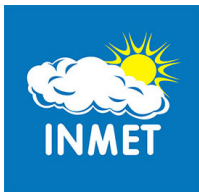




O ABBYY FlexiCapture® Engine ajuda a ler uma carga de 100 anos de dados climáticos

Customer Overview



Nome

Instituto Nacional de Meteorologia (INMET)

Sede

Brasília, Brasil

Indústria

Governo

Web

inmet.gov.br

Projeto

Extrair dados de documentos digitalizados que diferem em termos de estrutura, complexidade e condições físicas para um banco de dados históricos de meteorologia

Solução

ABBYY FlexiCapture Engine

Resultado

- 3 milhões de páginas processadas
- 4 bilhões de caracteres reconhecidos
- Maior velocidade e qualidade dos dados reconhecidos

O Instituto Nacional de Meteorologia (INMET) é uma agência responsável pelo monitoramento e previsão das condições climáticas no Brasil. Foi criado em 1909 e, desde então, as suas inúmeras estações de monitoramento espalhadas por todo o país, registram continuamente os fenômenos meteorológicos e climáticos.

Projeto

Não é preciso dizer que o INMET agora possui uma coleção impressionante de registros meteorológicos que datam do final do século XIX. Esses dados podem ser usados para determinar padrões climáticos a longo prazo, analisar mudanças nas condições climáticas e produzir previsões mais precisas. É inegável a sua importância para uso científico e prático, em particular nas atividades do Ministério da Agricultura, Pecuária e Abastecimento.

Infelizmente, até recentemente o acesso ao banco de dados meteorológicos era complicado, se não inviável, pelo fato de ele existir apenas no papel; era realmente difícil classificar e procurar partes específicas de informações sem benefícios das tecnologias modernas.

Solução

Em 2012, o INMET realizou um processo de licitação para o projeto de transferir as informações em papel (livros digitados em formato A3 e folhetos manuscritos em vários formatos) para um banco de dados eletrônico, a fim de garantir a continuidade das informações e permitir fácil tabulação e análise dos dados meteorológicos. O contrato foi assinado com a Flexdoc Tecnologia da Informação Ltda, uma empresa especializada em fornecimento de soluções completas end-to-end na área de automação de processos de fluxos de trabalho. Dispondo de uma área de 1500m² para o armazenamento de documentos, ela implementou projetos com um volume superior a 30 milhões de documentos tratados. A tarefa definida pelo INMET foi um desafio, devido à sua escala (mais de 3 milhões de páginas a serem processadas) e da variedade de documentos em papel que diferiam em termos de estrutura, complexidade e até mesmo condições físicas. No início, a Flexdoc testou uma série de soluções e finalmente optou pelo ABBYY FlexiCapture Engine, por graças às referências de sucesso na indústria, escalabilidade e flexibilidade, o que permite o processamento de vários tipos de documentos. Reconhecendo a necessidade de incorporar o OCR no processo, a Flexdoc testou uma série de soluções, antes de escolher o ABBYY FlexiCapture Engine. Foi uma combinação perfeita entre precisão excepcional de dados e avançada escalabilidade. Mas a característica decisiva foi a sua flexibilidade, pois o

ABBY FlexiCapture Engine fornece uma série de ferramentas e utilidades que permitem um fácil “zoneamento” de formulários e capacidade de definir e modificar os modelos. Desta forma, toda a gama de documentos produzidos ao longo de décadas pôde ser processada por um único sistema.

Em primeiro lugar, os documentos digitalizados são importados e imediatamente enviados para o ABBYY FlexiCapture Engine para o reconhecimento dos tipos de formulários e correspondência de templates, a fim de localizar cada campo.

RES. DO DIA	LEITURA	COR	NÁXIMA	LEITURA	COR	NÁXIMA	LEITURA	COR	MÍNIMA	LEITURA	COR	MÍNIMA	ONTEM	ATUAL	ONTEM	ATUAL	TOTAL										
12	228	101	148	010	-416	16	152	-223	101	25	1280	100	220	1234	100	234	046	0618	010	0618	041	0218	010	800	810	010	
18	219	110	158	010	-416	16	110	-223	101	25	1288	100	218	1236	100	236	052	0615	010	0615	019	0315	019	019	010	010	010
24	216	110	185	010	-417	16	138	-223	101	25	1253	100	213	1230	100	230	023	0812	019	0812	019	0415	012	010	010	010	010

5 - OBSERVAÇÃO AGROCLIMATOLÓGICA

TEMPERATURA DO SOLO °C - PROFUNDIDADE EM cm

12 HORAS UTC				18 HORAS UTC				24 HORAS UTC			
1	2	3	4	1	2	3	4	1	2	3	4
218	218	214	112	302	302	312	312	320	340	320	324

TANQUE DE EVAPORAÇÃO 12 HORAS UTC

TEMP. MÍN. RELVA. °C	RAD. GLOBAL CAL/cm²	ONTEM mm	ATUAL mm	ENCHIMENTO mm	RETRADA mm	TEMP. MÁXIMA °C	TEMP. MÍNIMA °C	VENTO UNID.
		30	43	6	04			

TANQUE DE EVAPORAÇÃO 18 HORAS UTC

TEMP. MÍN. RELVA. °C	RAD. GLOBAL CAL/cm²	ONTEM mm	ATUAL mm	ENCHIMENTO mm	RETRADA mm	TEMP. MÁXIMA °C	TEMP. MÍNIMA °C	VENTO UNID.

7-APARELHOS REGISTRADORES

APARELHOS	12 UTC	18 UTC	24 UTC
BAROGRAFO - mb (Hpa)	1017.0	1001.3	1010.7
HIGROGRAFO - %	68	55	82
TERMÓGRAFO - °C	29.2	29.6	25.4
PLUVIÓGRAFO - mm	0.0	0.0	0.0
PIRANÓGRAFO - Cal/cm²			
ANEMÓGRAFO - km/10min	1.7	2.1	0.7

A automatização deste passo acelera muito o processamento do documento, como há mais de 20 tipos de brochuras, cada uma composta por pelo menos 6 tipos de páginas. Algumas páginas contêm mais de 150 campos. Os dados impressos são extraídos utilizando as tecnologias de OCR da ABBYY. Quanto aos campos com dados escritos à mão, devido à má condição física (alguns dos documentos datam de 1900) e por serem ilegíveis para uma máquina, os mesmos são enviados para operadores que digitaram os dados manualmente. Ao detectar os campos e os seus tipos, o ABBYY FlexiCapture Engine acelera e facilita muito o trabalho dos operadores que se especializaram em determinados campos.

Depois disso, todos os campos vão à verificação. 100% de precisão é uma necessidade, porque os dados serão posteriormente utilizados para cálculos científicos, pesquisas e previsões. Em caso de divergências ou dúvidas, os campos com resultados incomparáveis são enviados para os supervisores para uma análise mais abrangente. A validação de situações conflituosas e o controle de qualidade antes da exportação ainda requerem a participação de profissionais capacitados e especializados em meteorologia.

Finalmente, os metadados são exportados para supercomputadores de previsão climática do INMET. Um total de 85 pessoas estão envolvidas em todo o fluxo operacional: desde a importação e extração dos dados até a verificação e tratamento dos erros.

“Precisávamos de uma solução flexível por causa da variedade de documentos digitalizados, dispersos em vários padrões e formatos. O uso do ABBYY FlexiCapture para reconhecer as coordenadas dos campos e localização exata nos trouxe um enorme ganho de desempenho no tratamento de folhetos.”

Carlos Flávio Barreto F. de Souza, Diretor de Tecnologia, Flexdoc Tecnologia

Acerca da ABBYY

ABBY é uma empresa líder mundial no fornecimento de tecnologias e soluções que ajudam às empresas acionar as informações de maneira efetiva.

ABBY 3A

Estados Bálticos, Ásia, Oriente Médio, América do Sul, África
 P.O. Box #32, Moscou, 127273, Rússia
 Telefone: +7 (495) 783 3700
 Fax: +7 (495) 783 2663
sales_3a@abby.com

Resultado

A incorporação do ABBYY FlexiCapture Engine no fluxo de trabalho trouxe um aumento significativo da produtividade. Graças à solução de automação da ABBYY, a Flexdoc reduziu o número de digitadores e especialistas envolvidos no projeto em 30%. Os documentos agora são analisados automaticamente, o que torna as tarefas dos digitadores tão fáceis que poderiam ser feitas por uma criança. O trabalho de profissionais qualificados é reduzido à mera verificação, o que lhes permitiu dedicar mais tempo a atividades criativas e valiosas. O projeto envolve processamento de mais de 3 milhões de páginas e 4 bilhões de caracteres foram processados. Graças às tecnologias da ABBYY, a digitalização de toda a informação de tempo e clima do Brasil, anteriormente considerada um empreendimento complicado a longo prazo, agora será concluída até o ano de 2017.

