Documentación extensiones CKAN





1. CONTEXTO

2. OBJETIVO DEL DOCUMENTO

3. DESCRIPCIÓN DE MÓDULOS DE CKAN

- 3.1. Ckanext-dge
- 3.2. Ckanext-dge-scheming
- 3.3. Ckanext-dge-harvest
- 3.4. Ckanext-dge-drupal-users
- 3.5 Ckanext-dge-dashboard
- 3.6 Ckanext-dge-ga
- 3.7 Ckanext-dge-ga-report
- 3.8 Configuración de las extensiones: Ckanext-dge-scheming, Ckanext-dge-harvest, Ckanext-dge, Ckanext-dge-drupal-users
- 3.9 Configuración de las extensiones: Ckanext-dge-dashboard, Ckanext-dge-ga y Ckanext-dge-ga-report

4. LISTA DE FIGURAS

- Figura 1- Formulario para la creación de un dataset
- Figura 2- Accediendo a "Mi federador"
- Figura 3- Formulario para crear una fuente de federación
- Figura 4- Accediendo a la sección administrador
- Figura 5- Comprobando el estado de la federación
- Figura 6- Comprobando el filtro categoría
- Figura 7-Validando "Mi cuadro de mandos"











Figura 8-Validando el Cuadro de mandos general



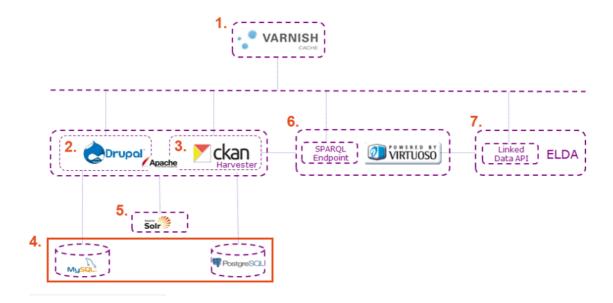


1. CONTEXTO

La página web *datos.gob.es* usa una tecnología basada en una arquitectura Open Source que combina diferentes componentes para cubrir las necesidades funcionales del proyecto.

La plataforma de la aplicación integra principalmente una herramienta de gestión de contenidos, un módulo multidioma que permite una navegación simétrica para cada uno de los idiomas, un gestor del catálogo de datos, un motor de federación para integrar los catálogos de datos de los organismos publicadores, una solución de Linked Data API y un punto de consulta SPARQL.

El diagrama de arquitectura del proyecto es el siguiente:



- 1. Proxy inverso: a través de Varnish se gestiona el enrutamiento a los cuatro servicios del proyecto: portal de contenidos, catálogo de datos, punto de consulta SPARQL y Linked Data API.
- 2. Portal de contenidos: desarrollado sobre el gestor de contenidos Drupal; gestiona el contenido editorial del proyecto.
- Catálogo de datos: implementado sobre CKAN; contiene el repositorio único de metadatos; integra el federador que permite sincronizar con los catálogos de todos los organismos publicadores.
- 4. Capa de persistencia: repositorio de datos para el portal de contenidos y el catálogo de datos; implementado en MySQL y PostgreSQL.
- 5. Motor de indexación: implementado sobre Apache Solr; provee de capacidades de búsqueda tanto para el portal de contenidos como para el catálogo de datos.
- 6. Punto de consulta SPARQL: implementado sobre Virtuoso de OpenLink Software como almacén de tripletas RDF; carga el grafo completo del catálogo de datos y provee de un punto de consulta SPARQL sobre el mismo.
- 7. Linked Data API: desarrollado sobre ELDA de Epimorphics; provee un RESTful API, actuando como Gateway sobre el punto de consulta SPARQL.









Como se puede observar, la aplicación utiliza CKAN, que es un catálogo de datos el cual contiene el repositorio único de metadatos y el cual tiene integrado un federador que permite sincronizar con los catálogos de todos los organismos publicadores.





2. OBJETIVO DEL DOCUMENTO

La finalidad de este documento es explicar detalladamente la funcionalidad de cada una de las extensiones de CKAN utilizada en la plataforma así como sus pasos de instalación.

Las extensions utilizadas por la aplicación son: ckanext-dge, ckanext-dge-drupal-users, ckanext-dge-scheming, ckanext-dge-harvest, ckanext-dge-ga, ckanext-dge-ga-report y ckanext-dge-dashboard.

Cabe recordar que el código fuente es propiedad de Entidad Pública Empresarial Red.es y está licenciado bajo GNU Affero General Public License, versión 3 o posterior.











3. DESCRPCIÓN DE LOS MÓDULOS DE CKAN UTILIZADOS POR EL PORTAL

datos.gob.es

3.1 Ckanext-dge

3.1.1 Descripción

La extensión "Ckanext-dge" modifica funcionalidades y estilos específicos que provee CKAN por defecto para adaptarlos a los de la página web datos.gob.es

3.1.2 Instalación

A continuación se presentarán los pasos para la instalación de esta extensión:

- Descargar de github el código fuente de la extensión. Este paquete se puede encontrar en: https://github.com/ctt-gobes/datos.gob.es/tree/master/ckan/ckanext-dge
- 2. Copiar dicho paquete en el directorio de trabajo que se quiera dentro de la máquina. Realizar este paso desde el entorno virtual donde se ha instalado CKAN. Para ello ejecutamos los siguientes comandos (en este ejemplo se ha copiado el paquete en el directorio /tmp):
 - . /usr/lib/ckan/default/bin/activate

 (default)root:/tmp# chown -R root:root ckanext-dge

 (default)root:/tmp# mkdir -p /usr/lib/ckan/default/src/ckanext-dge

 (default)root:/tmp# cp -pr ckanext-dge/. /usr/lib/ckan/default/src/ckanext-dge/
- 3. Una vez realizado lo anterior proceder a instalar ckanext-dge. Para ello ejecutamos el siguiente comando: (default)root:~# pip install -e /usr/lib/ckan/default/src/ckanext-dge
- 4. Para terminar la instalación de esta extensión es necesario realizar los pasos descritos en la sección 3.8 Configuración de las extensiones: Ckanext-dgescheming, Ckanext-dge-harvest, Ckanext-dge, Ckanext-dge-drupal-users









3.1.3 Comprobación de la instalación

Para corroborar que la extensión se ha instalado correctamente, se comprobará que los estilos de la página que ofrece CKAN por defecto han cambiado.

Además, se puede comprobar que el filtro "Categoría" existe ahora en la página web de CKAN. Este filtro no existe por defecto.



3.2 Ckanext-dge-scheming

3.2.1 Descripción

La extensión "Ckanext-dge-scheming" amplía el esquema de metadatos de CKAN a los requisitos de la Norma Técnica de Interoperabilidad de Reutilización de recursos de la información (NTI-RISP).

3.2.2 Instalación

A continuación se presentarán los pasos para la instalación de esta extensión:

 Descargar de github el código fuente de la extensión. Este paquete se puede encontrar en: https://github.com/ctt-gobes/datos.gob.es/tree/master/ckan/ckanext-dge-scheming





- 2. Copiar dicho paquete en el directorio de trabajo que se quiera dentro de la máquina. Realizar este paso desde el entorno virtual donde se ha instalado CKAN. Para ello ejecutamos los siguientes comandos (en este ejemplo se ha copiado el paquete en el directorio /tmp):
 - . /usr/lib/ckan/default/bin/activate

 (default)root:/tmp# chown -R root:root ckanext-dge-scheming

 (default)root:/tmp# mkdir -p /usr/lib/ckan/default/src/ckanext-dge-scheming

 (default)root:/tmp# cp -pr ckanext-dge-scheming/.

 /usr/lib/ckan/default/src/ckanext-dge-scheming/
- Una vez realizado lo anterior proceder a instalar ckanext-dge-scheming. Para ello ejecutamos el siguiente comando: (default)root:~# pip install -e /usr/lib/ckan/default/src/ckanext-dgescheming
- 4. Para terminar la instalación de esta extensión es necesario realizar los pasos descritos en la sección 3.8 Configuración de las extensiones: Ckanext-dgescheming, Ckanext-dge-harvest, Ckanext-dge, Ckanext-dge-drupal-users

3.2.3 Comprobación de la instalación

Para comprobar que esta extensión ha sido instalada de forma correcta hay que crear una nueva distribución en un dataset ya creado. A la hora de crear la distribución entre los datos a introducir debe aparecer un campo llamado "CA Nombre" que permita añadir un título para el dataset en Catalán tal y como se muestra en la imagen:









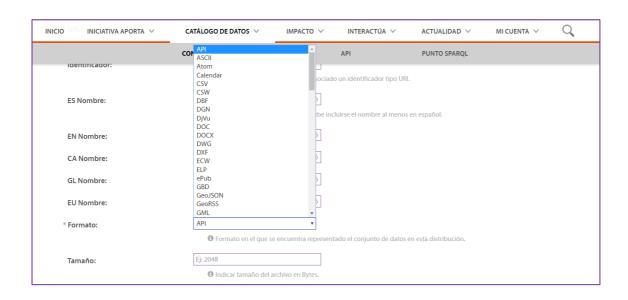


Figura 2- Formulario para la creación de un dataset





3.3 Ckanext-dge-harvest

3.3.1 Descripción

La extensión "Ckanext-dge-harvest" implementa el proceso de federación de los catálogos de datos de los organismos publicadores integrados en el sistema teniendo en cuenta las normas de interoperabilidad definidas en por la extensión "Ckanext-dge-scheming"

3.3.2 Instalación

A continuación se presentarán los pasos para la instalación de esta extensión:

- Descargar de github el código fuente de la extensión. Este paquete se puede encontrar en: https://github.com/ctt-gobes/datos.gob.es/tree/master/ckan/ckanext-dge-harvest
- 2. Copiar dicho paquete en el directorio de trabajo que se quiera dentro de la máquina. Realizar este paso desde el entorno virtual donde se ha instalado CKAN. Para ello ejecutamos los siguientes comandos (en este ejemplo se ha copiado el paquete en el directorio /tmp):
 - . /usr/lib/ckan/default/bin/activate
 (default)root:/tmp# chown -R root:root ckanext-dge-harvest
 (default)root:/tmp# mkdir -p /usr/lib/ckan/default/src/ckanext-dge-harvest
 (default)root:/tmp# cp -pr ckanext-dge-harvest/.
 /usr/lib/ckan/default/src/ckanext-dge-harvest/
- 3. Una vez realizado lo anterior proceder a instalar ckanext-dge-harvest y sus dependencias. Para ello ejecutamos el siguiente comando: (default)root:~# pip install -e /usr/lib/ckan/default/src/ckanext-dge-harvest (default)root:~# pip install -r /usr/lib/ckan/default/src/ckanext-dge-harvest/requirements.txt
- 4. Para terminar la instalación de esta extensión es necesario realizar los pasos descritos en la sección 3.8 Configuración de las extensiones: Ckanext-dgescheming, Ckanext-dge-harvest, Ckanext-dge, Ckanext-dge-drupal-users







3.3.3 Comprobación de la instalación

Para comprobar que esta extensión se ha instalado de forma correcta, se deben realizar las siguientes tareas:

1-Clicar en el menú "Mi cuenta" y posteriormente en el submenú "Federador". Posteriormente, se debe pulsar sobre el botón "Añadir fuente de federación".



Figura 3- Accediendo a "Mi federador"

2- Al pulsar sobre el botón "Añadir fuente de federación" se mostrará un formulario que se debe rellenar con los datos requeridos.

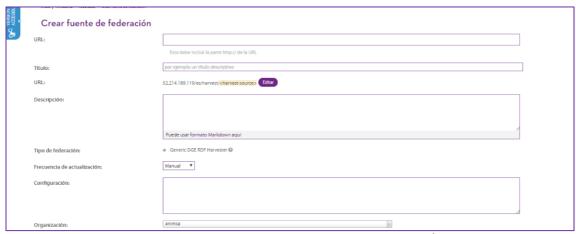


Figura 4- Formulario para crear una fuente de federación

Una prueba de que se ha instalado correctamente es que aparece en el formulario el campo "Tipo de federación" con el nombre "Generic DGE RDF Harvester" seleccionado.





3- A continuación, se debe cliclar sobre el botón "Administrador".



Figura 5- Accediendo a la sección administrador

4- Por último, es necesario seleccionar "Volver a federar" y esperar a que termine el proceso. Una vez el proceso ponga "Terminado", la extensión se habrá instalado correctamente.

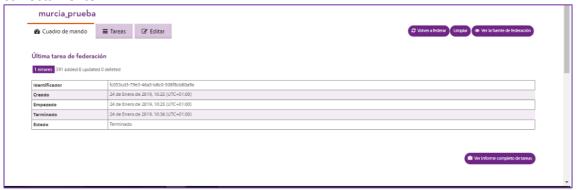


Figura 6- Comprobando el estado de la federación





3.4 Ckanext-dge-drupal-users

3.4.1 Descripción

La extensión "Ckanext-dge-drupal-users" permite mantener la sesión iniciada a los usuarios cuando pasan de la parte de DRUPAL del portal a la de CKAN.

3.4.2 Instalación

A continuación se presentarán los pasos para la instalación de esta extensión:

- Descargar de github el código fuente de la extensión. Este paquete se puede encontrar en: https://github.com/ctt-gobes/datos.gob.es/tree/master/ckan/ckanext-dge-drupal-users
- 2. Copiar dicho paquete en el directorio de trabajo que se quiera dentro de la máquina. Realizar este paso desde el entorno virtual donde se ha instalado CKAN. Para ello ejecutamos los siguientes comandos (en este ejemplo se ha copiado el paquete en el directorio /tmp):
 - . /usr/lib/ckan/default/bin/activate
 (default)root:/tmp# chown -R root:root ckanext-dge-drupal-users
 (default)root:/tmp# mkdir -p /usr/lib/ckan/default/src/ckanext-dge-drupal-users
 - (default)root:/tmp# cp -pr ckanext-dge-drupal-users /.
 /usr/lib/ckan/default/src/ckanext-dge-drupal-users /
- 3. Una vez realizado lo anterior proceder a instalar ckanext-dge-drupal-users y sus dependencias. Para ello ejecutamos los siguientes comandos:

(default)root:~# deactivate

root:~# apt-get install libmysqlclient-dev

(default)root:~# . /usr/lib/ckan/default/bin/activate

(default)root:~# pip install -e /usr/lib/ckan/default/src/ckanext-dge-drupalusers

(default)root:~# pip install -r /usr/lib/ckan/default/src/ckanext-dge-drupal-users/requirements.txt









4. Para terminar la instalación de esta extensión es necesario realizar los pasos descritos en la sección 3.8 Configuración de las extensiones: Ckanext-dgescheming, Ckanext-dge-harvest, Ckanext-dge, Ckanext-dge-drupal-users

3.4.3 Comprobación de la instalación

Para comprobar que la extensión se ha instalado correctamente, se puede comprobar que la sesión se mantiene si se accede a la sección de inicio de la aplicación y se pasa a la sección de "Catálogo de datos". En esta sección deberá aparecer un botón encima del buscador que ponga "Agrega conjunto de datos"





3.5 Ckanext-dge-dashboard

3.5.1 Descripción

La extensión "Ckanext-dge-dashboard" permite crear los cuadros de mandos que se muestran en la sección "Cuadro de mandos" del portal.

3.5.2 Instalación

A continuación se presentarán los pasos para la instalación de esta extensión:

- Descargar de github el código fuente de la extensión. Este paquete se puede encontrar en: https://github.com/ctt-gobes/datos.gob.es/tree/master/ckan/ckanext-dge-dashboard
- 2. Copiar dicho paquete en el directorio de trabajo que se quiera dentro de la máquina. Realizar este paso desde el entorno virtual donde se ha instalado CKAN. Para ello ejecutamos los siguientes comandos (en este ejemplo se ha copiado el paquete en el directorio /tmp):
 - . /usr/lib/ckan/default/bin/activate
 (default)root:/tmp# chown -R root:root ckanext-dge-dashboard
 (default)root:/tmp# mkdir -p /usr/lib/ckan/default/src/ckanext-dge-dashboard
 (default)root:/tmp# cp -pr ckanext-dge-dashboard /.
 - /usr/lib/ckan/default/src/ckanext-dge-dashboard /
- Una vez realizado lo anterior proceder a instalar ckanext-dge-dashboard.
 Para ello ejecutamos el siguiente comando:
 (default)root:~# pip install -e /usr/lib/ckan/default/src/ckanext-dge-dashboard
- 4. Para terminar la instalación de esta extensión es necesario realizar los pasos descritos en la sección 3.9 Configuración de las extensiones : Ckanext-dgedashboard, Ckanext-dge-ga y Ckanext-dge-ga-repo

3.5.3 Comprobación de la instalación









Para comprobar la correcta instalación de esta extensión hay que realizar dos pasos diferentes.

El primero de ellos consiste el autenticarse y verificar que en el menú "Mi cuenta" aparece el submenú "Mi cuadro de mandos".



Figura 7-Validando "Mi cuadro de mandos"

El segundo paso consiste en comprobar que sin estar autenticado aparece el submenú "Cuadro de Mandos" en el menú "Impacto".



Figura 8-Validando el Cuadro de mandos general





3.6 Ckanext-dge-ga

3.6.1 Descripción

La extensión Ckanext-dge-ga permite monitorizar los accesos que los usuarios hacen a las partes del portal relacionadas con CKAN, a través de Google Analytics.

3.6.2 Instalación

A continuación se presentarán los pasos para la instalación de esta extensión:

- Descargar de github el código fuente de la extensión. Este paquete se puede encontrar en: https://github.com/ctt-gobes/datos.gob.es/tree/master/ckan/ckanext-dge-ga
- 2. Copiar dicho paquete en el directorio de trabajo que se quiera dentro de la máquina. Realizar este paso desde el entorno virtual donde se ha instalado CKAN. Para ello ejecutamos los siguientes comandos (en este ejemplo se ha copiado el paquete en el directorio /tmp):
 - . /usr/lib/ckan/default/bin/activate
 (default)root:/tmp# chown -R root:root ckanext-dge-ga
 (default)root:/tmp# mkdir -p /usr/lib/ckan/default/src/ckanext-dge-ga
 (default)root:/tmp# cp -pr ckanext-dge-ga /.
 /usr/lib/ckan/default/src/ckanext-dge-ga /
- 3. Una vez realizado lo anterior proceder a instalar ckanext-dge-ga. Para ello ejecutamos el siguiente comando:

 (default)root:~# pip install -e /usr/lib/ckan/default/src/ckanext-dge-ga
- 4. Para terminar la instalación de esta extensión es necesario realizar los pasos descritos en la sección 3.9 Configuración de las extensiones : Ckanext-dgedashboard, Ckanext-dge-ga y Ckanext-dge-ga-repo

3.6.3 Comprobación de la instalación

Para comprobar que la extensión se ha instalado correctamente se debe acceder a la aplicación de Google Analytics y verificar que se están recogiendo el número de visitas del portal.





3.7 Ckanext-dge-ga-report

3.7.1 Descripción

La extensión "Ckanext-dge-ga-report" permite la generación de reportes basados en datos obtenidos de Google Analytics. En concreto, se descarga los datos almacenado en Google Analytics y realiza distintas operaciones para poder mostrarlos a través de las gráficas de la sección del cuadro de mandos del portal.

3.7.2 Instalación

A continuación se presentarán los pasos para la instalación de esta extensión:

- Descargar de github el código fuente de la extensión. Este paquete se puede encontrar en: https://github.com/ctt-gobes/datos.gob.es/tree/master/ckan/ckanext-dge-ga-report
- 2. Copiar dicho paquete en el directorio de trabajo que se quiera dentro de la máquina. Realizar este paso desde el entorno virtual donde se ha instalado CKAN. Para ello ejecutamos los siguientes comandos (en este ejemplo se ha copiado el paquete en el directorio /tmp):
 - . /usr/lib/ckan/default/bin/activate
 (default)root:/tmp# chown -R root:root ckanext-dge-ga-report
 (default)root:/tmp# mkdir -p /usr/lib/ckan/default/src/ckanext-dge-ga-report
 (default)root:/tmp# cp -pr ckanext-dge-ga-report /.
 /usr/lib/ckan/default/src/ckanext-dge-ga-report /
- 3. Una vez realizado lo anterior proceder a instalar ckanext-dge-ga-report. Para ello ejecutamos el siguiente comando: (default)root:~# pip install -e /usr/lib/ckan/default/src/ckanext-dge-ga-report
- 4. Para terminar la instalación de esta extensión es necesario realizar los pasos descritos en la sección 3.9 Configuración de las extensiones : Ckanext-dgedashboard, Ckanext-dge-ga y Ckanext-dge-ga-repo

3.7.3 Comprobación de la instalación







Para comprobar que la instalación ha sido correcta, se deberá comprobar que las gráficas del cuadro de mandos muestran los datos actualizados tal y como se presentan en la interfaz de Google Analytics.

3.8 Configuración de las extensiones: Ckanext-dge-scheming, Ckanext-dge-harvest, Ckanext-dge, Ckanext-dge-drupalusers

En esta sección se detallará cómo se deben activar las extensiones Ckanext-dgescheming, Ckanext-dge-harvest, Ckanext-dge y Ckanext-dge-drupal-users para su correcto funcionamiento.

1. (default)root:~# vi /etc/ckan/default/production.ini ckan.plugins = dcat dge_harvest dge_rdf_harvester harvest fluent scheming_datasets dge_scheming stats scheming.dataset_schemas = ckanext.dge_scheming:nti_dge_dataset.json scheming.presets = ckanext.scheming:presets.json ckanext.fluent:presets.json ckanext.dge_scheming:presets.json ckanext.dcat.rdf.profiles = dge_profile ckanext.dge.downloadable.formats = csv pdf kml #ckanext.dge_harvest.http_proxy= #ckanext.dge_harvest.https_proxy= ckan.harvest.mq.type = redis ckan.harvest.mq.hostname = IP_REDIS #ckan.views.default_views = image_view text_view recline_view ckan.locale default = es ckan.locale_order = es gl ca eu en ckan.locales_offered = es gl ca eu en

Debe sustituirse el siguiente parámetro:

IP_REDIS = IP del servidor Redis

2. Reiniciar el servidor:

(default)root:~# service apache2 restart











3. Crear el modelo de datos del federador (default)root:~# paster --plugin=ckanext-harvest harvester initdb -c /etc/ckan/default/production.ini

A continuación hay que configurar el federador. Los pasos para esta configuración son:

1.	Instalar el supervisor
	(default)root:~# apt-get install supervisor
2.	Activar las colas del federador de forma permanente:
	(default)root:~# mkdir /var/log/ckan (default)root:~# vi
	/etc/supervisor/conf.d/ckan_harvesting.conf;
	=======================================
	; ckan harvester
	;======================================
	[program:ckan_gather_consumer]
	command=/usr/lib/ckan/default/bin/pasterplugin=ckanext-harvest
	harvester gather_consumerconfig=/etc/ckan/default/production.ini
	; user that owns virtual environment.
	user=root
	numprocc_1
	numprocs=1
	stdout_logfile=/var/log/ckan/gather_consumer.log
	stderr_logfile=/var/log/ckan/gather_consumer.log
	autostart=true
	autorestart=true
	startsecs=10
	[program:ckan_fetch_consumer]









```
command=/usr/lib/ckan/default/bin/paster --plugin=ckanext-harvest
   harvester fetch_consumer --config=/etc/ckan/default/production.ini
   ; user that owns virtual environment.
   user=root
   numprocs=1
   stdout_logfile=/var/log/ckan/fetch_consumer.log
   stderr logfile=/var/log/ckan/fetch consumer.log
   autostart=true
   autorestart=true
   startsecs=10
   (default)root:~# supervisorctl reread
   (default)root:~# supervisorctl add ckan_gather_consumer
   (default)root:~# supervisorctl add ckan_fetch_consumer
3. Configurar la tarea que ejecuta el federador preiodicamente:
   (default)root:~# crontab -e
   # m h dom mon dow command
                     /usr/lib/ckan/default/bin/paster -plugin=ckanext-harvest
   harvester run -c /etc/ckan/default/production.ini
```

3.9 Configuración de las extensiones : Ckanext-dge-dashboard, Ckanext-dge-ga y Ckanext-dge-ga-report

En esta sección se detallará cómo se deben activar las extensiones Ckanext-dgedashboard, Ckanext-dge-ga y Ckanext-dge-ga-report para su correcto funcionamiento.

 En el servidor de CKAN, copiar el fichero de credenciales client_secret_***.json al directorio de la extensión ckanext-dgega-report: root:/tmp# cp client_secret_***.json /usr/lib/ckan/default/src/ckanext-dgega-report/credentials.json











Crear el token para la descarga de datos de la cuenta de Google Analytics: root:~# . /usr/lib/ckan/default/bin/activate (default)root:~#: cd /usr/lib/ckan/default/src/ckanext-dge-ga-report (default)root:/usr/lib/ckan/default/src/ckanext-dge-ga-report# paster -- plugin=ckanext-dge-ga-report dge_ga_report_getauthtoken -c /etc/ckan/default/production.ini (default)root:/usr/lib/ckan/default/src/ckanext-dge-ga-report# deactivate

Por línea de comandos se mostrará una URL y se solicitará una clave; la URL se debe ejecutar desde un navegador que devolverá un identificador; dicho identificador se debe insertar por línea de comandos; al finalizar el proceso, se habrá generado el fichero token.data en el mismo directorio de la extensión.

2. En el servidor de CKAN, crear las tablas del cuadro de mando en base de datos y cargar los datos históricos:

root:~# . /usr/lib/ckan/default/bin/activate

(default)root:~#: cd /usr/lib/ckan/default/src/ckanext-dge-dashboard (default)root:/usr/lib/ckan/default/src/ckanext-dge-dashboard# paster -- plugin=ckanext-dge-dashboard dge_dashboard_initdb -c

/etc/ckan/default/production.ini

(default)root:~#: cd /usr/lib/ckan/default/src/ckanext-dge-ga-report
(default)root:/usr/lib/ckan/default/src/ckanext-dge-ga-report# paster -plugin=ckanext-dge-ga-report dge_ga_report_initdb -c

/etc/ckan/default/production.ini

(default)root:/usr/lib/ckan/default/src/ckanext-dge-gareport# deactivate

3. En el servidor de CKAN,crear el alias para las rutas de los ficheros de datos del cuadro de mando:

root:/etc/apache2/conf-available# vi dashboard.conf

Alias /data /var/www/html/dashboard/data

root:/etc/apache2/conf-enabled# a2enconf dashboard

Crear el directorio para los ficheros de datos del cuadro de mando:

root:/var/www/html# mkdir -p dashboard/data









4. En el servidor de CKAN,copiar los ficheros de datos del cuadro de mando: root:/tmp/dashboard/data# cp * /var/www/html/dashboard/data/

