


The Common Data Architecture in Perspective, Michael Brackett, Data Resource Design and Remodeling




THE COMMON DATA ARCHITECTURE

IN PERSPECTIVE

DAMA INTERNATIONAL CONFERENCE
APRIL 25, 2006


MICHAEL H. BRACKETT

CONSULTING DATA ARCHITECT
DATA RESOURCE DESIGN & REMODELING



P.O. Box 151
Lilliwaup, Washington 98555
Voice: (360) 877 5507
Fax: None
mhbrackett@aol.com
members.aol.com/mhbrackett

© 2006 by Michael H. Brackett




CONTENTS

- ◆ *CURRENT SITUATION*
- ◆ *DATA RESOURCE CONCEPTS*
- ◆ *HALTING DATA DISPARITY*
- ◆ *RESOLVING DISPARATE DATA*
- ◆ *CONCLUSION & ANNOUNCEMENT*
- ◆ *PRESENTER'S BIBLIOGRAPHY*

© 2006 by Michael H. Brackett

The DAMA International Symposium & WILSHIRE Meta-Data Conference
Denver, Colorado ● April 23-27, 2006



CURRENT SITUATION

◆ **BUSINESS INFORMATION DEMAND**
AN ORGANIZATION'S CONTINUOUSLY INCREASING, CONSTANTLY CHANGING NEED FOR CURRENT, ACCURATE, INTEGRATED INFORMATION, OFTEN ON SHORT NOTICE OR VERY SHORT NOTICE, TO SUPPORT ITS BUSINESS ACTIVITIES

◆ **CURRENT BUSINESS IMPACT**


- THIS BUSINESS INFORMATION DEMAND IS NOT BEING MET
- PUBLIC & PRIVATE SECTOR BUSINESS IS IMPACTED
- BUSINESS DECISIONS ARE LESS-INFORMED
- CITIZENS & CUSTOMERS ARE IMPACTED
- OPPORTUNITIES ARE MISSED

◆ **PROBLEM**

- HIGH QUANTITIES OF DISPARATE DATA IN MOST ORGANIZATIONS
- DATA NOT THOROUGHLY UNDERSTOOD
- DATA RESOURCE NOT FORMALLY MANAGED
- DATA NOT READILY AVAILABLE
- DATA QUALITY LOW

BUSINESS INFORMATION DEMAND NOT BEING MET!

© 2006 by Michael H. Brackett



CURRENT SITUATION

◆ **DISPARATE DATA - A TRUISM**
DATA THAT ARE ESSENTIALLY NOT ALIKE, OR ARE DISTINCTLY DIFFERENT IN KIND, QUALITY, OR CHARACTER. THEY ARE UNEQUAL AND CANNOT BE READILY INTEGRATED TO ADEQUATELY MEET THE BUSINESS INFORMATION DEMAND.


◆ **BASIC DATA PROBLEM**

- UNKNOWN DATA EXISTANCE
 - ORGANIZATION NOT AWARE OF ALL DATA AT ITS DISPOSAL
 - USUALLY NOT EVEN INVENTORIED
- UNKNOWN DATA MEANING
 - CONTENT AND MEANING OF DATA NOT FULLY KNOWN
 - DATA NOT THOROUGHLY UNDERSTOOD
- HIGH DATA REDUNDANCY
 - DATA HIGHLY REDUNDANT AND INCONSISTENT
 - AVERAGE REDUNDANCY FACTOR OF 10 FOR LARGE ORGANIZATIONS
- HIGH DATA VARIABILITY
 - DATA HIGHLY VARIABLE IN FORMAT AND CONTENT
 - AVERAGE FACTOR OF 15 TO 20 FOR LARGE ORGANIZATIONS

THAT'S THE GOOD NEWS!

© 2006 by Michael H. Brackett

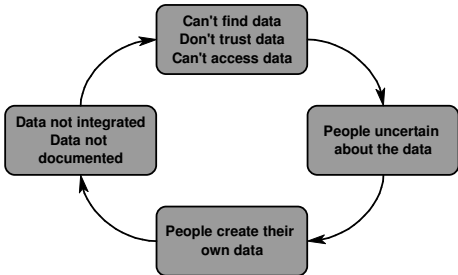
The Common Data Architecture in Perspective, Michael Brackett, Data Resource Design and Remodeling



CURRENT SITUATION


- ◆ **DISPARATE DATA RESOURCE**
A DATA RESOURCE THAT IS SUBSTANTIALLY COMPOSED OF DISPARATE DATA THAT ARE DIS-INTEGRATED AND NOT SUBJECT ORIENTED. A STATE OF DISARRAY WHERE THE LOW QUALITY DOES NO, AND CANNOT, ADEQUATELY SUPPORT THE BUSINESS INFORMATION DEMAND.

- ◆ **DISPARATE DATA CYCLE**
A SELF-PERPETUATING CYCLE WHERE DISPARATE DATA CONTINUE TO BE PRODUCED AT AN EVER-INCREASING RATE BECAUSE PEOPLE DO NOT KNOW ABOUT EXISTING DATA OR DO NOT WANT TO USE EXISTING DATA.



THAT'S THE BAD NEWS!

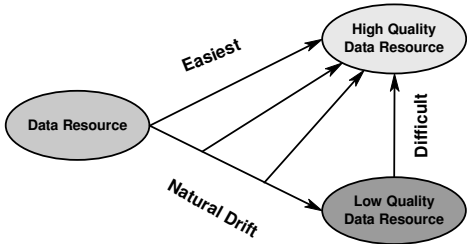
© 2006 by Michael H. Brackett



CURRENT SITUATION


- ◆ **DISPARATE DATA SPIRAL**
 - DISPARATE DATA INCREASING BY SEVERAL ORDERS OF MAGNITUDE
 - CREATING DISPARATE DATA FASTER THAN EVER BEFORE
 - SPREADING TO NON-TABULAR DATA AND DATA RESOURCE DATA
 - DISPARITY SPIRALING OUT OF CONTROL
 - NO FORESEEABLE END IN SIGHT!

- ◆ **DATA RESOURCE DRIFT**
THE NATURAL, STEADY DRIFT OF A DATA RESOURCE TOWARDS DISPARITY IF ITS DEVELOPMENT IS NOT PROPERLY MANAGED AND CONTROLLED.



THAT'S THE REALLY BAD NEWS!

© 2006 by Michael H. Brackett



CURRENT SITUATION

- ◆ **NO STATUS QUO**
 - THERE IS NO STATUS QUO FOR DEVELOPING A HIGH-QUALITY DATA RESOURCE
 - NATURAL DRIFT TOWARD DISPARITY WILL CONTINUE
 - SITUATION WILL GET WORSE

**A STATUS QUO LEADS TO
ORGANIZATIONAL FAILURE BY INFORMATION DEPRIVATION!**


- ◆ **SILVER BULLETS JUST DO NOT EXIST**

OBJECTIVE: *AN ATTEMPT TO ACHIEVE SOME GAIN WITHOUT ANY PAIN*
RESULT: *ENDURING CONSIDERABLE PAIN WITH MINIMAL GAIN*

THERE ARE NO SILVER BULLETS!

- ◆ **MUST CONSCIOUSLY ALTER THE NATURAL DRIFT**
 - TOWARD HIGH-QUALITY DATA RESOURCE
 - BREAK DISPARATE DATA CYCLE
 - STOP THE SPIRALING DISPARITY
 - CREATE A HIGH-QUALITY SHARABLE DATA RESOURCE

© 2006 by Michael H. Brackett




DATA RESOURCE CONCEPTS

- ◆ **SECOND LAW OF THERMODYNAMICS**
 - ENTROPY IS THE STATE OR DEGREE OF DISORDERLINESS
 - ENTROPY IS LOSS OF ORDER – INCREASING DISORDERLINESS
 - ENTROPY INCREASES OVER TIME
 - THINGS BECOME MORE DISORDERLY OVER TIME
 - MANY EVERYDAY EXAMPLES WHERE ORDER GIVES WAY TO DISORDERLINESS
 - IN A CLOSED SYSTEM
 - ORDERED ENERGY CAN BE CONSUMED ENTIRELY INTO DISORDERED ENERGY
 - BUT NEVER VISA VERSA
 - PRINCIPLE OF *INCREASING ENTROPY*
 - SYSTEM REACHES STATE OF MAXIMUM ENTROPY - *AN EQUILIBRIUM*
 - LAW IS INEVITABLE AND IRREVERSIBLE FOR CLOSED SYSTEMS
 - IN AN OPEN SYSTEM
 - ENTROPY CONTINUES TO INCREASE
 - AN EQUILIBRIUM IS NOT REACHED
 - ENTROPY CAN BE REVERSED WITH INPUT OF ENERGY
 - MUST DO WORK TO REDUCE ENTROPY AND INCREASE ORDERLINESS

DATA RESOURCE IS AN OPEN SYSTEM

© 2006 by Michael H. Brackett



DATA RESOURCE CONCEPTS

- ◆ **DISPARITY AND ENTROPY**
 - DISPARITY CAN BE EQUATED TO ENTROPY
 - DISPARITY IN THE DATA RESOURCE IS EQUATED TO INCREASING ENTROPY


DISPARITY = DISORDERLINESS = INCREASING ENTROPY

- ◆ **DISPARITY AND ENERGY**
 - RESOLUTION OF DISPARITY CAN BE RELATED TO ENERGY
 - RESOLUTION OF DISPARITY REQUIRES ENERGY
- ◆ **COST AND ENERGY**
 - EFFORT CAN BE EQUATED TO ENERGY
 - ENERGY CAN BE EQUATED TO COST
 - REDUCTION OF ENTROPY / DISPARITY HAS AN ASSOCIATED COST

EFFORT = ENERGY = COST

RESOLUTION OF DISPARITY COMES AT A COST

© 2006 by Michael H. Brackett




DATA RESOURCE CONCEPTS

- ◆ **'DATA' - 'INFORMATION' - 'KNOWLEDGE' STILL CONFUSED**
 - DATA
 - INDIVIDUAL FACTS OUT OF CONTEXT WITH LITTLE MEANING
 - 123.45
 - DATA IN CONTEXT
 - INDIVIDUAL FACTS WITH MEANING
 - ACCOUNT BALANCE AT NOON ON JANUARY 12, 2001
 - INFORMATION
 - SET OF DATA IN CONTEXT
 - RELEVANT TO ONE OR MORE PEOPLE
 - AT A POINT IN TIME OR FOR A PERIOD OF TIME
 - KNOWLEDGE
 - INFORMATION RETAINED BY INDIVIDUALS
 - COMBINED WITH EXPERIENCE

DATA ARE THE FOUNDATION FOR INFORMATION & KNOWLEDGE


© 2006 by Michael H. Brackett

The Common Data Architecture in Perspective, Michael Brackett, *Data Resource Design and Remodeling*



DATA RESOURCE CONCEPTS

◆ **BUSINESS INTELLIGENCE VALUE CHAIN**




◆ **I-ORGANIZATION**

- INTELLIGENT LEARNING ORGANIZATION - HUMAN RESOURCE REALM
- HUMAN RESOURCE REALM LINKS TECHNOLOGY REALM & BUSINESS REALM
- REQUIRES GOOD INFORMATION TO SUPPORT BUSINESS STRATEGIES & GOALS
- REQUIRES A HIGH-QUALITY SHARABLE DATA RESOURCE

MUST BEGIN WITH THE DATA RESOURCE!

© 2006 by Michael H. Brackett



DATA RESOURCE CONCEPTS


◆ **QUALITY DEFINED**

- **DATA RESOURCE QUALITY**
A MEASURE OF HOW WELL THE DATA RESOURCE SUPPORTS THE CURRENT AND THE FUTURE BUSINESS INFORMATION DEMAND
- **DATA QUALITY**
A SUBSET OF DATA RESOURCE QUALITY DEALING WITH DATA VALUES
- **ULTIMATE DATA RESOURCE QUALITY**
A DATA RESOURCE THAT IS STABLE ACROSS CHANGING BUSINESS AND CHANGING TECHNOLOGY SO IT CONTINUES TO SUPPORT THE CURRENT AND FUTURE BUSINESS INFORMATION DEMAND
- **INFORMATION QUALITY**
A MEASURE OF THE ABILITY TO GET THE RIGHT DATA, TO THE RIGHT PEOPLE, IN THE RIGHT PLACE, AT THE RIGHT TIME, IN THE RIGHT FORM, AT THE RIGHT COST, SO THEY CAN MAKE THE RIGHT DECISIONS, AND TAKE THE RIGHT ACTIONS.

**INTELLIGENT LEARNING ORGANIZATION
NEEDS A HIGH-QUALITY STABLE DATA RESOURCE!**

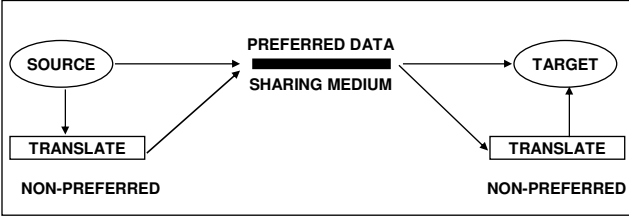
© 2006 by Michael H. Brackett

The Common Data Architecture in Perspective, Michael Brackett, Data Resource Design and Remodeling



DATA RESOURCE CONCEPTS

◆ **DATA SHARING VISION**
UNDERSTAND DATA IN A COMMON CONTEXT SO THEY CAN BE READILY SHARED TO MEET THE BUSINESS INFORMATION DEMAND




◆ **DATA SHARING PROCESS**

- DATA SHARED THROUGH PREFERRED VARIATIONS
- SOURCE ORGANIZATION TRANSLATES TO PREFERRED IF NECESSARY
- TARGET TRANSLATIONS TRANSLATES FROM PREFERRED IF NECESSARY
- SOURCE AND TARGET CAN CHANGE INDEPENDENTLY

SHARING PREFERRED DATA REQUIRES A THOROUGH UNDERSTANDING OF DATA!

© 2006 by Michael H. Brackett

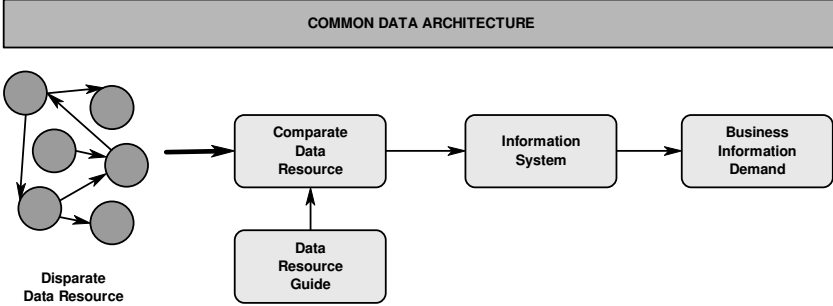


DATA RESOURCE CONCEPTS

◆ **COMPARE DATA VISION**


- COMPARE DATA
DATA THAT ARE ALIKE, SIMILAR IN KIND, QUALITY, AND CHARACTER, ARE EASILY UNDERSTOOD, AND CAN BE READILY INTEGRATED
- COMPARE DATA RESOURCE
SUBJECT ORIENTED, INTEGRATED, HIGH QUALITY, SHARABLE DATA RESOURCE THAT FULLY SUPPORTS THE CURRENT AND THE FUTURE BUSINESS INFORMATION DEMAND

COMMON DATA ARCHITECTURE



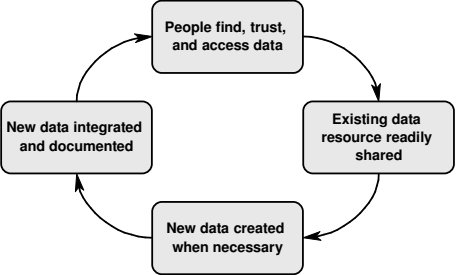
© 2006 by Michael H. Brackett

The Common Data Architecture in Perspective, Michael Brackett, Data Resource Design and Remodeling



DATA RESOURCE CONCEPTS

◆ **COMPARETE DATA CYCLE**
A SELF-PERPETUATING CYCLE WHERE THE USE OF COMPARTE DATA IS CONTINUALLY REINFORCED BECAUSE PEOPLE UNDERSTAND AND TRUST THE DATA.




```

graph TD
    A[People find, trust, and access data] --> B[Existing data resource readily shared]
    B --> C[New data created when necessary]
    C --> D[New data integrated and documented]
    D --> A
        
```

**STOP THE NATURAL DRIFT OF A DATA RESOURCE
 BY HALTING THE DISPARATE DATA CYCLE
 AND STARTING A COMPARTE DATA CYCLE!**

© 2006 by Michael H. Brackett



DATA RESOURCE CONCEPTS

◆ **COMPARETE DATA RESOURCE**

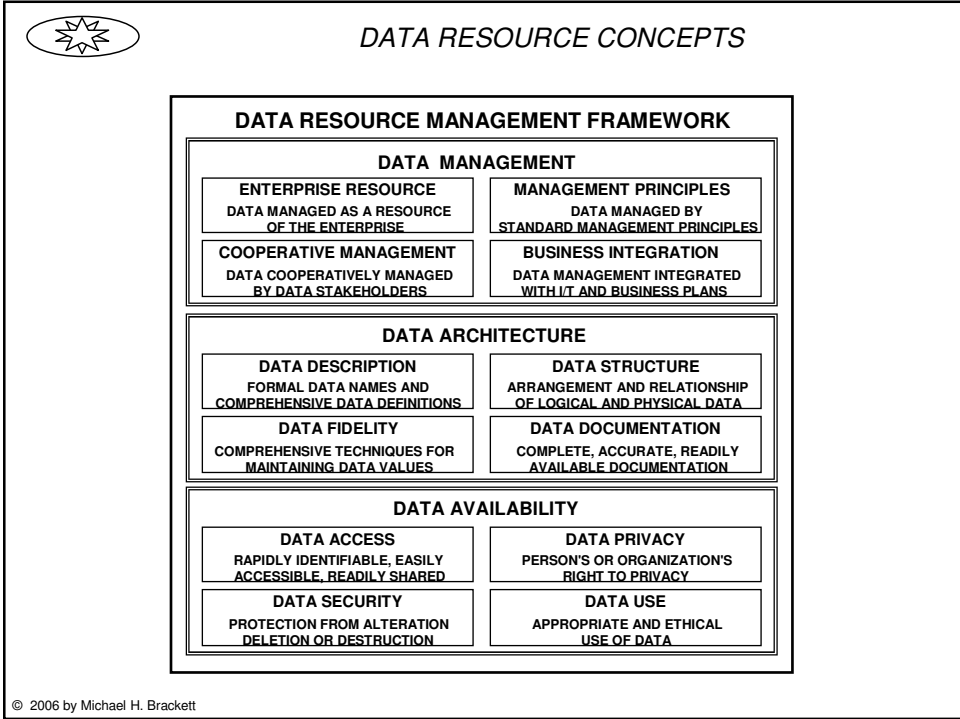
- **SUBJECT ORIENTED**
 - BASED ON BUSINESS OBJECTS AND EVENTS IN THE REAL WORLD
 - DATA SUBJECTS REPRESENT BUSINESS OBJECTS AND EVENTS
 - DATA RESOURCE STRUCTURED BY DATA SUBJECTS
 - ALL CHARACTERISTICS ABOUT A DATA SUBJECT STORED WITH THE DATA SUBJECT
- **INTEGRATED**
 - FULLY INTEGRATED WITHIN A COMMON DATA ARCHITECTURE
 - ONE ARCHITECTURE FOR THE ENTIRE DATA RESOURCE
 - ALL DATA MANAGED WITHIN THAT SINGLE ARCHITECTURE
 - DATA PROPERLY DEPLOYED TO MEET BUSINESS NEEDS
- **SUPPORTED BY DATA RESOURCE GUIDE**
 - COMPREHENSIVE INDEX TO DATA RESOURCE
 - A COMPLETE INTEGRATED UNDERSTANDING ABOUT THE DATA RESOURCE
- **INTEGRATED**
 - ACROSS ORGANIZATION UNITS
 - ACROSS BUSINESS ACTIVITIES

**AN INTEGRATED DATA RESOURCE
 IS ONE VERSION OF TRUTH ABOUT THE BUSINESS!**

© 2006 by Michael H. Brackett

The DAMA International Symposium & WILSHIRE Meta-Data Conference
 Denver, Colorado ● April 23-27, 2006

The Common Data Architecture in Perspective, Michael Brackett, Data Resource Design and Remodeling




DATA RESOURCE CONCEPTS

- ◆ **DATA ARCHITECTURE DEFINITION 1**
THE METHOD OF DESIGN AND CONSTRUCTION OF A DATA RESOURCE THAT IS BUSINESS DRIVEN, BASED ON REAL-WORLD SUBJECTS PERCEIVED BY THE ENTERPRISE, AND IMPLEMENTED INTO APPROPRIATE OPERATING ENVIRONMENTS. IT CONSISTS OF COMPONENTS THAT PROVIDE A CONSISTENT FOUNDATION ACROSS ORGANIZATIONAL BOUNDARIES TO PROVIDE EASILY IDENTIFIABLE, READILY AVAILABLE HIGH-QUALITY DATA TO SUPPORT THE BUSINESS INFORMATION DEMAND

- ◆ **DATA ARCHITECTURE DEFINITION 2**
THE COMPONENT OF THE DATA RESOURCE FRAMEWORK THAT CONTAINS ALL THE ACTIVITIES, AND THE PRODUCTS OF THOSE ACTIVITIES, RELATED TO THE IDENTIFICATION, NAMING, DEFINITION, STRUCTURING, INTEGRITY ACCURACY, EFFECTIVENESS, AND DOCUMENTATION OF THE DATA RESOURCE

DATA INTEGRATION REQUIRES A FORMAL DEFINITION OF DATA ARCHITECTURE!

© 2006 by Michael H. Brackett



DATA RESOURCE CONCEPTS

◆ **COMMON DATA ARCHITECTURE**


- COMMON CONTEXT FOR
 - INVENTORYING ALL DATA
 - UNDERSTANDING THE CONTENT AND MEANING OF DATA
 - IMPROVING DATA QUALITY
 - TRANSFORMING AND INTEGRATING DISPARATE DATA
 - DEFINING NEW DATA
 - MANAGING DYNAMIC DATA DEPLOYMENT

- EINSTEIN'S PRINCIPLE
 - A PROBLEM CANNOT BE SOLVED WITH THE SAME LEVEL OF TECHNOLOGY USED TO CREATE THAT PROBLEM
 - A HIGHER LEVEL OF TECHNOLOGY IS NEEDED
 - THE COMMON DATA ARCHITECTURE IS THE HIGHER LEVEL OF TECHNOLOGY FOR UNDERSTANDING AND MANAGING DATA

- DEVELOP CUSTOMIZED METHOD
 - ONE SIZE FITS ALL METHODS DO NOT WORK
 - COMMON DATA ARCHITECTURE HAS CONCEPTS - PRINCIPLES - TECHNIQUES
 - ORGANIZATION HAS PROBLEMS - NEEDS - ENVIRONMENT
 - CUSTOMIZE A METHOD TO SUIT ORGANIZATION

COMMON DATA ARCHITECTURE TRANSCENDS ALL DATA!

© 2006 by Michael H. Brackett



DATA RESOURCE CONCEPTS

◆ **ENTERPRISE DATA ARCHITECTURE**

- MANY DIFFERENT DEFINITIONS - CONFLICTING AND CONFUSING

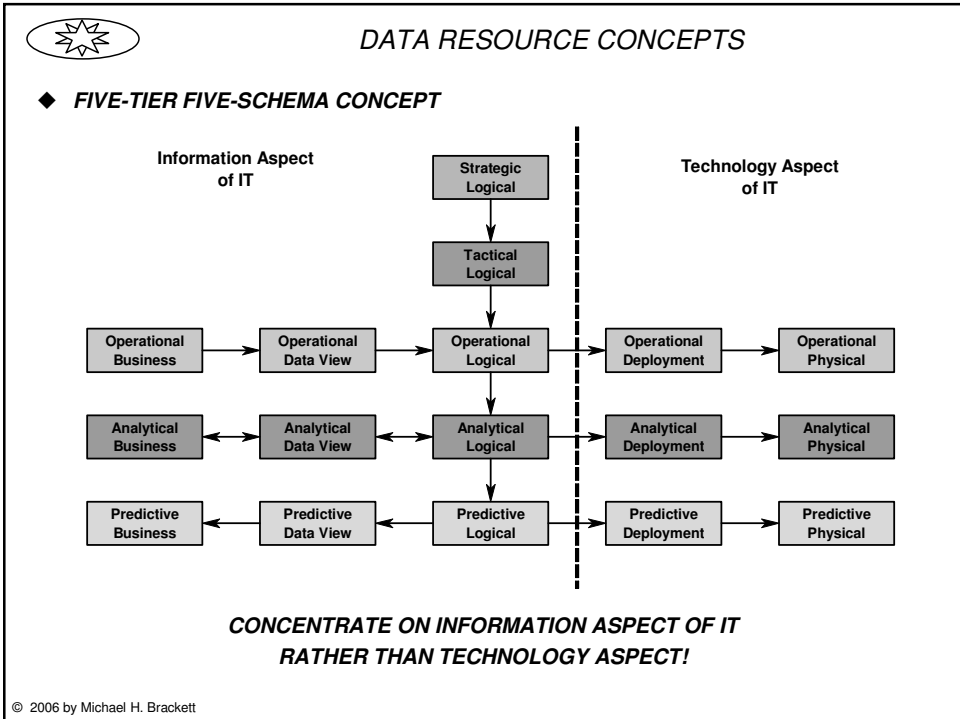
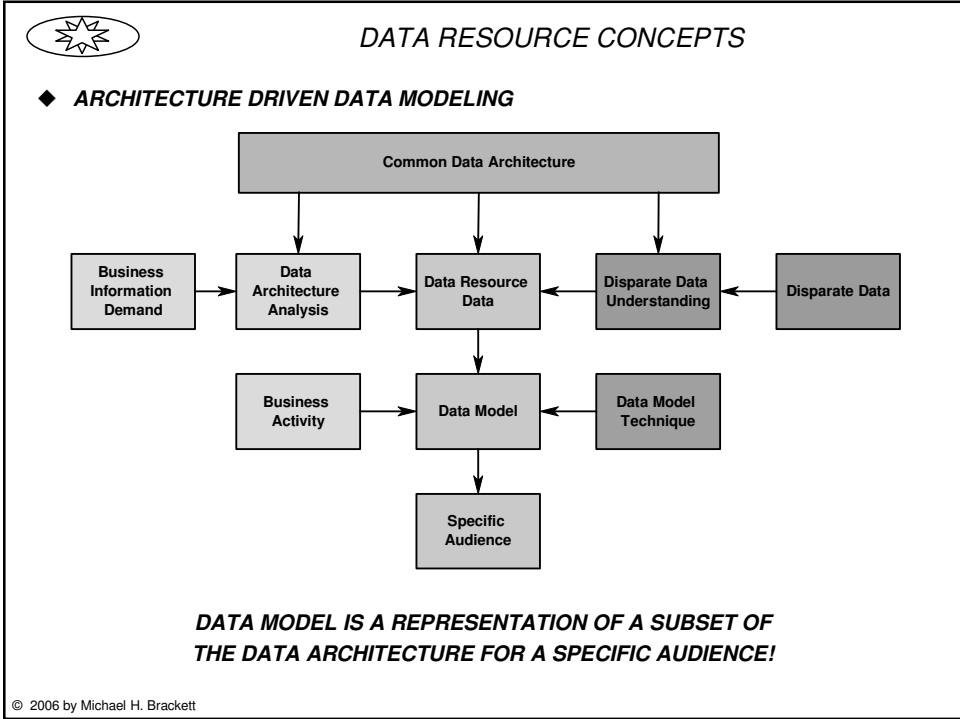
AN ENTERPRISE DATA ARCHITECTURE IS A SUBSET OF THE COMMON DATA ARCHITECTURE THAT REPRESENTS THE COMPARATE DATA RESOURCE.

- IT REPRESENTS THE DESIRED ENTERPRISE-WIDE DATA ARCHITECTURE FOR AN ORGANIZATION.
- IT IS THE PATTERN FOR BUILDING THE COMPARATE DATA RESOURCE FROM THE DISPARATE DATA


THE ENTERPRISE DATA ARCHITECTURE REPERESENTS THE COMPARATE DATA RESOURCE!

© 2006 by Michael H. Brackett

The Common Data Architecture in Perspective, Michael Brackett, Data Resource Design and Remodeling



The DAMA International Symposium & WILSHIRE Meta-Data Conference
Denver, Colorado ● April 23-27, 2006




HALTING DATA DISPARITY

- ◆ **CONCEPT - FORMAL DATA NAMES**
- ◆ **PRINCIPLES**
 - FORMAL DATA NAMING TAXONOMY
 - DATA NAMING VOCABULARY
 - PRIMARY DATA NAME
 - DATA NAME WORD ABBREVIATIONS
 - DATA NAME ABBREVIATION ALGORITHM

- ◆ **CONCEPT - COMPREHENSIVE DATA DEFINITIONS**
- ◆ **PRINCIPLES**
 - MEANINGFUL DATA DEFINITIONS
 - THOROUGH DATA DEFINITIONS
 - CORRECT DATA DEFINITIONS
 - FUNDAMENTAL DATA DEFINITIONS

- ◆ **CONCEPT - PROPER DATA STRUCTURES**
- ◆ **PRINCIPLES**
 - TECHNICALLY CORRECT & CULTURALLY ACCEPTABLE
 - STRUCTURALLY STABLE & BUSINESS FLEXIBLE
 - PROPER DETAIL FOR AUDIENCE

© 2006 by Michael H. Brackett




HALTING DATA DISPARITY

- ◆ **CONCEPT - PRECISE DATA INTEGRITY RULES**
- ◆ **PRINCIPLES**
 - DATA RULE CONCEPT
 - DATA INTEGRITY RULE NAMES
 - DATA INTEGRITY RULE NORMALIZATION
 - DATA INTEGRITY RULE NOTATION
 - DATA INTEGRITY RULE TYPES
 - FUNDAMENTAL DATA INTEGRITY RULES
 - DATA INTEGRITY RULE ENFORCEMENT
 - PROACTIVE DATA QUALITY MANAGEMENT - SELF CORRECTING PROCESS

- ◆ **CONCEPT - ROBUST DATA DOCUMENTATION**
- ◆ **PRINCIPLES**
 - DATA RESOURCE DATA
 - DATA RESOURCE DATA ASPECTS
 - COMPLETE DATA DOCUMENTATION
 - CURRENT DATA DOCUMENTATION
 - UNDERSTANDABLE DATA DOCUMENTATION
 - NON-REDUNDANT DATA DOCUMENTATION
 - READILY AVAILABLE
 - KNOWN TO EXIST
 - ANCILLARY DATA DOCUMENTATION

© 2006 by Michael H. Brackett




HALTING DATA DISPARITY

- ◆ **CONCEPT - REASONABLE DATA ORIENTATION**
- ◆ **PRINCIPLES**
 - BUSINESS SUBJECT ORIENTATION
 - BUSINESS CLIENT ORIENTATION
 - FIVE-TIER CONCEPT
 - DATA NORMALIZATION
 - SINGLE ARCHITECTURE ORIENTATION
 - SINGLE FACT ORIENTATION

- ◆ **CONCEPT - ACCEPTABLE DATA AVAILABILITY**
- ◆ **PRINCIPLES**
 - ADEQUATE DATA ACCESSIBILITY
 - ADEQUATE DATA PROTECTION
 - ADEQUATE DATA RECOVERY
 - PROTECTED PRIVACY AND CONFIDENTIALITY
 - APPROPRIATE DATA USE

- ◆ **CONCEPT - ADEQUATE DATA RESPONSIBILITY**
- ◆ **PRINCIPLES**
 - AUTHORIZED DATA STEWARDSHIP
 - REASONABLE MANAGEMENT PROCEDURES
 - CENTRALIZED CONTROL

© 2006 by Michael H. Brackett




HALTING DATA DISPARITY

- ◆ **CONCEPT - EXPANDED DATA VISION**
- ◆ **PRINCIPLES**
 - WIDER SCOPE
 - REASONABLE DEVELOPMENT DIRECTION
 - REALISTIC PLANNING HORIZON
 - COOPERATIVE ESTABLISHMENT

- ◆ **CONCEPT - APPROPRIATE DATA RECOGNITION**
- ◆ **PRINCIPLES**
 - TARGET VESTED INTEREST
 - DIRECT BUSINESS INVOLVEMENT
 - TAP THE KNOWLEDGE BASE
 - START WITHIN CURRENT BUDGET
 - INCREMENTALLY COST EFFECTIVE APPROACH
 - PROOF POSITIVE PERSPECTIVE
 - BE OPPORTUNISTIC
 - BUILD ON LESSONS LEARNED
 - NO BLAME - NO WHITEWASH ATTITUDE
 - NO UNNECESSARY JUSTIFICATION

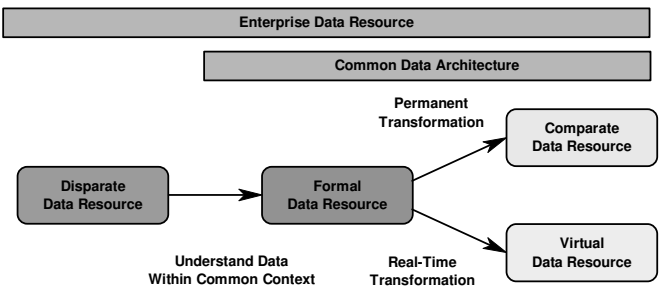
© 2006 by Michael H. Brackett

The Common Data Architecture in Perspective, Michael Brackett, *Data Resource Design and Remodeling*




RESOLVING DISPARATE DATA

◆ **CONCEPT - DATA RESOURCE TRANSITION**



- **DISPARATE DATA RESOURCE - CURRENT STATE**
 - DATA NOT WELL UNDERSTOOD
- **FORMAL DATA RESOURCE - NECESSARY STATE**
 - NECESSARY TO THOROUGHLY UNDERSTAND DISPARATE DATA
- **VIRTUAL DATA RESOURCE - DESIRED STATE**
 - REAL-TIME PRODUCTION TRANSFORMATION OF DISPARATE DATA
- **COMPARE DATA RESOURCE - IDEAL STATE**
 - PERMANENT TRANSFORMATION OF DISPARATE DATA

© 2006 by Michael H. Brackett



RESOLVING DISPARATE DATA


◆ **PRINCIPLE - FORMAL DATA RESOURCE**

- **DATA INVENTORY**
 - INVENTORY EXISTING DISPARATE DATA
 - RAISES AWARENESS OF THE DATA THAT EXIST
 - SOLVES THE FIRST PROBLEM WITH DISPARATE DATA
- **DATA CROSS REFERENCE**
 - CROSS-REFERENCE DISPARATE DATA TO COMMON DATA ARCHITECTURE
 - PROVIDES AN UNDERSTANDING OF DISPARATE DATA IN A COMMON CONTEXT
 - SOLVES THE SECOND PROBLEM WITH DISPARATE DATA
- **DATA VARIABILITY**
 - IDENTIFY DATA VARIABILITY
 - DESIGNATE PREFERRED DATA VARIATIONS
 - DEVELOP DATA TRANSLATION SCHEMES
 - PREPARATION TO SOLVE THE THIRD PROBLEM WITH DISPARATE DATA
- **DATA REDUNDANCY**
 - IDENTIFY DATA REDUNDANCY
 - DESIGNATE PREFERRED DATA SOURCES
 - PREPARATION TO SOLVE THE FOURTH PROBLEM WITH DISPARATE DATA

DEVELOPING A FORMAL DATA RESOURCE IS NON-DESTRUCTIVE!

© 2006 by Michael H. Brackett

The DAMA International Symposium & WILSHIRE Meta-Data Conference
Denver, Colorado ● April 23-27, 2006



RESOLVING DISPARATE DATA

◆ **DATA TRANSFORMATION**
THE PROCESS OF TRANSFORMING DISPARATE DATA TO COMPARE DATA WITHIN THE COMMON DATA ARCHITECTURE

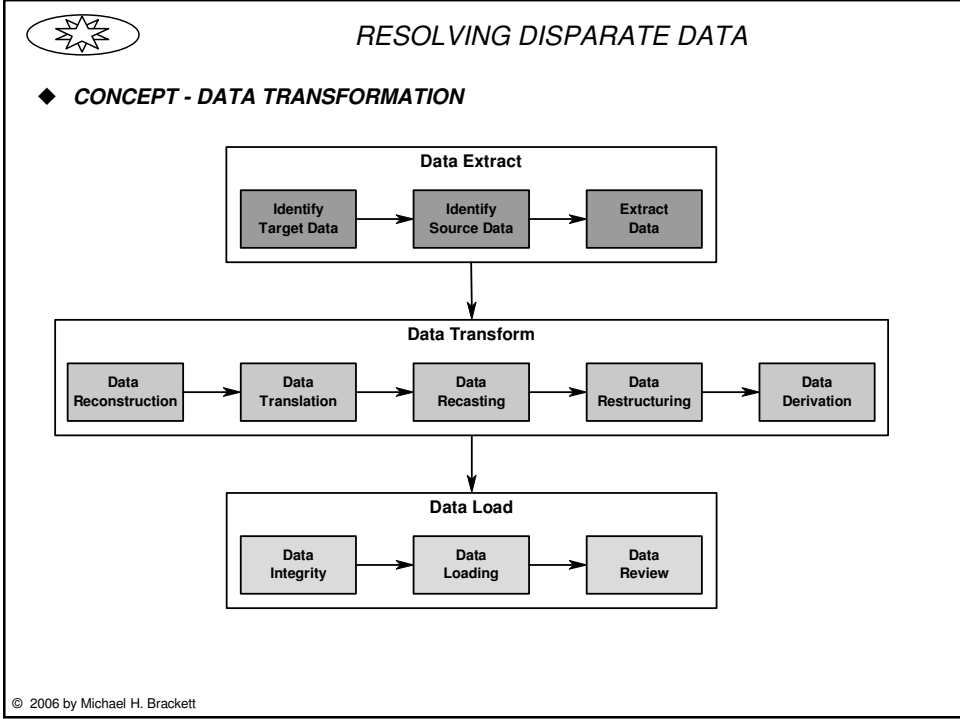
◆ **APPROACH**

- FORMAL DATA TRANSFORMATION WITHIN COMMON DATA ARCHITECTURE
- BASED ON FORMAL DATA RESOURCE - THOROUGH UNDERSTANDING OF DATA
- REQUIRES ROBUST DATA TRANSFORMATION
- NOT A TRIVIAL TASK - BUT NOT IMPOSSIBLE EITHER
- IGNORED IN MANY METHODS

- MANY ORGANIZATIONS NOT FORMALLY INTEGRATING THEIR DATA
- USUALLY INCOMPLETE / INCONSISTENT DATA INTEGRATION
- IDENTIFYING A SINGLE SYSTEM OF REFERENCE USUALLY NOT VALID
- USING *SUCK-AND-SQUIRT* METHODS TO OBTAIN DATA USUALLY DON'T WORK
- USING *BRUTE-FORCE-PHYSICAL* APPROACHES
- MORE THAN SIMPLE EXTRACT - TRANSFORM - LOAD
- MORE THAN MINOR DATA CLEANSING

COMPARE DATA RESOURCE BUILT WITHIN ENTERPRISE DATA ARCHITECTURE THROUGH DATA TRANSFORMATION!

© 2006 by Michael H. Brackett





CONCLUSION

- ◆ **COMMON DATA ARCHITECTURE**
 - TRANSCENDS ALL DATA
 - BOTH DISPARATE DATA AND COMPARETE DATA
 - CONTEXT FOR UNDERSTANDING AND MANAGING ALL DATA

- ◆ **ENTERPRISE DATA ARCHITECTURE**
 - SUB-SET OF COMMON DATA ARCHITECTURE
 - REPRESENTS THE DESIRED ENTERPRISE-WIDE DATA ARCHITECTURE
 - THE PATTERN FOR TRANSFORMING DISPARATE DATA TO COMPARETE DATA

- ◆ **DATA INTEGRATION**
 - FORMAL PROCESS
 - DEFINING COMMON DATA ARCHITECTURE
 - INVENTORYING DISPARATE DATA
 - CROSS REFERENCING DISPARATE DATA TO THE COMMON DATA ARCHITECTURE
 - ENHANCING THE COMMON DATA ARCHITECTURE
 - MAKING PREFERRED DESIGNATIONS
 - TRANSFORMING DISPARATE DATA ACCORDING TO THOSE DESIGNATIONS

THERE IS NO PLACE FOR BRUTE-FORCE-PHYSICAL APPROACHES!

© 2006 by Michael H. Brackett



ANNOUNCEMENT

- ◆ **JOHN ZACHMAN**
 - FRAMEWORK FOR INFORMATION SYSTEM ARCHITECTURES
 - EVOLVING FROM THE 1980'S

- ◆ **MICHAEL BRACKETT**
 - COMMON DATA ARCHITECTURE
 - EVOLVING FROM 1980'S

- ◆ **MERGING THE COMMON DATA ARCHITECTURE INTO FRAMEWORK COLUMN 1**
 - DISCIPLINE FOR MANAGING THE DATA RESOURCE
 - FOUNDATION FOR FORMAL DATA RESOURCE MANAGEMENT

© 2006 by Michael H. Brackett

The Common Data Architecture in Perspective, Michael Brackett, *Data Resource Design and Remodeling*



PRESENTER'S BIBLIOGRAPHY

DEVELOPING DATA STRUCTURED INFORMATION SYSTEMS
KEN ORR & ASSOCIATES, 1984

DEVELOPING DATA STRUCTURED DATABASES
PRENTICE HALL, 1987

PRACTICAL DATA DESIGN
PRENTICE HALL, 1990

DATA SHARING USING A COMMON DATA ARCHITECTURE
JOHN WILEY & SONS, 1994

THE DATA WAREHOUSE CHALLENGE: TAMING DATA CHAOS
JOHN WILEY & SONS, 1996

DATA RESOURCE QUALITY: TURNING BAD HABITS INTO GOOD PRACTICES
ADDISON WESLEY, 2000

DATA RESOURCE INTEGRATION: UNDERSTANDING & RESOLVING DATA DISPARITY
IN PROGRESS, TENTATIVE PUBLISHER: ADDISON WESLEY, TENTATIVE DATE: 2001

DATA RULES: KEY TO DATA RESOURCE QUALITY
IN PROGRESS, NO PUBLISHER OR DUE DATE

CONSULT PRESENTER'S WEB PAGE FOR ADDITIONAL DETAILS

© 2006 by Michael H. Brackett