

# IBM Datacap Taskmaster

Die Lösung für Scannen, automatisches Klassifizieren und intelligente Datenextraktion

Michael Vahland  
IT-Specialist ECM  
Software Group  
Michael.Vahland@de.ibm.com



# Agenda

- Einführung
- Datacap-Architektur
- Datacap-Oberflächen
- Stärken von Datacap

# Einführung - Betrachtung zum geschäftlichen Nutzen

## → Mehr Effizienz

- Unmittelbarer Zugriff auf erfasste Dokumente
- Aufbereitung von Metadaten während des Prozesses
- Erstellung von Statistikberichten (für dynamische Bedarfsvorhersagen)
- Einfacheres operatives Management – umfasst Tools für Implementierung und Anwendungsverwaltung

# Einführung - Leistungsfähigkeit

## Nutzenprogression

**Geschäftlicher Nutzen/ROI**

**Unterstützung mehrerer Kerne/CPU's**

**SOA/Web-Services-Architektur**

**Erweiterte Dokumentidentifikation**

**Verteiltes, webbasiertes Scannen**

**Komplexe OCR – Zeilenweise Verarbeitung**

**Wiederverwendbare Geschäftsregeln**

**ICR-Handschrifterkennung**

**Einfache OCR – Header-Level**

**Optical Mark Recognition (OMR)**

**Barcodeerkennung**

**Key-From-Image-Funktionalität**

**Scan/Cleanup-Funktionalität**

**Leistungsspektrum und Ausgereiftheit**

# Einführung - Problem : Mangelnde Flexibilität



- Tasks in Reihenfolge und Umfang oft fest vorgegeben

# Agenda

- Einführung
- Datacap-Architektur
- Datacap-Oberflächen
- Stärken von Datacap

# Datacap Architektur

## → Flexibilität

- Orchestration der passenden Tasks in gewünschter Reihenfolge
- Orchestration der Tasks basierend auf Regeln
- Wiederverwendung von Workflows in Teilen bei Änderung der geschäftlichen Anforderungen

## → Skalierbarkeit

# Taskmaster – Orchestrierung der Erfassungsprozesse

- Taskmaster : Die zusammenfassende Umgebung zur Orchestrierung der Capture - Prozesse
  - Anwendung
    - Beispiel : Hypothek
  - Workflow
    - Beispiel : Vollständige Bearbeitung eines Antrags auf eine Hypothek
  - Job
    - Instanz eines Workflows
    - Zusammenfassung aller Workflow-Schritte
  - Task
    - Beispiel : Formularerkennung
      - Abarbeitung einer logisch vollständigen Funktion
      - Beispiele : Virtuelles Scannen – Klassifizierung – Upload – Visuelle Identifizierung – Erkennung
      - Aufbau mit Hilfe von Regeln
- Jobs und Tasks werden von **IHNEN** erstellt – mit Hilfe von Regeln und des Designtools Datacap Studio



# Taskmaster Komponenten

## → Taskmaster-Server

- Service für die Authentifizierung von Benutzern, Überprüfung von Berechtigungen und das Management von Stapelwarteschlange und Workflow
- Unterstützung von MS SQL- und Oracle-Datenbanken

## → Taskmaster-Web-Server

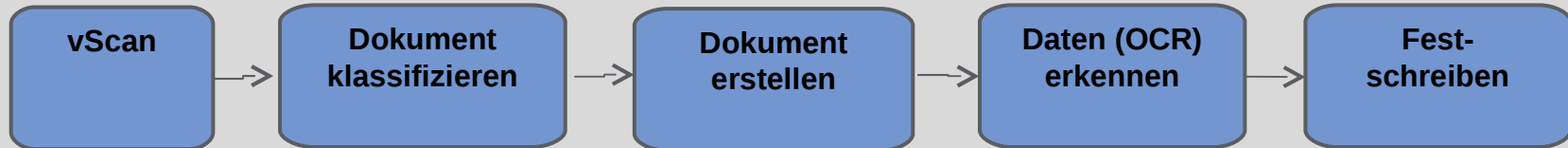
- Hosting der Taskmaster-Webanwendung
- Durchführung der Clientkommunikation
- Verwaltungsfunktionen – RV2-Berichterstellung und Dashboardüberwachung

# Rulerunner Enterprise

- Einheitlicher, regelbasierter Service
- Leistungsspektrum
  - Dateiimport – „virtuelles Scannen“
  - Imageverbesserung
  - Seiten- und Dokumentidentifikation
  - Datenextraktion mithilfe mehrerer OCR/ICR-Engines – Nuance, ABBYY, Recostar
  - Barcodeerkennung ( 1D und 2D )
  - Dokumentaufbau- und Integritätsprüfung
  - Datenprüfung, -bereinigung und -suche (Lookup)
  - Formatkonvertierung
  - Export von Dokumenten und Metadaten
- Multithreaded

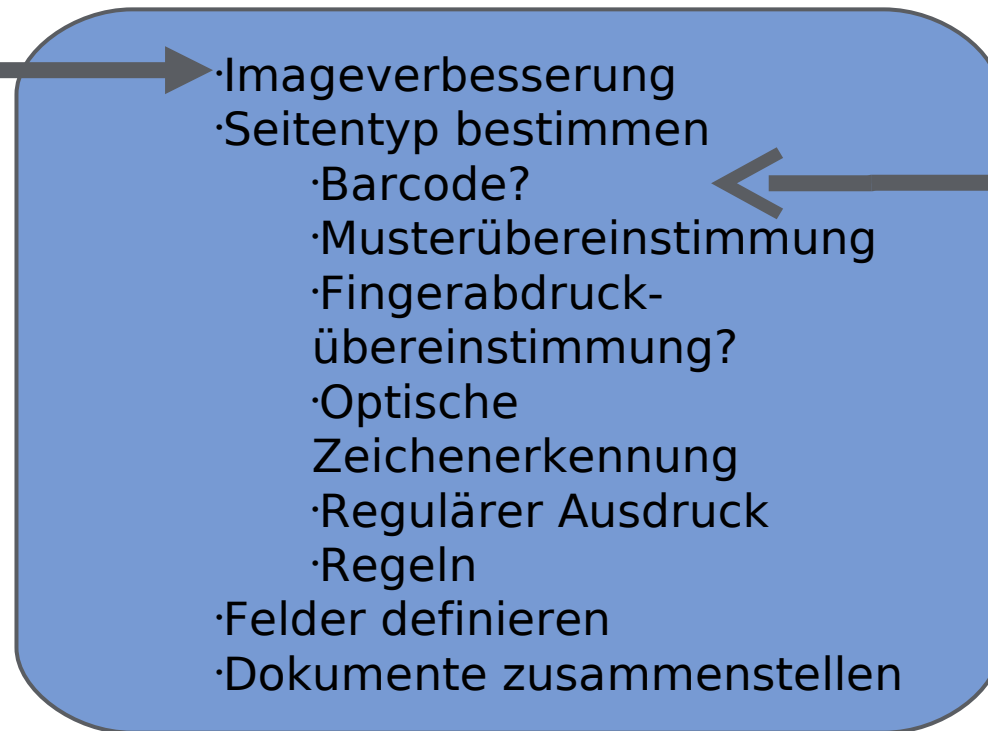
# Beispiel

**JOB**



*Regelbasierte Verarbeitung/regelbasierter Workflow*

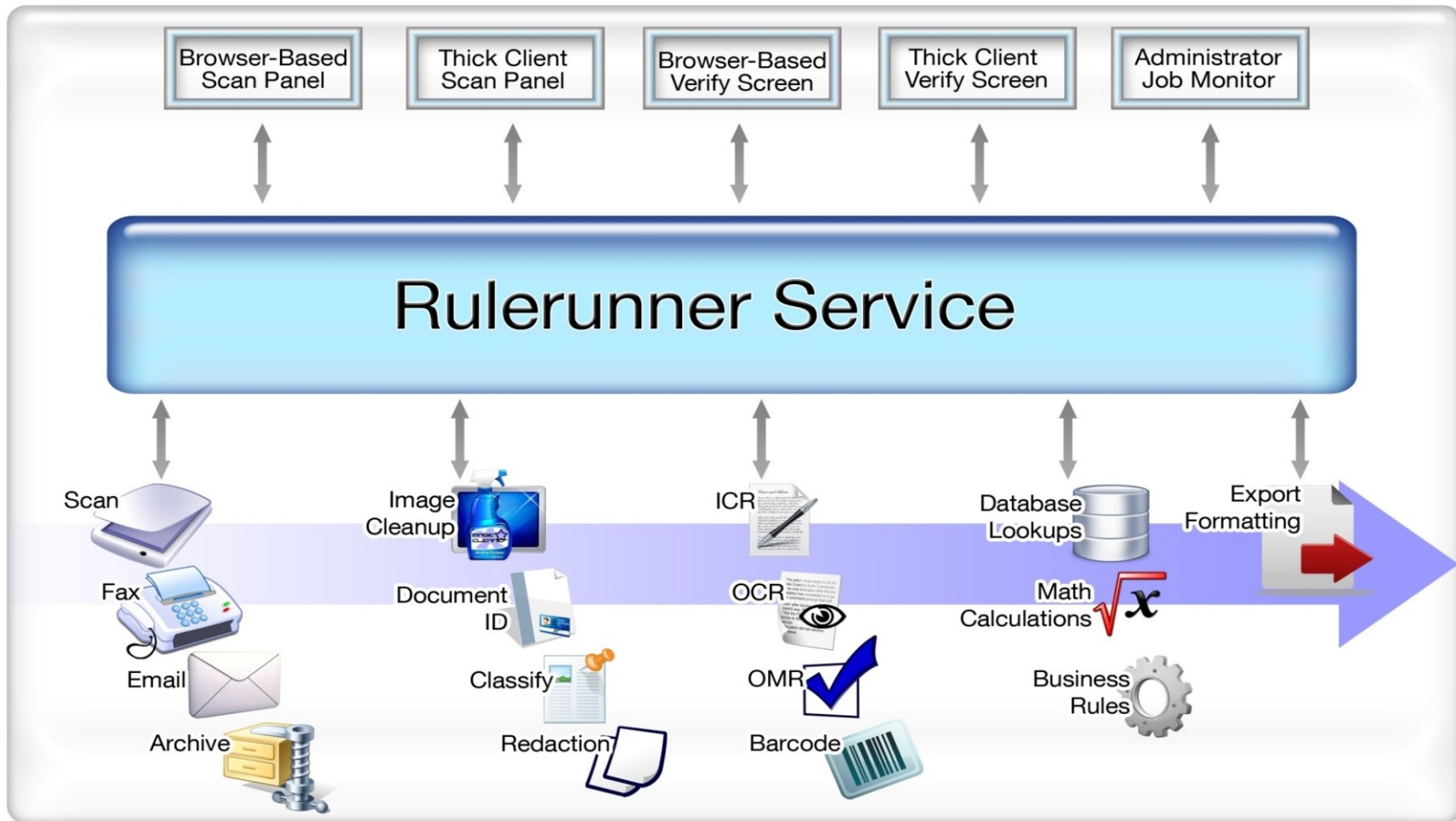
**TASK**



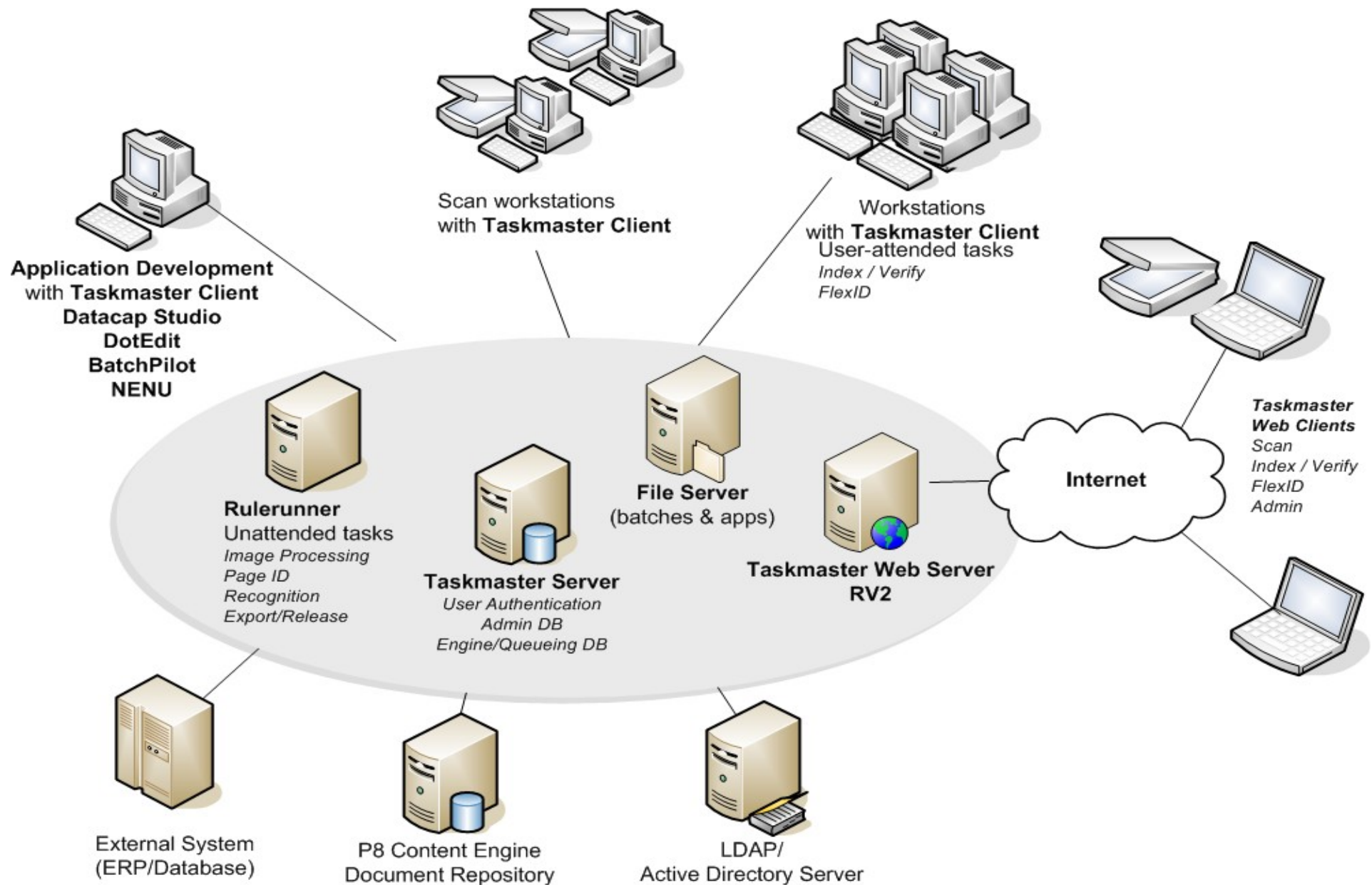
**REGELN**



# Rulerunner Service



# Physische Implementierung

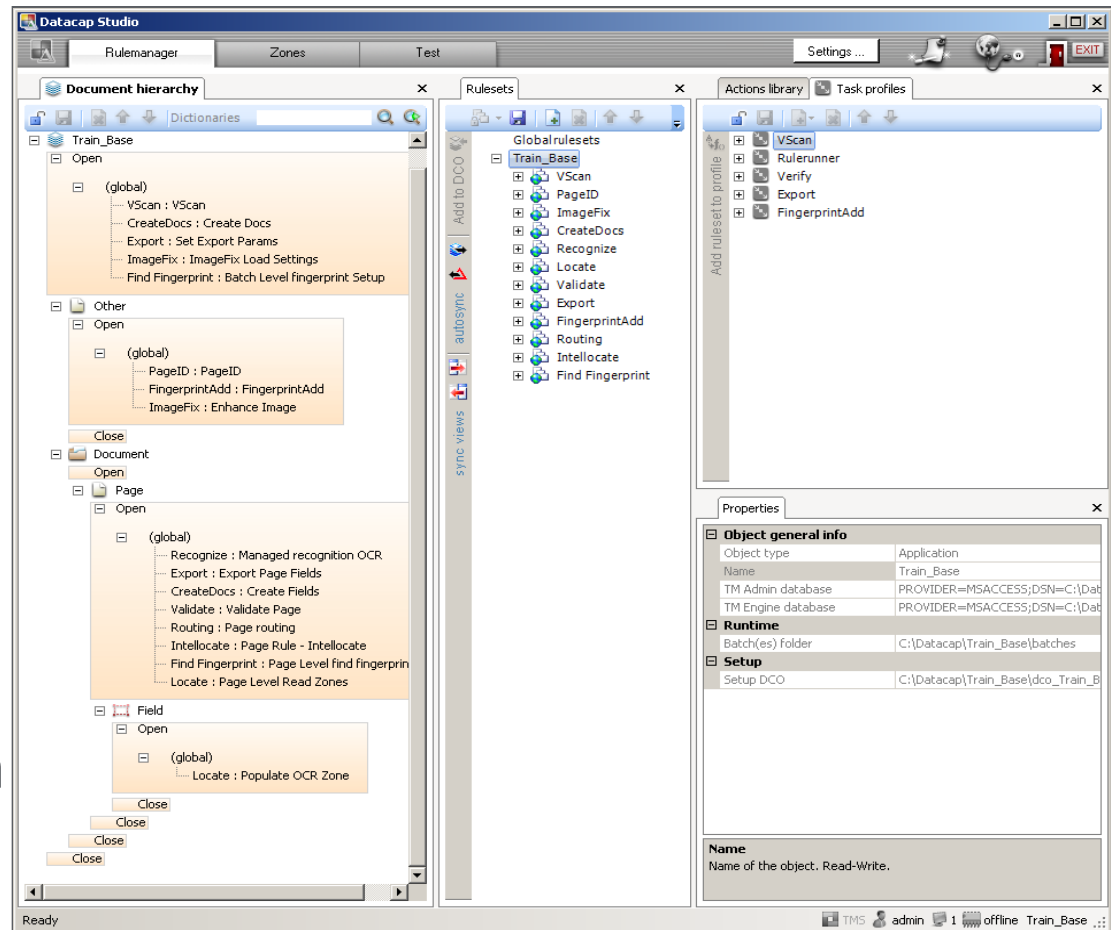


# Agenda

- Einführung
- Datacap-Architektur
- Datacap-Oberflächen
- Stärken von Datacap

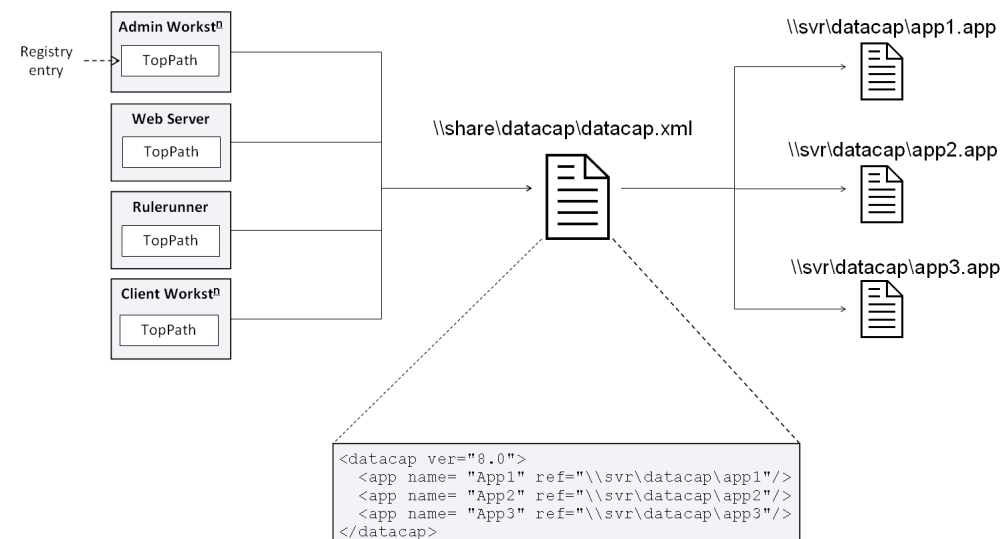
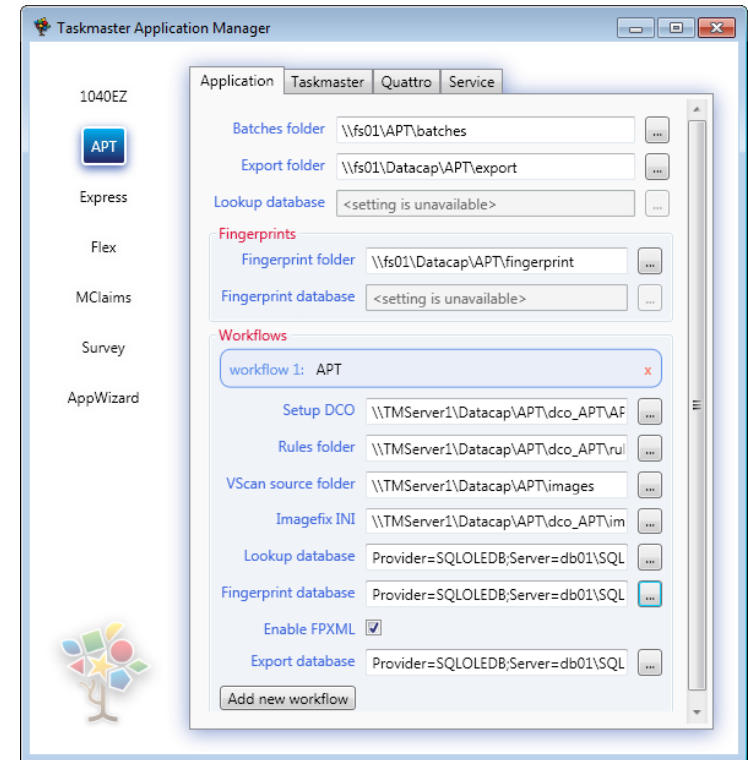
# Datacap Studio

- Anwendungsentwicklungs-umgebung
- Hunderte vordefinierter „Aktionen“ für das Konfigurieren von Anwendungen
- Aktionen sind Routinen, die alle Aspekte des Erfassungsprozesses steuern
- Offener Code für Entwickler verfügbar
- Angepasste Aktionen können der Bibliothek hinzugefügt werden



# Application Manager

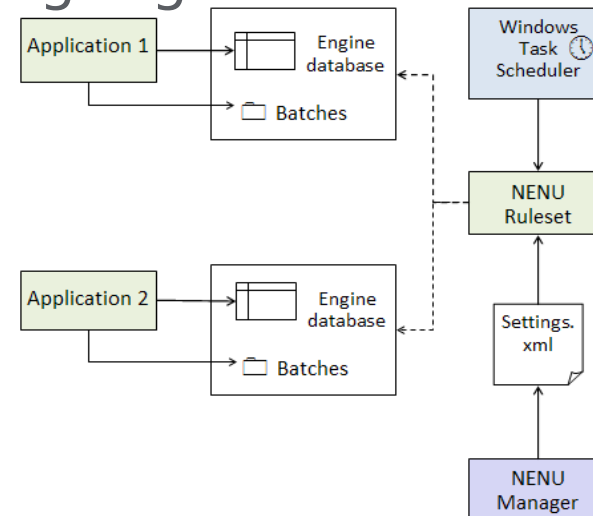
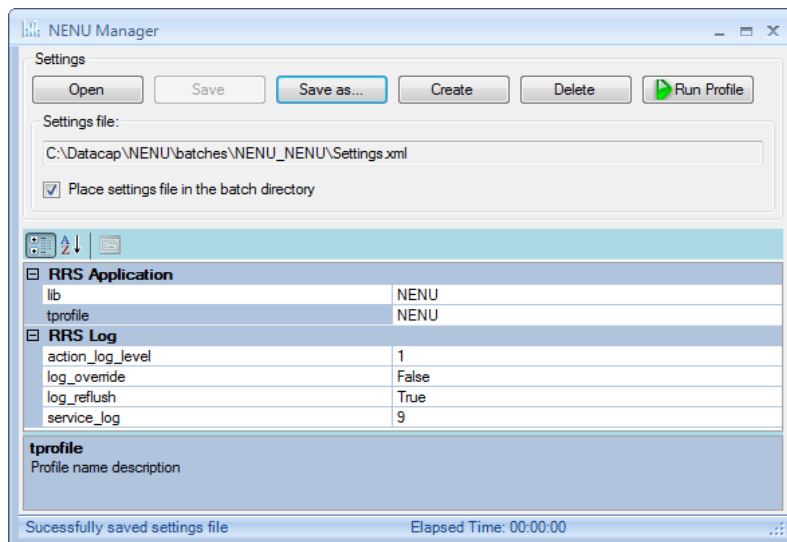
- Erlaubt die einfache Portierung/Hochstufung von Anwendungen zwischen Umgebungen (Entwicklung/Test/Produktion)
- Erlaubt die zentrale Konfiguration von verteilten Umgebungen
- Registriert die Position von Anwendungskomponenten und Ressourcen an einer zentralen Stelle
- Verschlüsselte Übertragung der Datenbankverbindungsdaten





# NENU

- Automatisiert wiederkehrende Systemtasks wie:
  - Batchüberwachung, Statusbenachrichtigung und automatisches Löschen von abgeschlossenen Batchoperationen (Housekeeping)
  - Erfassung von Snapshotdaten in einer Datenbank für RV2
  - Versand von E-Mail-Benachrichtigungen



# RV2

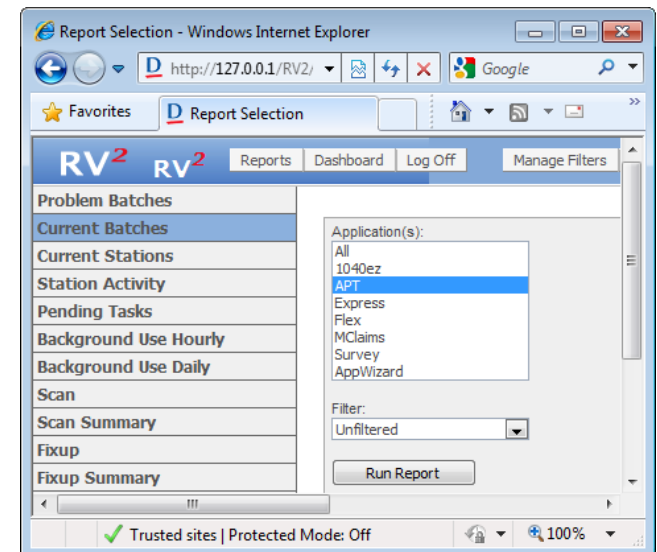
- Aktivitätenberichte in Echtzeit
- Webbasiert
- Erstellung von individuellen Berichten mit Visual Studio
- Anwendungsübergreifende Berichterstellung
- Kombinationsmöglichkeit von mehreren Berichten/Diagrammen in einem Dashboard
- Interaktiver Modus oder sequenzieller Modus mit NENU

The screenshot shows the RV2 Dashboard interface. At the top, there are navigation tabs for 'Reports', 'Dashboard', and 'Log Off', along with a 'Refresh: 1 Min' button and 'Add Remove' options. The main content is divided into two sections: 'Current Batches' and 'Background'. Both sections have a 'Filter: Unfiltered' indicator and a timestamp of '11/2/2010 8:25:08 AM'. The 'Current Batches' section contains a table with columns for Task, Job, Status, Batch, Priority, and Start. The 'Background' section contains a table with columns for Station, Batch, Job, Task, Status, Operator, Start Time, Time, Pages, and Docs. A 'Batch Count' summary is also present at the bottom of the Background section.

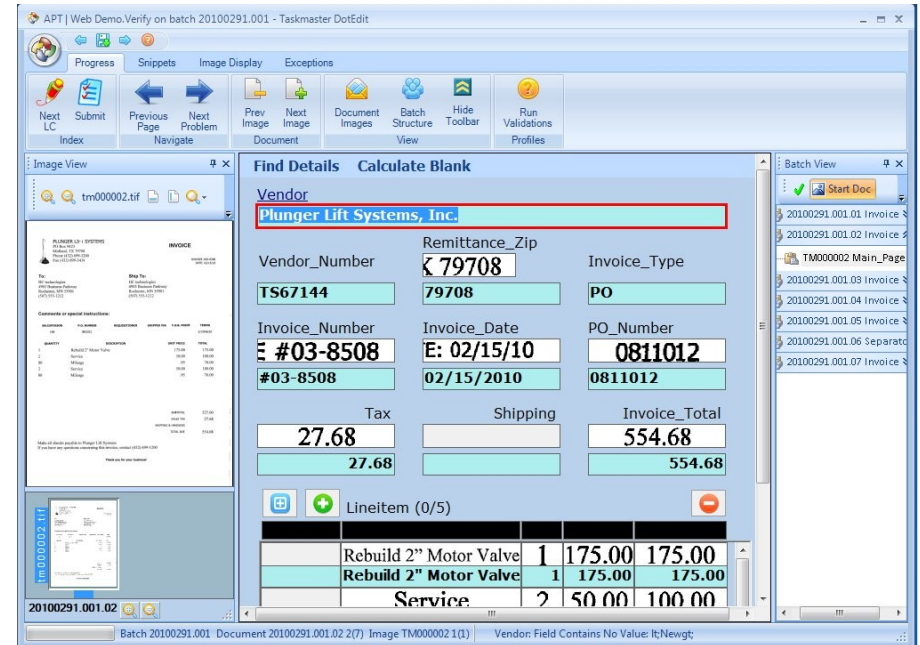
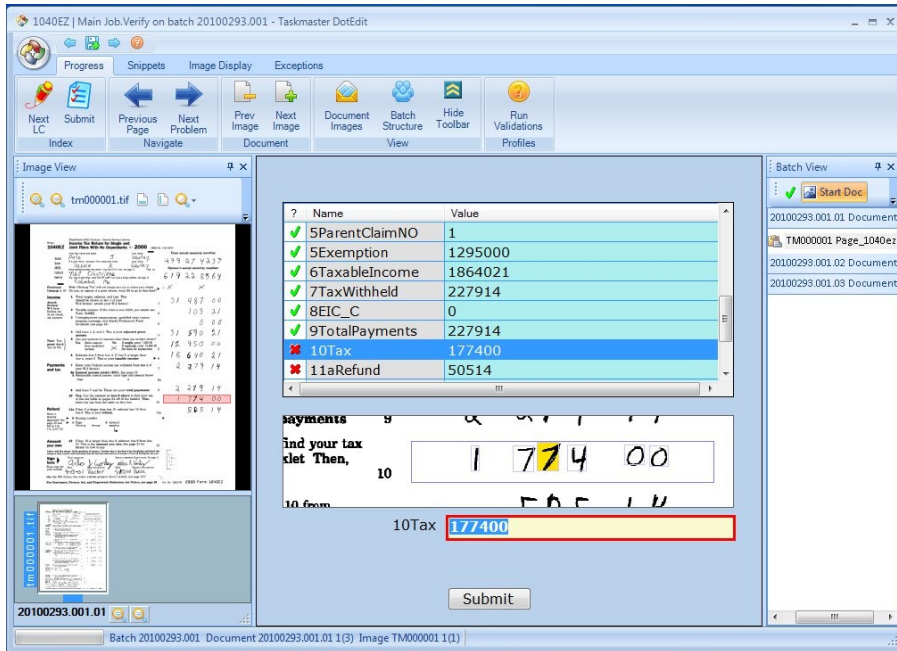
Task	Job	Status	Batch	Priority	Start
Batch Profiler	Demo	running	20100306.001	5	11/2/2010 8:16:44 AM
Verify	Demo	pending	20100306.002	5	11/2/2010 8:17:11 AM
Batch Profiler	Demo	pending	20100306.003	5	11/2/2010 8:18:57 AM
Batch Profiler	Demo	pending	20100306.004	5	11/2/2010 8:19:08 AM

Station	Batch	Job	Task	Status	Operator	Start Time	Time	Pages	Docs
1	5						00:01:09	40	7
	20100306.001	Demo	VScan	finished	admin	11/2/2010 8:16:47 AM	00:00:04	8	0
	20100306.002	Demo	VScan	finished	admin	11/2/2010 8:17:11 AM	00:00:02	8	0
	20100306.002	Demo	Batch Profiler	finished	admin	11/2/2010 8:17:26 AM	00:00:59	8	7
	20100306.003	Demo	VScan	finished	admin	11/2/2010 8:18:57 AM	00:00:02	8	0
	20100306.004	Demo	VScan	finished	admin	11/2/2010 8:19:08 AM	00:00:02	8	0
<b>Batch Count</b>							<b>Time</b>	<b>Pages</b>	<b>Docs</b>
5							00:01:09	40	7



# DotEdit



- Thick Client z.B. für Prüf-Task
- Automatisches Erstellen von „field-at-a-time“ Anzeigen
- Windows Forms / Visual Studio / C#

# DotEdit

APT | Web Demo.Verify on batch 20100291.001 - Taskmaster DotEdit

Progress Snippets Image Display Exceptions

Next LC Submit Previous Page Next Problem Prev Image Next Image Document Images Batch Structure Hide Toolbar Run Validations Profiles

Image View tm000002.tif

**Find Details Calculate Blank**

**Vendor**  
**Plunger Lift Systems, Inc.**

Remittance\_Zip  
 Vendor\_Number  Invoice\_Type

Invoice\_Number  Invoice\_Date  PO\_Number

Tax  Shipping  Invoice\_Total

Lineitem (0/5)

Quantity	Description	Unit	Price	Total
1	Rebuild 2" Motor Valve		175.00	175.00
2	Service		50.00	100.00
2	Service		50.00	100.00
2	Service		50.00	100.00

20100291.001.02

Batch 20100291.001 Document 20100291.001.02 2(7) Image TM000002 1(1) Vendor: Field Contains No Value: It;Newgt;



# Taskmaster Web Client & Flex

- Browserbasierter Client für Scan-, Prüf- und Verwaltungsoperationen
- Berichte und Dashboards
- Umfassend konfigurierbar
- Unterstützung mehrerer Browser

The screenshot displays the Taskmaster Web Client interface in a browser window. The main content area shows an invoice for JWS of Colorado Inc. with the following details:

**INVOICE**  
**JWS of Colorado Inc.**  
 1675 Broadway, Suite 2800  
 Denver, CO 80202-4628  
 Phone (303) 825-5467 Fax (303) 825-3434  
 INVOICE # 28100  
 DATE: 02/26/10

Below the invoice header, there are sections for 'TO' (HC technologies, 4903 Business Parkway, Rochester, MN 55901, (507) 555-1212) and 'SHIPPING TO' (HC technologies, 4903 Business Parkway, Rochester, MN 55901, (507) 555-1212).

A table lists items with columns for 'salesperson', 'PO', 'SHIPPING METHOD', 'shipping terms', and 'delivery d'. The items listed are:

salesperson	PO	SHIPPING METHOD	shipping terms	delivery d
195	678430			

Another table shows a list of items with columns for 'qty', 'item #', and 'description':

qty	item #	description
10	201	Trucking - Ticket 138432
130	204	Brine Water
130	203	Fresh Water
19	206	KCL
6	201	Trucking - Ticket 146611
130	203	Fresh Water
19	206	KCL

On the right side of the interface, there is a list of tasks with columns for task ID, description, and count. The tasks listed are:

Task ID	Description	Count
20110069.001	APT	1
20110069.001.01	Invoice	0
TM000001	Main_Page	1
20110069.001.02	Main_Page	0
TM000002	Main_Page	0
20110069.001.03	Invoice	0
TM000003	Main_Page	2
20110069.001.04	Invoice	0
TM000004	Main_Page	1
20110069.001.05	Invoice	0
TM000005	Main_Page	1
20110069.001.06	Separator	0
TM000006	Document_Separator	49
20110069.001.07	Invoice	0
TM000007	Main_Page	1
TM000008	Trailing_Page	49

The interface also includes a 'Submit' button and a navigation bar with 'Prev', 'Next', 'keys', and 'Hold' buttons. The browser window title is 'Demo-WebFlexID.WebFlex : 20110069.001'.

# Agenda

- Einführung
- Datacap-Architektur
- Datacap-Oberflächen
- Stärken von Datacap

## Stärken von Datacap

- Taskmaster – Orchestrierung der Erfassungsprozesse
- Rulerunner Enterprise – zusammenfassende Umgebung
- Web-Services
- Regelbasierte Klassifizierung und Erkennung
- Offene Schnittstellen
- Lizenzmodell

# Orchestrierung der Erfassungsprozesse

- Einfache Anpassung der Technik bei Änderung der Geschäftsprozesse
- Einfache Umsetzung der geschäftlichen Anforderungen mit Hilfe von Regeln
- Optimale Nutzung der Ressourcen durch inhaltsabhängige Auswahl der Tasks



# Rulerunner Enterprise

- Die zusammenfassende Management-Umgebung
- Zentrale Instanz
- Hochverfügbar auszulegen

# Offene Schnittstellen

- Web-Services : Flexible Schnittstelle zu anderen Systemen
- Visual Basic Script
- C#

# Lizenzmodell

- Nutzerbasierte Lizenzen – Web-Client eingeschlossen
- Keine Volumenabhängigen Kosten
- Bei sehr hohen Leistungsanforderungen  
Processorbasierende Lizenz

# Weitere Informationen

<http://www-01.ibm.com/software/data/content-management/datacap/>

Michael.Vahland@de.ibm.com

Tel. : 01 72/8 37 71 24