

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

BACKGROUND



- 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



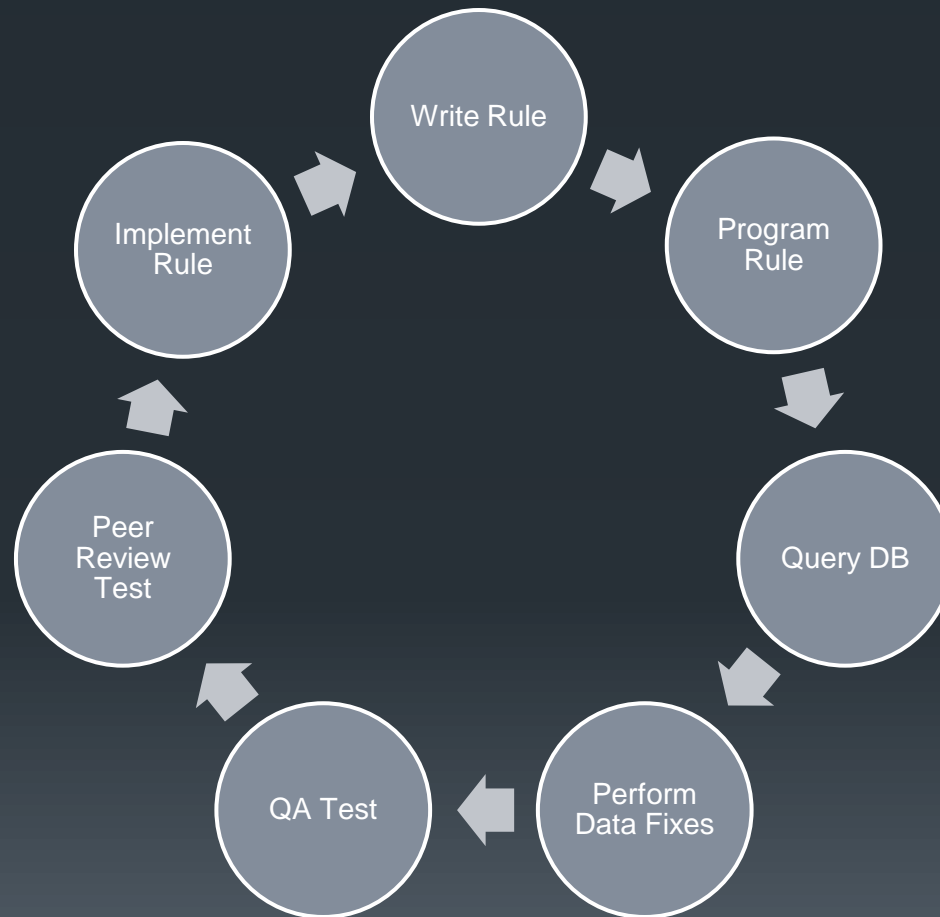
- Long Term Goal:
 - Automate manual consolidation

- 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



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

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

RESULTS

Implementation Date	Type of Rule	Description	Global Fix	Manual Fix	BRMS Fix	Total Corrections
6-2010	Class 49	DCO	1,554	91	16	1,661
6-2010	Class 38	Autopsy	0	518	16	534
3-2012	IF 631	Chemo Hosp vs Chemo Summ	5,045	0	1	5,046
3-2012	IF632	Hormone Hosp vs Hormone Summ	0	8	1	9
3-2012	IF633	BRM Hosp vs BRM Summ	7,269	0	0	7,269
3-2012	IF761	Location of Rad, Rad Date	6,288	0	145	6,433
3-2012	IF750	Breast: LNS exam, SSF 3	0	2	18	20
3-2012	IF837	Breast: CS TS, CS Ext	0	1	0	1
		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

Activity Type	Rule Number	Description	Count
Admission Created	Rule #16	c38-6 Change	1
	Rule #77	IF631-2 Change	1
	Rule #89	IF750-1-2 Change	2
Audit	Rule #45	DateFlagChange1 Hard Review	1
Edit Errors	Rule #83	IF671-2 Change	11
	Rule #89	IF750-1-2 Change	1
Manual Review	Rule #16	c38-6 Change	1
	Rule #83	IF671-2 Change	2
Miscellaneous	Rule #12	c38-2 Change	2
	Rule #13	c38-3 Change	2
	Rule #16	c38-6 Change	1
	Rule #83	IF671-2 Change	34
	Rule #89	IF750-1-2 Change	6
Region Active Followup	Rule #83	IF671-2 Change	1
SEER Edit	Rule #12	c38-2 Change	2
	Rule #16	c38-6 Change	3
	Rule #83	IF671-2 Change	40
	Rule #89	IF750-1-2 Change	7
Visual Editing	Rule #12	c38-2 Change	2
	Rule #13	c38-3 Change	2
	Rule #3	c49-3 Change	1
	Rule #78	IF632-2 Change	1
	Rule #83	IF671-2 Change	57
	Rule #89	IF750-1-2 Change	2
	Rule #9	c49-8 Change	14
Total	10 distinct rules		197

RESULTS

BRMS Activity by Site

Site	Rule Number	Description	Count
C349	Rule #3	c49-3 Change	1
	Rule #9	c49-8 Change	1
C421	Rule #16	c38-6 Change	1
	Rule #83	IF671-2 Change	52
C442	Rule #83	IF671-2 Change	2
C443	Rule #83	IF671-2 Change	14
C444	Rule #83	IF671-2 Change	2
C445	Rule #83	IF671-2 Change	17
C446	Rule #83	IF671-2 Change	15
C447	Rule #83	IF671-2 Change	5
C448	Rule #83	IF671-2 Change	1
C449	Rule #83	IF671-2 Change	1
	Rule #9	c49-8 Change	3
C504	Rule #83	IF671-2 Change	1
	Rule #89	IF750-1-2 Change	8

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
email: cmoody@ccr.ca.gov

QUESTIONS ?

