Enhancing Data Quality through Automation

Cheryl Moody, BA, CTR
California Cancer Registry

NAACCR 2012 Annual Conference
6/5/2012
OVERVIEW

- Background
- Automating Data Quality – The Process
- Impact on Data Quality
- Benefits
- Considerations
- Future Plans
**BACKGROUND**

- **What is Business Rules Management?**
  - The ability to automate an organization’s decision logic
• Visual editing was reduced from 100% to 40%

• A new approach to verifying data quality was required

• Business Rules Management Solutions project was developed to address data quality issues
  • Initiated in 2006; resurrected in 2009
BACKGROUND

• Objective:
  • Automation of manual processes where feasible

• Goal:
  • Maintain or improve data quality while decreasing the time spent performing manual tasks
BACKGROUND

- Long Term Goal:
  - Automate manual consolidation

- Short Term Goal:
  - Develop business rules for
    - Class of Case
    - Non-site related general rules based on existing edits
    - CS schema specific rules
Automating data quality: The Process
AUTOMATING BUSINESS RULES

Process

Write Rule

Program Rule

Query DB

Perform Data Fixes

QA Test

Peer Review Test

Implement Rule

Perform Data Fixes

6/4/12-6/7/12
NAACCR Educational Conference – California Cancer Registry
STEP ONE: Create Business Rules

- Experienced CTRs assembled to study data standards
  - Created actionable rules written in “IF, THEN” statements
  - Modeled after existing edits

- Actionable rules provided to Information Systems Programmers
STEP TWO: Program Rules

- Programmer develops code based on CTRs written documentation

- Rules programmed to recognize incorrect code

- System auto-corrects to appropriate value rather than requiring manual review and correction
STEP THREE: Data mine the database

- CTRs create & run data miner queries containing rule set conditions

- Existing cases that are identified as miscoded based on rule conditions are analyzed

- Based on query analysis, rule and/or programming is adjusted as appropriate
STEP FOUR: Perform data fixes

- Based on query findings, miscoded cases are corrected in the database
  - Global fix
  - Manually corrected
STEP FIVE: TEST, TEST, and then TEST AGAIN!

- Each rule condition is rigorously tested
- Programmers test rule integrity within the existing database structure
- CTRs test rule functionality and expected outcomes in the QA and UA environment
STEP SIX: Implement the rule

- Implement the rule within the existing system structure
- Notify users of the implemented rule(s)
  - Version Release Notes
  - Newsletter notification
  - Audit History Log
- Website
STEP SEVEN: Post-Implementation Test

- Post-implementation, re-run original queries
  - Identify cases in the database that were not corrected by the implemented rule
  - Analyze each case to determine why the rule did not fire on a particular case
    - Rule or programmed code requires revision
    - Rules were turned off by IT during time period 6/4/12 - 6/7/12
AUTOMATION: Example

- Edit (SEER IF14): If Age At Diagnosis < 15, Marital Status at Diagnosis must = 1 (single, never married)

- Rule is constructed to recognize cases where Age at Diagnosis < 15, but Marital Status <> 1
Automation Rule Example

- The resulting rule would read as follows:

  - If the following conditions are true:
    - Age at Diagnosis is less than or equal to 014
    - Marital Status <> 1

  - THEN CHANGE Marital Status to equal 1
AUTOMATION: Example

- System is programmed to run BRMS automation rules BEFORE edits
  
  - Auto-change occurs at any level of processing within the database
    
    - i.e., File Upload, Visual Editing, Active Follow-up, Corrections, Database Inquiry
  
  - Auto-correction accomplished in the absence of manual review or manual actions of any kind
<table>
<thead>
<tr>
<th>Implementation Date</th>
<th>Type of Rule</th>
<th>Description</th>
<th>Global Fix</th>
<th>Manual Fix</th>
<th>BRMS Fix</th>
<th>Total Corrections</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-2010</td>
<td>Class 49</td>
<td>DCO</td>
<td>1,554</td>
<td>91</td>
<td>16</td>
<td>1,661</td>
</tr>
<tr>
<td>6-2010</td>
<td>Class 38</td>
<td>Autopsy</td>
<td>0</td>
<td>518</td>
<td>16</td>
<td>534</td>
</tr>
<tr>
<td>3-2012</td>
<td>IF 631</td>
<td>Chemo Hosp vs Chemo Summ</td>
<td>5,045</td>
<td>0</td>
<td>1</td>
<td>5,046</td>
</tr>
<tr>
<td>3-2012</td>
<td>IF 632</td>
<td>Hormone Hosp vs Hormone Summ</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>3-2012</td>
<td>IF 633</td>
<td>BRM Hosp vs BRM Summ</td>
<td>7,269</td>
<td>0</td>
<td>0</td>
<td>7,269</td>
</tr>
<tr>
<td>3-2012</td>
<td>IF 761</td>
<td>Location of Rad, Rad Date</td>
<td>6,288</td>
<td>0</td>
<td>145</td>
<td>6,433</td>
</tr>
<tr>
<td>3-2012</td>
<td>IF 750</td>
<td>Breast: LNS exam, SSF 3</td>
<td>0</td>
<td>2</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>3-2012</td>
<td>IF 837</td>
<td>Breast: CS TS, CS Ext</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

|                     | TOTAL        |                      |            |            |           |                  |
|                     | TOTAL        |                      | 20,156     | 620        | 197       | 20,973           |

BRMS is activated anytime a case is loaded or a change is made to it.
1.3 million cases in the database; approximately 160,000 new cases annually
### RESULTS

**BRMS Rules by Activity Type**

January 2010 to May 2012

<table>
<thead>
<tr>
<th>Activity Type</th>
<th>Rule Number</th>
<th>Description</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admission Created</td>
<td>Rule #16</td>
<td>c38-6 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #77</td>
<td>IF631-2 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #89</td>
<td>IF750-1-2 Change</td>
<td>2</td>
</tr>
<tr>
<td>Audit</td>
<td>Rule #45</td>
<td>DateFlagChange1 Hard Review</td>
<td>1</td>
</tr>
<tr>
<td>Edit Errors</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #89</td>
<td>IF750-1-2 Change</td>
<td>1</td>
</tr>
<tr>
<td>Manual Review</td>
<td>Rule #16</td>
<td>c38-6 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>2</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Rule #12</td>
<td>c38-2 Change</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rule #13</td>
<td>c38-3 Change</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rule #16</td>
<td>c38-6 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Rule #89</td>
<td>IF750-1-2 Change</td>
<td>6</td>
</tr>
<tr>
<td>Region Active Followup</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>1</td>
</tr>
<tr>
<td>SEER Edit</td>
<td>Rule #12</td>
<td>c38-2 Change</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rule #16</td>
<td>c38-6 Change</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Rule #89</td>
<td>IF750-1-2 Change</td>
<td>7</td>
</tr>
<tr>
<td>Visual Editing</td>
<td>Rule #12</td>
<td>c38-2 Change</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rule #13</td>
<td>c38-3 Change</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rule #3</td>
<td>c49-3 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #78</td>
<td>IF632-2 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Rule #89</td>
<td>IF750-1-2 Change</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Rule #9</td>
<td>c49-8 Change</td>
<td>14</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>10 distinct rules</strong></td>
<td></td>
<td><strong>197</strong></td>
</tr>
</tbody>
</table>
## RESULTS
### BRMS Activity by Site

<table>
<thead>
<tr>
<th>Site</th>
<th>Rule Number</th>
<th>Description</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>C349</td>
<td>Rule #3</td>
<td>c49-3 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #9</td>
<td>c49-8 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #16</td>
<td>c38-6 Change</td>
<td>1</td>
</tr>
<tr>
<td>C421</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>52</td>
</tr>
<tr>
<td>C442</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>2</td>
</tr>
<tr>
<td>C443</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>14</td>
</tr>
<tr>
<td>C444</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>2</td>
</tr>
<tr>
<td>C445</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>17</td>
</tr>
<tr>
<td>C446</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>15</td>
</tr>
<tr>
<td>C447</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>5</td>
</tr>
<tr>
<td>C448</td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Rule #9</td>
<td>c49-8 Change</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Rule #83</td>
<td>IF671-2 Change</td>
<td>1</td>
</tr>
<tr>
<td>C504</td>
<td>Rule #89</td>
<td>IF750-1-2 Change</td>
<td>8</td>
</tr>
</tbody>
</table>
BENEFITS

- Consistent
  - Programmed to recognize a coding error based on existing edits

- Accurate
  - Performs the identical auto-correction every time a miscode is recognized

- Functions at every level of system processing on a 24/7 basis
  - File Upload
  - Visual Editing
  - Corrections
  - Active Follow-up
  - Database Inquiries
ENHANCE DATA QUALITY

• Check existing edit sets for automation potential
  • California-specific edits
  • SEER Edits
  • NAACCR edits

• Analyze Collaborative Staging site specific schemas for potential business rules

• Review monthly edit error listings
CONSIDERATIONS

- Identifying a project manager
- Developing a project team
- Finding experienced CTRs with an aptitude for writing business rules
- Rule set maintenance
FUTURE PLANS

- Automate Consolidation
- Analyze BRMS activity and determine educational opportunities
- Automate the post-implementation testing process
RECOGNITION

- Business Rules Management Solutions Team Members:
  - Cheryl Moody, BA, CTR – BRMS Project Manager
  - Marianne Schlecht, CTR – BRMS Rule Writer
  - Marilyn Scocozza, CTR – BRMS Rule Writer
  - Tanweer Ahmad, System Architect
  - Scott Riddle, BRMS Programmer
  - Ghenadie Ciornii, BRMS Programmer, Reports
CONTACT INFORMATION

Cheryl Moody, BA, CTR
Quality Control Specialist
California Cancer Registry
(916) 779-0311
e-mail: cmoody@ccr.ca.gov