

Applications, Standards, and Architecture in Practice

BioIT World
May 1, 2007
Boston, Ma

Richard Ferrante
RDF Software Group
rdf(at)rdfsg.com

The Goal of Architecture

- Allow timely solution of business problems
 - As the business, science and regulatory environments evolve
- Gather data in actionable form
 - Accurate, vetted and current
- Example: BI tool rev 0; *the evolution of a spreadsheet*
 - Beginning as a personal record keeping device
 - Evolution into a tool/metaphor
 - Arguably remains a good presentation technique for even the most sophisticated analyses
- Highlight the importance of Architecture and Standards as tools become more sophisticated.

Initially A Personal Tool

- Data Entry / Consolidation
 - A record of what has happened
- Data Analysis
 - Spreadsheet formulas
 - Computed flags

If variance > 2.0
red background

Compound	Lot	Cost	IC50	PK	Trial 1 endpoint	Variance	Trial 2 endpoint	Variance
A	1	50	0.01	20	7	2	8	5
A	2	100	0.01	23	7	1	9	1
A	3	20	0.01	15	5	1.5	6	1.2
A	4	10	0.01	18	6	1	6	1

Begins to be Shared

- Copying the spreadsheet to individual desktops
 - 3 users, 3 copies of the spreadsheet, 3 sets of data?

Compound	Lot	Cost	IC50	PK	Trial 1 endpoint	Variance	Trial 2 endpoint	Variance
A	1	50	0.01	20	7	2	8	5
A	2	100	0.01	23	7	1	9	1
A	3	20	0.01	15	5	1.5	6	1.2
A	4	10	0.01	18	6	1	6	1

Compound	Lot	Cost	IC50	PK	Trial 1 endpoint	Variance	Trial 2 endpoint	Variance
A	1	50	0.01	20	7	2	8	5
A	2	100	0.01	23	7	1	9	1
A	3	20	0.01	15	5	1.5	6	1.2
A	4	10	0.01	18	6	1	6	1

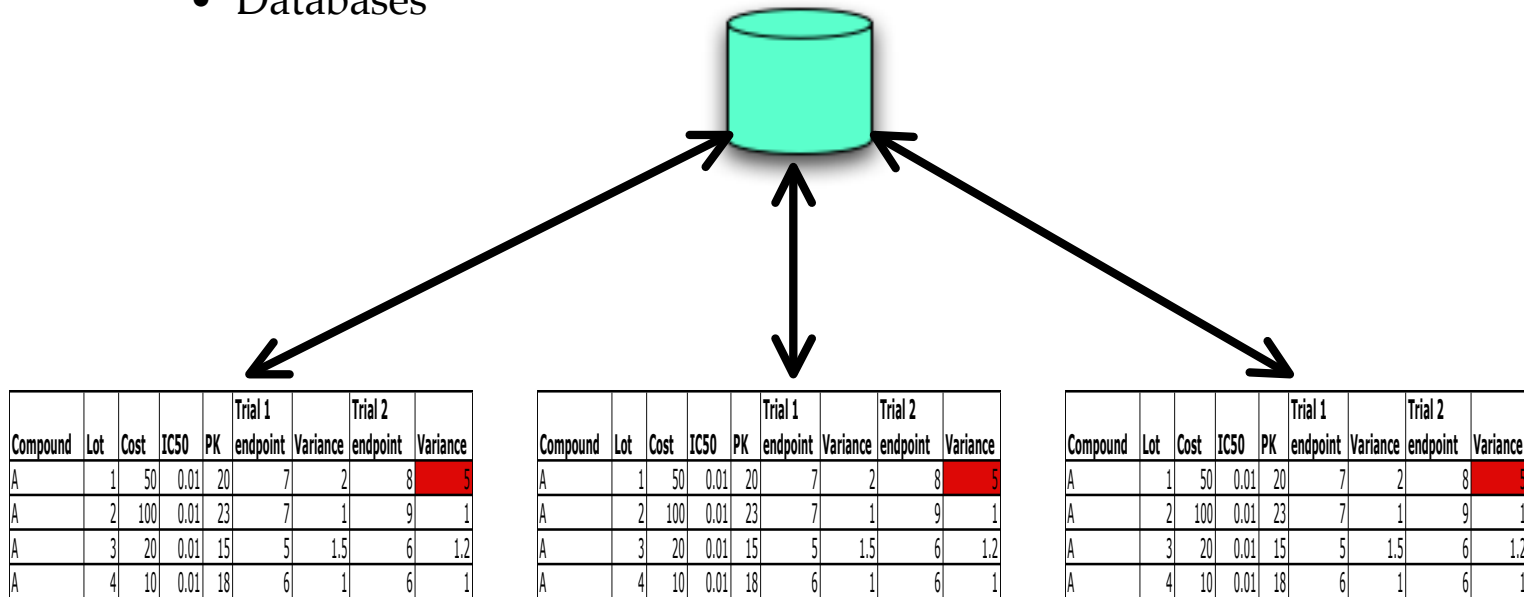
Compound	Lot	Cost	IC50	PK	Trial 1 endpoint	Variance	Trial 2 endpoint	Variance
A	1	50	0.01	20	7	2	8	5
A	2	100	0.01	23	7	1	9	1
A	3	20	0.01	15	5	1.5	6	1.2
A	4	10	0.01	18	6	1	6	1

Sharing Raises the Bar

- The Spreadsheet
 - Are the formulas in the spreadsheet correct?
 - Shared
 - Business processes
 - Analysis techniques
 - Business process modeling
 - Business process reengineering
 - Do we have the same version of the spreadsheet?
 - Revision control
 - Deployment processes
 - Arguably 'pre-architecture' but.....
 - How do you reference the columns for analysis?
 - How are expected/variant values designated
 - Via formula? e.g., 'variance > 2' in the example
 - Via cell?
 - Via analysis?
 - Becomes critical as the spreadsheet grows into a workbook etc.
 - Data identity
 - Analysis consistency

Sharing The Data

- Shared Data
 - Files
 - Databases

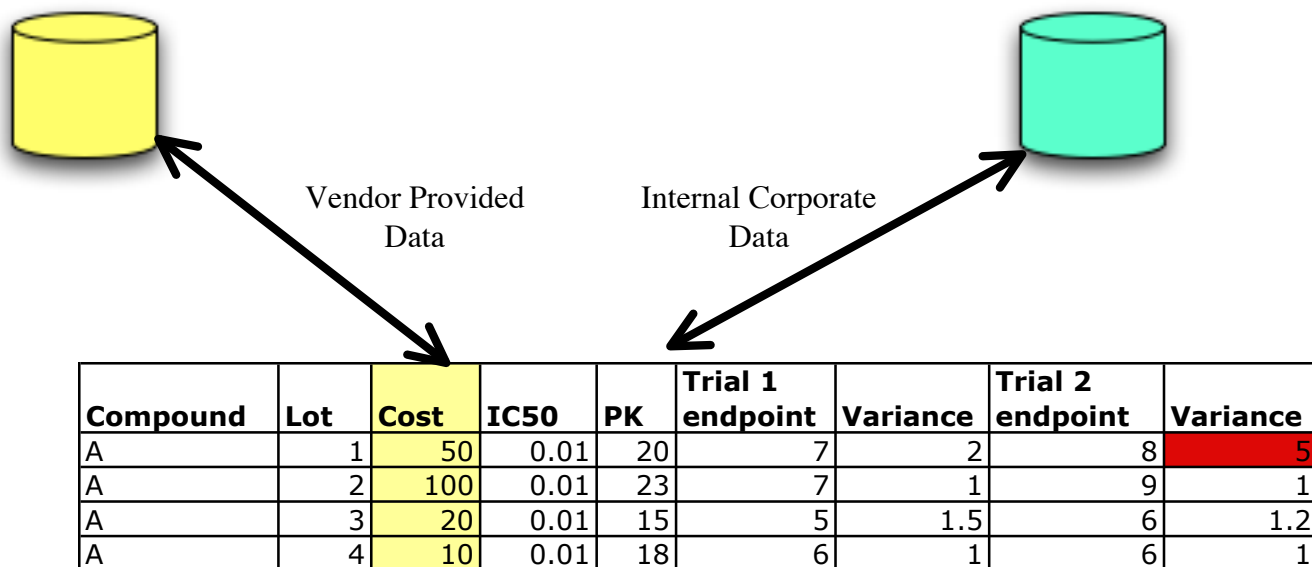


Are You Seeing What I'm Seeing?

- Shared data location
 - Files
 - Database access
- Agree upon the protocols for updating the data
 - Insertion
 - Alteration
- Agree upon the structure of the data
 - Cardinality
 - Evolution
 - Ownership
- One organization/ One (or more) standards
 - “Discussions” reflect conflicting experience/needs
 - Can be very individualized
- Adopting industry standards “not necessary”
 - May facilitate internal discussions
 - Mergers
 - Revisiting internal processes

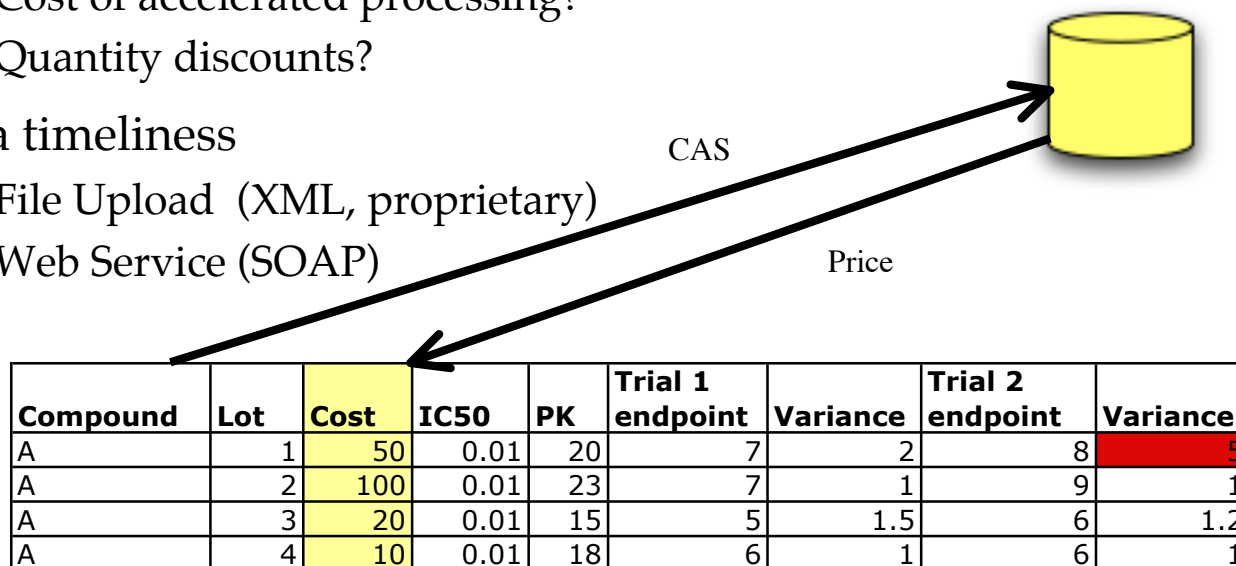
Sharing Across Organizations

- Our data + your data = better information
- Identifier based attribute access
 - Compound Prices



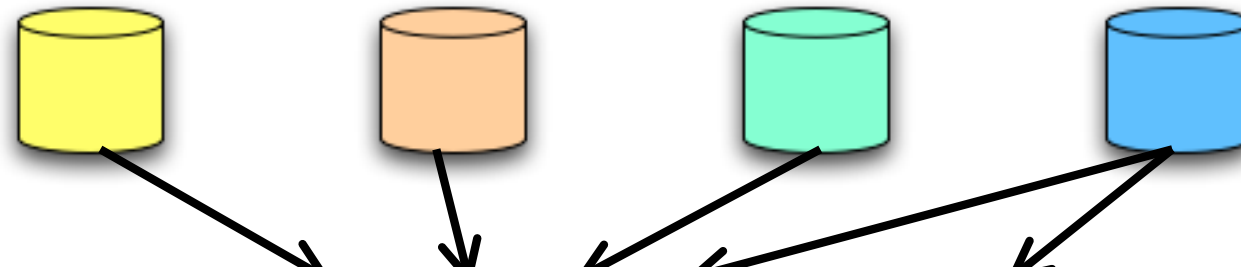
Sharing Data Across Organizations

- Vendor supplies more than one client (they hope!)
- Data format
 - CAS identifiers
 - “Request Quantum” single/batch compound query, business parameters
 - Cost of accelerated processing?
 - Quantity discounts?
- Data timeliness
 - File Upload (XML, proprietary)
 - Web Service (SOAP)



Sharing Processes

- We buy the compound
 - I do in-vitro
 - You do in-vivo
 - CRO for clinical



Compound	Lot	Cost	IC50	PK	Trial 1 endpoint	Variance	Trial 2 endpoint	Variance
A	1	50	0.01	20	7	2	8	5
A	2	100	0.01	23	7	1	9	1
A	3	20	0.01	15	5	1.5	6	1.2
A	4	10	0.01	18	6	1	6	1

Sharing the Analysis

- Communicate around the processes
- Standardize around a dataset
 - CDISC
 - HL7
- Data Insert
 - Incremental?
 - Batch
 - Per cell?
- Query Statement/Resolution
 - Shared analysis
 - Shared result
- Data Revision
 - Batch/Cell
 - Capture cause?
 - Capture corrective action?

Increasing Intensity of Business Interaction

- Tools
 - Share via copying
- Data (actionable information)
 - Copying / gathering individual elements
 - Data elements and values
- Processes (subcontracting)
 - Copying / gathering sets of data
 - Datasets
- Analysis (collaboration)
 - Processes and question answering
 - Interactions and workflows



Increasing Depth of Architectural Interaction

- Tools
 - Common platform
- Data (actionable information)
 - Common identifiers
- Processes (subcontracting)
 - Data set granularity / frequency
- Analysis (collaboration)
 - Common data structure
 - Common workflow patterns (timeliness, corrections)

CDISC SDTM
HL7

CDISC LSIA



RDF Software Group

Richard Ferrante
rdf (at) rdfsg.com