OpenText digital mailroom solutions and technologies

**Distributed and multichannel capture**
- OpenText Capture Center
- OpenText Captiva Cloud Capture Toolkit
- OpenText Captiva Mobile SDK
- OpenText TeleForm & Liquid Office
- OpenText LEAP Snap

**Automated document and data recognition**
- OpenText Captiva
- OpenText Capture Center
- OpenText Capture Document Recognition
- OpenText Capture Recognition Engine
- OpenText TeleForm
- OpenText LEAP Snap

**Enterprise scalability**
- OpenText Captiva
- OpenText Captiva Real-Time (REST APIs for capture services)

**Embed within business process**
- OpenText Captiva
- OpenText Captiva Real-Time (REST APIs for capture services)
- OpenText LEAP Capture Services
Enterprise capture figures

- **30 billion images** captured per year
  Covering a range of industries and geographical locations. OpenText is a market leader.

- **13 of top 20 banks** run OpenText capture
  Seen as the capture platform of choice for many FSI customers. OpenText is critical to many processes for the top banks.

- **7 million images** per day by one customer
  Our largest customer consistently captures over 7 million pages per day at a single site.

- **1,000,000s of mobile devices**
  First released in 2014, and now on version 2.6, Captiva’s mobile SDK is embedded within a number of consumer-facing mobile applications.
What do people look for?

Omnichannel
- Centralized capture
- Remote branch capture
- Mobile devices
- White mail
- Electronic documents
- Web portals
- Web services
- Email
- File drop
- Fax
- Chatbots
- Twitter

Turn unstructured into structured
- Invoice details
- Cheques
- Trade confirmations
- Sanction compliance
- Names
- Addresses
- Account information
- Tax details
- Medical records
- GDPR compliance
- PPI

Machine automation
- Fraud detection
- Credit checks
- Invoice payments
- Mortgage approvals
- Insurance claims
- Liability
- Money deposits
- Consumer loans
- Expense payments
- Account amendments
- Tax returns
Capture Processes

**Capture**
Increasing channels and formats of captured documents – system must handle new channels consistently & accurately

**Sorting & extraction**
Automated document recognition technologies accelerate dependent processes, reduce processing costs, and increase accuracy

**Process integration**
Deliver document data (and often documents) to business process engine (ECM, BPM, or business system)

**Capture throughout process**
Documents are not always received at process initiation. Integration at various points within process improves process accuracy, flexibility and collaboration
It’s about the “3A’s”

Automation, APIs & AI
Automation

• Automate business processes for faster response time
• Humans manage exceptions, robots do the rest
• Integrate with 3rd party systems
• Automate complex business rules for data integrity
• Can model very complex processes
APIs provide ubiquitous access to capture functionality

- Expose mailroom knowledge base to new channels
- Discrete REST services for specific tasks
- OpenText has a broad range of APIs, including on-device and scanner controls, which are also OEM’d by our competitors
Customer Onboarding Automation

• Mobile SDK & Captiva Real-time services
• SDK provides feedback on the device
• Captiva Real-time provides intelligent document recognition techniques
• Embeds into customer’s existing mobile application
Automation of Taxation, Fees & Contribution Payments

• Captiva Real-time services
  • When capture functionality is needed outside a mailroom (batch) system
  • Accelerates transactions by literally an order of magnitude
    • Seconds, instead of minutes or hours

• Real-time automated processing of documents from:
  • ATMs
  • Mobiles
  • Web portals

• Multiple variations of documents
Artificial Intelligence (AI)

• Techniques to improve accuracy and speed by learning document layouts

• Neural networks to improve automation from documents that are difficult to process by other means

• Add Magellan for advanced analytics and Natural Language Processing (NLP)
Using Content Analytics for Data Extraction

• Can help us improve recognition of Proper Nouns
• Reduce TCO of Capture solutions
• Automated & Assisted learning
• Normalisation of Data
  • Town > Borough > County > Country > Continent

Structure to the Unstructured

<topics/>
<concepts/>
<people/>
<organizations/>
<locations/>
<sentiment/>

Content Analytics uses semantic technologies to unlock business value from unstructured content and applying inferred metadata on the content.
Leveraging AI to enhance document capture

AI-Augmented Capture

- Process documents during capture
- Use Advanced Recognition for structured and semi-structured documents
- Use AI for unstructured documents and named entity extraction
- Automate routing and processing
Captiva + Real-Time + Magellan + xECM
Captiva Updates

Captiva Mobile SDK 2.6
Captiva Mobile SDK Overview

- Enables mobile applications to take optimal pictures for OCR
  - Photos can be a little blurry; document images must not be
  - SDK provides feedback to user on the device
  - Software waits to take pictures until the device is stable and focus is locked
  - Auto-reject low-quality images

- Self-contained APIs
  - Customers need not know anything about document capture
  - Extensible UI for visual alignment with the application
  - Also allows automatic multi-page capture using video feed from mobile camera

- Supports Captiva Real-Time Intelligent Document Recognition
  - Classification and data extraction in real-time while your customer is holding their phone
Captiva Mobile SDK

APIs for mobile image capture
- Light sensor
- Autofocus
- Motion sensor
- Torch control
- Image stabilization time control
- Document size
- Automatic multi-page capture
  - *Preview image quality scoring*
- White balance
- Perspective
- Glare
- Image guides
- Visual quality indicators
- Labels
- Replaceable UI
- *Resizable preview pane*

APIs for on-device image enhancement
- On-device barcode recognition
- Orientation correction
- Distortion correction
- Gray-scale & adaptive binary conversion
- Automatic cropping
- Resizing
- File format conversion
Captiva Updates

Captiva 7.7
Captiva Web Client

- Knowledge-worker capture
- Real-time automated classification & extraction
- Common UI for:
  - Scanning
  - Indexing/data validation
  - Rescanning
- Called mid-process
  - Indexing/validation
  - Rescanning
- Same UI used for LEAP Snap product

Updated
Captiva Web Client
Advanced PDF Processing

• Our customer are seeing large increases in born-digital documents
• Direct data extraction for text-based PDF documents
  • Improved accuracy
  • Higher throughput
• Special handling for PDFs with both text and images
  • Use text only
  • Use image only
  • Voting algorithm between text and image
Capture Center Update

OCC 16.4
Capture Center: Combining any input and any output

Enterprise Scan Client

Email

Fax

FTP site

Network Folder

SharePoint Content Server

Custom

Internet Information Server

OCC Web server

Monitoring

Configuration

Dispatch

Recognition

Validation

Open Text Capture Center Server

OTDS

SAP, Oracle

Content Suite, Process Suite

IDM, VRD, TCP, …

SharePoint

File System

Custom

OpenText Confidential. ©2018 All Rights Reserved.
New in OCC 16.4

• Highlights
  • Support for tabular data, lookup lists and scripting in web validation
  • Improved adaptive recognition
  • Mobile scan client

• Other function
  • Process Suite export now supports new Content Building Block
  • REST API to load documents into OCC
  • Improved Real-time response
  • More cluster node management tools
  • Logging and history view for changes in user / task management
  • Function to process only a subset of a PDF
Web validation: New features

Tabular data

Lookup list
Thank you
Leftover slides
LEAP Snap & LEAP Capture Services

SaaS Capture
LEAP overview
SNAP

A simple, powerful app to quickly convert documents into actionable data to automate your business. Capture content on the go with a mobile app, packing real-time advanced recognition in the palm of your hand.
Snap SaaS capture
What makes SNAP different?

Advanced recognition
• Intelligent auto-classification of any document, in real-time
• Automatic data extraction to quickly turn unstructured content into actionable information

Easy & intuitive
• Enterprise-grade advanced recognition with consumer-grade user experience
• Visual cues and a guided experience make it quick and easy to capture data and handle any exceptions

Simplified set-up
• Spin up new capture environments in minutes
• Design-as-a-service offering simplifies the development of new document templates
• Repository-agnostic with CMIS compatibility
LEAP Capture Services

A set of hosted JSON REST microservices allowing Developers to quickly enhance custom content-centric apps with advanced document & data recognition

Basic Imaging Services
- Validate data
- Image processing
- Barcode Detection
- Create Full Page OCR + Convert to PDF
- Export Services

Advanced Recognition Services
Customizable & Automatic
- Classify Document
- Extract metadata from documents
Capture Development Focus

Distributed Capture
Capture Anywhere
- Mobile
- Browser
- Cloud
- Real-Time

Advanced Recognition
Automatic and Effortless
- Accuracy
- Classification and Extraction Automation
- Machine Learning

Enterprise Platform
Trust and Scale
- Lower TCO
- Security
- Configuration & Deployment
Captiva Real-Time – how it works

Module Server is responsible for the lifecycle of a set of capture modules (starting, recycling, etc.) It is not part of the communications for task assignment.

The number of Module Servers per machine is configurable.

The number of each capture module per Module Server is configurable (global setting).

A 3rd party load balancer can be put in front of multiple REST Servers.

Real-Time Servers distribute work to the appropriate task using a round-robin approach.

Captiva 7.7 adds node lock/unlock to improve load balancing when 2 or more servers are running.
Captiva Real-Time – scaling up

Add new instances of Module Server to ensure there are always available instances for each capture module. No additional licensing required.

All servers and modules for a single RT system use a single file share for configuration and temporary data storage. Scaling is achieved by starting up another instance and pointing it to the file share.

Add new instances of REST server to handle several hundreds of simultaneous transactions. No additional licensing required.

Image Converter
Image Processor
Classifier
Extractor

Image Converter
Image Processor
Classifier
Extractor

Image Converter
Image Processor
Classifier
Extractor

Image Converter
Image Processor
Classifier
Extractor

Image Converter
Image Processor
Classifier
Extractor
Highly-available stand-alone Real-Time system

Line of Business Application

Mobile App

3rd party load balancer

REST Server web farm

Captiva Module Services
(as many instances as needed to meet peak throughput requirements)

One Real-Time Server license can do ALL this!!

2 stripped-down Captiva servers in a ScaleServer group
(included with Real-Time Server)

Configuration / profiles
User authentication
Real-Time Capture
Embedded into LOB applications

- Easily scales to support mobile capacity needs – just install additional module service instances
- Cloud & web-farm-friendly / designed for "cloud first"
- Call only services needed; no CaptureFlow required
- Integrated with Captiva Web Capture Client

Line of Business Application

Customer Web Server

Load Balancer

REST Server Web Farm

Captiva Module Services
- DocType Enumeration
- Image Conversion
- Image Processing
- Barcode Recognition
- OCR & PDF Creation
- Document Classification
- Data Extraction
- Validation

Captiva Module Services
SAMPLE WEB CLIENT IMPLEMENTATION
With Real-Time Advanced Recognition at Scan-time

User Scan pages

REST auto detects doc type & doc breaks

REST auto populates data entry form

User reviews data & submits
PLACE scan button on web form of Business App

Button invokes Web Client with config & metadata via URL

Captured docs placed in repository

User redirected back to business app

- Parameterized URL launching
- Pass SSO tickets
- Specify localization preference
- Specify Web Client Scan Profile to use
WEB CLIENT ARCHITECTURE (REST Real-time)

Client Tier
- Mobile App + SDK & Web Client

Web Tier
- Load Balancer
- REST Server
- REST Services
- DocType Enumeration
- Image Processing
- Barcode Recognition
- Document Classification
- Data Extraction
- Validation

App Tier
- REST API
- Repository
Key Benefits

- **Highly Scalable**
  - Only constrained by the load on the module servers
  - Easily directs new module servers to the shared configuration to add capacity

- **Highly Secure**
  - Multiple authentication options available
  - Network traffic is secure with HTTPS
  - Storage – secure the REST data directory
Summary

• Captiva Real-Time Capture Services
  – Standalone services for immediate & real-time capture processing
  – Designed for concurrent processing of large number of transactions
  – Robust under low-bandwidth conditions and on unstable networks
  – Simplifies integration of capture capabilities in other applications