

Santa Clara University

Development of the Institutional Research Data Mart

WASC Visiting Team Presentation

October 21, 2009

How do we know we need a data mart?

- *Organizations need people to understand them.*
- *For understanding, people need information.*
- *For information, people need easy access to reliable, quality data.*

Mission Related

How does the data mart support the mission and objectives of Santa Clara?

- Platform for 'evidence-based thoughtful decision-making'
- Enhance the University's quality assurance programs through improved processes for review of academic and co-curricular programs
- Deepen our understanding of student learning and develop more effective methods of assessing learning

Business Intelligence

How does the data mart support the business intelligence needs of Santa Clara?

- Develop metrics of diversity and inclusiveness
- Develop ways of assessing the Santa Clara teaching-scholar model
- Develop indicators for the assessment of 'competence, conscience, and compassion'
- Move towards a learning organization: business processes and data governance
- Integrate student-reported assessment with student-record measures

Why a data mart?

- PeopleSoft is a transaction system
- Wrong “tool” for the job
- PS Query doesn’t support our mission

The Switch from ISIS to PeopleSoft

What happened to reporting?

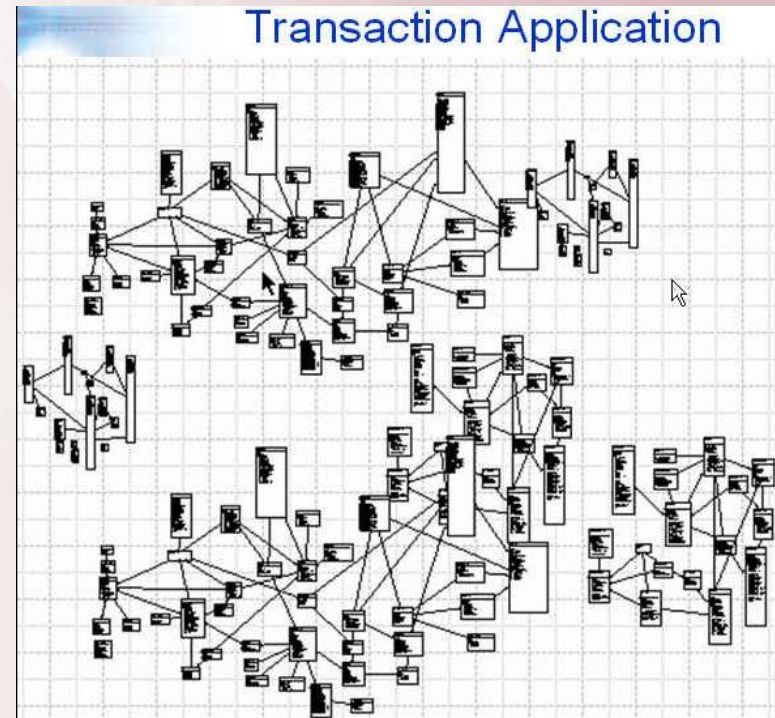
- **ISIS (1989-1999)**
 - Centralized Database; Centralized Reporting
 - Regular intervals
 - Snapshot data
 - Delivered (canned) reports
 - Data as image; 'Monarch' as verb

- **PeopleSoft (fall 1999-today)**
 - Centralized Database; De-centralized Reporting
 - Ad-hoc, independent timing
 - Transaction data
 - PS Query access for 'Power users'
 - Independent 'shadow' databases

'Reporting' from a Transaction System

Why is reporting from PeopleSoft so difficult?

- Challenges to reporting from a transaction system
 - Sophisticated use of SQL
 - Thorough understanding of business processes
 - Complete understanding of the relations among the tables of normalized data
 - Data are a 'moving target'
 - Transaction fields & field values rarely coincident with reporting categories and definitions
- There must be a better way.



Build Something!

Goals of the Data Mart:

- *Provide source for use in satisfying IR reporting obligations in the areas of admissions, student registration, degrees awarded, financial aid, and faculty headcount, course and class data*
- *Provide source for satisfying IR analytical obligations in the areas of student retention, admissions, and*
- *Identify business rules and use to structure data for reporting*
- *Audit PS data for consistency and completeness*

Our Path:

- Pragmatic, low-budget, immediate payoff, learn-as-you-go data mart
 - Iterative, limited-scope process.
 - Identify and rank key reporting needs
 - Identify the data required to support key reports
 - Build de-normalized table(s)
 - Clean, document, feedback
 - (semi) automate
 - repeat



Examples of the Data Mart in Action

Reporting is back!

- **WASC Accreditation**
 - 62 WASC exhibits prepared using DMIR and Cognos Reporting Tool
 - Data extracts from DMIR used in SPSS for more complex statistical analysis
 - Faculty data reports produced from DMIR for the first time using data from DMIR combined with imported Provost databases
 - Deadlines met even with 1 FTE shortage for 11 months in the Office of Institutional Research
- **Program Review Reports**
 - Class Size
 - Classes Taught by Faculty Type
 - Grade Distribution Reports by School/College and Department
 - Ethnicity and Gender, Headcount and FTE : Full-Time Faculty
 - Ethnicity and Gender, Headcount and FTE : Academic Year and Quarterly Part-Time Faculty
 - Headcount/FTE--Fall Term
 - Undergraduate Students by Major – Fall Terms
 - Undergraduate Students by Major, Ethnicity & Gender - Fall Terms
 - Graduate & Law Students by Major, Ethnicity & Gender – Fall Terms
 - Undergraduate Degrees Completed by Ethnic Group & Gender - Academic Year
 - Graduate Degrees Completed by Ethnic Group & Gender - Academic Year
 - Undergraduate Level – Degrees Conferred with Honors and GPA by Academic Year (First-time Full-time Freshmen)
 - Graduate Level – Average GPA and Units Earned at Degree Completion
 - Undergraduate Level Completions – Average High School GPA and SAT Scores (First-time Full-time Freshmen)
- **External Survey and Government Agency Reporting**
 - About 40 surveys completed/coordinated annually by IR
 - DMIR used in 2008-09 cycle
 - Programs easily updatable to future terms/years
 - Facilitates new staff orientation to using DMIR for external agency surveys
 - Responses consistently prepared
 - Data consistently stored
 - Response completion time shorter
 - Auditing results faster and easier



Moving forward...

Lessons learned, best practices adopted

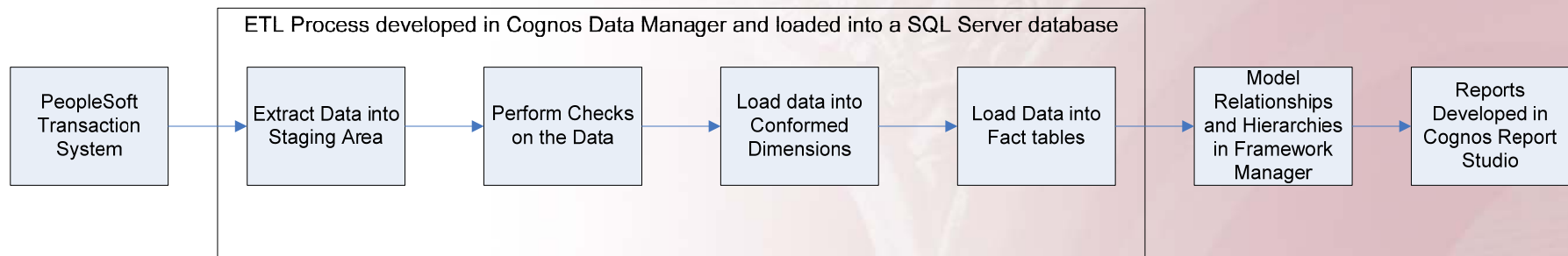
- Census-based to daily extracts
- Conformed dimensions
- Systematic approaches
- Analysis leads to more analysis

Challenges/Design Responses

A Sampling of Design Solutions in the Institutional Research Data Mart



Overview of IR Data Mart Design



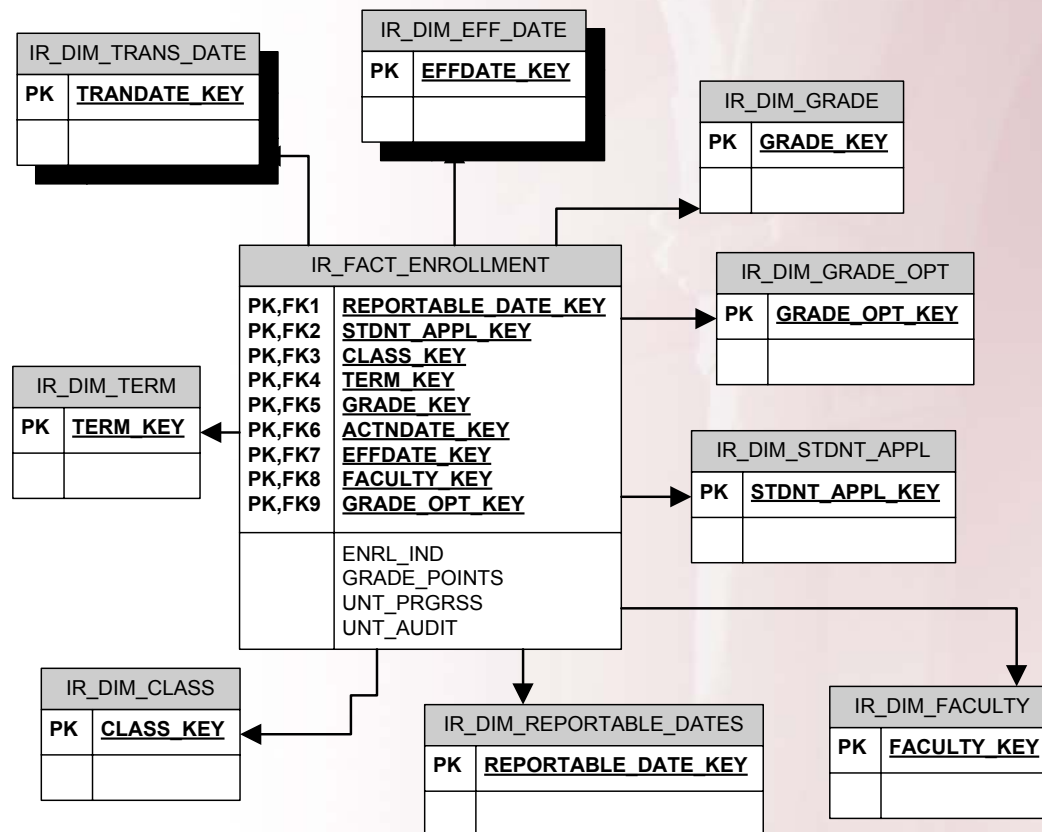
- ETL and Staging
- Dimensional Tables
- Reporting Tools

Reconciling Reporting and Analysis

Problem:

- Keeping reporting and analysis in sync
- Different views of the data

Solution:

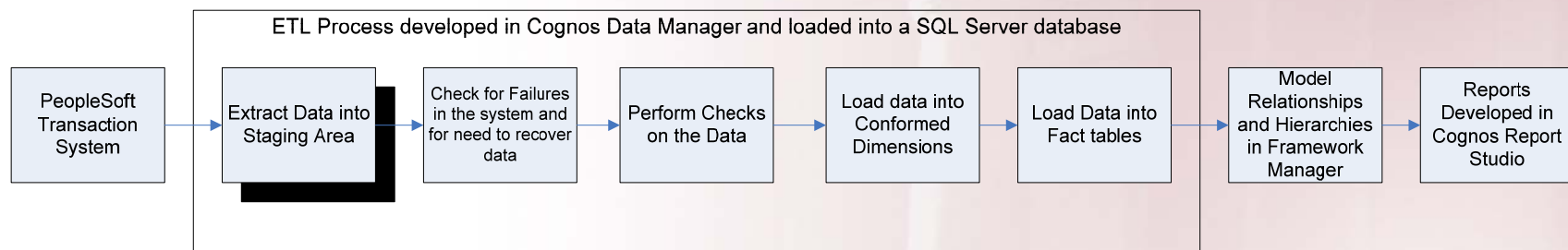


Correction Mode Changes

Problem:

- Correction Mode changes can be used on Key fields
- Correction Mode can be used to delete a record
- Correction Mode does not leave an audit trail

Solution:



Logical inconsistencies within the data

Problem:

- Data entry errors may create invalid combinations of data
- Data across areas in the system may not be logically consistent with one another

Solution:

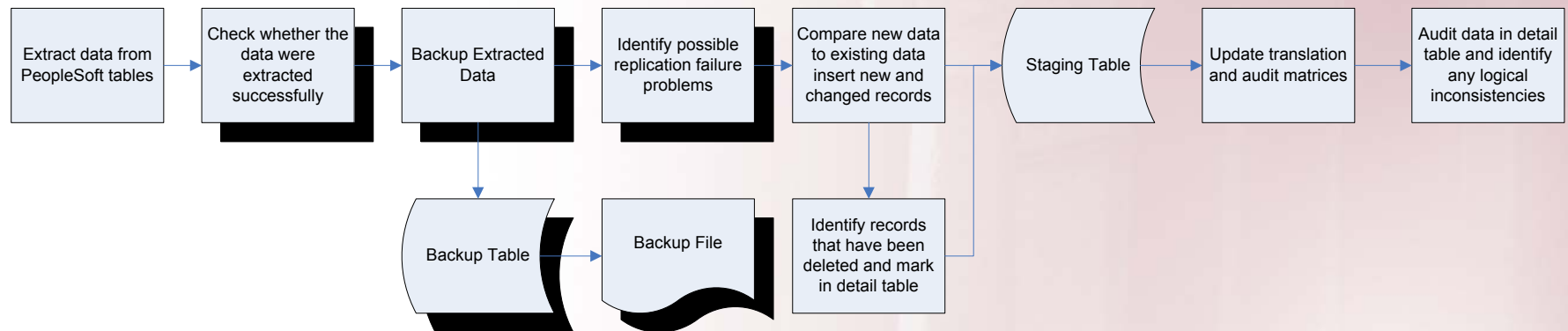
Career	Program	Plan	Plan Type	Plan Rank	Valid_yn
Undergraduate	Undergraduate Business	Marketing	MAJ	1	Y
Undergraduate	Undergraduate Business	Marketing	MAJ	2	N
Undergraduate	Undergraduate Arts & Science	Marketing	MAJ	2	Y
Undergraduate	Undergraduate Business	Accounting	MAJ	1	Y
Undergraduate	Undergraduate Business	Accounting	MAJ	2	N
Undergraduate	Undergraduate Arts & Science	Accounting	MAJ	2	Y
Undergraduate	Undergraduate Business	Physics	MAJ	1	N
Undergraduate	Undergraduate Arts & Science	Physics	MAJ	1	Y
Undergraduate	Undergraduate Arts & Science	Physics	MAJ	2	Y
Undergraduate	Undergraduate Business	Physics	MAJ	1	N
Undergraduate	Undergraduate Business	Physics	MAJ	2	Y

Access to Data may be interrupted

Problem:

- Access to PeopleSoft data is through a replication server
- User notification is relied upon to identify any problems with replication or other problems

Solution:



Future needs are not yet known

Problem:

- We can only anticipate future events based on our past experience
- Need for data will grow and things will change

Solution:

- Anticipate where you can
- Build flexibility into design