

kulturgut scanner

bürogut scanner

Marken der MIK-CENTER GmbH

Migration auf Production 3

Eine Erfolgsgeschichte

Einrichtung: Bauamt der Stadt Den Haag, NL

Vertragspartner:
NMS (National Multiple Services)

Servicepartner:
Zeuschel GmbH, Tübingen
repräsentiert durch Frank-Ulrich Weber

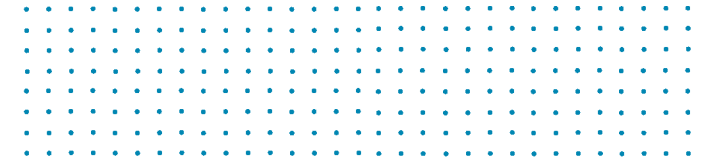
MIK-CENTER GmbH, Berlin
repräsentiert durch Matthias Ronge

NMS
NATIONAL MULTIPLE SERVICES

ZEUSCHEL
Kulturgut ist unsere Leidenschaft

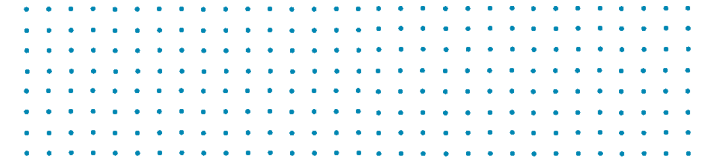
MIK
CENTER GmbH





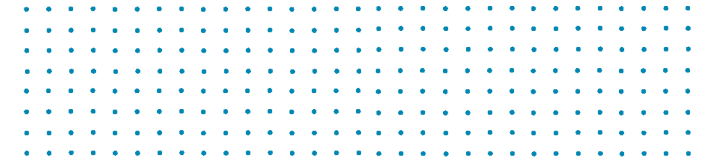
Das Bauamt der Gemeinde Den Haag

- Projekt: Digitalisierung aller Verwaltungsfälle von vor 1900–2010
- Laufzeit: 5 Jahre, Beginn 01/2020
- Zu scannen: 3 Regal-km
- Mai 2022: 150 Meter gescannt, 200.000 Images = 40 TB unkomprimierte TIFFs



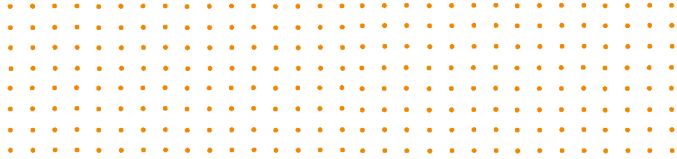
Bisheriger Arbeitsablauf

- Es wurde Kitodo.Production 2 mit Zeutschel ZED-Server eingesetzt
- Erstellung leerer Vorgänge für Fälle pro Flurstück (initial 10)
- Laufzettel drucken
- Einlegen der Laufzettel in die Akten in der Reihenfolge ihres Auffindens



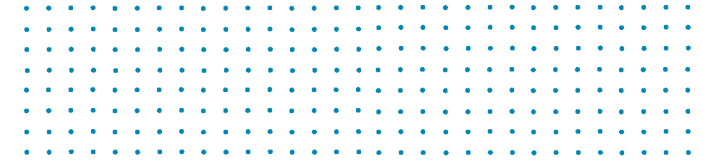
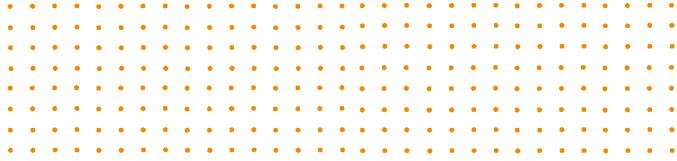
Bisheriger Arbeitsablauf: Nachteile

- Unkomprimierte TIFFs verbrauchen zu viel Speicherplatz
- JPEG-komprimierte TIFFs werden von OpenJDK nicht unterstützt; Oracle Java ist keine Option
- Keine weitere Entwicklung in Production 2
→ Entscheidung: Migration auf Production 3



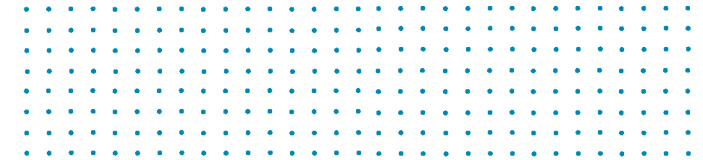
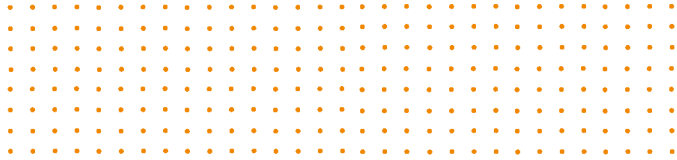
Sonderanpassung in Production 2

- Funktion, Straßenkarten-Metadaten im Metadateneditor herunterzuladen
 - RDF-Import-Knopf in Version 2 wurde umfunktioniert, um die Daten von einem selbstgeschriebenen Datenbankserver zu lesen
 - Ersatz in Version 3 erforderlich



Spezielle Anforderungen an Production 3

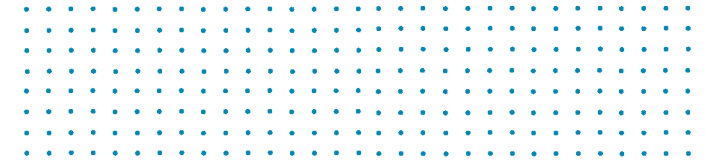
- Ersatz für den kundenspezifischen Datenbankzugriff im Metadateneditor
- Bei hierarchischen Vorgängen:
 - Zeige eine Zusammenfassung der Fortschritte der Unterordnungen bei der Überordnung (bisher weiß)
 - Export von oben nach unten ermöglichen
- Hierarchieverwaltungsfunktionen reaktivieren
- Automatische Vorgangstitelerstellung mit automatischer Nummerierung für Unter-Vorgänge



Ersatz für Adressdatenbank im Metadateneditor

- Bisher: Möglichkeit, im Regelsatz die Werte (<option >s) von Auswahlfeldern für ein Strukturelement zu begrenzen
→ Das Gleiche abhängig vom Wert eines anderen Metadatenfeldes

```
<restriction division="Monograph">
  <permit key="Person">
    <permit key="Role" unspecified="forbidden">
      <permit value="aut"/>
      <permit value="edt"/>
    </permit>
  </permit>
</restriction>
<restriction division="Photograph">
  <permit key="Person">
    <permit key="Role" unspecified="forbidden">
      <permit value="dpc"/>
      <permit value="pht"/>
    </permit>
  </permit>
</restriction>
```



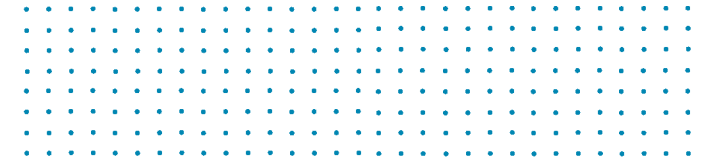
Ersatz für Adressdatenbank: Folgen

- ein Regelsatz mit 65 MB, 11 Millionen Zeilen XML
- AJAX-Aktualisierungen der Anzeige langsam (3 Sek.)
 - ergänzende Entwicklung: <include >s im Regelsatz ermöglichen
- später: Funktionalität sollte auch bei der Vorgangserstellung verfügbar sein (nicht stabil)
 - Entscheidung: Katalogimport mit kundenspezifischem SRU-Webservice mit den Adressdaten

Bedingte Werte- begrenzungen

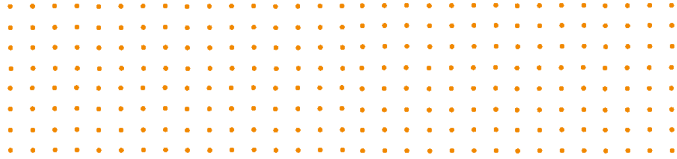
```
<declaration>
  <key id="area">
    <label>Area code</label>
    <label lang="de">Flurstücknummer</label>
    <option value="1234-2a"/>
    <option value="1234-2b"/>
    <!-- ... -->
  </key>
  <key id="address">
    <label>Address</label>
    <label lang="de">Adresse</label>
    <key id="street">
      <label>Street</label>
      <label lang="de">Straße</label>
      <option value="Buschweg"/>
      <option value="Pastor-Luis-Straße"/>
      <option value="Rosenstraße"/>
      <option value="Vereinsstraße"/>
      <!-- ... -->
    </key>
    <key id="house">
      <label>House no.</label>
      <label lang="de">Hausnummer</label>
      <option value="1"/>
      <option value="1a"/>
      <option value="1b"/>
      <option value="2"/>
      <option value="2 Rgb."/>
      <label>2, rear building</label>
      <label lang="de">2, Rückgebäude</label>
      </option>
      <option value="2a"/>
      <option value="3"/>
      <!-- ... -->
    </key>
  </key>
</declaration>
```

```
<correlation>
  <restriction key="address">
    <permit key="street" minOccurs="1" maxOccurs="1" unspecified="forbidden">
      <condition key="../area" equals="1234-2b">
        <permit value="Rosenstraße"/>
        <permit value="Vereinsstraße"/>
      </condition>
    </permit>
    <permit key="house" minOccurs="1" maxOccurs="1" unspecified="forbidden">
      <condition key="../area" equals="1234-2b">
        <condition key="street" equals="Rosenstraße">
          <permit value="3"/>
          <permit value="5"/>
          <permit value="7"/>
        </condition>
        <condition key="street" equals="Vereinsstraße">
          <permit value="2"/>
          <permit value="2 Rgb."/>
          <permit value="4"/>
          <permit value="4a"/>
        </condition>
      </condition>
    </permit>
  </restriction>
</correlation>
```



Bedingte Wertebegrenzungen: Beispiel

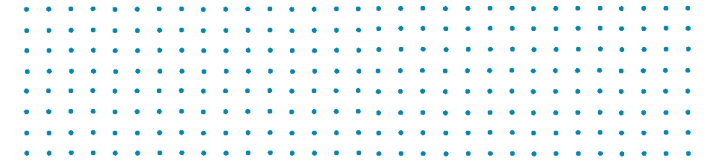
- `<mods:genre authority="...">...</mods:genre>`
 - `authority="aad"`: Gattungsbegriffe der Arbeitsgemeinschaft Alter Drucke
 - `authority="marcgt"`: MARC Genre Terms List



<include>s im Regelsatz

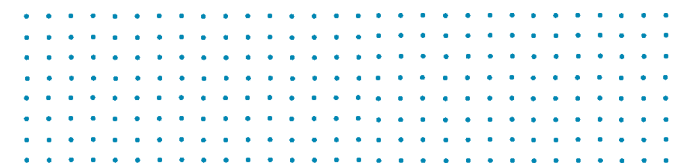
- Ein Regelsatz pro Projekt
- Regelsatz inkludiert gemeinsame Definitionen
- Nur spezielle Definitionen werden überladen (hier Kollektionen)
- Extrem lange Definitionen in separater Datei möglich

```
<ruleset xmlns="http://names.kitodo.org/ruleset/v2">  
  <include>bib_shared.xml</include>  
  <include>bib_shared_language-field.xml</include>  
  <declaration>  
    <key id="singleDigCollection">  
      <label>Digital collection</label>  
      <label lang="de">Digitale Kollektion</label>  
      <option value="eins">  
        <label>Collection 1</label>  
      </option>  
      <option value="zwei">  
        <label>Collection 2</label>  
      </option>  
      <option value="drei">  
        <label>Collection 3</label>  
      </option>  
      <preset>eins</preset>  
    </key>  
  </declaration>  
</ruleset>
```



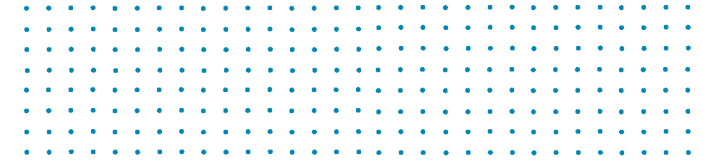
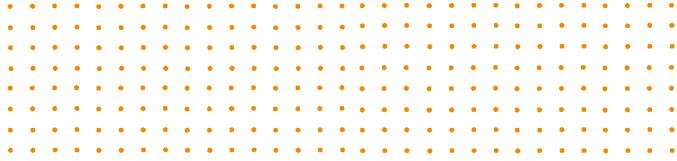
Export von Vorgängen von oben nach unten

- Wenn eine Überordnung exportiert wird, die Unterordnungen hat, die abgeschlossen sind aber noch nicht exportiert wurden, werden diese mit der Überordnung zusammen exportiert
- Verwendung: Kein Exportschritt im Workflow, der Projektverantwortliche exportiert die Überordnung irgendwann manuell



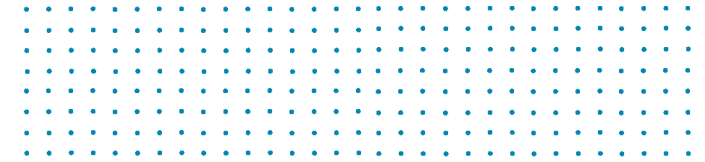
Eigener SRU-Webservice

- Suchen: Adresse auf Flurstück / alle Gebäude einer Straße auf dem Flurstück als Bereich
- Suchantwort: Minimaler PICA+-Datensatz nur mit Treffer-ID (003@.0) und Label (021A.a)
- Trefferantwort: gibt <kitodo:metadata> zurück
- XSLT: Datensatz unverändert durchleiten



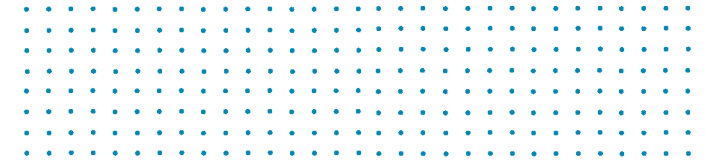
Automatische Nummerierung von Unterordnungen

- Titel der Überordnung als Vorsatz für den Titel der Unterordnung ermöglichen
- Automatische Belegung eines Metadatenwertes mit dem nächsten Wert der Bandzählung
- Integer-Werte mit führenden Nullen ermöglichen



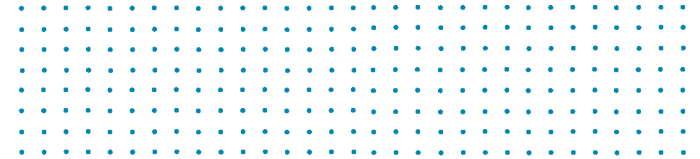
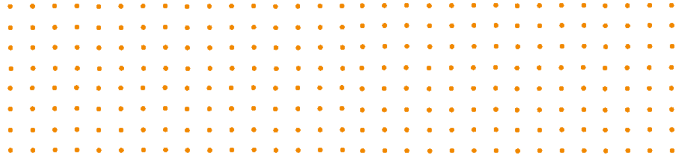
Titel der Überordnung als Vorsatz für den Bandtitel

- Den Vorgangstitel der Unterordnung mit dem Titel der Überordnung beginnen
- Vorgangstiteldefinitionen mit “+” zu beginnen funktionierte nur für Zeitungen
- Vorgangstiteldefinitionen mit “+” vorne wurde für Unterordnungen ermöglicht
- Beispiel: `<division id="request" processTitle="+'_'+DocumentNumber">`



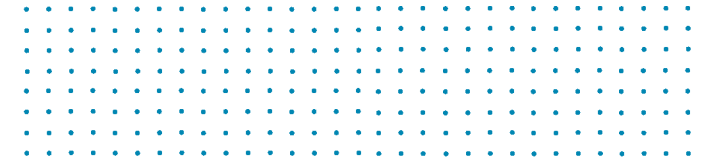
Automatische Vergabe der Bandzählung

- Beim Erstellen einer Unterordnung soll ein Metadatenfeld die nächste Bandzählung enthalten
- Nummer des nächsten Bandes := Momentane Anzahl Unterordnungen + 1
- Zu aktivieren mittels: `<key id="DocumentNumber" use="childCount">`



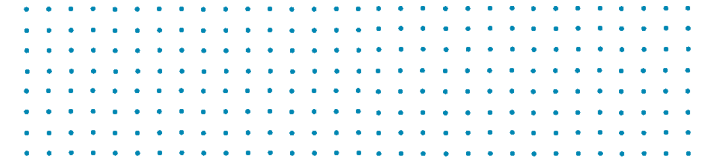
Integer-Werte mit führenden Nullen ermöglichen

```
<key id="DocumentNumber">  
  <label>Document number</label>  
  <codomain type="integer" minDigits="4"/>  
</key>
```



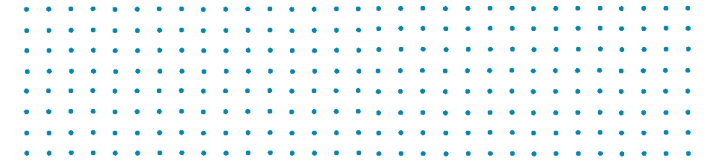
Stufen der Software-Inbetriebnahme

- OT: Umgebung für Probeinstallationen (operational technology)
- ACC: Umgebung für Abnahmetests (accptance test)
- PROD: Umgebung für Produktivsysteme
- jeweils mit zwei IP-Adressen, eine interne und eine externe



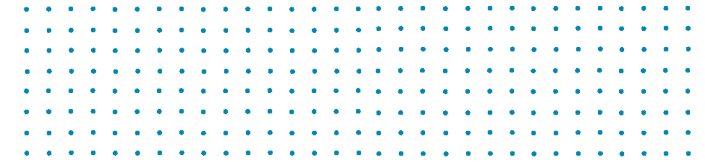
Weitere Anforderung: Umstellung auf JPEG

- Perl-Skript mit ImageMagick
 - TIFF-Dateien nach JPEG konvertieren
 - Vorschaubilder (Thumbnails) erstellen
 - Dateiendung und MIME-Typ in meta.xml ändern



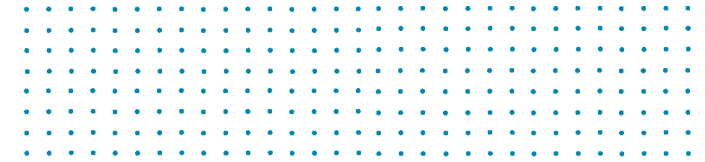
Migration des Produkktivsystems: Tag 1

- Datenbankmigration mit Reparaturen (PR #4049) problemlos
- Komponenten installieren und Webanwendung aktualisieren
 - Error connecting to ElasticSearch!



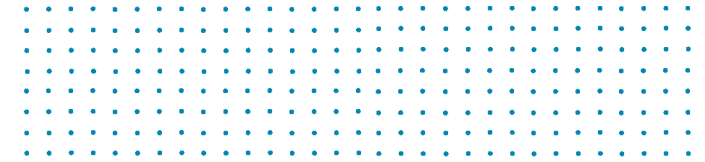
Das Elasticsearch-Problem

- Elasticsearch beschwerte sich über „bad login credentials“
- Ursache: Der Paketmanager installierte Elasticsearch v. 8 (nicht kompatibel)
- Problem bestand nach Deinstallation und erzwungener Installation von Version 7 weiter
- Lösung: deinstallieren, manuell alle verbliebenen Elasticsearch-Dateien und -Verzeichnisse löschen, dann Version 7 installieren



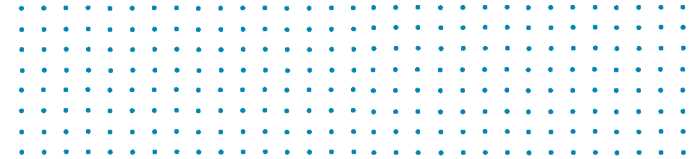
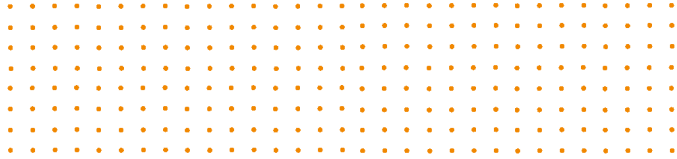
Migration des Produktivsystems: Tag 2

- LDAP-Server in Datenbank konfigurieren
- Rolle „Almighty“ erstellen und administrativen Benutzern zuordnen
- Warten, dass die Bildkonvertierung fertig war (lief bereits seit 8 Tagen)
- Start der Metadatenmigration



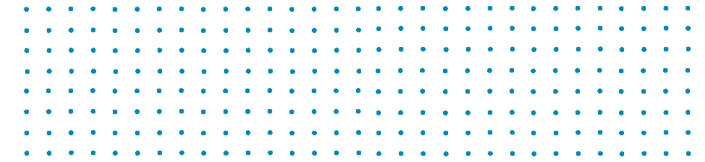
Migration des Produkktivsystems: Tag 3

- Suchindex aufbauen
- Ersten Workflow erstellen
 - Zuordnen des Workflow zu 395.000 Vorgängen dauerte 17 Stunden!
 - brachte mir einen freien Nachmittag am Strand ein



Migration des Produktivsystems: Tag 4

- In der Workflow-Migration erschienen plötzlich weitere 17.000 Vorgänge
- Neues Projekt für künftige Vorgänge angelegt
- Test
- Konfiguration anpassen



Gefundene Probleme

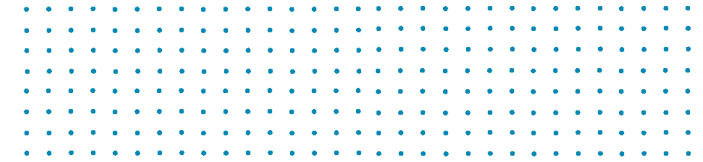
- Die Benutzerliste ist nicht sichtbar, bevor das Indexprofil an ElasticSearch gesendet wurde
- Viele Probleme beim Anlegen von Metadatengruppen bei der Vorgangserstellung!
- Unerwünschtes Tausender-Trennzeichen Komma erscheint in Zahleneingabe (z.B.: Jahr = 1,845)
- „DocType“ wird nicht übersetzt
- Presentation: tx_dlf_pageviewproxy erzeugt eine 10-sekündige Pause (Timeout) vor Bildanzeige



kulturgut scanner

bürogut scanner

Marken der MIK-CENTER GmbH



Kitodo.Presentation-Upgrade

- Upgrade von TYPO3 7.6 LTS auf TYPO3 9.5 LTS
- Upgrade von Kitodo.Presentation 2.2 auf 3.2.3
- Upgrade von Solr 3.6.2 auf 8.11.2
- Upgrade auf TYPO3 11.5 LTS geplant...

Den Haag **START** **ZOEK** **COLLECTIES**

Digitaal Bouwarchief 1900-2010 Gemeente Den Haag

In dit archief vindt u de documenten geordend in dossiers. Per dossier wordt één zaak behandeld en treft u zowel tekeningen als tekstdocumenten aan, die op het moment van digitaliseren zijn gevonden. Het kan dus voorkomen dat er op een later moment documenten worden teruggevonden die in het door u geraadpleegde dossier thuishoren.

Naast het zoeken op adres kan er ook gezocht worden op CR-nummer. Dit nummer werd vanaf 1900 aan een bouwplan gegeven: de bouw van één of meerdere bouwwerken op een perceel. Indien een zoekopdracht op één adres geen resultaten oplevert kan op CR-nummer het oprichtingsdossier worden geraadpleegd.

Indien uit de documenten de architect bekend is, is dit als aparte zoekmogelijkheid opgenomen. U dient er overigens wel rekening mee te houden dat hergebruikt van ontwerpen onder de Auteurswet vallen, dat nog levende architecten aanspraak op vergoeding kunnen maken, evenals nabestaanden tot 70 jaar na het overlijden van betrokken architect.

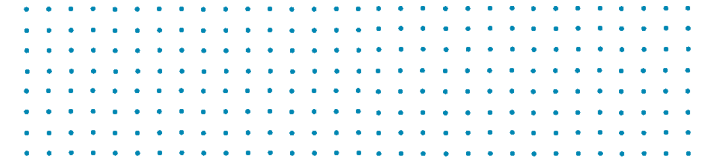
Aantal dossiers (2965 titles)
Op de hoogte blijven van wijzigingen

1062 10831 1122 113 1144 1158
117 1208 1230 1266 1353 1373
13744 13845 1393 1413 1423 1461
1476 17143 201 253 264 285 3157
416 44 456 481 595 626 715 7290
730 735 810 812 864 9198 95

Procesverbaal van oprichting

Zoeken op adres ...

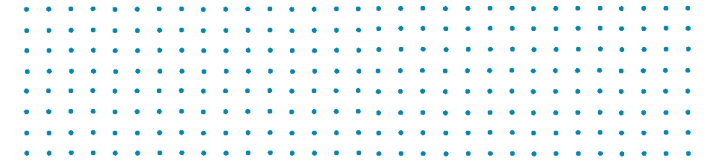
Zoek op CR nummer ...



Kitodo.Presentation

Einige Fakten

- Holländische/englische/deutsche Benutzeroberfläche
- Spezielle Suchfelder z.B. CR-Nummer (Flurstücknummer)
- Login-Funktionalität für Benutzer geplant
- Eine Migration auf MSSQL für Produktion und Präsentation war geplant, wurde aber wegen zu großer Komplexität und nicht gewährleisteter Lauffähigkeit verworfen.



Vielen Dank für Ihre Aufmerksamkeit!

Schlussfolgerungen:

- Die Migration auf Production 3 ist ein steiniger Weg, aber möglich.
- Je gründlicher Sie testen, desto reibungsloser verläuft die Produktivmigration.
- Die Community und kommerzielle Dienstleister stehen für Fragen und Unterstützung zur Verfügung.