize

Using COTS products will seldom bring immediate benefits; simply choosing a COTS solution will rarely have the instant effect of producing systems that are cheaper or better. Part of this is due to the realities of DoD systems: there are some unavoidable and military-unique aspects of systems that are not present in commercial systems that require tailoring of COTS products. (As a simple example, commercial organizations do not need capabilities for mobilization.) In addition, there are regulations and policies that government systems must follow that are not required of commercial systems. But in addition to these, it is also true that the immediate benefit an organization will gain from a COTS approach will, in the short term, be offset by the immediate cost of making a large paradigm shift. Also, there is currently little data that indicates the long-term costs when a product that requires extensive tailoring must be updated (and retailored) as its vendor makes new releases.

There is little doubt that long-term savings will result, and that the use of COTS products is both warranted and necessary. But it is equally probable that there will be occasions, the current program included, where use of COTS will be more costly in the short term than building a comparable system in-house.

Defined processes for using COTS in government systems are needed

The lack of a clear process, or even a simple roadmap, to guide the program was a drawback. The project, with little else to guide it, was essentially forced to invent its overall approach, its evaluation processes, and the kinds of decisions about tradeoffs. It also realized (only in mid-

course) the significance of several unexpected questions about process reengineering posed by the commercial application, and how they would be critical factors in evaluation. This led to delays and inefficiencies: even with a severe schedule constraint in place, it was twenty-four months from the initial market survey until the choice of product Y.

COTS products will change very rapidly

One of the major factors that led to the initial rejection of product Y was its vendor's pricing strategy; by the time the product was reevaluated almost two years later, this strategy had changed dramatically. This experience points out one of the major characteristics of a COTS approach: the time frame during which any product evaluation is valid is very brief. Vendors change both their products and their business practices rapidly, and what is true about a product today will very likely not be true several months from now. Decisions about COTS products must be made with the awareness of this reality and must anticipate that product change will be rapid. (In effect, use of a COTS approach means that the adage about "change is the only constant" becomes very true.)

Use of COTS products will necessitate business process reengineering

A key lesson of this study is that the use of COTS products very likely means that one is purchasing a business process as much as a product. This lesson was learned here very quickly: as the project unfolded, management became more and more aware of the need to change the organization's business processes to align more closely with the COTS products, first of product X, and then more drastically with product Y. One factor was the growing understanding that to reduce the tailoring and customization of the product (or rather, to reduce the time needed to put the product into service), the organization would need to change to fit the product. It was evident in this program that management understood the potential for resistance to such change and that resistance must be overcome. Resistance to change and overcoming that resistance will have associated costs.

Pilot programs are a useful mechanism for COTS evaluations

The pilots were extremely beneficial for this project. Because there is a large potential for misinterpretation and misunderstanding when presenting or discussing a commercial product, hands-on evaluation of promising commercial products is mandatory; pilot programs are a useful way to do this. The robustness of the commercial product, critical aspects of the system, and the tailoring and customization capabilities of the commercial product were all evaluated in the context of a pilot. Participation by the end users, in actual system conditions, allowed those users to help make tradeoff decisions. (It also had a reverse benefit, since this participation also set the end users' expectations for the delivered system.) Results of the pilots indicated potential trouble spots, verified (and disproved) vendor claims and promises, and allowed a better understanding of the resources necessary to complete the system.

(We note that this case study also shows that pilots are, in and of themselves, insufficient indicators. In this instance, the pilot use of product X led to the decision to choose product X. It was only after the fact that the vendor's inaction and unresponsiveness brought about the decision to reevaluate, which subsequently led to the new decision in favor of product Y.)

6 Conclusion

This case study represents a work in progress: the program has achieved Milestone 2, and the initial fielding of the system is expected before the end of 1998. The experiences of this project provide useful insight into the actualities that will be encountered by other government and DoD organizations as they move toward using COTS products as the norm. This project demonstrates the importance of careful evaluation practices, together with the need for understanding all of the implications of a commercial approach. To the extent that the issues encountered by this project will be repeated in other comparable projects, the lessons learned here may be valuable guidelines.

Feedback

Comments or suggestions about these monographs are welcome. We want this series to be responsive to the real needs of government personnel. To that end, comments concerning inclusion of other topics, the focus of the papers, or any other issues are of great value in continuing this series of monographs. Comments should be sent to:

Editor
SEI Monographs on COTS
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213
cots@sei.cmu.edu