

datos.gob.es

**Cómo publicar datos abiertos de manera rápida y sencilla (con CKAN)**

Versión: diciembre 2017

**ÍNDICE**

**1. Introducción**

**2. Selección de los datos que se van a publicar**

2.1. Criterios

2.2. Temáticas

2.3. Restricciones

**3. Licencias para facilitar la reutilización**

3.1. Establecer unas condiciones de uso simples

3.2. Adoptar una licencia predefinida

**4. Preparación de los datos**

4.1. Formatos

4.2. Nivel de detalle

4.3. Preparación de los metadatos

**5. Publicar los datos (con CKAN)**

5.1. Requisitos previos

5.2. Instalación básica

5.3. Instalación avanzada

5.4. Administración

5.5. Características avanzadas

5.5.1. Compatibilidad con la NTI

5.6. Carga de datos

**6. Evaluación de uso**

**7. Diálogo con la comunidad**

**8. Mantenimiento**

**9. Evolución**

**10. Referencias**

**11. Tablas y gráficos**

**1. INTRODUCCIÓN**

El objetivo que se persigue con los **datos abiertos** (open data) es aparentemente sencillo: la publicación de esos datos de forma que puedan ser reutilizados y que todo el mundo se pueda beneficiar de ellos. Esta guía está destinada a todos aquellos que deseen empezar a compartir sus datos pero que dispongan de recursos limitados para poder hacerlo. El objetivo de la guía es servir de ayuda y orientación para que cualquiera pueda abrir fácilmente sus datos, aunque no tenga grandes conocimientos sobre estrategias de apertura de la información o las tecnologías asociadas a ellas[[1]](#footnote-1).

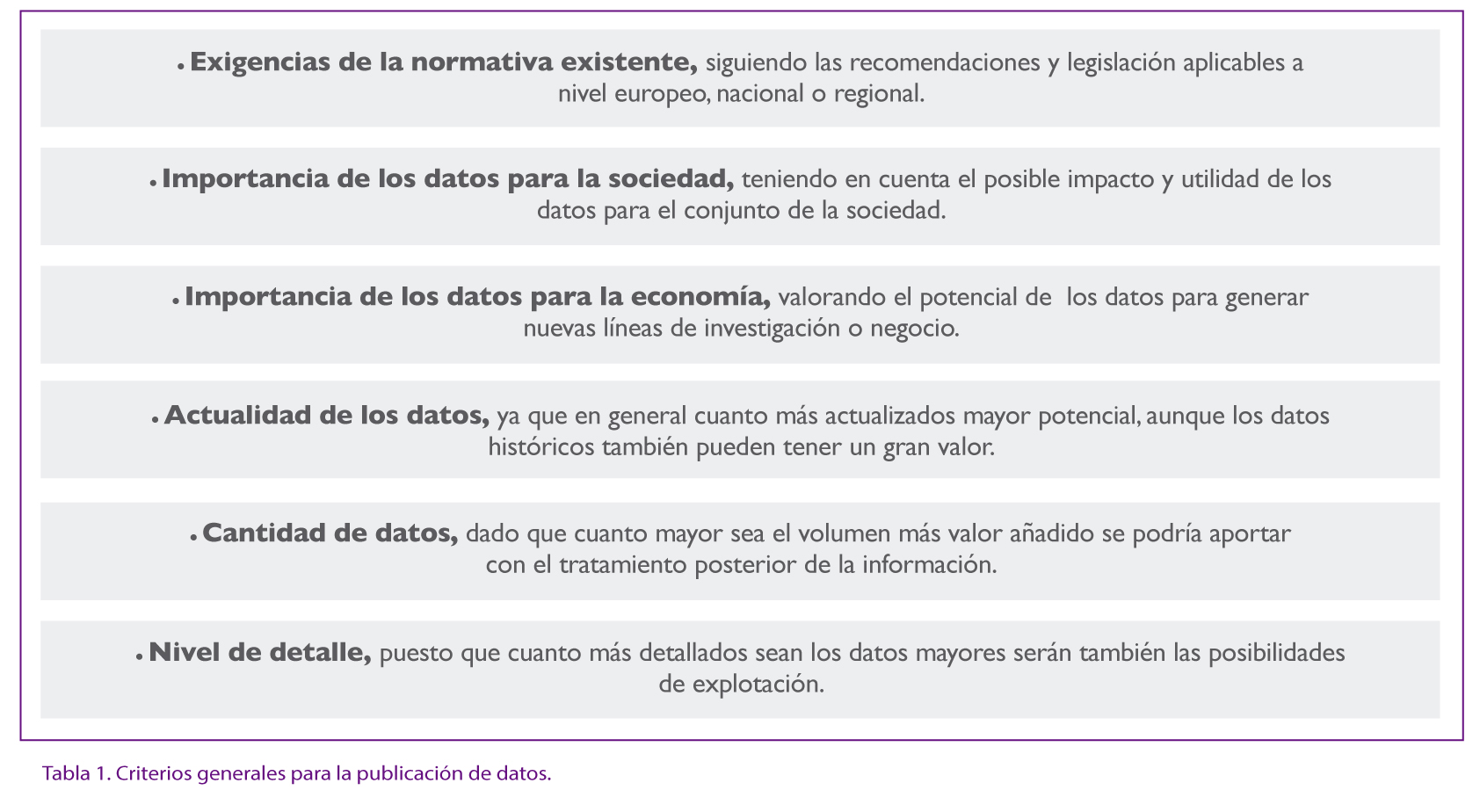
En la primera parte se establecen una serie de sencillas pautas que servirán de guía a la hora de localizar y preparar los datos para su publicación. A continuación se indicará como publicar los datos en la Web a través de un catálogo y de la forma más simple posible. Para ello utilizaremos como ejemplo [CKAN](http://ckan.org/), una plataforma de publicación de datos de código abierto disponible gratuitamente. No obstante, las recomendaciones generales de esta guía pueden ser de aplicación también con cualquier otra plataforma que se desee utilizar.

2**. SELECCIÓN DE DATOS QUE SE VAN A PUBLICAR**

El objetivo final de cualquier iniciativa de datos abiertos siempre será conseguir **publicar todos los datos disponibles**. Sin embargo, cuanto más limitados sean el tiempo y los recursos disponibles, más importante será determinar cuál es la información más valiosa desde el punto de vista de la *reutilización* para poder dar mayor prioridad a su publicación.

**2.1. Criterios**

La priorización debería adecuarse a las particularidades de cada caso. Los gestores de la información serán los mejores conocedores de los datos que manejan y de su potencial audiencia, pero algunos criterios generales que pueden servir de guía general tal y como se muestra en la siguiente tabla:



**2.2. Temáticas**

En cuanto a las temáticas más demandadas, existen varios criterios de referencia promovidos por organismos oficiales de relevancia en la materia tales como:

* Los resultados del estudio [MEPSIR](http://ec.europa.eu/information_society/newsroom/cf/document.cfm?doc_id=1198) de la Unión Europea;
* Las materias que generan negocio por parte del sector infomediario en España, según el [análisis del sector realizado por ONTSI;](http://www.ontsi.red.es/ontsi/es/content/estudio-de-caracterizaci%C3%B3n-del-sector-infomediario-2016)
* Los datos comunes a las distintas iniciativas open data en España;
* Los [conjuntos de datos prioritario definidos por el G8](https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex);
* Las [recomendaciones](http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?action=display&doc_id=6421) de la Comisión Europea.
* La [disponibilidad mundial de datos](http://opendatabarometer.org/report/analysis/implementation.html) analizada por el [Open Data Barometer](http://opendatabarometer.org/);
* La [clasificación de sectores temáticos](http://datos.gob.es/datos/sites/default/files/files/12_tax_02.pdf) utilizada por datos.gob.es.

En un primer análisis general se podrían sugerir como temáticas de alto interés los pertenecientes a los ámbitos de información del siguiente esquema.



En cualquier caso, cada publicador debe también completar y adaptar su propia lista de criterios e intereses de acuerdo a su actividad.

**2.3. Restricciones**

Aunque la norma general consiste simplemente en **publicar cuanto más mejor**, existen también limitaciones a la hora de publicar datos que se deben respetar, como se muestra a continuación.

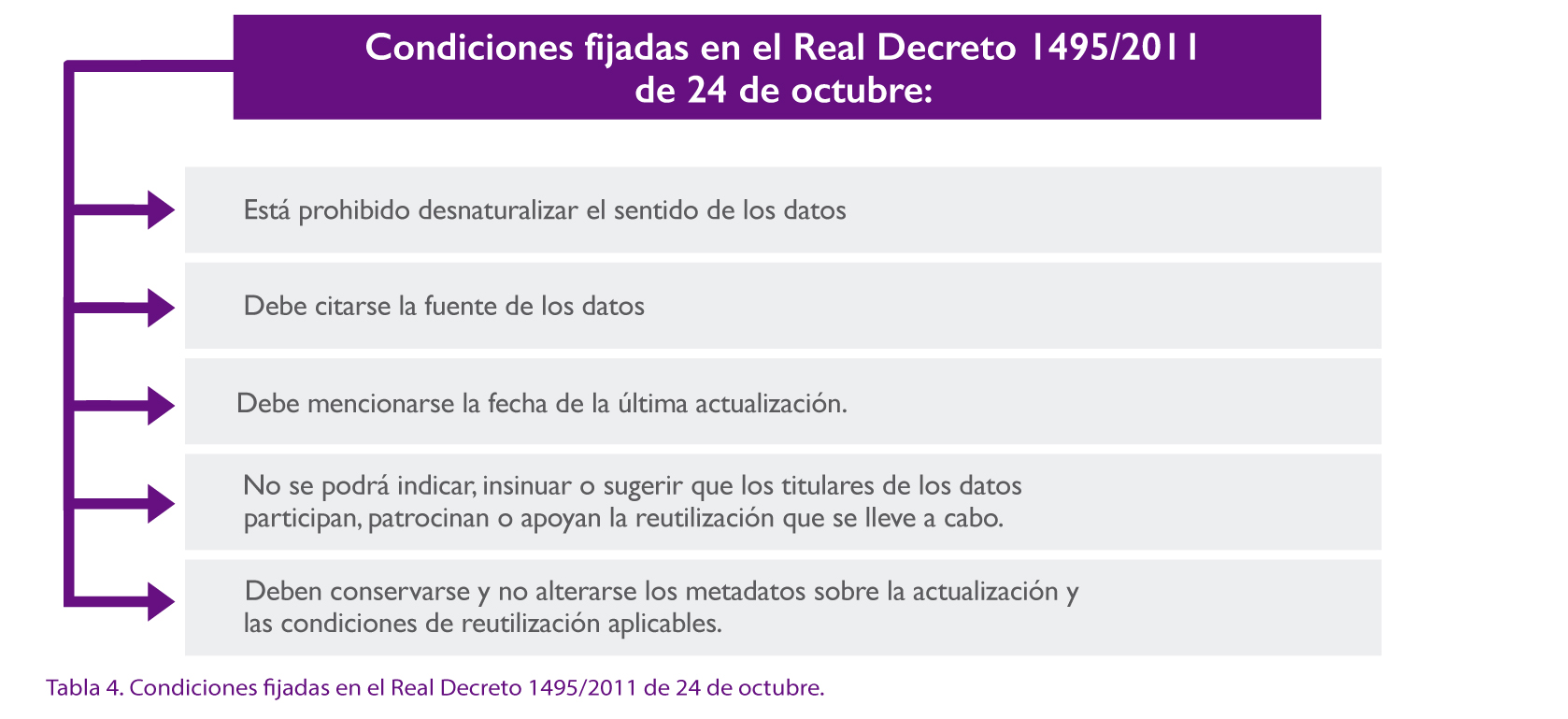
En cualquiera de estos casos la información restringida debería ser filtrada y eliminada **previamente a su publicación**.

**3. LICENCIAS PARA FACILITAR LA REUTILIZACIÓN**

Es fundamental que cualquier dato publicado incluya **información explícita sobre las condiciones de uso** de esa información para dejar claro a los potenciales usuarios cuáles son las posibilidades de reutilización de esos datos. Para hacerlo de forma simple contamos con dos opciones alternativas.

**3.1. Establecer unas condiciones de uso simples**

En este caso indicaremos que los datos están plenamente disponibles para su reutilización bajo una serie de condiciones mínimas, para las cuales podemos guiarnos por ejemplo por las indicadas en el [Real Decreto 1495/2011](http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-17560), de 24 de octubre.



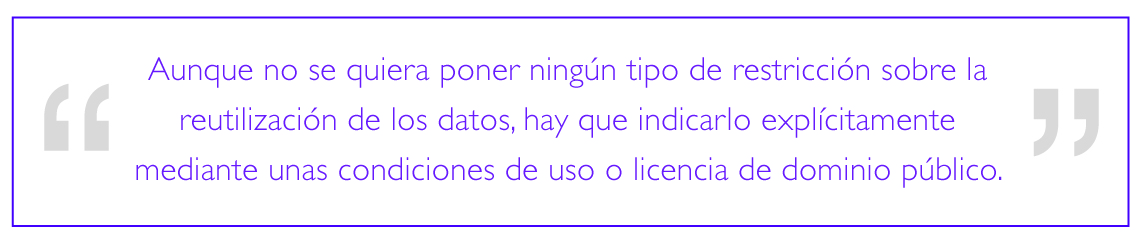
**3.2. Adoptar una licencia predefinida**

En este otro caso simplemente indicaremos que los datos están disponibles para su reutilización bajo las condiciones establecidas por alguna de las [licencias compatibles con los principios de los datos abiertos](http://opendefinition.org/licenses/).

Estas licencias podrán a lo sumo establecer restricciones relacionadas con la atribución de la fuente original y la obligatoriedad de compartir cualquier trabajo realizado con los datos bajo las mismas condiciones.

Podremos elegir por ejemplo entre las siguientes opciones:

1. Licencias de **dominio público**, que nos permiten compartir los datos sin ningún tipo de restricción:
   * + [Creative Commons Zero](https://creativecommons.org/publicdomain/zero/1.0/deed.es) (CC Zero).
     + [Public Domain Dedication and License](http://opendatacommons.org/licenses/pddl/summary/) (PDDL 1.0).
2. Licencias que únicamente obligan a **atribuir** la fuente original de los datos:
   * + [Creative Commons-Reconocimiento](http://creativecommons.org/licenses/by/4.0/deed.es) (CC BY 4.0).
     + [Open Data Commons Attribution License](http://opendatacommons.org/licenses/by/summary/) 1.0 (ODC-By 1.0).
3. Licencias que requieren **atribución** de la fuente original y compartir los resultados bajo las **mismas condiciones**:
   * + [Creative Commons-Reconocimiento Compartir Igual](http://creativecommons.org/licenses/by-sa/4.0/deed.es) (CC BY-SA 4.0).
     + [Open Data Commons Open Database License](http://opendatacommons.org/licenses/odbl/summary/) (ODbL 1.0).



Importante también señalar que, aunque no se quiera poner ningún tipo de restricción sobre la reutilización de los datos, hay que indicarlo explícitamente mediante unas condiciones de uso o licencia de dominio público, ya que de otro modo tendríamos dos problemas.

* Los potenciales reutilizadores no sabrían cuáles son las condiciones aplicables en la práctica por no tener ninguna indicación sobre ello.
* Podrían seguir aplicándose algunos derechos *por defecto* (por ejemplo, relacionados con los derechos de actor o de propiedad de las bases de datos) al no haber renunciado explícitamente a ellos mediante la declaración de dominio público u otro tipo de condiciones.

**4. PREPARACIÓN DE LOS DATOS**

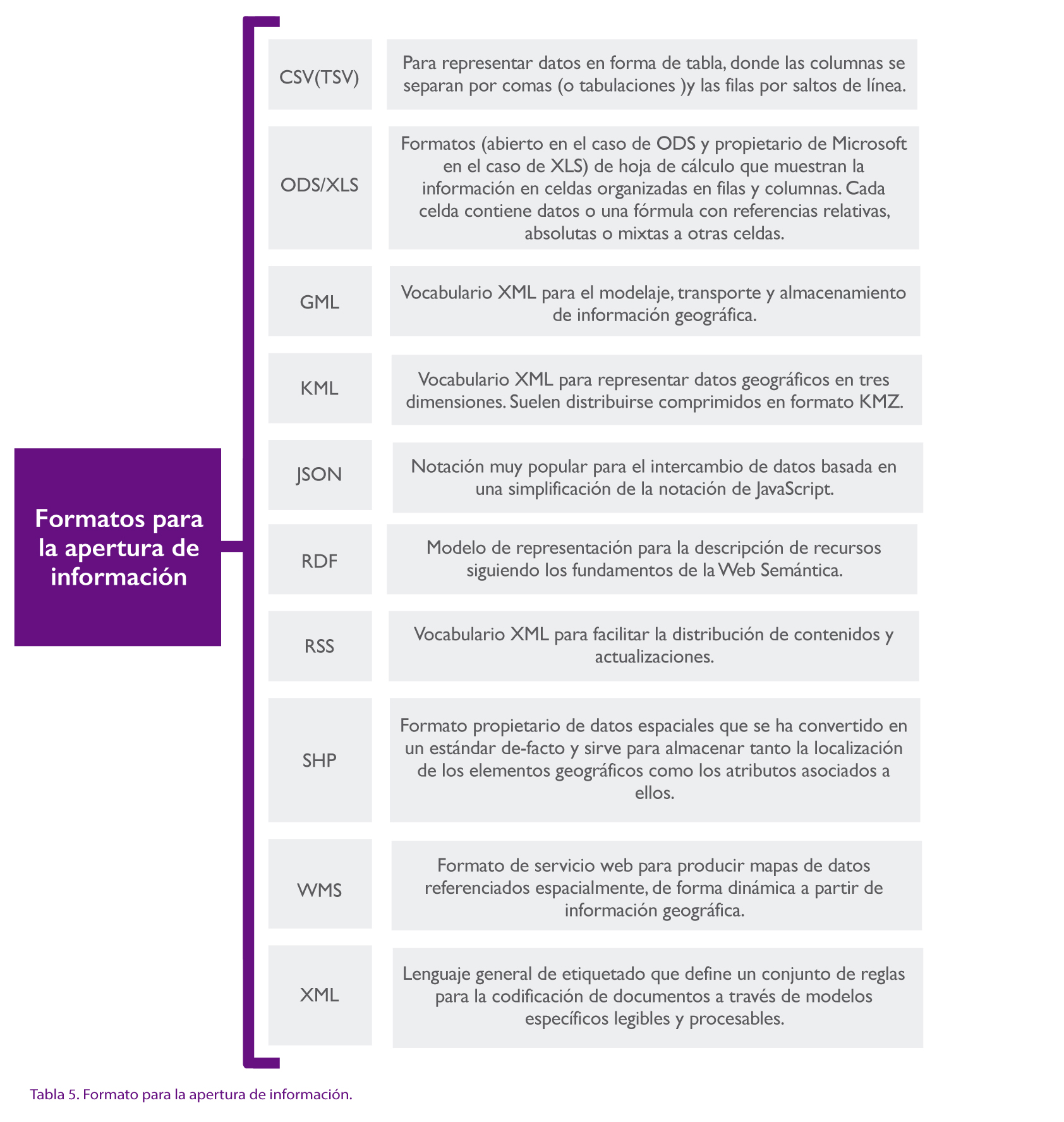
Uno de los principales requisitos para la reutilización es que el **procesamiento automático de los datos** sea posible, ya que esta es la única manera de trabajar con grandes cantidades de datos de forma eficiente. Para ello será necesario que la información esté publicada en los **formatos adecuados** y con el suficiente nivel de calidad y detalle para que los datos sean aprovechables.

**4.1. Formatos**

Algunos criterios a seguir a la hora de seleccionar los formatos adecuados son:

1. Formatos **estructurados** que siguen **modelos de datos estandarizados**.
2. Formatos directamente **legibles por las máquinas**.
3. Los formatos **más demandados** entre nuestros potenciales reutilizadores.
4. Formatos **abiertos** que permitan un uso **sin restricciones** para ampliar el público potencial.
5. Formatos que permitan publicar ***metainformación*** incorporada.
6. Formatos compatibles con lo establecido en la [**Norma Técnica de Interoperabilidad del Catálogo de estándares**](http://www.boe.es/boe/dias/2012/10/31/pdfs/BOE-A-2012-13501.pdf).

A continuación, podemos ver algunos ejemplos de formatos frecuentemente utilizados para la apertura de la información.



Asimismo, y siempre que se pueda, es recomendable también ofrecer los datos en varios **formatos distintos simultáneamente** para así intentar dar respuesta a las necesidades del mayor número de *reutilizadores* posible.

**4.2. Nivel de detalle**

Es también importante **respetar el nivel de detalle original de los datos,** tal y como se generan en su origen, evitando cualquier modificación y alteración previa a su publicación para que cada *reutilizador* pueda luego realizar el tratamiento de la información que más le convenga.

La única excepción a esta regla será el filtrado que pueda ser necesario realizar debido a las cuestiones de **privacidad**, **confidencialidad** **o** **seguridad** que se han citado anteriormente.

**4.3. Preparación de los metadatos**

El último paso previo a la publicación de los datos consistirá en preparar los **metadatos** que los acompañarán proporcionando información adicional para facilitar su localización, clasificación y reutilización. Es muy importante no sólo proporcionar el mayor número de metadatos posible, sino hacerlo también de la forma correcta, siguiendo los siguientes criterios

Por otro lado, a su vez los metadatos servirán a veces para describir los **conjuntos de datos** y otras para describir los distintos **recursos** (o archivos) que los componen. A continuación se muestra una tabla resumen de los metadatos que podemos proporcionar.



Para una descripción más detallada sobre cada uno de estos metadatos podemos consultar el Anexo III de la [NTI de Reutilización de recursos de información](http://www.boe.es/boe/dias/2013/03/04/pdfs/BOE-A-2013-2380.pdf) y su [guía de aplicación](http://datos.gob.es/saber-mas?q=node/2676).

**5. PUBLICAR LOS DATOS (CON CKAN)**

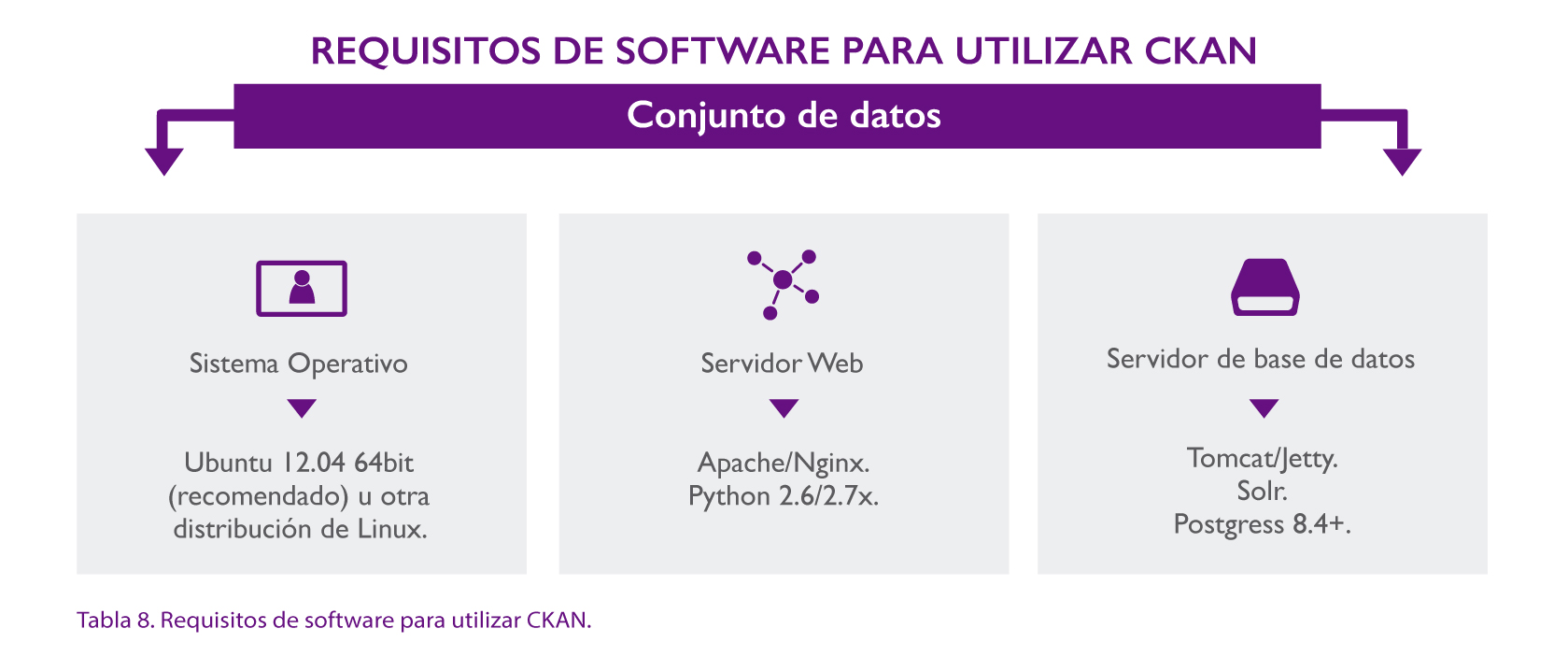
Para una gestión óptima de los datos de cara a su reutilización es necesario establecer una recopilación de los mismos en un punto de acceso común, que generalmente denominamos **catálogo de datos**, lo que permitirá una organización más eficaz y un acceso más fácil por parte de los potenciales *reutilizadores*.

En esta guía utilizaremos como ejemplo [CKAN](http://ckan.org/) porque es una plataforma de código abierto ampliamente utilizada y que está disponible gratuitamente, aunque existen también muchas otras opciones disponibles.

[CKAN](http://www.ckan.org) es una herramienta concebida para facilitar la publicación de catálogos de datos[[2]](#footnote-2). Se podría comparar con un sistema de gestión de contenido tradicional (CMS), pero orientado a la publicación de datos en lugar de otro tipo de contenidos. Nos permitirá publicar y administrar colecciones de datos con muy pocos recursos y, una vez publicados, los usuarios podrán explorarlos e incluso pre-visualizarlos a través de mapas, gráficas y tablas.

**5.1. Requisitos previos**

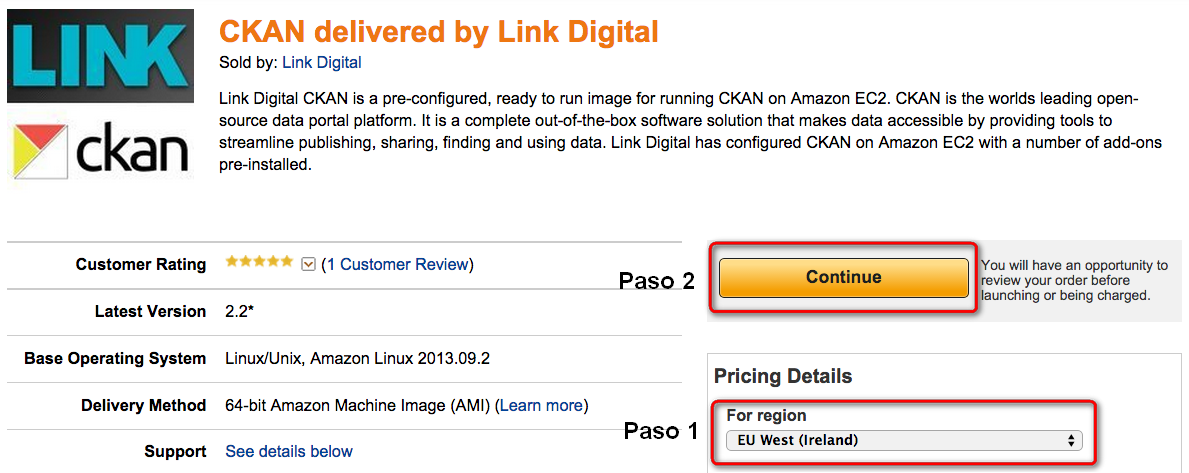
CKAN puede instalarse en cualquier máquina con una capacidad normal. Como ejemplo de requisitos mínimos para una instalación básica podríamos establecer una máquina con doble núcleo, 4 GB de RAM y 60 GB de espacio disponible en disco.



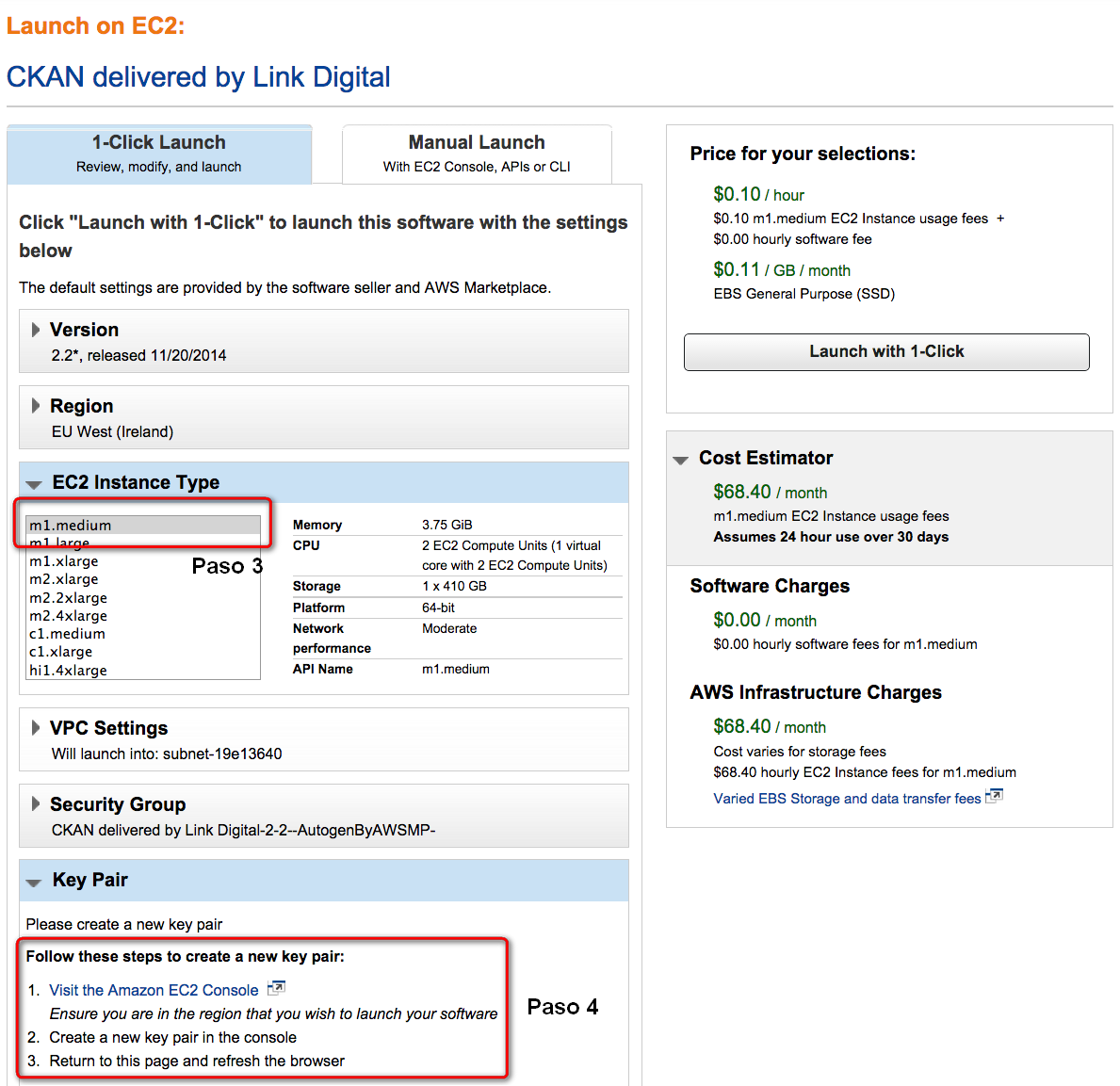
**5.2. Instalación básica**

Si nuestros conocimientos técnicos son limitados, la forma más rápida y sencilla de empezar a trabajar con nuestro catálogo de datos puede ser utilizar una versión pre-instalada de CKAN en un servicio en la nube como [Amazon Web Services](http://aws.amazon.com/es/) (AWS) con tan solo unos pocos pasos. Esta opción es sin duda la más simple, pero hay que tener en cuenta también que **conllevaría un coste por uso de los servicios de AWS desde 70$ mensuales[[3]](#footnote-3)** (para un uso de 24 horas al día/30 días al mes y la configuración básica que vamos a comentar).

Para ello una vez hayamos creado nuestra cuenta de usuario en AWS[[4]](#footnote-4), accederemos a la [versión pre-instalada](https://aws.amazon.com/marketplace/pp/B00JEF0278/ref=mkt_wir_ckan), elegiremos la región donde queremos desplegar la aplicación (*EU West–Ireland* para alojamiento de los datos en Europa) [Paso 1] y seleccionaremos la opción “Continuar” (*Continue*) [Paso 2].



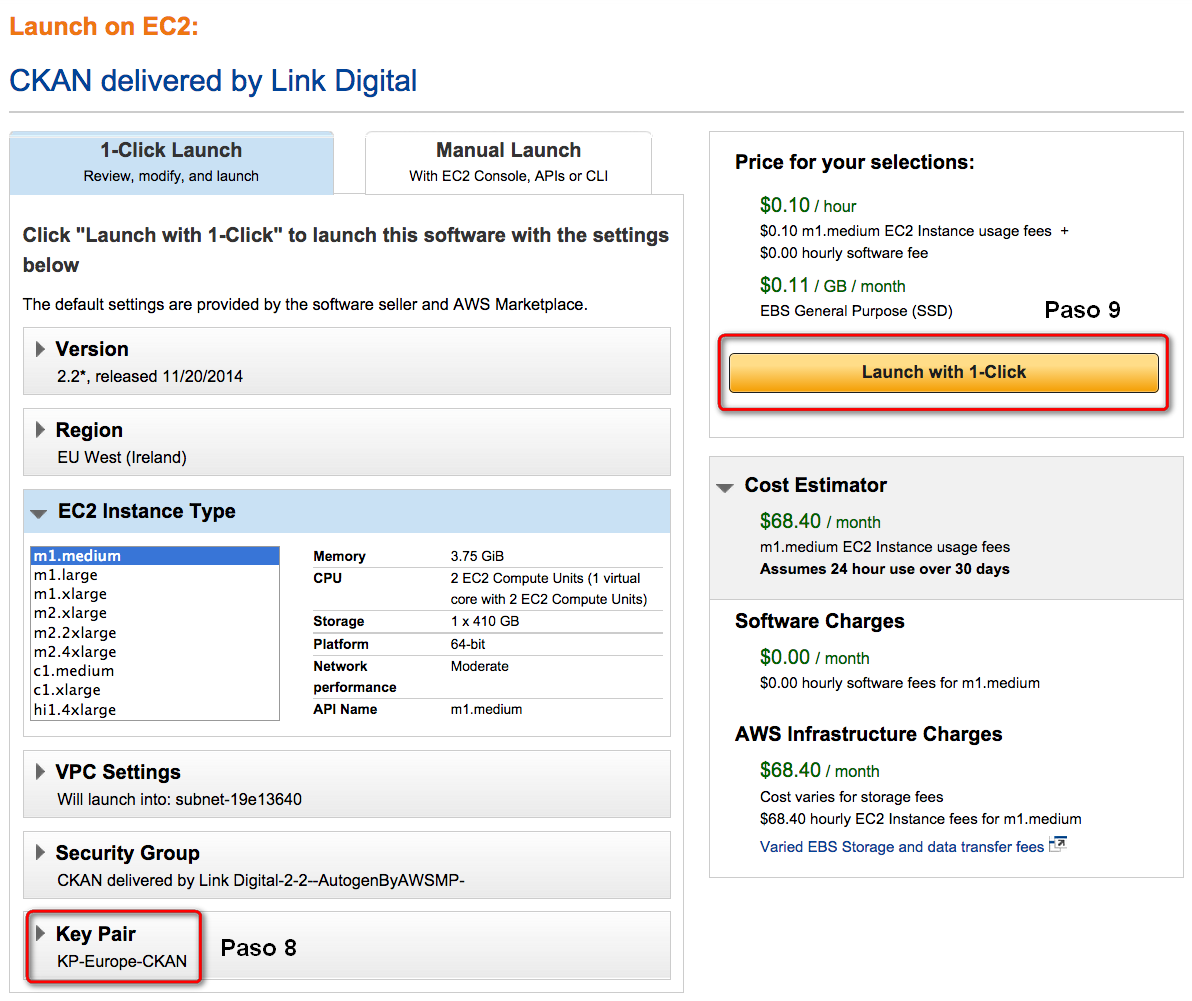
A continuación, escogeremos el tipo de servidor (*m1.medium* debería ser suficiente para empezar) [Paso 3] y tendremos que generar una nueva clave[[5]](#footnote-5) (*key pair*) para poder acceder luego al servidor. Usaremos entonces la opción “*Visit the Amazon EC2 Console*” [Paso 4] que abrirá la Consola de gestión en una ventana nueva.



Ya en la consola usaremos la opción “Create Key Pair” [Paso 5], introduciremos un nombre para la clave [Paso 6] y la generaremos con la opción “Create” [Paso 7]. 

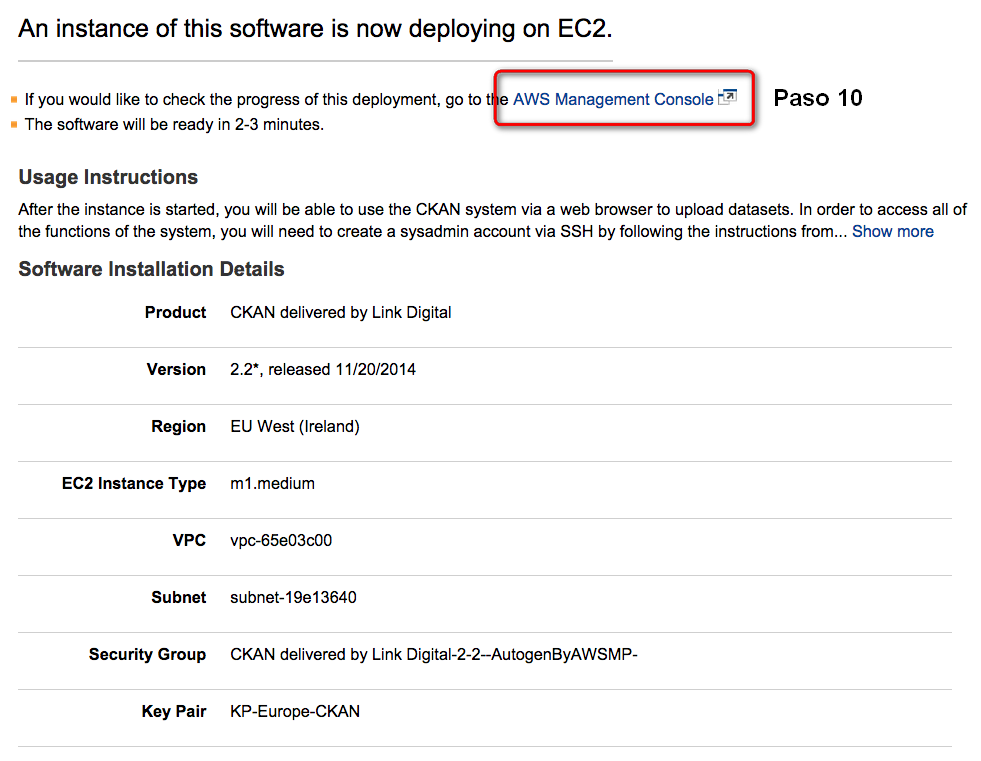
Se generará un archivo con el nombre de clave que hayamos definido y la extensión *.pem* que deberemos guardar y reservar en una localización segura.

Volveremos entonces a la página anterior y la recargaremos para comprobar que la clave ya está asignada [Paso 8] y podremos arrancar finalmente nuestro nuevo CKAN utilizando la opción “Launch with 1-Click” [Paso 9].

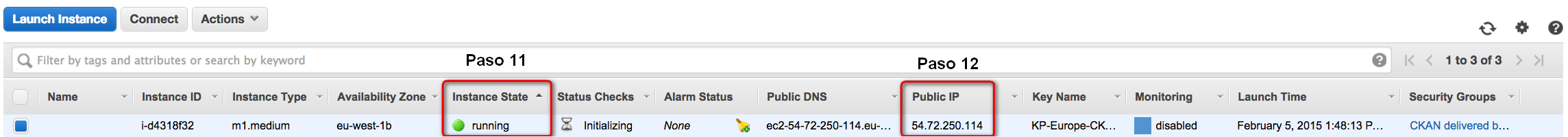


Inmediatamente aparecerá un mensaje indicándonos que la *instancia* de CKAN que acabamos de crear se está inicializando y en unos pocos minutos estará lista para ser utilizada.

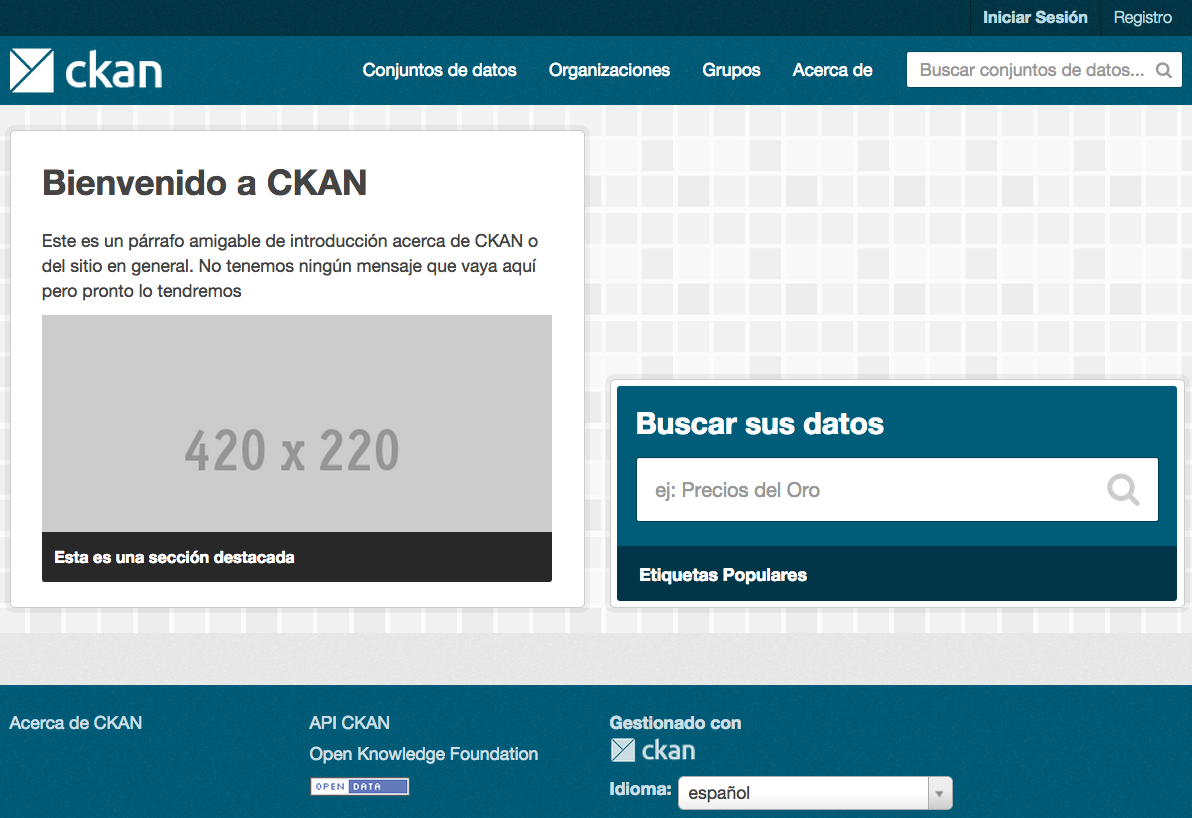
Podemos comprobar su estado en todo momento a través del acceso al panel de control [Paso 10].



Ya en el panel de control podremos ver cuándo CKAN está ya en marcha (el estado pasará a ser “*running*”) [Paso 11] y verificar que funciona correctamente accediendo a través de un navegador web utilizando de la IP pública que tiene asignada [Paso 12].



Si todo ha salido correctamente esto es lo que veremos en nuestro navegador al introducir la IP asignada en la barra de direcciones.



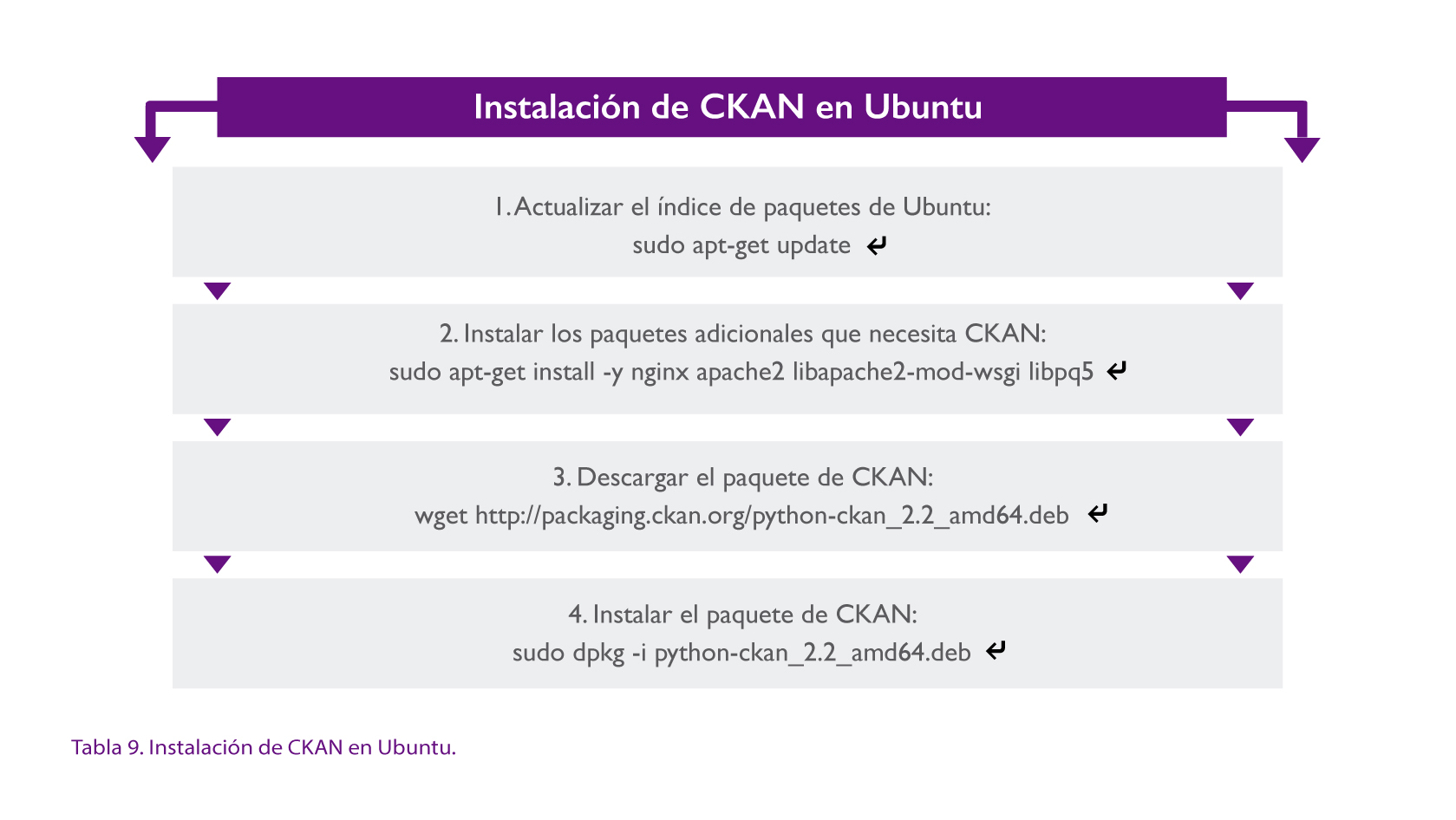
Ahora podríamos ya registrarnos directamente y empezar a utilizar CKAN como usuarios para publicar datos.

**5.3. Instalación avanzada**

Si contamos con unos conocimientos técnicos más amplios y queremos optar por una instalación más personalizada, la más cómoda entre las [opciones disponibles](http://docs.ckan.org/en/latest/maintaining/installing/index.html) sería utilizar la versión empaquetada para [servidores Ubuntu 12.04, 64-bit](http://www.ubuntu.com/download/server/install-ubuntu-server). Aquí contaríamos con dos alternativas respecto al *hardware* a utilizar:

1. Nuevamente podríamos optar por utilizar un servicio en la nube como AWS, aunque en esta ocasión partiríamos simplemente de una [preinstalación de Ubuntu](https://aws.amazon.com/marketplace/pp/B007Z5YWX4) y necesitaríamos crear un grupo de seguridad (*security group*) con permisos 0.0.0.0 para los puertos 22, 80 y 5000. Contaríamos además con la ventaja añadida de poder seleccionar servidores más asequibles para realizar nuestras pruebas (*m1.small por unos 35 dólares mensuales*[[6]](#footnote-6)) y, además, podríamos optar también por utilizar una [cuenta de prueba gratuita durante 1 año](http://aws.amazon.com/es/free/) disponible para los servidores más básicos[[7]](#footnote-7) (*t1.micro*).
2. La otra posible alternativa en este caso sería realizar la instalación en algún servidor propio donde pudiéramos instalar o tuviéramos ya instalada la versión de Ubuntu correspondiente.

En cualquier caso, una vez solucionadas las necesidades de hardware procederíamos a instalar CKAN en [cuatro sencillos pasos](http://docs.ckan.org/en/latest/maintaining/installing/install-from-package.html) desde la línea de comandos de nuestro Ubuntu ya instalado.



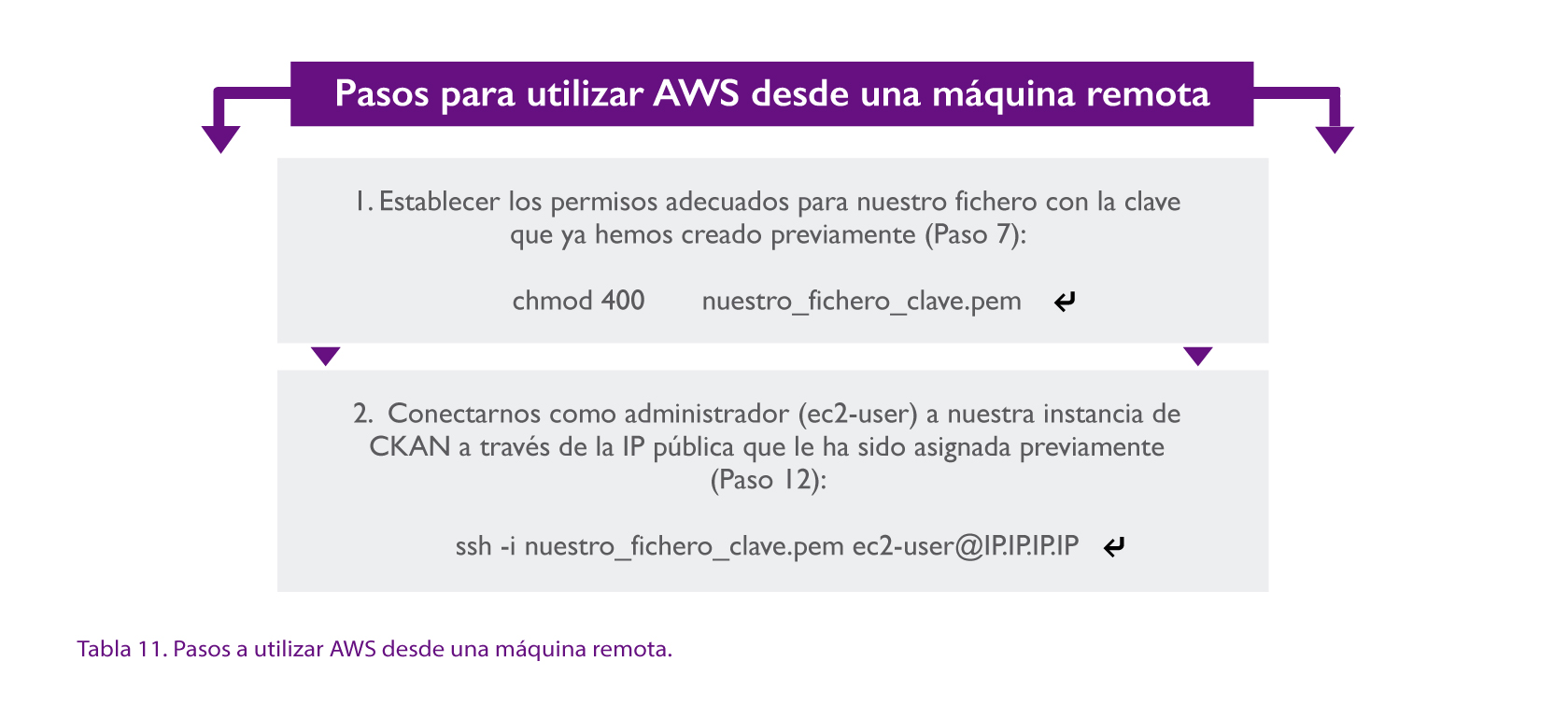
Además, debemos asegurarnos de que tanto la base de datos PostgreSQL como Solr están correctamente instalados y configurados.



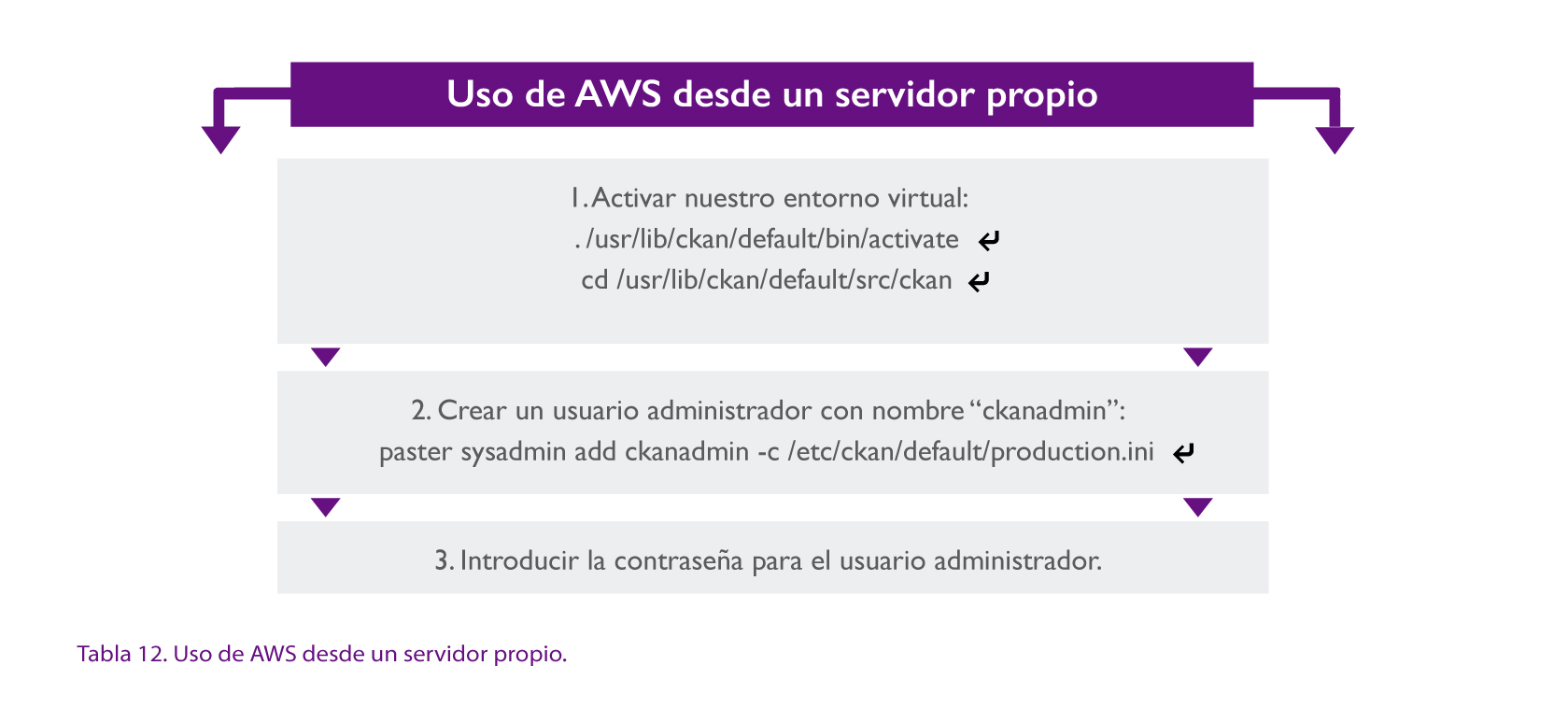
Si todos los pasos se han ejecutado correctamente ahora ya podremos acceder a nuestro CKAN a través de [*http://localhost*](http://localhost)o el DNS que le hayamos asignado.

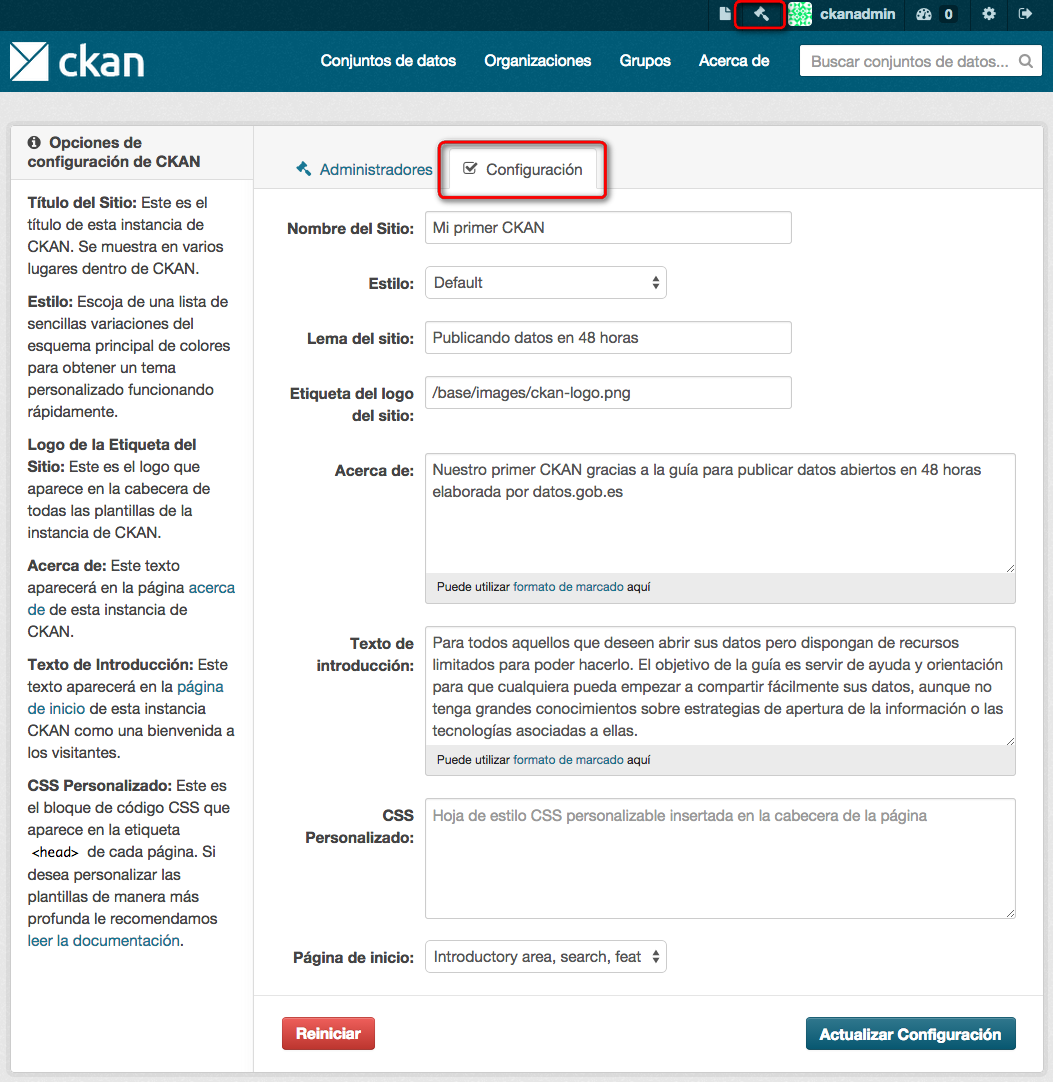
**5.4. Administración**

Una vez tenemos CKAN instalado y funcionando, el siguiente paso será crear un usuario administrador para poder realizar la configuración básica de nuestro catálogo. **Si estamos usando AWS** lo primero será conectarse a nuestra máquina remota siguiendo los siguientes pasos desde la línea de comandos.



Después, **tanto si estamos usando AWS o un servidor propio**, seguiremos los siguientes pasos nuevamente a través de la línea de comandos de Ubuntu.



Con esto ya podremos acceder a nuestro CKAN como administradores y cambiar la su configuración básica para adaptarla a nuestras preferencias.

Para ello, nos identificaremos a través de la opción “Iniciar Sesión” del menú superior y utilizaremos la opción con el icono del “martillo” (también en la parte superior) para acceder al menú de administración donde desde la opción “Configuración” podremos cambiar aspectos básicos como el nombre del sitio, el estilo de presentación, el lema y el tema de presentación. Ahora ya tenemos nuestro propio catálogo preparado y listo para empezar a compartir datos.

**5.5. Características avanzadas**

Para una personalización más avanzada de la apariencia del catálogo deberemos acudir a las recomendaciones de la [guía de personalización](http://docs.ckan.org/en/latest/theming/index.html) de CKAN y para conocer todas las posibilidades de configuración tenemos la [guía avanzada de administración](http://docs.ckan.org/en/latest/maintaining/configuration.html).

Además, CKAN cuenta con un amplio [catálogo de extensiones](http://extensions.ckan.org/) que nos permitirá ampliar considerablemente su funcionalidad de una forma limpia y sencilla. Incluso los usuarios más avanzados podrán crear sus propias extensiones para añadir nueva funcionalidad.

También disponemos de una completa [interfaz de programación](http://docs.ckan.org/en/latest/api/index.html) (API) que facilitará la creación de aplicaciones y servicios a partir de los datos publicados.

***5.5.1. Compatibilidad con la NTI***

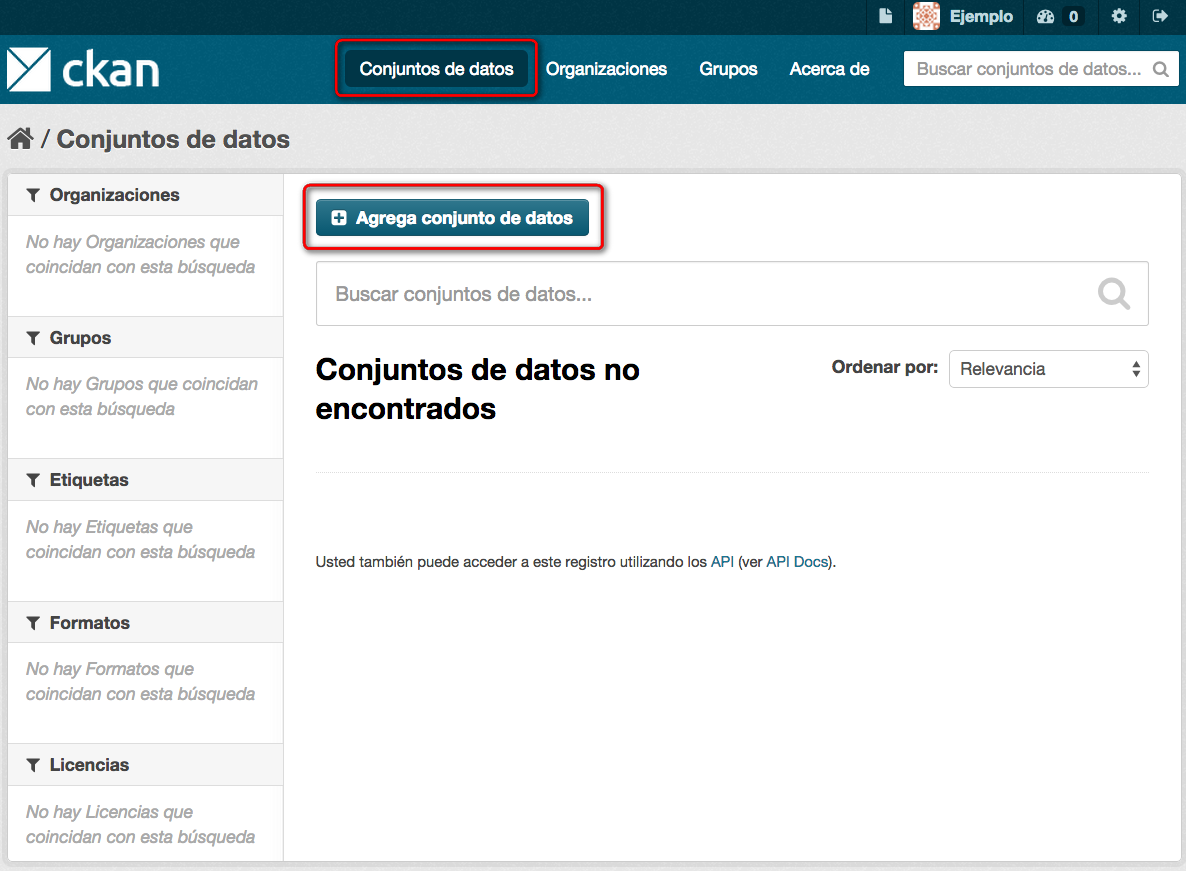
En este punto es importante también llamar especialmente la atención sobre la [Norma Técnica de Interoperabilidad para la Reutilización de recursos de información](http://www.boe.es/boe/dias/2013/03/04/pdfs/BOE-A-2013-2380.pdf), ya que es la normativa nacional de referencia a la hora de garantizar la interoperabilidad y facilidad de uso de los datos que publicamos para su reutilización.

Existe el inconveniente de que CKAN presenta varios [problemas de compatibilidad con la NTI](http://opendata.aragon.es/public/documentos/AnexoI_Analisis_Metadatos_Aragon_OpenData_v31-01-14.pdf)[[8]](#footnote-8), por lo que no sería posible cumplir completamente dicha norma sólo con una instalación de CKAN por defecto como la que hemos realizado en esta guía. Para tratar de paliar estos problemas en la medida de lo posible y con el mínimo esfuerzo podríamos hacer uso de las mejoras que otros ya han realizado previamente, como por ejemplo:

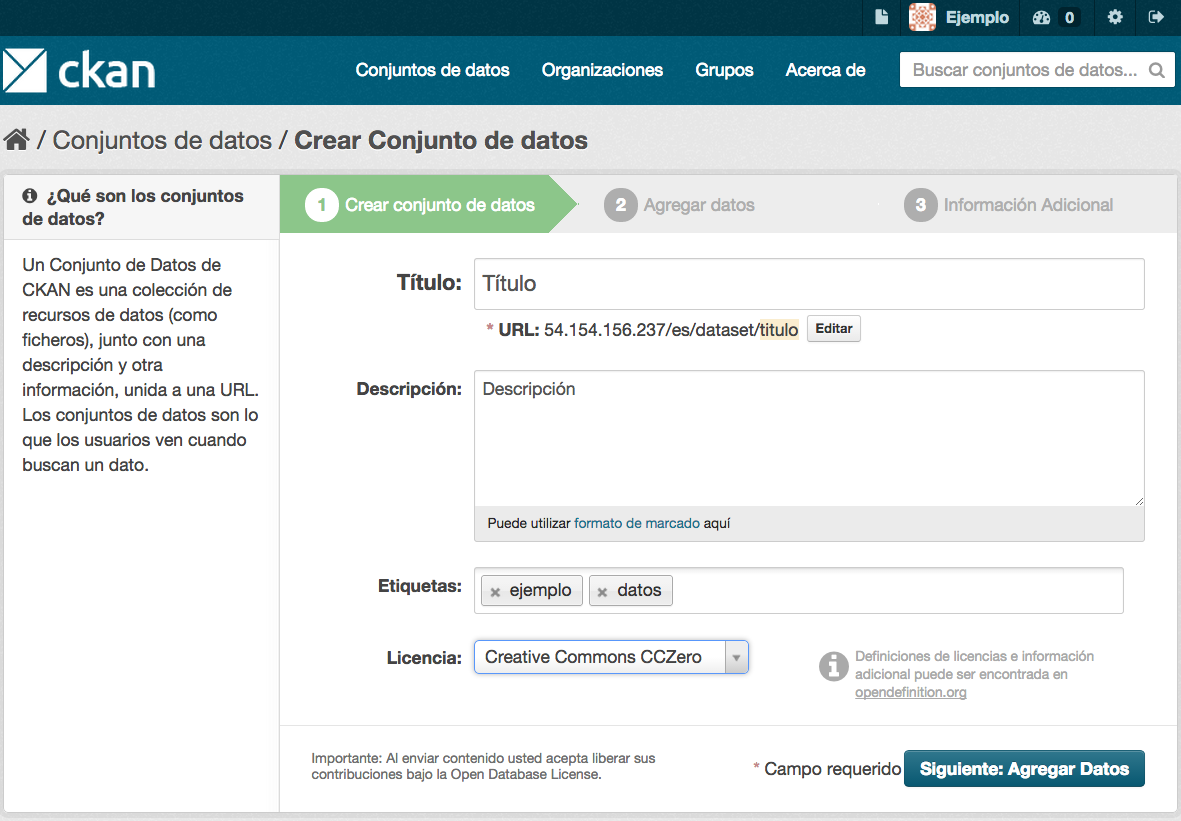
* Instalar la extensión [CKAN+DCAT](http://extensions.ckan.org/extension/dcat/) que mejora considerablemente el soporte de [DCAT](http://www.w3.org/TR/vocab-dcat/) en la plataforma (estándar de metadatos del W3C para facilitar la interoperabilidad de catálogos en el que se basa la NTI).
* Instalar la extensión desarrollada por el portal datos.gob.es [ckanext-dge-scheming](https://github.com/ctt-gob-es/datos.gob.es), que amplia los metadatos del catálogo CKAN a los necesarios para cumplir con la NTI. Para que esta extensión funcione es necesario instalar también [ckanext-scheming](https://github.com/ckan/ckanext-scheming).
* Reutilizar el código de alguna de las extensiones y mejoras realizadas por otras iniciativas previas de datos abiertos en nuestro país para mejorar la compatibilidad con la NTI y facilitar la *federación* de los datos de nuestra plataforma con el catálogo nacional de datos [datos.gob.es](http://datos.gob.es/):
  + [Extensión de Opendata Cáceres](https://github.com/ontoquercus/ckanext-federgob) para la federación con datos.gob.es
  + [Extensión del portal de datos abiertos del Ayuntamiento de Málaga](https://github.com/damalaga/ckanext-federador).
* [Personalizar los formularios](http://docs.ckan.org/en/latest/extensions/adding-custom-fields.html) de metadatos que se utilizan para introducir la información sobre los conjuntos de datos y poder así adaptarlos a las necesidades de la norma.

**5.6. Carga de datos**

Ahora que ya tenemos todo preparado, tan sólo queda empezar a añadir datos a nuestro catálogo. Lo primero que debemos hacer es registrarnos con nuestro usuario, seleccionar la opción “Conjuntos de Datos” del menú superior y luego la opción “Agrega conjunto de datos”.



A continuación, podemos empezar a crear nuestro primer conjunto de datos para compartir añadiendo la información correspondiente a los metadatos del “Título”, la “Descripción”, las “Etiquetas” y la “Licencia”.



Una vez hayamos terminado accederemos al siguiente paso utilizando la opción “Siguiente: Agregar Datos” de la esquina inferior derecha.

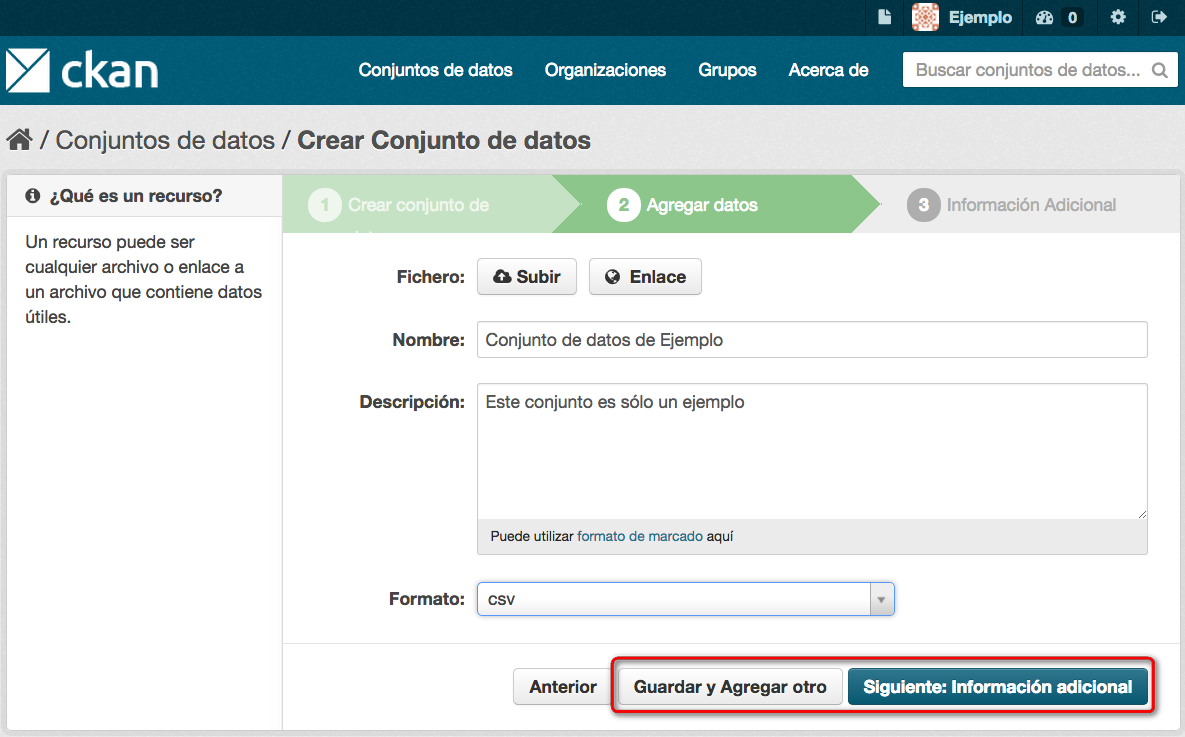
Ahora podremos ya elegir entre “Subir” directamente un fichero con los datos o añadir un “Enlace” directo a los datos si están ya disponibles en una dirección web.



Si elegimos, por ejemplo, añadir un “Enlace”, simplemente tendremos que incluir la dirección web del sitio donde estén alojados los datos.

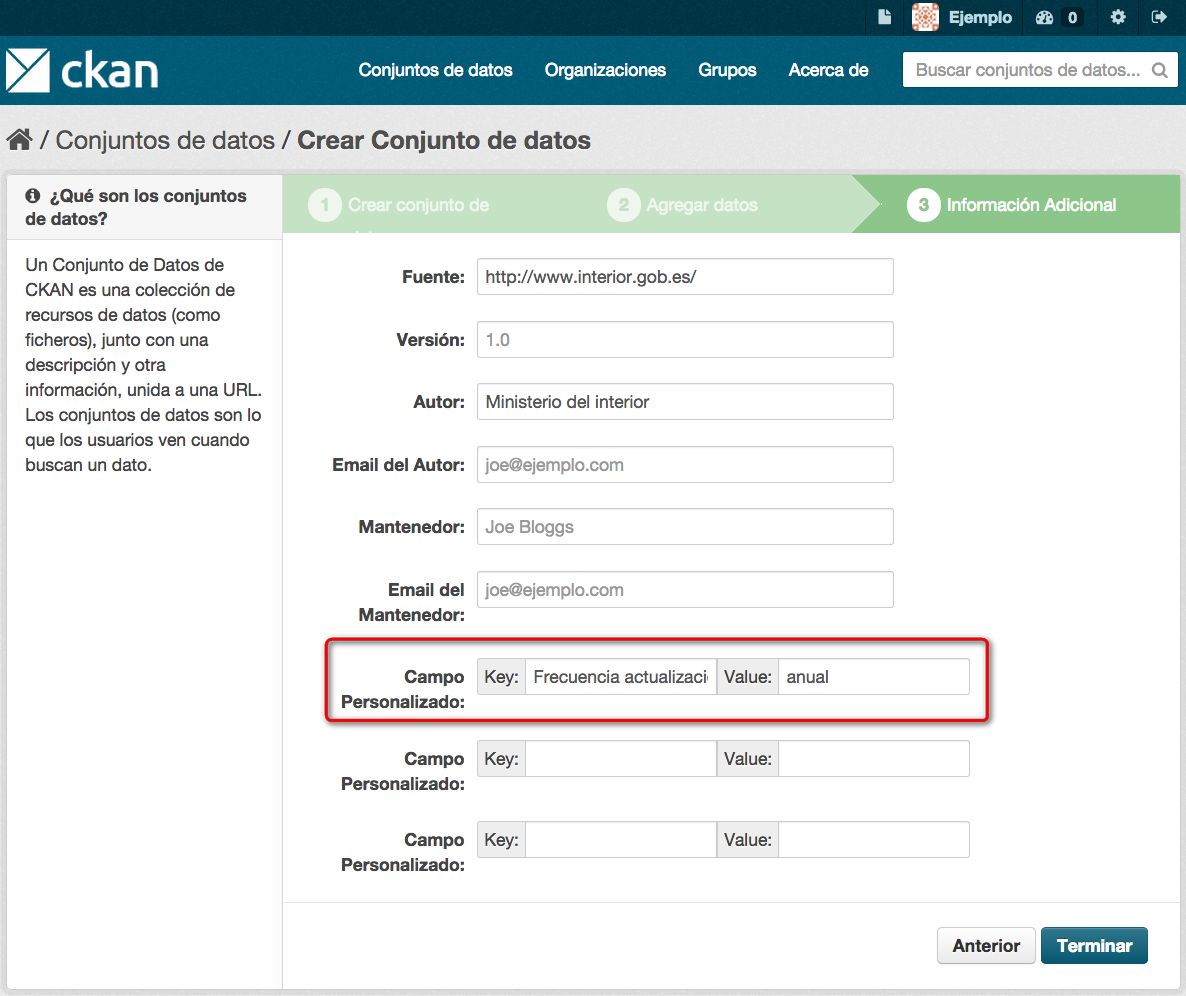
Por último, completaríamos con el “Nombre” y la “Descripción”, así como la extensión del formato en el que están guardados los datos (CSV, JSON, XML).

Una vez hemos finalizado podemos elegir entre guardar estos datos y “Agregar” más datos a este mismo conjunto o terminar de añadir datos y pasar a proporcionar algo más de “Información adicional”.



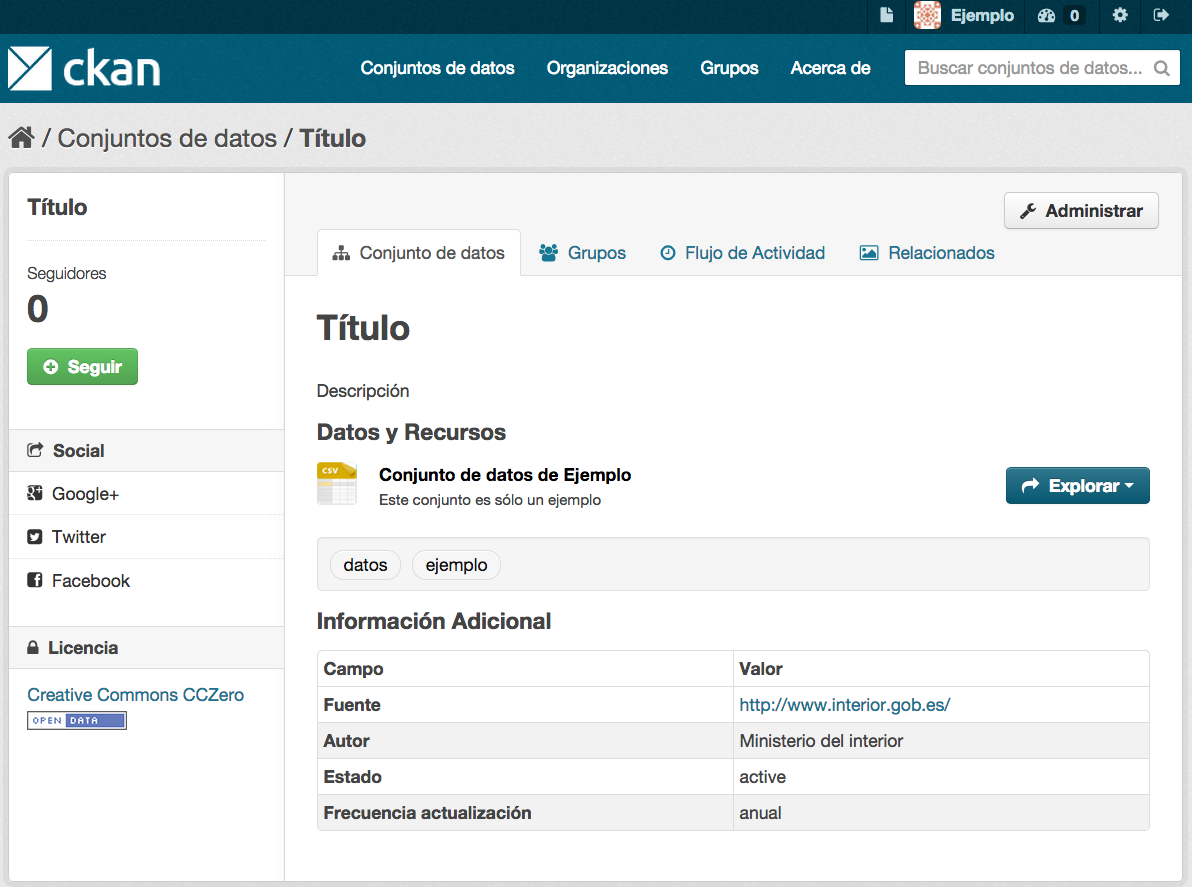
Cuando hemos terminado de añadir todos los datos tan solo nos quedará introducir algo más de información adicional, como la “Fuente” original de los datos, la “Versión” el “Autor” o el “email” de contacto del autor, entre otras cosas.

También podremos añadir otros “Campos Personalizados” que podemos aprovechar para incluir otros metadatos que necesitemos (por ejemplo porque así lo indique la normativa) y no estén recogidos entre los anteriores. En este caso deberemos añadir tanto el nombre del metadato (Key) como su valor (Value).



Este mismo proceso lo repetiremos tantas veces como sea necesario hasta que hayamos introducido todos los conjuntos de datos que tengamos previsto.

A partir de este momento los datos estarán también disponibles a través del catálogo para todo aquel que quiera usarlos y experimentar con las distintas opciones de clasificación y visualización con las que cuenta la plataforma.

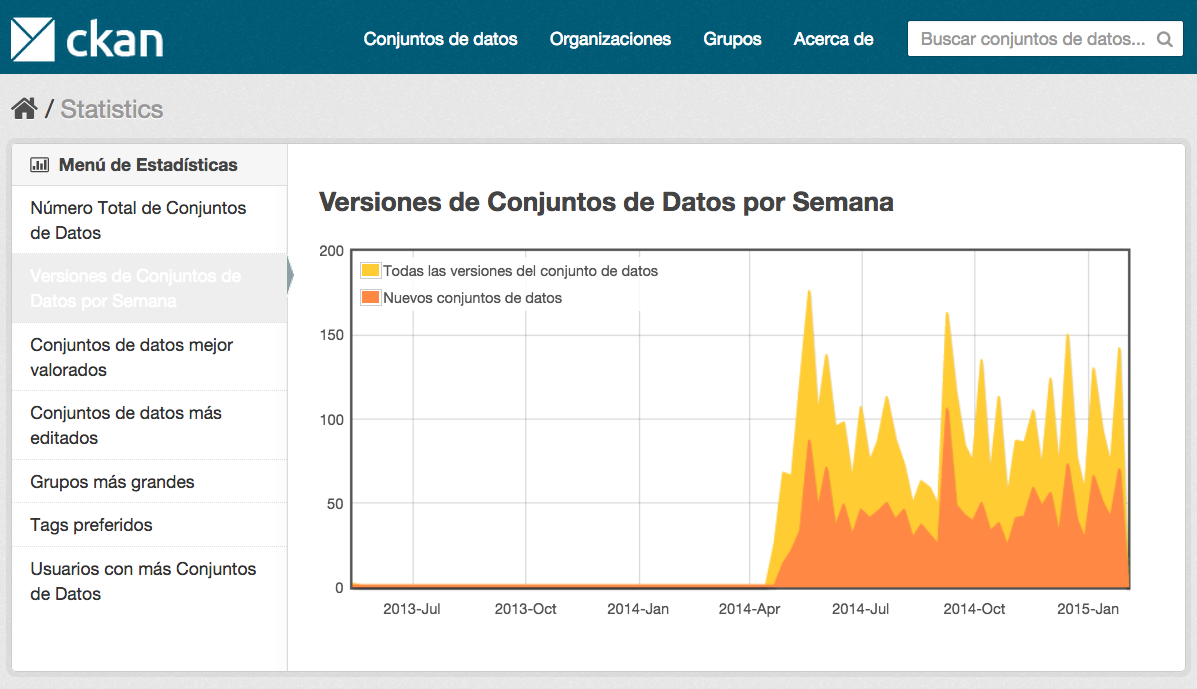


Para conocer más posibilidades sobre la gestión de conjuntos de datos u organizaciones se puede consultar la [guía de usuario](http://docs.ckan.org/en/latest/user-guide.html#using-ckan) de CKAN.

**6. EVALUACIÓN DE USO**

Una vez hemos publicado nuestros datos es interesante también tener cierta **medida del uso** que se les está dando para saber cuáles pueden ser los más interesantes desde el punto de vista de quienes los van a reutilizar. Esto nos ayudará también a definir mejor nuestros criterios de selección y priorización en la publicación de datos. El propio CKAN nos proporciona algunas herramientas que pueden también ser de utilidad para este objetivo, como:

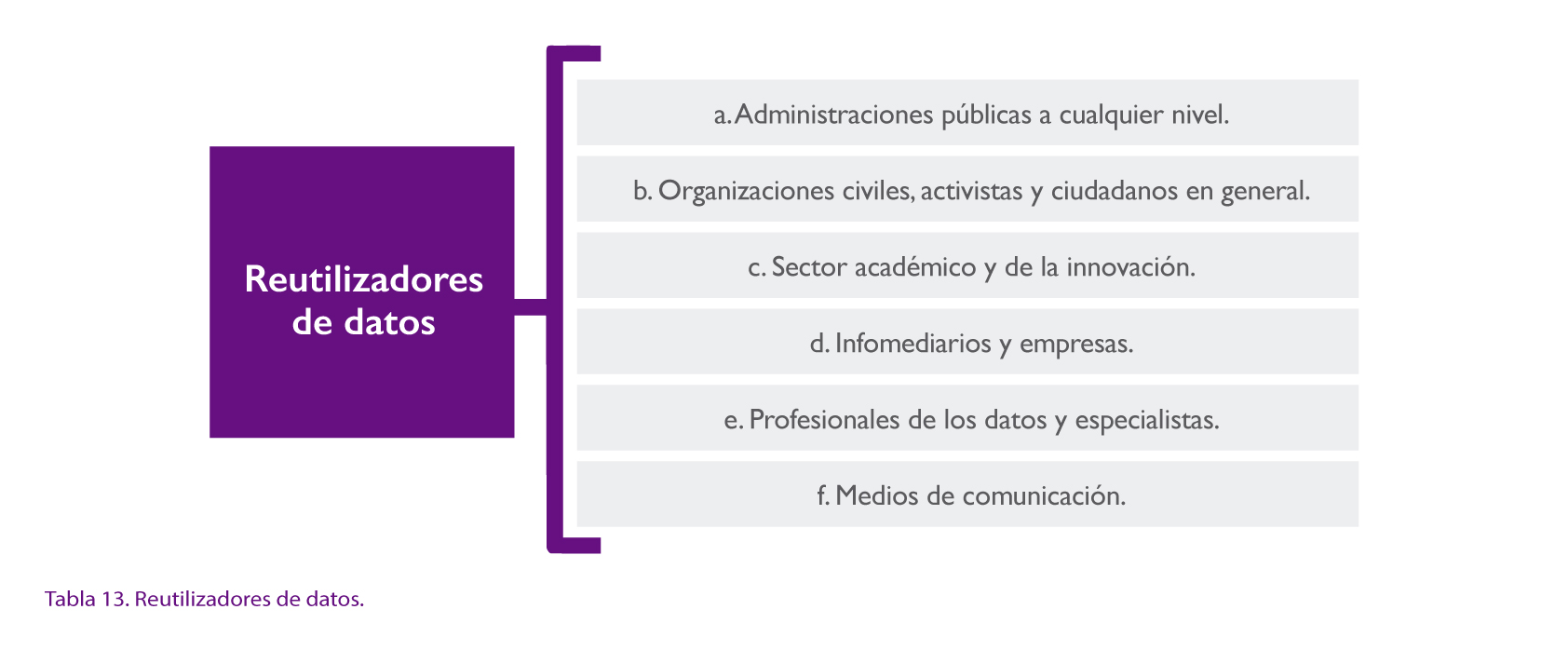
* La [extensión de estadísticas](http://docs.ckan.org/en/latest/maintaining/stats.html) del catálogo, accesible añadiendo */stats* a la dirección de nuestro catálogo, que nos permitirá conocer entre otros:
  + El número de conjuntos de datos publicados.
  + El número de actualizaciones semanales.
  + Los conjuntos de datos mejor valorados.
  + Los conjuntos de datos con más actualizaciones.



* La [extensión de Google Analytics](https://github.com/ckan/ckanext-googleanalytics) para CKAN, que nos ayudará a integrar ambos sistemas y contar con distintos indicadores como número de visitas, descargas, acceso a la API o distintas analíticas de datos, entre otros.
* El [monitor de visitas](http://docs.ckan.org/en/latest/maintaining/tracking.html) a páginas que, una vez activado, nos ayudará a:
  + Ordenar los conjuntos de datos por popularidad.
  + Destacar conjuntos de datos y recursos populares.
  + Mostrar contadores de visitas para conjuntos de datos y recursos.
  + Exportar los datos para su análisis.

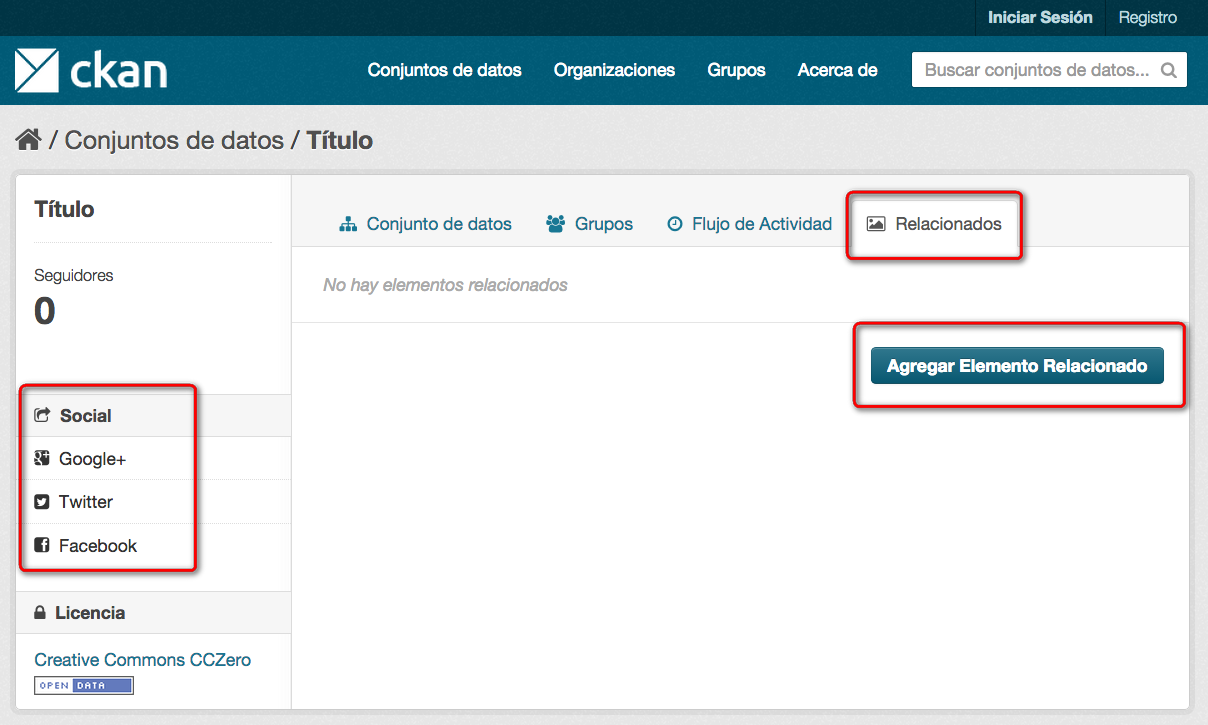
**7. DIÁLOGO CON LA COMUNIDAD**

Además de los indicadores de evaluación exclusivamente cuantitativos que acabamos de ver, no debemos olvidar que para ser capaces de comprender la demanda de datos en profundidad deberíamos también establecer un diálogo continuo con la **comunidad de *reutilizadores*** de los datos.



Nuevamente, CKAN nos ofrece algunas herramientas interesantes para ayudarnos con el diálogo y la comunicación, entre las que podemos destacar:

* Las funciones que proporciona para compartir conjuntos de datos en las redes sociales más populares (*Google+*, *Twitter* y *Facebook*).
* La posibilidad de comentar y discutir sobre los conjuntos de datos gracias a extensiones como [disqus](https://github.com/ckan/ckanext-disqus)[[9]](#footnote-9) o [comments](https://github.com/rossjones/ckanext-comments).
* La opción de [compartir ideas](http://docs.ckan.org/en/latest/maintaining/apps-ideas.html), aplicaciones o cualquier otro contenido relacionado con los datos publicados.



Sin embargo, conviene también recordar que, para poder llevar a cabo los objetivos que se planteen, hará falta combinar estas herramientas *online* con otras actuaciones *offline* necesarias para **facilitar la comunicación y colaboración continua**, así como para **conocer de primera mano sus necesidades** y obtener *feedback* útil a través de otros **indicadores cualitativos** como, por ejemplo, *cuestionarios, entrevistas* y *grupos de trabajo.*

**8. MANTENIMIENTO**

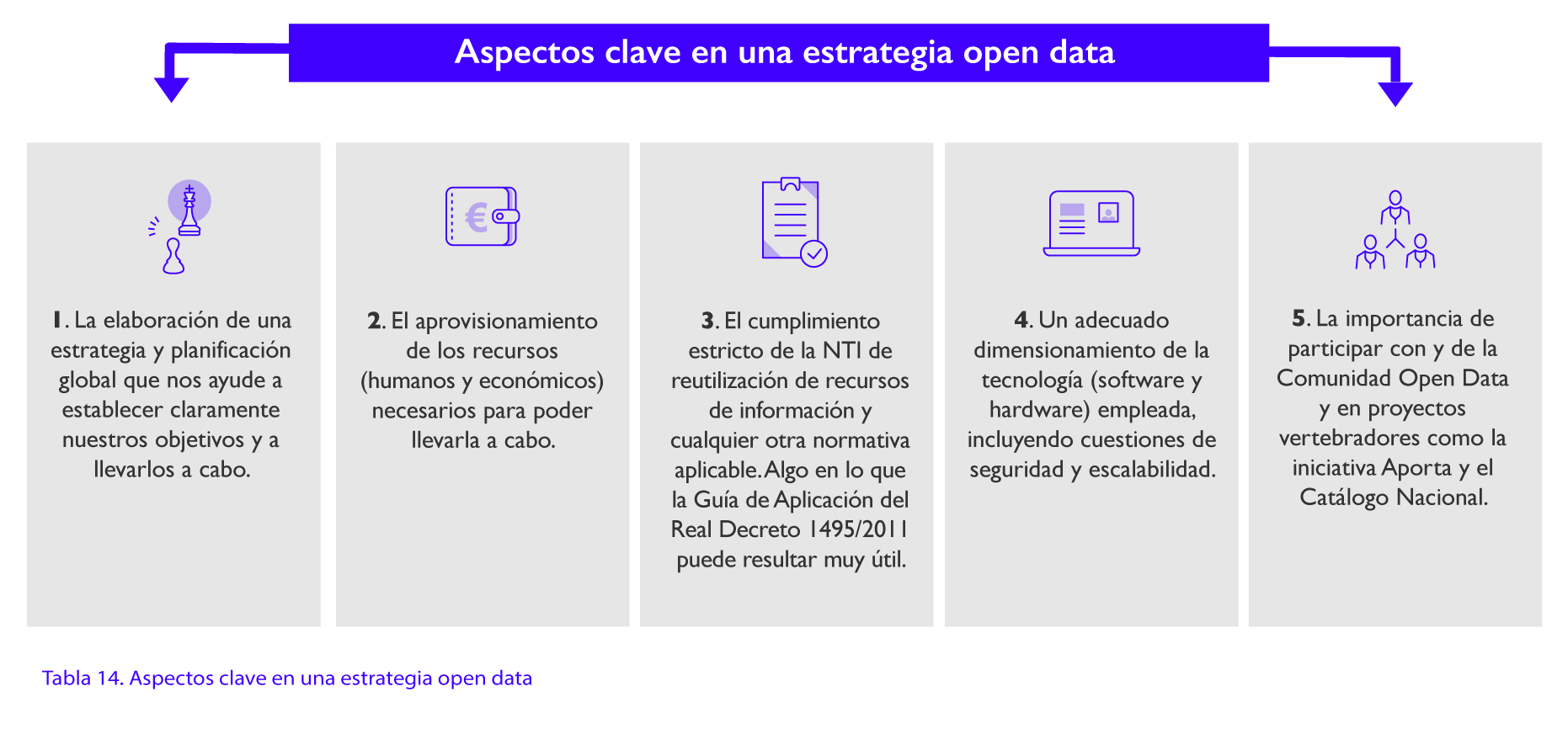
La carga inicial de los datos no es más que el primer paso en la apertura de la información. La incorporación de nuevos conjuntos de datos debería ser una constante a partir de ese momento, y especialmente la actualización frecuente de los conjuntos de datos que ya hayan sido incorporados a la plataforma. Para ello tendremos en cuenta que contamos con dos tipos de conjuntos de datos:

* **Estáticos**: Cuando los datos están disponibles a través de **archivos** en cualquier formato. Esta opción requerirá que alguien se encargue de su mantenimiento actualizando el conjunto de datos de forma manual *subiendo* un nuevo archivo de forma periódica con la frecuencia que se haya establecido en cada caso.
* **Dinámicos**: Cuando se pueda establecer una **conexión directa** con la fuente original desde donde se obtienen los datos a través de una dirección web (URL) donde se han publicado. En este caso cada vez que los datos se actualicen se reflejará directamente en la plataforma de catalogación sin necesidad de intervención alguna.

**9. EVOLUCIÓN**

Para finalizar, debe quedar claro que el objetivo de esta guía es principalmente ayudar a todos aquellos que cuenten con pocos recursos y conocimientos limitados a dar sus primeros pasos y experimentar con la publicación de datos abiertos. Las indicaciones aquí recogidas servirán para poder experimentar fácilmente en muy poco tiempo y sin apenas necesidad de inversión.

No obstante, si se quiere apostar por una estrategia continuada de apertura de datos a medio y largo plazo, será necesario pensar en llevar a cabo un proyecto más complejo prestando especial atención a los siguientes puntos:



**10. REFERENCIAS**

Aragón Open Data, Guía de catalogación de datos abiertos.

[**http://opendata.aragon.es/public/documentos/Guia\_Gestion\_v1.0.pdf**](http://opendata.aragon.es/public/documentos/Guia_Gestion_v1.0.pdf)

EU, Measuring European Public Sector Information Resources.

[**http://ec.europa.eu/information\_society/newsroom/cf/document.cfm?doc\_id=1198**](http://ec.europa.eu/information_society/newsroom/cf/document.cfm?doc_id=1198)

Datos.gob.es, Guía de aplicación del Real Decreto 1495/2011.

<https://datos.gob.es/es/documentacion/guia-de-aplicacion-del-real-decreto-14952011>

ONTSI, Estudio de caracterización del sector infomediario en España.

[**http://www.ontsi.red.es/ontsi/es/estudios-informes/estudio-de-caracterizaci%C3%B3n-del-sector-infomediario-en-espa%C3%B1-edici%C3%B3n-2012**](http://www.ontsi.red.es/ontsi/es/estudios-informes/estudio-de-caracterizaci%C3%B3n-del-sector-infomediario-en-espa%C3%B1-edici%C3%B3n-2012)

UK Cabinet Office, G8 Open Data Charter and Technical Annex.

[**https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex**](https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex)

EC, Guidelines on recommended standard licenses, datasets and charging.

[**http://ec.europa.eu/information\_society/newsroom/cf/dae/document.cfm?action=display&doc\_id=6421**](http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?action=display&doc_id=6421)

Open Data Barometer, OGD Implementation: data availability.

[**http://opendatabarometer.org/report/analysis/implementation.html**](http://opendatabarometer.org/report/analysis/implementation.html)

BOE, Real Decreto 1495/2011, por el que se desarrolla la Ley 37/2007 sobre reutilización de la información del sector público.

[**http://www.boe.es/diario\_boe/txt.php?id=BOE-A-2011-17560**](http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-17560)

Open Definition, Conformance licenses.

[**http://opendefinition.org/licenses/**](http://opendefinition.org/licenses/)

Creative Commons, About the Licenses.

[**http://creativecommons.org/licenses/**](http://creativecommons.org/licenses/)

Datos.gob.es, Normativa.

<http://datos.gob.es/es/documentacion?f%5B0%5D=field_doc_tx_type%3A316>

Datos.gob.es, Buenas prácticas de formatos para la publicación de conjuntos de datos.

<http://datos.gob.es/sites/default/files/doc/file/soporterisp-buenas_practicas_formatos_de_publicacion.docx>

Datos.gob.es, Guía metodológica para iniciativas de datos abiertos sectoriales.

<http://datos.gob.es/es/documentacion/guia-metodologica-para-iniciativas-de-datos-abiertos-sectoriales>

Datos.gob.es, Informe de Tendencias e Iniciativas de Datos Abiertos.

<http://datos.gob.es/es/documentacion/informe-de-tendencias-e-iniciativas-de-datos-abiertos>

Amazon, Amazon Web Services.

[**http://aws.amazon.com/**](http://aws.amazon.com/)

CKAN, User guide.

[**http://docs.ckan.org/en/latest/user-guide.html**](http://docs.ckan.org/en/latest/user-guide.html)

BOE, Norma Técnica de Interoperabilidad de Reutilización de recursos de información.

[**http://www.boe.es/boe/dias/2013/03/04/pdfs/BOE-A-2013-2380.pdf**](http://www.boe.es/boe/dias/2013/03/04/pdfs/BOE-A-2013-2380.pdf)

Datos.gob.es, Guía de aplicación de la NTI de Reutilización de recursos de información.

[**http://datos.gob.es/content/guia-de-aplicacion-de-norma-tecnica-de-interoperabilidad-de-reutilizacion-de-recursos-de**](http://datos.gob.es/content/guia-de-aplicacion-de-norma-tecnica-de-interoperabilidad-de-reutilizacion-de-recursos-de)

Aragón Open Data, Informe de conformidad con la NTI de Reutilización de recursos de información.

[**http://opendata.aragon.es/public/documentos/Informe\_NTI\_Aragon\_OpenData\_v31-01-14.pdf**](http://opendata.aragon.es/public/documentos/Informe_NTI_Aragon_OpenData_v31-01-14.pdf)

Wolters Kluwer, Monográfico especial Gobierno Abierto y Open Data

[**https://funkziuni.files.wordpress.com/2014/07/aaj\_7-8\_2014\_br.pdf**](https://funkziuni.files.wordpress.com/2014/07/aaj_7-8_2014_br.pdf)

**11. TABLAS Y GRÁFICOS**

Tabla 1. Criterios generales para la publicación de datos.

Tabla 2. Ámbitos de información.

Tabla 3. Limitaciones a la hora de publicar los datos.

Tabla 4. Condiciones fijadas en el Real Decreto 1495/2011 de 24 de octubre.

Tabla 5. Formatos para la apertura de información.

Tabla 6. Tipología de metadatos.

Tabla 7. Tabla de resumen de metadatos.

Tabla 8. Requisitos de software para utilizar CKAN.

Tabla 9. Instalación de CKAN en Ubuntu.

Tabla 10. Configuración de Postgre SQL y Solp.

Tabla 11. Pasos a utilizar AWS desde una máquina remota.

Tabla 12. Uso de AWS desde un servidor propio.

Tabla 13. Reutilizadores de datos.

Tabla 14. Aspectos clave en una estrategia open data.

1. La guía ha sido diseñada para ser útil incluso cuando no se tengan conocimientos técnicos, no obstante será mucho más fácil seguirla si se cuentan ya con ciertos conocimientos básicos. [↑](#footnote-ref-1)
2. Una versión de demostración de CKAN está disponible en <http://demo.ckan.org/es/>. [↑](#footnote-ref-2)
3. Para un cálculo personalizado utilizar <http://calculator.s3.amazonaws.com/index.html>. [↑](#footnote-ref-3)
4. Para crear la cuenta será necesario proporcionar un número de tarjeta de crédito. [↑](#footnote-ref-4)
5. En el caso de que ya seamos usuarios de AWS y tengamos alguna clave ya generada podremos simplemente seleccionarla en lugar de crear una nueva. [↑](#footnote-ref-5)
6. Para un cálculo personalizado utilizar http://calculator.s3.amazonaws.com/index.html [↑](#footnote-ref-6)
7. Los servidores básicos tienen unas características muy limitadas y es probable que no sirvan para cualquier cosa que no sean pruebas. [↑](#footnote-ref-7)
8. En esta comparativa se incluye un análisis del nivel de implementación de la NTI en la versión 2.2a de CKAN en el que se ve claramente que la compatibilidad es baja. [↑](#footnote-ref-8)
9. A través del servicio externo proporcionado por <https://disqus.com/> [↑](#footnote-ref-9)