



IBM Intelligent Forms Processing (IFP)

2D Bar Codes Increase Accuracy, Decrease Costs



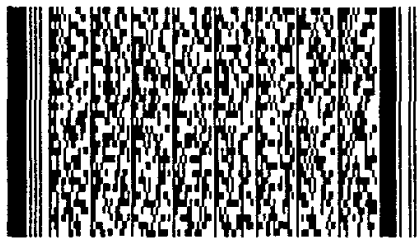
Susan Cargill, IFP Project Manager & Marketing Strategist
IBM Global Services, Business Consulting Services
800 North Frederick Avenue, Gaithersburg, MD 20879
301-240-2124, scargill@us.ibm.com

<http://www2.clearlake.ibm.com/gov/ifp>

Enabling Government & Industry to Automate Data Capture Processes

IBM Intelligent Forms Processing (IFP) automates the capture of data from paper documents. IFP employs Optical and Intelligent Character Recognition (OCR/ICR) and sophisticated data repair technologies to ensure highly efficient, virtually error-free scanning and processing of data, images, and two dimensional (2D) bar codes from simple and complex forms of multiple lengths, sizes, and shapes.

2D Bar Code Reading Improves Data Capture Accuracy



2D bar codes encode data both horizontally and vertically so that more characters are encoded than in traditional one dimensional (1D) bar codes. The most popular 2D format is PDF417, which can store approximately 1000 characters. Key benefits of using 2D bar codes include:

- Data security because bar codes cannot be read by the human eye
- 100% accuracy because the bar codes have built-in redundancy of characters

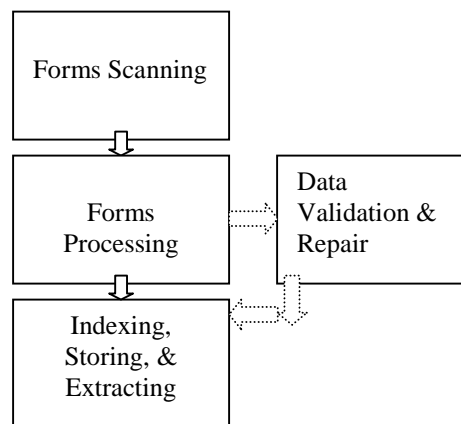
A PDF417 bar code includes several user-definable parameters that control the size and shape of the bar code:

- *Error correction level.* Number of ‘check digits’ encoded in the bar code. Higher levels improve readability by adding more check digits and typically do not make the bar code much larger.
- *X-value.* Width of the narrowest bars in the bar code. Higher values improve readability but make the bar code much larger.
- *Columns.* Number of columns of data in the bar code. Higher values make the bar code wider and shorter but reduce readability by making the decoder more sensitive to skew.
- *Y/X ratio.* Height of each row and the shape or aspect ratio of the smallest patterns in the bar code. Higher values improve readability but make the bar code taller.

Note that many forms contain both a 1D bar code for form recognition and a 2D bar code to be used instead of OCR/ICR. In contrast to 1D bar codes, 2D bar codes are used to

store data from forms, not just to help recognize forms, and 2D bar codes include built-in error detection and correction capability.

Manual key-entry of data from forms is a slow process prone to errors. And, while traditional forms processing of paper documents using a high-quality OCR/ICR system provides substantial cost savings and greater accuracy, data from 2D bar codes can be read as correct from the initial scan into the forms processing system. Therefore, use of 2D barcodes as a part of an automated forms processing and OCR/ICR system adds even greater speed and accuracy to the process. In fact, if the 2D barcode is readable, it is 100% accurate and data validation and repair is not necessary. 2D bar code reading means fewer data capture errors, and fewer errors mean fewer customer complaints.



Benefits of using 2D bar codes on forms include:

- High read rates
- No need to separate 2D bar code documents from documents without a bar code before scanning
- Data capture with high confidence and no misreads
- Much higher data accuracy than with data entry or OCR/ICR
- Significant reduction in manual resources required for data repair
- Faster processing of documents through IFP

How IFP Processes 2D Bar Code Data

IFP lets you define each data capture field using its powerful Forms Training Utility (FTU). A 2D bar code is defined as an input field like an OCR/ICR field, and the font type is “PDF417” instead of hand-print or machine-print used for OCR/ICR.

IFP Forms Processing (FP) reads the 2D bar code, places the desired results into output fields, and sets field coordinates to display to the user the expected OCR/ICR field rather than the 2D bar code (since the 2D bar code is not humanly readable). If the 2D bar code

is not readable, IFP performs OCR/ICR on the expected input field as if the 2D bar code was not present. So, the 2D bar code can only improve the OCR/ICR results.

As part of its System Services, IFP also prepares a 2D bar code administrative report that contains counts and read rates.

IBM first employed the 2D Bar Code function of IFP in 2002 for the Comptroller of Maryland's 2001 Spring Tax Season. Four percent of the Maryland returns or 114,000 returns contained 2D bar codes that were processed by the IBM IFP system. The results were 100% data capture accuracy when the bar codes were read successfully, 95% of the time. The implementation for Maryland includes Federal Tax Association (FTA) syntax validation, two different decoding algorithms that provide "voting" similar to engine voting in the OCR world, erasure support in the event that the bar code contains no value, and OCR backup in case the bar code read fails for any reason.

Of note to Tax customers, the FTA has collaborated with State governments, software developers (such as Intuit and Nelco), and system integrators to develop a standard recommended format for encoding taxpayer data. The recommendations include:

- Error correction level of 4
- X-value of 15 mils or greater if reading bar codes from 200 dpi images
- Columns with a 2:1 width/height ratio preferred
- Y/X ratio with a 2:1 width/height ratio preferred
- Data encoded as multi-line text with one field per line
- Omission of punctuation as much as possible

IFP fully complies with the FTA recommendations. Refer to the FTA Web site for more information at www.taxadmin.org.

IFP: Use IBM's Integration Advantage to Improve Your Bottom Line

IFP can help you:

- Reduce data-entry work and costs
- Increase accuracy and speed of data capture to meet tight deadlines
- Decreases paper handling, document management, and warehouse storage costs
- Improve customer satisfaction by providing instant access to forms and faster processing

A cost-benefit analysis for forms processing:

Studies have shown* that, for an average paper form, the cost of key-entry is \$2.00 per document. Therefore, processing a low volume of 1000 forms per day using IFP instead of labor-intensive key-entry resources can payoff in as short as one year. Consider this example:
\$2.00 x 1000 forms per day = \$2,000.00 per day to process 1000 forms
\$2,000.00 per day x 260 business days a year = \$520,000.00 cost per year

Assuming 75% of the typical cost for form key-entry is the cost of labor, 75% of \$520,000.00 is \$390,000.00. Assuming the cost of a key-entry person is \$35,000.00 per year, this \$390,000.00 is made up of 11 people. With IFP, you can reduce this number of people to no more than three:

Your labor cost becomes 3 x \$35,000.00 or \$105,000.00, and **your savings is \$390,000.00 - \$105,000.00 or \$285,000.00 per year**. Assuming an estimated investment of \$250,000.00 for a standard IFP system configuration customized for your unique forms:

\$285,000.00 savings - \$250,000 system cost = return on your investment in one year.

* Source: Fletcher Spaght Market Study 12/2000