

CKAN Workshop BKA



ckan

**CKAN, Erleichterung bei der Umsetzung von Open
Government Data Portalen, 04.10. 2011**
Dachfoyer des Bundeskanzleramtes am Minoritenplatz 1

Martin Kaltenböck, CMC

Semantic Web Company, OKFO & OGD Austria



These slides are published under :

<http://creativecommons.org/licenses/by/3.0>

CKAN Workshop BKA: Agenda

CKAN Workshop, BKA, 04.10. 2011, 09.00 -17.30 Uhr

09.00 - 09.30	Willkommen & Vorstellung der Cooperation OGD Österreich (BKA) Agenda und Ziele des CKAN Workshops (SWC)
09:30 - 10.00	Vorstellungsrunde der TeilnehmerInnen (Person, Organisation, Rolle, Zugang zu OGD & CKAN, Erwartungen)
10.00 - 11.00	Open Government Data - Grundlagen und Entwicklungen
11.00 - 11.15	<i>Pause</i>
11.15 - 12.30	CKAN Teil 1: Datenkataloge und CKAN - CKAN Grundlagen
12.30 - 13.30	<i>Mittagspause</i>
13.30 - 15.00	CKAN Teil 2: CKAN Details, Gruppenarbeit: OGD Portalfeatures
15.00 - 16.00	OGD und Meta Daten Management
16.00 - 16.15	<i>Pause</i>
16.15 - 17.00	Linked Government Data - Einführung & Grundlagen
17.00 - 17.30	Feedback Runde (SWC) und Verabschiedung (BKA)

CKAN Teil 2: Details zu CKAN

- **Details zu CKAN Features**
 - Erstellen und Editieren von Datenpaketen
 - Managing Data Packages in CKAN
 - Suchfunktionen in CKAN
 - CKAN & Atom Feeds
 - Die CKAN API
 - CKAN und RDF
- **CKAN für System Administratoren**
 - Extensions zu CKAN
 - Workflows
 - Möglichkeiten für Geodaten / -informationen
 - Administration (Guide)
- **CKAN: die LOD Cloud, Ausblick: CKAN im LOD2 Projekt & publicdata.eu**
- **Wichtige & hilfreiche Ressourcen zu CKAN**

Managing Data Packages in CKAN

<http://wiki.ckan.org/Managing-Data-Packages>

- CKAN erlaubt das **Registrieren UND den Upload** von Datensätzen
- **Umgang mit SPAM:** delete Funktion oder Tag Meldung: meta.spam
- **CKAN Bookmarklet** (<http://wiki.ckan.org/Managing-Data-Packages>)
- **Delete** (entfernt nur Sichtbarkeit) & **Purge Funktion** (entfernt Daten)
(Admin Bereichckan-admin/trash)
- **Entfernen von Revisions** (Purge a revision)
- **Umgang mit Duplikaten:** delete Status oder Tag Meldung: meta.duplicate
- **Tag Konventionen**
 - Tag Familien: zB meta.duplicate

Suchfunktionen in CKAN

CKAN Instanzen können auf 3 verschiedene Wege durchsucht werden:

- Via Web Interface
- Via API
- Via SPARQL endpoint <http://semantic.ckan.net/> (nur für RDF Daten möglich)

Suche via Web Interface - was wird durchsucht

- name
- title
- notes
- tags (i.e. searching for packages with a tag that matches the search term)

Suchparameter - wie kann gesucht werden

- title:chemistry
- tags:chemistry
- chemistry tags:ocean

Anmerkung: Solr/Lucene Erweiterung für optimierte Suchfunktion verfügbar

CKAN & Atom Feeds

CKAN bietet out of the box Atom Feeds (ähnlich wie RSS) für Alert Mechanismen an wie folgt:

- Individual package feeds - jedes Daten Packages bietet einen Atom Feed
- Main feed - bietet Änderungen aller Daten Packages einer Instanz an

Folgende Details werden im Atom Feed angeboten

- **updated / published:** Date & time of the change, given in ISO format.
- **author:** The author's login name or IP address.
- **link:** Link to Revision in the web interface.
- **summary:** Log message for the revision.

Die CKAN API

To use a the CKAN API, your application will make an HTTP request and parse the response. By default, the response format is JSON. Most of the API is RESTful and your methods will be the standard HTTP methods like GET, PUT, POST and DELETE.

Folgende API Endpoints werden angeboten:

- [Domain Model/Package](#)
- [Domain Model/Group](#)
- [Domain Model/Revision](#)

Es existieren API Client Libraries für: Javascript Client, Python Client (offiziell unterstützt) , Command-line Client, PHP Client

Detailinformationen für Developer

- <http://wiki.ckan.org/API>
- <http://docs.ckan.org/en/latest/api.html>
- http://wiki.ckan.org/Using_the_API

CKAN und RDF

- Die Metadaten von verschiedenen CKAN Instanzen werden in regelmässigen Abständen in RDF konvertiert (Resource Description Framework) und auf <http://semantic.ckan.net/> veröffentlicht und verfügbar gemacht.
- Weiters werden RDF Daten, welche in CKAN verfügbar sind via SPARQL Endpoint zur Abfrage angeboten
- Ein Python Script für die Konvertierung von Meta Daten nach RDF ist verfügbar

Siehe auch:

- http://wiki.ckan.org/RDF_and_CKAN
- <http://semantic.ckan.net/>

Extensions zu CKAN

CKAN bildet ein Plugin System für die Entwicklung von Extensions an

Folgende Extensions existieren derzeit (siehe: http://wiki.ckan.org/List_of_Extensions)

- **Disqus Comments** - Kommentarfunktion auf Data Packages
- **Data Preview** - Vorschaufunktion auf Resources (Datensichten)
- **Webhooks service notification** - Autom. Update Nachricht für Drittsysteme
- **Moderated Edits** - Workflow für moderiertes Editieren
- **Apache Solr search backend** - Solr Suchfunktion inkl. Volltext, Facetten und Statistiken
- **Queue support & worker interface** - Nachrichten Erweiterung für nachgelagerte Datenarbeiten
- **OGC OWS Support** - Unterstützung von OWS Services (OGC Standards) wie WMS und CSW
- **CKAN RDF** - RDF Serialisierung für Daten packages (Meta Daten)
- **Wordpresser** - Wordpress als CMS für CKAN
- **Storage** - Integriertes Speichersystem für Daten für CKAN
- **Todo Extension** - Issue und Todo Management für CKAN
- **Metadata Harvester** - Framework zum Harvesting von Metadaten anderer System in CKAN
- **Google Analytics Extension**

Für Developer

- Form API Tester
- Weitere - aber nicht getestete bzw nicht angenommene Extensions verfügbar

CKAN Workflows - Überblick

CKAN unterstützt verschiedene Workflow Modelle für das Management von Datensätzen (Metadaten) wie folgt

- **Wiki Workflow**
JEDER kann Inhalte bearbeiten
- **Moderated Workflow**
2 Rollen: Sysadmins / Moderators & Editors
Hierbei können beide Rollen editieren - aber die Änderungen müssen durch den Sysadmin / Moderator freigegeben werden
Anmerkung: benötigt die 'Moderated Edits Extension'
- **Aproved Publisher Model**
spezieller Workflow welcher für die International Aid Transparency Initiative entwickelt wurde - Code ist frei verfügbar: <https://bitbucket.org/okfn/ckanextiati> - es ist aber offen, ob dies in Version 1.5 unterstützt wird!
- **Sysadmin Only Workflow**
erlaubt das Editieren ausschließlich durch Sysadmins

CKAN Teil 2: CKAN Details und Ausblick

CKAN: Möglichkeiten für Geodaten / -informationen

Folgende Konventionen für das Ablegen von Geoinformationen / -daten werden in CKAN üblicherweise verwendet:

- spatial-text: Textual representation of the extent / location of the package
- spatial: [GeoJSON](#) representation of the extent of the package (Polygon or Point)
- spatial-uri: Linked Data URI representing the place name

Beispiele:

- spatial-text: United Kingdom
- spatial: { "type": "Polygon", "coordinates": [[[0.50, 49.74],[0.5, 59.25], [-6.88, 59.25], [-6.88, 49.74], [0.50, 49.74]]] }
- spatial-uri: <http://www.geonames.org/2635167>

oder

- spatial-text: Matsushima
- spatial: { "type": "Point", "coordinates": [38.36, 141.07] }
- spatial-uri: <http://www.geonames.org/2111964>

Bei der Verwendung dieser Konventionen können die vorhandenen Geo-Extensions in Folge verwendet werden, wie folgt

- **Geo Indexing:** spatial_query plugin of the [ckanext-spatial](#) - erweitert die CKAN DB für Geodaten
- **Bounding Box Queries für den API Search:** spatial_query plugin
- **Dataset Extent Map:** spatial_query & dataset_extent_map - ermöglicht Kartenansicht in Detailsicht des Datensatzes
- **WMS Previewing:** wms_preview plugin - Vorschaufunktion für WMS Layer in der Resource - Sicht (Beta)

Details siehe: http://wiki.ckan.org/Geospatial_Capabilities

CKAN Administration (Guide)

Der CKAN Administration Guide bietet folgende Informationen

- Installationsanleitungen (Package oder Source Installation)
- Post Installations Setup Anleitung
- Customisation Anleitung
- Load Datasets - Anleitungen für Datenimport (API oder Harvester)
- Anleitung für 'Paster Script' als Verwaltungsmöglichkeit
- Anleitung zur Rechteverwaltung
- Anleitung für die Verwendung von Extensions
- Anleitung für das Erstellen von Extensions
- Anleitung für das Arbeiten mit Formularen in CKAN
- Anleitung für DB Dumps
- Upgrade Informationen
- Information zur Sprachverwaltung in CKAN
- Referenzinformationen zu: Konfiguration, API, Testing, etc.

Siehe: <http://docs.ckan.org/en/latest/index.html>

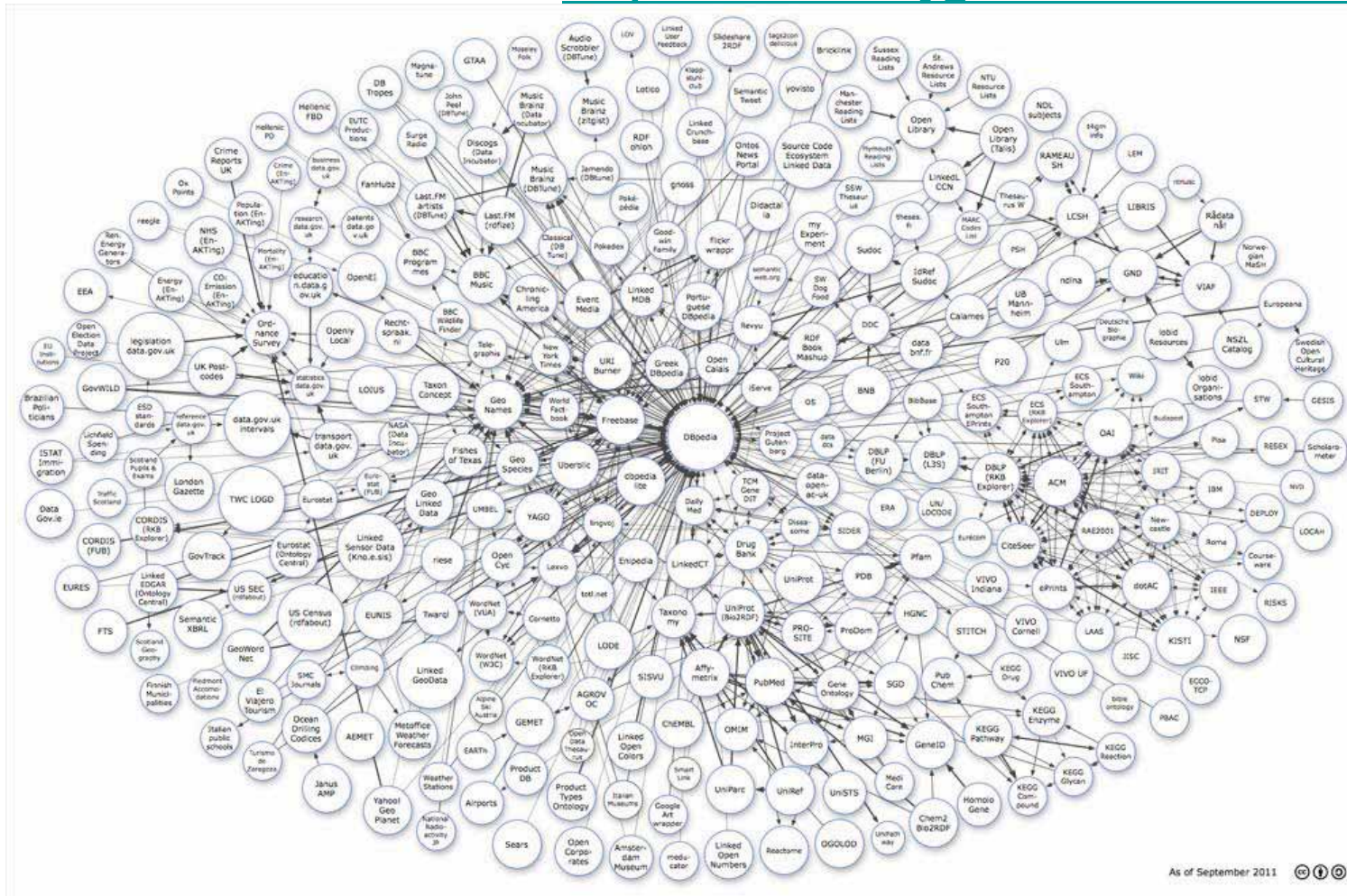
CKAN- Technische Grundlagen und Informationen

For developers working on CKAN core - technical background to the CKAN software.

- [Design and Architecture](#)
- [Domain Model](#)
- [User Stories](#)
- [Load Testing](#) - notes on how we load tested CKAN
- [Purpose](#)
- [Data Formats](#) - review of existing data exchange and querying formats
- [Related Software](#) - software related to CKAN (e.g. useful tools, similar software etc)

CKAN Teil 2: CKAN Details und Ausblick

CKAN & die LOD Cloud: <http://richard.cyganiak.de/2007/10/lo/>



CKAN & die LOD Cloud: Grundlagen zur Generierung des LOD Cloud Diagramms

Welche Daten werden in die Cloud integriert?

- Data items are accessible via de-referencable URIs
- The dataset sets at least 50 RDF links pointing at other datasets or at least one other dataset is setting 50 RDF links pointing at your dataset.

Wie kann ich meine Linked Data Sets in das Diagramm einspielen?

- Please register with CKAN before editing or adding any packages.
- Please confirm that your data set does not already exist on CKAN before adding a new data set.
- Add or edit your data set and describe it with the following minimum required information:
 - CKAN name (a unique id)
 - title
 - URL
 - number of triples
 - links to other data sets.
- Please tag newly added data sets with *lod*.
- If you are not aware of any in- or outlinks, tag it with *lodcloud.nolinks*.
- Please provide as much additional information as possible (e.g. SPARQL endpoint, void description, license, and the topic of the data set).

Weiterführend müssen folgende Konventionen erfüllt sein:

<http://www.w3.org/wiki/TaskForces/CommunityProjects/LinkingOpenData/DataSets/CKANmetainformation>

Ausblick: CKAN & LOD2 - publicdata.eu

Das LOD2 Projekt - Creating Knowledge out of Interlinked Data - <http://lod2.eu>

publicdata.eu - find and reuse datasets from local, regional and national public bodies across Europe from a single place



Konkrete Ziele & Arbeiten:

- Harmonisierung der Meta Daten (Strukturen) der EU Datenkataloge
- Bereitstellung von Werkzeugen für die Konvertierung von Daten zu RDF und LOD
- Unterstützung der Publikation von LOGD von Daten der öffentlichen Verwaltung in Europa
- Umsetzung von publicdata.eu

Wichtige & hilfreiche Ressourcen zu CKAN

- Offizielle Produktwebsite: <http://ckan.org/>
- CKAN WIKI: http://wiki.ckan.org/Main_Page
- CKAN Mailing Liste: <http://lists.okfn.org/mailman/listinfo/ckan-discuss>
- CKAN Weblog: <http://ckan.org/blog/>
- CKAN Partner: in Österreich zB Semantic Web Company - <http://www.semantic-web.at>
- CKAN Workshop am OGD Camp 2011 in Warschau (in Planung)

CKAN Ressourcen für Developer

- [Submitting a code patch](#)
- [Becoming a CKAN Developer](#) - how to get the code, how to contribute (coding standards etc)
 - [Source Code Repository \(bitbucket\)](#)
- [Making changes to model code](#) - What to be aware of
- [Development Overview](#) - Roadmaps etc
- [User Stories Spreadsheet](#)
- [CKAN developer trac](#) - Tickets
- [Release process](#)
- How to deploy: <http://wiki.ckan.org/Deployment>

Die CKAN Community - wie kann ich beitragen...

Get involved in CKAN development

- [Meetups](#)
- [Project Steering Committee](#)
- [People](#)
- [List of all CKAN instances](#)
- [Ideas for CKAN screencasts](#)
- Related components:
 - [Webstore](#) maintains tabular data
 - [Helmut](#) the reconciliation server
 - [datapatterns.org](#), a data wrangling handbook

CKAN Teil 2: Gruppenarbeit OGD Portalfeatures

- Bitte bilden Sie 2 Arbeitsgruppen
- Fragestellung: Welche Funktionen ausser den 'Basisdaten' in einem Katalog soll ein OGD Portal (lokal - regional - national) Ihrer Ansicht nach aufweisen / anbieten?
- 15-20 Minuten Zeit
- Danach 2 x 5 Minuten Präsentation
- Die Ergebnisse werden der Cooperation OGD Österreich aufbereitet (und anonymisiert) übergeben.

Infos & Kontakt



Martin Kaltenböck, CMC
Semantic Web Company GmbH (SWC)

Mariahilferstrasse 80
1070 Wien



Semantic Web Company (SWC)

Web

<http://www.semantic-web.at>

<http://blog.semantic-web.at>

m.kaltenboeck@semantic-web.at

Phone: +43-1-402 12 35–25

LOD2

Creating Knowledge out
of Interlinked Data

Web

<http://www.lod2.eu>

Blog

<http://blog.lod2.eu>

OKFO

OGD Austria

Web

<http://gov.opendata.at>

OGD Camp 2011 Warsaw

<http://www.ogdcamp.org>

