



“The OCR solution has moved well beyond our expectations for improved data quality. In some cases, we’ve reduced report processing time from a half hour to five minutes. That’s huge for us. In managing contractors, we’re now moving away from junior-level ‘data entry operators’ to fewer senior-level staff who will be able to spend more time on initiatives such as trend analysis and quality control. Our organization is also pursuing other OCR solutions for applications such as the reporting of pipeline incidents and training registration.”

– Michelle Glode
Operations Research Analyst
U.S. DOT

Electronic government is changing the way agencies operate and interact. The Government Paperwork Elimination Act (GPEA), which provides for electronic submission of forms, also requires that systems protect confidentiality, ensure that information is not altered in an unauthorized way, and make it available when needed. Information received on paper must be safeguarded through a reliable process for capture, transfer, and storage. The Pipeline and Hazardous Materials Safety Administration (PHMSA) of the U.S. Department of Transportation recently mastered such a challenge in managing information pertaining to hazardous materials incidents. Through a Kofax advanced forms recognition solution implemented by GxP Partners, PHMSA tripled its processing capacity of HMIS 5800.1 forms submitted by carriers of hazardous materials while increasing accuracy and more efficiently deploying consulting staff.

The Situation: Hazardous material incident reports (5800.1) are used to identify problems with packaging; loading, handling, and unloading operations; and inadequacies within transport vehicles, freight containers, and cargo holds. Potentially serious problems can trigger regulatory or enforcement actions. Although reports can be submitted electronically, PHMSA had seen a steady increase in the volume of 5800.1 forms received by U.S. mail. Incident reports were entered manually into the Hazardous Materials Information and Reporting System (HMIS). The contractor staffing model for this activity assumed data entry of 1,600 reports per month; PHMSA was receiving an average of 2,400 per month – 50 percent more than the staffing model supported. Increasing the administration’s contracting resources was not in the budget. A backlog could result in delayed or inaccurate reporting of transportation safety data to Congress, the public, the media, and other stakeholders.

The Solution: Given the complexity and varying format of the 5800.1 form, PHMSA needed an advanced optical character recognition (OCR) technology to capture information based on the content of the text rather than its physical location on the page. GxP Partners provided a turnkey solution including: scanning and data capture software, installation services, configuration to extract and automatically validate fields from the form 5800.1 into the HMIS Oracle database, end-user training, and technical support. The solution provided faster scanning time, automatic page separation based on bar codes, automated and more accurate data extraction based on keywords rather than zones, built-in validation, and the ability to capture both printed and hand-written data. Since the operators’ role shifted from data entry to limited data entry and verification, the processing time per form was reduced by more than 65% from 30 to 10 minutes – saving over 200 hours a month.

The Process

After developing a working prototype in two weeks, GxP Partners delivered on a 55-day production plan for the full system. Services included: project management; installation and configuration of Kofax workstations, servers, and remaining HMIS form fields; installation of software and five validation workstations; creation of system installation documentation and validation scripts; testing; training; and on-site support for week one.

The Products

The GxP Partners' turnkey solution enables PHMSA to extract more than 200 fields from 5800.1 forms with more than the guaranteed 90% accuracy rate. The staff now validates the data electronically before it is released to the HMIS Oracle database. GxP Partners implemented the INDICIUS forms processing software for scanning 25,000 images per month in a production environment with a back-up scanner. Developed by Kofax, the leading provider of technology for Intelligent Capture & Exchange, the solution features:

Kofax Ascent Capture – delivers powerful, production-level document and data capture. Ascent Capture collects paper documents, forms, and e-documents from scanners and other sources; transforms them into reusable information; and delivers it all into line-of-business applications, databases, and archives. It can be deployed quickly and easily for a simple application or dramatically enhanced with modules and connectors to create sophisticated business solutions.

Virtual ReScan (VRS) Image Enhancement Module – an add-on to Ascent Capture that is the de facto standard for increasing scanning productivity. VRS automatically performs sophisticated correction and enhancement on scanned documents, so they emerge straight and clear. It offers: auto rotation, auto brightness, auto contrast, auto page length detection, despeckling, deskew, advanced clarity, cropping (black border removal), and blank page detection.

INDICIUS – a suite of flexible, customizable software modules that intelligently automates the classification, sorting, and separation of paper and electronic documents. It then extracts and validates information contained in the documents.

The solution met all PHMSA requirements, including the ability to process complex forms with varying formats and extract data with a minimum 90% accuracy rate. It also extracted hand-written and machine-printed data, as well as bar codes, check boxes, and tabular data.

Repeatable Successes

Ask GxP Partners to help you calculate the ROI on a Kofax solution if your organization needs to:

- Unify information from multiple paper and electronic sources and formats
- Automate the time-consuming tasks of processing, extracting, and classifying information from a wide variety of documents
- Integrate extracted information automatically into multiple business or government applications and processes
- Exchange information with constituents, customers, partners, and vendors in a broad range of formats
- Enable stakeholders to easily interact with your organizational processes to provide and receive correct, complete information
- Increase document processing capacity without increasing headcount
- Ensure data accuracy with built-in validation throughout the capture process, improving document case handling and customer service

