
pydora

Release 0.2

Maxime Borry

Feb 10, 2022

CONTENTS:

1	PyDora	3
1.1	Installation	3
1.2	Quick start	3
1.3	Documentation	4
2	Python API	5
3	CLI - Command Line Interface	7
3.1	pydora	7
4	Example input files	9
4.1	Credentials file	9
4.2	Projects file	9
4.3	Tags file	9
5	Connecting outside MPI-EVA	11
6	Indices and tables	13
	Python Module Index	15
	Index	17

Homepage: github.com/sidora-tools/pydora

PYDORA

PyDora is a toolkit to retrieve samples and metadata from the MPI-EVA Pandora internal database. PyDora is the Python cousin of [Sidora](#).

You can use PyDora both as a [command line tool](#), or directly [from Python](#).

1.1 Installation

- Install using pip (most people)

1. If you **don't** have set up your GitHub ssh keys

```
$ pip install git+https://github.com/sidora-tools/pydora
```

1. If you have set up your GitHub ssh keys

```
$ pip install git+ssh://git@github.com/sidora-tools/pydora.git
```

- Install in dev environment

```
$ git clone git@github.com:sidora-tools/pydora.git
$ cd pydora
$ conda create -f environment.yml
$ conda activate pydora_dev
$ pip install -e .
```

1.2 Quick start

```
$ pydora -c credentials.json -t assets/example_tags.txt
Successfully Connected to Pandora Database
Making request to Pandora SQL server
Downloaded table
Samples and metadata have been written to /Users/maxime/Documents/github/pydora/
↪pandora_samples.csv
```

1.3 Documentation

The documentation of PyDora is available here: pydora.rtfld.io

PYTHON API

`pydora.get_credentials(credentials)`

Get credentials to access SQL server

Parameters `credentials` (*str*) – Json formatted files with credentials for accessing Pandora

Returns {host:'server_address', login:'login', password:'pwd'}

Return type dict

`pydora.retrieve_samples(host, port, login, password, projects=None, tags=None, output=None, join='sql')`

Retrive samples having projects or tags from Pandora DB

Parameters

- **host** (*str*) – Address of SQL server
- **port** (*int*) – Port of SQL server
- **login** (*str*) – login
- **password** (*str*) – password
- **projects** (*list*) – list of projects to include (one per line)
- **tags** (*list*) – list tags to include (one per line)
- **join** (*str*) – Table join method, either pandas (local) or sql (server)

Returns (pandas dataframe) Table of retrieved samples and metadata

CLI - COMMAND LINE INTERFACE

To access the help menu:

```
$ pydora --help
```

The list of arguments of options is detailed below

3.1 pydora

PyDora: Retrieve samples and metadata from MPI-EVA Pandora internal database

Author: Maxime Borry

Contact: <maxime_borry[at]eva.mpg.de>

Homepage: github.com/sidora-tools/pydora

```
pydora [OPTIONS]
```

Options

--version

Show the version and exit.

-c, --credentials <credentials>

Default credentials.json

-p, --projects <projects>

File listing projects to include (one per line)

-t, --tags <tags>

File listing tags to include (one per line)

--join <join>

Join method

Default sql

Options sqlpandas

-o, --output <output>

Warinner samples metadata information

Default pandora_samples.csv

EXAMPLE INPUT FILES

For the CLI usage of PyDora, you need up to 3 files:

4.1 Credentials file

An example `credentials.json` file. For real credentials, please ask on the [Sidora mattermost channel](#)

4.2 Projects file

An example `projects.txt` file

4.3 Tags file

- An example `tags.txt` file

CONNECTING OUTSIDE MPI-EVA

When connecting from outside the MPI-EVA servers (e.g.) from your laptop, through the VPN, you have to establish a ssh tunnel

```
ssh -L 10001:pandora.eva.mpg.de:3306 <yourusername>@daghead1
```

You will need to slightly modify the `credentials.json` file to account for the ssh tunnel. An example `credentials.json` file when working through a ssh tunnel can be found here [assets/example_credentials_ssh_tunnel.json](#)

INDICES AND TABLES

- `genindex`
- `modindex`
- `search`

PYTHON MODULE INDEX

p

pydora, 5

Symbols

`--credentials <credentials>`
 pydora command line option, 7
`--join <join>`
 pydora command line option, 7
`--output <output>`
 pydora command line option, 7
`--projects <projects>`
 pydora command line option, 7
`--tags <tags>`
 pydora command line option, 7
`--version`
 pydora command line option, 7
`-c`
 pydora command line option, 7
`-o`
 pydora command line option, 7
`-p`
 pydora command line option, 7
`-t`
 pydora command line option, 7

G

`get_credentials()` (*in module pydora*), 5

M

module
 pydora, 5

P

pydora
 module, 5
 pydora command line option
 `--credentials <credentials>`, 7
 `--join <join>`, 7
 `--output <output>`, 7
 `--projects <projects>`, 7
 `--tags <tags>`, 7
 `--version`, 7
 `-c`, 7
 `-o`, 7
 `-p`, 7

`-t`, 7

R

`retrieve_samples()` (*in module pydora*), 5