

# Kofax Informe sobre produtos

“Fazer negócios é conversar”

– David Weinberger

The Cluetrain Manifesto:

The End of Business as Usual;

por Christopher Locke, Rick Levine,

Doc Searls e David Weinberger

## Classificação e separação

### A captura de alto volume de documentos exige classificação e separação

Com frequência, as tecnologias de automação de documentos são vítimas de seu próprio êxito. Por exemplo, o layout original do teclado de uma máquina de escrever, que ainda é usado praticamente em todos os teclados de hoje, foi especificamente projetado para compensar as falhas dos componentes mecânicos da época.

Dificuldades semelhantes continuam surgindo à medida que a automação gera problemas no sentido ascendente e descendente do fluxo de trabalho. A velocidade de rendimento nominal dos dispositivos modernos de captura de documentos em alta velocidade raramente é atingida, já que gargalos na preparação, classificação e correção de erros de documentos em geral tornam os processos de trabalho lentos. Além disso, o que parece ser intuitivo para pessoas pode ser extremamente difícil de reproduzir com os softwares.

A captura de imagens de documentos – o processo de obter documentos em papel e convertê-los em uma imagem digital ao digitalizá-los – abriu um mundo de vantagens de automação de processos. Os documentos podem ser transmitidos eletronicamente, ao invés de serem transportados fisicamente. Grandes quantidades de informações podem ser armazenadas em espaços menores, enquanto as funções de análise de dados, fluxo de trabalho dos processos, pesquisa e recuperação usufruem de uma melhoria exponencial na produtividade. Contudo, as

melhorias na velocidade de conversão de documentos enfrentaram uma série de desafios que, inicialmente, não haviam sido previstos ou que não foram superados com facilidade.

A tecnologia de Reconhecimento Óptico de Caracteres (sigla OCR, em inglês) é uma conquista surpreendente, oferecendo a capacidade de extrair dados de formulários. Para o olho humano, “2” é dois e “5” é cinco; fazer com que o software chegue à mesma conclusão – com precisão comparável – não é um feito simples. Embora os mecanismos de reconhecimento tenham sido bastante aprimorados, dispor de um software que compreenda o conteúdo e o contexto de uma página impressa – e que seja capaz de tomar decisões sobre classificação, encaminhamento e separação de documentos – é ainda mais complexo.

## O alto custo de documentos

Apesar do aumento de uso das transações eletrônicas, os volumes de papel continuam crescendo. No relatório: Summary of Worldwide Capture Software 2006–2010, a Harvey Spencer Associates prevê uma necessidade contínua de capturar mais informações de entradas não estruturadas em sistemas de gerenciamento de documentos e processos de negócios específicos. “Há uma necessidade maior de gerenciar os negócios com mais eficiência e de compreender o valor que esse software traz”, relata Spencer. “Estimamos que o setor de captura irá crescer para US\$ 2,42 bilhões em 2010 – uma taxa de crescimento anual composta de 16,4% – excedendo as taxas de crescimento mundial de TI e ECM.”

A pesquisa da Spencer também aborda o alto custo do manuseio de documentos. Os pesquisadores relataram a taxa média de desempenho do manuseio de correspondências para um cargo administrativo de preparação de documentos: Alguém que seja habilidoso no manuseio de documentos pode

processar entre 750 e 1000 páginas por hora – ou seja, retirar de um envelope, classificar, empilhar e preparar para a digitalização em lotes. Isso não se refere à digitalização propriamente dita. De acordo com Spencer, esse número está sujeito a alterações com base na quantidade de reparos que os documentos possam exigir.

Esses números destacam a enorme despesa relacionada ao processamento manual. Quando multiplicada pelos milhões de documentos manuseados por muitas organizações, os custos poderão ser espantosos. Embora o manuseio mecânico tenha progredido ao permitir a abertura de envelopes e a digitalização de imagens dos documentos para substituir o que historicamente era uma operação manual, os avanços dos softwares têm sido mais graduais.

Embora scanners para um alto volume de documentos fizeram com que a conversão de milhões de páginas de papel pareça rotina, uma vez que os volumes de documentos processados aumentaram, as expectativas quanto aos sistemas também cresceram. Por vezes, as demandas sobre os sistemas de processamento determinam suas capacidades.

Um desafio que sobrecarregou os processadores de alto volume de documentos foi a necessidade de diferenciar os documentos individuais entre si. Isso é particularmente difícil ao lidar com grupos de documentos de várias páginas, como aqueles encontrados com frequência em formulários de processamento de empréstimos ou na sala de correspondência típica. Apesar de normalmente ser apenas uma questão de alguém dizer “aqui é onde um documento de formulário de empréstimo termina e aqui é onde começa o relatório de crédito”, essa determinação se mostra bastante complexa para um sistema automatizado.

Porém, devido à necessidade de um volume grandioso, os processadores são obrigados a alimentar grandes quantidades de páginas não classificadas em scanners de documentos, organizadas por critérios que não vão muito além do que a ordem na qual elas são recebidas. A separação de páginas é um processo manual que requer muita mão-de-obra, seja ao digitalizar cada documento individualmente ou ao inserir “folhas de separação” entre cada documento. Outra opção é a separação posterior à digitalização, na qual um auxiliar administrativo deve fazer uma associação de documentos de várias páginas, assim que os mesmos tiverem sido digitalizados (vínculo pela imagem).

Além de ser uma forma dispendiosa e ineficiente de separar documentos misturados, o trabalho humano está sujeito a erros.

## Automatização da extração de dados

Na maioria dos sistemas automatizados de processamento de formulários, dados específicos devem ser extraídos. Por exemplo, os sistemas de contas a pagar devem saber o valor em dólares a ser pago, portanto esses sistemas – sejam manuais ou automatizados usando a tecnologia de Reconhecimento Óptico de Caracteres (OCR) – procuram campos de dados como: “Saldo devedor” ou “Total da fatura”. A ampla variedade de termos para campos específicos, sua localização em um documento não estruturado e o número praticamente ilimitado de tipos de dados tornaram a automatização do processo de extração particularmente desafiadora. Além disso, esses desafios estão presentes mesmo depois que o tipo de documento é conhecido. (Voltando ao exemplo, nenhuma pessoa nem o OCR pode começar a procurar pelo “valor devido” até que o documento seja reconhecido como uma fatura.) Portanto, a primeira etapa na extração de dados é o processo de identificação de documentos.

## Identificação de documentos

Além de atuar como um meio de extração de dados, a classificação precisa de documentos também serve para identificar documentos que fazem parte das transações comerciais diárias. Se um documento for uma correspondência que deva ser respondida, um formulário para atendimento ou um formulário de pedido de vendas, um documento será utilizado para iniciar (ou dar continuidade) a um fluxo de trabalho e, com frequência, deverá ser encaminhado para ação.

Para o gerenciamento de arquivos, o arquivamento de documentos e a retenção de registros, a identificação e a classificação precisa de documentos se tornam fundamentais. O local de um dado documento no sistema de gerenciamento de registros em uma organização é determinado pelo seu conteúdo. Na verdade, os documentos são frequentemente categorizados, arquivados e retidos em diversos repositórios, dependendo das características específicas das informações que eles contêm.

Embora a classificação de tipos específicos de documentos, como documentos de contabilidade, seja relativamente simples, em um ambiente com documentos diversos – como uma sala de correspondências – há possibilidade de que uma ampla variedade de itens recebidos não seja identificada de imediato. Em setores regulamentados de modo rigoroso, como a indústria farmacêutica ou petroquímica, talvez haja inúmeros documentos altamente técnicos ou altamente não estruturados que exijam especialistas treinados (e caros) para catalogar com precisão.

Mesmo em ambientes de processamento onde a triagem de documentos tenha sido realizada (por exemplo, o uso de endereços de caixas postais específicas), geralmente há uma variedade de tipos de documentos recebidos. Digitalizar esses documentos, que geralmente chegam em lotes de diversos tipos,

requer bastante intervenção manual. Ou cada documento é digitalizado de maneira individual especificando a cada vez o tipo de documento para o software, ou o papel é pré-classificado em lotes de documentos idênticos.

## Separação manual de documentos

Os processos de trabalho que implementam a separação e a identificação manuais de documentos estão repletos de ineficiências. Os gargalos atrasam o fluxo de trabalho e comprometem a eficiência das tecnologias de informação implementadas no sentido descendente do fluxo. Os scanners de documentos e os sistemas de compartilhamento de arquivos em rede permanecem ociosos enquanto aguardam que operadores manuais processem os documentos.

Em conjunto com esses problemas está o fato de que todos os processos manuais estão sujeitos a erros, o que em última instância gera a necessidade de outra separação e identificação, dessa vez correta. Essas “exceções” que devem ser encaminhadas outra vez ou duplicadas, porque simplesmente estão perdidas, são a parte mais cara de qualquer processo de manuseio de documentos.

Em resposta a esses desafios, as organizações implementaram uma série de técnicas, sendo que a maioria delas envolve um alto grau de mão-de-obra manual:

- **Limitar a digitalização para documentos de tamanho fixo com um número específico de páginas**

Isso requer a classificação de lotes em documentos idênticos de tamanho fixo; na verdade, algumas organizações optam por fotocopiar todos os documentos antes da digitalização de forma que os scanners sejam alimentados com páginas de tamanho idêntico.

- **Adicionar folhas de separação entre os documentos**

O software de captura automática pode identificar páginas em branco de maneira confiável, de forma que

os documentos podem ser divididos com a inserção de uma “folha de separação” entre documentos individuais ou lotes. Com vários documentos e diversos números de páginas, o uso de folhas de separação requer bastante trabalho manual e é caro, pois elas devem ser inseridas antes da digitalização e, se tiverem de ser reutilizadas, devem ser removidas posteriormente. Em operações de alto volume, esses custos podem ser assustadores. Por exemplo, um arquivo de empréstimo típico pode exigir até 50 folhas de separação. Multiplicados por centenas ou milhares – ou, em alguns casos, milhões de arquivos de empréstimos – esses custos se tornam proibitivos. Estima-se<sup>1</sup> que os custos de mão-de-obra sejam de 1,6 centavos de dólar para inserir uma folha de separação e mais 1,6 centavos de dólar para removê-la. Como os custos de impressão são de aproximadamente 1 centavo de dólar por página, é mais barato deixar as folhas de separação no lote e imprimir novas para outros lotes.

- **Uso de folhas de separação com código de barras**

Em algumas operações, as folhas de separação contêm códigos de barras para indicar o tipo de documento, permitindo que o software de captura separe e classifique documentos com mais facilidade. Isso proporciona identificação mais precisa dos documentos, mas também adiciona custos. Na verdade, o custo adicional apenas de papel e tinta para os códigos de barras pode ser significativo. E, mais uma vez, o acréscimo de um código de barras adiciona um período de preparação aos documentos quando as folhas são inseridas, já que é necessário tomar bastante cuidado para assegurar que a folha de separação correta seja utilizada. (Poucos operadores são capazes de distinguir um código de barras de outro.)

<sup>1</sup> White Paper – Auto Classification Saves Cost in High Volume Scanning Environments by Harvey Spencer Associates

Infelizmente, a separação manual de documentos, que requer um funcionário envolvido na preparação do documento para vislumbrar documentos com o intuito de determinar seu conteúdo, é cara, morosa e apresenta probabilidade de erros.

Para processadores de lotes de alto volume, a vantagem inicial da automatização da preparação de documentos é obtida com a economia de mão-de-obra da eliminação das folhas de pré-classificação e da separação de documentos. Outras vantagens resultam do processamento mais rápido, gerando maior produtividade e a capacidade de derivar mais vantagens com base nos aplicativos instalados.

## Automatização do processo de separação de documentos: difícil, não impossível

Em resposta às limitações e despesas adicionais da separação manual de documentos, as empresas estão ansiosas por aplicativos de captura que possam automatizar o processo de identificação e de separação de documentos. Isso permitiria que as organizações utilizassem scanners de alta velocidade e outros dispositivos de captura nas velocidades para as quais foram projetados. Ao eliminar a necessidade de separar documentos antes da digitalização, não só a mão-de-obra é reduzida, mas os processos de trabalho são dinamizados. Infelizmente, para muitos desenvolvedores de software, isso é extremamente difícil de ser realizado. De forma mais desafiadora, os sistemas para separação automática de documentos primeiro devem identificar e classificar páginas individuais de forma bem-sucedida.

Há dois métodos básicos de classificação automatizada de documentos: com base em imagem e com base em texto. Ambos utilizam softwares de reconhecimento para analisar documentos e layouts de páginas com o objetivo de identificar padrões recorrentes para categorizar documentos.

## Classificação de documentos com base em imagem

A classificação com base em imagem depende de documentos que tenham uma aparência semelhante em contrapartida ao seu conteúdo. Nos casos em que há pouco ou nenhum texto legível, essa metodologia pode ser bastante eficaz. A classificação com base em imagem pode ser muito rápida e, com algoritmos modernos, sua velocidade não é afetada por um número grande de tipos de documentos.

- **Classificação de imagens aprendida pelo exemplo**

A classificação com base em imagem “aprende pelo exemplo” ao comparar a aparência das páginas dos documentos. Esse método é mais adequado para documentos estruturados, como formulários. O software funciona dividindo a imagem de uma página em uma matriz de quadrados e, em seguida, calculando a probabilidade de encontrar pixels dentro de cada quadrado. Esses sistemas estão se tornando cada vez mais sofisticados, analisando não apenas os segmentos individuais da matriz, mas a imagem como um todo.

Para funcionar com eficácia, esse método requer documentos recorrentes para “treinar” o software. Quanto mais exemplos o sistema encontrar, mais preciso se tornará.

Anteriormente, o poder computacional extra exigido para a análise de páginas completas exigia muito dos processadores. No entanto, com a capacidade computacional aumentada dos sistemas modernos, há pouca vantagem em termos de velocidade com o uso de uma abordagem somente por região para a classificação das imagens.

- **Classificação de imagens com base em regras**

Com a classificação de imagens que aprende pelo exemplo, deve haver um número mínimo de

documentos semelhantes para fornecer uma amostra representativa para treinar o sistema. Além disso, as diferenças nas fontes de captura de entrada (por exemplo, diferentes scanners ou aparelhos de fax) podem causar impacto na eficácia desse método.

A classificação de imagens com base em regras depende da determinação de características definidas para pesquisa pelo software. Esse método também é adequado para localizar códigos de barras ou logotipos de empresas, que sempre estão presentes em documentos específicos. Isso pode funcionar como uma maneira excelente para filtrar previamente tipos de documentos padrão, antes da implementação de métodos mais sofisticados.

As desvantagens da classificação de imagens com base em regras incluem a necessidade de selecionar manualmente os recursos definidos, o que pode consumir tempo durante o processamento de grandes conjuntos de documentos. Outra preocupação é o fato de que a velocidade de processamento poderá ser afetada se um grande número de tipos de documentos for definido desse modo. Entretanto, o processamento com base em imagem é mais rápido do que o baseado em texto, pois o último exige OCR de páginas completas.

### Classificação de documentos com base em texto

A classificação de documentos com base em texto pode partir do conteúdo de uma região definida ou do documento inteiro. Além disso, essa técnica de identificação pode ser multifacetada; por exemplo, uma vez classificada como solicitação de seguro, outra classificação poderá se basear no nome de um cliente ou número de conta. A classificação com base em texto abrange de métodos simples que buscam palavras ou frases específicas no documento a sistemas de auto-aprendizagem mais complexos que são aprimorados ao longo do tempo.

Nos sistemas mais rudimentares, as palavras ou frases-chave devem ser especificadas e ajustadas manualmente, já que novas palavras-chave se tornam necessárias. Em alguns sistemas, as palavras ou frases podem receber pesos para aprimorar a precisão da identificação. Somente um número limitado de palavras, frases e pesos podem ser utilizados, uma vez que todos devem ser especificados manualmente, de forma que a precisão desse método é bastante limitada. A totalidade de todas as palavras, frases e pesos de um determinado sistema é denominada base de conhecimento. É difícil manter a base de conhecimento manualmente, pois a adição de uma nova palavra-chave pode violar as classificações existentes.

Sistemas mais sofisticados utilizam técnicas de auto-aprendizagem para identificar automaticamente as palavras e frases, e seus pesos. Uma dessas técnicas avançadas é denominada Support Vector Machine (SVM). As SVMs podem compor bases de conhecimento de dezenas de milhares de palavras automaticamente por meio da auto-aprendizagem, o que é muito maior e, portanto, muito mais preciso do que pode ser feito pelos meios manuais. Além disso, como essa máquina aprende automaticamente, a precisão é aperfeiçoada ao longo do tempo e a manutenção é simples. Um método de auto-aprendizagem é fornecer à SVM uma hierarquia existente de pastas contendo documentos que já estão classificados, denominada conjunto de aprendizagem. A SVM trabalha pelas pastas e analisa todos os documentos para criar sua base de conhecimento.

O conjunto de aprendizagem contém documentos que já foram classificados manualmente. Infelizmente, alguns dos documentos no conjunto de aprendizagem não serão classificados com precisão porque o processo de classificação foi manual e todos os

processos manuais apresentam alguns erros. A tecnologia mais recente na classificação de documentos utiliza uma técnica denominada Maximum Entropy Discrimination (MED), que gera sua base de conhecimento da mesma forma, porém é mais receptiva a documentos classificados incorretamente no conjunto de aprendizagem.

Diante de uma ampla variedade de tipos de documentos, a classificação de texto de auto-aprendizagem pode ser bastante eficaz. O período de configuração é menor, uma vez que não há necessidade de personalizar regras.

A classificação do texto será mais eficaz quando houver textos amplos para serem analisados. Portanto, a técnica não é adequada para documentos que contenham pequena quantidade de texto. Além disso, para obter bons resultados, um exemplo representativo (em geral, pelo menos 20) de cada tipo de documento deve estar presente para treinar o software adequadamente. Como a classificação de textos de auto-aprendizagem depende de OCR de páginas completas para cada imagem, essa técnica não é tão rápida quanto aquelas que se baseiam em imagem. Atualmente, pode demorar alguns segundos por página. Porém, à medida que as velocidades de processamento forem aprimoradas, isso se tornará menos problemático.

## Manuseio de documentos com várias páginas

Um dos maiores impedimentos para a automação da captura de documentos é o fato de que as páginas devem ser digitalizadas uma por vez, embora os documentos possam ser compostos por várias páginas. É especialmente difícil para o software compreender com precisão onde um documento termina e o próximo começa.

À medida que os aplicativos de captura se tornam mais sofisticados, há progressos recentes que

visam solucionar esse paradoxo. Esquemas simples reconhecem a primeira página do documento e o software pode ser treinado por meio de exemplos (ou seja, mostrando-se ao software de captura a página frontal de cada tipo de documento). Sempre que o sistema reconhecer uma das páginas frontais de qualquer um dos documentos, irá considerar que um novo documento foi iniciado. Esse processo classifica e separa em uma única operação.

Esse método de separação inteligente de documentos depende muito da precisão da classificação das páginas. No caso de uma das páginas frontais não ser reconhecida corretamente, o problema será duplicado; o segundo documento será vinculado ao primeiro como um anexo e será efetivamente eliminado do fluxo de trabalho.

Para assegurar que as páginas do documento tenham sido analisadas com precisão, as mesmas devem ser inspecionadas manualmente. A sofisticação da plataforma de captura de documentos determina a facilidade com a qual as correções podem ser feitas. Trabalhando com um sistema de captura modular, um operador conserva a capacidade de separar manualmente páginas dos documentos conforme necessário. No entanto, nem todos os aplicativos de captura oferecem esse recurso, sendo que nesse caso o documento deve ser liberado para o back-end e corrigido lá – um processo mais caro, com probabilidade de erros e que consome tempo.

Outras limitações da superfície básica da identificação de documentos em documentos altamente não estruturados, onde não há evidências gráficas ou contexto suficiente para identificar a primeira página de um documento. Além disso, em casos nos quais páginas misturadas tiverem sido digitalizadas fora de ordem – uma situação frequente em ambientes de alto volume – os documentos aparentemente terão páginas faltantes ou páginas de outros documentos inseridas.

## Classificação global

Mais recentemente, sistemas altamente sofisticados foram desenvolvidos de forma que estão empregando várias técnicas de reconhecimento para identificar não apenas a primeira página em um documento com várias páginas, mas também a última página e aquelas do meio. Além das evidências fornecidas por páginas individuais, esses sistemas analisam constantemente o encadeamento contíguo de páginas e como elas se relacionam entre si. O software pode analisar a semelhança de cada página que pertence ao início, meio ou fim de um documento. Mediante cada determinação, o posicionamento mais provável é feito – conforme mensurado em relação a outras possibilidades.

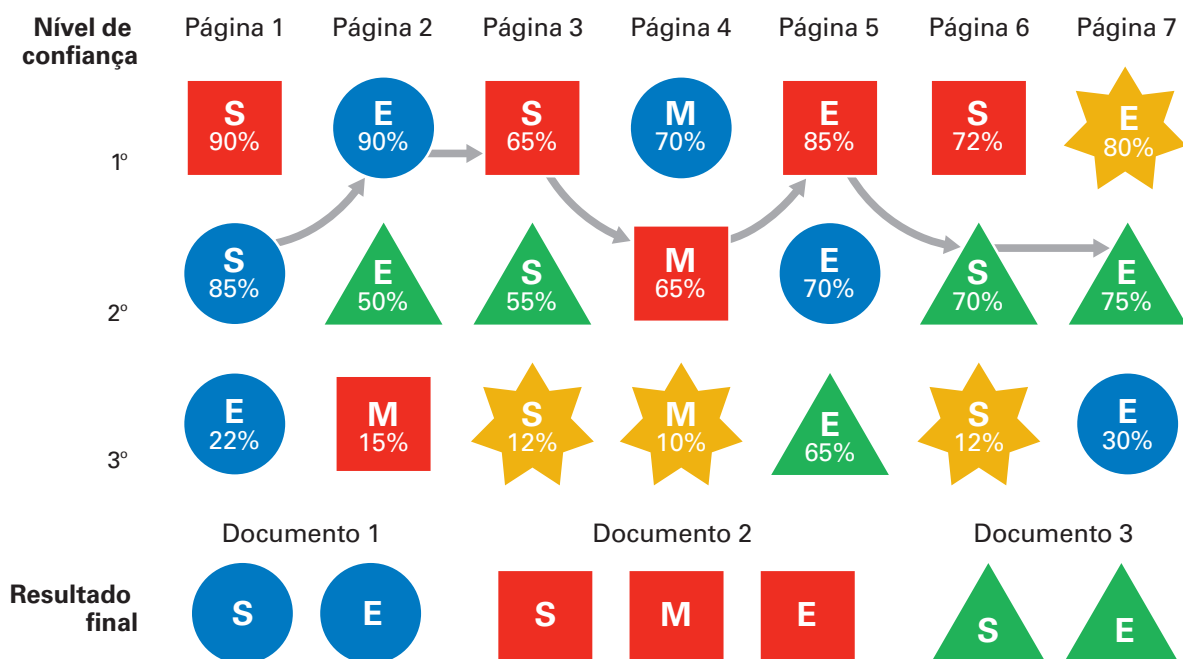
Essas tecnologias de reconhecimento robustas podem identificar limiares entre vários documentos semi-estruturados e não estruturados em um único lote, substituindo o processo tradicional de inserir páginas de separação e aprimorando bastante a precisão dos esquemas rudimentares de classificação da primeira página.

Como exemplo, considere que há três tipos de documentos possíveis representados pelas páginas com formato de círculo azul, quadrado vermelho, triângulo

verde e estrela amarela no diagrama abaixo. Cada página é digitalizada e classificada, e um nível de confiança é associado a cada tipo possível de página. No diagrama abaixo, o encadeamento de imagens digitalizadas é mostrado no diagrama e três classificações possíveis são exibidas para cada página, com a mais confiável no topo e a menos provável na parte inferior. A porcentagem representa o nível de confiança da classificação. O S representa uma página inicial, o M uma página do meio e o E uma página final. Há uma série de regras de negócios nesse exemplo simples:

- Todos os documentos começam com uma página S
- Todos os documentos terminam com uma página E
- Se um documento tiver mais de duas páginas, todas as páginas entre as páginas S e E serão páginas M
- Uma página E de um documento é seguida por uma página S de outro documento

No exemplo, as regras de negócios, além da exibição global do fluxo de páginas, significam que as páginas digitalizadas são classificadas como documentos do tipo círculo azul, quadrado vermelho e triângulo verde. (As páginas possuem formas diferentes – quadrados, círculos, triângulos e estrelas – para que seja possível diferenciá-las quando forem impressas em preto



e branco.) Se apenas as primeiras páginas mais confiáveis tivessem sido utilizadas para classificar os documentos, a escolha teria sido incorretamente três documentos vermelhos.

Um modelo de Classificação global proporciona uma série de vantagens em relação a outros métodos. Como um sistema de Classificação global observa cada página no contexto daquelas que estão ao seu redor, mesmo nos casos em que uma página tiver baixa probabilidade de reconhecimento, o documento ainda poderá ser classificado e separado de forma correta. O sistema também poderá reordenar as páginas que foram alimentadas incorretamente ou misturadas durante o manuseio dos documentos.

Os documentos podem ser encaminhados para o processo ou fluxo de trabalho correto da transação pretendida sem classificação prévia. Isso reduz bastante os custos de manuseio de documentos associados ao processamento manual e aumenta a produtividade dos processos de negócios.

## Na prática hoje

Atualmente, as soluções da Kofax estão ajudando empresas de muitos setores diferentes a processar documentos com maior rapidez e eficiência. Observe a seguir uma pequena amostra das soluções e suas vantagens imediatas.

### Processamento de empréstimos e hipotecas

O processamento de hipotecas é um dos setores que mais envolve papel e, ao longo dos últimos anos, a Kofax implementou soluções para mais da metade das 25 principais empresas de hipoteca. Um dos clientes está obtendo economia que atinge US\$ 35 mil por dia (cerca de US\$ 700 mil por mês) graças a uma redução na mão-de-obra manual. Outro cliente na área de hipotecas reduziu os custos do processamento de cada empréstimo em US\$ 15. (Com períodos de

pico de processamento chegando a cerca de 30 mil empréstimos por mês, a economia ultrapassa meio milhão de dólares mensalmente.) Essa economia se reflete diretamente no resultado.

Com a tecnologia integrada automática de separação de documentos, bancos e outros processadores de empréstimos podem aumentar a eficiência e reduzir os custos relacionados ao processamento de centenas de milhares de hipotecas e documentos de empréstimo que recebem de escritórios de empréstimo e clientes.

Embora essas vantagens se estendam a vários setores, a Harvey Spencer Associates cita os exemplos de processamento de empréstimos hipotecários, que contêm uma combinação de documentos, alguns deles com uma única página e outros com várias páginas. “Os usuários podem confirmar automaticamente se todos os documentos exigidos estão na pilha e na ordem correta, inclusive corrigindo-os eletronicamente quando for necessário,” relatam os analistas. “Quando alguém compra uma pilha de hipotecas, os documentos de suporte devem ser transferidos. No mercado de hipotecas secundárias, estamos falando de inúmeras páginas, já que é provável que o comprador esteja adquirindo milhares de hipotecas como um pacote.”

As ferramentas de automação de documentos permitem uma análise muito mais profunda dessas grandes quantidades de documentos. Por exemplo, proporcionar a capacidade de extrair informações eletronicamente que ofereçam um melhor valor para o hipotecado – tais como evidências da probabilidade de ser pago novamente.

Uma empresa de processamento de hipotecas residenciais que lida com 12 milhões de imagens por mês implementou as soluções da Kofax. Com a pasta de cada cliente contendo cerca de 100 páginas e até 80 tipos de documentos diferentes, esse processador era um candidato ideal para a separação automatizada

de documentos. Antes da automação, o volume de trabalho exigia uma equipe de 60 pessoas para executar a preparação e a separação dos documentos. Uma equipe extra de 16 pessoas era necessária para o controle de qualidade.

Graças à separação automatizada de documentos, o processador foi capaz de reduzir significativamente o processamento manual. Agora, a mesma carga pode ser manuseada com uma equipe de dez pessoas para fazer a separação e a preparação de documentos e três pessoas para revisar.

Spencer acredita que essas vantagens se estendam a todos os setores. “Inicialmente, essas implementações estão sendo utilizadas para coisas como processamento de faturas e automação da sala de correspondências”, afirma ele. “Mas também há a oportunidade no segmento ad-hoc – nesse caso, o usuário não deseja perder tempo selecionando quais botões irá pressionar – se o sistema tomar as principais decisões, será possível reduzir significativamente o período que o usuário ad-hoc caro gasta no dispositivo de digitalização.”

## Governo

Outro setor que consome muito papel é o governo, que possui uma variedade enorme de tipos e processos de documentos. O Utah Department of Human Services é uma organização governamental desse tipo e se dedica a oferecer serviços de apoio infantil e para crianças carentes. A entidade desejava capturar os documentos em papel de mais de 90 mil casos de apoio infantil que estavam armazenados em armários de arquivos em todo o estado de Utah. Isso representava de seis a sete milhões de páginas de documentos em papel. Utilizando a tecnologia avançada de classificação e separação da Kofax, o órgão foi capaz de concluir essa tarefa em oito meses, ao invés dos três a cinco anos esperados, o que representou uma economia de tempo de mais de 80%.

Agora que os arquivos estão armazenados eletronicamente, vários funcionários podem trabalhar no mesmo caso ao mesmo tempo, ajudando a aumentar a produtividade e a atender melhor os clientes.

Agora o ORS pode assegurar que todos os arquivos de casos estão em conformidade total com as regulamentações do setor e que todos os formulários foram capturados com precisão. Em arquivos de papel, os arquivos de casos frequentemente se tornam uma área de despejo para todos os tipos de conteúdo desnecessário. A solução da Kofax realmente automatiza a classificação e a identificação de documentos elegíveis dos arquivos de casos, e os documentos desnecessários e ineleáveis podem ser facilmente eliminados do armazenamento no arquivo digitalizado. Além de manter todos os arquivos de casos limpos, esse processo também gera eficiências que permitem aos funcionários obter as informações adequadas em tempo hábil.

## Outros casos

Os casos a seguir demonstram os níveis de economia que uma organização que esteja implementando a tecnologia eficiente de classificação, extração e separação de documentos poderá atingir. Estes são os resultados reais dos clientes da Kofax:

- A Pathology Network reduziu significativamente os custos relacionados à impressão de folhas de separação, como papel, tinta e toner ao utilizar a tecnologia de separação automática de documentos para os formulários de solicitação de amostras de patologia recebidos. Sua utilização também diminuiu bastante os erros humanos relacionados à separação manual de documentos, já que o sistema extrai automaticamente as informações escritas à mão e impressas do formulário de solicitação dos pacientes para a obtenção de um grau elevado de precisão.

- Uma empresa de processamento de empréstimos foi capaz de eliminar praticamente toda a preparação manual de documentos de seu processo, resultando em uma redução de 145 para 15 funcionários em período integral. Apesar da redução, a produtividade aumentou em quase 90%. O período de processamento diminuiu de seis minutos por documento para menos de um minuto. Em última análise, a empresa pode dobrar o número de empréstimos processados diariamente.
- Uma empresa de serviços de informação processa cerca de 400 mil documentos de recursos humanos mensalmente. O bureau de serviços buscou uma solução que pudesse reconhecer áreas de páginas formatadas para extrair dados específicos para preencher campos de índice. A tecnologia da Kofax não só executou o reconhecimento da área requisitada, mas também proporciona a vantagem de ser capaz de reconhecer e separar automaticamente diferentes tipos de formulários.
- Outro cliente estimou que seriam necessários 50 funcionários em tempo integral para desempenhar o mesmo trabalho manualmente. Ao implementar a classificação e a separação automatizadas, a empresa pode dobrar a capacidade utilizando somente dez pessoas.
- Uma empresa de processamento de empréstimos estima uma economia anual de US\$ 420 mil em mão-de-obra reduzida devido à diminuição no processamento manual de documentos, com economia anual líquida de US\$ 100 mil do processador de documentos somente em consumíveis (tinta e papel) ao eliminar folhas de separação; um aumento na precisão de 98% utilizando folhas de separação para 99,5% com o software.
- Um processador de hipotecas registrava um período de recuperação de documentos de oito minutos. A

empresa calculou que obteria retorno do investimento total se pudesse diminuir o período de recuperação para dois minutos. Agora, o tempo atual de recuperação de documentos é inferior a um minuto como resultado da categorização e da recuperação de documentos mais rápidas proporcionadas pela automação.

A automação de documentos percorreu um longo caminho desde que a máquina de escrever foi intencionalmente projetada para tornar a produtividade mais lenta. Ao mesmo tempo em que parece não haver limites para a capacidade de uma organização em gerar documentos em papel, o processamento desse papel tem sido aprimorado de forma notável. Com o advento da tecnologia intuitiva de identificação e separação de documentos, o que anteriormente compunha uma das etapas mais caras e mais demoradas do processamento de documentos agora pode ser automatizado de modo confiável.

## Sobre a Kofax

A Kofax plc (LSE: KFX) é a principal fornecedora de soluções de automação de processos comerciais baseados em documentos. Há mais de 20 anos, a Kofax fornece soluções premiadas, que facilitam o fluxo de informações dentro de uma organização, por meio do gerenciamento da captura, transformação e troca de informações empresariais críticas em papel, fax e formatos eletrônicos de uma maneira mais precisa, ágil e econômica. Essas soluções proporcionam um rápido retorno do investimento a milhares de clientes dos setores de serviços financeiros, administração pública, terceirização de processos de negócios, saúde, cadeia de suprimentos e outros.

A Kofax fornece essas soluções por meio de suas próprias organizações de vendas e serviços e por meio de uma rede global de mais de mil parceiros autorizados em mais de 60 países nos cinco continentes.

Para mais informações, visite [www.kofax.com](http://www.kofax.com).