Add Decision Analysis to Your COTS Selection Process

Barbara Cavanaugh Phillips and Susan M. Polen
Software Productivity Consortium

Processes for evaluating, comparing, and selecting commercial off-the-shelf (COTS) products are often nonexistent, or when performed, are vague, poorly documented, nonrepeatable, and inconsistent. At best, these deficiencies can decrease confidence in the selection decision; at worst, this lack of forethought leads to poor decisions that delay a project's development, increase life-cycle costs, and reduce quality. It is critical to a project's success that the most appropriate COTS product is selected. This article describes an answer to this common problem and a decision-making process specifically for COTS evaluation, and it provides some lessons learned from the application of this process.

How can you select commercial off-the-shelf (COTS) software from a market of more than 50 available options? Which option has the best value for your project's features? Should you believe everything the vendor says? How can you account for that when making decisions? How can you convince your company that you have little confidence in decisions made quickly based on a few data points? When is it valuable to follow a structured process? What process and method should you follow? How can you save money and time in future evaluations?

In response to its members' need for a repeatable and systematic process for evaluating and selecting components and COTS software, the Software Productivity Consortium (the Consortium) developed the Comparative Evaluation Process (CEP). This article describes the Consortium's process so that you may add some ideas to your own process.

It is critical to a project's success that the most appropriate or "right" component be selected. Both the budget and schedule of a project are affected if the right component is not selected at the start. Off-the-shelf component use and integration are often required by clients and are written into contracts. The CEP is an instance of the Decision Analysis and Resolution (DAR) process area of the Capability Maturity Model® Integration℠ (CMMI℠) and a structured decision-making process. The relationship between CMMI's DAR and CEP is discussed at the end of this article. The process provides detailed guidance to alleviate the typical problems that occur during an evaluation. For example, managers often find that evaluations never end and are quite costly to the program. CEP's first activity is to scope the effort and schedule the activities.

℠ CMM Integration and CMMI are service marks of Carnegie Mellon University.

Several features separate CEP from similar processes. One feature is a suggested set of criteria that expands the understanding and evaluation of characteristics beyond commonly evaluated characteristics such as function or cost. Categories of criteria include basic (e.g., maintainability and installation), management (e.g., vendor viability, costs, and required training), architecture (e.g., platform and framework), strategic (e.g., information technology goals and business goals), and functional, which is specific to COTS class and project context.

> *"Selections are often second-guessed. By following a systematic process for evaluating and selecting COTS products, such as CEP [Comparative Evaluation Process], you adequately capture and document the information necessary to defend the selection."*

Another feature is the contextual focus that is explicitly part of CEP. The context focus enables effective use of resources to ensure a deliberate and calculated evaluation. This is in contrast to generically exercising an alternative's functionality, which provides little or no insight into how well each alternative aligns with your needs.

Finally is the credibility feature. Confidence in the data gathered during the evaluation is gained by knowing and rating the credibility of the data source. The selection decision includes this credibility factor on how well the evaluator knows the data values.

Comparative Evaluation Process Activities

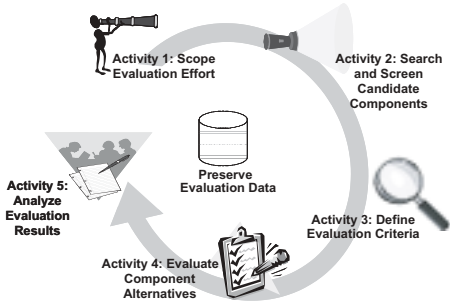
A systematic and repeatable process for evaluating and selecting COTS products provides the rationale necessary to support selection decisions made (e.g., pool of candidates, search criteria, minimum screening thresholds, alternatives to evaluate in depth, detailed evaluation criteria, and analysis). Selections are often second-guessed. By following a systematic process for evaluating and selecting COTS products, such as CEP, you adequately capture and document the information necessary to defend the selection.

CEP is made up of five top-level activities, which are explained below and depicted in Figure 1. Each activity has three to five sub-activities, which are explained in detail in the technical report [2].

Activity 1: Scope Evaluation Effort

This activity sets the expectations for the level of effort and schedule for the remaining activities within CEP. It provides the expected number of COTS products to search, screen, and evaluate. Feedback from future activities often requires redefining the scope for one or more of the activities. Feedback may

Figure 1: *CEP Diagram*



Credibility	Value	Description
Verified	10	Verified in-house using hands-on experience.
Demonstrated	7	Witnessed in a focused demonstration.
Observed	5	Seen but not studied.
Heard / Read About	3	Described by a user or vendor or seen in vendor or third-party literature.

Table 1: *Ordered List of Values Assigned to Credibility*

indicate that there were too many or too few possible candidate components located during the search or that the addition of criteria should change the scope. This activity allows you to plan resources while minimizing and identifying potential overruns.

Activity 2: Search and Screen Candidate Components

The search for candidates first requires that the initial search criteria and thresholds (the "must haves") be defined. The search criteria typically are based on required functionality and key constraints. Keep the criteria broad so that the search is not limited by too many constraints. Using the search criteria, perform a search for possible candidates from sources both internal and external to the project or organization. After locating the candidates, screen them by applying qualified minimum thresholds to the search criteria for each candidate. This allows the most promising candidates to be evaluated fully during Activity 4. Candidate screening is fundamental and cost effective because projects rarely have sufficient resources, budget, and schedule to evaluate every possible candidate.

Activity 3: Define Evaluation Criteria

This activity produces the detailed criteria necessary to support a repeatable and systematic evaluation. The definition of criteria refines, formalizes, and expands on the search criteria and addresses functional, architectural, management, strategic, performance, and financial characteristics of the candidates. Weights are established for all of the evaluation criteria with respect to each project's importance. The selection is based on criteria priority.

Activity 4: Evaluate Component Alternatives

The Evaluate Component Alternatives activity is conducted to assess how well the alternatives meet the defined criteria. Evaluation scenarios are developed to evaluate the alternatives within your particular context rather than generically exercising the alternative's functionality. Results are documented for analysis. While not all alternatives can or must be evaluated in the same manner, evaluation results are based on the available data. The available data may be from hands-on experience, witnessing vendor demonstrations, observing a user, and reading third-party literature or vendor's literature. Each type of data is given a rating value (Table 1). Credibility – rating the confidence in what the evaluator knows about an alternative – is then incorporated in the simple weighted average.

Activity 5: Analyze Evaluation Results

The evaluation produces data on how well each alternative meets the defined criteria. The analysis consists of activities to compare and contrast rankings of alternatives based on the priorities. Sensitivity analysis, using a decision-support method, is performed to determine the impact of criteria or groupings of criteria on the ranking of alternatives. More confident decisions may be made when the impact of the criteria is analyzed.

Decision Model

We developed an easy-to-use spreadsheet called the Decision Model that you can create yourself in a spreadsheet to hold the decision information (e.g., criteria, alternatives, priorities, ratings, and data charts). The Decision Model aids in decision making when comparing similar products using discriminating criteria.

- Software – Microsoft Excel.
- Decision Theory – Simple weighted average.
- Rows – Criteria.
- Columns – Alternatives.
- Cells – Criteria ratings for each alternative.

The following describes the Decision Model's basic features.

Decision Theory Model

The decision theory model behind the Decision Model is simple weighted averages. Simple weighted-average theory applies a weight to each criterion. The global weight is determined by multiplying weight, in percentages, by the weights of the criteria in the hierarchy. Assume the criteria hierarchy was as indicated in Table 2. Criteria 2.2 Vendor Viability has a local weight of 75 percent and is a subcriterion of 2.0 Management, which also has a local weight of 75 percent. To determine its global weights multiply 75 percent by 75 percent to equal 56.25 percent.

Weighting

Weights are applied to the evaluation criteria so that decisions can be made based on the results of the component evaluations. The weights are subjective and dependent on the particular project emphases. The decision-maker must provide a set of weights that are believed to be appropriate for the situation at hand. For the Decision Model, the weights in a level of the hierarchy must add up to 100 percent for normalization purposes. Additional averaging techniques such as dividing 100 points among the criteria or assigning them high, medium, and low values may be used and converted to a normalized scale.

Credibility

The purpose of credibility value scoring, as discussed above, is to include how well the evaluator knows the criteria value in the scoring equation. Often vendor-supplied information is not considered as valid as that verified through hands-on experience. The assignment of credibility values should reflect a level of confidence of the information contained in the criteria ratings value. To achieve this, Table 1 shows an example of an ordered list with the greatest confidence at the top of the list. This is reflected in the value assigned to each credibility scale item. The values are based on the experience of the evaluator. It is an attempt to quantify what is essentially qualitative.

Table 2: *Global Weight Calculation Example*

First Hierarchy Level Criteria (local weight)	Second Hierarchy Level Criteria (local weight)	Global Weight
1. Basic (25%)		
	1.1 Usability (50%)	12.5%
	1.2 Maintainability (50%)	12.5%
2. Management (75%)		
	2.1 Suggested Training (25%)	18.75%
	2.2 Vendor Viability (75%)	56.25%
TOTAL		**100%**

Calculating the Result

Using simple weighted average, the Decision Model calculates the results based on the criteria value and credibility ratings. Table 3 shows an example of the values entered into the simple weighted average.

The scoring uses a 10-point scale to normalize the data. For simplicity, the average is divided by 100 putting the result on a 100-point scale. Words used in the rating scale are converted to numbers using the Microsoft Excel function VLOOKUP. The set of values are named (Insert, Name, and Define) and referenced in the formula. In the example below, Excellent_Good_Fair_Value and Credibility_Value are defined names. The formula below would replace cell E3 in Table 3 if the named values are used and need to be converted to numbers. Alternatively, the numbers could be used directly.

E3=B3*VLOOKUP(C3,Excellent_Good_Fair_Value,2,FALSE)*VLOOKUP(D3,Credibility_Value,2,FALSE)/100

How to Interpret Results

A special alternative in the Decision Model is the one named the Perfect. Its criteria rating and credibility ratings are set at the maximum values. For the bar chart showing the cumulative scores for each alternative, the Perfect is set at 100 percent. When the criteria are grouped, the Perfect allows comparison between the highest possible score and the alternative's score. For example, Figure 2 shows that Alt C has the highest ranking for the Functional Criteria Category at 18 percent and the Perfect score for the category is 35 percent. Clearly, none of the alternatives performed very well in this category. Strategic criteria was not pertinent to the evaluation and the category was dropped. The evaluator now knows the selected COTS product is not going to have all the desired functionality and may consider refining the criteria, negotiating the requirements, or finding another source to provide the missing functionality.

Sensitivity analysis is a method for determining confidence in the results. This enables decision making based on the impact that the criteria have on the selection of the COTS product. The sensitivity analysis may include operations such as reviewing the weights of the evaluation criteria, making adjustments to the weights, and observing the effect on the results. This activity may be performed multiple times depending upon what is observed or uncovered while doing the sensitivity analysis.

	A	B	C	D	E
1			ALTERNATIVE A		
2	Criteria	Global Weight	Criteria Rating	Credibility Rating	Weighted Average
3	2.2 Vendor Viability	56.25%	Excellent (convert to 10)	Verified (convert to 10)	(B3*C3*D3)/100

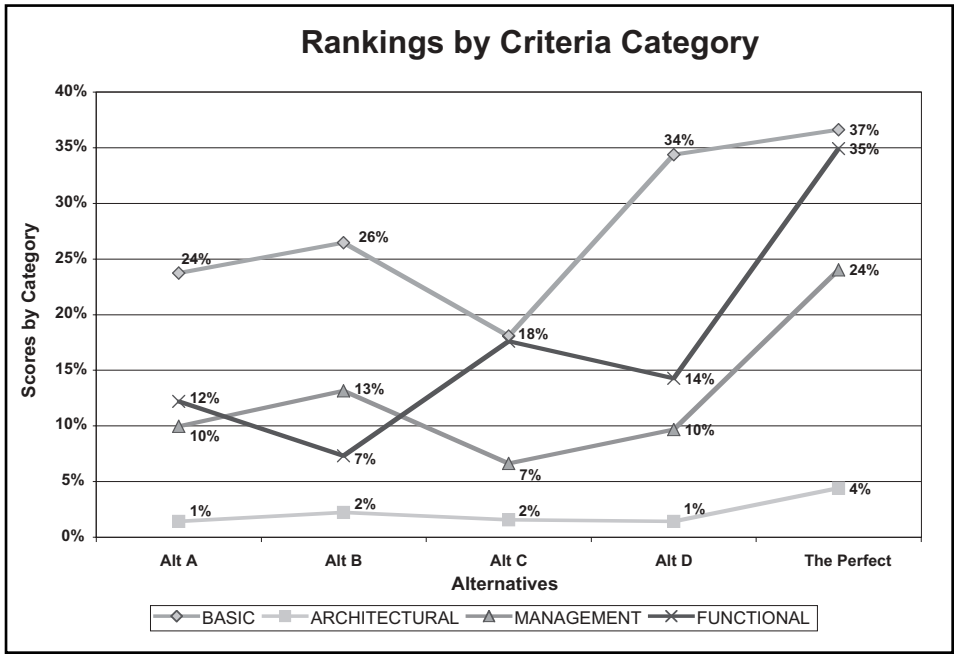
Table 3: *Calculations in the Decision Model*

Lessons Learned

Many lessons may be learned while applying CEP; all are of equal importance:

- Early and Effective Vendor Contact. Making contact and getting results to inquiries from vendors is a long and laborious effort. Do not underestimate effort and schedule for this activity. Staying on top of the communication flow helps prevent schedule slips. Smart vendors see the benefit of participating in an evaluation. The ones who have been most responsive are those who provided an explanation of how well their product performed against the criteria. We have made it a policy not to give vendors a copy of the evaluation report. It subjects us to too many unnecessary questions. Our process is focused on finding a product to fit the specific context of the evaluation, not a best in class, and vendors have a hard time understanding this.
- Look to Training Requirements for Information. Training requirements could be an indicator of the size or scope of installation and actual hands-on evaluation time. Typically, this ranges from none to one week. We attempted to install the products for one evaluation and finding it very difficult, discovered it required a week of administrator training.
- User Observation. Interviewing or observing a user may be a more practical and beneficial method of data collection over witnessing a vendor demonstration. The credibility factor can account for the difference in the source of ratings.
- Use of Evaluation Scenario and Data. To get the most out of a vendor demonstration, request that vendors provide a focused demonstration with materials from your evaluation scenario and data. Vendors typically have a set of features they want to show you, but they may not be the features in which you are interested.
- Subject Matter Experts. By obtaining the services of a subject matter expert to assist with the class of COTS products under evaluation, you can more efficiently identify possible candidates, define criteria, and develop an evaluation scenario and data.
- Estimation Data. Data used to estimate and scope the effort along with actual tracking data of the evaluation should be retained within the repository. It will provide historical data to be

Figure 2: *Ranking by Criteria Category Example*



used for estimating future evaluations.

- Demo Forum. Allowing stakeholders to learn and witness a demonstration of the final alternatives proves an excellent means for getting input and buy-in. More forums could be held to provide hands-on experience. A forum also may be appropriate for criteria identification.
- Advocate. Assigning an evaluation team member to be responsible for installing and learning a final alternative is an effective use of limited resources. The advocate becomes an expert in an assigned alternative and may rate the criteria with confidence.
- Software Installation. For those tools to be evaluated hands on, consider borrowing a vendor-owned laptop with the software already installed if installation problems become overwhelming. Capture this problem in the criteria ratings.
- Team Size. The ideal size for the evaluation team is between three and five. A smaller team may allow bias while bigger teams make communication complex and scheduling difficult.

Additional lessons may be learned while applying CEP to make confident COTS selection. Below are answers to the questions that began this article.

- Use a systematic and repeatable process such as CEP, which can be tailored and refined with each use to maximize its benefit.
- To ensure the best value for your desired features, translate your features into measurable criteria, assign priority to your criteria, rate your alternatives according to the criteria, and let simple weighted averages (or other decision-support method) provide the answer.
- Vendors are driven by current or potential profits. They can be cooperative and responsive when it is in their perceived interest to be so [3]. Never confuse selling with installing. Salespeople speak of the product's strengths but not the weaknesses. Factor in your data source (e.g., hands on, vendor demonstration, and vendor literature) when scoring the alternative criteria.
- A systematic and repeatable process for COTS evaluation and selection provides the rationale necessary to support decisions. The basis for the decision is available for review, which increases the confidence in the results.
- From a project management perspective, if the decision is important to the overall success of the project, then it should be given adequate resources. Those resources should be used efficiently and effectively, as is the case with CEP.
- Save all documentation (e.g., pool of candidates, search criteria, minimum acceptable threshold values, detailed evaluation criteria, alternatives to evaluate in depth, and analysis). The evaluation may need to be reviewed because of a new entry in the market or a new version of an existing tool. Maintain the evaluation data in a repository. It is helpful to see the artifacts from a completed evaluation when starting a new one to get ideas as candidate sources and criteria.

Table 4: *Comparison of the CEP and the CMMI DAR Process Area*

CEP	DAR Specific Practices (SP) and Pertinent Generic Practices (GP)
CEP	The purpose of Decision Analysis and Resolution is to analyze possible decisions using a formal evaluation process that evaluates identified alternatives against established criteria. GP 2.3 Provide adequate resources for performing the decision analysis and resolution process, developing the work products, and providing the services of the process. GP 3.1 Establish and maintain the description of a defined decision analysis and resolution process. GP 3.2 Collect work products, measures, measurement results, and improvement information derived from planning and performing the decision analysis and resolution process to support the future use and improvement of the organization's processes and process assets.
Activity 1: Scope Evaluation Effort	SP 1.4 Select the evaluation methods. GP 2.2 Establish and maintain the plan for performing the decision analysis and resolution process. GP 2.7 Identify and involve the relevant stakeholders of the decision analysis and resolution process as planned. GP 2.8 Monitor and control the decision analysis and resolution process against the plan for performing the process and take appropriate corrective action.
Activity 2: Search and Screen Candidate Component	SP 1.2 Establish and maintain the criteria for evaluating alternatives, and the relative ranking of these criteria. SP 1.3 Identify alternative solutions to address issues. SP 1.5 Evaluate alternative solutions using criteria and methods.
Activity 3: Define Evaluation Criteria	SP 1.2 Establish and maintain the criteria for evaluating alternatives, and the relative ranking of these criteria.
Activity 4: Evaluate Component Alternatives	SP 1.5 Evaluate alternative solutions using criteria and methods.
Activity 5: Analyze Evaluation Result	SP 1.6 Select solutions from the alternatives based on the evaluation criteria.

Comparative Evaluation Process and CMMI

Table 4 compares CEP with the DAR process area of the CMMI to show their relationship. The purpose of DAR is to make decisions using a structured approach that evaluates identified alternatives against established criteria.

Conclusion

In collaboration with our membership, the process has been successfully applied to select the following:

- Change Management Tools.
- Decision Analysis Tools.
- Knowledge Management Portals.
- Process Modeling and Simulation Applications.
- Voice Recognition Software.

In summary, a systematic approach to COTS evaluation was developed to help avoid common pitfalls associated with evaluations and trade studies. This approach assists evaluators with component selection. It is generally applicable to components and particularly to COTS software. It adapts decision-support methods to assist with implementation. The approach stresses the creation and maintenance of a repository for capturing evaluation data and lessons learned for future use. The Consortium has collaborated successfully with many members to select COTS products using CEP and is rapidly building a repository of completed evaluations.◆

References

1. Capability Maturity Model Integration (CMMI℠) for Systems Engineering/Software Engineering Integrated Product and Process Development. Ver. 1.1. Pittsburgh, Penn.: Software Engineering Institute, Dec. 2001. 530. Continuous Representation. Decision Analysis and Resolution (DAR).
2. Polen, Susan M., Louis C. Rose, and Barbara C. Phillips. Component Evaluation Process (SPC-98091-CMC, Version 01.00.02). Herndon, Va.: Software Productivity Consortium, 1999 <www.software.org/pub/darpa/darpa.html>.
3. Lessons Learned in Developing Commercial Off-the-Shelf (COTS) Intensive Software Systems. Federal Aviation Administration. Washington, D.C.: Software Engineering Resource Center, 2000.

About the Authors

Barbara Cavanaugh Phillips, certified Project Management Professional, is a senior member of the technical staff at the Software Productivity Consortium. Phillips is co-author of the Consortium's Comparative Evaluation Process (CEP) and has worked with the Consortium's membership to apply CEP. She has a bachelor's degree in American studies from George Washington University and a master's degree in information systems from the George Mason University. Phillips is a member of the Institute of Electrical and Electronics Engineers.

Software Productivity Consortium
SPC Building
2214 Rock Hill Road
Herndon, VA 20170
Phone: (703) 742-7309
Fax: (703) 742-7200
E-mail: phillips@software.org

Susan M. Polen is a senior member of the technical staff at the Software Productivity Consortium. Polen is co-author of the Consortium's Comparative Evaluation Process (CEP) and its supporting training courses, Web site, and Repository of Evaluations and has worked with a number of Consortium members to apply CEP. She has 13 years of software and system development experience including Motorola and Allied Signal. Polen has a bachelor's degree in computer science from Mary Washington College.

Software Productivity Consortium
SPC Building
2214 Rock Hill Road
Herndon, VA 20170
Phone: (703) 742-7178
Fax: (703) 742-7200
E-mail: polen@software.org

WEB SITES

Software Technology Support Center

www.stsc.hill.af.mil

The Software Technology Support Center (STSC) is an Air Force organization established to help other U.S. government organizations identify, evaluate, and adopt technologies to improve the quality and efficiency of their software products and their ability to predict delivery cost and schedule. The STSC Web site now provides mappings of the Capability Maturity Model for Software Version 1.1 to and from the Capability Maturity Model Integration℠ for Systems Engineering/Software Engineering/Integrated Product & Process Development Version 1.1.

Risk Management

www.acq.osd.mil/io/se/risk_management/index.htm

This is the Department of Defense (DoD) risk management Web site. The Systems Engineering group within the Interoperability organization formed a working group of representatives from the services and other DoD agencies involved in systems acquisition to assist in the evaluation of the DoD's approach to risk management. The group will continue to provide a forum that provides program managers with the latest tools and advice on managing risk.

INCOSE

www.incose.org

The International Council on Systems Engineering (INCOSE) was formed to develop, nurture, and enhance the interdisciplinary approach and means to enable the realization of successful systems. INCOSE works with industry, academia, and government in these ways:

- Provides a focal point for disseminating systems engineering knowledge.
- Promotes collaboration in systems engineering education and research.
- Assures the establishment of professional standards for integrity and in the practice of systems engineering.
- Encourages governmental and industrial support for research and educational programs to improve the systems engineering process and its practices.

Center for Software Engineering

http://sunset.usc.edu/index.html

Dr. Barry W. Boehm founded the Center for Software Engineering (CSE) in 1993. It provides an environment for research and teaching large-scale software design and development processes, generic and domain-specific software architectures, software engineering tools and environments, cooperative system design, and the economics of software engineering. One of CSE's main goals is to research and develop software technologies that can help reduce cost, customize designs, and improve design quality by doing concurrent software and systems engineering. It also aims for research topics that will facilitate the training and education of skilled software leaders.

The Open Group's Architectural Framework

www.opengroup.org

The Open Group is a vendor-neutral, international, member-driven standards organization. It focuses on the development of software standards that enable enterprise integration. The Open Group is a global network of information technology customers and vendors who are developing multi-vendor integration solutions through open standards, testing, certification, and branding. Members' benefits include the following:

- Advanced knowledge of technology and standards developments.
- The opportunity to participate in or lead the development of standards.
- Access to information on real-world implementations and proven practices.
- Better procurement practices supported by well-defined brands and standards.

Project Management Institute

www.pmi.org

The Project Management Institute (PMI) claims to be the world's leading not-for-profit project management professional association. PMI provides global leadership in the development of standards for the practice of the project management profession throughout the world.