



RBC Capital Markets

Institutionalized Impediments to Data Quality

The MIT IQ Industry Symposium Program
Session 1B: Federal Data Architecture

Thursday, 19 July 2007



Institutionalized Impediments to Data Quality



Organizations are the building blocks of our society. Yet every day these organizations engage in actions that inhibit their own effectiveness and those around them.

Five of the most glaring institutionalized impediments to data quality are noted below. All are major speed bumps on the road to data transparency and meaningful data quality assessment.

1. **Incompetent:** firms don't know the business that they are involved in.

Case: a prominent financial data services firm markets a product as a security master database that does not contain intra-day option strikes for all assets covered.

Challenge: this product is widely used by internet vendors of financial information, thus providing an incomplete picture of the breath of the markets.

Tactic: Using the guiding philosophy of “You can only play the game well if you know the rules”, step one involves identifying the source or sources of the rules. For listed options traded domestically this involves going to the Options Clearing Corporation (OCC). Step two involves finding one or more data sources that conform with the rules; the data sources would serve as benchmarks and as **insurance** against incompetent business partners.

Take-away

- Insurance

Institutionalized Impediments to Data Quality

2. **Malicious**: firm provides incorrect information to clients on processing of trade orders.

Case: a prominent brokerage firm provides false information on the reasons behind its order handling process.

Challenge: clients feel locked in (needlessly) to the brokerage firm because of the apparent complexity of the services provided.

Tactic: Step one involves articulating the problem that you need solved; this should be done at the lowest level of granularity needed for its solution. Step two should match the business partner solution to the articulated problem, element by element. Steps three and four involve verifying the claims made by the business partner, both the **positive claims** (what they can do) and the **negative claims** (why the problem is so difficult – why they can do what you cannot). In the case of trade order processing, this involves going to the exchanges and detailing how they handle orders – again, their rules and regulations. Once this has been done then the value added by the business partner can be clearly defined, and assessed against its cost. And most importantly, knowing the mechanics allows for opportunities to be correctly evaluated and taken advantage of, unlocking the firm from an unnecessary dependency.

Take-away

- Investigate both the positive claims and the negative claims.

Institutionalized Impediments to Data Quality

3. **Lazy:** firm provides outdated information on their web site regarding services provided and the performance (operational details) of the services.

Case: a prominent financial organization has extensive information on its web site detailing its internal system structure, services provided by the system, and performance attributes; when asked about the posted information the exchange says it is two years out of date.

Challenge: decreased use of the exchange and significant slippage in market share.

Tactic: Step one is to develop a document explaining the structure and services provided by the firm under review; the document should be crafted to address the needs of your audience. Step two involves getting the offending firm to review the document. And step three is obtaining a sign-off on the document from the firm being reviewed. Knowing the actual capabilities of the firm will provide you with a competitive advantage – it will allow you to correctly assess the role that this firm can play in your overall business plan. Or not play, as the case may be.

Take-away

- The negative side of the Principle of Least Effort.

The positive side: "The Principle of Least Effort: 'Progress doesn't come from early risers--progress is made by lazy men looking for easier ways to do things.'" Robert Heinlein in *Time Enough for Love*

Institutionalized Impediments to Data Quality

4. **Adversarial**: organization exploits information sources for its own benefit.

Case: This situation is found at both the inter-organizational level and at the intra-organizational level.

Challenge: incomplete information distribution and weaker than necessary analysis of problem situations.

Tactic: Data transparency through full disclosure is the goal that should be demanded by information consumers. In general, suppliers will attempt to provide the lowest cost product acceptable to the demand side. For example, providing a market analysis tool that doesn't explain its underlying algorithm in detail will not allow the tool user to correctly utilize the output of the tool. Or, not explaining how a price is created, puts the user of the information at a disadvantage.

Step one is for consumers, especially those within a firm, to demand products exhibiting data transparency. Step two is for you to carry your demand for data transparency beyond the walls of the business and into the community at large.

Take-away

- Information flow is demand driven.

Institutionalized Impediments to Data Quality

5. **Bureaucratic**: organization is compartmentalized into information silos (Balkanized).

Case: a prominent financial firm handles client information processing via a silo architecture, where the commercial side is segregated from the ecommerce side (hybrid model). This can occur in exclusively ecommerce businesses as well as hybrids.

Challenge: the firm has a security and cost exposure - the customer can game the system once the disconnect is identified. Lack of integrated client processing and differential processing paths allow savvy individuals to exploit the deficiency.

Tactic: Addressing this problem involves shifting the focus of the business from product development to process development. Step one is to understand that when speaking about data we are not speaking about free-floating atoms of information; each data element is the connecting link in a three part chain (data chain) of input process-data-output process. Step two is to identify data chain redundancies within the organization. And step three is to determine which data chain redundancies are critical to the business and which are not.

Take-away

- Data chain.