

# 2009 Solid Data Foundation Webinar Series #1

## Building Solid Data Foundations for Successful Business Intelligence Implementations



**John Poonnen**

**Senior Vice President  
GDI Infotech Solutions  
Division**

**Jim Diggs**

**Corporate Development  
GDI Infotech**

# Today's Panelists

## ❖ John Poonen – Guest Speaker

**Senior Vice President, GDI Infotech Solutions Division**

- **25 years of international IT consulting, program management, product management and systems engineering experience**
- **Led many Data Management initiatives for Fortune 500 companies with strong focus on Data Warehousing (DW), Business Intelligence (BI), Information Architecture**
- **Former VP of Health Care Payer Practice at CSC-Covansys and Director of Information Strategy/Enterprise Architecture at Blue Cross Blue Shield of North Carolina.**

## ❖ Jim Diggs – Host

**Corporate Development, GDI Infotech**

- **30 years in technology companies**
- **Former VP, Janeeva, ADP, IBM, BlueGill Technologies**

# **The Solid Data Foundation Webinar Series #1 Objectives:**

- ❖ Introduce the Principles of building a Solid Data Foundation**
- ❖ Learn the major data challenges of Business Intelligence implementations**
- ❖ Learn how Data Quality and Profiling can lead to a successful BI implementation**

# The Solid Data Foundation

## The five tenets of a Solid Data Foundation










- 1. Data quality is baselined, assessed and continually monitored across key dimensions**
- 2. Data issues and anomalies are known early enough, so that risks can be addressed**
- 3. Data issues are remediated and data quality is enhanced**
- 4. Data is ready to be integrated into one authoritative place**
- 5. Data Quality principles and processes are expanded enterprise-wide with Data Governance and Data Stewardship**



 COMPLETENESS  CONFORMITY  ACCURACY  DUPLICATION  INTEGRITY  CONSISTENCY

# Why build Solid Data Foundations...Business Drivers

- ❖ **The most common business drivers cited were....**
  - Shrinking timelines to implement BI in the face of competition
  - Capped cost of BI delivery – get it right first time
  - Increased breadth of information – more complexity of data
  - Need for information consistency – across the enterprise
  - Manage spiralling data growth and reduce data redundancy
  - Realize savings due to consolidation efforts and “reuse” of data
  
- ❖ **The typical organization has multiple data silos that are out of control....need to reduce costs**
  - Often distributed globally across many business units
  - Many DWs/DMs may be redundant and cost \$\$ to keep in sync
  - Durable business success directly depends on building SDFs.

# Five Common Data Challenges in Business Intelligence Implementation

1. **Lack of standards**  
2. **Hidden Data surprises**  
3. **Information buried in silos** 
4. **Data myopia – or the lack of integrated data**
5. **The data anomalies nightmare**    

 COMPLETENESS    CONFORMITY    ACCURACY    DUPLICATION    INTEGRITY    CONSISTENCY

# Do Not Underestimate The Impact Of Data Quality On the Business!

## ❖ A key example – Inflated Home Value

- Effect : County Budget Slashed
- Effect: Months before root cause is discovered
- Effect: Irreparable harm is done

## ❖ Remedy: Build a Solid Data Foundation



GIGO – Garbage in -Garbage out

# Representative View Of Data Quality Problems

Erroneous entry of inflated house value

District	Block	Lot	Owners Name	Property Location	City, State	Zip	Net Taxable Value
9	312	8	BROWN ALFREDA & CHARLES JR	77 PRINCETON AVE	ANN ARBOR, MI	48105	\$239,100
24	8303	24	BROWN, ADRIENNE	37 BUDD ST	ANN ARBOR, MI	48113	\$277,000
29	443	7	BROWN, A J & D E	9 MAIN DR	ANN ARBOR, MI	48106	\$332,500
19	127	11	BROWN RICHARD H III & MARGARET Y	1 POST LN	ANN ARBOR, MI	48105	\$400,000,000
10	42	18	DAVIS / DAVIS, WILLIAM & VIRGINIA	11 KENNEDY TER	ANN ARBOR, MI	48103	\$294,900
9	1102	14	DAVIS WILLIAM L/HELEN J	7BAKER AVE	ANN ARBOR, MI	48104	\$206,300
9	1807	26	DAVIS WALTER W & RUTH E	19 NEW ST	ANN ARBOR, MI	48105	\$1,019,000
9	310	1	DAVIS, AGNES MRS.	27 AUDREY PL	ANN ARBOR, MI	48107	\$399,000
38	1	11	DAVIS, A L & D D DOCKEY	54 EAST AVE	ANN ARBOR, MI	48108	\$182,800
9	1102	14	DAVIS WILLIAM L/HELEN J	SEVEN BAKER ST	ANN ARBOR, MI	48104	\$206,300
7	8	14	DAVIS, ALFRED G/MICHELE	SHEEPHILL DR	ANN ARBOR, MI	48113	\$262,200
19	141	9	JONES , THOMAS/KELLY B	15 DOGWOOD DR	ANN ARBOR, MI	48107	\$300,400
38	43	9	JONES, ALAN S & SUSAN L	127 KINGS HWY	ANN ARBOR, MI	48109	\$309,600
19	144	35	Attn: JONES, ALTON C	5 GOLF LN	ANN ARBOR	48103	\$206,300
29	265	2	JONES, ANITA A & ALBERT C JONES	205 ATLANTIC DR	ANN ARBOR, MI	48106	\$323,100
1	105	14	SMITH, CATHERINE ANNE	496 MORRIS AVE	ANN ARBOR, MI	48109	\$400,700
1	36	4	SMITH, EDWARD J/CATHERINE	117 LAKE AVE	ANN ARBOR, MI	48113-1113	\$317,100
1	34	1	SMITH	312 MAIN ST	ANN ARBOR, MI	48103	\$277,000
1	83	1	SMITH, GEORGE R/KATHLEEN S	232 HOLMES ST	ANN ARBOR, MI	99999	\$205,300
1	69	12	SMITH, RICHARD/DEBRA G	243 ROOSEVELT AVE	UNKNOWN	48108	\$197,900
9	712	12	Williams Elizabeth J	106 BAKER ST	ANN ARBOR, MICH.	48113	\$320,100
9	89	11	WILLIAMS RODNEY & ANDREA	43 LIVINGSTON AVE	DETROIT, MI	48108	\$799,000
19	131	9	WILLIAMS, ADRIAN & MARGARET C	62 WEST MAIN ST	ANN ARBOR, MI	48106	\$229,000
35	9406	9	WILLIAMS, ALAN L JR & KRISTINE L	66 STEPHEN PL	ANN ARBOR, MI		\$199,400
36	1160	44	WILLIAMS, ALBERT G	8 YELLOW BARN AVE	ANN ARBOR, MI	48108	\$599,000

■ COMPLETENESS 
 ■ CONFORMITY 
 ■ ACCURACY 
 ■ DUPLICATION 
 ■ INTEGRITY 
 ■ CONSISTENCY

# Anomalies in Data can sink any BI Initiative

*Who is our best customer and what is our revenue from them last month? What can we sell them next? When?*

CUSTNUM	NAME	ADDRESS	SALES \$
90328574	IBM	187 N.Pk. Str. Salem NH 01456	8,494.00
90328575	I.B.M. Inc.	187 N.Pk. St. Sarem NH 01456	3,432.00
90238495	International Bus. M.	187 No. Park St Salem NH 04156	2,243.00
90233479	Int. Bus. Machines	187 Park Ave Salem NH 04156	5,900.00
90233489	Inter-Nation Consults	15 Main St. Andover MA 02341	6,800.00
90234889	Int. Bus. Consultants	PO Box 9 Boston MA 02210	10,243.00
90345672	I.B. Manufacturing	Park Blvd. Boston MA 04106	15,999.00

No common key

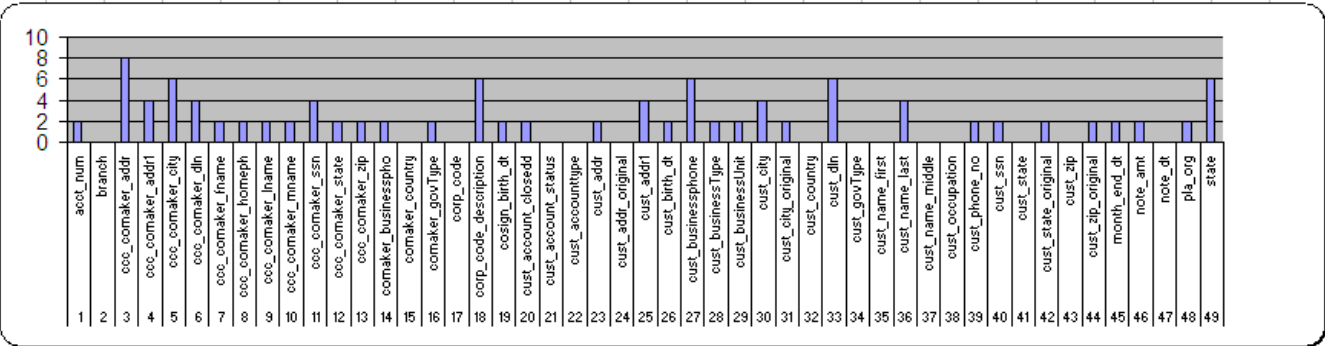
Anomalies

Lack of Standards

Spelling Errors

# Data Profiling: Data Quality Complex Analysis

Field Name	acc_num	branch	ccc_comaker_addr1	ccc_comaker_addr1	ccc_comaker_city	ccc_comaker_dln	ccc_comaker_fname	ccc_comaker_homphone	ccc_comaker_hname	ccc_comaker_mname	ccc_comaker_ssn	ccc_comaker_state	ccc_comaker_zip	comaker_businesspho	comaker_country	comaker_govType	corp_code	corp_code_description_aka_legal_v	ccosign_birth_dt	ccosign_birth_dt	ccosign_aka_legal_vehicle	cust_account_closeddate	cust_account_status	cust_accounttype	cust_addr
Uniqueness	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Completeness	Green	Green	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red
Conformity	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
Structure	Green	Green	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red	Red
Validity	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green



Order on the report/  
in the file

Fields  
Cumulative DQ Index

Critical

Warning

Clean



# This is a challenging time... the Right Time to build Solid Data Foundations

- ❖ **Applications need to be delivered quicker**
  - Applications in the midst of development will face data quality related problems which will delay and jeopardize their completion
- ❖ **Business Intelligence requirements keep growing**
  - Financial Services, Insurance, Healthcare and others are under growing pressure for better data quality to drive operations and marketing effectiveness
- ❖ **Compliance, Regulatory and Risk pressure**
  - The financial stability and business exposures need better data quality for reporting and compliance
- ❖ **Mergers and Acquisitions are growing**
  - M&A will drive even more data and information integration projects increasing the need for better data quality
- ❖ **Data Management Resources are limited**
  - Much of client SME time continues to be consumed by resolving data problems and not addressing tasks critical for business success

***The time is now.***

***...build Solid Data Foundations for BI***

# Data Quality Audit & Profiling (DQAP): The Vital Step in Building Solid Data Foundations



- **Audit and profile business data to get a holistic view**
- **Identify triage and evaluate Data Quality anomalies**
- **Assess and develop action plan for resolving issues that undermine business processes.**