

Managing Volatile Data within a Data Warehouse: It Can Be Done  
Ed Schaefer  
Senior Applications Developer  
Enrollment & Marketing Research, Enrollment Management & Marketing  
DePaul University

Association for Institutional Research, 2010 Annual Forum  
Chicago, Illinois  
June 1, 2010

Background...



Large, urban, private, Catholic



Nine colleges/schools



Six campuses; overseas programs



# Enrollment Management & Marketing (EMM)

Admissions

Career and Money Management

Center for Access and Attainment

Enrollment & Marketing Research

Enrollment Systems & Operations

Financial Aid

Office of Institutional Planning and Research

Student Records

University Marketing Communications

University Marketing Strategy



# EMM Data Warehouse

Admissions and Financial Aid

Research and Analysis

Office of Institutional Planning and Research  
Enrollment & Marketing Research

What we know about DW...

Large, complex...



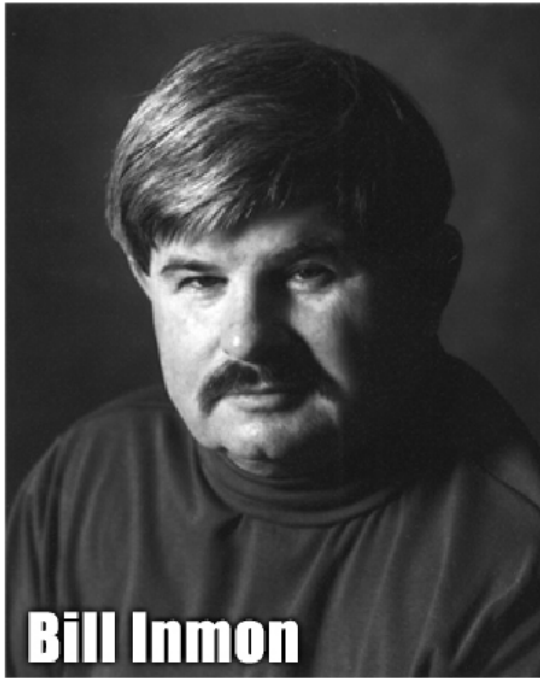
...and pricey technology

*Time consuming to build...*





# DW Concepts...



Integrated 

Time-variant

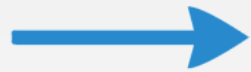
Subject oriented

Non-volatile 

So, what's my issue with non-volatile...

Data changes are the T of ETL...

Extract



Tranform



Load

Copyright © 2010 Oracle Corporation. All rights reserved.  
Oracle and the Oracle logo are trademarks of Oracle Corporation and/or its affiliates. Other brands and product names are trademarks of their respective owners.



## Types of slowly changing dimensions (SCD)

- Type 1: Truncate and reload; history rewritten daily
- Type 2: New rows written with data changes; full historical capture
- Type 3: Previous and new value columns; partial historical capture

When it comes to non-volatile...

Type 1



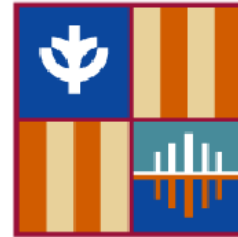
Type 2



# New business process impacts history



- Biology
- Mathematics
- Journalism
- Communication
- Etc...

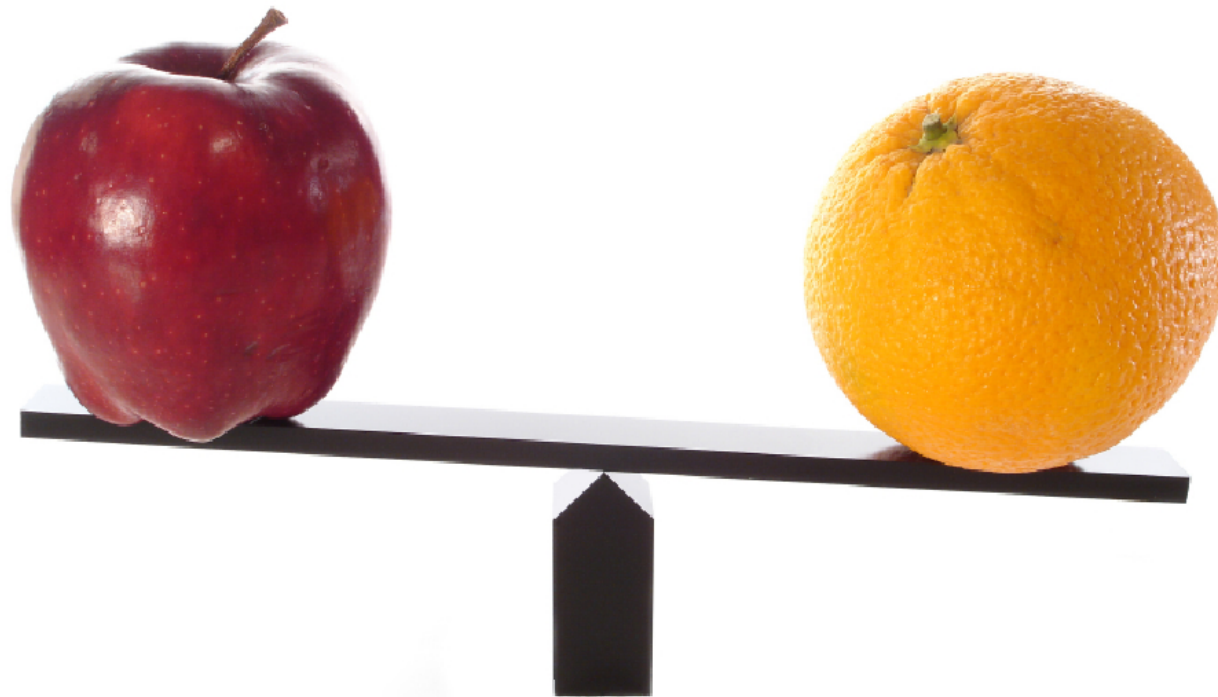


- Biology
- Mathematics
- Etc...



- Journalism
- Communication
- Etc...

New logic  
impacts history



History is just  
wrong

Mr.





When it comes to non-volatile...

Type 1



Type 2



New business process impacts history



- Biology
- Mathematics
- Etc...

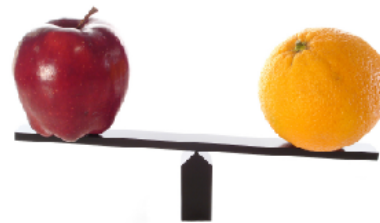


- Biology
- Mathematics
- Journalism
- Communication
- Etc...



- Journalism
- Communication
- Etc...

New logic impacts history



History is just wrong

Mr.



EMM looks for DW partner in Information Services (IS)...

EMM DW needs:

- Wish to utilize university DW...but not maintain

Information Services (IS) DW:

- ✓ Maintain university DW

EMM DW needs:

- Maintain near daily captures of cleaned data

Information Services (IS) DW:

✓ Daily loads

EMM DW needs:

- Supply history for YTD and trends

Information Services (IS) DW:

✓ Type 2 SCDs

EMM DW needs:

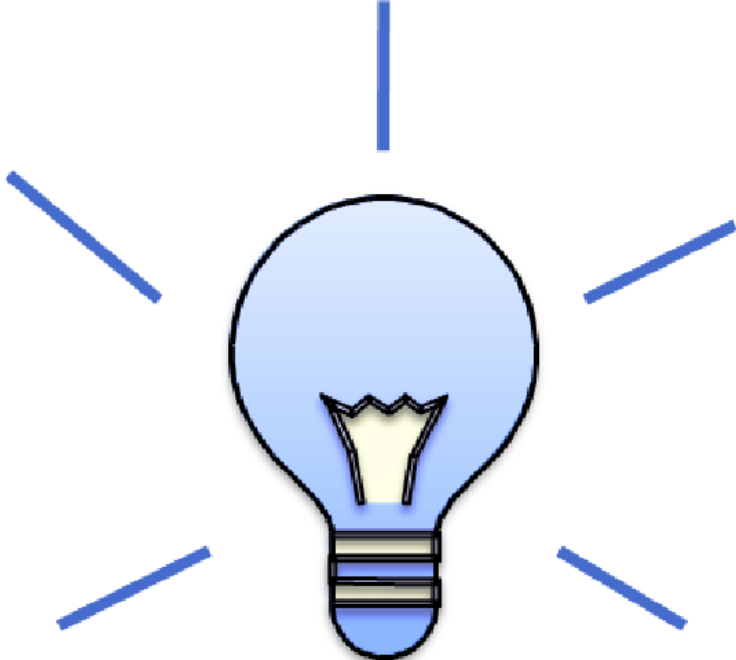
- Must be able to modify history when needed

Information Services (IS) DW:

~~X~~ Non-volatile



Our solution...





Data from DW-Q may feed into the transformation process

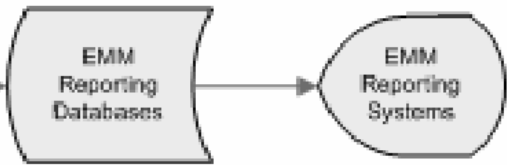
E/I: Errors and Inconsistencies

**Quality Feedback Process**

Apply OIPR, EMR, and FA transformation logic

**REPORTING EXTRACTS**

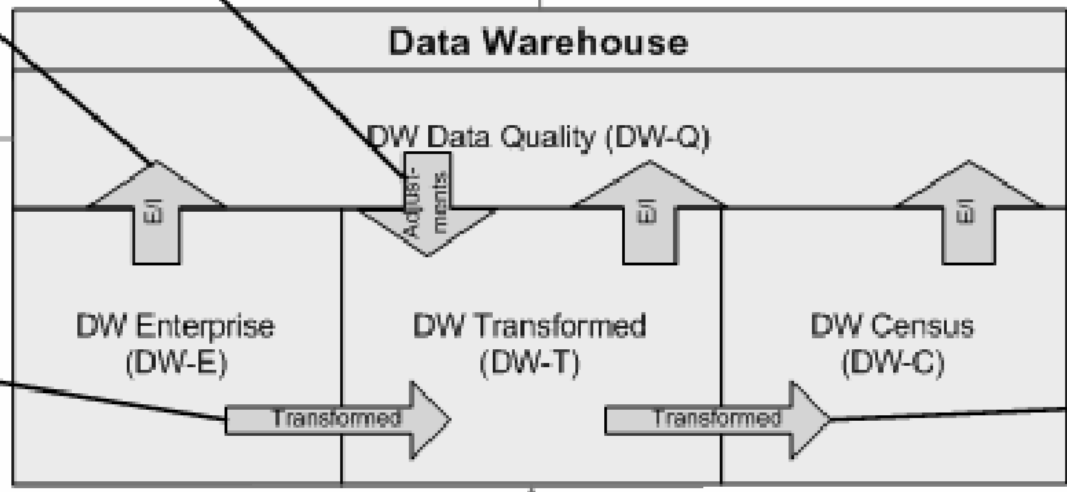
- Copy
- Restructure data but data values remain unaltered



- EMM Reporting Systems include:
- Admission Summary
  - Enrollment update report
  - HCRC data extract
  - EMM report suite
  - OIPR Fact File & cubes

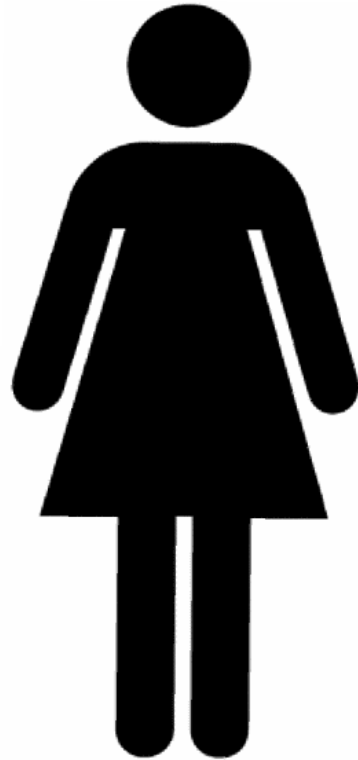
Reporting systems vary between departments but data from the DW DOES NOT

Further census data checks and, if required, transformations

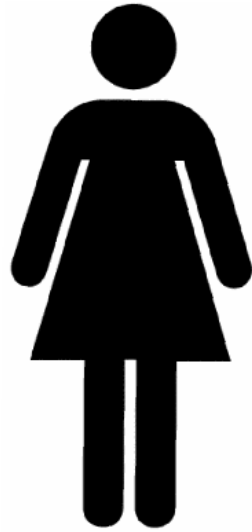


Example of how it works...

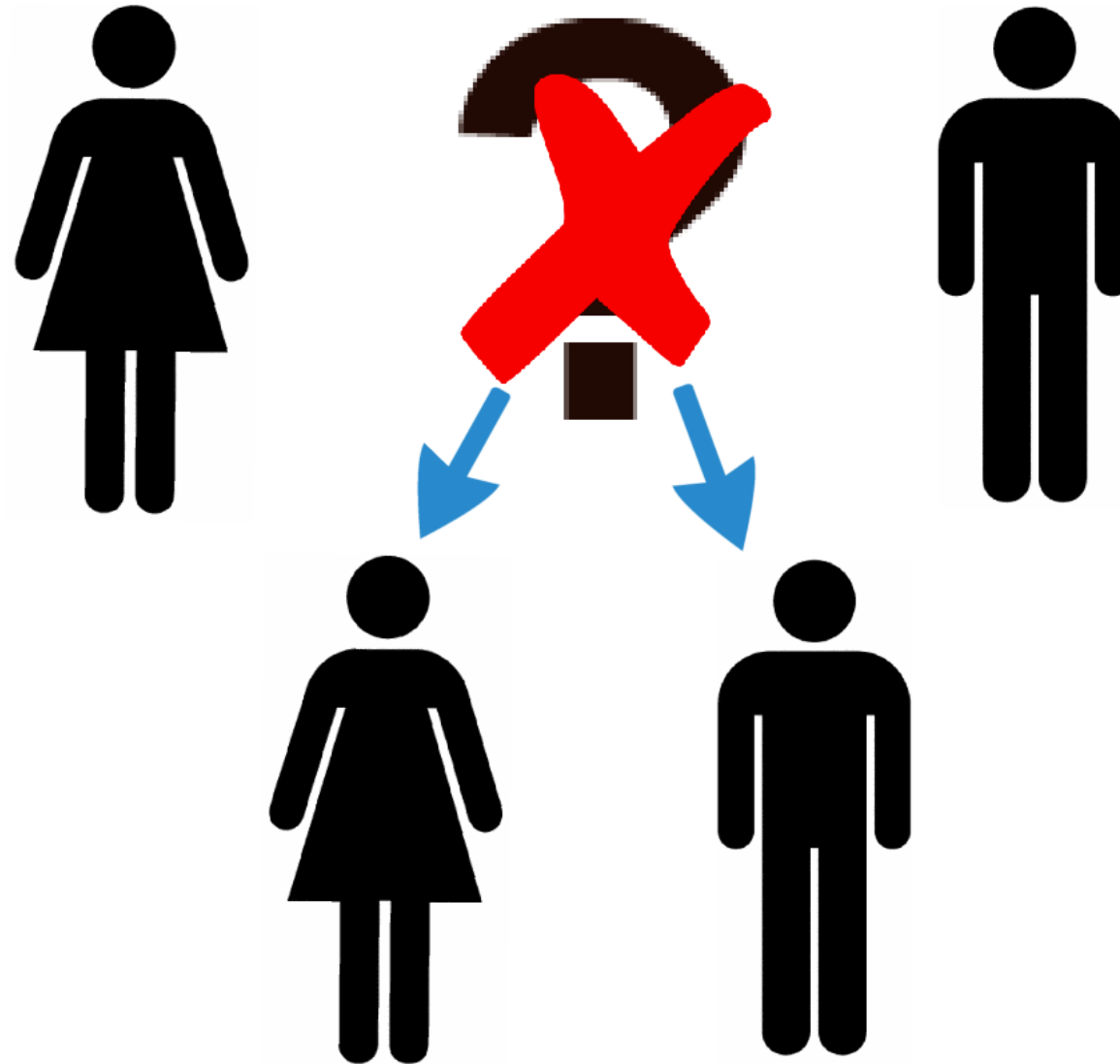
D\_PERSON in DW has GENDER



D\_PERSON in DW has GENDER



EMM builds and maintains  
D\_PERSON\_AUX with R\_GENDER



# Integrated: PERSON\_KEY in D\_PERSON\_AUX

D\_PERSON

D\_PERSON\_AUX



Since EMM controls  
D\_PERSON\_AUX...

D\_PERSON

D\_PERSON\_AUX



# Groundrules...

- UPDATE non-key fields only
- INSERT cannot split a record
- DELETE with extreme caution
- Only IS can execute DELETE



## Groundrules...

- UPDATE non-key fields only
- INSERT cannot split a record
- DELETE with extreme caution
- Only IS can execute DELETE

## Likelihood...

- Somewhat likely
- Very unlikely
- Extremely unlikely

Thank you!



For further questions or comments:

Ed Schaefer  
eschaefer@depaul.edu

