Can You Trust Your Data? Measurement and Analysis Infrastructure Diagnosis

October 2008
David Zubrow
SEI
Can You Trust Your Data? Measurement and Analysis Infrastructure Diagnosis

Carnegie Mellon University, Software Engineering Institute (SEI), Pittsburgh, PA, 15213

Approved for public release; distribution unlimited

Security classification of:
- Report: unclassified
- Abstract: unclassified
- This page: unclassified

Limitation of abstract: Same as report (SAR)

Number of pages: 39

Name of responsible person: unclassified
Dave Zubrow is Manager of the Software Engineering Measurement and Analysis (SEMA) initiative within the Software Engineering Institute (SEI). Prior to joining the SEI, Dave served as Assistant Director of Analytic Studies for Carnegie Mellon University. He is a SEI certified instructor and appraiser, member of several editorial boards of professional journals, and active in standards development. Dave is a senior member of the American Society for Quality. Dave earned his PhD in Social and Decision Sciences and an MS in Public Policy and Management from Carnegie Mellon University.
The benefit and value of measurement comes from the decisions and actions taken in response to analysis of the data, not from the collection of the data.
Measurement and Analysis in Action

1. Data collection
2. Data stored
3. Data analyzed, interpreted, & stored
4. Data & interpretations reported
5. Decision-making
6. Corrective actions to improve process

- Measurements collected
- Data repository
- Corrective actions
- Decision-making
- Data collection
- Data stored
- Data analyzed, interpreted, & stored
- Data & interpretations reported
- Measurement Report

David Zubrow, October 2008
© 2008 Carnegie Mellon University
Information Needs

Information Product

Interpretation

Indicator

Analysis Model

Derived measure

Derived measure

Measurement Function

Base measure

Measurement Method

Entity

Attribute

Analysis & Reporting

Data Collection

Measurement Concepts

Adapted from ISO 15939 Information Model
Polling Question

To what extent does your organization take steps to ensure it is getting value from its project data?

- Not at all
- Somewhat
- A great deal
Outline

The Need for a Measurement and Analysis Infrastructure Diagnostic (MAID)
  • Measurement errors and their impact

MAID Methods
  • Process Diagnosis
  • Data and Information Product Quality Evaluation
  • Stakeholder Feedback

Summary and Conclusion
Where do Measurement Errors come From

Data Entry Errors

• Manual data entry

• Lack of integrity checks

Differing Operational Definitions

• Project duration, defect severity or type, LOC definition, milestone completion

Not a priority for those generating or collecting data

• Complete the effort time sheet at the end of the month

• Inaccurate measurement at the source

Double Duty

• Effort data collection is for Accounting not Project Management.
  — Overtime is not tracked
  — Effort is tracked only to highest level of WBS
Where do Measurement Errors come From?

Dysfunctional Incentives
- Rewards for high productivity measured as LoC/Hr
- Dilbert-esque scenarios

Failure to provide resources and training
- Assume data collectors all understand goals and purpose
- Arduous manual tasks instead of automation

Lack of priority or interest
- No visible use or consequences associated with poor data collection or measurement
- No sustained management sponsorship

Missing data is reported as a valid value
- Can’t distinguish 0 from missing when performing calculations
Purpose for Measuring is Understood

Source: CMU/SEI-2006-TR-009
Are Documented Processes Used?

- Frequently: 876 responses (47.3%)
- Occasionally: 559 responses (30.2%)
- Rarely: 269 responses (14.5%)
- Never: 87 responses (4.7%)
- I don't know: 18 responses (1.0%)
- N/A: 43 responses (2.3%)

Source: CMU/SEI-2006-TR-009
What is Measurement Error?

Single Value: Deviation from the “true” value

- Distance is 1 mile, but your odometer measures it as 1.1 miles
- Effort really expended on a task is 2.75 hours, but it is recorded as 3

Data Set: Error introduced as a result of the measurement process used

- Not as defined, but as practiced
Gold Standard: Accuracy and Precision

(σ) Accurate but not precise
(μ) Precise but not accurate
Both accurate and precise
Cost of Poor Data Quality to an Enterprise – Typical Issues and Impacts

**Typical Issues**

- Inaccurate data [1-5% of data fields are erred]
- Inconsistencies across databases
- Unavailable data necessary for certain operations or decisions

**Typical Impacts**

<table>
<thead>
<tr>
<th>Operational</th>
<th>Tactical</th>
<th>Strategic</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Lowered customer satisfaction</td>
<td>• Poorer decision making &amp; decisions take longer</td>
<td>• More difficult to set strategy</td>
</tr>
<tr>
<td>• Increased cost</td>
<td>• More difficult to implement data warehouses</td>
<td>• More difficult to execute strategy</td>
</tr>
<tr>
<td>• Lowered employee satisfaction</td>
<td>• More difficult to engineer</td>
<td>• Contribute to issues of data ownership</td>
</tr>
<tr>
<td></td>
<td>• Increased organizational mistrust</td>
<td>• Compromise ability to align organization</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Divert management attention</td>
</tr>
</tbody>
</table>

Source: Redman, 1998
Impacts of Poor Data Quality

Inability to

• manage the quality and performance of software or application development
• Estimate and plan realistically

Ineffective

• process change instead of process improvement
• and inefficient testing causing issues with time to market, field quality and development costs

Products that are painful and costly to use within real-life usage profiles

Bad Information leading to Bad Decisions
Why a Measurement and Analysis Infrastructure Diagnostic

Quality of data is important

- Basis for decision making and action
- Erroneous data can be dangerous or harmful
- Need to return value for expense

Cannot go back and correct data once it is collected – opportunity/information lost

Need to get the quality information to decision makers in an appropriate form at the right time

Keep from collecting the wrong type of data
Polling Question

To what extent does your organization take steps to ensure the quality of its project data?

- Not at all
- Somewhat
- A great deal
Outline

The Need for a Measurement and Analysis Infrastructure Diagnostic (MAID)

- Measurement errors and their impact

MAID Methods

- Process Diagnosis
- Data and Information Product Quality Evaluation
- Stakeholder Feedback

Summary and Conclusion
MAID Objectives

Compare an organization’s current measurement and analysis activities against a defined set of criteria

- Are we doing the right things in terms of measurement and analysis?
- How well are we doing those things?
- How good is our data?
- How good is the information we generate?
- Are we providing value to the organization and stakeholders?

Make recommendations for improvement

- How can identified gaps or weaknesses be addressed?
- How can we prepare for achieving higher maturity?
  - Many mistakes made in establishing M&A at ML2 and 3 that do not create a good foundation for ML4 and 5
Methods Overview

The MAID approach includes

- a thorough review of measurement-based planning documents, processes/procedures, analysis results, and management reports
- a series of individual and group interviews with personnel who
  - collect measurement data
  - analyze, interpret and report the measurement information
  - use the reported data to make decisions
- a briefing and detailed report describing the strengths and weaknesses of the measurement program
Criteria for Evaluation: Measurement Planning Criteria

Measurement Objectives and Alignment

• business and project objectives
• prioritized information needs and how they link to the business, organizational, regulatory, product and/or project objectives
• necessary organizational and/or software process changes to implement the measurement plan
• criteria for the evaluation of the measurement process and quality assurance activities
• schedule and responsibilities for the implementation of measurement plan including pilots and organizational unit wide implementation

Adapted from ISO 15939.
Measurement Planning Criteria

Measurement Process

- definition of the measures and how they relate to the information needs
- responsibility for data collection and sources of data
- schedule for data collection (e.g., at the end of each inspection, monthly)
- tools and procedures for data collection
- data storage
- requirements for data validation and verification procedures
- confidentiality constraints on the data and information products, and actions/precautions necessary to ensure confidentiality
- procedures for configuration management of data, measurement experience base, and data definitions
- data analysis plan including frequency of analysis and reporting

Adapted from ISO 15939.
Criteria for Evaluation: Measurement Processes and Procedures

Measurement Process Evaluation

• Availability and accessibility of the measurement process and related procedures
• Defined responsibility for performance
• Expected outputs
• Interfaces to other processes
  — Data collection may be integrated into other processes
• Are resources for implementation provided and appropriate
• Is training and help available?
• Is the plan synchronized with the project plan or other organizational plans?
## Documenting Measurement Objectives, Indicators, and Measures

<table>
<thead>
<tr>
<th>Objective</th>
<th>Questions</th>
<th>Visual Display</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Data Collection
- **How**
- **When/How Often**
- **By Whom**
- **Form(s)**

### Data Reporting
- **Responsibility for Reporting**
- **By/To Whom**
- **How Often**

### Specify Measures

### Collect Data

### Communicate Results

### Specify Data Collection Procedures

### Establish Measurement Objectives

### Communicate Results

### Analyze Data

### Specify Analysis Procedures

### Store Data & Results

### Data Storage
- **Where**
- **How**
- **Security**

### Algorithm

### Assumptions

### Interpretation

### Probing Questions

### Analysis

### Evolution

### Feedback Guidelines

### X-reference

---

**Software Engineering Institute**

---

David Zubrow, October 2008

© 2008 Carnegie Mellon University
Criteria for Evaluation: Data Definitions

Data Definitions (meta data)

• Completeness of definitions
  – Lack of ambiguity
  – Clear definition of the entity and attribute to be measures
  – Definition of the context under which the data are to be collected
• Understanding of definitions among practitioners and managers
• Validity of operationalized measures as compared to conceptualized measure (e.g., size as SLOC vs. FP)
Criteria for Evaluation: Data Collection

Data collection

• Is implementation of data collection consistent with definitions?
• Reliability of data collection (actual behavior of collectors)
• Reliability of instrumentation (manual/automated)
• Training in data collection methods
• Ease/cost of collecting data
• Storage
  – Raw or summarized
  – Period of retention
  – Ease of retrieval
Criteria for Evaluation: Data

Quality

- Data integrity and consistency
- Amount of missing data
  - Performance variables
  - Contextual variables
- Accuracy and validity of collected data
- Timeliness of collected data
- Precision and reliability (repeatability and reproducibility) of collected data
- Are values traceable to their source (meta data collected)

Audits of Collected Data
Criteria for Evaluation: Data Analysis

Data analysis

- Data used for analysis vs. data collected but not used
- Appropriateness of analytical techniques used
  - For data type
  - For hypothesis or model
- Analyses performed vs. reporting requirements
- Data checks performed
- Assumptions made explicit
Appropriate Analysis: Types of Hypothesis Tests

- **Data Type**
  - # Samples (Data groups)
    - 1 Sample
    - 2 Samples
    - 3+ Samples

- **Interval or Ratio (Parametric Tests)**
  - Mean
  - Variance

- **Ordinal (Non-Parametric Tests)**
  - Median
  - Variance / Fit

- **Nominal**
  - Similarity
    - >2 cells
      - Chi-Square
      - Binomial Sign Test
      - =2 cells

- **Proportion**
  - Similarity
    - 1 Proportions test
    - 2 Proportions test

- **1 Sample**
  - 1-sample t test
  - Chi-Square test
  - 1-sample Wilcoxon Signed Ranks test
  - Kolmogorov-Smirnov Goodness of Fit test

- **2 Samples**
  - Independent 2-sample t test
  - Paired t test
  - Normal F test
  - Paired
  - Not Normal
  - Mann Whitney U test
  - Wilcoxon matched Paired

- **3+ Samples**
  - ANOVA (1 & 2 way ANOVA; Balanced ANOVA; GLM)
  - MANOVA (General & Balanced)
  - Independent Bartlett test
    - Levene test
    - Not Normal
  - 1-way ANOVA
  - Friedman 2-way ANOVA
  - Van der Waerden Normal scores test

- **Proportion**
  - Fisher Exact test (1-way ANOVA; Chi-Square test)

**Software Engineering Institute | Carnegie Mellon**

David Zubrow, October 2008
Analysis Evaluation: Appropriate Modeling

Quantifying Relationships of X Factors with Y Outcomes

ANOVA & MANOVA

Y (Continuous)
- 1 Variable, 2 levels: t-Test, Hotelling’s T², Discriminant Analysis
- 1 Variable, ≥ 2 levels: 1-Way ANOVA, 1-Way MANOVA
- ≥ 2 Variables: 2-Way ANOVA, 2-Way MANOVA
- Mixtures of Discrete & Continuous: ANCOVA, MANCOVA

ANOVA & MANOVA in Minitab

Y (Continuous)
- 1 Variable, 2 levels: Stast-Basic Stats> sample t or paired t
- 1 Variable, ≥ 2 levels: Stast-ANOVA/1-way or 1-way (unstacked)
- ≥ 2 Variables: Stast-ANOVA/2-way, Stast-ANOVA/Balanced or General MANOVA

Continuous
- Correlation & Regression
- ANOVA & MANOVA

Discrete
- Chi-Square & Logit
- Correlation & Regression

You Begin Here

X (Discrete)
- Continuous
- Discrete

X (Continuous)
- Correlation & Regression
- ANOVA & MANOVA

X (Continuous)
- Correlation & Regression in Minitab
- Bivariate Regression
- Multiple Regression
- Canonical Analysis (CANONA)

X (Continuous)
- 1 Variable
- ≥ 2 Variables

X (Continuous)
- 1 Variable
- ≥ 2 Variables

Software
Criteria for Evaluation: Reporting

Reporting

- Evidence of use of the information
- Timing of reports produced
- Validity of measures and indicators used
- Coverage of information needs
  - Per CMMI
  - Per Stakeholders
- Inclusion of definitions, contextual information, assumptions and interpretation guidance
Criteria for Evaluation: Stakeholder Satisfaction

Stakeholder Satisfaction

• Survey of stakeholders regarding the costs and benefits realized in relation to the measurement system

• What could be improved
  — Timeliness
  — Efficiency
  — Defect containment
  — Customer satisfaction
  — Process compliance

Adapted from ISO 15939.
Polling Question

Do you feel your organization views measurement and analysis as a process?

• Not at all
• Somewhat
• A great deal
Outline

The Need for a Measurement and Analysis Infrastructure Diagnostic (MAID)

• Measurement errors and their impact

MAID Methods

• Process Diagnosis
• Data and Information Product Quality Evaluation
• Stakeholder Feedback

Summary and Conclusion
Summary

Measurement and analysis is a process

- It needs to be supported to be institutionalized and effective
- Some measurement error and diminished utility will result from choice of measurement infrastructure elements, procedures and instrumentation

Measurement Infrastructure Diagnostic:

- Characterizes performance of measurement system
- Identifies improvement opportunities for:
  - Measurement processes and data quality

Good information from high quality measures and analyses to support decision making
In God We Trust,
All Others Bring Good Data.

[Attributed to W. Edwards Deming, father of quality revolution]
SEMA Curriculum

Implementing Goal-Driven Measurement
- Feb 24-26 in DC, June 9-10, September 15-17, December 1-3 in DC

Analyzing Project Management Indicators
- March 10-11, July 14-16, October 6-8

Improving Process Performance using Six Sigma
- January 26-30, April 20-24, November 2-6

Designing Products and Processes using Six Sigma
- May 18-22, December 7-11 in DC
Questions?
References


