The Journal of
IMAGING SERVICES
ISSN1055-8098
Professional Development Information for Production Document Imaging Operations
Volume 17, No. 1

Data Entry Best Practices

by Bill Krautter

Accuracy of data has been a life-long achievement for our industry. As far back as the old 'garbage-in/garbage-out' days, we have recognized the importance of accurate information. Many statistical models have been developed and presented while people were aghast with the results and costs of inaccurate data. Ever since the 'verify-punch' feature was added on card-punch stations, technology has developed between human and machine to assist data accuracy. When technology alone could not assist people to produce the desired results, rules and policies were developed and implemented. For over fifty years, data accuracy among people and technology has evolved into industry practices, and those which produce the desired results are the best.

Defining Data Entry Accuracy

Companies achieving data entry accuracy are successful when they define the client-desired data accuracy. This success depends on management's ability to develop realistic goals for accuracy, stated in terms that all producers and users of the data can understand and visualize. Often times "quality" is a significant part of the company's mission statement. Then this becomes the mission of the data processing organization when converting and processing information. As broad goals are developed to assist in the deployment and attainment of the company's mission, specific goals are used at the task and operational levels of the organization to carry out the mission.

Data accuracy goals are an integral part of carrying out the company's quality mission. Data accuracy goals should be stated as a percent of accurate data converted. This might look like, "The company's data accuracy goal for non-verified characters is 96% of all characters converted," or "The company's data accuracy goal for verified characters is 99% of all characters converted." The goal statement must be clear to all people within the organization and also those outside the company who are affected by data accuracy.

The definition of the accuracy goal would be useless without a process of measuring data involved in achieving the goal. Part of the measurement process is developing an accuracy standard. The standard should reflect the minimum acceptable accuracy rate by the company. Each data entry operator's data or keying accuracy rate is compared against the standard.

It is very important for the operator to know exactly how their accuracy rate is calculated. For example, if every single keystroke or if every *nth* (factor of 1 to many) keystroke is measured in the calculation. The calculation should include the correct number of keystrokes produced, divided by the total number of keystrokes required, and produce an accuracy rate as a percentage of 100 percent.

Accuracy Rates

All accuracy rates are based on some method of verification. Methods can range from double-key entry verification, sight verification, program edits, field validation, and post-processing result reports. The method of verification may vary between data conversion applications, but data entry operators should know what constitutes an error. Errors should be defined by types: keystroke, field, or application interpretation. For example,

continued →

does a transposition of two numbers count as one or two keystroke errors? Do three characters keyed incorrectly in the same field count as one field, or three keystroke errors? It is also important to define when an error becomes an error, in cases where an operator corrects their own error before the verification process occurs.

Publishing the measurement process has the greatest positive effect on defining accuracy. The mere printed text of the defined measurement process causes all stakeholders to take ownership of data accuracy. Once documented, all people know how data is converted, measured, and reported. The publication of the process and results should extend to all users of the data, both internal and external to the organization. A satisfied customer often is one who understands how a less-than-perfect result can occur.

Best Practices Begin with the Employment Process

In any precision application, the human element will be the most vulnerable to change. These human changes or inconsistencies will directly affect the accuracy of the desired results. Reducing the frequency of inconsistencies is important to producing dependable accurate data conversion. People need to know what the expected results are, and see the environment where they are expected to perform. This begins with the employment process of the individuals interested in a data entry position.

Companies should thoroughly describe and allow the applicant to view the working environment. Data entry position applicants should see a typical workstation, understand how documents flow in and out of the work area, have a typical work day described, and inform them how long they can expect to be seated and working without interruptions or breaks. Part of the job interview process should include a review of typical data entry conversion projects that a new employee would be expected to work.

It is also helpful to show operator applicants the progression of how data entry work assignments are made and how they can graduate to more challenging conversion assignments. The pre-employment process should include a data accuracy test. This test should be a good example of the type of data conversion the data entry applicant is expected to perform during their employment. The results of the test are shared with the applicant and compared against the standard data entry accuracy rate to establish if the applicant meets the company's minimum acceptable accuracy rate.

The most important aspect of the pre-employment process (as it relates to data accuracy) is communication to the operator applicant of the statistical goals, data conversion processes, and measurements used by the company. Since data accuracy should be a component of the employee's wage, the operator applicant should be interested in how this measurement is calculated and figured into their salary. The applicant should be given the minimum acceptable data accuracy rate to maintain their employment during the pre-employment process.

As with other employment policies, the minimum acceptable data accuracy rate, measurement procedures and policies, wages related to maintaining the rate, and any possible disciplinary actions should be in writing and signed by each operator applicant when they are hired.

The Practice of Continuous Measurement

Measurement validates the progress towards achieving data accuracy goals. The practice of constant, consistent measurement helps attain and maintain data entry accuracy. When measurement is a reliable process and an active part of data capture/conversion tasks, the resulting information is valuable to the organization and its customers. Besides producing a rate of accuracy, this process motivates employees and identifies improvement opportunities.

When a data entry operator knows transcribed data will be continuously checked for accuracy, there is a height-

ened attention to correct keying, data interpretation rules, and exception data processing.

A constant, but non-intimidating reminder about doing an accurate job will improve operator morale. The more frequent the publication of the measurement results, the better an operator feels about doing a consistent, accurate job. Frequency of measurement also gives more opportunity to correct inaccurate conversion processes or tasks before too many errors enter the resultant data files.

The process of measurement also includes identification and categorization of errors so corrective action can be defined and initiated. Errors are separated into groups based on overall corrective strategies. It is helpful to statistically measure the number of errors in each of these groups, using a method to determine the significance of the error relative to other errors. This will be helpful when prioritizing improvement opportunities later during implementation. The corrective strategies usually include: operator interface to equipment (i.e., transpositions, keyboard mapping), re-training (application interpretation, data rules) and automation (data edits, table validation and custom programming).

Evaluate Opportunities in Specific Areas for Improvement

Every error discovered during the measurement process is an opportunity for improving data accuracy. Each error should be evaluated and a strategy developed to prevent the error from re-occurring. The strategy is developed into an improvement process that is implemented based on a priority set by the significance of the data error. Improvement processes maybe designed to correct multiple error types in the same group. For example, a re-training process could correct multiple data conversion rule errors and form field interpretation errors. Table validation program routines could correct transposition and "data out of range" errors.

The magnitude and frequency of the error will also determine a strategy for process improvement. This could occur when a field interpretation error is made by a majority of operators, but re-training will take too long or become too expensive. A programmatic software change to the data entry program might be faster to deploy at a lower overall cost. There will be times when error correction opportunities will require input from the various stakeholders in the data conversion application. These people could be part of adjoining data processes, or users of the resulting data. A data preparation stakeholder might be consulted to insert a process to identify (to the data entry operator) an often-omitted field of information.

The intermediate stakeholder and user of the converted data may implement a post-process which will scrub a data field more accurately and consistently, rather than relying on human rules-based interpretations. All incremental costs of each error improvement processes should be identified and quantified. If stakeholders are involved in the changed process, those costs should be included. Once quantified, each error correction opportunity should be ranked in order of greatest impact to maximize actual overall improvement of data accuracy when compared to the data accuracy standard.

Implement the Improvements and Measure Incremental Results

The implementation of a new data conversion process should be deployed in a test environment to isolate the effects of the error improvement. This allows each new process improvement to be measured independently and record its statistical relevance to improved data accuracy. This approach will allow the results of each error improvement process to be measured against the results of the original process. Once the improvements are measured and results compared individually, these improvements can be combined using a layering implementation strategy to maximize the total improvement in overall data accuracy. This will identify the error improvement processes that have a diminishing effect on overall data accuracy when combined with other processes.

It is possible that some error improvement processes will have a negative effect and actually decrease data accuracy. An example of this is when a validation table is added to the data entry program to improve keying accuracy. The values of the table are dynamic and the table update interval is less frequent than the new values appear on data conversion documents. A negative effect occurs when a correct table value is keyed from the document, but a table validation process marks it as incorrect because the program tables are not updated in a timely manner with new values.

Using the layering strategy of combining error improvement processes, the optimum pairing of these processes can be identified and implemented into a controlled test processing environment. Once implemented under test conditions and the results are calculated, the best combination of improvement processes are determined. The costs of deploying the combined error improvement processes are identified and tallied.

Review Results and Cost Justify Improved Data Accuracy

The practice of data entry accuracy is a continuous mission for most organizations. Multiple opportunities for error improvement are frequently identified and evaluated. Statistical data related to the data entry conversion accuracy is paramount to the success of the improvement processes implemented. Original actual data accuracy rates, company standard data accuracy rates, and new process improved data accuracy rates must be maintained for proper evaluation. The improved accuracy rate results are measured separately from the original actual rates, and the company standard accuracy rates. Costs associated with original, standard, and improved accuracy rates should also be part of the statistics gathered when justifying improved data accuracy.

Cost justification models should be developed when evaluating various single or combined error correction process improvements. When presented with multiple opportunities for data accuracy improvement, benefits of the improved data accuracy coupled with the cost of the correction process should be presented in a comparable form. This will allow a clear comparison of varying levels of improved data accuracy along with the related costs. When evaluating multiple opportunities in a justification model, the opportunity creating the greatest improved data accuracy rate should be prioritized first. In most justification models the cost of accuracy at the point of conversion will be less than the cost of correcting inaccurate data after that point.

Bill Krautter, President, Statistical Service Corporation of Austin, 8870 Business Park Drive, Austin, TX 78759; Phone: 512-795-5000, Email: wfk@statco.com