

Taking Data Quality Metrics to the Boardroom; a Case Study

Lowell Fryman, Enterprise Integration Solutions LLC

Taking Data Quality Metrics to the Boardroom: A Case Study

Lowell Fryman

President

Enterprise Integration Solutions LLC



April 23-27, 2006 Denver, Colorado

1

Agenda

- Case Study Background
- Traditional IT Metrics
- Measuring Data Assets
- Example Metrics
- Presentation is King
- Approach Summary
- Q&A



© 2006 EIS LLC

DAMA-I Conference April 2006

2

The DAMA International Symposium & WILSHIRE Meta-Data Conference
Denver, Colorado ● April 23-27, 2006

Background

- For 20+ years Data Management professionals have considered data as an enterprise asset
- Today many business executives agree
- Data critical to business initiatives should be monitored and reported to leadership
- Yet, enterprises do not measure and improve the value of the data assets



© 2006 EIS LLC

DAMA-I Conference April 2006

3

Case Study Background

- Established Media and Yellow Page Publisher
- Sales force of 1,000
- Annual revenue of \$1.5 billion USD
- CIO publicly declares “our data is a competitive advantage”
- True, but no metrics on the data assets existed



© 2006 EIS LLC

DAMA-I Conference April 2006

4

Case Study Background

- One strategic business initiative was stated as “increase market share of Internet advertising”
- Changing the company Internet site to provide greater search capabilities
- Partnering with other Internet advertising firms
- Expanding the media delivering the advertising services they provided to their customers



© 2006 EIS LLC

DAMA-I Conference April 2006

5

Case Study Background

- The Internet requires mapping and advanced “search” functions
- Advertiser Address data supported the business initiative
- Data captured for the Yellow page Printing business may not have the level of completeness and accuracy for Internet mapping
 - The same data may have differing quality requirements for different business usages



© 2006 EIS LLC

DAMA-I Conference April 2006

6

Traditional IT Metrics

- ✱ Total Cost of Ownership (TCO)
- ✱ Return on Investment (ROI)
- ✱ Service Level Agreement (SLA)
- ✱ Service Level Operations (SLO)
 - ✱ Up-Time
 - ✱ Availability
- ✱ Other Metrics
 - ✱ Bug fixes
 - ✱ Help Desk metrics



© 2006 EIS LLC

DAMA-I Conference April 2006

7

Traditional IT Metrics

- ✱ Financial Metrics
 - ✱ TCO
 - ✱ ROI
- ✱ Performance Metrics
 - ✱ SLA
 - ✱ SLO
 - ✱ Help Desk
 - ✱ Application Fixes
- ✱ Not Data Asset metrics!
- ✱ We do not commonly measure Data Assets



© 2006 EIS LLC

DAMA-I Conference April 2006

8

The old adage

- ✱ “If you can’t measure it, you can’t manage it”
- ✱ Managing by gut fill is still prominent

- ✱ Data is critical to most business initiatives
- ✱ Data quality tends to degrade as data ages or the more data is touched
- ✱ All data does not have the same value to the organization

- ✱ So how do we find the “valuable” data assets?

© 2006 EIS LLC

DAMA-I Conference April 2006

9

Case Study – Customer Address

- ✱ Ability to map an Advertiser Address on the Internet
- ✱ Requires a longitude and latitude from geocoding software
- ✱ Optimal is a complete and CASS certified address
- ✱ Look at businesses in a Yellow page book
 - ✱ (800) GOODPIE
 - ✱ (720) PLUMMER
 - ✱ Corner of 104th & I-25 (no City)
 - ✱ 375 E County Line Rd. (no City)

© 2006 EIS LLC

DAMA-I Conference April 2006

10

Case Study – Customer Address

- Data Profiling software to determine complete and missing data
- Quantify the population of Addresses
- Quantify the Addresses that could be measured and improved
 - Eliminate National and telephone only
 - Fill in data that could be
 - Worked with USPS and Vendors to improve processes
- Geocoding software to assign Long/Lat at rooftop and or centroid levels

© 2006 EIS LLC

DAMA-I Conference April 2006

11

Measuring Data Assets



- What Data Assets should be measured?
- What are the business initiatives that must be supported now? What are the priorities?
- What categories (subjects) of data are created or used by that business initiative?
- What are the usage business rules versus the capture business rules?
- What are the quality requirements for each usage?
- What level of quality exists in the data?
- What level of quality can be defined as a target for improvement?

© 2006 EIS LLC

DAMA-I Conference April 2006

12

Measuring Data Assets

- What are the components of a quality asset
- Define what is important to your organization
- *Case study*
 - *Value of data = quality of data*
 - *Quality=Accuracy x Completeness x Timeliness x Relevancy x Accessibility*
- Many data quality approaches consider difference components for quality
- **Accuracy** and **Completeness** are the basis for all



© 2006 EIS LLC

DAMA-I Conference April 2006

13

Examples of Data Asset Metrics

- *Both absolute values and percentages are valuable*
- Count of Customers
- Customers with addresses
- Customers with complete addresses
- Customers with phone numbers only
- Percentage of Customers with phone only
- Percentage of Customers with addresses
- Percentage of Customers with complete addr.

© 2006 EIS LLC

DAMA-I Conference April 2006

14

Examples of Data Asset Metrics

- ✱ Count of addresses geocoded
- ✱ Count of addresses not geocoded
- ✱ Count of addresses failed geocoding
- ✱ Count of geocodes at rooftop
- ✱ Count of geocodes at centroid
- ✱ Percentage of addresses geocoded
- ✱ Percentage of addresses not geocoded
- ✱ Percentage of failed geocodes
- ✱ Percentage of geocodes at rooftop
- ✱ Percentage of geocodes at centroid

© 2006 EIS LLC

DAMA-I Conference April 2006

15

Presentation is King

- ✱ Many different presentation methods may be used to different audiences
- ✱ Senior executives respond better to less detail and more pictures
- ✱ Line managers are used to detail and numbers/reports
- ✱ “A picture is worth a thousand words”
- ✱ Charts are best (of any kind)

© 2006 EIS LLC

DAMA-I Conference April 2006

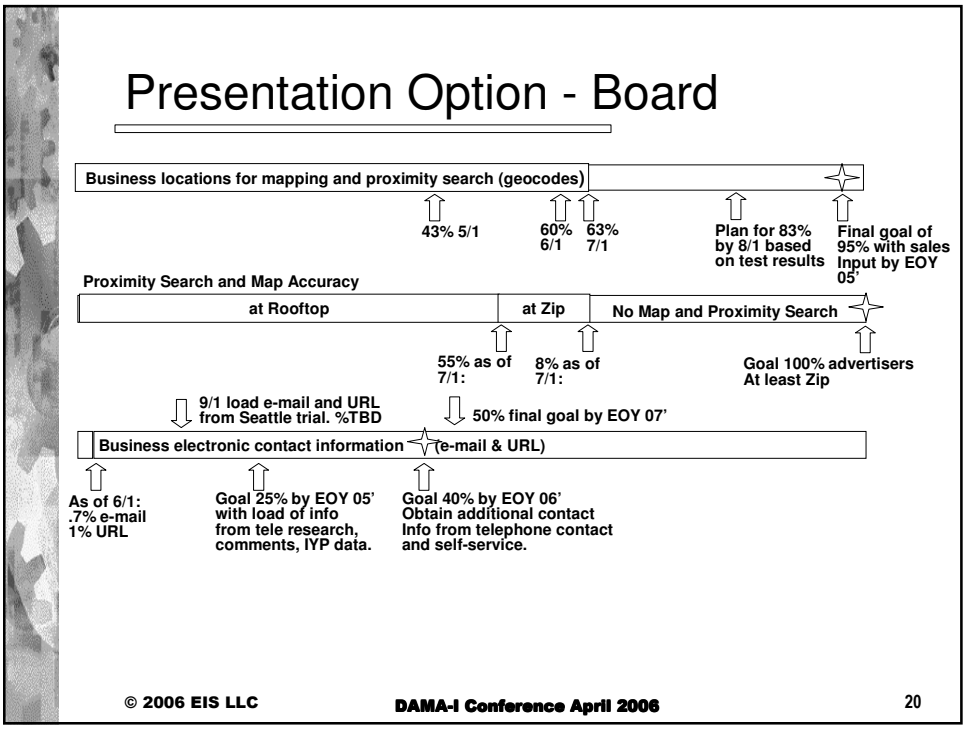
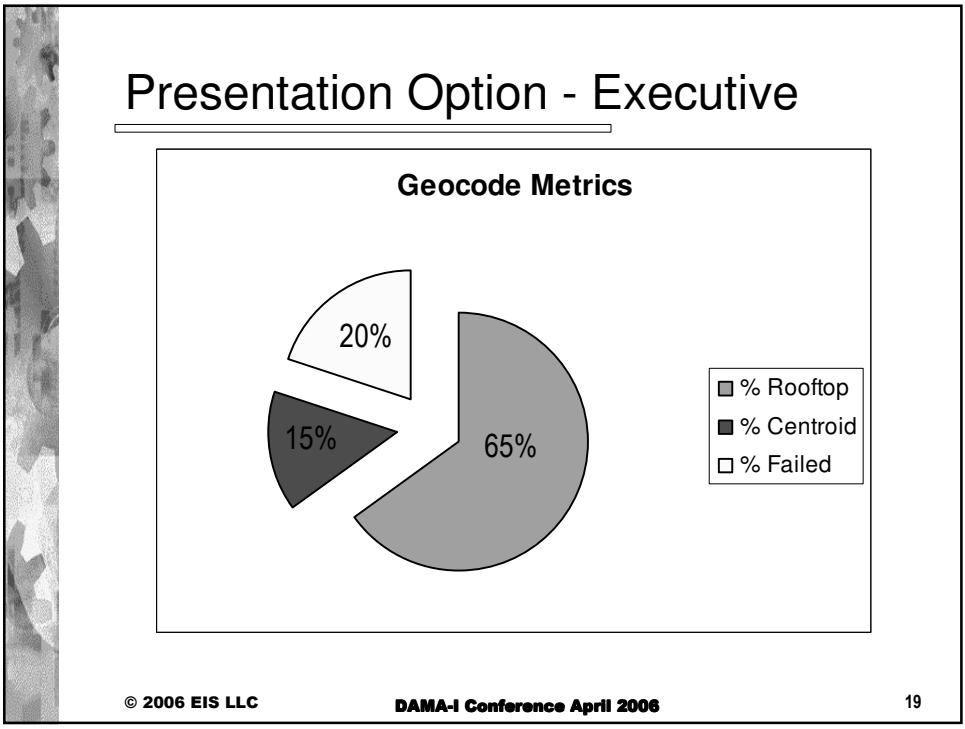
16

Presentation Option - Analyst

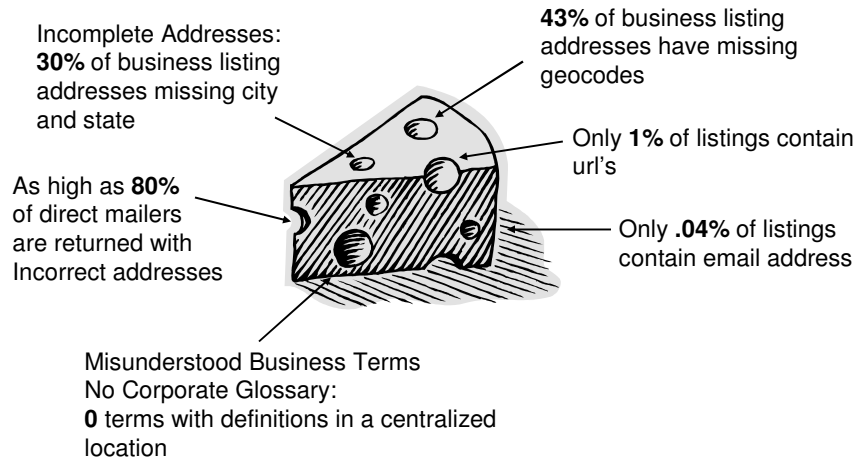
GEOCODE VALUE	11/5/05	11/12/05	11/19/05	11/28/05
0	2,795,259	2,797,494	2,799,031	2,801,374
1	493,281	493,505	493,946	494,360
2	335,431	354,940	365,157	375,488
4	100,630	118,482	118,547	118,646
5	110,692	130,330	130,402	130,511
8	183,155	183,784	183,872	184,253
9	154,664	155,196	155,270	155,591
10	69,192	69,429	69,463	69,607
Blank	181,820	182,461	182,547	182,726

Presentation Option - Manager

Count of businesses with addresses	3,701,543
Percentage with Geocoded address	80
Percentage with Rooftop Geocode	65
Percentage with Centroid Geocode	15
Percentage Failed Geocode	20



Presentation Option - Executives



© 2006 EIS LLC

DAMA-I Conference April 2006

21

Approach Summary

- Executive sponsorship of the program is critical
- Start with the highest priority business initiatives (ROI or political)
- Identify the data usage for the initiative
- Identify the quality components most critical for measuring success
- Define the desired target for each quality component (100% is generally not realistic)
- Establish the people and processes for the quality program, as well as the tools

© 2006 EIS LLC

DAMA-I Conference April 2006

22

Approach Summary

- Data profiling tools are great for initial analysis of the data
- Define the timing and life-cycle for each metric
- Identify causes for poor quality data
- Create a database to store metrics values over time for trending
- Resolve conflicting business rules that may exist between data capture and data usage
- Never underestimate the importance of graphical presentations
- For executives, less is best
- You may have to manually create the charts for your presentations, that's OK

Q & A

