

OpenCayle

caso de uso

Ángel Manuel Guerrero Higuera
am.guerrero@unileon.es



Marzo 2021

Distributed under: Creative Commons Attribution-ShareAlike 4.0 International



- Un repositorio de información se considera *abierto* (*open data*) si la información que contiene se difunde a todo aquel que lo requiera **sin restricción** alguna.
- Open Definition (<http://opendefinition.org/>), un proyecto de Open Knowledge International, precisa el significado de *open*:
 - ▶ Abierto significa que cualquiera puede acceder, usar, modificar y compartir libremente con cualquier propósito (sujeto, como mucho, a restricciones para preservar la autoría y la apertura)
- El concepto *Open Data* surge de la idea de que determinados tipos de datos deberían estar a disposición de cualquiera que desee utilizarlos de cualquier forma, sin restricciones de copyright, patentes u otros mecanismos de control.
 - ▶ Soren Auer et al. (2007). Dbpedia: A nucleus for a web of open data. In *The semantic web* (pp. 722-735). Springer, Berlin, Heidelberg.



- El problema surge cuando estos datos tienen valor comercial.
- El acceso o la reutilización de los datos, controlado por diversas organizaciones, tanto públicas como privadas, se realiza mediante restricciones como licencias, copyright, patentes y cargos por uso o reutilización.
- El concepto Open Data mantiene que estas restricciones van en contra del bien común y que estos datos deberían estar disponibles sin ninguna restricción o coste.
- En consecuencia, una parte importante de la comunidad científica aboga por el acceso libre a la información, argumentando que ello favorecería un avance científico mayor y más rápido.
 - ▶ Jennifer C. Molloy (2011). The open knowledge foundation: open data means better science. PLoS Biol, 9(12), e1001195.
- Organizaciones como Open Knowledge International (<https://okfn.org/>) tratan de favorecer y dar a conocer esta forma de pensar y trabajar.



- Los objetivos del movimiento Open Data son similares a los que persiguen otros movimientos *abiertos*, como son Open Source, Open Hardware, Open Content u Open Access.
- La filosofía detrás del movimiento Open Data existe desde hace tiempo, sin embargo, el termino en sí mismo es reciente, ganando popularidad con el auge de Internet e iniciativas como Data.gov o Data.gov.uk.
- La filosofía Open Data puede aplicarse a información de cualquier tipo: científica, médica, datos de administraciones públicas privadas, etc.



OpenCayle

Plataforma de servicios para almacenamiento masivo y publicación de datos científicos, aunque puede ser usado para datos de cualquier naturaleza.

Servicios:

1 Publicación

- ▶ Portal CKAN → <https://open.scayle.es/>

2 Almacenamiento masivo

- ▶ El servicio de almacenamiento, base de la plataforma, está basado en el concepto de almacenamiento de objetos y buckets del protocolo S3.
 - 100 buckets.
 - 1Tb de almacenamiento, distribuido como el usuario considere en objetos asociados a buckets.





OpenCayle. Guía de usuarios

Español English

OpenCayle. Guía de usuarios

OpenCayle es una plataforma de servicios para almacenamiento masivo, publicación y tratamiento de datos científicos, aunque puede ser usado para datos de cualquier naturaleza. El servicio de almacenamiento, base de la plataforma, está basado en el concepto de almacenamiento de objetos y buckets del protocolo S3.

Los datos del servicio OpenCayle por defecto son:

- 100 buckets.
- 1Tb de almacenamiento, distribuido como el usuario considere en objetos asociados a los buckets.

<https://www.scayle.es/manual/es/opencayle>

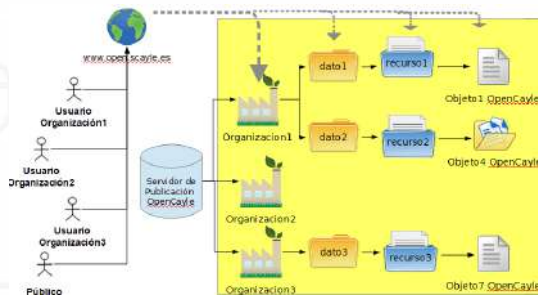


1. Servicio de publicación



Visión general del servicio de publicación

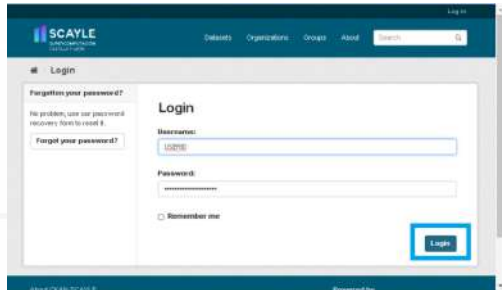
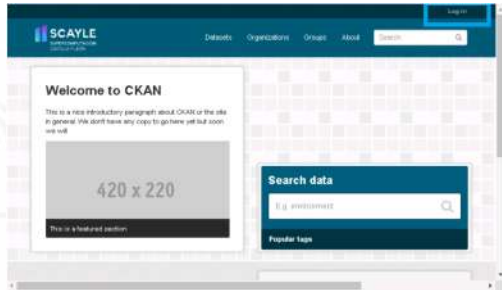
OpenCayle proporciona una plataforma de publicación de datos abiertos almacenados en su servicio, de modo que las organizaciones puedan hacer público los datos asociados a sus investigaciones, publicaciones, etc.



- **Organización** → agrupan colecciones de datos.
- **Dato** (Dataset) → colección de recursos de datos (como ficheros).
- **Recurso** → enlace a un objeto que contiene información.



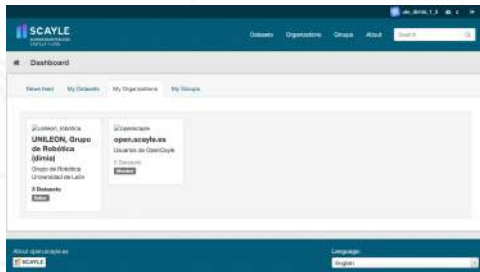
Acceso al servicio de publicación de OpenCayle



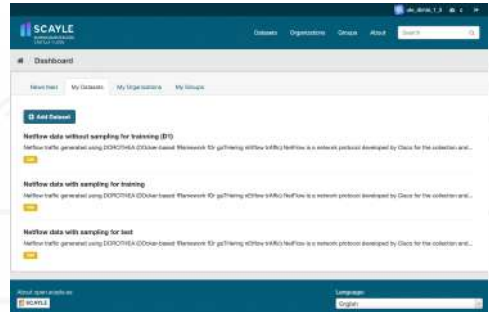
<https://open.scayle.es/>



Dashboard del usuario



Mis organizaciones



Mis datasets



Crear un dataset y añadir recursos

1 Create dataset 2 Add data

Title:

* URL: open.acalya.es/dataset/c/dataset [Edit](#)

Description:

You can use Markdown formatting here

Tags:

License:
 License information and additional information can be found at [creativecommons.org](#)

Organizations:

Visibility:

Paso 1: añadir metadatos

1 Create dataset 2 Add data

URL:

Name:

Description:

You can use Markdown formatting here

Format:

* This will be guessed automatically. Later mark it as wrong.

[Previous](#) [Save & add another](#) [Finish](#)

Paso 2: añadir recursos



■ Modificar o eliminar un recurso:

1 Seleccionar organización

2 Seleccionar dataset

3 Seleccionar un recurso de la lista → Explore → Edit

- Para modificar el recurso → Editar los datos en el formulario → Update resource
- Para eliminar el recurso → Delete

■ Modificar o eliminar un dataset:

1 Seleccionar organización

2 Seleccionar un dataset de la lista → Manage

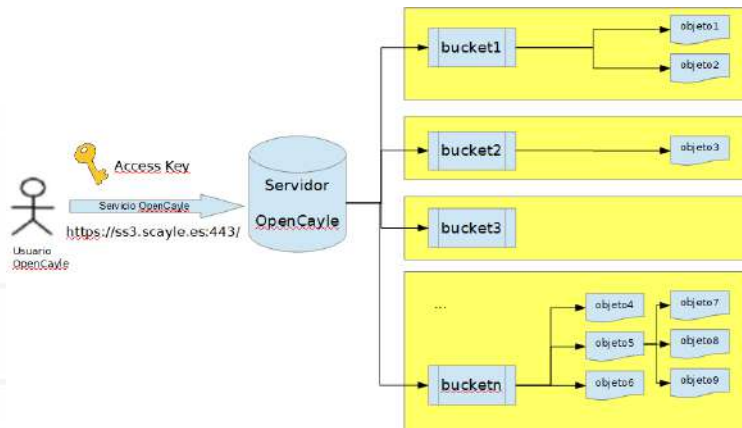
- Para modificar el dataset → Editar los datos en el formulario → Update Dataset
- Para eliminar el dataset → Delete



2. Servicio de almacenamiento masivo



Visión general del servicio de almacenamiento



- **Bucket** → Contenedor de objetos.
- **Objeto** → Fichero (*key*) guardado en un bucket.
- **AccessKey** → Token de conexión al servicio OpenCayle.



Acceso al servicio usando Python I

- Instalación y descarga del software necesario → <https://www.scayle.es/manual/es/opencayle/servicio-almacenamiento-con-s3/uso-servicio-con-python>
- Probar la conexión → `s3CheckConnection.py`

```
PARAMETERS: Access Key <XXXX>.
```

```
$ python3 s3CheckConnection.py ewogICAgIlJHV19UT0tFTiI6I...  
Your accessKey connects successfully to OpenCayle
```

- Información sobre los recursos OpenCayle del usuario → `s3Resources.py`.

```
PARAMETERS: Access Key <XXXX>.
```

```
$ python s3Resources.py ewogICAgIlJHV19UT0tFTiI6I...  
PARAMETERS: Access Key <ewogICAgIlJHV19UT0tFTiI6I...>
```

```
-----  
OpenCayle Resources associated to user.  
-----
```

```
...
```

```
-----  
END OpenCayle Resources
```



Acceso al servicio usando Python II

■ Crear bucket → s3CreateBucket.py

PARAMETERS: Access Key <XXXX> Bucket <YYYY>.

```
$ python3 s3CreateBucket.py ewogICAgIlJHV19UT0tFTiI6IHsKICAgICAgICAidmVyc2lvbiI6IDE... test-bucket
PARAMETERS: Access Key <ewogICAgIlJHV19UT0tFTiI6I...> Bucket <test-bucket>.
Bucket test-bucket created successfully.
```

■ Crear key → s3CreateKey.py

PARAMETERS: Access Key <XXXXX> Bucket Name <YYYY> Key Name <ZZZZ> File Name <PATH-TO-LOCAL-FILE> Is Private (0 Not/1 Yes) <0> ExpirationTime (0 not expiration) <280>.

```
$ python3 s3CreateKey.py ewogICAgIlJHV19UT0tFTiI6I... test-bucket netflow-data.csv netflow-data.csv 0
280
Key netflow-data.csv on bucket testing created on mode public-read and security False. If it is
secure, its expiration time is 0.
URL associated is: https://ss3.scayle.es:443/testing/netflow-data.csv
Key netflow-data.csv associated to Bucket test-bucket created successfully.
```



■ Borrar una key → s3DeleteKey.py

PARAMETERS: Access Key <XXXX> Bucket Name <YYYY> Key Name <ZZZZ>.

```
$ python3 s3DeleteKey.py ewogICAgIlJHV19UT0tFTiI6I... test-bucket netflow-data.csv  
Key netflow-data.csv deleted from bucket test-bucket successfully
```

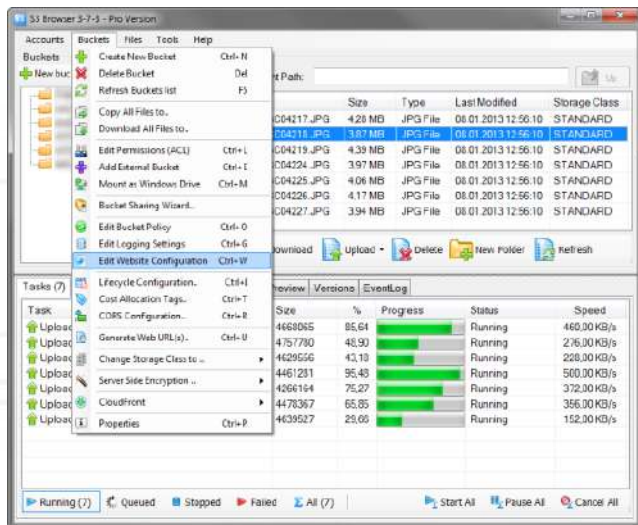
■ Borrar un bucket (y las keys asociadas) → s3DeleteBucket.py

PARAMETERS: Access Key <XXXX> Bucket <YYYY>.

```
$ python3 s3DeleteBucket.py ewogICAgIlJHV19UT0tFTiI6I... test-bucket  
Bucket test-bucket deleted successfully
```



Uso servicio con Interfaz gráfica (S3Browser) I



<https://s3browser.com/>

Uso servicio con Interfaz gráfica (S3Browser) II

Edit Account online help

Edit account details and click Save changes

Account Name:
usuario@es3.acayla.es
Assign any name to your account.

Account Type:
S3 Compatible Storage
Choose the storage you want to work with. Default is Amazon S3 Storage.

REST Endpoint:
es3.acayla.es:443
Specify S3-compatible API endpoint. It can be found in storage documentation. Example: test.server.com:8080

Signature Version:
Signature V2
Choose the supported signature version. Default value is Signature V2

Access Key ID:
accessKeyCodificadaEnBase64
Required to sign the requests you send to Amazon S3, see more details at <https://s3browser.com/keys>

Secret Access Key:
accessKeyCodificadaEnBase64
Required to sign the requests you send to Amazon S3, see more details at <https://s3browser.com/keys>

Encrypt Access Keys with a password:

Menú Accounts → Manage accounts

Configuración del acceso a OpenCayle



Uso servicio con Interfaz gráfica (S3Browser) III

Add New Mapped Drive

Specify new drive properties and click Add new drive

Storage Type:
S3 Compatible Storage
Choose the storage you want to work with. Default value is Amazon S3 Storage.

REST Endpoint:
s3.scayle.es:443
Specify S3-compatible API endpoint. It can be found in storage documentation. Example: rest.server.com:3000

Signature Version:
Signature V2
Choose the supported signature version. Default value is Signature V2.

Access Key ID:
XXXXXX
Access Key ID can be found here: https://console.aws.amazon.com/iam/home?#security_credentials

Secret Access Key:
XXXXXXXX
Secret Access Key can be found here: https://console.aws.amazon.com/iam/home?#security_credentials

Amazon S3 Bucket:

Select or specify bucket name and optional path. Leave this field empty to mount entire account.

Mapped Drive Letter:
Z
Select available drive letter for your drive.

Menú Buckets → Mount as Windows Drive → TntDrive

Uso del servicio como unidad de disco



Estadísticas de uso



<https://extranet.scayle.es/>



OpenCayle

caso de uso

Ángel Manuel Guerrero Higuera
am.guerrero@unileon.es



Marzo 2021

Distributed under: Creative Commons Attribution-ShareAlike 4.0 International

