



Processing Corrections: How do we efficiently get the most important data?

Scott Wood, BA

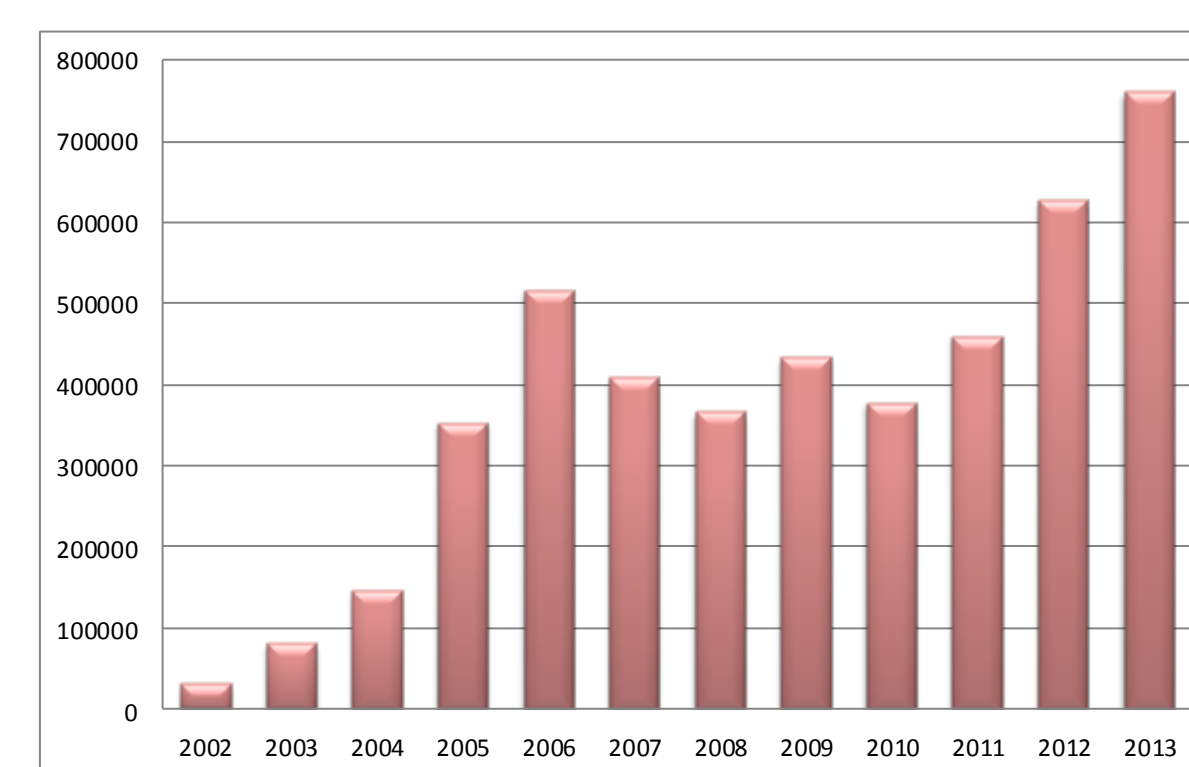
Automation Programmer IV, Production Automation & Quality Control
California Cancer Reporting and Epidemiologic Surveillance (CalCARES) Program*
Institute for Population Health Improvement, UC Davis Health System



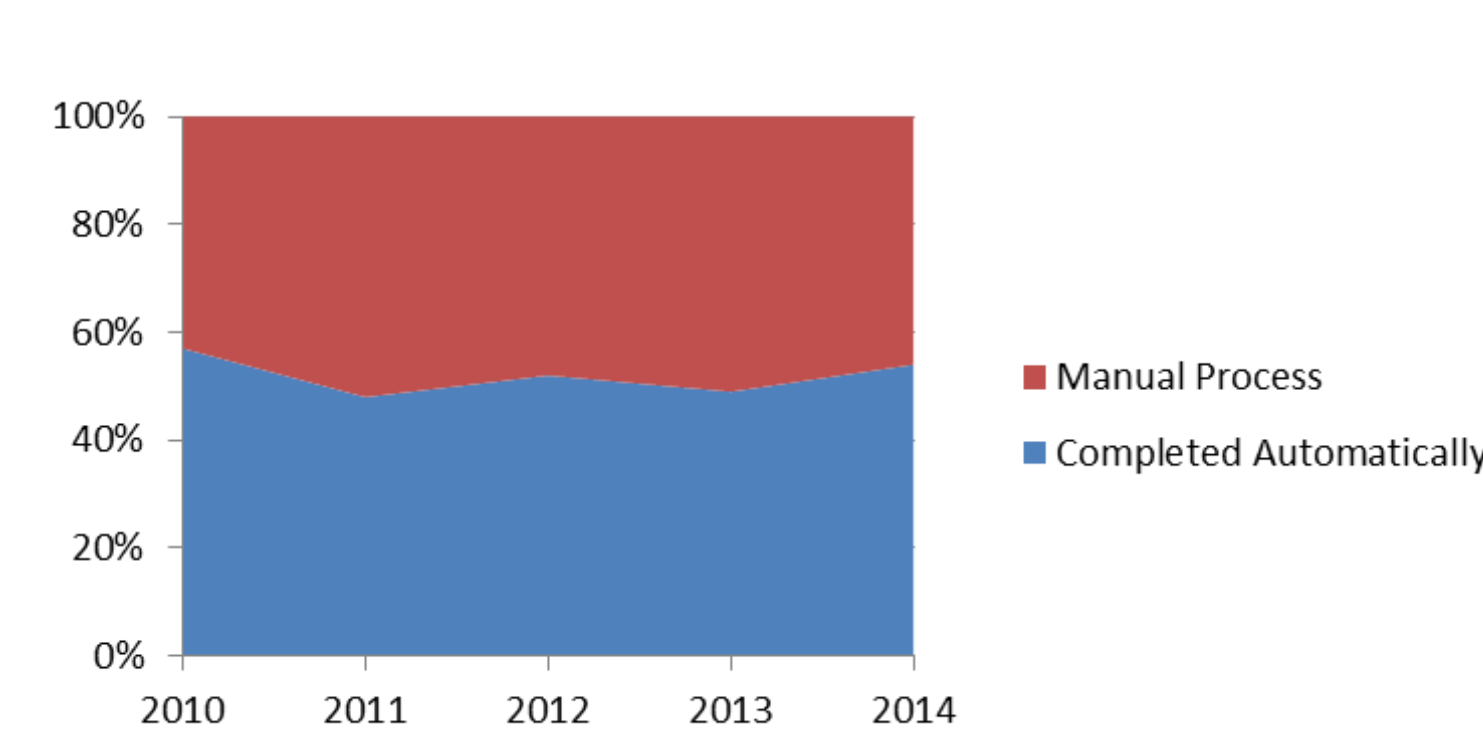
Background

In 2013 the California Cancer Registry processed nearly 775,000 correction records created by abstractors to update cancer abstracts that have already been transmitted to Eureka (our database management system). Since implementation in 2002, Eureka has processed almost 5 million correction records. While some corrections update the data without any user interaction, the rest of the corrections require user interaction to either link to correct patient or to consolidate into the existing data. If each correction/update takes one minute for a registry user to process, this equates to over 6000 hours of work, more than three full time employees statewide per year. In order to be more efficient, the California Cancer Registry has created a project to analyze corrections and determine what can be done to reduce manual work effort without compromising the quality of the data.

Corrections Uploaded Per Year



Percentages of Manual Versus Automated Work



Objective

To analyze the incoming update/correction data, frequency of data types, and current business processes to determine how to best reduce the manual work effort required to consolidate the incoming data.



Methodology

The correction/update records used in the analysis were all uploaded to the Eureka DMS between 2010 and 2014. This sample consists of over 2.2 million records comprised of 408 distinct data items. For the purpose of simplification the data items have been grouped into six categories:

- ◊ Name/Demographic
- ◊ Diagnosis
- ◊ Hospital/Physician
- ◊ Treatment/Surgery
- ◊ Staging
- ◊ Other

Overall Results

The data fields involved in the update/correction records from 2010 -2014 show the distribution of categories:

- ◊ 51 % Treatment/Surgery
- ◊ 27.9 % Hospital/Physician
- ◊ 7.9 % Diagnosis
- ◊ 7.5 % Staging
- ◊ 5.5 % Demographic
- ◊ 0.2 % Other

Analysis of all data fields currently waiting to be manually resolved show the distribution of categories:

- ◊ 57.4 % Treatment/Surgery
- ◊ 14.3 % Hospital/Physician
- ◊ 7.5 % Diagnosis
- ◊ 14.4 % Staging
- ◊ 6.3 % Demographic
- ◊ 0.1 % Other

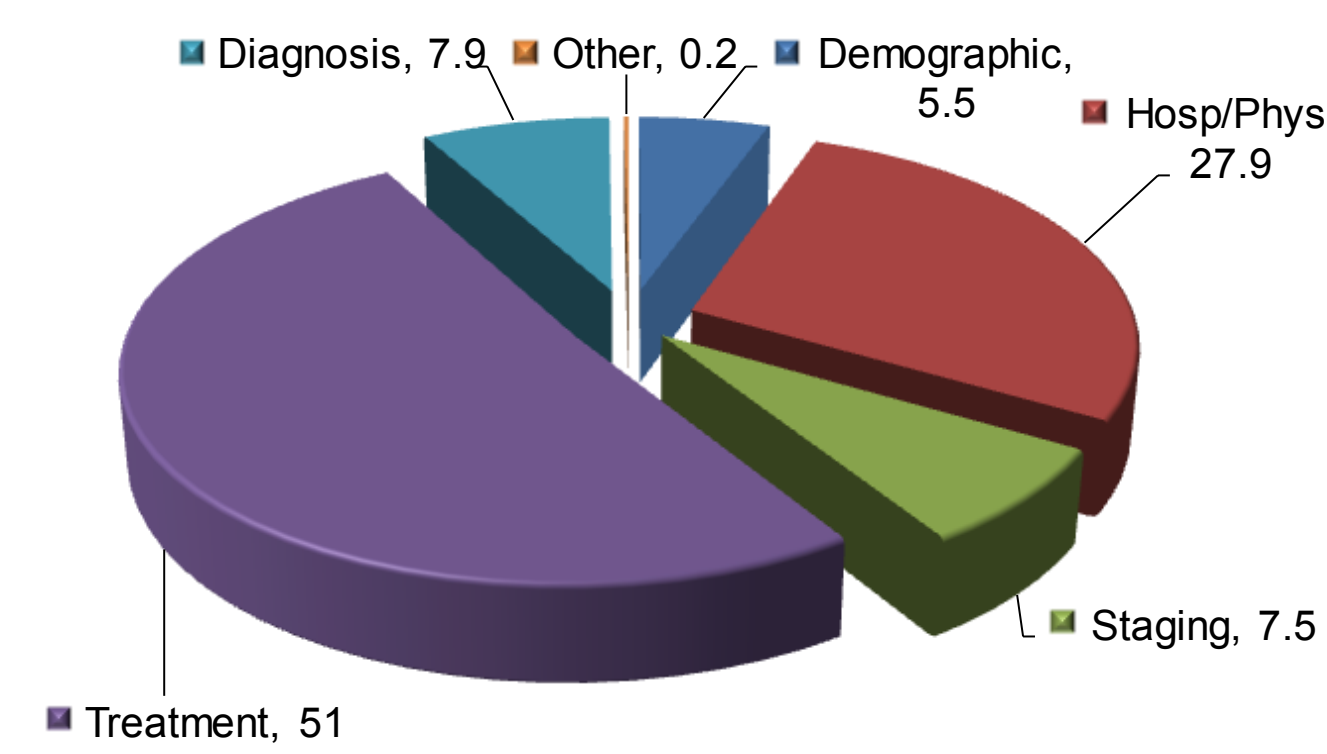
The difference between the distribution of submitted fields versus fields that are involved in manual processes to resolve are as follows:

- ◊ +6.4 % Treatment/Surgery
- ◊ -13.6 % Hospital/Physician
- ◊ -0.4 % Diagnosis
- ◊ +6.9 % Staging
- ◊ +0.8 % Demographic
- ◊ -0.1 % Other

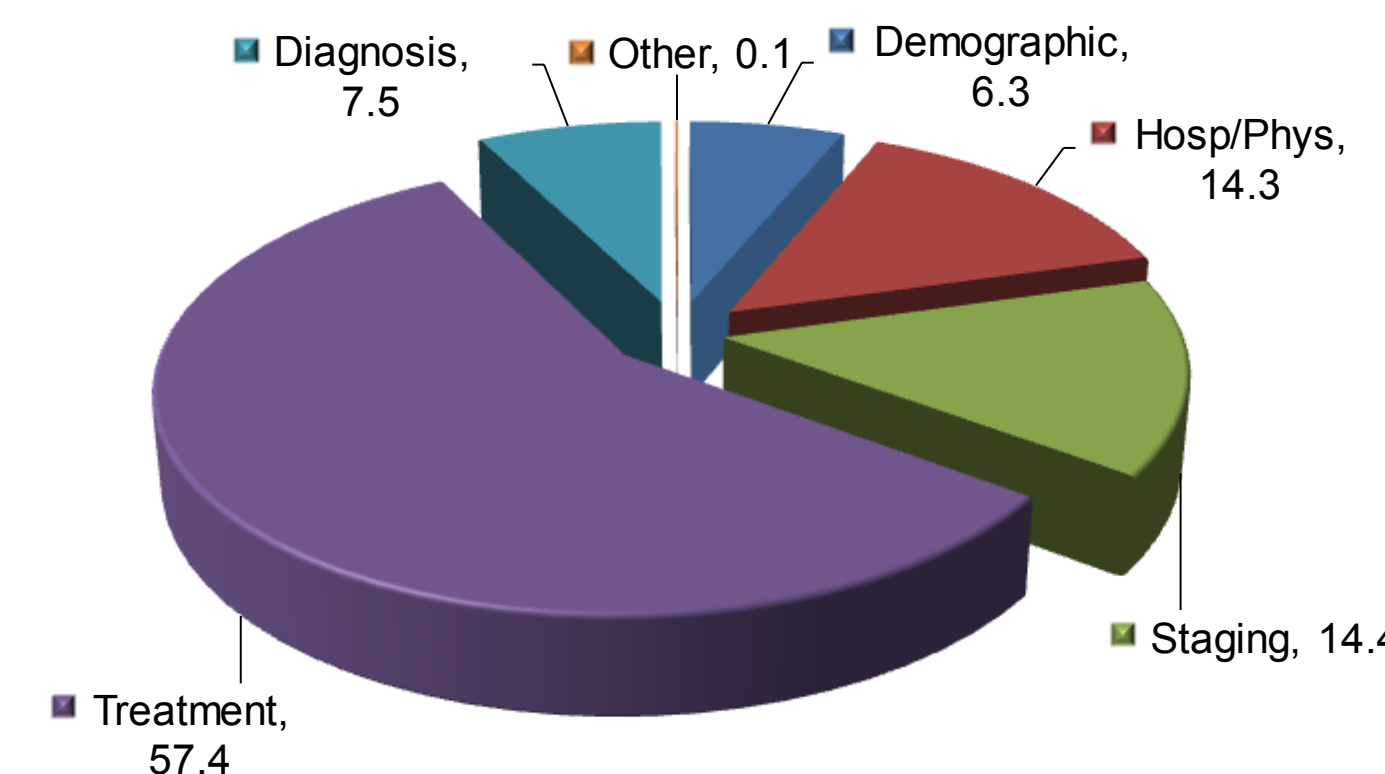
When a correction/update records falls into a manual process the functions to be performed fall into the following categories:

- ◊ Linkage Resolution - requires manual process to link to patient or admission (identifiers changed or update proceeds case)
- ◊ Manual Review - updating abstract level data requires manual process due to edit error or updates fields designated "Quality Control" fields
- ◊ Consolidation- requires manual process to consolidate data at patient or tumor level.

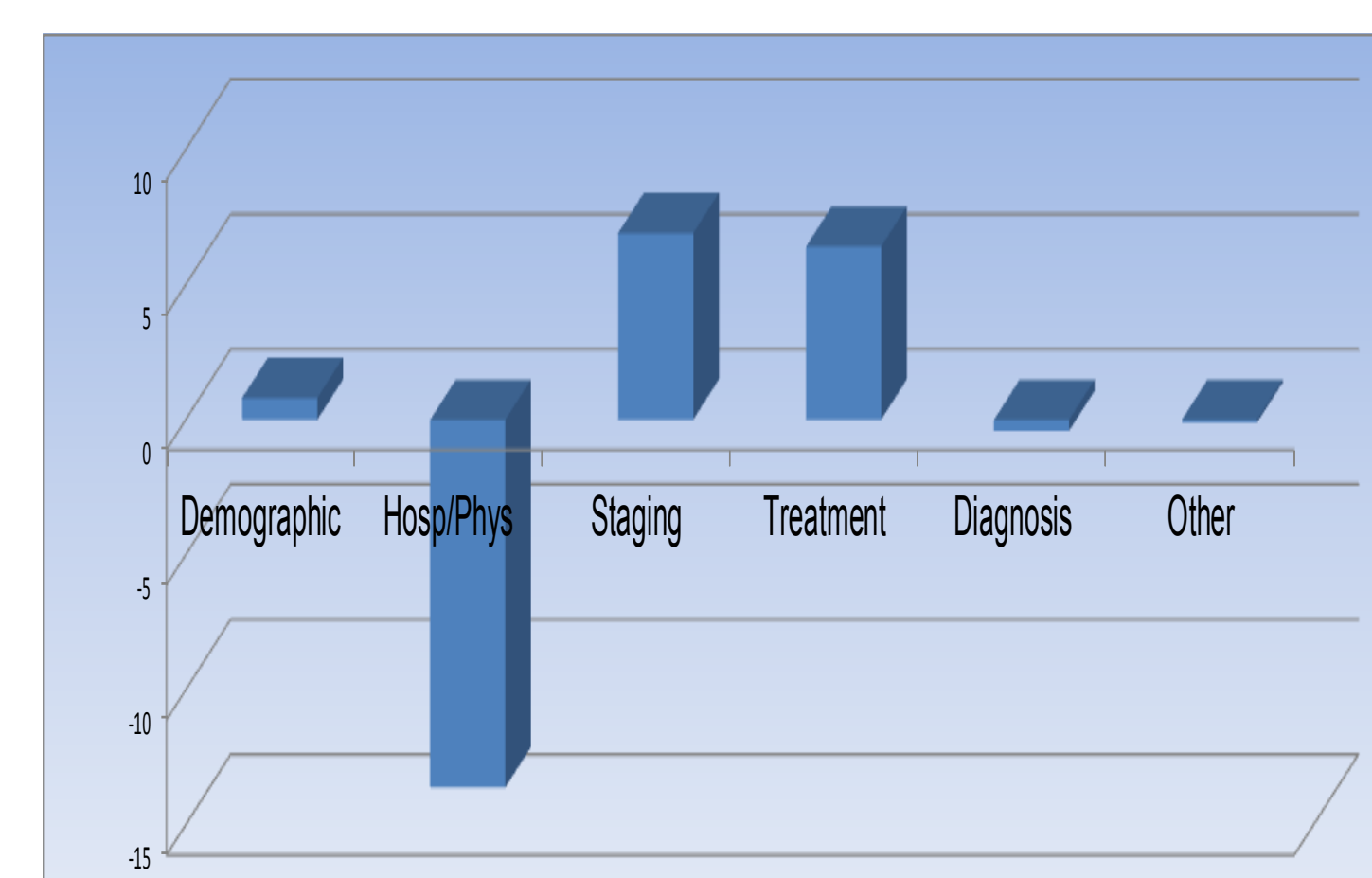
Distribution of Data Categories for all Updates/Corrections



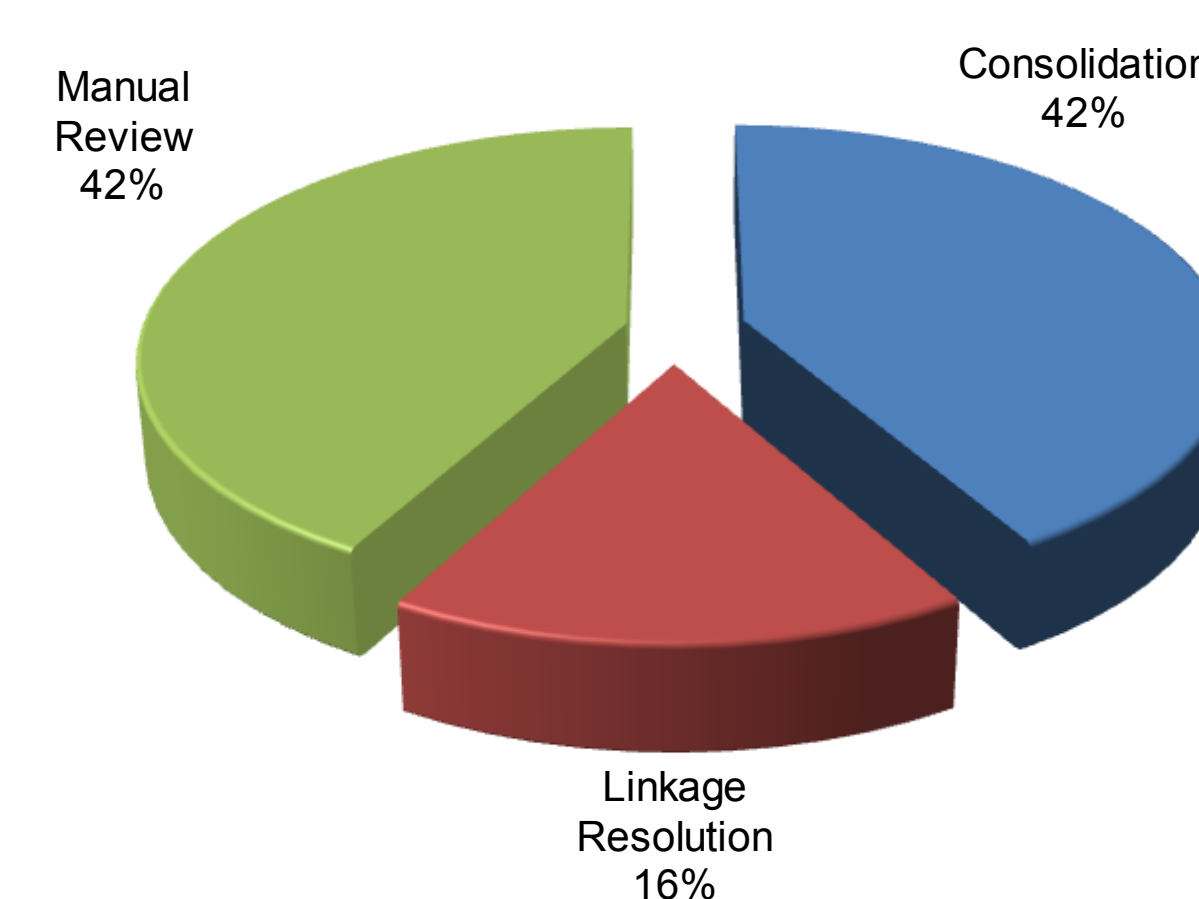
Distribution of Data Category for Manual Processes



Change between Submitted and Manual Process.



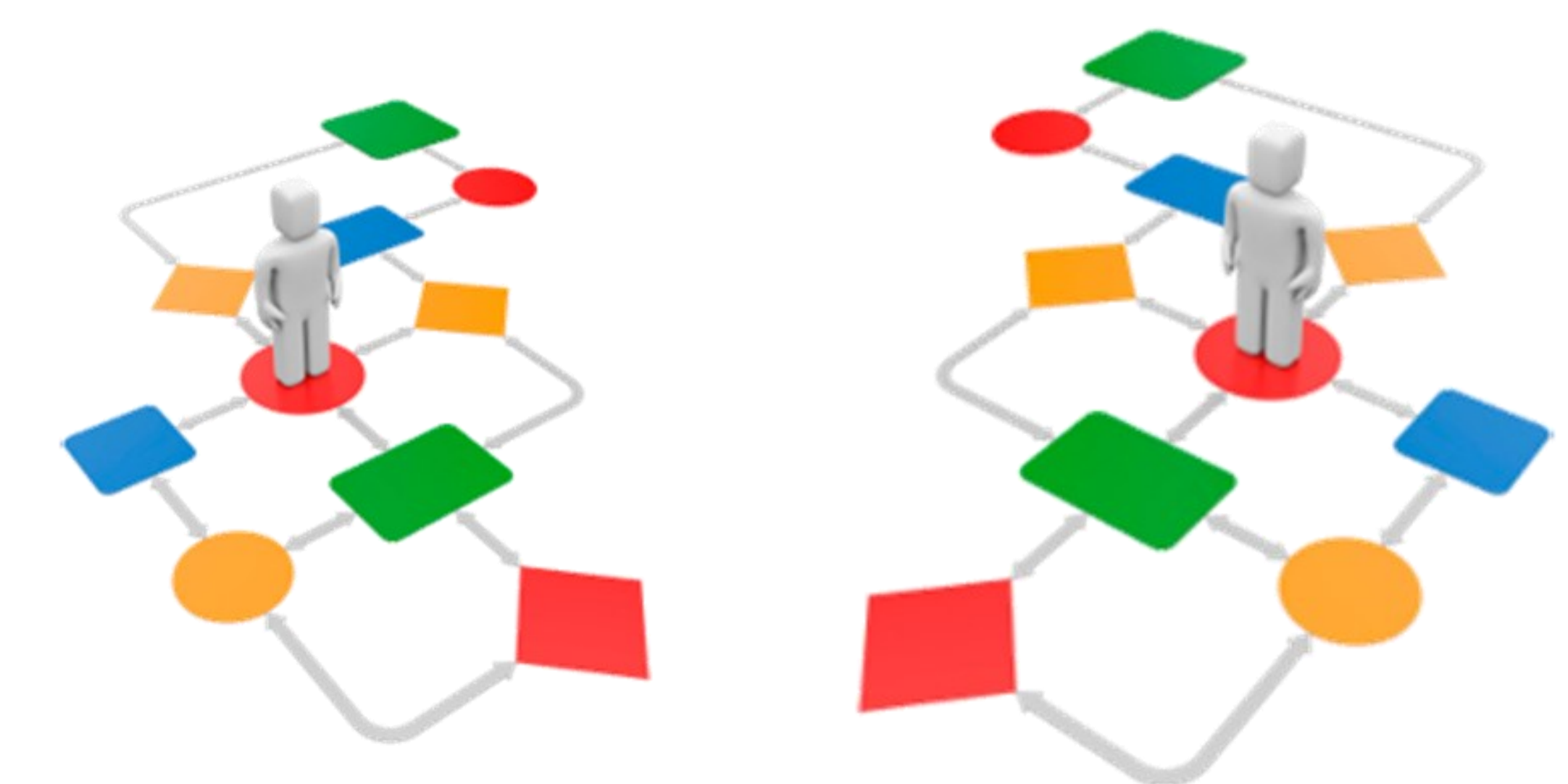
Distribution of Category of Manual Processes



Current Business Process

The Eureka DMS programmatically attempts to apply all corrections automatically. Manual work effort is created when any of the following scenarios occur:

- ◊ Does not link to an abstract already in the database
- ◊ Links to an abstract that has been deleted
- ◊ Introduces new edit errors at abstract or patient/tumor level
- ◊ Updates a case that has undergone Visually Editing (manual quality control)
- ◊ Effects a field that is designated for Quality Control
- ◊ Unable to consolidate to patient or tumor level data automatically as the incoming data does not meet the specifications that cover automated consolidation.



Conclusion

Analysis of the data involved and a review of the entire correction/update process helps to identify the areas where most of the manual work efforts are required. From this data, the following process improvements are recommended in order to increase the amount of automatically applied correction/update data; reducing the amount of manual work required to incorporate this updated data into the consolidated data sets.

- ◊ The Production, Automation and Quality Control Unit has already identified the need to only review the items if they have been quality controlled. Even if a field is designated as a "Visually Edited" data item, the system should attempt to update the data without the requirement to be manually reviewed.
- ◊ Analyze and Identify reasons cases go to linkage resolution, these cases can be indicators of missing abstract reports.
- ◊ Since over 50% of the incoming correction/update data uploaded to the Eureka DMS is related to treatment and/or surgery, and 57% of the manual work falls into this category, a major opportunity for automation and process improvement would come from analyzing treatment and surgery data to determine what decisions need to be made in order to consolidate the data. From this analysis we can add additional business rules to the consolidation specifications.
- ◊ Data fields related to staging should also be analyzed to determine what is the most common causes for the requirement of manual review. If these fields (or any others) commonly cause edits to fire, then the Production, Automation and Quality Control Unit can determine if these edits can be resolved automatically. If so, these business rules can be developed and added to the existing set of automated business rules that correct the data without the need for manual work.

*The CalCARES program partners with the California Department of Public Health (CDPH) to manage the operations of the state mandated California Cancer Registry program