

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 WS BKA: Vorstellungsrunde

- Wer bin ich & was ist meine Rolle in der Organisation?
- Wie ist mein Zugang zum Thema Open Government Data bzw. Linked Open Data (LOD)?
- Kenne ich CKAN? Wenn ja, wie gut kenne ich CKAN?
- Was sind meine Erwartungen an diesen Workshop?
- Optional: Was sind meine Erwartungen für den Einsatz von (Linked) Open Data bzw. CKAN in meiner Organisation?

CKAN WS BKA: ‚Workshop Infos & Regeln‘

- Fragen sind immer & jederzeit willkommen!
- Dieser Workshop ist ein Basis WS zu OGD & CKAN
Technische Detailfragen et al. werden selbstverständlich gerne aufgenommen und in Nachbereitung geklärt
- Den Fragen- & Ideenpool bitte aktiv nützen & einfordern
- Die Workshop Unterlagen werden digital für alle TeilnehmerInnen zur Verfügung gestellt.

CKAN Teil 1: Datenkataloge und CKAN

- Was ist ein Datenkatalog?
- Wofür ein Datenkatalog? Einsatzgebiete von Datenkatalogen?
- Datenkataloge weltweit: ein kurzer Marktüberblick
- CKAN - the data hub - Definition & Historie
- Einsatz von CKAN im Bereich Open Data weltweit, in Europa und in Österreich

Was ist ein Datenkatalog?

Wikipedia

http://de.wikipedia.org/wiki/Data_Dictionary

Ein **Data Dictionary** - in deutscher Übersetzung auch **Datenwörterbuch**, **Datenkatalog** oder etwas unscharf *Datenverzeichnis* genannt - ist ein Katalog von Metadaten, der die Definitionen und Darstellungsregeln von Datenelementen enthält.

Zusatz:

Es beschreibt alle Anwendungsdaten eines Unternehmens und die Beziehungen zwischen den verschiedenen Datenobjekten mit dem Ziel einer redundanzfreien und zentralen Definition der Datenstrukturen und entspricht somit einer Darstellung des anwendungsspezifischen Datenmodells.

Wofür ein Datenkatalog - Einsatzgebiete & OGD

Im Bereich Open Government Data werden Datenkataloge verwendet um:

- A) Metadaten zu Daten / Datensätzen zu verwalten - die Daten selbst werden hierbei NICHT im Datenkatalog gespeichert / gehalten, sondern es existieren nur Verweise / Links auf die Daten (Datenquellen) selbst.
- B) Metadaten UND Daten / Datensätze zu verwalten - die Metadaten UND die Daten selbst werden im Datenkatalog gehalten (möglicherweise existieren zusätzlich Verweise / Links auf die Daten (Datenquellen) selbst .
- C) Mischformen: Metadaten UND/ODER Daten / Datensätze zu verwalten. Hierbei kann seitens der UserInnen entschieden werden, ob nur Metadaten gehalten werden und/oder auch die Daten / datensätze selbst.

Eigenschaften von Datenkatalogen

- **Auffindbarkeit:** Datensätze sollten für Entwickler und Endkunden auffindbar sein
- **Standardschnittstellen:** Datensätze sollten via web-basierte Standard-APIs ansprechbar sein
- **Quellenübergreifende Standardformate:** Datensätze sollten in Standard-Formaten verfügbar sein, im Idealfall auf syntaktischer *und* semantischer Ebene
- **Verknüpfbarkeit:** Daten sollten referenzierbar, möglichst eindeutig und damit quellenübergreifend verknüpfbar sein
- **Community Features:** Datenmngt (Einpflege, Wartung, Löschen), weite Mechanismen zur Qualitätssicherung, Rating, Kommentarfunktionen - Tagging Funktionen - oftmals in Form von Crowd Sourcing Mechanismen.

Unterschied / Abgrenzung: Datenkatalog - Datenportal

Datenportale (Open Data Portale) bieten als Kern einen Datenkatalog entsprechend der vorangegangenen Definition (Verwaltung von Metadaten und/oder Daten / Datensätze) PLUS eine Vielzahl zusätzlicher Angebote und Funktionalitäten.

Folgende Merkmale & Funktionen wurden in der LOD2 OGD Stakeholder Survey 2010/2011 EU-weit erhoben (siehe: <http://survey.lod2.eu>)

- Raw Data (Basisdaten)
- Suchen, Browsen, Gruppieren und Clustering von Daten
- Kommentarfunktionen, Fehlerreporting, Tagging
- Analyse- & Visualisierungs Werkzeuge für Daten
- Alert Mechanismen (regionale Daten) & Crowd Sourcing Mechanismen (zB data repair)
- Support Center, wie Chat, Forum, Ansprechpartner, Call Center etc
- UND: Apps, News, Events, Wettbewerbe, Best Practise, Success Stories ...

CKAN Teil 1: Datenkataloge & CKAN - CKAN Grundlagen

Datenkataloge / -portale weltweit: kurzer Marktüberblick

Name	Kurzbeschreibung	Ownership	Meta daten & Date n	Visualisier- ungs Tools	API	5 Sterne Modell
Socrata	They target strongly on providing a toolset for Government Institutions (US).	Commercial	Ja	Ja	Socrata Open Data API (SODA)	Bis 4
CKAN	A mainly community driven portal for collecting, rating and grouping data (UK & EU).	NGO	Ja	Nein	RESTful	Bis 5
Factual	Factual lays a focus on get the given data in a structure. So they have thematical and spacial groups. Customers: media industry like Newsweek, and target Bloggers and Apps-Developers (US).	Commercial	Ja	Geo (ja)	RESTful requests	k.a.
Windows Azure	Clientbased Cloud Solutiun, with extra add-ons for Open Government Data (US, EU).	Commercial	Ja	Nein	Storage Services API Blob Service API Queue Service API Table Service API	Bis 4
Google Fusion Tables	Comming from Google Spreadsheet, GFT offers visualisation and collaboration tools. The data storage is a kind of add-on, to run those services (US).	Commercial	Ja	Ja	Fusion Tables API	Bis 3
Tableau Public	Client based toolset for visualisation. Stores data and visualistion online. Targets on media and blogging, with providing Widgets (US).	Commercial	Ja	Ja	Keine API	1
Linking Open Government Data Project	Run by the Tetherless World Constellation . We are translating government-related datasets into RDF, linking them to the Web of Data and providing demos and tutorials on mashing up and consuming linked government data.	R & D	Ja	Nein	SPARQL	Bis 5
Fluid Ops Information Workbench	Lots of features on top of a data catalogue for BI, data integration, visualisation, analysation, additional search etc.	Commercial	Ja	Ja	API für datenimporta	k.a.

CKAN - the data hub - Definition & Historie

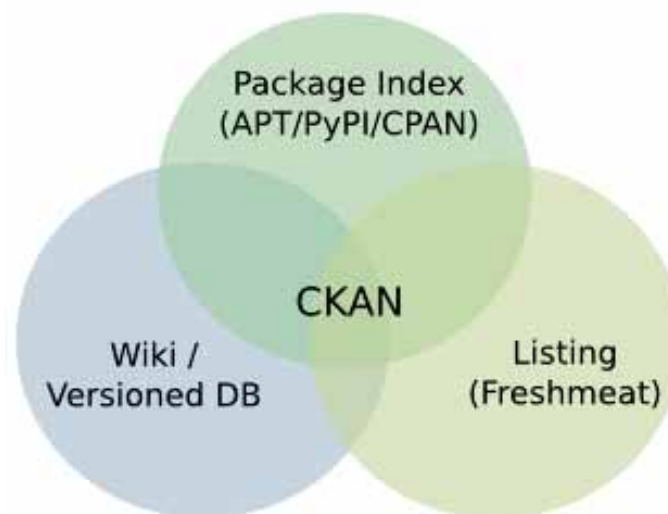
CKAN - the Open Source Data Hub Software

Publish, share and find and manage data



ckan

- CKAN is the Comprehensive Knowledge Archive Network, a **registry** of open knowledge datasets and projects (and a few closed ones).
- CKAN makes it easy to **find, share and reuse open content and data**, especially in ways that are machine automatable.
- As a system CKAN functions as a synthesis of several different services:



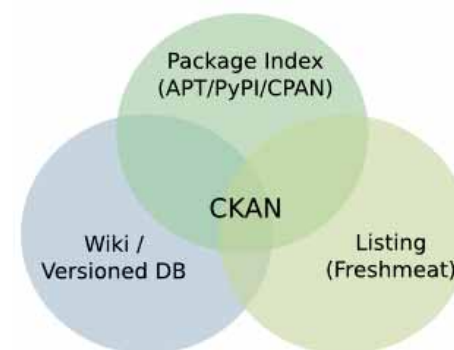
CKAN - the data hub - Definition & Historie

CKAN - the Open Source Data Hub Software
Publish, share and find and manage data



CKAN hat sich stark an vorhandenen Systemen der Verwaltung von Open Source Software orientiert!

Wie im Diagramm ersichtlich kombiniert CKAN die Funktionen: listing/registry, dataset index & wiki.



Listing / registry:

basiert auf der Idee von Open Source Registries wie z.B.: freshmeat (<http://freshmeat.net/>)



Wiki Funktionen: die Basis bildet das integrierten 'Version Domain Modell' welches aus der Python Library / Welt stammt; somit Versionierung und Wiki-ähnliches Interface

Dataset Index: Speichern von strukturierten (Meta) Daten und seit kurzem auch Daten (in progress) und damit Funktionen wie Suchen - Browsen etc...

CKAN - the data hub - Definition & Historie

CKAN - the Open Source Data Hub Software

Publish, share and find and manage data



ckan

Versioned Domain Model (vdm)

Versioned Domain Model is python library which allows you to ‘version’ your domain model in the same way that source code version control systems such as subversion allow you version your code. It is designed for use with SQLAlchemy.

- Code home: <http://pypi.python.org/pypi/vdm>
- Documentation: <http://packages.python.org/vdm/>
- Source: <http://bitbucket.org/okfn/vdm/>
- Mailing list: <http://lists.okfn.org/mailman/listinfo/okfn-help>

CKAN - the data hub - Definition & Historie

CKAN - the Open Source Data Hub Software
Publish, share and find and manage data



Die OKFN - Promoting Open Knowledge in a Digital Age



Founded in 2004, we're a not-for-profit organization promoting open knowledge: any kind of data and content - sonnets to statistics, genes to geodata - **that can be freely used, reused, and redistributed.**

We promote open knowledge because of its potential to deliver far-reaching societal benefits.

We **create standards** like the Open Definition, **organize events** like OKCon and Open Government Data Camp, **run projects** like Where Does My Money Go and Open Shakespeare, and **develop tools** like CKAN and the Annotator **to help people create, find and share open material.**

CKAN - the data hub - Definition & Historie

CKAN - the Open Source Data Hub Software
Publish, share and find and manage data



The Open Definition defines openness in relation to data and content.

<http://opendefinition.org/>

OPEN KNOWLEDGE

OPEN DATA

OPEN CONTENT

“A piece of content or data is open if anyone is free to use, reuse, and redistribute it – subject only, at most, to the requirement to attribute and share-alike.”.

In addition: Open Software Service Definition (OSSD) which defines ‘openness’ in relation to online (software) services. It can be summed up in the statement that “A service is open if its source code is Free/Open Source Software and non-personal data is open as in the Open Knowledge Definition (OKD).”.

Einsatz von CKAN im Bereich Open Data: Weltweit, in Europa und in Österreich



Best Practise Beispiele für den Einsatz von CKAN als Datenkatalog / (Teil eines) Open Data Portals:

- UK Open Data Portal: <http://data.gov.uk/>
- Helsinki Region: <http://data.hri.fi/>
- Greater Manchester councils: <http://datagm.org.uk/>
- International Aid Transparency Initiative: <http://iatiregistry.org/>
- Norwegisches Open Data Portal: <http://data.norge.no/>
- Niederländisches Open data Portal: <http://register.data.overheid.nl/>
- Belgisches Open data Portal: <http://openbelgium.be/>
- Berliner Open data Portal: <http://daten.berlin.de/>
- Linzer Open Data Portal: <http://data.linz.gv.at>

Liste aller CKAN Instanzen: <http://wiki.ckan.org/Instances>

Einsatz von CKAN im Bereich Open Data



Andrew Stott, UK Government Director of Transparency & Digital Engagement, responsible for launch of data.gov.uk: “Using Open Source was the best decision we ever made. A big thumbs up to the Open Knowledge Foundation and CKAN.”

Seit Juli 2011:  datacatalogs.org

- 141 gelistete Datenkataloge weltweit
- von führenden Open Data ExpertInnen gepflegt (weltweit)
- Wichtiges Instrument für zukünftige Zusammenarbeit bzw für die Interoperabilität von verschiedenen Datenkatalogen
- Personen dieser Gruppe arbeiten auch am DCAT Vokabular - einem Austausch Vokabular für Datenkataloge!

Siehe: <http://vocab.deri.ie/dcat-overview> und <http://vocab.deri.ie/dcat>

Und: http://www.w3.org/egov/wiki/Data_Catalog_Vocabulary

CKAN Teil 1: Grundlagen zu CKAN

- **Basisbegriffe / Begriffsklärung und FAQ zu CKAN**
- **Feature Tour zu CKAN**
 - Core Features
 - Additional Features
- **Technische Spezifikationen & Requirements**
- **Community Features und Mechanismen**
- **CKAN und die Integration in Drittsysteme (Bspl: WordPress, Drupal)**

Basisbegriffe / Begriffsklärung & FAQ zu CKAN

- **Resource** = a single file/API/etc (name taken from URL = Uniform Resource Locator!)
- **Package** = a collection of Resources that you would want to keep together = (roughly) a Dataset
- **Packages exist because** it is usual that a 'dataset' actually consists of multiple files or APIs, e.g. because
- The **dataset is split** into lots of files (e.g. statistics by month)
- There are **multiple versions of the same resource** e.g. single file and an API or the 'same' data in different formats (e.g. RDF and CSV).

Basisbegriffe / Begriffsklärung & FAQ zu CKAN

Ausblick

In the near future we intend to **replace the single concept of package** with three related concepts: **Dataset, Distribution and Collection**.

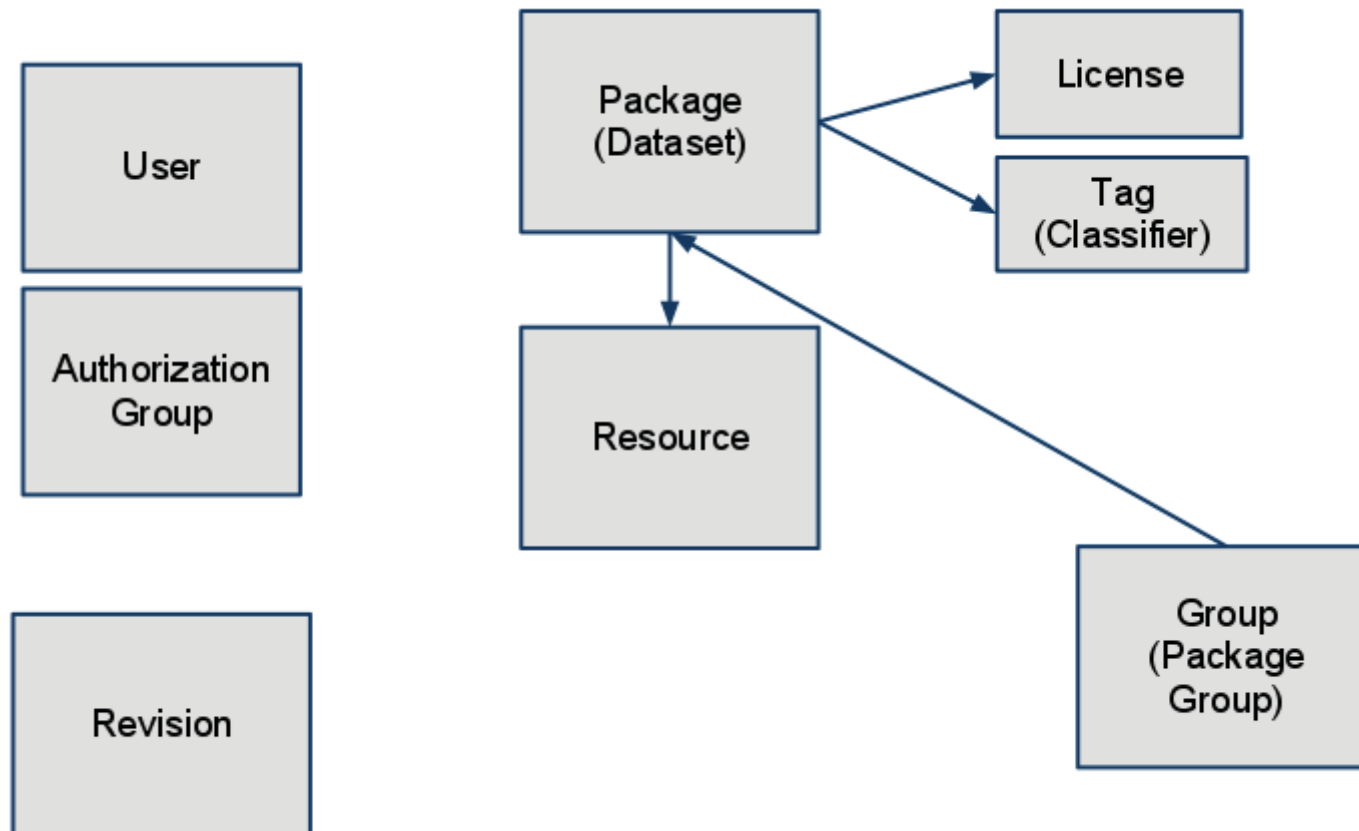
- A "**Dataset**" represents information about a dataset such as its **name, license, description, publisher** etc
- A "**Distribution**" of the dataset is the way the dataset is actually published, for example as a downloadable file, an API you can query etc. A dataset will often have more than one distribution because people often like to publish the same data in lots of ways.
- A "**Collection**" is a set of datasets that have some sort of relationship to one another and that are worth considering together.
For example if I have a dataset released each month representing a government department's spending I might group them together into a "timeseries" collection. If lots of departments release the same data I could group all spending for March into a "March Spending" collection. By creating a collection you save someone having to manually search for each dataset they are interested in.
- **At the moment** packages serve each of the three above roles so you should feel free to use them in any way you like to publish the data you want so that others can find it on CKAN.

Basisbegriffe / Begriffsklärung & FAQ zu CKAN

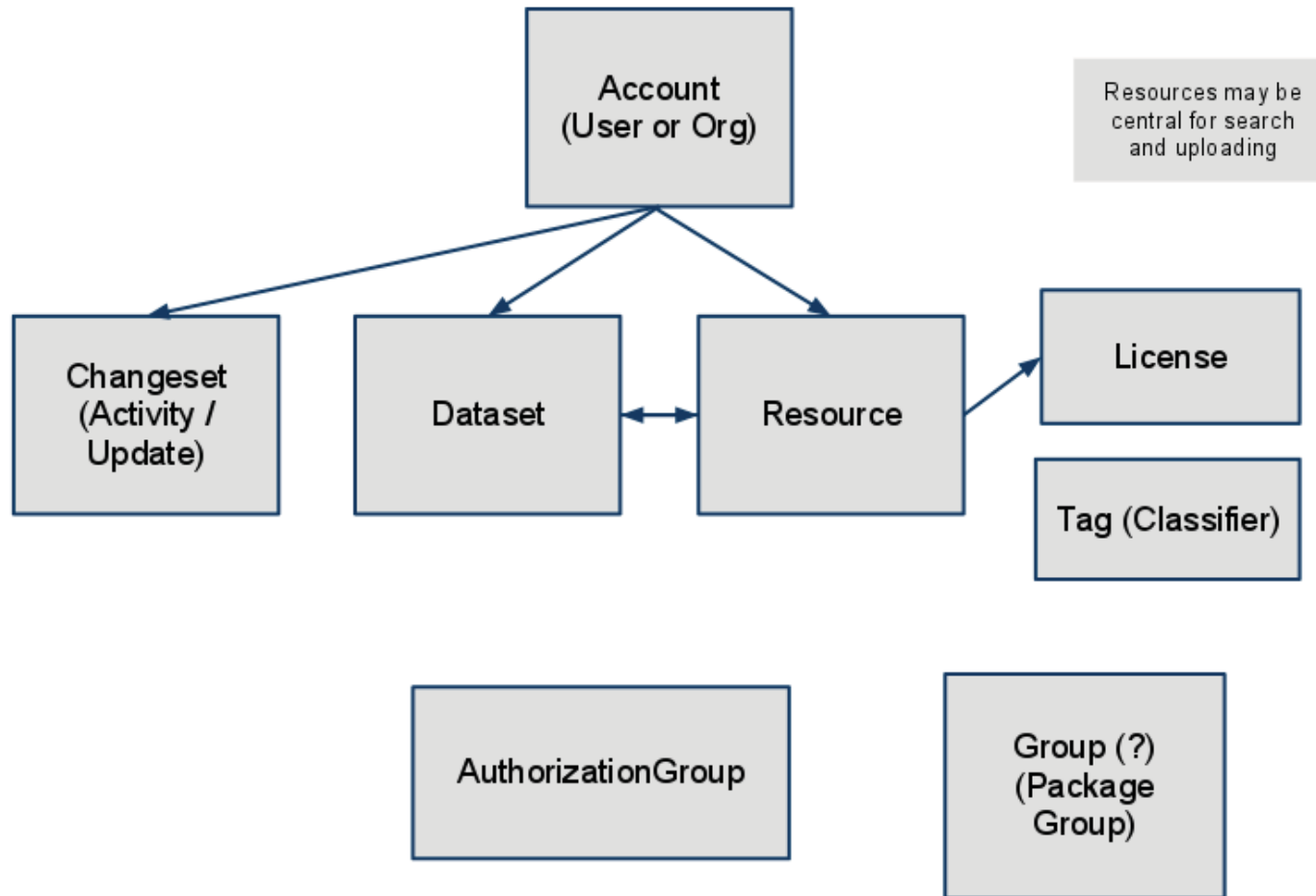
Interoperabilität von CKAN mit anderen Datenkatalogen bzw Datensystemen

- **Yes!** CKAN has a **well designed and robust harvesting interface** that exposes three simple API calls so that third party catalog vendors can easily add support for their catalog.
- In addition to its own API, CKAN already has **harvesting interfaces that support standards such as CSW/INSPIRE for interoperability** with other geospatial cataloges such as GeoNetwork. CKAN is used for the UK portal for INSPIRE compliance as part of a federated network. CKAN can both import and export INSPIRE and GEMINI 2.1 records.
- More broadly it has **support for importing data from a number of catalogues** and sites across Europe.

Aufbau: CKAN Domain Modell



Derzeit in Diskussion - Aufbau: CKAN Domain Modell V2



Aufbau: Domain Modell: Package

Übersicht

- Basic Information
- Resources
- Groups
- Details
- Extras

Plus

- History
- Authorization
- Subscribe (RSS)
- View bzw. Edit Ansicht

The screenshot shows a web browser window displaying a CKAN dataset page. The title is '2000 U.S. Census in RDF (rdfabout.com)'. The page includes a navigation bar with 'View', 'Edit', and 'History' buttons, and a 'Subscribe' button. The main content area contains a description of the dataset, a list of resources, and additional information. The right sidebar features a 'First time at the Data Hub?' message, a 'Source' field, an 'Author' field, a 'Maintainer' field, a 'Tags' list, and a 'Groups' list.

2000 U.S. Census in RDF (rdfabout.com)

View Edit History Subscribe

2000 U.S. Census converted into over a billion RDF triples.

Population statistics at various geographic levels, from the U.S. as a whole, down through states, counties, sub-counties (roughly, cities and incorporated towns)

Names: also found in the of SPARQL Endpoints.

From home page:

- For the detailed Census statistics, you'll have to download the raw Census data files from the Census Bureau, my Perl script and the patch file below and run it yourself because the files are too big for me to offer as a download!
- The data and scripts can be reused under Creative Commons Attribution-NonCommercial-ShareAlike

Resources

URL	Name/Description	Format
http://www.rdfabout.com/demolencensus/	Download	
http://www.rdfabout.com/sitemap.xml	XML Sitemap	metasitemap
http://www.rdfabout.com/spargl	SPARQL endpoint	api/spargl
http://www.rdfabout.com/rdf/us/gov/pe/us/mr	Example (RDF/XML)	examplerdf.xml

Additional Information

Field	Value
namespace	http://www.rdfabout.com/rdf/us/gov/pe/
shortname	US Census (rdfabout)
triples	1002848918

First time at the Data Hub?
The Data Hub is a catalogue for data
Click here to find out more ...

Source
<http://www.rdfabout.com/demolencensus/>

Author
Joshua Tauberer

Maintainer
Maintainer not given

Tags

- census
- data
- demographics
- deref-vocab
- format-dc
- format-geonames
- format-politico
- format-rdf
- geographic
- linkeddata
- lod
- lodcloud.coltrix
- no-licence-metadata
- no-provenance-metadata
- no-vocab-mappings
- population
- published-by-third-party
- rdf
- statistics
- us

add tags +

Groups

- LOD Cloud
- Linking Open Data

Aufbau: Domain Modell: Package

- Basic Information
 - Title
 - Slug
 - Home Page
 - Description
 - License
 - Tags

The screenshot shows the CKAN 'Edit' page for a dataset titled '2000 U.S. Census in RDF (rdfabout.com)'. The page is divided into several sections:

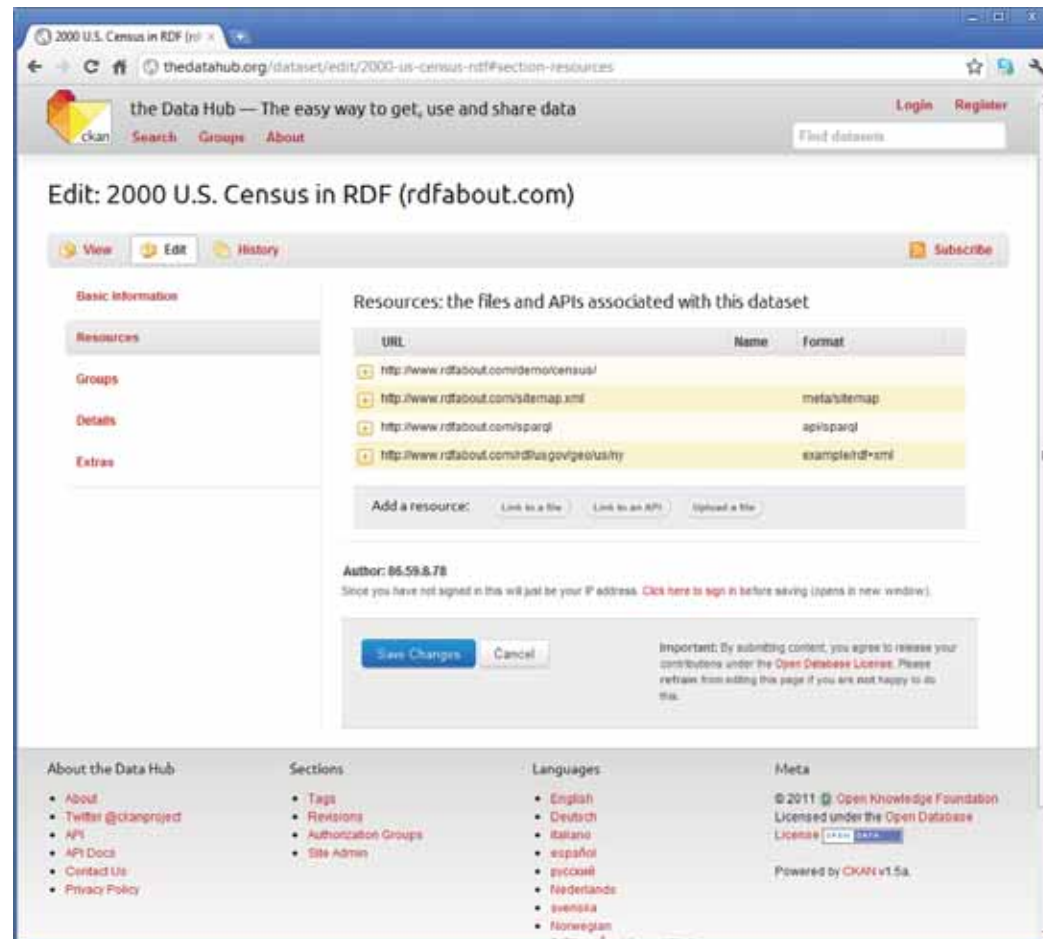
- Title:** 2000 U.S. Census in RDF (rdfabout.com)
- Slug:** 2000-us-census-rdf
- Home Page:** http://www.rdfabout.com/demo/census/
- Description:** 2000 U.S. Census converted into over a billion RDF triples. Population statistics at various geographic levels, from the U.S. as a whole, down through states, counties, sub-counties (roughly, cities and incorporated towns). Notes: also found in the of SPARQL Endpoints. From home page: > * For the detailed Census statistics, you'll have to download the raw Census data files from the Census Bureau, my Per... You can use [Markdown formatting](#) here.
- License:** Non-ORF Compliant Creative Commons Non-Commercial
- Tags:** census data demographics deref-vocab format-dc format-geonames format-politico fo...

Aufbau: Domain Modell: Resource

- Resource (multiple)
 - Name
 - Description
 - URL
 - Format
 - Resource Type
 - Size (Bytes)
 - Mimetype
 - Last Modified
 - Mimetype (inner)
 - Hash
 - ID (autom.)

Siehe dazu Details:

http://wiki.ckan.org/Domain_Model/Resource



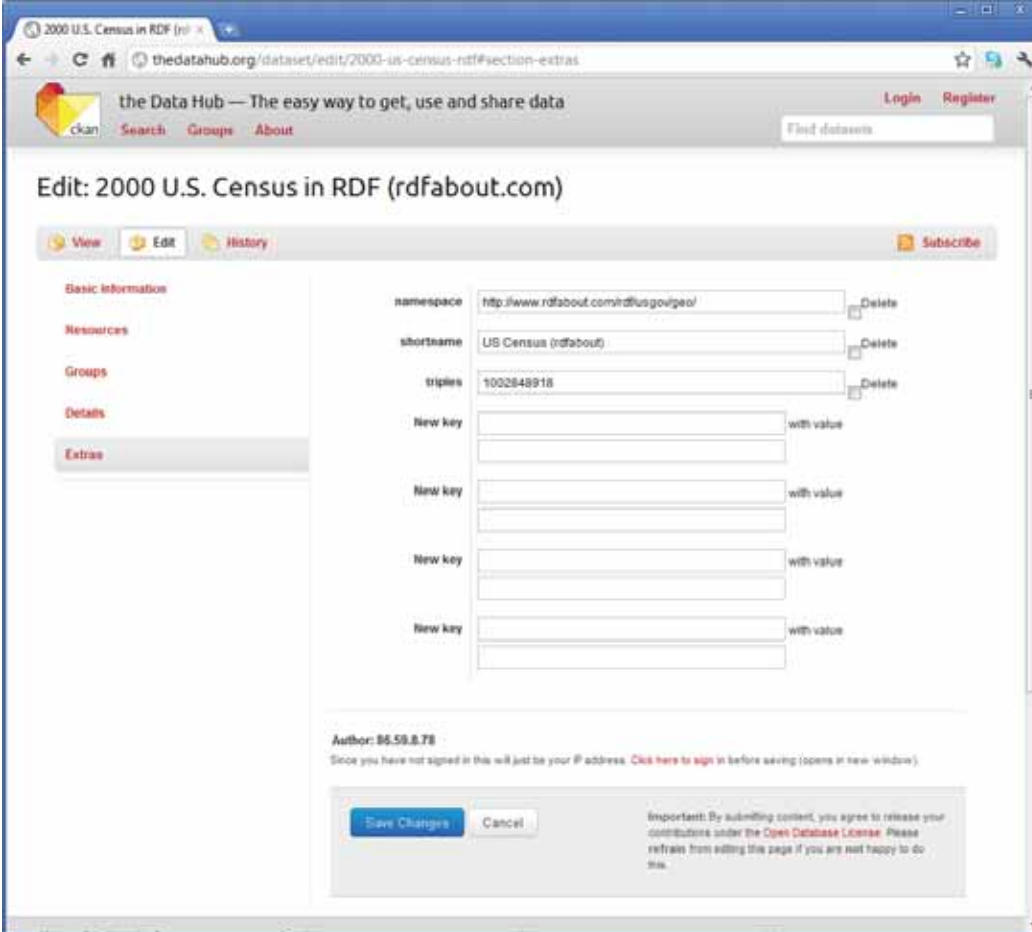
The screenshot shows the CKAN interface for editing a dataset. The page title is "Edit: 2000 U.S. Census in RDF (rdfabout.com)". The main content area is titled "Resources: the files and APIs associated with this dataset". It contains a table with columns "URL", "Name", and "Format". The table lists three resources:

URL	Name	Format
http://www.rdfabout.com/demo/census/		
http://www.rdfabout.com/sitemap.xml		meta/sitemap
http://www.rdfabout.com/sparql		api/sparql
http://www.rdfabout.com/rdf/usgov/geo/us/hy		example/rdf+xml

Below the table, there are buttons for "Add a resource:" with options "Link to a file", "Link to an API", and "Upload a file". At the bottom of the main content area, there is a "Save Changes" button and a "Cancel" button. A warning message states: "Important: By submitting content, you agree to release your contributions under the Open Database License. Please refrain from adding this page if you are not happy to do this."

Aufbau: Domain Modell: Groups & Extra

- **Groups (multiple)**
 - Erlaubt das Gruppieren von Packages
- **Extras**
 - Erlaubt das Anlegen von n zusätzlichen Meta Daten Attributen



The screenshot shows the CKAN 'Edit' page for a dataset titled '2000 U.S. Census in RDF (rdfabout.com)'. The page is divided into several sections: 'Basic Information', 'Resources', 'Groups', 'Details', and 'Extras'. The 'Extras' section is currently active, displaying a list of existing extras and a form to add new ones. The existing extras include 'namespace' (http://www.rdfabout.com/rdf/usgov/geo/), 'shortname' (US Census (rdfabout)), and 'triples' (1002848918). Below these, there are four 'New key' entries, each with a text input field and a 'with value' label. At the bottom of the page, there is a 'Save Changes' button and a 'Cancel' button. A warning message states: 'Author: 86.58.8.78 Since you have not signed in this will just be your IP address. [Click here to sign in](#) before saving (opens in new window). Important: By submitting content, you agree to release your contributions under the [Open Database License](#). Please refrain from editing this page if you are not happy to do this.'

Aufbau: Domain Modell: Details

- Details
 - Author
 - Author Email
 - Maintainer
 - Maintainer Email
 - Version

The screenshot shows a web browser window displaying the CKAN 'Edit' page for a dataset titled '2000 U.S. Census in RDF (rdfabout.com)'. The page is titled 'the Data Hub — The easy way to get, use and share data'. The 'Details' section is active, showing the following fields:

- Author:** Joshua Tauberer
- Author email:** http://razor.occams.info/
- Maintainer:** (empty field)
- Maintainer email:** (empty field)
- Version:** (empty field)

Below the fields, there is a warning: 'Author: 06:59:78. Since you have not signed in this will just be your IP address. Click here to sign in before saving (opens in new window)'. At the bottom of the form, there are 'Save Changes' and 'Cancel' buttons, and a disclaimer: 'Important: By submitting content, you agree to release your contributions under the Open Database License. Please refrain from editing this page if you are not happy to do this.'

Aufbau: CKAN: History

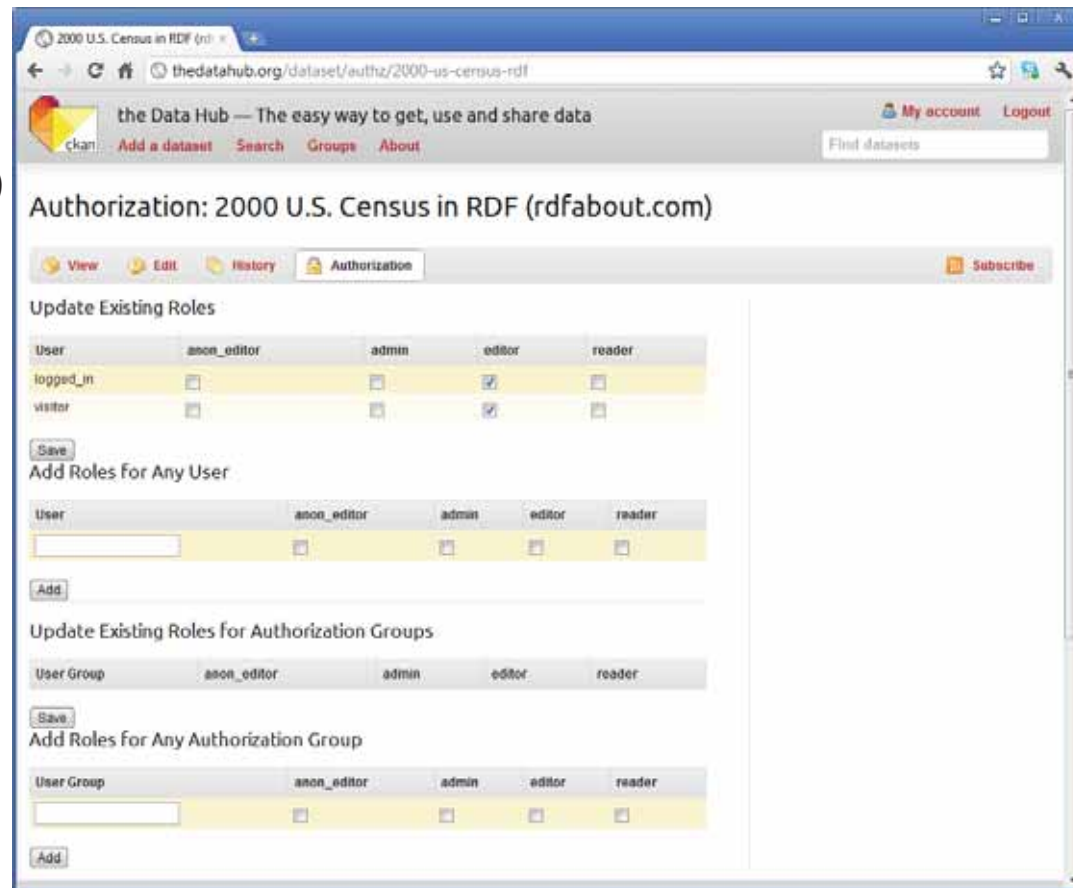
- Ansicht von Revisions & Vergleiche von Versionen möglich

The screenshot shows the CKAN interface for the dataset '2000 U.S. Census in RDF'. The page title is 'History: 2000 U.S. Census in RDF (rdfabout.com)'. Below the title, there are tabs for 'View', 'Edit', and 'History', and a 'Subscribe' button. The main content is a table of revisions with the following columns: Revision, Timestamp, Author, and Log Message.

Revision	Timestamp	Author	Log Message
886d...	2011-08-10 12:52	Arija Jertzsich	REST API: Update object 2000-us-census-rdf
852d...	2011-07-18 23:07	Arija Jertzsich	REST API: Update object 2000-us-census-rdf
883b...	2010-10-05 23:03	https://me.yahoo.com/viaDggC1e1TvdapVpZvF1Zn8g7q0-#bae73	REST API: Update object 2000-us-census-rdf
e445...	2010-09-27 17:53	https://me.yahoo.com/viaDggC1e1TvdapVpZvF1Zn8g7q0-#bae73	REST API: Update object 2000-us-census-rdf
8ba...	2010-09-27 17:44	https://me.yahoo.com/viaDggC1e1TvdapVpZvF1Zn8g7q0-#bae73	REST API: Update object 2000-us-census-rdf
95c...	2010-09-27 17:26	https://me.yahoo.com/viaDggC1e1TvdapVpZvF1Zn8g7q0-#bae73	REST API: Update object 2000-us-census-rdf
fcc...	2010-09-24 11:55	10.227.3.101	8ba, shorthand
02de...	2010-09-23 17:24	https://me.yahoo.com/viaDggC1e1TvdapVpZvF1Zn8g7q0-#bae73	added self-descriptiveness metadata
a51a...	2010-09-23 17:14	https://me.yahoo.com/viaDggC1e1TvdapVpZvF1Zn8g7q0-#bae73	added self-descriptiveness metadata

Aufbau: CKAN: Authorization

- Verwaltung von Berechtigungen für Daten Packages (nur für Admin möglich)



The screenshot shows the CKAN interface for managing permissions on a dataset. The browser address bar shows the URL: `thedatahub.org/dataset/authz/2000-us-census-rdf`. The page title is "Authorization: 2000 U.S. Census in RDF (rdfabout.com)".

The interface includes a navigation bar with "View", "Edit", "History", and "Authorization" tabs. Below this, there are three main sections for managing roles:

- Update Existing Roles:** A table with columns for User, anon_editor, admin, editor, and reader. The "editor" and "reader" columns have checkboxes checked for the "logged_in" and "visitor" users.
- Add Roles For Any User:** A form with a text input field and checkboxes for the same roles.
- Update Existing Roles for Authorization Groups:** A table with columns for User Group, anon_editor, admin, editor, and reader.
- Add Roles For Any Authorization Group:** A form with a text input field and checkboxes for the same roles.

Features Tour zu CKAN -Core Features (Übersicht)

- Core catalog based around Datasets and Resources (Files and APIs)
 - Tagging
 - Ratings
 - Arbitrary metadata
 - Dataset relationships
- Web user interface (WUI)
 - Dataset adding, editing, listing etc
 - Wiki features such as "Recent Changes", edit histories, purging of changes etc
 - User management and user home pages
- Full JSON-based REST API with clients in Python, PHP, Perl ...
 - RDF version also available
 - CKAN is easy to use as your "catalogue" backend
- An Extension and Plugin system
- Integrated Storage for storing data

Features Tour zu CKAN - Additional Features (Übersicht)

- Search: Full searchability (including full-text support) via API and WUI
- Access control: fine-grained access control for packages and group
- Run CKAN in different modes (Security Concept)
 - Wiki mode - anyone can contribute
 - Restricted mode - only designated users can edit
 - Publisher Mode - for organizations with multiple publisher groups
- INSPIRE support
- Federated structure: easily set up new instances with common search etc.
- Additional interfaces:
 - Excel importer (upload dataset/package information direct from a spreadsheet)
 - Fully featured command line client (datapkg)
- Visual Customisation possible

Technische Spezifikationen und Requirements

- CKAN (Aktuelle Version: 1.4.3) ist durchgängig in Python programmiert (Python: major releases 2.5 to 2.7)
- Läuft optimal auf Ubuntu 10.04 (oder auf Ubuntu in einer Virtualisierungsumgebung)
- CKAN verwendet PostgreSQL (PostgreSQL RDBMS ab Version 8.4)
- Mögliche Such-Erweiterung: Apache Solr

System Requirements

Software

- Ubuntu GNU/Linux 10.04 LTS (either 32 or 64bit for x86)
- enabled okfn-CKAN add-on Package Repository

Hardware

- works with 512MB of RAM, yet more is better
- up-to-date 64bit, Intel-compatible CPU

Details und Developer / Admin Guide: <http://readthedocs.org/docs/ckan/en/latest/>

Community Features & Mechanismen

CKAN ist grundsätzlich für den Einsatz im Bereich Crowd Sourcing entwickelt worden - dies bedeutet, es ist grundsätzlich ein sehr offenes System (es existiert aber ein klares Sicherheits und Berechtigungskonzept).

Folgende Ansätze werden von CKAN daher angeboten

- Klar definierter Core an Attributen für die Meta Daten
- Aber: wenige Pflichtfelder, um Barrieren möglichst niedrig zu halten
- Eintragen von Datensets auch ohne Registrierung / Login
- Kommentarfunktion auf Daten packages
- Gruppierung von Datensets zur Unterstützung einer besseren Zusammenarbeit an 'Daten Sammlungen'
- CKAN bietet 3 verschiedene Produktiv Modes an:
 - Wiki mode - anyone can contribute
 - Restricted mode - only designated users can edit
 - Publisher Mode - for organizations with multiple publisher groups

CKAN und die Integration in Drittsysteme

Starke Integration in Drupal und WordPress bereits realisiert, um rund um CKAN Datenportal mit zusätzlichen Inhalten und Funktionen anbieten zu können!

Drupal

Im Zuge der Entwicklung von data.gov.uk wurde ein Drupal - CKAN Modul entwickelt, welches Inhalte von CKAN nach Drupal synchronisiert, um Drupal in Folge als Frontend zu verwenden. Das Drupal Modulo ist frei verfügbar.

Siehe: <http://drupal.org/project/ckan>

Und: <http://data.gov.uk/blog/datagovuk-releases-open-source-code>

WordPress

Ermöglicht das Zusammenspiel von CKAN als Datenkatalog und von WordPress als CMS im Zusammenspiel - WordPress wird hierzu für zusätzliche (CMS) Informationen wie News, Blog oder Hilfe et al. verwendet.

Siehe: http://wiki.ckan.org/Extensions#Wordpress_-_use_Wordpress_as_a_CMS_for_CKAN

Und: <http://ckan.org/2011/04/04/wordpress-extension-released/>

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>

