

# Special Interest Group Oracle WebCenter

Eric Bos  
Oracle ECM Consultant  
28 Oktober 2013

# Oracle WebCenter Capture

1. Webcenter Capture vs OFR (Perceptive IDC)
2. WebCenter Capture
3. Workspaces en andere begrippen
4. DC-Console: Workspace
5. DC-Client: Scannen
6. Oracle Forms Recognition / Brainware

- **Webcenter Capture**
  - “Dom” scannen
  - OCR t.b.v. textsearchable maken PDF
  - Barcodes:
    - Document Scheiding
    - Classificatie en data extractie
  - Handmatig indexering
  - Scripts: JavaScript
- **Oracle Forms Recognition (Perceptive IDC)**
  - Intelligent Document Capture
  - OCR t.b.v. data extractie & textsearchable maken PDF
  - Classificatie op basis van o.a. layout
  - Documentscheiding door scanner of via barcodes
  - Indexering op basis van data extractie
  - Script: Visual Basic

- **Overeenkomst Webcenter Capture & OFR**
  - Beiden kunnen gebruik maken van database lookup velden
  - Indexering via DC-Client (Capture) / (Web)Verifier (OFR)
  - Oplevering output in diverse vormen

- **Webcenter Capture**
  - Geen desktop applicatie meer
  - Alleen webapplicatie in Weblogic domain
  - Geen servers meer
    - Import Server → Import Processor
    - Recognition Server → Recognition Processor
    - Commit Server → Commit Processor
  - VB → JavaScript
  - Gebruikers nu in Weblogic, EM en met OPSS
  - Capture rollen:
    - Capture User: scannen en/of indexeren
    - Capture Workspace Manager: beheren workspaces
    - Capture Workspace Viewer: bekijken workspaces
  - Batch Processing Data in EM
  - Let op:
    - Linux: geen Recognition Processor en geen Searchable PDF
    - Bug in UCM Commit Processor: server op IP koppelen, niet op servernaam!!

- Workspaces: scan & indexeer omgeving

**ORACLE WebCenter Capture Workspace Console** Preferences Help Sign out About Signed in as Alice.Fransen

**Capture Workspaces** + [ ] ✕ ?

Filter

Name	Description
Scan2Archive	Documenten scannen en opslaan in ucm

**Workspace: Scan2Archive**

Security **Metadata** Classification Capture Processing Commit Advanced

**Metadata Fields** + [ ] ✕ ?

Name	Field Type	Max Length	Required	Locked	Display Format	Input Mask	Auto Populate
Afdeling	Alpha Numeric	0	No	No			(none)
Barcode	Alpha Numeric	0	No	No			(none)
BarcodeData	Alpha Numeric	0	No	No			(none)
BatchNaam	Alpha Numeric	0	No	No			Batch Name
Behandelaars	Alpha Numeric	0	No	No			(none)
BehandelaarsID	Alpha Numeric	0	No	No			(none)
Bestandsnaam	Alpha Numeric	0	No	No			(none)
DocumentSoort	Alpha Numeric	0	No	No			(none)
DocumentBeschrijving	Alpha Numeric	0	No	No			(none)
DossierDocument	Alpha Numeric	0	No	No			(none)
DossierStatus	Alpha Numeric	0	No	No			(none)
DossierType	Alpha Numeric	0	No	No			(none)
Scandatum	Date		No	No			Scan Date

**User Defined Choice Lists** [ ] ✕

Type	Source Name	Choice List
<input checked="" type="checkbox"/> Document Capture	Capture Source	
<input type="checkbox"/> Document Capture	Capture Source	DossiersDocumenten
<input type="checkbox"/> Document Capture	Capture Source	DossierType

- **Workspaces:**
  - Security: rechten op workspace
  - Metadata:
    - Metadata velden
    - Choice lists
    - Database choice lists
    - Dependent choice lists
    - Database lookup profiles
  - Classification:
    - Batch Statussen
    - Document Profiles
  - Capture:
    - Client Profiles
    - Import Processor Jobs
      - Import van directory
      - Import van email
      - Import van lst-files

- Workspaces: scan & indexeer omgeving
  - Processing:
    - Recognition Processor Jobs
      - Barcode pagina scheiding
      - Barcode lezen
    - Document Conversion Jobs
  - Commit:
    - Commit to Webcenter Content
    - Commit to Webcenter Content Image
    - Commit to Textfile
  - Advanced: scripts



- DC-Console:
  - Beheren workspaces
  - DC-Console:
  - Beheren workspaces
  - <http://server:16400/dc-console>
- DC-Client:
  - Capture (scannen) van documenten
    - Scanner
    - Filesystem
  - <http://server:16400/dc-client>

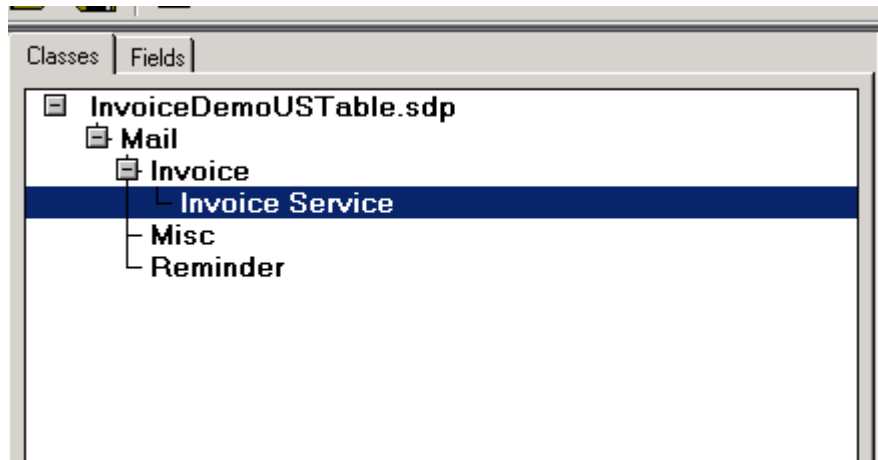
- Oracle Forms Recognition – Perceptive IDC
  - OFR is OEM-versie Perceptive IDC
  - Verschillen:
    - OFR lagere versie dan Perceptive IDC
      - Dus ook lagere ABBY Recognition Engine
    - OFR werkt met filesystem, Perceptive IDC met database
    - OFR client verifieer applicatie, Perceptive IDC WebVerifier
    - OFR integratie met Webcenter Capture & Webcenter content, Perceptive IDC heeft dit niet

- Perceptive IDC
  - Diverse processen, aan & uitzetten

Input State:	Process Step:	Output State:
	Import	succeeded: 100 failed: 50
100	→ OCR	succeeded: 200 failed: 150
200	→ Document Separation	valid: 230 invalid: 215
200	→ Classification	classified: 300 not classified: 250
300	→ Extraction	valid: 700 invalid: 550
700	→ Export	valid: 800 invalid: 750
650	→ Custom Processing	succeeded: 700 failed: 690
450	→ Database Export	succeeded: 800 failed: 760
800	→ Clean Up	

Perform advanced import failure processing  
 Perform folder based classification+extraction step  
 Perform folder based serial processing  
 Corrupted document's failure state: 0

- Perceptive IDC
  - Classificatie op basis van
    - Layout
    - Woorden
    - Taal
  - Subclassificatie



- Perceptive IDC
  - Extractie op basis van
    - Zones
    - Simple & Regular Expressions
    - Woorden
    - Tabel extractie

**Classes** | Fields

[1] Date  
3/9/00

[2] Total  
23,728.60

[3] Net  
23,379.10

[4] InvoiceNumber  
6-323-785

[5] Table

Code	Text
1	745-00 BMX Toddler
2	367-10 BMX Mountain
3	854-21 Road Warrior

**Surf Side Ride**

Customer Remittance:      Date: 3/9/00  
Invoice #: 6-323-785

Ride-A-Mile Club  
541 Surf Side Drive  
Clearwater Beach, CA 90632

Please Remit:  
Surf Side Ride  
33 Seagull Lane  
Akron, CA 65690

Number	Description	Qty	Price	Amount
745-00	BMX Toddler Bike	10	499.95	4,999.50
367-10	BMX Mountain Bike	20	699.98	13,999.60
854-21	Road Warrior Bike	4	1,095.00	4,380.00

**Analysis** | Evaluation | Field | Validation

Available Templates  
<No template> Copy Template

Available Analysis Engines  
Format Analysis Engine Save as Template

Format Analysis | General | Regions

Analysis Method  
 Find Format  
 Find Designator

Prefix Suffix  
Ignore Characters

Designator Type  
Next Word Syntax help  
# numeric 0-9  
@ alpha a-z  
? alpha or numeric  
Simple Expression start of word  
end of word

? Format Strings

0	###@1#2###0.0
1	#[3.8]
2	

- Perceptive IDC: Extra voorbeelden extractie

?	Format Strings
0	[-]\{0,1\}[ ]\{0,1\}[0-9]\{0,2\}[-]\{0,1\}[0-9]\{0,2\}[-]\{0,1\}[0-9]\{0,2\}[-]\{0,1\}[0-9]\{2\}[ ]\{0,1\}[C]\{0,1\}[-]\{0,1\}
1	

?	Format Strings
0	relatienummer
1	rc-nummer
2	rek.courantr.
3	RC-Nummer
4	[0-9]\{4,6\}
5	***[0-9]\{7\}
6	[0-9]\{1\}
7	atienummer
8	[0-9]\{2\}[-][0-9]\{3\}
9	[0-9]\{1,4\}[ ]\{0,1\}[0-9]\{1,5\}
10	

?	Format Strings
0	[a-gi-pr-v]\{3,8\}[ ]\{0,10\}[0-9]\{2,4\}
1	

- Perceptive IDC: Output voorbeeld

```
PRI_3493_045DA_2011_12 - Kladblok
Bestand  Bewerken  Opmaak  Beeld  Help
3493;S079;01-12-2011;00;;;-24,47;-24,47;33758
3493;S079;01-12-2011;00;;;-6.650,33;-6.650,33;33758
3493;S079;01-12-2011;00;;;-39,28;-39,28;33758
3493;S079;01-12-2011;10;L201     ;;;-2,72;-2,72;33758
3493;S079;01-12-2011;10;L201     ;;;2,62;2,62;33758
3493;S079;01-12-2011;10;L201     ;;;-2,71;-2,71;33758
3493;S079;01-12-2011;10;L201     ;;;4,34;4,34;33758
3493;S079;01-12-2011;10;L201     ;;;-2,33;-2,33;33758
3493;S079;01-12-2011;50;;;0,00;0,00;6.712,92;33758
```

- Demo
  - Webcenter Capture
  - Perceptive IDC