Electronic Raw Data and the Use of Electronic Laboratory Notebooks

Ron Shoup
Outline for Today

- **Raw Data** - defining it, creating it, and keeping it for posterity
- **Hybrid Data** - paper, electronic
- **Data Integrity** - regulatory concerns from recent inspections
- **Electronic Laboratory Notebooks**
  - Introducing the Concept
  - Capabilities and Advantages
  - Implementation - one case study
  - Future View
Defining Raw Data

- Simplest form—visual reading of a measurement, recorded on paper, with attribution.

- **Raw data:**
  - Earliest original recorded observation of a process, event, or state.
  - Primary measurement from an instrument or tool.
  - Measurement that is not changed or replaced by secondary operations performed later, through human intervention.
  - The point of origin for any post-measurement processing.
Examples of Raw Data

• **Transcription of mass reading** from a balance to a lab notebook or paper form by the operator in real time

• **Observation by phlebotomist** that blood sample from Subject 123 was drawn at 10:30 AM

• **Operating conditions for a mass spectrometer** as recorded in lab notebook

• **Temperature and humidity data** in archives data logger
Detector Output Signal (mV)

Raw Data Points

Time
<table>
<thead>
<tr>
<th>Filename</th>
<th>Sample ID</th>
<th>Response</th>
<th>RT</th>
<th>S/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>09Feb2012_020_120208175129</td>
<td>Site Check 10pL</td>
<td>35924.23</td>
<td>0.50</td>
<td>30710.368</td>
</tr>
<tr>
<td>09Feb2012_021_120208175129</td>
<td>Control Negative</td>
<td>121</td>
<td>0.50</td>
<td>3.047</td>
</tr>
<tr>
<td>09Feb2012_022_120208175129</td>
<td>Control Negative</td>
<td>9</td>
<td>0.96</td>
<td>3.246</td>
</tr>
<tr>
<td>09Feb2012_023_120208175129</td>
<td>Control Positive</td>
<td>24</td>
<td>1.02</td>
<td>5.835</td>
</tr>
<tr>
<td>09Feb2012_024_120208175129</td>
<td>CS 5 pg/mL</td>
<td>1611</td>
<td>0.53</td>
<td>83.581</td>
</tr>
<tr>
<td>09Feb2012_025_120208175129</td>
<td>CS 10 pg/mL</td>
<td>5342</td>
<td>0.53</td>
<td>173.966</td>
</tr>
</tbody>
</table>

**Method**

09Feb2012_025_120208175129 - mz= 147.04-147.04  RT: 0.45 - 1.20  NL: 2.8E3
F+ cEB (BRM 1972) [147.037-147.047]
CDS and MS Systems: Raw Data

- **Instrument Data Acquisition Parameters** recorded at time of sample injection

- **Detector Response vs. Time**
  - Simplest, earliest record of the run
  - Detector signal intensity for a given mass transition, wavelength, or potential
  - Captured according to the Data Acquisition parameters before any further manual processing
  - Permits operator to separate processed data files from raw data for complete transparency
  - Allows any subsequent operator to reconstruct processed data from the same original datafile.
Data Integrity

- Capture the original observation as raw data, at the moment it was observed.
- Capture it once—no transcriptions, no copies.
- Who was the observer and when did they record the data?
- Any changes? Reasons? Who made the change? Who reviewed the change?
- Audit trails!
- How to lock the data and preserve for posterity?
21 CFR Part 11

- Compliance for electronic records:
  - Access must be limited and controlled
  - Systems validation appropriate for the level of risk
  - Electronic signatures
  - Electronic records equivalent to paper records
  - Written policies and SOP’s to hold users accountable
  - Training, education, and experience of staff
  - Quality System to control systems and data
21 CFR Part 11 systems gone awry: paper vs electronic

Q: When may be paper records be used as an alternate means of preserving electronic records?
A: The printouts must preserve the accuracy, completeness, content, and meaning of the electronic source record—FDA August 2003 Guidance on Part 11

*Hybrid systems are the most difficult to organize and maintain for complete transparency.*
Interpretation Problems with Part 11

• We define our printouts as the raw data. The electronic files are really just the means to create the raw data.
• We archive only the PDF’s of the paper outputs as our raw data.
• We have a Part 11 compliant system. All files are therefore trustworthy.
• We don’t routinely review audit trails in QAU. Audit trails are too detailed for even the simplest data.
• We had to give administrator access to all lab staff in order to use the software effectively.
FDA Inspection Comments

- Failure to use audit trails to track any changes to the study data.
- Failure to indicate the reason for change to raw data.
- The data system allows users to erase and overwrite raw data due to widespread administrator rights.
- Non compliance with electronic signatures policy
- Lack of controls to prevent tampering with data
- Lack of management oversight to insure data integrity
- No process for using audit trails for the review of original and processed data
- Biased integrations without traceability; disabling audit trails
Electronic Laboratory Notebooks

- An electronic lab notebook is a computer program designed to replace paper laboratory notebooks.
The issues with paper:

- Searching for information in paper systems
- Printed methods and companion paper forms
- Detecting disconnects between various paper systems (equipment calibration expiration vs. use in study)
- Knowing when data is truly created in a paper system
- Knowing whether results are built from secured and validated resources
- **Exhaustive QC checks to verify data**
• Why implement an ELN?:
  – Attribution of every data entry
  – Versions of every form during completion
  – Electronic signature of originator and reviewer
  – Integrity of data inside DB, no loose files to collect
  – Completeness - all records at your fingertips
  – Searchability, filtering
  – Ability to:
    – Control forms and process
    – Control access to the process
    – Supervise data entries in real time for errors and non-completes
# One Implementation

<table>
<thead>
<tr>
<th>4Q 2009</th>
<th>Completed evaluation of 3 ELN products, purchased IDBS E-Workbook</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Q 2010</td>
<td>Extensive initial training; wrote user requirements for approximately 20 workflows to create templates</td>
</tr>
<tr>
<td>2Q-4Q 2010</td>
<td>Template creation, administrative operations decisions</td>
</tr>
<tr>
<td>1Q-2Q 2011</td>
<td>Validation of templates and hierarchy; final SOPs</td>
</tr>
<tr>
<td>2013</td>
<td>Over 60 templates now validated</td>
</tr>
</tbody>
</table>
ELN Terminology

• **Validated Template**- a predefined set of tables, requiring the same information every time (similar to a paper form) that has been tested and approved.

• **Experiment**-
  – Any information placed in the ELN; placing a word document into the ELN is an “experiment”
  – Any process performed using a validated template; for example: balance checks, audits, equipment maintenance

• **Version History**- the history of saves performed on an experiment
Empty template ready for use in experiment:
Completed experiment:
Hierarchical controls:

Only approved templates allowed for regulated studies

Only designated method(s) can be applied to the study

Folders define one’s ability to execute, and where to execute

All staff can view any experiment, at any stage
• New sponsor in Watson ➔ sponsor hierarchy built in ELN
• New study in Watson ➔ study hierarchy built in ELN
• New analytical run in Watson ➔ new run built in ELN under Run Organizer folder of study
• Standardized sample names for all types of normal and validation samples
Watson data autoloaded into validation experiment:

1. Complete the Template Modification Table.
   Fill in the expected final concentration of the Selectivity samples in the Template - Selectivity VS Final Concentration.
   The rows that have NA in them do not need a concentration.

2. Scan each matrix lot into the Matrix Scan Table and scan the Spiking Solution barcode in the Spiking solution Query Table.
   a) Click the Run Associated Searches button to run both the Matrix and the Spiking Solutions queries.
   b) If the wrong compound is found in the Spiking Solution table open the Analyte table to see the wrong compound highlighted in Red.

3. Scan or select the barcodes of the appropriate matrices and enter all mandatory information in the Selectivity Table.
   Add any Pipettes.
   a) If more than 6 matrices are needed, click on the "+" item and hit enter to add the appropriate amount of matrices.
   b) Matrices selected for Spiked samples are automatically selected for corresponding Unspiked samples.
   c) If a red triangle appears in the cell then the matrix is an invalid matrix.
   Choose another matrix.

Selectivity Table

<table>
<thead>
<tr>
<th>Select Matrix Lot</th>
<th>Expiration Date of Matrix Lot</th>
<th>Data Field Check</th>
<th>Spiking Solution_A</th>
<th>Total Volume (ml)</th>
<th>Spiking Solution Percent</th>
<th>Spiking Solution Value (g/l)</th>
<th>Spiking Solution Value (g/l)</th>
<th>Sample Barcode Label ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20-Nov-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>20-Nov-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>20-Nov-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>20-Nov-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>20-Nov-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>20-Nov-2011</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Matrix Scan

<table>
<thead>
<tr>
<th>Scan Matrix</th>
<th>Species</th>
<th>Matrix</th>
<th>Anticoagulant</th>
<th>Experiment Hyperlink</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Human</td>
<td>Plasma</td>
<td>K2 EDTA</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Human</td>
<td>Plasma</td>
<td>K2 EDTA</td>
<td>Executed Common Practices Experiment (K2 EDTA)</td>
</tr>
<tr>
<td>C</td>
<td>Human</td>
<td>Plasma</td>
<td>K2 EDTA</td>
<td>Executed Common Practices Experiment (K2 EDTA)</td>
</tr>
<tr>
<td>D</td>
<td>Human</td>
<td>Plasma</td>
<td>K2 EDTA</td>
<td>Executed Common Practices Experiment (K2 EDTA)</td>
</tr>
<tr>
<td>E</td>
<td>Human</td>
<td>Plasma</td>
<td>K2 EDTA</td>
<td>Executed Common Practices Experiment (K2 EDTA)</td>
</tr>
<tr>
<td>F</td>
<td>Human</td>
<td>Plasma</td>
<td>K2 EDTA</td>
<td>Executed Common Practices Experiment (K2 EDTA)</td>
</tr>
</tbody>
</table>

Selectivity Blank Matrix - Peak 1

<table>
<thead>
<tr>
<th>Matrix Lot</th>
<th>Analyte Area</th>
<th>Analyte Area (Analyst Mean)</th>
<th>ISTD Area</th>
<th>ISTD Area (ISTD Analyst Mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOT A</td>
<td>75</td>
<td>0.1%</td>
<td>10</td>
<td>0.0%</td>
</tr>
<tr>
<td>LOT B</td>
<td>5</td>
<td>0.0%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>LOT C</td>
<td>10</td>
<td>0.1%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>LOT D</td>
<td>0</td>
<td>0.0%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>LOT E</td>
<td>0</td>
<td>0.0%</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>LOT F</td>
<td>0</td>
<td>0.0%</td>
<td>0</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Selectivity ID | SEL_BAM.0056-000161

Template Approved For User?: Yes
Template Formulas Complete?: Yes
All Mandatory Fields Complete?: Yes
Passed all Field Checks?: Yes
Recovery calculations, from knowing the run number and sample identifiers
Components for “Audit By Exception”
4 Versions of an Evolving Record:

Version: 1, User: bousuma At: Sep 19, 2011 9:46:35 AM (Created EXPERIMENT)
Version: 2, User: bousuma At: Sep 19, 2011 9:49:07 AM (Data Added)

EXPERIMENT

Reason: Data Changed
Additional Comments: Replaced Completed Vendor Survey Form (insurance information added)
## Audit Trail Example: Equipment Maintenance

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>User</th>
<th>Type</th>
<th>Description</th>
<th>Sheet</th>
<th>Source</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun 21, 2011</td>
<td>12:57:51 PM</td>
<td>engelb</td>
<td>Map Variables</td>
<td>Variable(s) mapped to catalog</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:00 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>&lt;Use Tools: Auditing; Show Audit Description to see description&gt;</td>
<td>Equipment Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:06 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>Value &quot;Preventative Maintenance&quot; added to cell 'Calibration or Reverification, Preventative Maintenance, or Repair': Entry</td>
<td>Maintenance</td>
<td>Maintenance Information</td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:08 PM</td>
<td>engelb</td>
<td>Search Run</td>
<td>Associated searches run</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:11 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>Value &quot;EQM-BEQ.1112-03-Jan-2011-003&quot; added to cell Expt ID:Entry</td>
<td>Maintenance</td>
<td>Maintenance Information</td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:17 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>&lt;Use Tools: Auditing; Show Audit Description to see description&gt;</td>
<td>Preventative</td>
<td>Preventative Information</td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>1:09:44 PM</td>
<td>engelb</td>
<td>Data Changed</td>
<td>System updated Experiment Properties from E-WorkBook</td>
<td>Experiment</td>
<td>Equipment Information</td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>1:09:51 PM</td>
<td>engelb</td>
<td>Data Changed</td>
<td>Value changed from &quot;2011-06-21T12:58:13 (40715.5404355556)&quot; to &quot;2011-06-09T00:00:00 (40703)&quot; in cell Current Date:Entry</td>
<td>Maintenance</td>
<td>Maintenance Information</td>
<td>Wrong Date-Update current date to reflect date of oil change.</td>
</tr>
</tbody>
</table>
Audit Trail Example: Equipment Maintenance - oil change in pump

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
<th>User</th>
<th>Type</th>
<th>Description</th>
<th>Sheet</th>
<th>Source</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun 21, 2011</td>
<td>12:57:51 PM</td>
<td>engelb</td>
<td>Map Variables</td>
<td>Variable(s) mapped to existing Variable(s)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:57:55 PM</td>
<td>engelb</td>
<td>Data Changed</td>
<td>System updated E from E-Workbook</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:50:00 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>Use Tools: Auditing&gt; Description to see description&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:06 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>Value &quot;Preventative to cell Calibration or Preventative Maintenance Entry&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:10 PM</td>
<td>engelb</td>
<td>Search Run</td>
<td>Associated search entered</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:11 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>Value &quot;EQM-BEQ: 1003&quot; added to cell Maintenance Information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:13 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>Value &quot;2011-06-21: 5404355556&quot; added to cell Maintenance Information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>12:58:17 PM</td>
<td>engelb</td>
<td>Data Added</td>
<td>&lt;Use Tools: Auditing&gt; Description to see description&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>1:09:44 PM</td>
<td>engelb</td>
<td>Data Changed</td>
<td>System updated Experiment Properties from E-Workbook</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun 21, 2011</td>
<td>1:09:51 PM</td>
<td>engelb</td>
<td>Data Changed</td>
<td>Value changed from &quot;2011-06-21T12: 58:13 (40715.5404355556)&quot; to &quot;2011-06-09T00:00:00 (40703)&quot; in cell Current Date: Entry</td>
<td></td>
<td></td>
<td>Wrong Date - Update current date to reflect date of oil change.</td>
</tr>
</tbody>
</table>
## Sources of Bioanalytical Data

<table>
<thead>
<tr>
<th>Facility Records</th>
<th>Training Records</th>
<th>Study Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software development and validation</td>
<td>Regulatory training</td>
<td>Protocols, sample analysis and validation plans</td>
</tr>
<tr>
<td>Software change control and administration</td>
<td>Qualifications, CV, job description</td>
<td>Sample management</td>
</tr>
<tr>
<td>Metrology records</td>
<td>SOP and other internal training</td>
<td>Reagents, sample preparation</td>
</tr>
<tr>
<td>Environmental controls</td>
<td>Meetings, workshops, shortcourses</td>
<td>Instrumental analysis and data processing</td>
</tr>
<tr>
<td>Archives- access, in/out</td>
<td></td>
<td>Scientific review and acceptance</td>
</tr>
<tr>
<td>Reverifications depending on risk assessment</td>
<td></td>
<td>Reporting</td>
</tr>
<tr>
<td>Quality audits to management</td>
<td>Quality audits to management</td>
<td>Quality audits to management</td>
</tr>
</tbody>
</table>
Evolution of Good Data Practices

• 1990: Lab notebooks, chromatography integrators, one-shot integrations.

• 1995: CDS, diskettes with raw and processed files, paper printouts, bankers boxes

• 2000: LIMS, data files on local drives, PDFs, CDs

• 2005: SDMS/database to hold raw and processed files and metadata, searchable.

• 2007: ELN/database to capture all data and records, no paper forms, with complete version control and audit trails in place.
Ron Shoup
rshoup@aitbioscience.com
317-715-8826

AIT Bioscience, LLC
7840 Innovation Blvd
Indianapolis, IN 46278