

ROI Models for DQ

Diby Malakar

Senior Director – Engineering & Product Management
Cloud9 Analytics



Agenda

- What to expect?
- Need for ROI
- 15 Principles of ROI Development
- Financial Valuation Models
- Quality Metrics & Quality SLA
- Case Study
- Q & A

What to expect?

- Understand the background and need for ROI
- Understand what works and what does not work
- Learn techniques for developing easy-to-comprehend ROI Models
- Use the above to build business case for instituting a DQ Management program

Does DQ matter?

- Cost of poor data may be 10-25 % of the total revenues of any organization
- Topmost technology reason for CRM projects to fail or miss expectations
- Data Warehousing Institute estimates that data quality problems cost U.S. businesses more than \$600 billion a year

...Most organizations get this...

So why are we here?

It is NOT the lack of awareness as regards to the value of high-quality information,

but

rather how critical they are to the business and how bad data really affects the business

How does ROI fit in?

- Like any other project, DQ and Data Governance initiatives require funding to proceed
- Management will not sign off on anything unless value of investment can be quantified
 - No action happens until they know what they are going to get

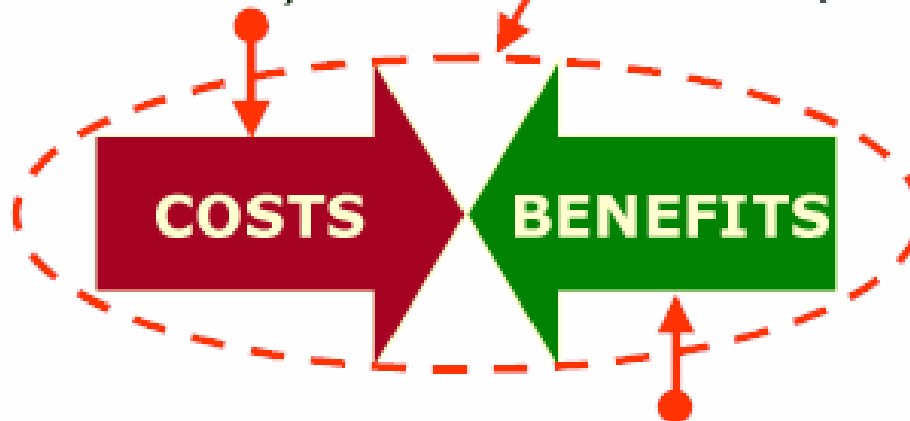
What is ROI?

- Short for *Return On Investment*
- ROI is
 - An accounting formula
 - Used to obtain an actual or perceived future value of an expense or investment
 - Source: <http://www.webopedia.com/TERM/R/ROI.html>
- ROI measures how effectively a business uses its capital to generate profit
 - The higher the ROI, the better

ROI: Looking at it visually

WHAT WE GIVE
(INVESTMENT)

ROI
(WORTH)



WHAT WE GET
(RETURN)

ROI Example

- ROI is a return ratio that compares the net benefits of a project verses its total costs.
- $ROI = (\text{Net Benefits} / \text{Total Costs}) * 100$
- Example
 - Net benefits = \$20000, Total Costs = \$10000
 - $ROI = (\$20000 / \$10000) * 100 = \mathbf{200\%}$
 - In this case, Net benefits is *twice* the Total Expected Costs

ROI – Art or Science

Securing funding for DQ projects more of an *Art* than a true science

...But...there is hope...!

Basic principles can be followed to achieve greater success

Principles of ROI Development

- Principle 1 – Know where the money is
 - Identify the costs and opportunities that are associated with processes that depend on data
 - Find out as to who the stakeholders are in terms of who would pay the costs and who needs to understand the costs
 - Evaluate how these costs or revenues can be measured and tracked

Principles of ROI Development

- Principle 2 – Know your data
 - Understand data flow within organization
 - Understand usage and purpose of data within applications
 - Understand structure, meaning and quality of data
 - Understand how bad affects the business

Principles of ROI Development

- **Principle 3 – Keep the model simple**
 - Include 1 or 2 numbers that really matter
 - Don't be tempted to fill up the model with too many numbers or valuations
 - One shouldn't need to be a financial wizard to understand the numbers

Principles of ROI Development

- Principle 4 – Study your stakeholders
 - Identify stakeholder with budget and resource responsibilities
 - Identify what things are important to them
 - Understand what model they like while approving projects
 - If possible, identify the key stakeholder and tailor presentation to his or her needs
 - Do not try to please all styles

Principles of ROI Development

- Principle 5 – Ensure investment in all 3 areas
 - Time – Ensure management is aware of % time commitments from key individuals
 - Funding – Identify \$\$\$ needed for hiring or capital expenditure
 - Resource – Identify virtual or full-time resources for DQ Project

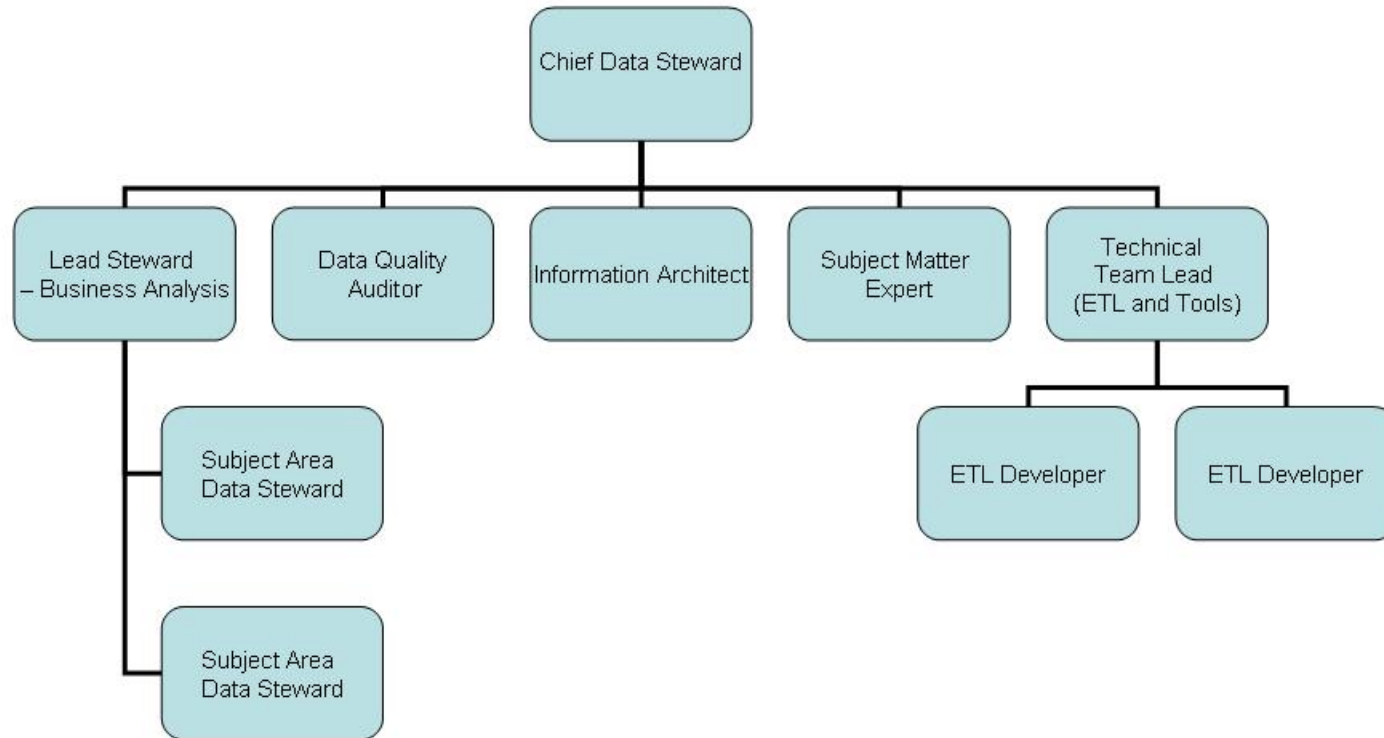
Principles of ROI Development

- Principle 6 – Pick a data problem where the problem is apparent
 - Quantify the impact of bad data
 - Identify if the data is putting the organization's goals at risk

Principles of ROI Development

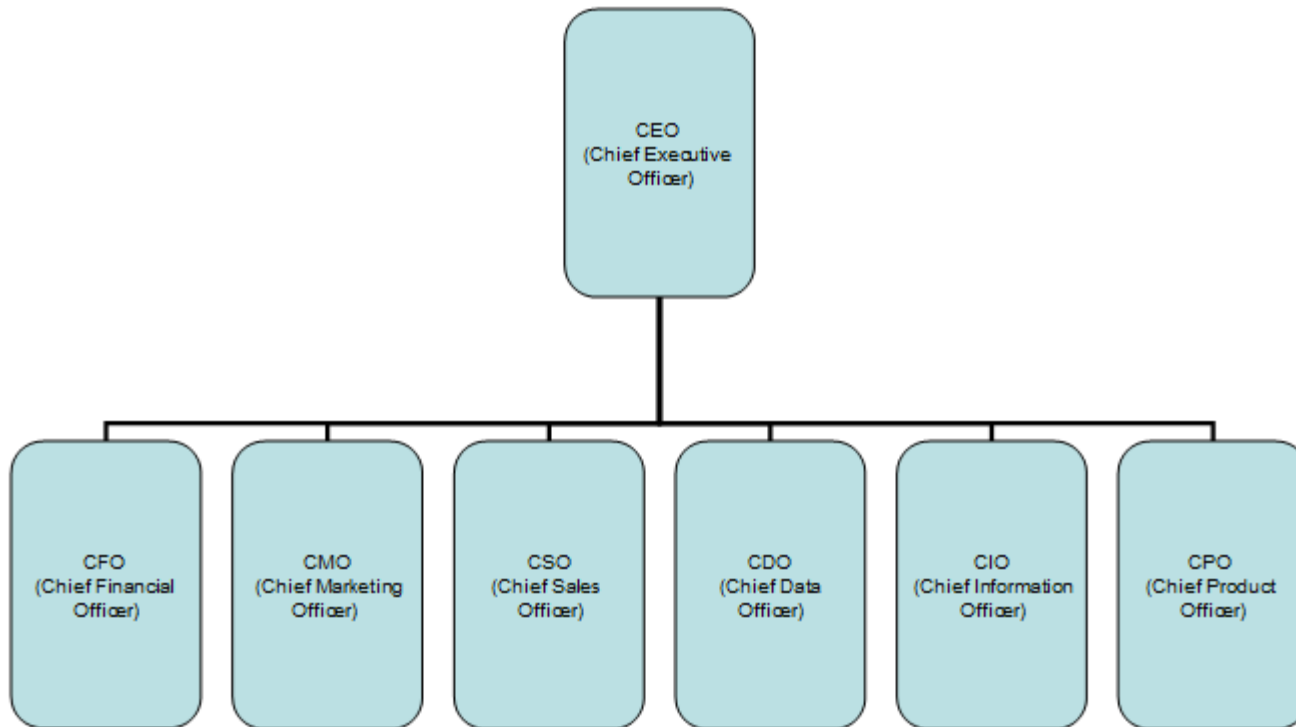
- Principle 7 – Establish Organization and structure
 - Understand downstream impact to teams
 - Define High-Level Data Governance structure with roles and responsibilities
 - Include suggestions regarding who should play the respective roles
 - Define high-level modus-operandi
 - Recommend Chief Data Steward

Chief Data Steward



Ref: Diby Malakar (TDAN Jan'06) - <http://www.tdan.com/i035hy02.htm>

Chief Data Steward becomes CDO



Ref: Diby Malakar (TDAN Jan'07) - <http://www.tdan.com/i039hy04.htm>

Principles of ROI Development

- Principle 8 – Build momentum around the state of DQ
 - Ensure company is not in denial around the DQ problem
 - Getting money for a unrecognized problem is difficult at best
 - Have meetings 1 on 1 to communicate message prior to the ROI meeting
 - Sync up with lowest level managers whose operations are impacted

Principles of ROI Development

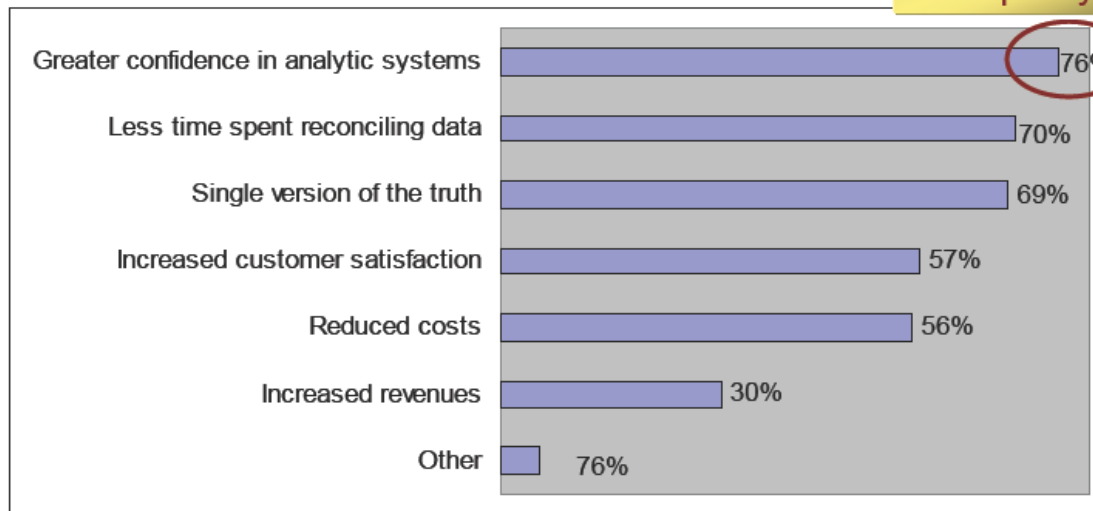
- Principle 9 – Gather statistics from external sources around DQ benefits
 - Eg: TDWI Data Quality Survey
 - Eliminates ambiguity and raises awareness around DQ

Benefits of DQ - TDWI

TDWI Data Quality Survey

Benefits Derived from High Quality Data

The #1 of high quality data



The Data Warehousing Institute, March 2006. Based on 750 respondents, multiple responses allowed

Principles of ROI Development

- Principle 10 – ROI Model development is an ongoing effort
 - ROI calculations and the ROI model itself need to be adjusted and refined at every toll gate of the Data Governance program
 - Make financial analysis a living and changing part of the program
 - Buy off on ROI is probably the most important task of the CDO

Principles of ROI Development

- Principle 11 – DQ Initiatives should almost always be tied to Compliance
 - Tie it up with SOX Compliance and it will get all the attention
 - As compliance is mandatory, there would be a higher willingness to pay for it
 - Couch it as an opportunity for improving the business instead of a “have to do” project

Principles of ROI Development

- Principle 12 – Pick a Pilot Project where results can be demonstrated quickly
 - Project duration should be less than 2 months
 - Manage scope throughout the project
 - Scale up the strategy slowly as the scope increases

Principles of ROI Development

- Principle 13 – Establish how you would measure progress using simple metrics to establish and sustain Quality SLA
 - Use concepts like EIQ (Enterprise Information Quality Quotient) to measure success of program
 - EIQ Trending (before and after) is very helpful

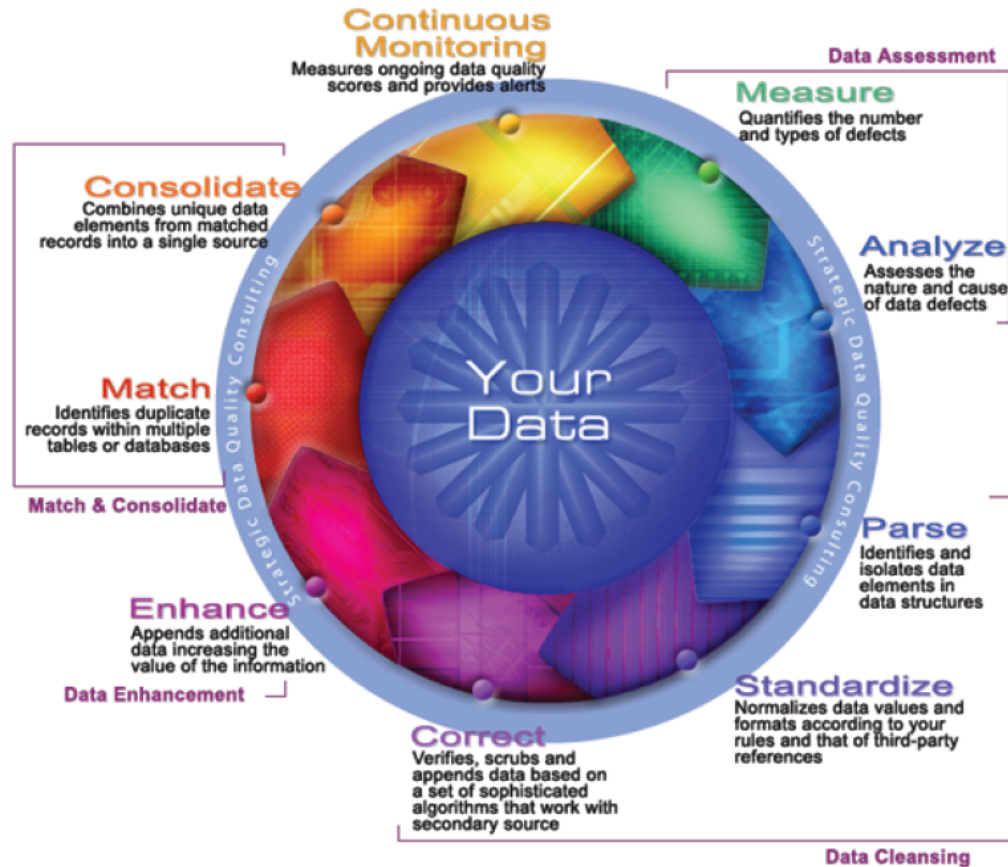
Principles of ROI Development

- Principle 14 – DQ Management
Components should include reactive and pro-active components
 - Show the mindset that the program will “react” to problems that already exist
 - Show the mindset that the DQ program framework will diminish the potential for new problems to arise
 - Both mindsets are needed

Principles of ROI Development

- Principle 15 – Provide a high-level overview of DQ Methodology to be followed
 - Business Objects' DQ Methodology
 - Larry English's TIQM Methodology

BO DQ Methodology



Alternate Financial ROI Metrics

- **Payback Period.** The amount of time required for the benefits to pay back the cost of the project
- **Net Present Value (NPV).** The value of future benefits restated in terms of today's money
- **Internal Rate of Return (IRR).** The benefits restated as an interest rate

Financial ROI Metrics: Payback Period

- **Formula: Investment/Cash Flow (Year)**
 - = \$1500/\$1000 = 1.5 years
 - Project Cost = \$1,500, Yearly Savings = \$1,000
- **Pros**
 - Easy to understand for most individuals
 - Shorter the payback period better the investment
- **Cons**
 - Does not properly account for the time value of money, inflation, risk, financing

Financial ROI Metrics: NPV

- Formula:

$$NPV = \text{initial investment} + \frac{\text{Cash flow Year 1}}{(1+r)^1} + \dots + \frac{\text{Cash flow Year n}}{(1+r)^n}$$

- Net Cash Flow = Benefits – Costs
- R = % Rate of return of comparable investment alternatives

- Pros

- Useful for financial appraisal of long-term projects

- Cons

- Calculations are a little involved

NPV - Example

Net Present Value — [Equipment] [Date]

Operational Costs	Value
Cost of equipment	\$14,500
Fixed costs	\$0
Vehicle insurance	\$1,600
Driver pay	\$16,200
Miscellaneous	\$0
Maintenance	Variable
Depreciation	N/A
New sales	125
Profit per sale	\$5
Courier delivery charge	\$16
Number of deliveries	1,250
Sale of equipment	\$2,300
Interest rate	5%

Gray cells will be calculated for you and do not require any entry.

[Company Name] CONFIDENTIAL

Term in years	Expenses			Income				Cash flow	Cumulative cash flow
	Fixed costs	Other costs	Total	Money saved by project	New sales generated by project	Equipment sales	Total		
0	\$14,500		\$14,500					\$14,500	\$14,500
1	17,800	\$200	18,000	\$20,000	\$625		\$20,625	2,625	11,875
2	17,800	250	18,050	20,000	625		20,625	2,575	9,300
3	17,800	300	18,100	20,000	625		20,625	2,525	6,775
4	17,800	400	18,200	20,000	625		20,625	2,425	4,350
5	17,800	550	18,350	20,000	625		20,625	2,275	2,075
6	17,800	700	18,500	20,000	625	2,300	22,925	4,425	2,350
7			0					0	0
8			0					0	0
9			0					0	0
10			0					0	0
								2,350	
							NPV =	\$404	

Other Financial Valuations

- **Cost-Benefit Analysis**
 - Includes issues other than cash such as time savings, loss of reputation, market penetration
- **Real Option Valuation**
 - Used to evaluate investments under conditions of uncertainty such as future product demand or future value of an asset

Other Financial Valuations

- IRR (Internal Rate of Return)
 - Calculates rate of return of a project with no assumptions around reinvestment of cash flows
 - IF $IRR >$ project's **cost of capital**, or *hurdle rate*, the project will add **value** for the company
- Modified IRR or Growth Rate of Return
 - Similar to IRR but makes explicit assumptions about reinvestment of cash flows

Quality Metrics

“One size fits all” set of metrics is never a good solution...so customize

But...

- Don't let team get too stuck on semantics
- Pick the top 5 quality metrics that are
 - Most relevant to the organization
 - Generally agreed upon

Quality Metrics: All

No	Dimension	Definition of Dimension
1	Accessibility	the extent to which data is available and quickly retrievable
2	Accuracy/Reliability	the extent to which the data is correct and reliable
3	Believability	the extent to which data is regarded as true and credible
4	Completeness	the extent to which data is of sufficient breadth and depth for the task at hand
5	Concise Representation	the extent to which data is compactly represented
6	Consistent Representation	the extent to which data is presented in the same format
7	Ease of manipulation	the extent to which data is easy to manipulate and be applied to various tasks
8	Interpretability	the extent to which data is in appropriate languages, symbols, and units, and the definitions are clear
9	Objectivity	the extent to which data is unbiased, unprejudiced, and impartial
10	Relevancy	the extent to which data is applicable and helpful for the task at hand
11	Reputation	the extent to which data is highly regarded in terms of its source or content
12	Security	the extent to which access to data is restricted appropriately to maintain its security
13	Timeliness	the extent to which the data is sufficiently up-to-date for the task at hand
14	Understandability	the extent to which data is easily comprehended
15	Value-Added	the extent to which data is beneficial and provides advantages from its use

Quality Metrics: Consolidated

No	Dimension	Definition of Dimension
1	Accuracy	the extent to which the data is correct, reliable, credible, unbiased, unprejudiced and impartial
2	Completeness	the extent to which data is of sufficient breath and depth for the task at hand
3	Consistency	the extent to which data is presented in the same format and compactly represented and is interpretable (i.e. in appropriate languages, symbols, and units, and the definitions are clear)
4	Currency	the extent to which the data is sufficiently up-to-date for the task at hand.
5	Relevance	the extent to which data is applicable and helpful for the task at hand

EIQ: Definition

- EIQ (Enterprise Information Quality) quotient is a ...
 - Single-valued enterprise level metric to measure state of DQ and IM (Information Maturity Level)

Benefits of EIQ

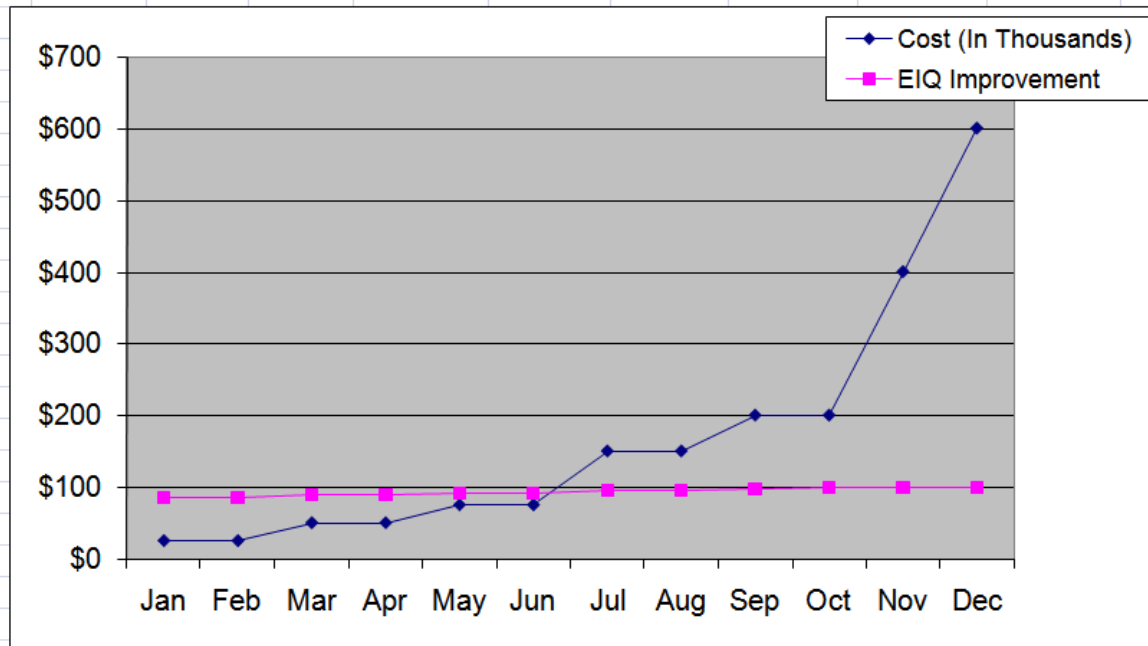
- EIQ progression can be used as a succinct measure by the Chief Data Steward
 - To communicate the state of quality to senior management
 - Provide comparative assessments over time

Management loves numbers!

EQ Trending

EQ Trending

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Cost (In Thousands)	\$25	\$25	\$50	\$50	\$75	\$75	\$150	\$150	\$200	\$200	\$400	\$600
EQ Improvement	85	86	89	90	91	92	96	96.5	97	99	99.5	99.8



Cost Categories – Hard Costs

- Hard Costs are those whose effects can be estimated and/or measured
 - Customer attrition
 - Error detection and rework
 - Error prevention
 - Customer service issues
 - Delays in processing
 - Delayed or cancelled projects

Cost Categories – Soft Costs

- Soft Costs are those that have an effect on productivity but are difficult to measure
 - Difficulty in decision making
 - Organizational mistrust
 - Lowered ability to effectively compete
 - Lowered employee satisfaction

Ref: David Loshin (Feb'02) -<http://www.tdan.com/i023fe02.htm>

Components of ROI Model

- Category
- Problem
- Quantification of Business Impact
- Yearly Cost Incurred
- Revenue or Potential Value Lost
- Total DQ Costs
- Estimated Funding
- ROI Metrics

ROI Model: Case Study

ROI Model for DQ Program

Last Updated: February 25, 2005

Category	Problem	Quantification of Business Impact	Incurred (\$)	Potential Value (Lost)
IT Infrastructure & Processes	Decentralized and redundant DQ and QA processes in the Medical-Surgical DW	Increased Cost (Staff Time). Approximately 5% of DW Medical-Surgical DW Resource Budget spent on duplicated DQ/QA efforts. Note: 2004 DW Budget: \$4.5M (30 Resource @\$150K Each)	\$225,000	
	Redundant processes around transaction feeds being handled in multiple destinations (Xconnect Feed to System A, System B, System C)	Increased Cost (Staff Time). Approximately 5% of Application Team Resource Budget spent on redundant processes Note: 2004 Application Budget: \$2.25M (15 Resource @\$150K Each)	\$112,500	
	Lengthy learning curve for new resources and occurrence of significant rework during Business Analysis phase for DW projects due to misunderstanding and misinterpretation of information.	Increased Cost (Staff Time). External Consultants spent extra time (at least 2 Weeks @ \$100/Hour) in Business Analysis and Requirements phases of DW Implementations. Note: 3 projects happen annually with 3 new consultants working on each project (80 Hours * \$100/Hour * 3 Consultants * 3 Projects)	\$72,000	
	DW Hardware Infrastructure supporting redundant processes	Increased Cost (Hardware). Approximately 5% of DW hardware budget for 2004/2005 spent on supporting redundant processes. Note: DW Hardware Budget for 2004 was \$4M	\$200,000	
Time-to-market & Customer Value	Lack of ability to support accurate "spend analytics". Price Parity Reports (available for over 1 year now) have low usage due to low confidence in accuracy of data as well as low utilization of ~30% of available data due to data quality issues.	Potential Lost Value due to low accuracy in "spend analytics and reports". Estimated savings (value) range for the top 106 hospitals from the set of 2005 (planned) reports for 2 upcoming project releases is \$40M - \$160M. These figures assume we meet the DQ goal of getting an 80%++ product match for Healthcare Supplychain Marketplace transactions. According to XYZ Study, savings per hospital are estimated to be \$2M. Note: Assuming at least 1% increase in company revenue by introduction of reports to other GPO's (potential marketplace customers) due to minimal annual savings of \$40M (0.01 * \$40M = \$400K).		\$400,000
	Cleansed Product, Orders and Unit of Measure (UOM) Data not available for transactional applications	Increased Cost (Staff Time). Estimated cost of additional work in DW to make cleansed product, orders and Unit of Measure (UOM) data available for transactional applications to be about 16 Man Weeks (3 Resources * 6 Weeks * \$100/Hour)	\$72,000	
	Redundant processes around "Contract Eligibility Rules" for pricing discounts	Increased Cost (Staff Time). Approximately 5% of Application Team Resource Budget spent on redundant contract eligibility processes (4 repositories to be synced for every business rule change) Note: 2004 Application Budget: \$2.25M (15 Resource @\$150K Each).	\$112,500	
	Inability to tie orders to contracts	Lost Value. Potentially a loss of opportunity to provide strategic value to hospitals by delivering solutions that allowed them to accurately, granularly, consistently assess their spend and contract compliance. Note: Assuming at least 1% increase in company revenue by introduction of reports to other GPO's (potential marketplace customers) due to minimal annual savings of \$40M (0.01 * \$40M = \$400K).		\$400,000
TOTAL			\$794,000	\$800,000
Other Project Costs		Overhead cost for managing non-conforming DQ-related projects, teams, and QA cycles (15% of Total Cost mentioned above - 5% Product Management, 5% Project Management, 5% Architecture)	\$119,100	
TOTAL (current DQ costs)			\$913,100	
Estimated Funding for Proposed Data Governance Program Note: 1 Business Analyst, 1 QA/Auditor, 2 Engineers @ \$150K Per Resource			\$600,000	
ROI (= (Benefits/Costs) * 100)			152.18	
Payback Period (in years) (=Investment/Cash Flow)			0.66	

Q & A



Speaker Contact Info

Diby Malakar

Senior Director – Engineering & Product Management
Cloud9 Analytics

diby.malakar@cloud9analytics.com

dibymalakar@hotmail.com

Cell: 408-375-2570