Embedded R&D for Cybersecurity in an Operational Environment

Luanne Burns Goldrich
Lauren Davis
Daniel Miller
Johns Hopkins University
Applied Physics Lab
11100 Johns Hopkins Road
Laurel, MD 20723
luanne.goldrich, lauren.davis, daniel.miller
@jhuapl.edu

Robert Gatlin*
Brian Gattoni*
Department of Homeland Security
4601 Fairfax Drive
Arlington, VA
robert.gatlin, brian.gattoni@hq.dhs.gov

*The opinions expressed in this article are the author’s and do not necessarily represent the position of the United States, the Department of Homeland Security, or any other entity

ABSTRACT

This paper describes a paradigm shift from how cybersecurity research and development (R&D) is traditionally applied in an operational environment. The methodology is referred to as embedded R&D (eR&D); cybersecurity researchers are tightly coupled with the operational stakeholders. This tight-knit relationship allows the researchers to elicit R&D requirements from the stakeholders seamlessly on a continuous basis, and gives the researchers immediate access to the tactical environment in which the analysts and operators work; this helps close the gap in the age-old disconnect between the research and operational communities.

Tools may be employed to enhance, augment, and advance the mission of an eR&D team. One such tool, REQcollect [5], was developed through several agile development iterations and the through the transition of other projects. Multiple federal agencies have sponsored the work and subsequently transitioned the technologies into use. The predecessors to REQcollect are REQdb (REQuirements Database) and DART3 (Department of Homeland Security Assistant for R&D Tracking and Technology Transfer) [6]. REQcollect combines the best components of these two systems: a requirements elicitation and collection tool and a Google-like matching algorithm to identify potential transitions of R&D projects with similar or identical requirements.

1. INTRODUCTION

The rapid proliferation of threats to the nation’s critical information networks and infrastructure demands that the nation develop new concepts and technologies for cybersecurity. To be complete, the discussion about integrating R&D into operations must include an overview of work done in the technology transition space, the inherent problems encountered there, and the low probability of success because of those problems.

Communication is one challenge for technology transition; not only must researchers and transition partners find and connect with each other, they must learn to speak a common language in order to facilitate transitions. Researchers tend to speak in technical terms (e.g. how they solved the problem, equations, mathematics) while transition partners, venture capitalists (VCs), and vendors speak in more business-oriented terms (e.g. return on investment (ROI), value-add, risk).

The goal of the eR&D team is to promote projects for possible transition and integration into operational environments. Technologies, tools, and knowledge are all subjects of interest for integration into operations. Because the eR&D team is an enabler of operations, it has both a perspective on current operations and an R&D vision that can assist in advancing current or future operations. Figure 1 shows the relationship of eR&D to operations and to the research community.

2. TECHNOLOGY TRANSITION

To be complete, the discussion about integrating R&D into operations must include an overview of work done in the technology transition space, the inherent problems encountered there, and the low probability of success because of those problems.
Both entities want to reach out across the valley of death and build the transition bridge. The bridge construction requires that they communicate and translate; only then is it possible to traverse the valley.

The work of Briggs et al. [1] extended the Technology Acceptance Model (TAM) by introducing the Technology Transition Model (TTM). TTM posits that actual system use is a function of behavioral intention \(B\), in which \(B\) is a multiplicative function of perceived net value \(V\) and perceived frequency of net value \(F\). Perceived net value is defined as an attitude, a subjective assessment based on the perceived consequences of changing from the existing technology to the newly proposed technology. Note that \(V\) is not a measure of how different life may be with the new system, but how the user feels about the likely difference. Users also consider how frequently \(F\) they expect to derive value from the new technology. TTM suggests that \(F\) and \(V\) combine multiplicatively to cause \(B\). Thus, the relationship between \(F\), \(V\), and \(B\) may be expressed:

\[
B = f(F \times V)
\]

\(F\) may be zero or positive, it cannot have a negative value because there is no frequency that occurs less than zero times. No matter how high \(F\) or \(V\) becomes, if either is zero, then \(B\) is zero. If \(V\) becomes negative, \(B\) may also become negative, and the user may avoid system use altogether [5].

TTM introduces another construct to consider when determining technology transition. TTM posits that users consider the perceived value of transition \(T\) when deciding whether to accept a new technology. While \(F\) and \(V\) represent the value derived from comparing the new system to the old system, \(T\) represents the derived value of the transition activity itself. The two primary dimensions of \(T\) are switching costs and switching benefits. For example, the learning curve associated with the new system would represent a switching benefit. On the other hand, being regarded as the champion of a game changing technology may be regarded as a switching benefit. Thus, \(T\) is a moderating construct on \(F\), \(V\), and \(B\):

\[
B = f(F \times V + T)
\]

One of the findings of Briggs et al. was that in order for perceived net value to be positive, a repeatable process must be introduced in support of the proposed technology for transition. This is the equivalent of the Standard Operating Procedures (SOPs) that are critical for Security Operations Centers (SOCs) to perform effectively today [7]. This supports the eR&D methodology proposed in this paper, and applies to the dynamically changing cybersecurity landscape.

3. MODEL for eR&D TEAMS

3.1 Overview

To maximize the potential for a proposed methodology to operate as intended, developing a comprehensive model will clarify challenges in the problem space and enhance the likelihood of a successful implementation and integration (Figure 3).

![Figure 3 eR&D Methodology](image)

In the Identify cycle, the eR&D team collaborates with operational areas within the organization to determine and elicit the operational, or stakeholder, R&D requirements. As these requirements are refined through the Identify process, the requirements are fed into a repository (REQcollect, described in section 4) where the requirements can be stored, queried, categorized, prioritized, archived and maintained over time. As progress is made in the Identify cycle, it enables movement or activity in the other cycles.

In the Discover cycle, the team is outwardly facing and is tasked with discovering ongoing research projects that are being conducted across the cybersecurity spectrum. Universities, consortia, and other research organizations may also perform research of interest. The team discovers technologies by collaborating with other cybersecurity organizations, attending conferences, and researching potential projects. Once an accurate, concise summary of a project is written, the Discover cycle feeds the technologies into the repository. Note that the Discover cycle is driven by a detailed knowledge of the R&D requirements elicited during the Identify cycle.

As these two cycles are executed, validated requirements and technologies are added into the repository; the built-in Google-like correlation algorithm in the Filter cycle automatically finds closely related or similar requirements with technologies that will potentially satisfy the requirements, in whole or in part. Once matches are provided, the Filter cycle continues; the eR&D team uses its unique knowledge of the requirements and operational environments of the stakeholders to further examine and possibly select a candidate technology for promotion to a stakeholder.

Lastly, the Integrate cog turns as the eR&D team facilitates the integration of filtered/matched technologies to the stakeholder; appropriate subject matter experts (SMEs) are involved with the research team. The eR&D team guides the technology and stakeholders through the integration, and ultimately relinquishes...
the technology to the stakeholder, as shown in Figure 4 below. As the Integrate cycle is executed, lessons learned and important process improvement information are fed back into the repository.

![Figure 4 R&D Integration Stakeholder, PI, eR&D Involvement](image)

The tools employed for each cycle of the eR&D process help narrow the Valley of Death by moving either the start or finish lines of the technology transition process. Figure 5 illustrates the toolset used by eR&D and indicates the specific tools that have the effect of narrowing the Valley of Death by moving the start or finish lines. The embedded R&D process, the REQcollect [5] repository and Agile Research [3] affect the starting point by moving Research closer to Operations; REQcollect, Agile Research, Analyses of Research (AoR), and R&D Shepherding shorten the distance by moving Operations closer to Research. Agile Research and AoRs will be discussed in the following sections.

![Figure 5 Moving the Transition Start and Finish Lines](image)

### 3.2 REQcollect

The Requirements Collection Repository (REQcollect) [5] is used by the eR&D team to gather and store requirements and project information, facilitates correlations and assists in launching transitions. It is the central repository for storing information about elicited requirements, discovered projects and technologies, matches between requirements and technologies, integration activities and lessons learned. The Johns Hopkins University Applied Physics Laboratory (JHU/APL) developed REQcollect to further enhance the R&D mission set of Federal Departments and Agencies. REQcollect runs an automated Apache Lucene [2] matching algorithm to complete a Google-like full-text search over project descriptions and requirement keywords to suggest prioritized lists of matches between requirements and projects.

After matches are made, users may select elements and characteristics for technology transition. The centralization and standardization of project and requirement data provide easy access and automated, suggested matches and discovery. [5]

Before the development of REQcollect, matching an R&D requirement to a research project was a time-consuming, manual process. The algorithm used by REQcollect streamlines the approach and supports more objectivity in the selection process by eliminating human bias in selection of projects to investigate. For more information on REQcollect, see [5]. The web-based interface of REQcollect, allows for easy insertion and editing of requirements and projects. The reporting utilities provide an easy interface for extracting requirements and generating reports. Report generation allows user-specified sorting: for example, requirements by priority, by organization, by category or by multiple fields. Requirements can be deprecated and reports can be run to include or exclude deprecated requirements. Figure 6 shows the REQcollect home page. The following sections elaborate the use of REQcollect for each of the methodology cycles.

![Figure 6 REQcollect Web-based Frontend](image)

### 4. eR&D PROCESS DETAIL

#### 4.1 Identify Cycle

There are four phases in the eR&D Identify process: elicitation (through embedded participation), analyze requirements, stakeholder feedback and requirements validation. These phases are depicted in Figure 7 as elements of the cyclical process. The eR&D team meets annually with the operational teams to ensure a comprehensive set of R&D requirements is gathered; requirements submission by the operational teams is also encouraged throughout the year.

During the elicitation phase, the eR&D team uses established requirements engineering techniques to gather requirements from the stakeholders. The techniques include but are not limited to use cases, interviews, and site visits. The techniques are selected based on stakeholder availability and complexity of the requirements.

During the requirements analysis phase, the eR&D team pays careful attention to the wording and exact meaning of the statements of requirements; these must be scrutinized so that the interpretation is unambiguous. Requirements must be testable, unambiguous to the extent possible at this level of abstraction, unique, and robust. This ensures that attempts to satisfy the requirements will result in finding the most appropriate technologies for operational integration.
Once the eR&D team determines the requirements are clearly stated with all these aspects satisfied, the stakeholder is asked to review them in the feedback phase. During the feedback phase, the stakeholders review the requirements for accuracy and completeness, and prioritize them. The stakeholders then relay comments and concerns about the requirements to the eR&D team.

The eR&D team then adjudicates any comments gathered during the validation phase. The final part of the Identify cycle is the validation phase; the requirements are peer reviewed by objective SMEs outside the operational areas and outside of the eR&D team. This may result in additional discussions with the stakeholders to address any concerns. Using the peer review process, the eR&D team ensures more concise and valid R&D requirements specification. Once comments have been adjudicated and all stakeholders are in agreement, the requirements are placed under configuration management and are ready to be entered into the REQcollect, where they are matched with potential technologies in the filter and integration cycles.

### 4.2 Discover Cycle

The Discover cycle, like the Identify cycle, is ongoing -- it is accomplished once for each relevant project identified. There may be many projects in this process simultaneously. There are three phases in the Discover process: exposure, investigation, and summary (Figure 8).

**Figure 8 eR&D Discover Cycle**

The eR&D team is first exposed to or becomes aware of a federally funded research project. This occurs when collaboration yields information about relevant research or when eR&D team members travel to other Federal agencies, academic institutions, Federally Funded R&D Centers (FFRDCs), University Affiliated Research Centers (UARCs), National Labs, symposia, conferences, and seminars in order to discover the most recent developments in federally funded R&D efforts. This allows the team to find those projects that best meet the needs of its stakeholders.

Visits to organizations are based on high-priority R&D requirements. Other events are considered as they apply to specific areas of operational need. Visits to organizations produce projects or action items that support the proceeding phase.

Once the eR&D team becomes aware of a federally funded research project in the cyber security domain, the project enters into the investigation phase. During this phase, the eR&D team learns more about the research, including funding, sponsors, maturity level, current results etc. This information is obtained by reading white papers, meeting with PIs, and reviewing presentation materials. Once the eR&D team has a basic understanding of the project, the project proceeds to the summary phase, where an Analysis of Research (AoR) document is initiated (a portfolio of documents about a research investigation).

### 4.3 Filter Cycle

The first phase of the Filter cycle is to use REQcollect’s automatic matching capability to identify possible matches between projects and requirements. The matches are a starting point for pairing requirements with projects; the automated keyword matcher is run at regular intervals to keep the matches current, in the event projects or requirements are added or changed. An eR&D team member reviews the suggestions and then investigates further to determine if they are viable. The Filter process is illustrated in Figure 9.

**Figure 9 eR&D Filter Cycle**

Once REQcollect has chosen the projects that best correspond to a requirement, the eR&D team, using its unique knowledge of the stakeholder’s operational environment and requirements (information obtained in the Identify phase, but not captured in the text of the requirements document), further investigates the top matches from the perspective of what would likely satisfy the operational requirements of the stakeholder. Information gained in this process is added to REQcollect, regardless of whether or not a project is selected for operational integration. If this process results in more than one viable technology, the eR&D team consults with the stakeholder to determine which project is the best candidate to satisfy their requirements.

The second phase of filtering is for selecting relevant matches. The automated algorithm suggests reasonable matches, but it is then the eR&D team selects matches that are relevant for high-
priority requirements. Note that the output of the filter cycle may be an empty set; this would mean no technology/project was found for a given requirement. Possibly the automated matcher did not find any matches or the ones found were eliminated by the eR&D team as irrelevant. In this case, impromptu discovery must be conducted if the requirement is of high priority or the stakeholder needs dictate that an R&D input is essential. This is discussed in the next cycle, Integrate.

4.4 Integrate Cycle

There are four phases in the Integrate process: investigate, initiate, increment, and integrate (Figure 10). If the REQcollect contains an existing technology that has been determined to match an R&D requirement (in whole or in part), the eR&D team would use the investigate phase to validate that both the stakeholder and the project principal investigator (PI) are in agreement with the match quality.

![Figure 10 eR&D Integrate Cycle](image)

Figure 10 eR&D Integrate Cycle

If no reasonable matches are found, the investigation would require a more in-depth approach because it would involve an impromptu discovery or survey of the available R&D to determine if a project exists that may have not been previously discovered and placed in REQcollect. This would require an investigation by relevant SMEs. The two types of investigation are summarized below.

- Technology found in REQcollect that potentially satisfies an R&D requirement
  - Investigation with PIs - Initiation – Integration
- No Technology found in REQcollect to satisfy an R&D requirement
  - Investigation by SMEs – Initiation - Integration

An AoR may contain a target recommendation, plan, status report(s) and a final report. It is a useful compilation of the artifacts produced in each phase of the integration process and resides in a central location. The eR&D director, AoR manager, and eR&D team may then leverage and determine the status of an AoR at any given time. The first document added to an AoR is an AoR Target Recommendation; it is created for each AoR investigation.

When a project has been selected to move forward, work on the AoR plan begins. The AoR plan is a high-level assessment of what it means to integrate a specific technology for a specific stakeholder. At this stage, SMEs from the stakeholder organization become actively involved in detailing what they need the technology to do in order for it to satisfy the stakeholder’s requirements. This early interaction is key to developing a system that allows the eR&D team to meet its goal of helping stakeholders take advantage of advances in technology research. This interaction results in requirements and acceptance test criteria by which the technology will be evaluated when the integration is complete.

The eR&D team arbitrates the SMEs’ requirements with the current design and/or functionality of the technology, the notional schedule and budget, and external constraints that may be placed on the researchers (e.g., conflicting requirements from another funding source). The eR&D team also arbitrates a notional schedule and budget between the stakeholder and the researchers. Once the stakeholders, SMEs, and researchers agree on the requirements and acceptance test plan, and that the requirements are in harmony with the notional schedule and budget, the AoR plan can be completed. At this point, all stakeholders must agree to the contents of the plan, and a Go/No-go decision is made about continuing the integration.

If the decision is made to not move forward with the integration, any additional project information gathered in support of the integration is captured in the REQcollect, lessons learned are used to inform the eR&D processes, and the project’s portfolio is archived (operational use). If a decision is made to proceed with the integration, an integration manager is selected to oversee the integration.

Table 1 represents a list of artifacts that are complete by the end of the investigation phase and

Table 2 is a list of artifacts that are complete at the conclusion of the initiation phase.

**Table 1: Artifacts at the conclusion of the Investigation Phase**

<table>
<thead>
<tr>
<th>Artifact</th>
<th>When</th>
<th>Responsible Party</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Requirements</td>
<td>Match/No Match</td>
<td>eR&amp;D Team</td>
<td>AoR Recommendation</td>
</tr>
<tr>
<td>Survey of Research</td>
<td>No match</td>
<td>SMEs</td>
<td>AoR Recommendation</td>
</tr>
<tr>
<td>Suggested Directions</td>
<td>Match/No Match</td>
<td>SMEs</td>
<td>AoR Recommendation</td>
</tr>
<tr>
<td>Proposed Team</td>
<td>Match/No Match</td>
<td>SMEs</td>
<td>AoR Recommendation</td>
</tr>
<tr>
<td>Other questions to consider</td>
<td>No Match</td>
<td>SMEs</td>
<td>AoR Recommendation</td>
</tr>
</tbody>
</table>

**Table 2: Artifacts at conclusion of the Initiation Phase**

<table>
<thead>
<tr>
<th>Artifact</th>
<th>Responsible Party</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirements</td>
<td>eR&amp;D Team</td>
<td>AoR Plan</td>
</tr>
<tr>
<td>Acceptance Test Criteria</td>
<td>eR&amp;D Team</td>
<td>AoR Plan</td>
</tr>
<tr>
<td>Notional Schedule/Budget</td>
<td>eR&amp;D Team</td>
<td>AoR Plan</td>
</tr>
<tr>
<td>Technology Information</td>
<td>eR&amp;D Team</td>
<td>AoR Plan</td>
</tr>
</tbody>
</table>
The integration cycle continues with *incremental*-Agile Research to integrate the technology with a step-wise, organized approach. The Agile Research [3] process is fast, incremental, and transparent; it is designed to produce actionable, impactful results. Agile Research is based on the following principles:

- Resources and logistics for cyber- security research can be predefined for fast response when needed
- Research can be organized into iterative, accumulating increments that produce actionable results
- Visibility and transparency with sponsors permit fast responses to evolving research findings

These increments are designed to keep the stakeholder aware of progress and allow changing direction as needed. At each phase in integration, there is a go/no go decision made on whether to continue. The artifacts for each increment are an AoR Status Report, any testing results, documentation and a demonstration.

The final phase is integration, which consists of fully testing the technology and deploying it to the operational environment. The artifacts at the completion of integration are the same as each incremental status report but also pull together a summary of each increment, an executive summary and lessons learned.

The technology integration continues by transferring control to the stakeholder organization that is receiving the technology. Various organizations have specific processes in place for technology insertion and, at this point in the integration; those processes must be invoked to determine funding, scheduling and staffing. Figure 11 illustrates the relationship and communication between the researchers, the stakeholder organization and the eR&D team.

**5. Summary**

The eR&D mission is achieved through four processes -- Identify, Discover, Filter, and Integrate -- which work together through the REQcollect in order to provide operational areas with the most recent advances in research technology. The complete eR&D process is shown in Figure 12.

![Figure 11: Integration Communications](image)

### 6. Future Work

The Integration cycle of the eR&D process can be refined by incorporating lessons learned during ongoing integrations; using Agile Research [3], transitions can be accelerated and allow operational areas to employ state-of-the-art technologies. While R&D may be initially less costly than commercial products, over time, ongoing operations and maintenance may be more difficult. Establishing a cooperative and collaborative infrastructure and agreements with research organizations and those to whom they license their technologies will be essential to long-term success.

### 7. References


