Techniques for COTS Software Evaluation

John C. Dean, CD, CST
Software Engineering Group
Institute for Information Technology
National Research Council Canada

Outline

- Introduction
- Fundamentals of COTS Product Evaluation
- PECA - An Evaluation Process
- Techniques for Evaluation
- Last Words
Background

- COTS - Commercial Off-The-Shelf

- Tutorial co-developed by NRC and SEI
  - NRC (Software Engineering Group)
  - SEI (COTS-Based Systems Initiative)

- One and a half years in development
  - Only one of the results of a 5-year collaboration

What Is a COTS Product?

- A product that is
  - sold, leased, or licensed to the general public
  - offered by a vendor trying to profit from it
  - supported and evolved by the vendor, who retains the intellectual property rights
  - available in multiple, identical copies
  - used without modification of the internals
Almost-COTS Products

- Freeware/Shareware
  - Vendor “profit” may not be dollars but some sort of brand recognition.
- Open source
  - The “vendor” is the community.
- Restricted distribution products
  - “General public” may have restricted meaning.
  - Ex: In the U.S., products with strong encryption can only be sold to U.S. companies.

What Is a COTS-Based System?

- Any system partially or completely constructed using COTS software products as integral components.
  - not measured in percentages
- Examples include
  - enterprise resource planning (ERP) systems
  - Army C3/C4 systems constructed using a combination of custom and COTS software
A Spectrum of COTS-Based Systems

Introduction

COTS-Solution Systems
One substantial product (suite) used (and sometimes tailored) to provide significant system functionality
- tailoring focus
- vendor maintained

COTS-Aggregate Systems
Multiple products from multiple suppliers integrated to collectively provide system functionality
- integration focus
- project maintained

The Fundamental Change

Traditional Development Approach

System Context
Architecture & Design
Implementation

Required COTS Approach

System Context
Simultaneous Definition and Tradeoffs
Marketplace
Architecture & Design

This change applies to business and contractual activities, as well as to engineering activities.
Outline

- Introduction
- Fundamentals of COTS Product Evaluation
- PECA - An Evaluation Process
- Techniques for Evaluation
- Last Words

What Is COTS Product Evaluation

- The examination of COTS products for the purpose of determining the products’ fitness for use in a particular context.

- Fitness for use means such things as
  - Does it provide adequate functionality?
  - Does it interoperate with other components?
  - Can you adapt it to fit your need?
Simple View of COTS-Based Systems

Fundamentals

- Building a COTS-based system is like completing a puzzle. Filling each hole in the puzzle requires an evaluation.
- The shape and size of the remaining hole changes with each evaluation. The shape and size of the pieces also change with time!

Discovery vs. Matching

COTS product evaluations have different goals depending on the phase of the CBS project.

- Discovery: multiple objectives, flexible, many unknowns
- Matching: well-defined objective, firm technical certainty

© 2001 by Carnegie Mellon University/National Research Council Canada
Outline

- Introduction
- Fundamentals of COTS Product Evaluation
- PECA - An Evaluation Process
- Techniques for Evaluation
- Last Words
There Is More Than Just Process

- The evaluation process is a high-level description of what we need to do.
- Techniques are low-level descriptions of how to do it.

Tailored Process + Techniques

Instantiated Process

Outline

- Introduction
- Fundamentals of COTS Product Evaluation
- PECA - An Evaluation Process
- Techniques for Evaluation
- Last Words
An effective team is critical for a successful evaluation.

Diverse skills are necessary, such as those held by
- developers and technical people
- domain experts
- business people

Good balance of power is important.

Are dedicated teams a good idea?

---

**Picking the Approach:**

**First Fit vs. Best Fit**

**First Fit**

![Diagram of First Fit]

**Best Fit**

![Diagram of Best Fit]
Picking the Approach: Filters

Techniques

- As selection is narrowed the evaluation techniques become more rigorous.

Factors to consider
- Size of candidate field
- Discriminating criteria
- Evaluation budget

Requirements vs. Criteria

Techniques

- Evaluation requirements are stated in terms of needs.
  - “Information transfer shall be protected from unauthorized access.”

- Criteria are stated in terms of capabilities to satisfy those needs.
  - “support for secure sockets or equivalent security mechanism”

- Criteria are quantifiable, evaluation requirements may not be.
Defining Criteria

- Two ways to define criteria
  - reuse preexisting criteria
  - generate new criteria

- We can reuse criteria to cover aspects that are not system specific.
  - product feature checklist
  - organizational checklist

- We must generate criteria to cover system-specific requirements.

Product Feature Checklists

- Feature checklists are standard fare for product comparisons; they are
  - a kind of specification for a class of products
  - often reflective of only common, not unique, features
  - as reliable as the source

<table>
<thead>
<tr>
<th>Internet Server</th>
<th>Netscape Enterprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can require password</td>
<td>✔</td>
</tr>
<tr>
<td>Supports SSL v. 2</td>
<td>✔</td>
</tr>
<tr>
<td>Supports SSL v.3</td>
<td>✔</td>
</tr>
<tr>
<td>Supports S-HTTP</td>
<td>✔</td>
</tr>
<tr>
<td>Supports PCT</td>
<td>✔</td>
</tr>
<tr>
<td>Prohibit by domain name</td>
<td>✔</td>
</tr>
<tr>
<td>Prohibit by IP address</td>
<td>✔</td>
</tr>
</tbody>
</table>

Relevant for your context? Why?
Organizational Checklists

Categories of Criteria (Level 1)
- Technical Requirements
- Compatibility with other COTS
- Adaptability, Flexibility
- Reliability, Maintainability
- System Integration
- System Integrity
- Security
- Vendor Support
- Training
- Documentation
- License Restrictions

Techniques

- Organizations need (want) consistency in IT.
  - represent corporate interests
  - avoid incomplete criteria
- Organizational checklists provide some uniformity and predictability.
  - in coverage of corporate needs
  - in the overall evaluation process

Example of Technical Criterion

Techniques

- Requirement
  - “The system shall calculate longitude.”

- Capability statement
  - “Given input data accurate to five decimal places, the system shall calculate longitude to an accuracy of four decimal places.”

- Quantification method
  - “Provide various values of valid input to the system and confirm that the longitude is accurate to four decimal places.”
Feature driven rather than requirements driven
- Products are reviewed to compile a list of features.

Focuses on risk rather than fitness
- Features are analyzed to determine the risk associated with the presence or absence of that feature in a product.
- The risk statement is the explicit mapping of a product feature to a system need.

Techniques

Step 1: identify interesting product features
Step 2: assert risk to system of a product not exhibiting a feature
Step 3 (optional): categorize and quantify risk
Step 4 (optional): identify mitigations
**Example: Risk-Driven Generation**

**Techniques**

- Product features are the source of criteria.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatic load balancing of servers based on user connects</td>
<td>Response time worse than required by users</td>
</tr>
<tr>
<td>Vendor is foreign-owned</td>
<td>None: “buy America” not required</td>
</tr>
</tbody>
</table>

**Collecting Data**

**Techniques**

- You collect data to provide the basis for your analysis.

- Less obvious but equally important
  - degree of confidence in data
  - architecture and design implications of the product
  - limitations/conditions placed on use of product
  - options for use of the product uncovered during measurement
  - deficiencies in assessment methods and in evaluation requirements or criteria
Data Collection Techniques

- There are several different families of techniques.
  - literature review
  - vendor appraisals
  - hands-on experiments

- You may want to use multiple data collection techniques for a single criterion.

- For example, data collection techniques for “transactions per second” include
  - trust the provided documentation
  - extrapolate from pre-existing benchmarks
  - perform a specific experiment for the operational environment

Literature Review

- Vendor sources
  - user manuals, marketing brochures
  - release notes, version history
  - vendor references

- Internet survey
  - on-line research of reference materials
  - gossip and rumor (beware of unchecked references)

- Consumer report and user groups
  - experiences of other parties with a particular product
Vendor Appraisal

Techniques

• Vendor appraisal is the analysis of the vendor organizations’ processes, personnel, and organizational capabilities as they affect the COTS product.

• Sources include
  – basic vendor business and capability information (e.g., interviews, vendor literature, capability evaluations)
  – independent financial analyses (e.g., Standard & Poor’s)
  – strategic information and lists of users (e.g., trade journals, vendor information)
  – customer kudos and complaints (e.g., web sites, user groups)

Hands-on Techniques

Techniques

• Unless the product choice you are making is trivial, you will need to use hands-on techniques with the COTS product to
  – verify claims
  – determine interactions with other components
  – determine feasibility of proposed architectures and designs
  – determine performance, reliability, etc. in your context
  – identify assumptions made by the product

• Techniques include product probe, prototypes, and scenario-based testing.
Product Probe

Techniques

- An experiment in which specific features of a product are investigated
  - guided experimentation
  - discovery (a.k.a. “playing around”)
- Types of guided experimentation
  - error testing
  - stress testing
  - data stream fault insertion
  - reactions at or near the boundaries

Even “playing around” requires discipline to record the sequence of activities that led to a result.

Prototypes

Techniques

- A prototype is a small-scale version of the system to demonstrate critical features and design decisions.
- Prototypes allow experimentation with the COTS product in the real context.
- Prototypes are useful to observe:
  - interfaces with major components of the system
  - user interaction
  - critical product features (performance, robustness...)
  - areas of high technical risk (integration with legacy system, interoperability with external sources...)

© 2001 by Carnegie Mellon University/National Research Council Canada
Scenario-Based Evaluation

- A *scenario* is a step-by-step description of a function that a system must perform.

Scenarios focus on a specific system pattern of use.
- are most useful for determining “fitness for use”
- help identify products that exceed evaluation requirements

- We can define scenarios based on Use Cases.

Use Cases

- A *use case* is a sequence of interactions between an Actor and the system.

- An Actor is anything that communicates with, but is external to, the system.
  - can play a single role only
  - is not necessarily a user

- A single user may adopt multiple roles.
Steps for Scenario-Based Evaluation

Techniques

- Step 1: Define scenario by isolating a particular (sub)process.
  - may involve multiple Use Cases
- Step 2: Define specific test procedures for scenario.
  - include expected outcomes
- Step 3: Establish an environment.
  - emulates or simulates area of inspection
- Step 4: Insert product(s) into environment.
- Step 5: Run tests and record results.

Other Techniques

Techniques

Benchmark: Experiment to determine some quantifiable characteristic of the product

Product insertion: Experimental (and usually limited) fielding of the product

Demonstration: Vendor guided preview of product capability
Analyze Data

Techniques

- Consolidation *summarizes* data to extract information.

- Analysis *reasons* about that information in order to make a recommendation.

- Reasoning is by nature a highly creative task. However, some techniques may help:
  - sensitivity analysis
  - gap analysis
  - cost of repair

Sensitivity Analysis

Techniques

- Every evaluation is subject to uncertainty.
  - Are my measured values accurate?
  - What if my assumptions concerning weights are invalid?
  - How dependent are the results on a particular criterion?

- Uncertainty creates risk.

- Sensitivity analysis shows how your selection results react to changes in your weights or scores.
Gap Analysis

Techniques

• Gap Analysis provides a sense of best fit.
  – especially useful to understand end-user process mismatches
  – doesn’t use mathematical aggregation
  – can provide a basis for cost of fulfillment calculations

• A gap analysis typically uses a matrix
  – criteria down the side
  – candidates across the top
  – cells may be filled with text on
    • how a product feature provides the function
    • how well the product fulfills the criterion

Cost of Fulfillment

Techniques

• One way to view results is to consider the cost of fulfilling a deficit in a product.
  – “deficit” is judged with respect to the system needs

• Understanding the cost of fulfillment requires you to make assumptions about the use of the product in the system:
  – architecture and design
  – impact on maintenance
  – business considerations

• This requires expertise outside the scope of evaluation.
  – Evaluators must team with others to determine possible approaches to the fulfillment and their estimated costs.
Cost of Fulfillment

Evaluation Requirement(s) → Candidate's Capabilities

Evaluation Requirement(s) → Cost of Fulfillment

Candidate's Capabilities → Cost of Fulfillment

Outline

- Introduction
- Fundamentals of COTS Product Evaluation
- PECA - An Evaluation Process
- Techniques for Evaluation
- Last Words
A Little Dose of Reality

The process is a means, not the end.
- Boundaries between steps may become blurred.
- There is no “one size fits all” in COTS evaluation.
- The process should never get in the way of making a good recommendation.

COTS evaluation is an ongoing activity.
- Evaluation is not once and done.
- Evaluations build on previous evaluations.

A Little More Reality

Uncertainty is a part of COTS evaluation.
- Even a great job will leave doubt and risk.
- Sometimes there is no clear winner.
- Time erodes confidence.

COTS evaluation is not an isolated task.
- It is performed in the context of a larger process.
- It is performed in the context of the system.
Contact

John Dean  
E-mail: John.Dean@nrc.ca  
Phone: 991-6975  

Software Engineering Group  
Institute for Information Technology  
National Research Council Canada  
Ottawa, Ontario K1A 0R6  

Contact  

ICCBSS 2002  

http://www.iccbss.org