BEYOND BUSINESS PROCESS REENGINEERING (BPR): DATA QUALITY ENGINEERING

Anton Raneses*

Defense Energy Support Center; anthony.raneses@dla.mil

For presentation at the Massachusetts Institute of Technology, Nov. 5-7, 2004

Executive Summary/Abstract: Process improvement tools such as Six Sigma could cost organizations like GE as much as 0.4% of its revenues. Consequently, organizations must constantly evaluate the cost-effectiveness of data quality tools. This presentation finds that beyond business process reengineering (BPR), data quality engineering in the form of failure mode and effects analysis (FMEA) could be utilized as a relatively inexpensive means to address data quality issues. The portrayed system integration example shows that on one hand, BPR provides a sound method of identifying data quality problems by way of gap analysis. On the other hand, FMEA allows user prioritization of such gaps as a function of risk, which likely leads to decreasing the bias in the calculation of the costs of systems change requests (SCRs) and problem trouble reports (PTRs) and providing a clearer picture of the bottom line.

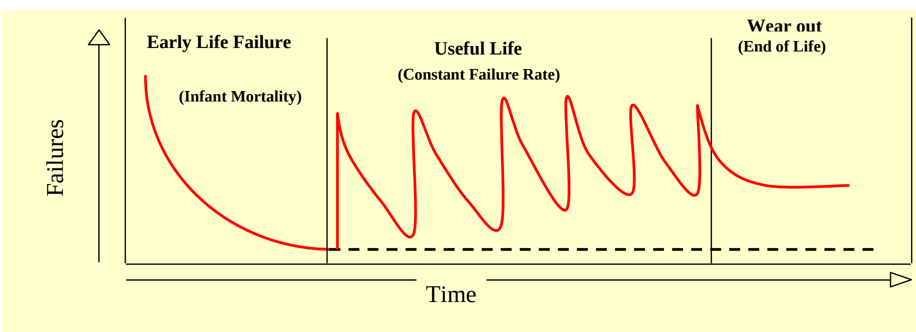
*The views of the author do not necessarily represent those of the Department of Defense

1

9th International Conference on Information Quality, 2004

Motivation #1

From a Total Data Quality Management (TDQM) standpoint, we are interested in capturing <u>failure modes</u> (data quality problems) in the software life cycle trend



Source: John Best, ETM 5291, Oklahoma State University

2

Motivation #2: Data quality problems in systems integration are, for the most part, captured via the SCR/PTR process, which is in serious need of improvement.

System Change Request (SCR) Form

TO: DSIO-EA (CM)	REQUESTOR PRIORITY:2	REQUESTED DATE: 08/04/2004	DSIO-EA (CM USE ONLY) DATE RECEIVED: ___ DATE CLOSED: ___
	PROBLEM AREA: OED - Oracle Energy Downstream		CONFIGURATION IDENTIFICATION NUMBER: ___-___-___-___
TITLE: LOC DEV - Upload Daily OPIS Prices			
ASSOCIATED NUMBERS (*TAR/DESC/Issue/Scenario/Support Magic*): DESC2516			
REQUESTOR: NAME (*last, first*): Thompson, Mik / Zahorchak, Theresa ORGANIZATION: DESC-P			PHONE COMMERCIAL: (703) 767-9378 DSN: ___-___ x___

3

List of Software Quality Risks*

Use cases: working features fail
Robustness: common errors are handled improperly
Performance: slow system performance
Localization: problems with time zones, currencies, etc.
Data integrity: dbase becomes corrupted/accepts improper data
Usability: software's interface is cumbersome or inexplicable
Volume/capacity: the system fails at peak or sustained loads
Reliability: at peak loads, the system crashes

We suggest that the data quality problems identified in our SCRs/PTRs are likely to be highly correlated with software quality risks

4

*Source: Rex Black, 2002

9th International Conference on Information Quality, 2004

What is missing from this SCR/PTR table?

Quality Risk	*Failure Mode(s)*	*Priority*
Functionality	Can't edit text.	1
	Can't format text.	1
	Can't handle tables.	2
	Can't insert pictures.	3
Performance	Display more than two keystrokes behind.	1
	File ops longer than two seconds for large typical file.	1
	File ops longer than five seconds for large atypical file.	3
Compatibility	Can't import Word files.	1
	Can't import WordPerfect files.	3

Lacks notions of risk

Severity: How dangerous is a failure of the system stemming from this area?
Priority: How much does a failure of the system in this area compromise the value of the product to customers and users?
Likelihood: What are the odds that a user will encounter a failure in this area?
Detection: What are the odds that a failure in this area will escape detection?

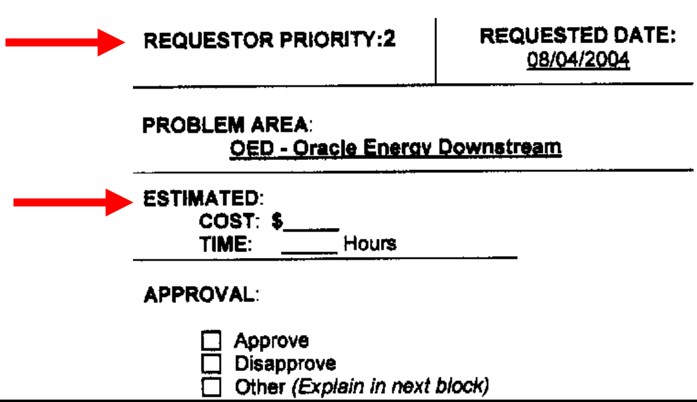
5

*Source: Rex Black, 2002

9th International Conference on Information Quality, 2004

Two Important Questions:

- How important is the PTR/SCR (e.g., is the priority number realistic?)
- Not only should we consider how good the fix is, but also inquire "how much is the fix truly worth to us"?


REQUESTOR PRIORITY:2

REQUESTED DATE: 08/04/2004

PROBLEM AREA:
OED - Oracle Energy Downstream

ESTIMATED:
COST: $____
TIME: _____ Hours

APPROVAL:

☐ Approve
☐ Disapprove
☐ Other *(Explain in next block)*

6

Problem Statement – how could we improve the very mechanism (SCR/PTR) that addresses data quality problems in systems integration?

✓ BPR + gap analysis = a way/path to create and rank System Change Requests and Problem Trouble Reports (SCRs/PTRs)

✓ A method to rank SCRs/PTRs by their relative risk importance through data quality engineering (FMEA)

✓ Have the capability to obtain reliable pecuniary costs estimates of SCRs/PTRs

7

9th International Conference on Information Quality, 2004

Objectives of this presentation

✓ Contribute to the literature on Information Quality (IQ)

✓ Showcase BPR & gap analysis efforts at a DoD agency

✓ Utilize the literature on FMEA and explain its usefulness to data quality

✓ Portray how might FMEA affect the bottom-line – the extent vendor-pricing estimates are likely biased

8

The Organization (DESC) – *Executive Agent* Candidate for Fuel Buys

- ✓ DESC handled over 5.5 million sales transactions in FY 2002. Net sales for FY 2002 topped $5.7 billion, and DESC awarded $ 6.2 billion in contracts in FY 2003.
- ✓ Consequently, a new automated fuel management system called **FAS** was obtained and implemented.

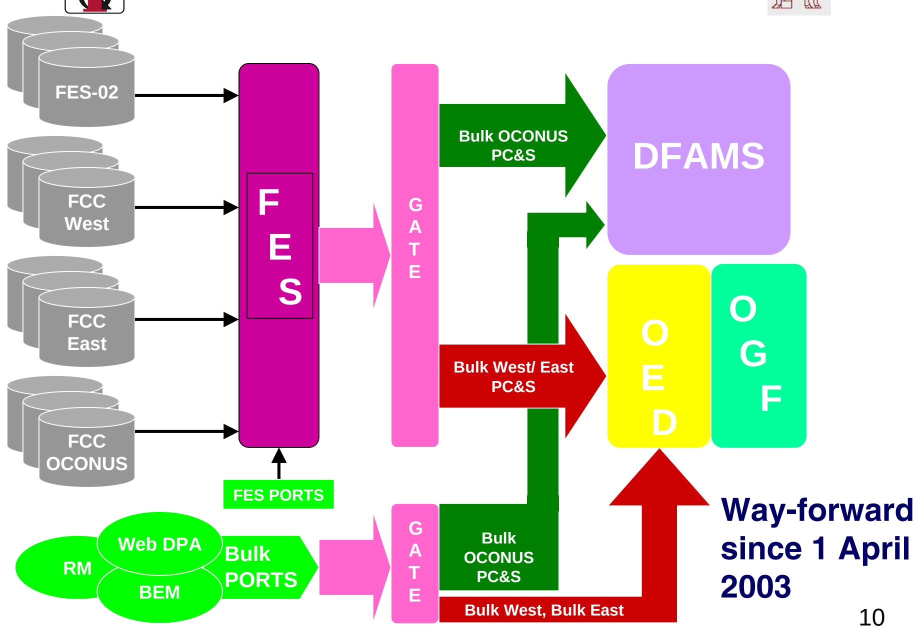
Oil Energy Downstream (OED) + Oracle Government Financial (OGF) + Other = Fuel Automated System (FAS)

- ✓ Systems Integration means transitioning from a Fortran-coded system to an ERP-based COTS, **FAS**.

Source: DESC Fact Book, 2003; Raneses and others (2004)

9

9th International Conference on Information Quality, 2004



10

Source: Petroleum Management Consultants (2004)

9th International Conference on Information Quality, 2004

TDQM literature emphasizes DMAIC Phases

- ✓ **D**efine the business process
- ✓ **M**easure the performance of the core business process
- ✓ **A**nalyze the process map to determine root causes of defects and means of improvement
- ✓ **I**mprove by designing sustainable solutions to fix and prevent problems
- ✓ **C**ontrol the improvements; keep new processes on check

11

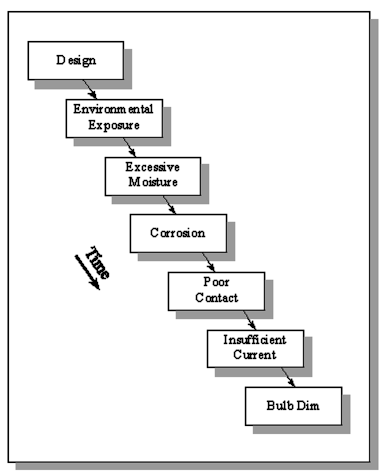
9th International Conference on Information Quality, 2004

Suggested Methodology: "Data Quality Engineering"/FMEA

FMEA is the study of failure mode and effects.

Failure Modes are sometimes described as categories of failure. A potential Failure Mode describes the way in which a product or process could fail to perform its desired function (design intent or performance requirements) as described by the needs, wants, and expectations of the internal and external Customers.

An **Effect** is an adverse consequence that the Customer might experience. The Customer could be the next operation, subsequent operations, or the end user.



Source: Haviland Consulting Group

12

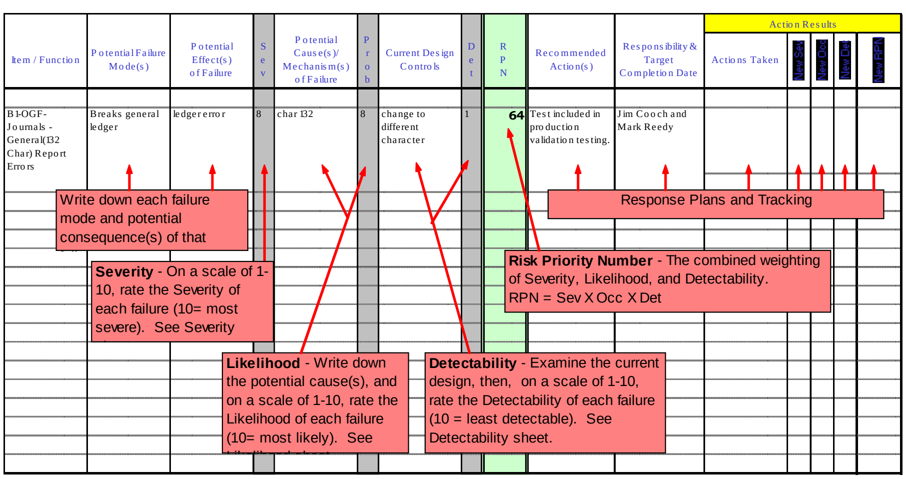
9th International Conference on Information Quality, 2004

Severity Effect	Probability of Failure	Detectability	Ranking
Hazardous without warning	Very High: Failure is almost inevitable	Absolute Uncertainty	10
Hazardous with warning	Medium High	Very Remote	9
Very High	High: Repeated failures	Remote	8
High	Low High	Very Low	7
Moderate	Moderate: Occasional failures	Low	6
Low	Low Moderate	Moderate	5
Very Low	High Low	Moderately High	4
Minor	Low: Relatively few failures	High	3
Very Minor	High Remote	Very High	2
None	Remote: Failure is unlikely	Almost Certain	1

13

9th International Conference on Information Quality, 2004

DMAIC + Risk Analysis = FMEA



14

9th International Conference on Information Quality, 2004

	Oracle Local Development Work in Progress (04/12/04)											
	TAR#	Description	Duration	Start	Finish	Need By	Phase Ne	Priority	Owner	Status	Develope
1	DESC1983	Returns Without Credit	1 day?	03/18/04	03/18/04	04/01/04	PC&S-B E	1000	Davis, Pat	BUS	Earl
2	DESC2516/F1	SIP/Price Series Code Upload	1 day	01/23/04	01/23/04	04/01/04	PC&S-B E	1000	Thompson	DEV	Earl
3	DESC2914	EDI Rollup/Breakout Process Change for Tax Rate	1 day?	03/30/04	03/30/04	04/30/04	PC&S-B E	1000	Coffel, Bet	LOG	Julius
4	DESC2785	Modification to the Contract Activity Report	5 days	02/09/04	02/13/04	02/27/04	Bulk Produ	600	Comar, Bil	LOG	Julius
5	DESC2844	Modification to the Contract Activity Report	3 days	02/04/04	02/06/04	03/01/04	Bulk Produ	600	Comar. Bil	LOG	Julius
6	DESC2907	Add Clin to Contract Activity Disbursement Report	1 day?	03/29/04	03/29/04	04/30/04	Production	600	Coffel, Bet	LOG	Julius
7	DESC2911	Timeouts and Return Values for Custom Programs	1 day?	03/30/04	03/30/04	04/20/04	Production	600	Smith, Ma	LOG	Earl
8	DESC2651	Lowest Projected Inventory includes Estimates	10 days	01/19/04	01/30/04	09/30/03	Bulk Produ	470	Barnett, Cl	LOG	Dave/Tracy
9	DESC1868	OED to FES interface validation Super User Book Inventory V	10 days	01/08/04	01/21/04	NA	Production	400	Cerda, Joh	LOG	Wally
10	DESC2852	Retiring of scripts	1 day	02/09/04	02/09/04	NA		400	Smith, Ma	LOG	Julius
11	DESC2728	OED Tanker Lift Schedule reports adding Lift Area	4 days	02/04/04	02/09/04	01/07/04	Bulk Produ	256	Barnett, Cl	LOG	Dave
12	DESC2722	BargeNet Report adding Sort parameter	2 days	02/02/04	02/03/04	01/07/04	Bulk Produ	255	Barnett, Cl	LOG	Dave
13	DESC2851	Issues Consumption Chart using Loc of Mvt Id vs Load L	2 days	02/09/04	02/10/04	02/09/04	Bulk Produ	235	Barnett, Cl	LOG	Tracy
14	DESC2397	Weighted Average Price (waiting on requirements)	0 days	01/22/04	01/22/04	03/31/04	Bulk Produ	230	Todd, Barb	BUS	Tracy
15	DESC2883	Transaction by Mvt-Disc Report	1 day?	03/10/04	03/10/04	03/10/04		230	Brooks, Al	LOG	Tracy
16	DESC2802	Need to Re-establish the load process from DAAS to the DOD	10 days	03/22/04	04/02/04	NA		225	Tolbert, Ri	LOG	Earl
17	DESC2841	Update DAAS Master File for Sales	5 days	02/04/04	02/10/04	05/15/04		225	Brooks, Al	LOG	Earl
18	DESC2908	Improve Inv Compare to Guides Upload	1 day?	03/29/04	03/29/04	05/03/04		220	Barnett, Cl	LOG	Dave
19	DESC2681	Daily Inventory and Movements improvements for book inv, rec	5 days	02/18/04	02/24/04	01/29/04	Bulk Produ	160	Barnett, Cl	LOG	Dave
20	DESC2301	Forecasting Sales for the Projected Inventory Report	1 day?	02/10/04	02/10/04	12/05/03	Bulk Produ	155	Barnett, Cl	LOG	Dave
21	DESC2687	Distribution System Projected Inventory Sheet Type	15 days	02/25/04	03/16/04	01/07/04	Bulk Produ	150	Barnett, Cl	LOG	Dave
22	DESC2694	OED Interface to run Overdue Tanker Moves Report	1 day?	03/17/04	03/17/04	02/02/04		150	Brooks, Al	LOG	Dave
23	DESC2739	New S/H version of the Unmatched Stock Transfers Report	3 days	01/30/04	02/03/04	02/02/04		150	Brooks, Al	LOG	Tracy
24	DESC2740	New S/H version of the Unmatched Purchases Report	3 days	02/04/04	02/06/04	02/02/04		150	Brooks, Al	LOG	Tracy
25	DESC2468	Redwood Report Repository Process Improvments	10 days	02/16/04	02/27/04	09/30/03		140	Weber, Joy	LOG	Julius
26	DESC2753	Multiple Combined Liability Report	2 days	02/10/04	02/11/04	03/15/04		140	Brooks, Al	DEV	Tracy
27	DESC2834	Report Gating History Report	3 days	01/29/04	02/02/04	01/30/04	Production	140	Brooks, Al	LOG	Tracy
28	DESC2821	Modification to Remittance Address batch job	3 days	02/04/04	02/06/04	03/15/04	Production	138	Comar, Bil	LOG	Earl
29	DESC2830	Quarterly Sales Quarter Definitions and Cross tab re	3 days	01/22/04	01/26/04	03/15/04		137	Barnett, Cl	LOG	Tracy
30	DESC2689	Issues Consumption Chart add Region and Distn Sys	2 days	02/12/04	02/13/04	02/11/04	Bulk Produ	135	Barnett, Cl	LOG	Tracy
31	DESC2896	Modification to Fuel Receipts Summary Detail tab	1 day?	03/22/04	03/22/04	NA		134	Barnett, Cl	LOG	Tracy

From gap analysis, rank failure mode by priority

15

9th International Conference on Information Quality, 2004

Limitations of FMEA

- ✓ Examination of human error is limited
- ✓ Focus is on single-event initiators of problems
- ✓ RPNs (risk priority numbers), on their own, are not enough (e.g. not **volume or** economically-based)
- ✓ Other more sophisticated methodologies such as research work at MIT (e.g., Navarro and others, 2000) address the reduction the effects of requirements changes through system design

But suppose we could use FMEA data (e.g. RPN) to obtain better costs estimates…this could likely impact the bottom line

Source: John Best, ETM 5291, Oklahoma State University

16

Expected Cost Model

Suppose that:

- ✓ p_f = probability of a fault
- ✓ p_d = probability that it escapes detection

Assuming that p_f and p_d are independent, the probability that the user receives the data quality problem or defect is = p_f * p_d

- ✓ These probabilities can be estimated from the DESC Help Desk/ASG:
- ✓ production records
- ✓ customer records
- ✓ sample inspection results

Source: John Best, ETM 5291, Oklahoma State University

9th International Conference on Information Quality, 2004

Expected Cost Model

- ✓ Estimate cost per fault ***C***
- ✓ Start with rough estimate based on
  - Internal scrap, rework
  - Warranty costs
  - Other Cost-Of-Poor-Quality factors
  - **Proxy for data quality failure (e.g. manual work-around)**
- ✓ If *n* items are produced (yearly, monthly)
- ✓ Expected cost of data quality line items (PTRs/SCRs):

$$EC = Cn*p_f*p_d$$

Source: John Best, ETM 5291, Oklahoma State University

9th International Conference on Information Quality, 2004

Estimating Occurrence

Fault Occurrence Probability		Data Quality1	Data Quality2	Data Quality3
5/10	0.5			
1/10	0.1			
5/100	0.05			
1/100	0.01	X		
5/1000	0.005			
1/1000	0.001			X
5/10,000	0.0005		X	
1/10,000	0.0001			
5/100,000	0.00005			
1/100,000	0.00001			
5/1,000,000	0.000005			

Source: John Best, ETM 5291, Oklahoma State University

19

9th International Conference on Information Quality, 2004

Estimating Undetection

Undetection (escape) Probability	Data Quality1	Data Quality2	Data Quality3
10/10			
8/10			
6/10			
3/10			
1/10	X		
5/100		X	
1/100			
5/1000			X
1/1000			
1/10,000			

Source: John Best, ETM 5291, Oklahoma State University

20

Total Expected Cost Comparison

Costs

Failure Mode	Cost per Item	Month Volume	Probabilities Occurrence	Probabilities Detection	Expected Cost
SCR1	50	20,000	0.01	0.1	$1,000
SCR2	100	80,000	0.0005	0.05	$200
SCR3	30	100,000	0.001	0.005	$15
				Total Cost	$1,215

Source: John Best, ETM 5291, Oklahoma State University

21

9th International Conference on Information Quality, 2004

Results

Data Quality Capture + Cost Effectiveness Tool

BPR ***reveals data quality problems by way of gap analysis; leads to SCR/PTR documentation***

FMEA ***allows the user to incorporate risk in SCR/PTR costing***

Business Process	Gap Analysis	Nominal Cost ($)	Estimated Cost ($)
Ordering	Suspension SCR	biased/unbiased	closer to true costs
Receipting	Payment problems	biased/unbiased	closer to true costs
Pricing	Failed escalator	biased/unbiased	closer to true costs
Taxes	Tax change	biased/unbiased	closer to true costs

22

9th International Conference on Information Quality, 2004

Further Work

Next Steps

BPR & gap analysis were large steps for DESC, but it is just the beginning...

- ✓ **Gather and compile the data (e.g. collect** *n* items that are produced yearly, monthly, etc.) for the purposes of **Data Quality Capture**
- ✓ **Utilize Application System Group (ASG)/Help Desk Ticket info**
  - Direct Delivery and Bulk Fuel transactions
  - Inventory/Stock Control statistics
  - Other operational work
- ✓ **Integrate Data Quality Engineering (FMEA) in Gap Analysis**
  - Conduct frequency analysis
  - Use sensitivity analysis

23

9th International Conference on Information Quality, 2004

Summary

Problem Statement
How could we improve the very mechanism (SCR/PTR) that addresses data quality problems in systems integration?

Methodology
BPR + Gap Analysis = Part of the solution
Data Quality Engineering (FMEA) = Proposed solution

Results
Incorporates notion of risk in PTR/SCR costs
Likely decreases biased estimates

Further work/Challenges at DESC
This is merely a process presentation. The challenge ahead is in the execution of the proposed methodology.

References

[1] Best, J. ***ETM 5291 - Failure Mode & Effects Analysis***. **OSU, 2001.**

[2] Black, R. ***The Risks to System Quality: Investing in Software Testing***. **2002. http://www.stickyminds.com.**

[3] DESC. ***DESC Factbook 2003***. **Defense Energy Support Center, 2003.**

[4] General Electric Company. ***GE Investor Relations Annual Reports***. **22 July 2002.**

[5] Navarro, I., Leveson, K, and K. Lundqvist. ***Reducing the Effects of Requirements through System Design***. **MIT, Cambridge, MA, 2000.**

[6] Haviland Consulting Group. ***FMECA***. **http://www.fmeca.com/**

[7] Petroleum Management Consultants, Inc. ***PC&S FAS and Management Information Overview***. **PowerPoint slide presentation. Defense Energy Support Center, 2004.**

[8] Raneses, A; Thompson, Mik; Rayburn, K; and G. Winston. ***Business Process Re-Engineering at DESC***. **Powerpoint slide presentation, Defense Energy Support Center, 2004.**