Army Software Sustainment: Righting Our Cost Estimating Assumptions

Cheryl Jones
Presented to ICEAA Workshop
2021
Topics

• Purpose of this presentation
• Initiative details
• Data demographics
• Time phased analysis results
• Conclusions
Purpose

• A challenge in estimating software Total Ownership Costs at the beginning of a software acquisition is estimating the cost of software after development.
• Rules-of-thumb are often used to estimate the annual change in the software during the sustainment lifecycle.
• This presentation will present data analysis on software sustainment profiles that show a rise and fall in the number of changes over long periods of time contrary to the practice of using a constant annual change.
SWS Initiative Objective and Strategy

Accurately estimate Army system Software Sustainment (SWS) costs to:
- Effectively estimate and justify software and system life cycle costs
- Objectively evaluate Army system software sustainment execution costs
- Inform and optimize the allocation of available sustainment resources across the Army

Collect and evaluate SWS cost and technical data for all Army operational systems (Phase I, II, & III data calls)
Generate and validate cost estimating relationships from data collection
Implement systemic Army SWS data collection via the SRDR-M: Populate cost and technical data repository
Improve Army SWS policy, business, and technical requirements

Effective software sustainment cost estimation is the basis for Army system software life cycle cost management
Summary of Accomplishments

• Established Software Sustainment Data Collection Mechanisms
  – Army Software Data Collection Questionnaire
  – Software Sustainment WBS Used to Collect Sustainment Costs
  – Annual Data Collection

• Created Comprehensive Software Sustainment Data Repository
  – 214 Unique Systems (including DBS systems)
  – 855 Capability Releases
  – 676 IAVA Releases
  – 3,840 records on software license data

• Established Robust Foundation for Software Sustainment Fact-Based Decisions
  – Allocations of Costs by WBS Elements
  – Continue to improve Software Sustainment Cost Estimating Relationships

• Data and Analysis Results provided to DoD Community
**System Demographics**

Number of Systems by Commodity

<table>
<thead>
<tr>
<th>Category</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
<th>New Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aviation</td>
<td>26</td>
<td>23</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Intel</td>
<td>11</td>
<td>15</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Fires</td>
<td>25</td>
<td>17</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>C5ISR</td>
<td>18</td>
<td>14</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Missiles</td>
<td>25</td>
<td>15</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Network</td>
<td>17</td>
<td>14</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>SATCOM</td>
<td>18</td>
<td>15</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Vehicles</td>
<td>14</td>
<td>12</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Businesss</td>
<td>12</td>
<td>12</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Simulation</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Comms</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>MissionCMD</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>ChemBio</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Space</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Unique Systems: 208
Phase 1 Systems: 41
Phase 2 Systems: 151
Phase 3 Systems: 115
New Systems in Phase 3: 22
94 systems in the previous database that are not in Phase 3

Data Fields: 2,107

Systems by ACAT Level

<table>
<thead>
<tr>
<th>ACAT Level</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACAT I</td>
<td>52</td>
<td>33</td>
<td>88</td>
</tr>
<tr>
<td>ACAT II</td>
<td>33</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>ACAT III</td>
<td>88</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>ACAT IV</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>ACAT UKN</td>
<td>25</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>Non-POR</td>
<td>9</td>
<td>9</td>
<td>6</td>
</tr>
</tbody>
</table>

Distribution of System Age

<table>
<thead>
<tr>
<th>Age Range</th>
<th>PPSS</th>
<th>PDSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 10 Years</td>
<td>78</td>
<td>53</td>
</tr>
<tr>
<td>11 - 20 Years</td>
<td>47</td>
<td>20</td>
</tr>
<tr>
<td>21 - 30 Years</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>31 - 50 Years</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

Super Domains

<table>
<thead>
<tr>
<th>Domain</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td>99%</td>
</tr>
<tr>
<td>ENG</td>
<td>53%</td>
</tr>
<tr>
<td>AIS</td>
<td>40%</td>
</tr>
<tr>
<td>SUP</td>
<td>16%</td>
</tr>
</tbody>
</table>

16.2%
Software Releases

Average Capability Release Costs
- RT: $1,022,176
- ENG: $637,762
- AIS: $361,823
- SUP: $445,595

Average IAVA Release Costs
- RT: $73,603
- ENG: $92,518
- AIS: $150,430
- SUP: $1,433

Releases by Super Domain
- RT: 384
- ENG: 210
- AIS: 181
- SUP: 51

Releases by Size Measure
- SCs: 637
- SLOC: 200
- IAVAs: 185
- Requirements: 135
- Other: 22
- Agile Story Pts: 18
- RICE: 17

Change Type
- All Releases: 1,478
- Capability: 802
- IAVA: 676
- Maintenance: 361
- Hybrid: 224
- Enhance: 174
- Cyber: 28
- Other: 15

Average Software Changes
- RT: 84
- ENG: 69
- AIS: 47
- SUP: 131

Presented for the ICEAA 2021 Online Workshop - www.iceaaonline.com
Software Licenses

Licenses by Commodity
- Aviation: 682
- Intel: 621
- Network: 489
- Simulation: 332
- SATCOM: 298
- Fires: 285
- C5ISR: 272
- Missiles: 210
- Business: 145
- Vehicles: 108
- MissionCMD: 91
- ChemBio: 26
- Comms: 20
- Test: 9

Average License Cost by Super Domain
- RT: $32,565/PSS, $15,754/PDSS
- ENG: $47,637/PSS, $142,599/PDSS
- AIS: $35,568/PSS, $128,198/PDSS
- SUP: $23,984/PSS, $35,199/PDSS

Number of Licenses by Super Domain
- RT: 499/PSS, 1,085/PDSS
- ENG: 127/PSS, 624/PDSS
- AIS: 442/PSS, 469/PDSS
- SUP: 320/PSS, 22/PDSS
Previous Data Analysis

• Preparing data for analysis
  - Data collection
  - Data repository
  - Data normalization
  - Data imputation (backfill, online searches, averages)
  - Analysis
    o Outlier labeling rule
    o Histograms
    o Boxplots
    o Scatter plots
      - Loess smoothing
      - Non-linear Trend lines
    o Log Ordinary Least Squares regression
    o Non-linear regression
    o Regression trees
    o K-nearest neighbor

• Results
  - Cost distributions
  - Sensitivity analysis
  - Causal analysis
  - Cost Estimating Relationships (CER) using different segmenting strategies

• Current Analysis Focus
  - Changing release cost base over time in sustainment
Notional Software Sustainment Release Profile

Maintenance: normal repairs, continuously ongoing
Extra Bug Fixes: due to a major release
Updates: obsolescence, new functionality
Major Upgrades: architectural changes, new hardware, major functional changes
Analysis Data Selection

- Capability Release data records from data collection phases 1, 2, and 3.
- Only systems in the O&S phases were studied: 987 records
- Removed records with:
  - DBS: 16
  - No O&S phase start date: 39
  - No Release end data: 63
  - No Release Hours: 262
  - No SW Change (SC) counts: 452
- Records remaining: 345
- Time in the O&S Phase is defined as:
  2019– (Phase start date)
- Both All-data and Trimmed-data are presented
  - Data was trimmed using an Outlier Labeling Rule
Outlier Labeling Rule

- In Exploratory Data Analysis [1], Tukey includes a rule of thumb for labeling (as “outside”) observations that are extreme enough to be potential outliers. In terms of the lower fourth, $F_L$, and the upper fourth, $F_U$ (approximate quartiles of the sample), an observation is “outside” if it does not fall in the interval between the cutoffs:

  $$ F_L - k(F_U - F_L) \quad \text{and} \quad F_U + k(F_U - F_L) $$

  (Note: $F_U - F_L$ is often referred to as the Inter-Quartile Range – IQR)

  - Tukey’s $k$ is 1.5

- A subsequent study [2,3] shows that
  - For a low $n$ of 5, $k$ is above 3.0.
  - For an $n$ above 10, the median value for $k$ is 2.1 for 90% and 2.3 for 95% probabilities that no observations are beyond the cutoffs.
Software Changes over Time

**SC w/Outliers**

- Count: 329
- Median: 17.0
- Min: 1.0
- Max: 4,079.0
- 1st Quartile: 6.0
- 3rd Quartile: 42.0
- IQR: 36.0
- K: 2.1
- Outlier-L: -69.6
- Outlier-H: 117.6

**SC w/o Outliers**

- Count: 298
- Median: 14.0
- Min: 1.0
- Max: 99.0
- 1st Quartile: 5.0
- 3rd Quartile: 33.8
- IQR: 28.8

Presented for the ICEAA 2021 Online Workshop - www.iceaaonline.com
Release Hours over Time

**Hours w/Outliers**
- Count: 329
- Median: 2,940.0
- Min: 2.0
- Max: 731,605.0
- 1st Quartile: 1,225.4
- 3rd Quartile: 10,560.0
- IQR: 9,334.6
- K: 2.1
- Outlier-L: -18,377.2
- Outlier-H: 30,162.6

**Hours w/o Outliers**
- Count: 295
- Median: 2,614.0
- Min: 2.0
- Max: 99,298.0
- 1st Quartile: 1,094.4
- 3rd Quartile: 7,606.2
- IQR: 6,511.8
Hours / Software Change over Time

**Hrs/SC w/Outliers**
- Count: 329
- Median: 225.9
- Min: 2.0
- Max: 22,242.1
- 1st Quartile: 60.8
- 3rd Quartile: 756.6
- IQR: 695.8
- Outlier-L: -1,400.4
- Outlier-H: 2,217.8

**Hrs/SC w/o Outliers**
- Count: 299
- Median: 181.6
- Min: 2.0
- Max: 1,910.0
- 1st Quartile: 56.3
- 3rd Quartile: 521.6
- IQR: 465.3
Hours / Software Change over Time RT Domain

**Hrs/SC w/Outliers**
- Count: 103
- Median: 458.1
- Min: 4.0
- Max: 9,768.0
- 1st Quartile: 108.6
- 3rd Quartile: 1,404.8
- IQR: 1,296.2
- K: 2.1
- Outlier-L: -2,613.4
- Outlier-H: 4,126.8

**Hrs/SC w/o Outliers**
- Count: 95
- Median: 407.4
- Min: 4.0
- Max: 3,926.5
- 1st Quartile: 78.8
- 3rd Quartile: 1,130.6
- IQR: 1,051.8
Implications for Software Cost

Unit cost changes over time
• Indicates that different CERs are needed at different times in the O&S phase
DevSecOps
• If everything in DevSecOps could be considered a Software Change (Enhancements & Repairs), at what level do you staff your Software Factory?

![Graph showing hours per software change over years in O&S phase]
Conclusions

• The notional concept of the sustainment release profile appears to be valid although the rise and fall of sustainment action is different from the notional concept as shown by the data.

• In the past, sustainment CERs were developed assuming sustainment actions were relatively constant over time, i.e., annual change traffic, but this appears to not be the case.

• Budgeting approaches based on improved estimating methods need to be developed and applied
  - Otherwise, Program sustainment budgets will not reflect the actual work needed to be performed, which can lead to detrimental effects for the warfighter.

• The implications of variable software changes during sustainment must be explored further for the DevSecOps environment.
References

1. Tukey, J.S. Exploratory Data Analysis, Addison-Wesley, Reading MA, 1977


## Contributors

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheryl Jones</td>
<td>Measurement Analyst</td>
<td>US Army AFC, 973-724-2644</td>
</tr>
<tr>
<td>James Judy</td>
<td>NISEC Division Chief</td>
<td>ODASA-CE, 703-697-1612</td>
</tr>
<tr>
<td>Paul Janusz</td>
<td>Measurement Analyst</td>
<td>US Army AFC, 973-724-4849</td>
</tr>
<tr>
<td>Dr. Bradford Clark</td>
<td>Vice President</td>
<td>Software Metrics, Inc.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(703) 754-0115</td>
</tr>
<tr>
<td>Dr. Robert Charette</td>
<td>President</td>
<td>ITABHI Corporation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(540) 972-8150</td>
</tr>
<tr>
<td>Kimberly Roye</td>
<td>Senior Analyst</td>
<td>Galorath Federal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(703) 966-3192</td>
</tr>
</tbody>
</table>
Backup
## Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACAT</td>
<td>Acquisition Category</td>
</tr>
<tr>
<td>AIS</td>
<td>Automated Information Systems Super Domain</td>
</tr>
<tr>
<td>ENG</td>
<td>Engineering Super Domain</td>
</tr>
<tr>
<td>IAVA</td>
<td>Information Assurance Vulnerability Alert</td>
</tr>
<tr>
<td>ICEAA</td>
<td>International Cost Estimating and Analysis Association</td>
</tr>
<tr>
<td>IQR</td>
<td>Inter-Quartile Range</td>
</tr>
<tr>
<td>O&amp;S</td>
<td>Operations and Sustainment</td>
</tr>
<tr>
<td>PDSS</td>
<td>Post-Deployment Software Support (PM management, development and/or production funding)</td>
</tr>
<tr>
<td>PPSS</td>
<td>Post-Production Software Support (LCMC management, OMA funded)</td>
</tr>
<tr>
<td>RT</td>
<td>Real-Time Super Domain</td>
</tr>
<tr>
<td>SRDR-M</td>
<td>Software Resources Data Report for Maintenance</td>
</tr>
<tr>
<td>SUP</td>
<td>Support Super Domain</td>
</tr>
<tr>
<td>SWS</td>
<td>Software Sustainment</td>
</tr>
<tr>
<td>THrs</td>
<td>Total Release Hours</td>
</tr>
<tr>
<td>TSC</td>
<td>Total Software Changes</td>
</tr>
<tr>
<td>WBS</td>
<td>Work Breakdown Structure</td>
</tr>
</tbody>
</table>