A Man with a Watch: The Convergence of Authoritative Data Sources and Data Quality

ABSTRACT

It is a little-known fact that Mark Twain was the first promoter of Authoritative Data Sources (ADS). Twain observed that "a man with a watch knows what time it is. A man with two is never sure." Eliminating the redundant watch reduces ambiguity, reduces frustration, and saves money. Likewise, in today's world, designation of authoritative data sources (ADSs) is a powerful means to reduce any enterprise's data redundancy, and by so doing, to reduce ambiguity, improve consistency, ease the burdens and uncertainties imposed on end users resulting from the need to evaluating competing data sources, and ultimately, to reduce cost.

Twain was also an early proponent of data quality, observing that "the difference between the right word and the almost right word is like the difference between lightning and a lightning bug." Again, today's methodologies for data quality management provide an equally powerful means to improve the accuracy and usability of data, and in so doing to improve operations, enable more confident decision-making, and again, ultimately, to reduce costs.

More important, these two approaches…deliberate designation of ADSs and methodical data quality management… are mutually-reinforcing. Data quality management enables potential or designated authoritative data sources to meet rigorous quality standards, including standards on the associated metadata.   Conversely, the evaluation and designation of ADSs narrows the scope of the demanding data quality management regimen.

This is a management talk. It:

- Briefly summarizes the management structures needed to effect ADS and DQ separately,
- Synthesizes the two into a powerful whole,
- Lays bare the organizational barriers to success in this unified endeavor, and
- Summarizes the benefits of this powerful convergence.

BIOGRAPHY

Thomas C. Redman
President
Navesink Consulting Group

Dr. Thomas C. Redman, "the Data Doc," is President of Navesink Consulting Group, based on the Jersey Shore. He has helped hundreds of organizations and thousands of people understand the importance of data, start their data programs, and make order-of-magnitude improvements. In doing so, they lower costs, increase revenues,

improve customer satisfaction, and more make more confident planning decisions. His fourth book, Data Driven: Profiting from Your Most Important Business Asset (Harvard Business Press) is a Library Journal Best Business Book of 2008. Prior to forming Navesink, in 1996, Tom started and led the Data Quality Lab at Bell Labs. There, he and his team were first to extend quality principles to data and information. Tom is an internationally known lecturer and the author of dozens of papers. He holds two patents.

Andy Nodine Jr.
Windmill International Inc.

Andy Nodine is employed by Windmill International, Inc., of Nashua, NH, as a consultant. Since 2006, he has been serving the office of the Director of Plans and Integration within the Headquarters Air Force Deputy Chief of Staff, Manpower, Personnel and Services organization, with a focus on implementation of the Air Force's Information and Data Management Strategy. The goal of this initiative is to supply users across the Air Force Enterprise with on-demand access to authoritative, relevant and sufficient data in support of the Air Force mission by means of designated authoritative data sources, information discovery based on semantic technology, and governance through communities of interest aligned around business domains. Andy also works in the areas of data quality management and data governance. Andy's education includes a Bachelor of Science in Electrical Engineering from Union College in Schenectady, New York, and a Master of Science in Electrical Engineering from Michigan State University. He also attended a number of professional military education schools as part of his military career.

A Man With a Watch: The Convergence of Authoritative Data Sources and Data Quality
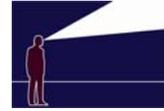
Wright A. Nodine Jr.

Windmill International, Inc.

Thomas C. Redman, Ph.D.

Navesink Consulting Group

MIT Information Quality Industry Symposium

July 13-15, 2011

tomredman@dataqualitysolutions.com

wanodine@worldshare.net

The Problem/Opportunity

Organizations are awash in redundant, inconsistent, disparate, competing, siloed data stores.

This leads to bad decisions, increases cost, confuses customers, and generally contributes to the prevalent data chaos.

Mark Twain said:

"A man with a watch knows what time it is.

A man with two is never sure."

Our Message

Designation of **Authoritative Data Sources** has been proposed as a means to address the problem.

However, this is only half of the solution!

Just because a source is designated doesn't mean the data it provides is of good quality!

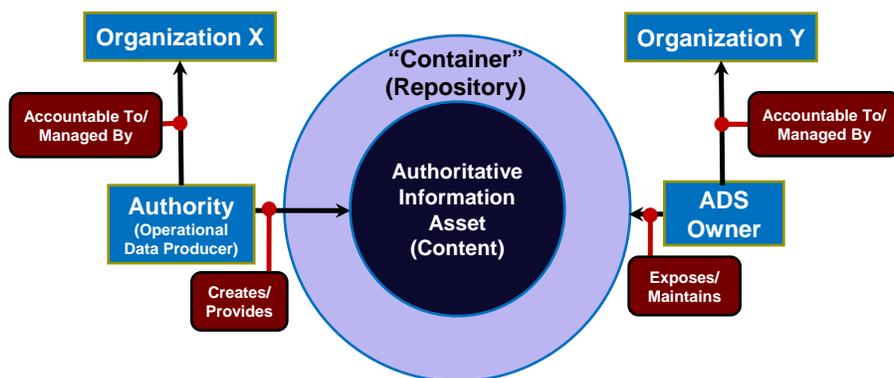A more complete solution requires assurance of **High Quality Data…**

…AS WELL!

Herein lies the convergence!

What is an Authoritative Data Source?

Credit for these concepts goes to Scott Renner and Frank Ponzio, who first posited and developed the ideas we present here.

ADSs in the Real World

Authority-based:

Example: The Government of Libya

Pros: A powerful authority can pull it off

Cons: Without proven quality, no one believes it

Marketplace-based:

Example: Bloomberg (in the front office)

Pros: High market demands and rich rewards for success

Cons: Few organizations operate like a market
   Weak basis for comparison

Do These Strategies Really Differ?

In the final analysis…

☐ A data source must <u>earn</u> the designation "authoritative" by deeds, not dictum.

☐ Data customers will use sources that WORK!

How Do We Apply this Thinking Inside an Organization (like the Air Force)?

Major Lessons:

□ A data source must earn the designation by deeds, not management dictum.

□ Markets can help, but are no panacea.

□ Internally, management can create conditions that lead to competition among data sources.

To make them work in a Hierarchical Bureaucracy:

□ Management must establish a governance regime with criteria and a process for vetting/selection.

□ Accountability must be maintained.

Redman-Nodine-Trusted Data-MIT-July2011   © Navesink Consulting Group, 2000-2011   WAN, TCR p. 7

The Point:

Quality Matters!
DQ must be moved
to the front and center

Redman-Nodine-Trusted Data-MIT-July2011   © Navesink Consulting Group, 2000-2011   WAN, TCR p. 8

The Point, Part Two:

In order to be accepted as an ADS,
a source must contain data of
quality that is
KNOWN, ADVERTISED and PROVEN

…based on parameters such as accuracy,
precision, and currency

What is the Basis for Trust?

Customers trust data when they a plot like this one.
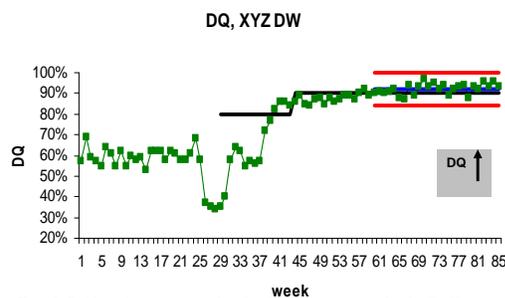

DQ, XYZ DW

Figure 2. Decision makers come to trust data when they become accustomed to plots like this one.

There is a lot going on here:
1. Understanding of customer needs (implicit)
2. Measurement
3. Early detection of a big issue
4. Continuous improvement
5. Statistical control
6. TRANSPARENCY

In the Hierarchical Bureaucracy, There Are Things Senior Leadership Can Do

Some things management ought to do anyway:

- Establish governance
- Set minimum standards for "authoritative:"
  - Clear management accountability
  - Follow relevant standards (i.e., "corporate definitions)
  - Minimum DQ-level and rate of improvement
- Require that all data sources measure and publish quality levels
- Require that all data be shareable to authorized users across the enterprise
- Establish and enforce clear accountability for compliance
- Teach best practices (see Appendix)
- Create and maintain the "data directory" to facilitate all of the above

The Real Issues Involve Incentives

Our series of questions:

1. How are people and organizations motivated?
   - Fear (e.g., loss of funding or status)
   - Fame (e.g., recognition, career advancement)
   - Fun (e.g., freedom to do things their own way)
   - Fortune (e.g., big bonuses)
   - Sense of Duty (e.g., loyalty, patriotism, greater good)
2. How do we evoke those motivations in a manner consistent with the type of "culture" we want to advance?
3. Can senior leadership implement the needed steps?

Each Organization will Answer Differently

Air Force:

- ☐ Alignment with mission
- ☐ Opportunity to garner investment (conversely, fear of loss of funding or status
- ☐ Recognition

Investment Bank:

- ☐ Fortune in the form of big bonuses
- ☐ Advancement

"Don't try this at home, kids"

- ☐ This is not simple, nor for the timid
- ☐ It requires…
  - ■ Visible leadership with long-term commitment
  - ■ Sense of urgency
  - ■ New organizational alignments and capabilities
  - ■ Consistent support and emphasis
  - ■ "Blocking and Tackling"
  - ■ Culture change
- ☐ But if we fail to act, we will fail, period.

Again, Our Message

**Designation of Authoritative Data Sources
holds the potential to ease the
prevalent data chaos.**

But this is only part of the solution.

**ADSs must ALSO deliver data of quality that is
high, known, and proven.**

**Getting there will take equal measures of
brains, brawn and guts.**

What Did They Say?

**This concludes
my briefing…**

Questions?

Thomas C. Redman, Ph.D.
+1 732-933-4669
tomredman@dataqualitysolutions.com

Wright A. (Andy) Nodine Jr.
+1 210-573-4242
wanodine@worldshare.net

Appendix 1:
Those with the Best Data "Think Prevention"

A Database is Like a Lake
To Clean Up The Lake, One Must First
Eliminate The Sources Of Pollutant

Redman-Nodine-Trusted Data-MIT-July2011   © Navesink Consulting Group, 2000-2011   T. C. Redman, Page 17

Appendix 2, The Ten Habits: 1-5

1. Focus on the most important needs of the most important *customers*.
2. Apply relentless attention to *process*.
3. Manage all critical sources of data, including *external suppliers*.
4. *Measure* quality at the source and in business terms.
5. Employ *controls* at all levels to halt simple errors and establish a basis for moving forward.

Redman-Nodine-Trusted Data-MIT-July2011   © Navesink Consulting Group, 2000-2011   T. C. Redman, Page 18

Appendix 2, The Ten Habits: 6-10

6. Develop a knack for *continuous improvement*.
7. Set and achieve *aggressive targets* for improvement.
8. Formalize *management accountabilities* for data.
9. Lead the effort using a *broad, senior group*.
10. Recognize that the hard data quality issues are soft and *actively manage the needed cultural changes*.

Redman-Nodine-Trusted Data-MIT-July2011   © Navesink Consulting Group, 2000-2011   T. C. Redman, Page 19