

# Geo-Metadaten in OpenData Portalen

- Ausgangslage
- Allgemeines zur ckan Software
- Synchronisierung GeoPortal.rlp ↔ ckan
- Ergebnis
- Lessons learned
- Diskussion

## Ausgangslage (1)

- Oktober 2012: Politische Entscheidung ein eigenes rheinland-pfälzisches OpenData Portal zu erstellen
- Ziel für Freischaltung CeBIT 2013
- Auftrag an Fa. aus Mainz vergeben
- **Prinzip:** Anpassung der Software ckan 1.8.1b an Bedürfnisse in Rheinland-Pfalz (Layout, Schnittstellen, Editor mit speziellen Pflichtfeldern)

Ende Januar war Prototyp fertiggestellt

## Ausgangslage (2)

### Probleme:

- Nur wenige Daten - ausschließlich aus dem Bereich des Innenressorts - vorhanden
- GDI-RP war nicht in Konzeptionsphase integriert

**Erste Idee:** *Automatische* Integration von Datensätzen aus der GDI-RP über *ckan Harvest Plugin*

- FOSS – Metadaten Portal Software der Open Knowledge Foundation



- Postgres Datenbank als Backend

PostgreSQL



- Einfache json basierte API

- Apache Solr Search Engine integriert



- Viele plugins verfügbar (harvest, spatial, ...)

- Weit verbreitet, EU, GB, ...

- Einfaches, erweiterbares Meta-Datenmodell

- In python programmiert (nach MVC Prinzip)  python™

# Synchronisierung (1)

## Erster Ansatz:

- Lokale Installation eines ckan 1.8.1b
- Installation der spatial Erweiterung (Postgis, ckan plugin)
- Testen des Harvesting Mechanismus über die CSW Schnittstelle (zunächst lokaler Test mit geonetwork Instanz)

## Synchronisierung (2)

### Aufgetretene Probleme:

- CSW Harvesting basiert auf einfachen Daten-Metadatenätzen – Daten-Service Kopplung ist **work in progress** (Großbritannien)
- Harvesting Server cached Informationen in einer Warteschlange – gibt oft Probleme
- Lizenzinformationen nicht sauber abbildbar
- Harvesting Mechanismus etwas kompliziert zu automatisieren – mehrere laufende Prozesse
- WMS Viewer plugin (OL Basis) fehlerhaft implementiert

## Synchronisierung (3)

Nach ein bis zwei Tagen des Testens - Entschluss folgenden Weg einzuschlagen:



- **Aktives Fortführen** des ckan von Rheinland-Pfalz über dessen API – 'Push-Verfahren'

### Vorbereitung:

- Erweiterung des GeoPortal.rlp Datenmodells um **isopen** Eigenschaft bei den vordefinierten Lizenzbedingungen + **Information der Anbieter**
- Script zur Abgabe der Informationen aus dem GeoPortal.rlp im ckan json Format
- Script zur Synchronisierung einer ckan Instanz mit der ckan json Exportliste des GeoPortal.rlp über ckan API
- Anlegen einer **Gruppe GDI-RP** zur späteren Identifizierung der Datensätze





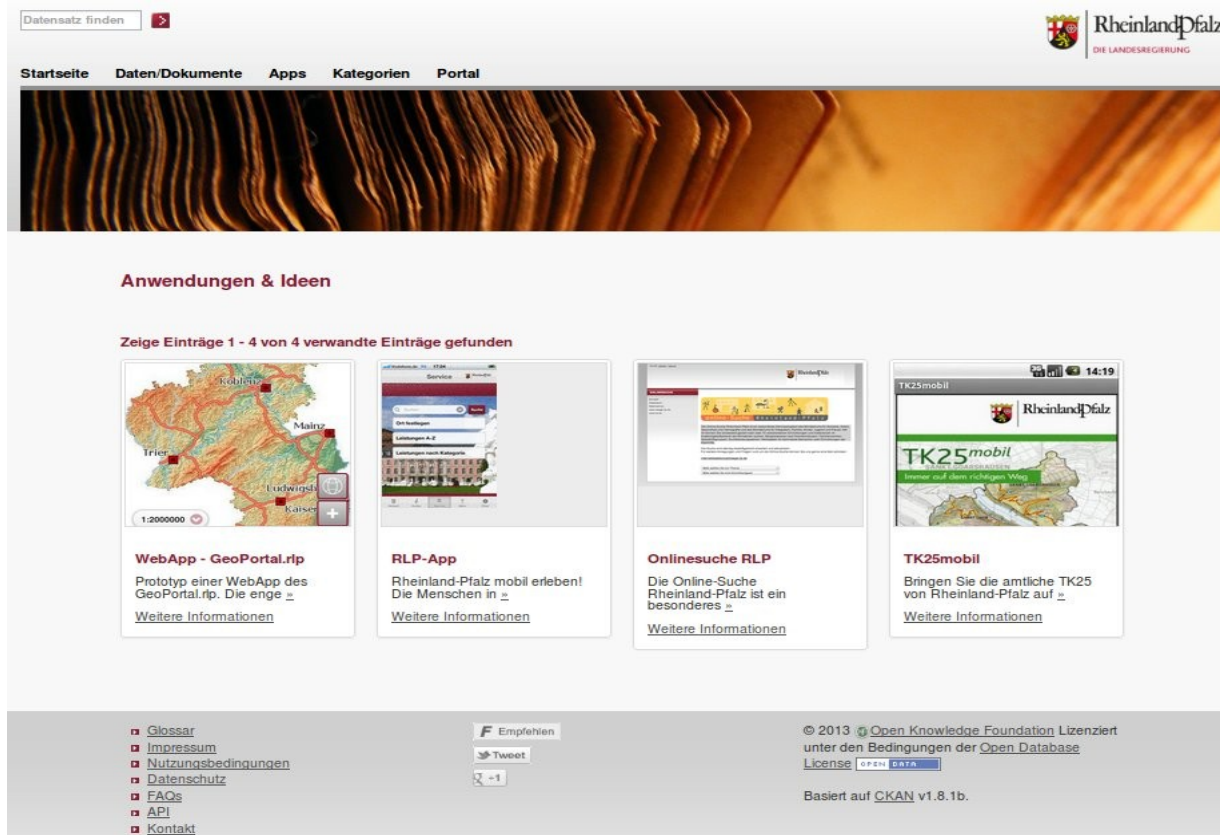
# Synchronisierung (4)

## Praktische Vorführung

- Lokale ckan Instanz:
  - <http://localhost:5000>
- GeoPortal Metadateneditor:
  - [http://localhost/mapbender\\_trunk/frames/index.php?gui\\_id=admin\\_wms\\_metadata](http://localhost/mapbender_trunk/frames/index.php?gui_id=admin_wms_metadata)
- Aufruf des sync Prozesses:
  - [http://localhost/mapbender\\_trunk/php/mod\\_syncCkan.php](http://localhost/mapbender_trunk/php/mod_syncCkan.php)

# Ergebnis (1)

- Bei Freischaltung auf der CeBIT stammten 670 der insgesamt 1365 Datensätze aus der GDI-RP
- Auf der Pressekonferenz wurde das GeoPortal.rlp gezeigt:  
<http://www.youtube.com/watch?v=rY97ydJyoxo>



The screenshot shows the GeoPortal.rlp website interface. At the top, there is a search bar labeled 'Datensatz finden' and the Rheinland-Pfalz logo. Below the search bar is a navigation menu with 'Startseite', 'Daten/Dokumente', 'Apps', 'Kategorien', and 'Portal'. A large banner image of a book spine is displayed. The main content area is titled 'Anwendungen & Ideen' and shows 'Zeige Einträge 1 - 4 von 4 verwandte Einträge gefunden'. There are four featured entries:

- WebApp - GeoPortal.rlp**: Prototyp einer WebApp des GeoPortal.rlp. Die enge ... [Weitere Informationen](#)
- RLP-App**: Rheinland-Pfalz mobil erleben! Die Menschen in ... [Weitere Informationen](#)
- Onlinesuche RLP**: Die Online-Suche Rheinland-Pfalz ist ein besonderes ... [Weitere Informationen](#)
- TK25mobil**: Bringen Sie die amtliche TK25 von Rheinland-Pfalz auf ... [Weitere Informationen](#)

At the bottom left, there is a sidebar with links: [Glossar](#), [Impressum](#), [Nutzungsbedingungen](#), [Datenschutz](#), [FAQs](#), [API](#), and [Kontakt](#). In the center, there are social media buttons for 'Empfehlen', 'Tweet', and '+1'. At the bottom right, there is a copyright notice: '© 2013 Open Knowledge Foundation Lizenziert unter den Bedingungen der Open Database License' and 'Basiert auf CKAN v1.8.1b.'

# Ergebnis (2)

- Einzelner Bebauungsplan im Open-Government-Data-Portal:

**Rockeskyll - Ebene: Sauerwasen - Auf der Nauerwies**

Bebauungspläne der Ortsgemeinde Rockeskyll Bebauungsplan "Sauerwasen - Auf der Nauerwies" der Ortsgemeinde Rockeskyll

**Zusätzliche Informationen**

Feld	Wert
Pflegebeauftragter	<a href="#">Verbandsgemeinde Gerolstein</a>
Lizenz	<a href="#">Datenlizenz Deutschland Namensnennung nicht-kommerziell</a>
Typ des Inhalts	Kartenebene
Ist ohne Entgelt nutzbar	Nein
Kontaktinformation Pflegebeauftragter	Ferdinand-Sauerbruch-Str. 15 56073 Koblenz
Ansprechpartner	Verbandsgemeinde Gerolstein
Ansprechpartner E-Mail	rathaus@gerolstein.de
Kontaktinformation Ansprechpartner	Kylweg 1 54568 Gerolstein 06591/ 13-0
Zeitraum von	2012-06-15 00:00:00
Zeitraum bis	2012-09-06 00:00:00

**Tags / Stichworte**

[Bauleitplan](#) [Bebauungsplan](#) [Bolan](#)  
[Rockeskyll](#)

**Kategorien**

[GDI-RP](#)

**Ressourcen / Datenquellen**

**Anzeige im GeoPortal**

Anzeige im GeoPortal.rlp  
URL: [http://www.geoportal.rlp.de:80/portal/karten.html?LAYER\[zoom\]=1&LAYER\[id\]=35037](http://www.geoportal.rlp.de:80/portal/karten.html?LAYER[zoom]=1&LAYER[id]=35037)  
Format: Kartenviewer

**WMS Capabilities Link zur Integration in GIS oder Webapplika...**

WMS Capabilities Link zur Integration in GIS oder Webapplikationen  
URL: [http://www.geoportal.rlp.de/mapbender/php/wms.php?layer\\_id=35037&REQUEST=GetCapabilities&VERSION=1.1.1&SERVICE=WMS](http://www.geoportal.rlp.de/mapbender/php/wms.php?layer_id=35037&REQUEST=GetCapabilities&VERSION=1.1.1&SERVICE=WMS)  
Format: WMS

**Metadaten zur WMS Kartenebene**

Metadaten zur WMS Kartenebene  
URL: [http://www.geoportal.rlp.de/mapbender/php/mod\\_showMetadata.php?languageCode=de&resource=layer&layout=tabs&id=35037](http://www.geoportal.rlp.de/mapbender/php/mod_showMetadata.php?languageCode=de&resource=layer&layout=tabs&id=35037)  
Format: HTML

Aufruf des GeoPortal.rlp

WMS Capabilities Link

HTML Metadatenanzeige

## Lessons learned (1)

- Wichtige Dinge:



(1) ckan verwendet eigene UUIDs für die interne Verwaltung – man braucht eigene persistente Bezeichner (e.g. auch UUIDs) als s.g. **name** des ckan packages ( ≈ Datensatz)



(2) ckan hat eine Vollhistorie – vor dem Löschen der Datensätze (state = deleted) *muss* der **name** abgeändert werden, sonst kann man diesen Datensatz später nicht mehr anlegen

(3) Die Updates aktualisieren direkt den Suchindex (solr) – nicht besonders schnell aber dafür speicherintensiv ;-)

(4) Derzeit sind im ckan nur Lizenzen auf Datensebene definierbar (bei den Geo-Metadaten auch auf Service Ebene möglich)

## Lessons learned (1)

- Pitfalls:



(1) ckan verwendet eigene UUIDs für die interne Verwaltung – man braucht eigene persistente Bezeichner (e.g. auch UUIDs) als s.g. **name** des ckan packages ( ≈ Datensatz)



(2) ckan hat eine Vollhistorie – vor dem Löschen der Datensätze (state = deleted) *muss* der **name** abgeändert werden, sonst kann man diesen Datensatz später nicht mehr anlegen



(3) Die Updates aktualisieren direkt den Suchindex (solr) – nicht besonders schnell aber dafür speicherintensiv



(4) Derzeit sind im ckan nur Lizenzen auf Datensatzebene definierbar (bei den Geo-Metadaten auch auf Service Ebene möglich)

## Lessons learned (2)



(5) Man sollte auf keinen Fall anfangen das interne ckan Datenmodell so abzuändern, dass eigene Validierungen durchgeführt werden (gilt auch allgemein bei für Verwendung von OSS Projekten). Leider in RP passiert – daher müssen verschiedene APIs zur Synchronisierung verwendet werden. Ein ckan2ckan Harvesting wird ggf. unmöglich.



(6) Bei Implementierung des SOLR ist auf ausreichende Speicherzuteilung unter Tomcat achten. In RP kommt es derzeit beim Update größerer Mengen von Datensätzen zu Überläufen – und damit Ausfällen des Systems

## Lessons learned (3)

- Grundsätzliches Problem:



Mehrfachnennungen von realen Datensätzen durch verschiedene Schnittstellen (z.B. HTML Editor / GDI) – wie kann man Identitäten sichern – Beispiel Soziale Einrichtungen RLP

- Mögliche Lösung:

Vorsehen einer OpenData Katalog UUID im Geometadatenatz

Weitere Frage:

Sind Apps wirklich immer nur statische Anwendungen – Beispiel WebApp as a Service! - Mobile Portal RLP



## Lessons learned (3)

### Vorteile Push gegenüber Pull:

- Aktualität höher
- Kann ggf. auch getriggert werden – **onchange** im Ursprungskatalog

### Im Hinblick auf das Metadatenhandling innerhalb der GDI-DE / INSPIRE zu überlegen:

- Übergang von Pull zu Push Mechanismen! Dadurch erhält man hochaktuelle Informationen, die für den Betrieb einer richtigen GDI eigentlich unabdingbar sind! Alternativ: Handling eines zusätzlichen Katalogeintrag-Zeitstempels und hochfrequente Überprüfung der angeschlossenen Kataloge.
- Einheitliche Abbildung von Lizenzinformationen – ggf. Anpassung der ISO19115.



# Diskussion

Vielen Dank für die Aufmerksamkeit

Armin Retterath  
Zentrale Stelle GDI-RP

0261/492-466  
[armin.retterath@lvermgeo.rlp.de](mailto:armin.retterath@lvermgeo.rlp.de)