



UNIVERSITY OF CENTRAL FLORIDA

INSTITUTIONAL RESEARCH

Data Integrity Reports:

*A Proactive Approach to
Data Quality Challenges*

FAIR Annual Conference

Cocoa Beach, FL

February 8, 2007

Presenters

- Antionette Lee
Coordinator, Computer Applications
- Linda Sullivan
Associate Director

Overview

- Maintaining quality, accuracy, and integrity of institutional data is a continual challenge to data administrators. UCF developed a proactive approach to solving this problem after years of dealing with the same data issues over and over again.
- Major data clean-up issues prior to each submission of state reports became further aggravated after the legacy to ERP conversion.
- Data integrity reports were developed to identify data errors for the support offices to correct on a regular basis. These reports have evolved with use and system upgrades.
- This proactive approach to monitoring and maintaining data quality has resulted in overall cleaner data, less time required to prepare state reports for submission, and improved business processes in the functional offices responsible for data entry.

University of Central Florida (UCF)

Metropolitan Research University

First classes – 1968

Fall 2006 enrollment – 46,700

12 colleges

5 regions with 16 distinct campus locations



Agenda/Contents

- Why a data integrity solution was needed
- Evolution of report design & delivery
- Reporting and operational benefits

Why Data Integrity Solution Needed

- Significant amounts of data to clean up in a short amount of time
- Same data errors recurring
- More involvement by functional offices in resolving data entry issues

Data Reports – Students

By career and status (Grad, Ugrad, Readmit)

- SSNs – all 9's
- Invalid race, citizenship, gender
- Invalid total transfer hours
- Current/former students without classification
- Birthdate issues
- Missing tuition residency
- Missing ACTs / SATs

Data Reports – Employees

- Active DROP participants
- Missing citizenship
- Birthdate issues

Evolution of Report Design & Delivery

- Two system upgrades
- User-friendly delivery and convenient access
- Accuracy of reports

Design Considerations

Objective: Keep it simple for the end-user!

- Automation
- Controlled Report Distribution
- Easy End-user Access

Design Considerations

Objective: Keep it simple for the end-user!

■ Automation

- The reporting process must be automated to avoid an increase in work load for Institutional Research and their IT technical support staff.

■ Controlled Report Distribution

- Reports should only be distributed if they actually contain data issues.

■ Easy End-user Access

- End-users should be able to easily access the reports.

Reporting Process Evolution

Query/Crystal Reports (Phase I)



Query (Phase II)



SQR (Phase III)
(Structured Query Report)

Phase I

Query/Crystal Process Design

- Create queries using ERP Query tool.
- Create Crystal reports, using the query results as the data source.
- Schedule the Crystal reports to run automatically as batch processes.
 - Email the report to the recipients.

Phase I

Query/Crystal Process Design

Report Sample (Query Selected Fields)

Query: Active Employees with Birthdate Issues

Selected Fields:

- Employee ID
- Effective Date of Transaction
- Name
- Birthdate
- Job Code
- Pay Plan
- Payroll Status

Phase I

Query/Crystal Process Design

Report Sample (Query Selection Criteria)

Query: Active Employees with Birthdate Issues

Selection Criteria:

- Payroll Status = 'Active'
- Age = < 9 or > 85
(Age based on birth year)

Phase I Query/Crystal Process Design Report Sample (Crystal Report)



Report ID: FXR6926P

University of Central Florida Institutional Research

Active Employees Who Have Invalid Birthdates

Page No. 1

Run Date: 1/23/2007

Run Time: 9:02:17AM

<u>Effective Date</u>	<u>EmpID</u>	<u>Name</u>	<u>Job Code</u>	<u>Salary Admin Plan</u>	<u>Birthdate</u>
10/30/200	1111111	Donald Duck	040	00	11/21/1918
12/30/200	2222222	Mickey Mouse	168	23	8/4/1921
7/1/2006	3333333	Pluto Pup	9314	00	10/9/1916
7/7/2006	4444444	Elmo Baglins	012	00	12/14/1921
10/31/200	6666666	Roger Rabbit	011	00	5/24/2008

Total Selected: 5

Phase I

Query/Crystal Process Evaluation

	Consideration Met?	Comments
Automation	Yes	ERP had functionality for automatic submissions.
Controlled Report Distribution	No	ERP had functionality to automatically disburse reports, but reports were distributed even if no errors were listed in the detail.
Easy End-user Access	Yes	Reports were e-mailed directly to the recipient's mailbox.

Phase II

Query Process Design

Report Sample (Query Selected Fields & Criteria)

- Query: Active Employees with Birthdate Issues
 - Selected Fields:
 - Employee ID
 - Effective Date of Transaction
 - Name
 - Birthdate
 - Job Code
 - Pay Plan
 - Payroll Status
 - Selection Criteria:
 - Payroll Status = 'Active'
 - Age = < 9 or > 85

Phase II

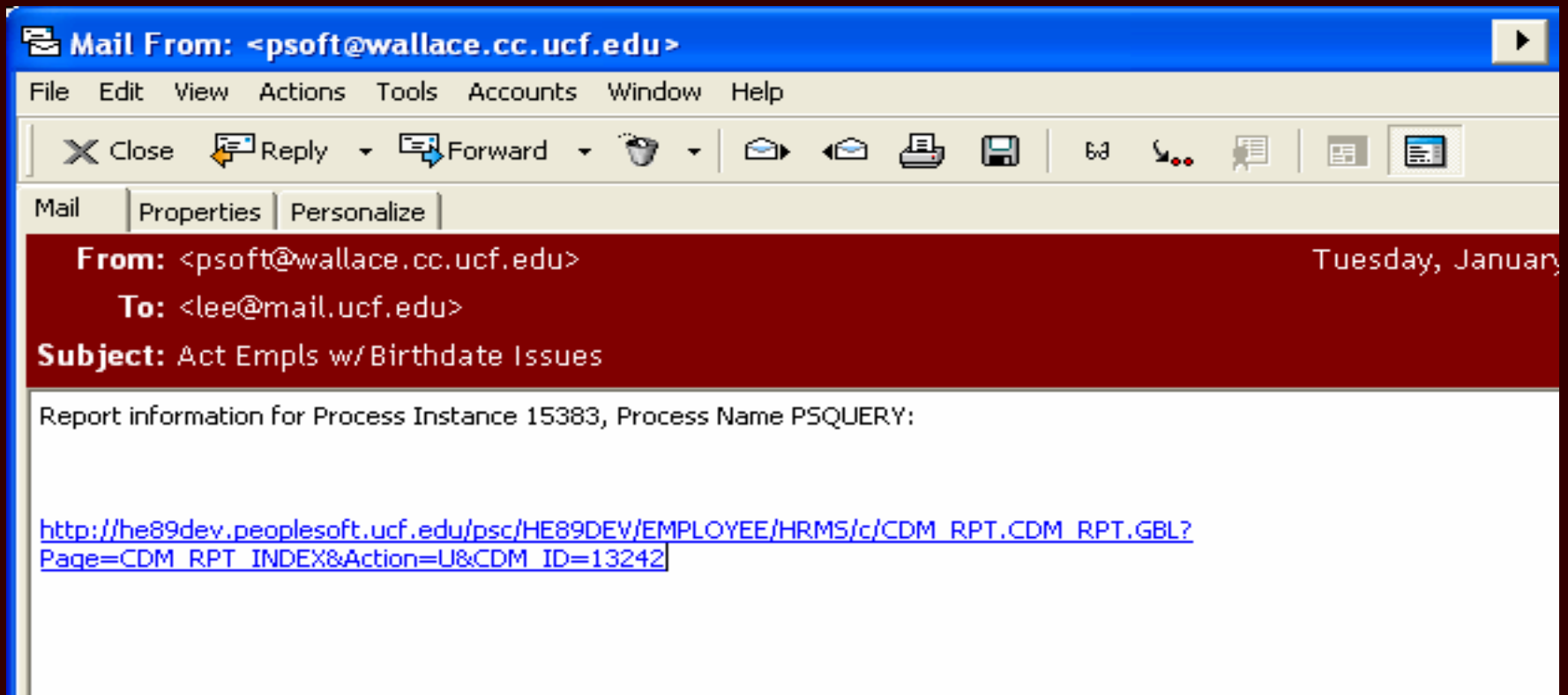
Query Process Design

- Create queries in ERP Query tool.
- Schedule the queries to run automatically as batch processes.
 - Email recipients link to Excel spreadsheet in the report repository.

Phase II

Query Process Design

Report Sample (E-mail Notification)



Phase II

Query Process Design

Report Sample (Excel Spreadsheet)

	A	B	C	D	E	F	G
1	Act Empls w/Birthdate Issues	5					
2	ID	Eff Date	Name	Birthdate	Job Code	Sal Plan	Pay Status
3	1111111	10/20/2006	Donald Duck	11/21/1918	040	00	A
4	2222222	12/30/2005	Mickey Mouse	8/4/1921	168	23	A
5	3333333	7/1/2006	Pluto Pup	10/9/1916	9314	00	A
6	4444444	7/7/2006	Bilbo Baggins	12/14/1921	012	00	A
7	5555555	10/31/2006	Roger Rabbit	5/24/2006	011	00	A
8							
9							
10							

Phase II

Query Process Evaluation

	Consideration Met?	Comments
Automation	Yes	ERP has functionality for automatic submissions. Cumbersome for multiple query submissions. It was more efficient to use a Job Stream to submit multiple queries.

Phase II

Query Process Evaluation (cont'd)

	Consideration Met?	Comments
Controlled Report Distribution	No	ERP has functionality for auto disbursement of reports, but it disburses the reports even if no errors are listed in the detail.
Easy End-user Access	No	The query reports are not e-mailed directly to the recipient's mailbox. Users are sent a link to the reports. They have to go through a lot of clicks to access the reports in the report repository.

Phase III

SQR (Structured Query Report)

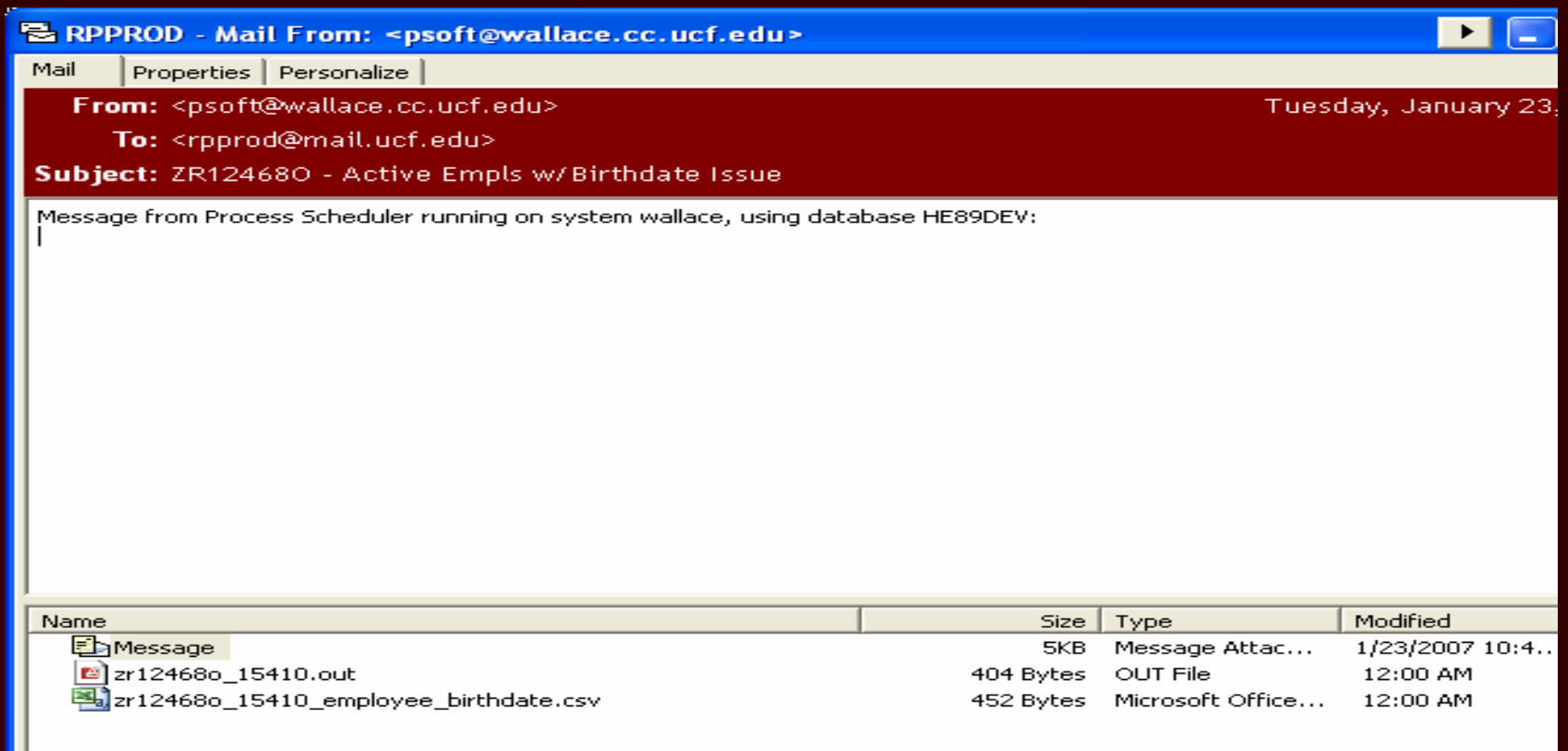
Process Design

- Create SQRs to output data errors to CSV (Comma Separated Value) files for Excel.
- Schedule the SQRs to run automatically
 - Option to mail output file to production e-mail account
 - Rules disburse output to appropriate recipient's mailbox

Phase III

SQR Report Sample

(Production Mailbox Notification)



RPPROD - Mail From: <psoft@wallace.cc.ucf.edu>

Mail | Properties | Personalize

From: <psoft@wallace.cc.ucf.edu> Tuesday, January 23, 2007 12:00 AM

To: <rpprod@mail.ucf.edu>

Subject: ZR124680 - Active Empls w/Birthdate Issue

Message from Process Scheduler running on system wallace, using database HE89DEV:

Name	Size	Type	Modified
Message	5KB	Message Attac...	1/23/2007 10:4..
zr12468o_15410.out	404 Bytes	OUT File	12:00 AM
zr12468o_15410_employee_birthdate.csv	452 Bytes	Microsoft Office...	12:00 AM

Phase III

SQR Report Sample

(Mailbox Disbursement Rule)

Edit Rule

Rule name:

When event is And items are: Received Sent Posted Draft

If conditions are (optional)

Item types:

- Mail
- Appointment
- Task
- Reminder note
- Phone message

Act on items where (Item Type is Mail) and Subject contains 'ZR124680' and Attachments contains '.csv'

Appointment conflict exists:

Then actions are

-
-
-

Move To Folder: zr12468 - Data Integrity Queries
Forward To: Antionette Lee; Linda Sullivan
Mark As Read

Phase III

SQR Report Sample (Recipient Instructions)

Forward

From: RPPROD CC:

To: Antionette Lee; Linda Sulli(...) BC:

Subject: Potential Data Integrity Issue: Files:

Message: Please review the latest Active Empls w/Birthdate Issues report and correct the data as needed.

OK
Cancel
Address Book...

Phase III

SQR Report Sample

(Forwarded E-mail)

Mail | Properties | Personalize

From: RPPROD Tuesday, January 23, 2007
To: Antionette Lee; Linda Sullivan
Subject: Fwd: ZR124680 - Active Empls w/Birthdate Issue (Potential Data Integrity Issue: Active Empls w/Bi

Please review the latest Active Empls w/Birthdate Issues report and correct the data as needed.

Name	Size	Type	Modified
Message	5KB	Message Attac...	1/23/2007 10:5...
ZR124680 - Active Empls w/Birthdate Issue	4KB	Mail Attachment	1/23/2007 10:4...

Phase III

SQR Report Sample (Excel Spreadsheet)

G35 fx

zr12468o_15410_employee_birthdate.csv

	A	B	C	D	E	F	G	H	I	J
1	Emplid	EmplRcd	Effdt	Last Name	First Name	MI	Job Code	Salary Admin Plan	Birthdate	
2	3333333	4	1-Jul-2006	Pup	Pluto	N	9314	00	9-Oct-1916	
3	1111111	5	20-Oct-2006	Duck	Donald	E	040	00	21-Nov-1918	
4	5555555	4	31-Oct-2006	Rabbit	Roger	M	011	00	24-May-2006	
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										

Phase III

SQR Process Evaluation

	Consideration Met?	Comments
Automation	Yes	ERP has functionality to automatically submit SQRs as a batch process.

Phase III

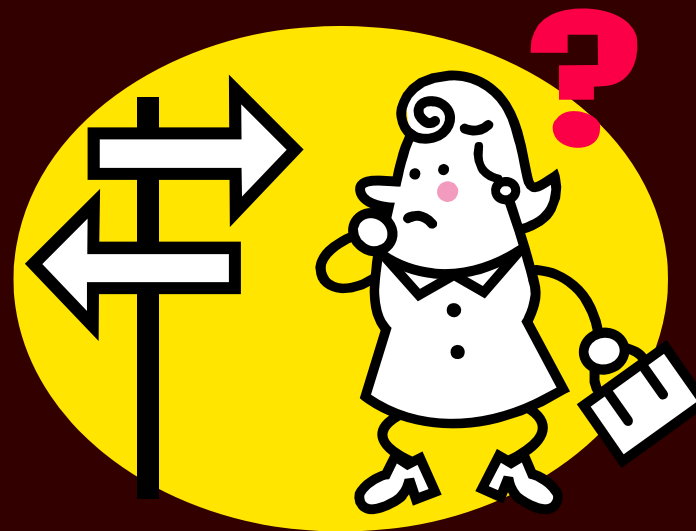
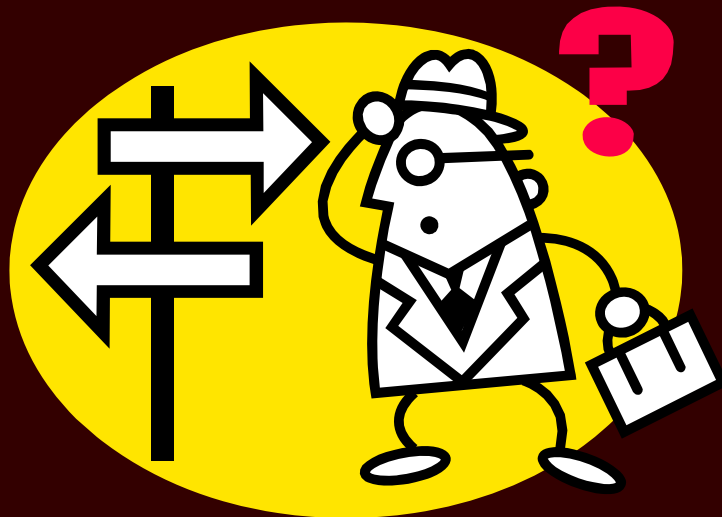
SQR Process Evaluation (cont'd)

	Consideration Met?	Comments
Controlled Report Distribution	Yes	<p>The SQRs only create CSV files if there are errors.</p> <p>ERP has functionality to automatically e-mail the output files to a production mailbox.</p> <p>The recipients will receive notification only if there are data errors.</p>
Easy End-user Access	Yes	The output files are e-mailed directly to the recipient's mailbox.

Reporting & Operational Benefits

- Clean-up of conversion data issues
- Reduced time in report preparation
- Smaller data sets to fix on a regular basis
- Business process analysis
- Departmental performance measurement tool
- Improved, consistent data quality

Questions??



How to Contact Us

- Office of Institutional Research
 - Email: iroffice@mail.ucf.edu
 - Web: <http://www.iroffice.ucf.edu>
 - 12424 Research Parkway, Suite 215
 - Phone: (407) 823-5061 Fax: (407) 823-4769
- Antionette Lee: lee@mail.ucf.edu
- Linda Sullivan: lindas@mail.ucf.edu