

ckanapi and ckanext-scheming

CKANCon 2015

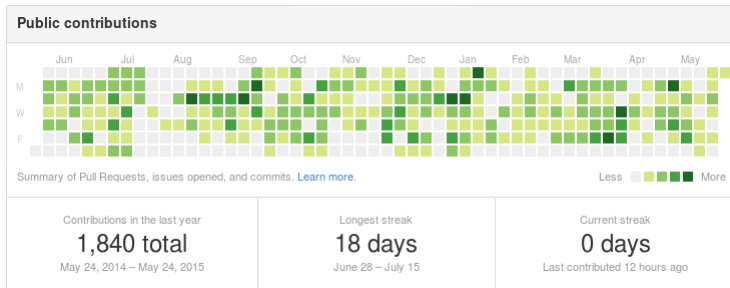
Ian Ward

open.canada.ca

www.datacats.com






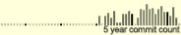














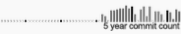


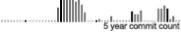





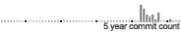
Python + Open Source developer

github.com/wardi




CKAN core

Top Contributors

Name	Kudos	12 Month Commits	All Time Commits	5 Year Trend	Primary Language	First Commit	Last Commit
 amercader		321	915		Python	over 3 years ago	5 days ago
 Ian Ward		188	284		Python	over 2 years ago	about 1 month ago
 Vitor Baptista		154	460		Python	over 2 years ago	8 months ago
 David Read		144	1026		Python	over 4 years ago	12 days ago
 Brook Elgie		101	101		Python	8 months ago	5 days ago
 Sean Hammond		93	1749		Python	over 3 years ago	3 months ago
 joetsoi		53	163		Python	about 2 years ago	10 days ago
 Ross Jones		48	574		Python	over 3 years ago	3 months ago
 Nigel Babu		18	235		Python	over 2 years ago	9 months ago
 nigelb		16	42		Python	about 1 year ago	4 months ago

open.canada.ca



Government of Canada
Gouvernement du Canada

Français

Search Canada.ca

Jobs

Immigration

Travel

Business

Benefits

Health







Taxes

More services

[Home](#) → Open Government

Open Government

Welcome to Canada's new Open Government portal. Here you can explore how the Government of Canada is working with the national and international open government community to create greater transparency and accountability, increase citizen engagement, and drive innovation and economic opportunities through Open Data, Open Information, and Open Dialogue.



Follow:      

Search

Open DataOpen InformationOpen Government

Search through our Open Data Portal.

Winners Showcase

  **CODE**
CANADIAN OPEN DATA EXPERIENCE

And the #CODE2015 winners are...

Item 1 of 4

Play

Latest News

- One Week until #IODC15
- "Career Path" App Wins! (CODE 2015)
- Minister Clement Starts the Clock for CODE 2015
- Canada to Host International Open Data Conference
- Minister Clement promotes CODE 2015
- Announcement of CODE 2015
- Wait time for priority procedures in Canada
- Government of Canada Employee Contact...

ckanext-fluent

github.com/open-data/ckanext-fluent

EN Books:

FR Books:

Field	Value
Books	<div><div>en</div><div>Franklin</div><div>fr</div><div>Benjamin</div></div>

```
"books": {"en": "Franklin", "fr": "Benjamin"}
```

ckanext-recombinant

github.com/open-data/ckanext-recombinant

GitHub

[Explore](#) [Features](#) [Enterprise](#) [Blog](#) [Sign up](#) [Sign in](#)

[open-data](#) / **ckanext-recombinant** Watch 2 Star 0 Fork 4

Create datastore tables for organizations and provide combined output

119 commits

1 branch

0 releases

4 contributors

branch: **master** **ckanext-recombinant** / +

Merge pull request #26 from wardl/master

wardl authored 17 days ago latest commit 34bde17d65

ckanext	treat blank numeric fields as 0s	17 days ago
.gitignore	add sample atl tables.json file	2 years ago
README.md	fix and forgive dirty numeric input; skip empty lines on reading .xls...	2 months ago
requirements.txt	remove old req	2 months ago
setup.py	proper namespace package for ckanext	5 months ago

README.md

<> Code

Issues 0

Pull requests 0

Pulse

Graphs

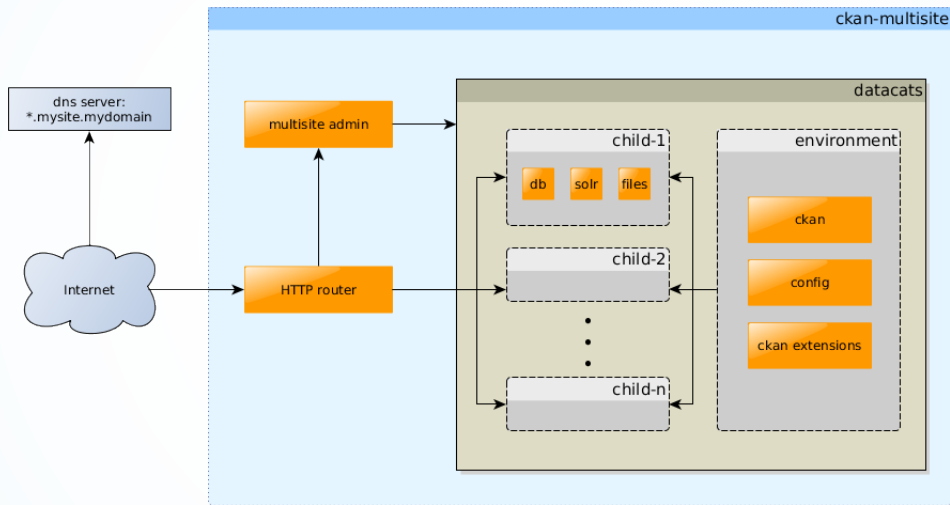
HTTPS clone URL

You can clone with **HTTPS** or **Subversion**.

Download ZIP

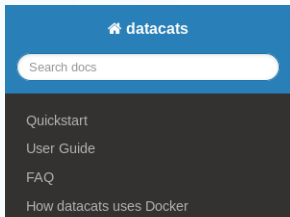
CKAN Multisite

github.com/boxkite/ckan-multisite



datacats

github.com/datacats/datacats



Command Reference

- help
- create
- deploy
- info
- init
- install
- list
- logs
- open
- paster
- pull
- purge
- reload
- shell

[Docs](#) » Command Reference

Command Reference

help

datacats command line interface

Usage:

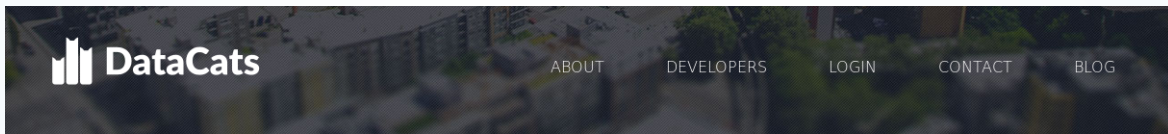
```
datacats COMMAND [ARGUMENTS...]  
datacats --help [COMMAND]  
datacats --version
```

The datacats commands available are:

create	Create a new environment
deploy	Deploy environment to production DataCats.com cloud service
info	Display information about environment and running containers
init	Initialize a purged environment or copied environment directory
install	Install or reinstall Python packages within this environment
list	List all environments for this user
logs	Display or follow container logs
open	Open web browser window to this environment
paster	Run a paster command from the current directory
pull	Download or update required datacats docker images
purge	Purge environment database and uploaded files
reload	Reload environment source and configuration
shell	Run a command or interactive shell within this environment
start	Create containers and start serving environment
stop	Stop serving environment and remove all its containers

Fork me on GitHub

www.datacats.com



ABOUT

DataCats provides Open Data software for cities. Our software helps governments increase transparency and accountability, better engage with citizens, and make meaningful decisions based on the data they collect.

BENEFITS



Create APIs around



Enable civic technology

ckanapi

github.com/ckan/ckanapi

A [command line interface](#) and [Python module](#) for accessing the [CKAN Action API](#)

`build passing` tested under Python 2.6, 2.7, 3.2, 3.3 and pypy

ckanapi CLI

The ckanapi command line interface lets you access local and remote CKAN instances for bulk operations and simple API actions.

Actions

Simple actions with string parameters may be called directly. The response is pretty-printed to STDOUT. e.g.:

```
$ ckanapi action group_list -r http://demo.ckan.org
[
  "data-explorer",
  "example-group",
  "geo-examples",
  ...
]
```

ckanext-scheming

github.com/open-data/ckanext-scheming

This extension provides a way to configure and share CKAN schemas using a JSON schema description. Custom template snippets for editing and display are also supported.

build passing coverage 100%

Requirements

This plugin is compatible with CKAN 2.3 or later.

Configuration

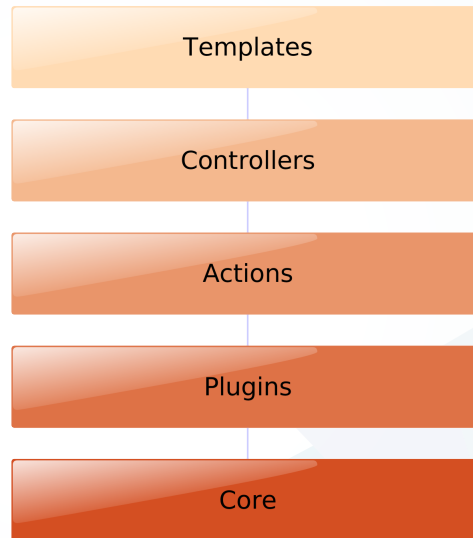
Set the schemas you want to use with configuration options:

```
ckan.plugins = scheming_datasets

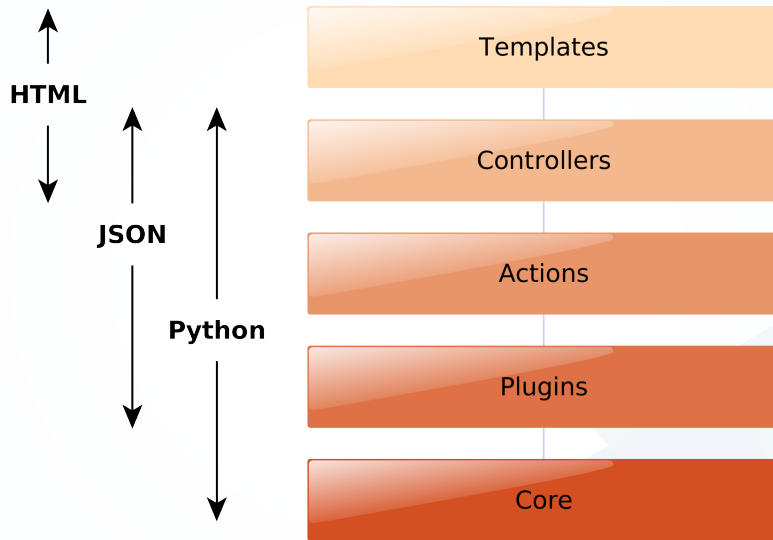
# module-path:file to schemas being used
scheming.dataset_schemas = ckanext.spatialx:spatialx_schema.json
                           ckanext.spatialx:spatialxy_schema.json
# will try to load "spatialx_schema.json" and "spatialxy_schema.json"
# as dataset schemas
#
# URL to schemas to be used can be specified as follows:
```

I ♥ CKAN

CKAN layers



Extending CKAN



ckanapi

github.com/ckan/ckanapi

1. Python library for calling ckan actions
2. Command-line client for calling ckan actions

Python library

Pass strings, lists, dicts, file objects like normal method calls

```
import ckanapi

ckan = ckanapi.RemoteCKAN('http://exampleckan', apikey=...)

with open('mydata.csv') as csv:
    ckan.action.package_create(name='best-dataset', title='Best evar')
    ckan.action.resource_create(package_id='best-dataset', upload=csv)
```


Safe

- Raises exceptions on errors, no manual error checking
- Clean context and data_dict for every call

Universal

- Python 2.6, 2.7, 3.2+ (source-compatible)
- Generic, no per-action custom code

Consistent interface

- LocalCKAN for core or extensions
- RemoteCKAN for remote API calls
- TestAppCKAN for tests

ckanapi Python library

```
pip install ckanapi
```

```
import ckanapi
```

github.com/ckan/ckanapi

[CKAN Action API Reference](#)

Command-line client

- Single and bulk actions
- Easy to pipeline
- Local and remote

Single actions

Pass strings as simple parameters

```
ckanapi action package_create name=best-dataset title="Best evar"  
ckanapi action package_list
```

Bulk actions

```
ckanapi dump datasets  
ckanapi dump organizations  
ckanapi dump groups
```

Bulk actions

```
ckanapi load datasets  
ckanapi load organizations  
ckanapi load groups
```


Bulk actions

```
ckanapi load datasets -p 3  
ckanapi load organizations -p 3  
ckanapi load groups -p 3
```

Pipelining

```
ckanapi dump datasets | ssh otherbox ckanapi load datasets -p 3
```

Local and remote

```
ckanapi dump datasets -r http://sourceckan | ckanapi load datasets -p 3
```

Bulk data format

- JSON Lines
- Consistent order with sorted keys

Track metadata in git

```
ckanapi dump datasets > datasets.jsonl  
git diff datasets.jsonl --stat
```

```
datasets.jsonl | 52 ++++++++++++++++++++++++++++++++++++++-----  
1 file changed, 36 insertions(+), 16 deletions(-)
```

Distributed load

```
split -n 1/3 datasets.jsonl part  
ckanapi load datasets -r http://web1 -a ... < partaa &  
ckanapi load datasets -r http://web2 -a ... < partab &  
ckanapi load datasets -r http://web3 -a ... < partac &
```

Summaries

```
head -5 datasets.jsonl | jq .title
```

```
"Exports of Dairy Products by Country of Destination Dairy Year 2001/02"
```

```
"Historical Border Wait Times"
```

```
"Number of Head Slaughtered in Federally Inspected Plants 2009 Hogs"
```

```
"Canadian Dairy Exports Month Calendar Year 2001 September"
```

```
"Federal Corporations"
```

Reports

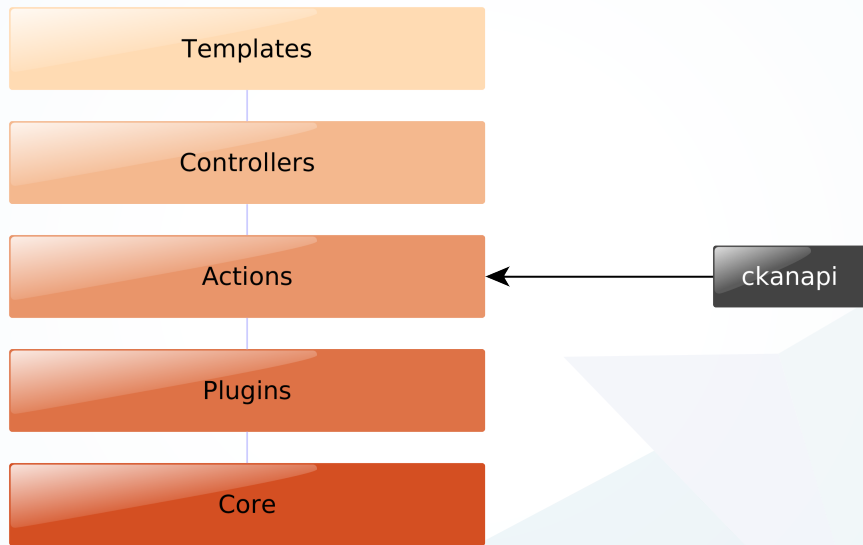
```
jq 'select(.organization.name!="nrcan-rncan")' -c datasets.jsonl | wc -l
```

```
8613
```

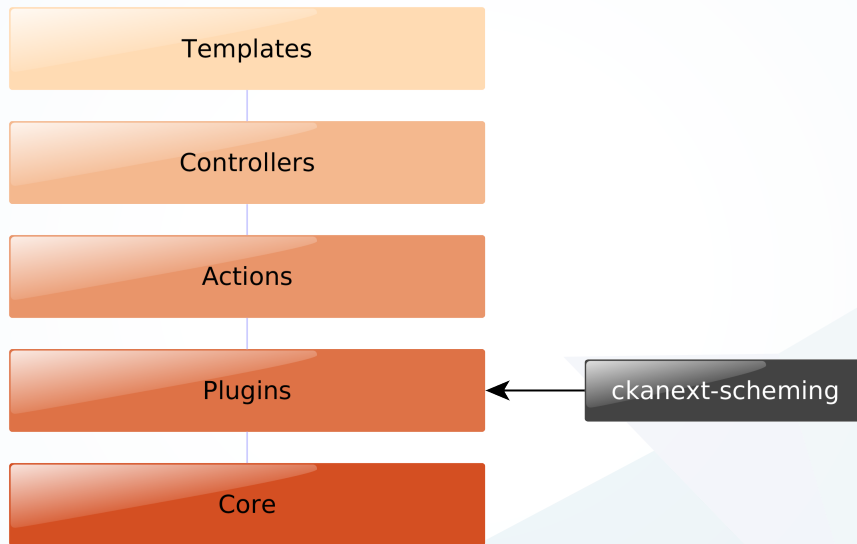

Command-line client

```
pip install ckanapi  
ckanapi -h
```

CKAN layers



CKAN layers



IDatasetForm is awesome

- Define dataset types
- Custom templates for edit + show
- New fields + validation rules for datasets + resources

IDatasetForm is hard

- Python plugin code
- Nested Jinja2 templates
- navl schema + validators

ckanext-scheming is easier

- Code optional
- Templates included
- Combined JSON schema
- Custom validators with IValidators

ckanext-scheming is sharable

- `scheming_dataset_schema_list`
- `scheming_dataset_schema_show`

Example schema

```
{
  "dataset_type": "camel-photos",
  "dataset_fields": [
    {
      "field_name": "title",
      "label": "Title",
      "form_placeholder": "eg. Larry, Peter, Susan",
      "form_snippet": "large_text.html",
      "form_attrs": { "data-module": "slug-preview-target" },
      "validators": "if_empty_same_as(name) unicode"
    },
    {
      "field_name": "name",
      "label": "URL",
      "form_placeholder": "eg. camel-no-5",
      "form_snippet": "slug.html",
      "validators": "not_empty unicode name_validator package_name_validator"
    },
  ],
}
```


Example schema

```
{
  "dataset_type": "camel-photos",
  "dataset_fields": [
    {
      "field_name": "title",
      "label": "Title",
      "form_placeholder": "eg. Larry, Peter, Susan",
      "preset": "title"
    },
    {
      "field_name": "name",
      "label": "URL",
      "form_placeholder": "eg. camel-no-5",
      "preset": "dataset_slug"
    }
  ],
}
```


Example schema


```
{
  "field_name": "category",
  "label": "Category",
  "help_text": "Make and model",
  "help_inline": true,
  "preset": "select",
  "choices": [
    { "value": "bactrian", "label": "Bactrian Camel" },
    { "value": "hybrid", "label": "Hybrid Camel" },
    { "value": "f2hybrid", "label": "F2 Hybrid Camel" },
    { "value": "snowwhite", "label": "Snow-white Dromedary" },
    { "value": "black", "label": "Black Camel" }
  ]
},
```

Example schema

```
{
  "field_name": "personality",
  "label": "Personality",
  "preset": "multiple_checkbox",
  "choices": [
    { "value": "friendly", "label": "Often friendly" },
    { "value": "jealous", "label": "Jealous of others" },
    { "value": "spits", "label": "Tends to spit" }
  ]
},
{
  "field_name": "a_relevant_date",
  "label": "A relevant date",
  "preset": "date"
},
}
```

Example form

 / Datasets / Create Dataset

 What are datasets?

A CKAN Dataset is a collection of data resources (such as files), together with a description and other information, at a fixed URL. Datasets are what users see when searching for data.

1 Create dataset


2 Add data

Title:

* URL: 127.0.0.1:5335/dataset/<dataset>

Humps:

Category:

 Make and model

Personality: ☒ Often friendly
☐ Jealous of others
☐ Tends to spit

A relevant date:

Other:

Future

- Groups and organizations
- Search facets
- Schema registry?

ckanapi and ckanext-scheming

- github.com/ckan/ckanapi
- github.com/open-data/ckanext-scheming

Questions?

ian@datacats.com