Improving Incident Response Through Simplified Lessons Learned Data Capture

Andrew Baze

Improving Incident Response Through Simplified Lessons Learned Data Capture

GIAC GCIH Gold Certification
Author: Andrew Baze, abaze@hotmail.com
Advisor: Dr. Johannes Ullrich

Accepted: February 3, 2021

Abstract

The Lessons Learned portion of the cybersecurity incident response process is often neglected, resulting in unfortunate missed opportunities that could help teams mature, identify important trends, and improve their security. Common incident handling frameworks and compliance regimes describe time-consuming and relatively complex processes designed to capture these valuable lessons. While an extensive and resource-heavy process may be necessary in some cases, it is often difficult for incident response teams to dedicate sufficient time to capture this lesson data at the end of an incident. Dedicating time is even more difficult when the team is simultaneously handling other incidents. This paper addresses the planning and implementation of a simplified approach to capturing Lessons Learned data at any time, as opposed to at the conclusion of an incident. This approach includes a tagging schema and demonstrates how identification of lesson type, sub-type, and associated work items can provide valuable data to further an organization's original Lessons Learned goals.

1. Introduction: Why Incident Response Guidelines are Often Not Followed

Most security professionals are familiar with the two most common cybersecurity incident response (IR) frameworks and their phases. Both the NIST and SANS frameworks begin with a planning and preparation phase and end with an incident review and learning phase. Table 1 below compares the language from both models.

NIST 800-61 Rev. 4 "Incident Response Life Cycle"	**SANS "PICERL"**	**Activities Per Phase**
Preparation	Preparation	Prepare documentation, plans, resources, training, etc. before an incident takes place
Detection & Analysis	Identification	Determine an incident has taken place, conduct triage
Containment, Eradication & Recovery	Containment	Stop the incident from spreading, continuing, or becoming worse
	Eradication	Eliminating the incident's effects and in some cases, the cause
	Recovery	Getting the business back to a functioning state
Post-Incident Activity	Lessons Learned	Reviewing the incident to identify areas for improvement or validate existing processes

Table 1: NIST 800-61 vs. SANS Incident Handling Phase Comparison

This study focuses on the last phase of the incident handling process. The "Post-Incident Activity" (NIST terminology) (Cichonski, Millar, Grance, & Scarfone, 2004) and the "Lessons Learned" activity (SANS terminology) (Skoudis & Strand, 2018) are intended, in summary, to identify areas for improvement after each incident, with the ultimate goal of extracting additional value from each incident and directing improvements into the Preparation phase of the Incident Response Lifecycle. Both the

Andrew Baze, abaze@outlook.com

"Post-Incident Activity" and "Lessons Learned" phrases will be referred to as "Lessons Learned" or "LL" in this document.

From the SANS Incident Handling curriculum, specific LL recommendations include the following:

- Develop a follow-up report, encourage all affected parties to review the draft, attempt to reach consensus and get signoff.
- After the report has been reviewed, schedule a Lessons Learned meeting… to get consensus on the Executive Summary of the report.
- Do this within two weeks of resuming production, with a maximum length of half a day. (Skoudis & Strand, 2018).

Figure 1 below further illustrates the relationship between Lessons Learned (Post-Incident Activity) and Preparation phases.

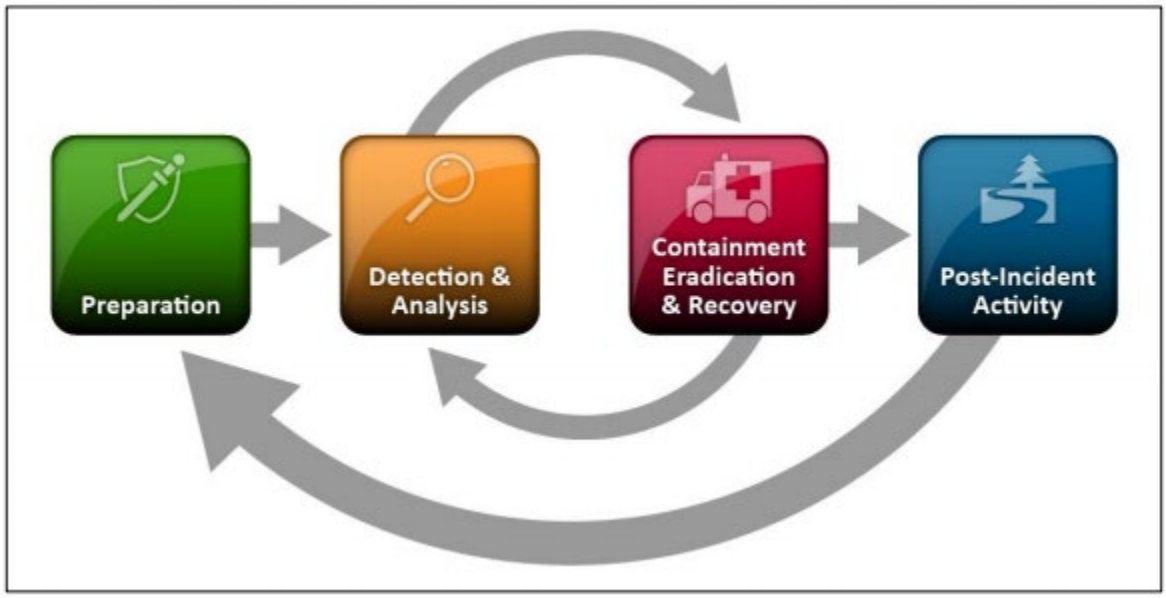

Figure 1: The Incident Response Life Cycle from NIST.SP.800-61r2 (Cichonski, Millar, Grance, & Scarfone, 2004)

The largest, bottom arrow in the illustration links "Post-Incident Activity" with "Preparation." This arrow represents the importance of learning from each incident and subsequently informing incident response plans.

Andrew Baze, abaze@outlook.com

2. The Problem – Lessons Learned Data Is Not Captured

Even though the use of Lesson Learned data is specified in NIST's Computer Security Incident Handling Guide (SP 800-61 Rev. 2), the guide also states, "One of the most important parts of incident response is also the most often omitted: learning and improving" (Cichonski, Millar, Grance, & Scarfone, 2004).

NIST is not alone in that assessment. A study guide for the well-known GCIH (GIAC Certified Incident Handler) certification states "Unfortunately, such learnings are often not collected at all, and when they are collected, they aren't consistently incorporated into the Preparation phase" (Mitropoulos, 2020). Even though the Post-Incident Activity phase is explicitly called out, it appears that it is frequently neglected. Determining why will help organizations find an appropriate solution. Some maintain that it is simply forgotten. As Thompson states, "One of the most important and often forgotten elements of the incident response program and execution of the incident response plan is conducting lessons learned" (Thompson, 2018). However, if a process is clearly described and enforced, it will not be forgotten. In other words, the lack of a Lessons Learned process likely indicates nonexistent or unenforced policy.

Even if a policy is clearly described, there is an inherently reactive and time-sensitive nature to incident response. As a result, there is often a perceived need to move on to the next incident as soon as the most critical steps of the current incident have been concluded. As Thompson further states, "Entities do not find the time to go over the incident response process and document what was effective and what needs improvement" (Thompson, 2018). When time is of the essence and the incidents keep coming in, something must give. Two common options are to reduce the time spent at the start or end of the process, by neglecting: 1) planning documentation improvements (at the beginning) or 2) Lessons Learned data capture and review (at the end).

In addition to the IR guidance called out in the SANS curriculum in the previous section, incident responders should consider this similar guidance:

> In addition to compiling written documentation for each incident, all relevant stakeholders should gather for an after-action meeting within a week or so of the

> incident being resolved. The meeting should include all members of the incident response team, relevant members of management, as well as other impacted groups, including development, operations, legal, public relations, and others as may be appropriate" (Anson, 2020).

What should an incident response team do when its team members are actively engaged in Identification, Containment, Eradication and Recovery activities? For many, such guidance may not be practical to follow. Suppose an Incident Response lead must choose between tracking down and coordinating feedback meetings for development managers, lawyers, the public relations team, etc., versus handling a newly identified security incident. In that scenario, the new incident will most likely take precedence. In many other cases, not only will there be no Lessons Learned meeting, but there may not be any capture whatsoever of LL data.

When an incident response team does not conduct a realistic LL session (whether in a formal meeting or otherwise) and is does not appropriately manage the outcomes from a mature LL process, that team does not comply with most incident response standards and misses critical opportunities to improve. Writer George Santayana states, "Those who cannot remember the past are condemned to repeat it." In other words, an incident response team that does not effectively record and fix problems is condemned to revisit them.

In the context of continuous improvement and the Capability Maturity Model, a team that is not able to embrace a consistent Lessons Learned approach to incident management will never reach Level 4 ("Quantitatively Managed"), with team processes defined in part as "refined and adapted." Reaching Level 3 ("Defined") will also be difficult to achieve since it has a defined and proactive approach informed by Lessons Learned. Figure 2 below illustrates key differences between maturity levels.

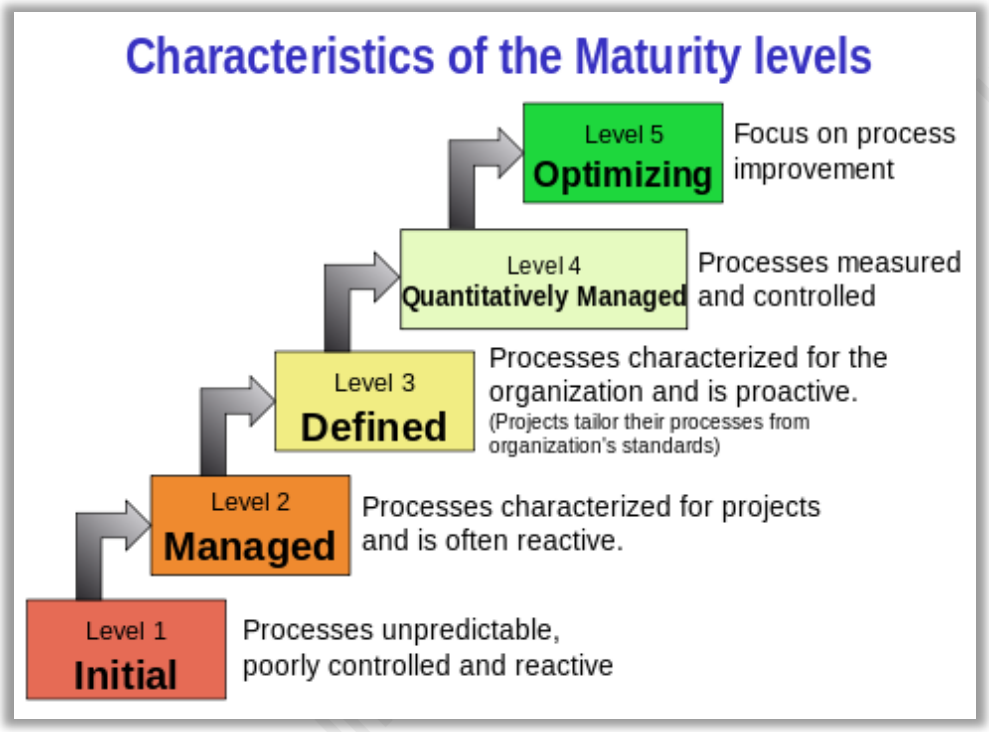

Figure 2: Characteristics of the Maturity Levels (Capability Maturity Model Integration, 2020)

Conversely, if an incident response team can capture LL data efficiently and ensure that problems are prioritized and proactively managed, even while handling overlapping incidents, continuous improvement and increased maturity are within its grasp.

Even a team that primarily performs reactive tasks, with little time or staff available for proactive process improvement work, could theoretically achieve Level 4. In these cases, capturing and measuring a variety of data types is likely happening already. Many teams track measures and metrics such as the time it takes to acknowledge an incoming incident, to close an incident, or the number of incidents per incident type by an organization over time. If an incident response team can capture Lessons Learned measurably, they are aligned with the spirit of this approach to process improvement: "Quantitative objectives for quality and process performance are established and used as criteria in managing the process. Quality and process performance is understood in

statistical terms and is managed throughout the life of the process" (Chrissis, Konrad, & Shrum, 2007).

The crux of this example is to capture and analyze the LL data, which facilitates improvements. How can it be captured efficiently, analyzed, and used to cause necessary change?

This paper describes a method by which a team of any maturity level can identify Lessons Learned efficiently and inexpensively. A case study was conducted using a three-phased approach with varying levels of data capture, as described in the next section.

3. Proposed Solution – Three-Phased Implementation of LL Tagging and Analysis

For this study, a three-phased approach was used. Phase 0 provided a baseline dataset to compare against in later phases. Phase 1 implemented simplified LL tagging and reporting. Phase 2 implemented more detailed tagging LL data during any phase of the incident (as opposed to only at an incident's conclusion) in addition to capturing associated work item data. While full incident data was available during the study, many details such as incident type breakdown, severity, incident titles and customer information cannot be shared publicly due to their sensitive nature.

Identifying and reporting on work items related to incidents and lessons was primarily facilitated by the change in Phase 2 that enabled the capture of associated work item links for each LL item via the "work item" tag in the expanded tagging. The work items captured apply to the incident response team, its customers, or anyone else in the organization. Creating the expectation that relevant work items that address each incident's root cause (when applicable) was intended to make the problem-solving conversations much easier, as they would be founded on solid incident data.

Data was reviewed across all phases to assess differences in LL quantity and quality using each approach. Each phase was built on top of the prior, expanding on both

quantity and quality of data captured, and increasingly exposed richer reporting capabilities.

This approach was only a first step and by no means intended to be a well-vetted, complete program. Such an endeavor would require several more months of additional analysis and adjustments to the process, as described in "Implications for Future Research."

A checklist of the steps a team can follow to deploy a similar program or related updates is provided in Appendix C: Checklist for Implementing Improved Lessons Learned Data Capture and Reporting.

3.1 Phase 0 Approach

Implementing Phase 0 required no changes and already used the policy, resources, and measures described as follows.

3.1.1 Phase 0 Policy

The policy that existed in the first phase of this experiment read as follows, with ***emphasis*** added in areas of the policy that introduced ambiguity (e.g., "may" versus "must").

> The incident response team shall capture Lessons Learned for every incident after the incident. For Severity 2 or Severity 3 cases, this ***may*** take place in an ad hoc manner, e.g., by capturing issues after each incident in the case management system. Data ***could*** be obtained via ad hoc conversations with key stakeholders, via email, a formal meeting, or any other communication method. For Severity 1 incidents, a meeting will be held with relevant stakeholders. Team members will enter all Lessons Learned data into the case management system. In all cases, team members will update incident response preparation materials (e.g., the IR [Incident Response] playbook) and enter applicable work items into the work management system.

Not only is the process described in this phase intended to represent the situation for many IR teams, but it is also intended to present a feasible mechanism for conforming to NIST 800-53 Rev. 4. expectations, specifically the following two controls from the Incident Response Control Family (Dempsey, Witte, & Rike, 2014):

- IR-4.c: Incorporates lessons learned from ongoing incident handling activities into incident response procedures, training, and testing, and implements the resulting changes accordingly;
- IR-8.d: Updates the incident response plan to address system/organizational changes or problems encountered during plan implementation, execution, or testing.

Also, the process should meet the expectations of these controls in the NIST Framework for Improving Critical Infrastructure Cybersecurity, aka the "CSF," or "Cybersecurity Framework" (Keller, n.d.):

- PR.IP-7: "Protection processes are continuously improved;"
- RS.IM-1: "Response plans incorporate lessons learned;"
- RS.IM-2: "Response strategies are updated."

3.1.2 Phase 0 Policy Flaws

Regardless of whether this approach meets internal or external compliance requirements, using an "as needed" approach to capturing LL data for Severity 2 and Severity 3 incidents without a dedicated process for reviewing the data and associating it with tracked work is flawed in at least these three ways:

1. Collecting LL data only after the incident is largely concluded is not necessarily more efficient than capturing issues real-time, and could also result in responders forgetting issues, especially in long-running incidents.
2. A required Lessons Learned meeting for Severity 1 incidents (without a similarly explicit requirement for Severity 2 and 3 incidents) will likely result in only capturing data for those Severity 1 incidents. This focus also presumes that the best learning occurs only in high-severity incidents, which is not necessarily true.

3. A requirement to capture data with no regular review or reporting is not likely to drive the desired behavior. An axiomatic approach to scorecard use is that "What you measure is what you get" and "[an] organization's measurement system strongly affects the behavior of managers and employees" (Kaplan & Norton, 1992).

3.1.3 Phase 0 Resources and Data Capture

No new resources were required to capture data in Phase 0. LL items were captured in an open text field in the case management tool, and they were not reported on in any automated way. Improvements made were the result of careful attention by incident handlers during work item planning and execution efforts, during the rare times that proactive work was possible. A meeting was only required for the Severity 1 cases (as per guidelines).

3.1.4 Phase 0 LL Tagging

No tagging was in place during Phase 0.

3.2 Phase 1 Approach

Implementing Phase 1 required specific policy, resources, and measures, as follows.

3.2.1 Phase 1 Policy

The policy was updated in Phase 1 of the study as follows (with changed areas ***emphasized***):

> "The incident response team shall capture Lessons Learned for every incident at the conclusion of the incident. For Severity 2 or Severity 3 cases, this may take place in an ad hoc manner, e.g., by capturing issues after each incident in the case management system. Data could be obtained via ad hoc conversations with key stakeholders, via email, a formal meeting, or any other communication method. For Severity 1 incidents, a meeting will be held with relevant stakeholders. All

> Lessons Learned data will be entered into the case management system ***using the required tagging format***.
>
> Team members will enter all Lessons Learned data into the case management system. In all cases, team members will update incident response preparation materials (e.g., the IR [Incident Response] playbook) and enter applicable work items into the work management system. ***The team will review Lessons Learned regularly via a team reporting dashboard."***

3.2.2 Phase 1 Policy Flaws and Improvements

Flaws #1 and #2 from Phase 0 (to capture data at the end of the incident and only conduct a dedicated meeting for Severity 1 incidents) were not addressed by Phase 1 policy changes.

Flaw #3 in Phase 0 regarding lack of reporting was remediated with the Phase 1 change in policy, which required tagging. Implementing tagging enabled basic data measurement. Reporting on that data drove more attention to the problem space and supported better data capture.

3.2.3 Phase 1 Resources and Data Capture

As mentioned above, LL data was captured for all incidents using an updated tagging system. Given the legitimate concern that this could add a significant workload to a team handling many incidents daily, the system had to be lightweight. As in Phase 0, a meeting was only required for the Severity 1 incidents.

3.2.4 Phase 1 LL Tagging

The case management toolset used in this case contains free-form text fields, so JSON (JavaScript Object Notation, a relatively simple data interchange format) markup was used to capture LL data as follows:

```
{
        "LessonsLearned": [
                { "Lesson": "first lesson text goes here",
                },
                    { "Lesson": "next lesson",
                    }
                    { "Lesson": "last lesson – no comma"
                    }
          ]
}
```

This approach facilitated extraction and reporting using a PowerBI (a Microsoft data visualization platform) dashboard.

The only reporting expected in Phase 1 was a month-end review of the previous month's data. There was no expectation to show a link between the LL data and major organizational efforts.

3.2.5 Phase 1 Measures

The measurements that accompanied the Phase 1 are as follows:

1. Count of incidents.
2. Count of LL items.
3. Count of LL items per incident.

The incident counts and LL items per incident were simple to calculate. However, there were neither references to work items in the LL data nor any automated way to extract work item data from the case management tool, since the case management and work tracking systems did not work together natively. As a result, there was no way to measure associated work items, much less their states of "not started," "started," or "complete" using automated reporting.

Also, given the ad hoc nature of ongoing improvements (e.g., a real-time playbook update made right after LL data was entered), there was no way to systematically measure this approach's effectiveness in terms of resulting changes. These gaps were by design. Phase 1 was not intended to be an end state, as it could not support

measurements required to support an increase in organizational maturity as called out in the CMMI model. These issues were addressed in Phase 2.

3.3 Phase 2 Approach

Phase 2 contained several important updates, including changes to policy and new tagging requirements, facilitating improved reporting and analysis capabilities.

The updated approach required these key modifications to address flaws identified in the Phase 1 policy:

1) Track LL data during any phase of the incident, facilitated by a simple tagging system in the case management toolset. Adding final or aggregated LL data may still occur during the LL phase of the incident, but it is not restricted to that phase. This will enable the incident handler to capture LL data while it is fresh and can still be reviewed, updated, and used in aggregate during the LL phase as needed.
2) LL data will be programmatically extracted from the case management system and reviewed weekly or bi-weekly, including incidents that are still in progress (which may already have LL data captured). This review will reinforce the behavior that LL data is to be captured on an ongoing basis. Using this data in discussions with key stakeholders may also reduce the need for a specific meeting if appropriately socialized, reviewed and approved by key stakeholders.
3) LL items will also be updated as applicable with associated work item links from the work tracking system. Those associated work items and their status (not started, started, complete) will be reviewed with the rest of the LL data. This approach will drive one of the ultimate lagging indicators for LL success: completed, LL-related work items.

Among other things, these modifications were intended to help combat the issue of the amount of time required to schedule a formal meeting, especially when the incident has been effectively mitigated and other incidents need attention.

It was hypothesized that this approach of rapid and low-cost capture of LL data, along with the accountability enabled by reporting the LL data and associated work items

per incident, would result in a more effective LL program. Effectiveness is measured by the amount of work that addressed root cause issues and otherwise improved IR planning, which is the realization of the arrow that connects the post-incident activity phase with the preparation phase.

3.3.1 Phase 2 Policy

The Phase 1 Policy was amended at the beginning of Phase 2 to the following (with changed areas ***emphasized***):

> "The incident response team shall capture Lessons Learned (LL) ***during every incident as lessons/problems are identified, and at the conclusion of each incident as needed. Regardless of severity, this may take place in an ad hoc manner***, e.g., by capturing issues after each incident in the case management system. Data could be obtained via ad hoc conversations with key stakeholders, via email, a formal meeting, or any other communication method. For Severity 1 incidents***, if no LL data has been captured previously or if other LL data needs to be captured,*** a meeting will be held with relevant stakeholders. All LL data will be entered into the case management system using the required format [***see tagging format updated for Phase 2***].
>
> Team members will enter all Lessons Learned data into the case management system. In all cases, team members will update incident response preparation materials (e.g., the IR [Incident Response] playbook) and enter applicable work items into the work management system. The team will review Lessons Learned data regularly via a team reporting dashboard."

These minor policy changes were intended to encourage data capture as it surfaced, as opposed to later in the incident. Also, it should have especially incentivized that behavior for Severity 1 cases, since a separate meeting may not be necessary (or at least, such a meeting may be conducted more quickly) if the necessary LL data had been captured throughout the incident.

3.3.2 Phase 2 Policy Flaws and Improvements

The requirement for real-time data capture for LL data could be a possible flaw. However, long-term trend analysis, team interviews, and other methods would need to be implemented to determine this.

For improvements, the Phase 2 policy removed the restriction of capturing LL data only during the final incident phase, ideally enabling more comprehensive data capture overall.

The second policy update removed the requirement for a dedicated meeting if LL data had already been captured from the relevant stakeholders. This was intended to encourage data capture from all stakeholders during the incident versus only at the end.

Lastly, the enhanced tagging schema would enable more than simple measurements (e.g., count of LL per incident). With this change, much more valuable analysis can take place, based on LL type, sub-type, and associated work items.

3.3.3 Phase 2 Resources and Data Capture

To expand on the Phase 1 approach, this LL data was captured using tagging in the case management system with various modifications to facilitate better data quality and to support root-cause analysis, double-loop learning (see Appendix A: Double-Loop Learning), and subsequent organizational action via work item tracking.

The following LL types (again using JSON markup) in Table 2 were supported.

LL Type	**Change needed, yes/no**
What went well	No change needed
What didn't go well – Stop doing	Change needed
What didn't go well – Start doing	Change needed

Table 2: Lessons Learned type fields

The following sub-types in Table 3 were also supported.

LL Sub-Types (for "didn't go well" items)	Description
Playbook update (change existing)	Modify an existing playbook entry
New playbook entry	Only for new playbook entries
Training needed (IR team)	Team or team member needs training
Better reporting needed	The incident could have been identified or reported better (includes monitoring and alerting issues)
Partner communication issue	Challenges identified when communicating with a customer or other partner
Other partner issue	Any issue with partners aside from communications
NA (for "Keep doing")	Not Applicable - use when other sub-types do not apply for any reason
Internal tooling issue	Issue related to the IR team's tools
External tooling issue	Issue related to customer or external partner tools
Other dependency	Any other IR team dependency
Work item link (contains title)	Link to work item in the IR team's work tracking system

Table 3: LL sub-type fields

To translate the lesson data into planned, tracked work in the organization's work tracking system, the following work item data fields in Table 4 were supported.

LL work item data	Description
Work item link (contains title)	For all items that need a work item, put the link(s) here
NA	If not applicable – already fixed, or work item(s) not needed

Table 4: LL Work item data fields

Regarding LL type in Table 2, in addition to capturing "What didn't go well," other important data such as "What went well" could be captured. The two "What didn't go well" options also forced the responder to choose whether something that needed to

start or stop. The intent with this approach was to cause more thought to be given to the problem before it was entered.

There were several options for LL sub-types in Table 3, and only one could be selected, which facilitated better reporting and analysis, since this forced the person entering the data to possibly break out complex issues into their more useful, individual "problems."

Lastly, a work item entry from Table 4 for an associated work item was included, which contained a link to an entry in the separate work tracking system. This helped identify problems or areas for which there was no planned work.

3.3.4 Phase 2 LL Tagging

The JSON schema for the Phase 2 Lessons Learned data was as follows:

```
{
        "LessonsLearned": [
                { "Lesson":"first lesson text goes here, and repeat at this level as needed",
                        "LLType":"Type goes here",
                        "LLSub-Type":"Sub-type goes here",
                        "LLWI":"work item /ADO link goes here",
                        "LLWI":"add work items here as needed, one per line"
                        },
                { "Lesson":"last lesson text goes here – no comma after curly brace",
                        "LLType":"Type goes here",
                        "LLSub-Type":"Sub-type goes here",
                        "LLWI":"work item link goes here"
                        "LLWI":"add work items here as needed, one per line"
                }
          ]
}
```

3.3.5 Phase 2 Measures

Just as the data capture for Phase 2 is more detailed, the reporting becomes richer.

While Phase 1 reporting only showed LL counts per incident, Phase 2 reporting allowed for aggregation of data that would eventually allow examination of trends, such as most common LL sub-types per incident type, most common LL sub-types overall,

incident sub-types with most or fewest associated work items, and LL items with no associated work items.

The use of three phases in this study was intended to meet two goals. The first was to ease the team into the tagging model over two months (during Phase 1 and Phase 2). The second was to demonstrate that the simple Phase 1 data capture approach was not sufficient for a realistic LL process. This also helped support buy-in with the team since it became clear to them in Phase 1 that providing such a limited set of data was not going to be as useful.

3.4 Observed Results – Phase 0

While Phase 0 was not a phase that required any unique planning, it provides a baseline to compare against.

For the month before Phase 1, incident data was examined to support the initial statements that LL data is unlikely to be captured without a convenient approach and clear policy requirements. In the case of the business in this study, the policy indicated that it was required to capture LL data for Severity 1 incidents and that capturing LL data for lower-severity incidents was optional.

3.4.1 Issues in Phase 0

The following issues were observed in Phase 0:

- There was no easy way to track the cases via simple query into the case management system (known issue).
- Little data was captured overall since the policy only mandated a dedicated LL process for Severity 1 incidents, and there was only one in this period.

In that context, the following Phase 0 Data content should come as little surprise. There were 68 incidents for this month. Only one of them had a dedicated Lessons Learned review related to a Severity 1 incident.

3.5 Observed Results – Phase 1

The second month for which data was observed and compared, and the first month for which LL tagging was enabled, is when the idea of "simplified" LL data capture was implemented.

3.5.1 Issues in Phase 1

The following issues were identified in Phase 1:

- Using JSON-formatted data in the case management system resulted in reporting challenges when any syntax errors were made, resulting in the requirement for additional, time-consuming manual review.
- The simplified tagging options (as expected) limited options for any detailed analysis beyond LL counts per incident.

At the beginning of Phase 1, the team reviewed the policy. Expectations were set regarding reporting, namely an end-of-month review of LL data.

At the end of the first month, additional time was required to do normal cleanup work for closed cases, which included making sure that LL data was added when applicable.

The number of incidents handled by the team in this month was fewer than expected when compared to the recent case volumes as well as the same timeframe the year before. Regardless, of the 50 incidents handled during this timeframe, 14 of them had associated LL data.

3.5.2 Results in Phase 1

Phase 1 was focused on meeting "compliance" requirements by only capturing LL data at the end of the incident and only reporting LL counts per incident.

Total Incidents	46
Incidents with LL data	14
Incidents with 1 LL only	11
Incidents with 2 LL	3
Total LL Count	21

Table 5: Phase 1 incident and LL counts

As these lessons were (by design) not broken down into any sub-categories, only general analysis can be done with this set of data. The intention of the incident handler can be inferred from some of the text in the LL statements, but these conclusions cannot be relied upon when performing analysis over time.

No work items were created specifically tied to any of these lessons. Since there was no way to link work items to incidents through either system or any way to capture a work item directly via tagging, there was no easy way to report on the work item information.

3.6 Observed Results – Phase 2

Phase 2 was equal in duration to Phase 1, approximately one month, and included the new expectations for enhanced tagging and real-time LL data capture.

3.6.1 Issues in Phase 2

Increasing the complexity (the number of required tags) in the JSON-formatted data in the case management system resulted in the ongoing reporting challenges when any syntax errors were made, resulting in the continued requirement for additional, time-consuming manual review.

3.6.2 Results in Phase 2

For purposes of comparison with Phases 0 and 1, Table 6 shows an apples-to-apples view of the shared data points.

Measure	Phase 0	Phase 1	Phase 2
Total Incidents	68	46	61
Incidents with LL data	1	14	9
Incidents with 1 LL only	1	11	5
Incidents with 2 LL+	0	3	4
Total LL Count	1	21	14

Table 6: Incident data compared between Phases 0, 1 and 2

Due to the increased detail captured with the Phase 2 tagging options in "LLSub-Type," Table 7 shows the additional details available:

Incidents with work items	8
Not started	7
Started	1
Complete	0
Incidents by type:	
Went well	3
Start doing something different	11
Stop doing something	0
Incidents by sub-type	
Playbook update (change existing)	4
New playbook entry	2
Training needed (IR team)	4
Better reporting needed	0
Partner communication issue	0
Other partner issue	0
NA (for "Keep doing" or for "NA" type)	3
Internal tooling issue	1
External tooling issue	0
Other dependency	0
Lessons with tracked work items captured	8

Table 7: Phase 2 Incident LL Data

Andrew Baze, abaze@outlook.com

This metadata allowed the team to make the following observations about the Phase 2 Lessons Learned.

Positive lessons (three of 14 lessons in this month) were captured (the "Went well" category), which enabled reinforcement of good behavior.

No incidents fell into the "Stop doing something" category. This could indicate that 1) nothing needed to stop or that 2) something needed to stop, but the category was misunderstood, and all such lessons were inappropriately categorized as "Start doing something different." This requires further investigation.

All incidents in the "Start doing something different" (11 of 14) category fell into the "Playbook update," "New playbook entry," "Training needed," "NA," or "Internal tooling issue" categories.

No lessons were identified for "Better reporting needed," "Partner communication issue," "Other partner issue," "External tooling issue," or "Other dependency" categories. However, anecdotal discussion with the team suggested that these categories are still valid and that lessons will likely be assigned such values in the future.

Visible for the first time in the work tracking system, work items related to some lessons were created, assigned owners, and tracked. During this relatively early stage in the introduction of these kinds of work items and due to the desire to treat them with flexibility, the Kanban format was used (vs. the other more "planned" other work the team had on the docket). Work items were not mandated for each LL entry. At the time this document was written, eight tasks had been captured. Of those eight, two were completed, one was marked "Active" (in progress) with an assigned owner and five were "New" (not started) with no owners assigned.

3.7 Review of Findings

Three months (which comprised of a baseline comparison month, a month of adjustment to improved but limited data tracking, and a month of expanded metadata tracking) is an insufficient amount of time to conduct a thorough evaluation of reasons for either data quality or quantity changes. Such changes were observed regardless, and

opportunities to evaluate the causes are explored in more detail in "Implications for Future Research."

Within the short duration of this study, providing the option for real-time LL data capture did not cause an increase in the volume of lessons. The number of lessons captured went from 21 in Phase 1 to 14 in Phase 2. Given the multiple, uncontrolled variables (see "Implications for Future Research" in this paper), this decrease is not an indictment of real-time capture, tagging requirements, or other policy changes.

As noted in Phase 0, not setting clear expectations for capturing LL data on all relevant cases resulted in not capturing LL data for any cases other than those mandated by policy.

Mandatory tagging with appropriate (attempted in Phase 2) metadata fields significantly increased the ability to analyze the LL data. The same cannot be said for the limited (Phase 1) fields, which added little value beyond the ability to report LL count data (given the constraints of the case management tool that otherwise did not facilitate such reporting).

Note that there was no way to capture whether a lesson was captured during or at the end of an incident, though anecdotal evidence suggests some lessons were captured in real-time. Adding tagging or otherwise querying case management historical updates to determine how this may have taken place was not feasible.

The ability to tag "things that went well" added the expected value. While the tagging itself cannot be credited for things going well or not, the ability to specifically capture such cases gave management the opportunity to reinforce expected behavior and praise team members in LL reviews.

While there was insufficient time in the case study to observe trends of LL data after Phase 2 tagging was implemented, a simple trend was observed for the captured Phase 2 work items, as shown in Figure 3 below.

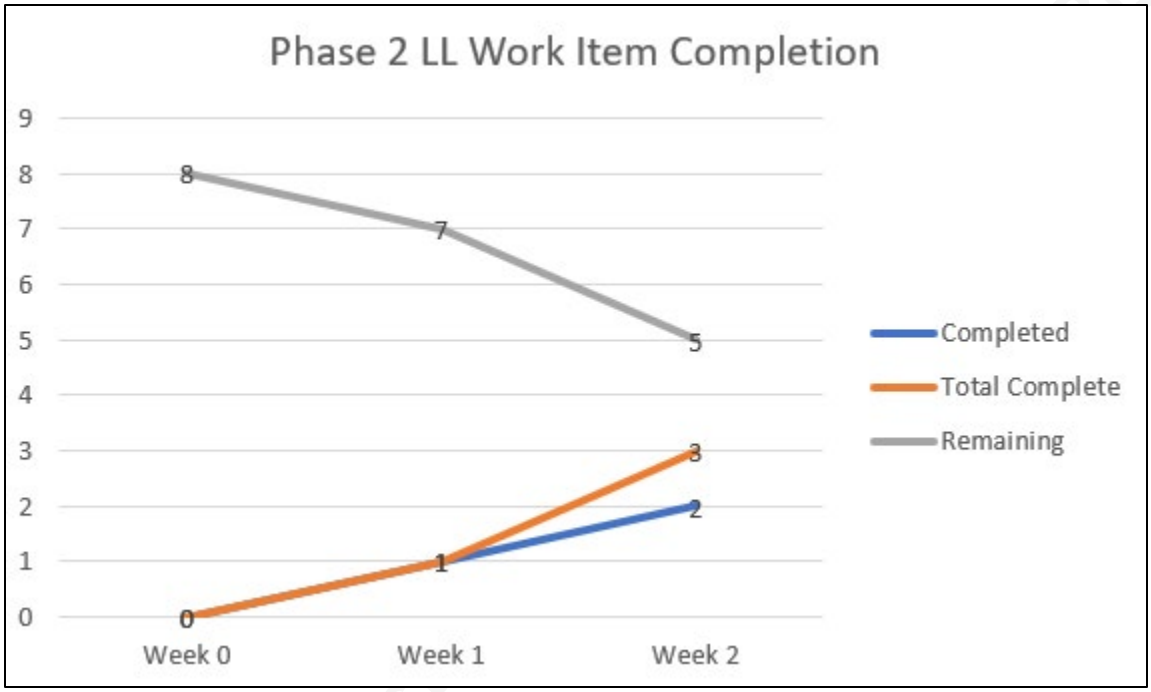

Figure 3: Phase 2 work items completed overall and per week

Simply tracking open and completed work items is a noteworthy improvement, especially for a team that didn't previously track any LL data in a way that could be pulled and reported programmatically.

Cross-referencing this data with the various Lessons Learned sub-types over longer periods of time should provide a useful view into the type of work dedicated to each type of problem.

4. Recommendations

The key learnings from this study are reflected in the following recommendations:

1) **Treat the Lessons Learned process as an absolute requirement**. The capture of LL data must be required, regardless of whether the caseload is high. Even a couple of minutes of LL data entry can result in a rich, useful dataset over time that can be used to support a variety of critical efforts. It does not necessarily require a time-consuming meeting or special reports.

Andrew Baze, abaze@outlook.com

2) **Make it easy to enter data.** Ensure that data can be captured, tagged, and analyzed with little or no significant cost to the incident handlers (beyond typing the lessons and selecting some key metadata). Onerous formatting will hamper data capture and reporting efforts.
3) **Review the data regularly.** Continuous (e.g., bi-weekly or monthly, depending on an organization's reporting rhythm) review of LL data will set a clear expectation for ongoing LL data capture, build organizational muscle around capturing such data and support the creation and growth of a valuable dataset.
4) **Put the data to work.** While this study did not have time to evaluate the longer-term analysis and use of the overall LL dataset, the completion rates of associated work items, and potential cultural changes (e.g., a possible improvement in security awareness among served customers), putting this data to work productively is the whole reason for the process. Aside from obvious improvements to the Incident Response team's processes and tools, broader impacts should be expected, tracked, evaluated, and fed back into the overall process. Examining trends identified in the LL metadata as well as trends within the work item data will also help facilitate a double-loop learning approach as described in Appendix A: Double-Loop Learning.

4.1 Implications for Future Research

In this study, given its time and other constraints, many variables could not be practically isolated or controlled. Since multiple independent variables were simultaneously implemented (e.g., capturing data real-time and using expanded metadata tagging), clearly attributing outcomes to those variables was not possible. It is also difficult to control other variables that play a key role in such a dynamic environment, such as the types and frequency of cybersecurity incidents over any given time. Therefore, dedicating far more time to such a study, combined with implementing one variable process improvement at a time, could result in the ability to measure potential causes and effects more closely.

Other variations could be explored to tease out differences that result from changes in management expectations regarding data volume and quality, incident phase, and frequency of reporting. Pre-planned communications, schemas and reporting dashboards for all expected phases could speed up the process by reducing the time to plan or re-plan between phases.

Depending on the organization and its unique challenges, the Phase 2 metadata tagging schema could also be adjusted (at least for the sub-type categories) to focus on the most relevant options, potentially simplifying the number of choices an incident handler needs to make when categorizing lessons and potentially reducing the data capture burden.

Another area where more time and energy could be spent is in interviewing team members during each phase of the process to determine their attitudes and perceptions of effectiveness, ease of use, and time spent during the LL process and the tracking of associated work items. A similar approach could be taken with the key stakeholders consuming the LL data to determine the perceived value of that data and how it is used by customers, partners, and others.

5. Conclusion

As demonstrated, while making relatively rapid changes across multiple variables in a dynamic cybersecurity incident response environment to improve data quantity and quality has many challenges, the outcomes of improved process and higher quality data should serve any cybersecurity team well.

An organization's cybersecurity Incident Response function is often directly involved with understanding and mitigating the worst of the exploited vulnerabilities affecting business. The value of all data generated by Incident Response teams is valuable, and Lessons Learned is one of the most valuable datasets that this team can generate. Supporting continuous maturity improvements and better overall analysis will help enable prioritized customer education on critical security risk areas and are key to an incident response team's success. No organization can afford to ignore this data.

Andrew Baze, abaze@outlook.com

Ongoing analysis can and must take place regardless of the case management toolset abilities, a frequently overwhelming caseload, or the current maturity (or lack thereof) of a team's processes. Since Lessons Learned data can (as demonstrated) be captured with a relatively small effort, it is imperative that management set and enforce data capture expectations, then support the regular reporting, analysis, and productive use of that data.

This study demonstrates how simple policy changes, minor technical adaptations, changed management expectations regarding data capture, reporting, and associated work item tracking can result in a significantly increased ability to review LL data and to track work that represents important potential improvements to incident response tools and processes.

Andrew Baze, abaze@outlook.com

References

Anson, S. (2020). *Applied Incident Response.* Hoboken: Wiley.

Argyris, C. (1977, September). Double Loop Learning in Organizations. *Harvard Business Review*.

Capability Maturity Model Integration. (2020, December 7). Retrieved from Wikipedia.org: https://en.wikipedia.org/wiki/Capability_Maturity_Model_Integration

Chrissis, M. B., Konrad, M., & Shrum, S. (2007). *CMMI : guidelines for process integration and product improvement.* Boston: Pearson Education, Inc.

Cichonski, P. R., Millar, T., Grance, T., & Scarfone, K. A. (2004). *Computer Security Incident Handling Guide*. Retrieved 11 3, 2020, from https://csrc.nist.gov/publications/detail/sp/800-61/rev-2/final

Dempsey, K. L., Witte, G. A., & Rike, D. (2014). *Summary of NIST SP 800-53, Revision 4: Security and Privacy Controls for Federal Information Systems and Organizations*. Retrieved 11 5, 2020, from https://csrc.nist.gov/publications/detail/white-paper/2014/02/19/summary-of-nist-sp-800-53-rev-4-security--privacy-controls/final

Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard - Measures That Drive Performance. *Harvard Business Review*.

Keller, N. (n.d.). *Framework for Improving Critical Infrastructure Cybersecurity*. Retrieved 11 5, 2020, from National Institute of Standards and Technology: https://www.nist.gov/cyberframework

Mitropoulos, N. (2020). *GCIH GIAC Certified Incident Handler All-in-One Exam Guide.* New York: McGraw-Hill.

Skoudis, E., & Strand, J. (2018). *Incident Handling Step-by-Step and Computer Crime Investigation.* North Bethesda: SANS Institute.

Thompson, E. C. (2018). *Cybersecurity Incident Response: How to Contain, Eradicate, and Recover from Incidents.* Berkeley: Apress.

Andrew Baze, abaze@outlook.com

Appendix A: Double-Loop Learning

Incident data needs to be captured and analyzed to mature the organization beyond a purely reactive model. Problems then need to be identified and fixed. However, since an incident response team is usually neck-deep in security incidents (another kind of problem), it is often easy to miss the opportunity to capture problem metadata that would enable more in-depth analysis. One way to compare the reactive and proactive learning approach is to illustrate the approaches through the lens of organizational learning. The simple definition of organizational learning used in this case is "a process of detecting and corrective error" (Argyris, 1977).

Understanding single-loop learning is important context for understanding double-loop learning. Single-loop learning is generally defined as a repeated attempt to solve a problem without changing the approach and without questioning the overall goal.

The venerable Harvard Business Review example of single-loop learning applies just as well today as it did in 1977: "Single-loop learning can be compared with a thermostat that learns when it is too hot or too cold and turns the heat on or off. The thermostat is able to perform this task because it can receive information (the temperature of the room) and therefore take corrective action" (Argyris, 1977).

Compare this to an incident response team working with customers who are continuously plagued by cross-site scripting attacks. The incident response team learns of the vulnerability, engages the customer, supports the fix, and closes the incident. The data (vulnerability) was acted upon until the state changed to "non-vulnerable." This repeats for each report of that type of vulnerability.

Another example of a single-loop learning approach to solving problems related to out-of-date playbook entries is to adjust the playbook whenever an inaccuracy was noted. While this is by no means inherently deficient or otherwise negative, it is a simple example.

Andrew Baze, abaze@outlook.com

Why is this approach insufficient? At first glance, it appears that important problems (i.e., inaccuracies in the playbook or vulnerability remediation) are being solved. And some simple problems may be solved with a single-loop approach. However, many problems have layers, and if a root cause is not determined, then problems are likely to persist even if some symptoms are addressed. An organization must distinguish between these symptoms and an underlying root cause.

Double-loop learning adds the additional "loop" of questioning the approach of the problem and the goal of the solution. This approach is much more likely to surface root-cause issues and enable more profound improvements that lead to organizational maturity improvements.

Using the earlier thermostat example in this context, "if the thermostat could question itself about whether it should be set at 68 degrees, it would be capable of not only detecting error but of questioning the underlying policies and goals as well as its own program" (Argyris, 1977).

In the case of the cross-site scripting vulnerabilities plaguing the customer organization, additional questions could be asked, such as:

- Why do these problems appear in this organization in the first place?
- What training do web developers receive regarding sanitizing data and only using approved libraries?
- How do we set and enforce policy for secure coding practices?

In the case of the "problem" of a playbook containing inaccurate data, a double-loop learning approach would include asking such questions as "Why did these inaccuracies exist?", "Would a different playbook update help reduce future inaccuracies?" or "What can we do to prevent inaccuracies in the first place?"

In both cases, it is the questioning of the problem itself, as part of root cause analysis, that allows the incident response team to identify work items that will enable fundamental improvements and increase organizational maturity.

How does this model factor into the approach described in this case study? The recommended approach focuses on the ability to analyze the lessons learned in aggregate,

which enables the organization to ask better questions. And while the three-phased approach described in this paper does not provide enough time and data to perform a long-term trend analysis, a continuation of Phase 3 (with its additional metadata capture) will enable that view and its corresponding double-loop learning approach.

Andrew Baze, abaze@outlook.com

Appendix B: Case Study Planning

The following basic planning approach was used to determine the timeline for the project.

Stakeholders in both phases primarily included the incident handlers in the organization and the management chain. In the context of RACI (Responsible, Accountable, Consulted and Informed), the incident response team manager was accountable for setting policy, ensuring that the JSON tagging and associated reporting mechanisms functioned as expected, that the data was reviewed in the expected timeframe, and that adjustments to the process were communicated clearly. The incident response team members were responsible for adding incident metadata as specified, including adding more data in the Phase 2 approach. The incident response team was consulted on the overall plans, the tagging options, the JSON format, and the perceived effectiveness of the various phases of the project. Management stakeholders were informed of the results.

Month 1 – Phase 1

For Phase 1, the following dependencies were required to proceed:

- Specification of JSON tagging schema;
- Verification that JSON data could be extracted and reported;
- A reporting dashboard;
- Policy regarding use of tagging for LL data communicated to team members;
- Phase 1 data review booked as part of the normal business rhythm.

Phase 1 data capture was expected to last for approximately 30 days.

Month 2 – Phase 2

For Phase 2, the following changes and additional dependencies were required to proceed:

- Expanded JSON tagging schema specified;

- Updated reporting dashboard to show expanded dataset;
- Updated policy communicated to team members;
- Phase 2 ongoing data review scheduled.

Phase 2 data capture was also expected to last for approximately 30 days. For both phases, the intention was to collect a roughly approximate set of data to facilitate the comparison.

Month 3 – Review and Analysis

The focus for the third month was intended to be the analysis of Phase 2 data, followed by the comparison of Phase 0, 1, and 2 data. Final documentation of the process, data analysis, and review with the team and management were intended to occur toward the end of the third month.

The three monthly plans took place as expected during the study.

Appendix C: Checklist for Implementing Improved Lessons Learned Data Capture and Reporting

The following checklist provides a high-level guide for enabling an incident response team to improve their Lessons Learned capability. This includes considerations for the LL data capture policy, the mechanisms by which LL data is captured given the current tools in use by the team, and the appropriate reporting of incident and work item data.

Implementing Improved LL Data Capture and Reporting	
Implementation Step	**Considerations and Key Questions**
Consider evaluating the team's maturity level using the CMMI model.	One introductory reference: https://uccs.edu/Documents/tboult/cmmi-overview05.pdf
	After going through the rest of this template and allowing for time to analyze data over several months, re-evaluate the team's maturity level and report any maturity improvements to management.
Assess current LL data capture and use policy.	Is capturing LL data currently required?
	If not, what changes are needed to make this required?
Assess the expected versus actual results of the policy.	Is the current policy being followed?
	If not, how can that be addressed?
Draft new or updated policy statements.	Would an updated policy be followed? Consider issues present in previous policy and make sure those are addressed in the updates.
Conduct a meeting to review updated policy statements.	Make updates as needed based on stakeholder feedback.
Conduct a recurring meeting to review the captured LL data.	Consider running this at least monthly, depending on case volume.

Andrew Baze, abaze@outlook.com

Assess software and tools used to capture LL data and associated work items.	Does current LL data capture also enable reporting?
	Can LL sub-types (or other important metadata) and work item links be captured?
	If not, what can be done to facilitate this?
	What other stakeholders may need to be engaged to make tooling updates?
Make tooling or other configuration updates as needed.	Coordinate with any engineering or operations stakeholders to help with updates.
	Ensure that LL data entry options are easy to use and that formatting issues will not prevent accurate reporting.
Conduct a meeting to review updates and train the team as needed.	This may take place simultaneously with the policy review meeting, depending on when tooling updates are made.
Report new LL data.	Work with reporting stakeholders to ensure that LL data is reported as expected.
	Work with work item tracking stakeholders to ensure all work items tied to LL data are being reported.
	If possible, provide a reporting view that cross-references LL data and associated work items, even if work items are tracked in a different system from incidents.
	Analyze LL and work item data and trends, then share as appropriate with management and other key stakeholders.

Table 8: Checklist for Implementing Improved Lessons Learned Data Capture and Reporting

The above checklist may be adjusted as needed to support any team's unique requirements, capabilities and environment. Also, since these improvements should part of an ongoing program (versus a single, time-bound project), some of the steps in the checklist should be repeated as policy, processes and tools are updated over time.

Andrew Baze, abaze@outlook.com