



**GRADO EN INFORMACIÓN Y DOCUMENTACIÓN**

**FACULTAD DE FILOSOFÍA Y LETRAS**

**UNIVERSIDAD DE LEÓN**

**Curso Académico 2019/2020**

**LA CALIDAD DE LOS DATOS ABIERTOS:  
ANÁLISIS DEL PORTAL DE DATOS  
ABIERTOS DE LA JUNTA DE  
CASTILLA Y LEÓN**

**OPEN DATA PORTALS QUALITY:  
ANALYSIS OF THE GOVERNING COUNCIL  
OF CASTILE-LEON OPEN DATA PORTAL**

**Alumna: M<sup>a</sup> Lourdes Calavia Álvarez**

**Tutora: M<sup>a</sup> Luisa Alvite Díez**

## **Resumen**

El presente trabajo tiene como objetivo analizar la calidad del portal de datos abiertos de la Junta de Castilla y León mediante un marco de evaluación de elaboración propia. Este marco sintetiza diferentes modelos evaluativos propuestos en la literatura y tiene en cuenta las recomendaciones de la propia Junta y de la Federación Española de Municipios y Provincias (FEMP) respecto a los elementos básicos y recomendables de un portal de datos abiertos. Los resultados obtenidos del análisis de una muestra de datasets del portal indican que tanto el portal como los datos ofrecidos son de calidad y representan un buen ejemplo de la aplicación de los principios y estándares de datos abiertos a la Administración. No obstante, se han detectado debilidades que pueden afectar a dicha calidad: excesiva sencillez del buscador del portal, falta de consistencia en las palabras clave, cierto desaprovechamiento del *feedback* de los usuarios y de las redes sociales, falta de programación en la actualización de algunos datasets, existencia de celdas vacías y valores nulos, carencias en metadatos, y dispersión de la información relativa a la rendición de cuentas, entre otros aspectos. Finalmente, se presentan diversas mejoras para solventar las debilidades detectadas y reforzar la calidad del portal.

## **Palabras clave**

Calidad, Datos abiertos, Evaluación, Gobierno abierto, Portal de datos abiertos de la Junta de Castilla y León, Portales de datos abiertos, Rendición de cuentas.

## **Abstract**

This work aims to analyse the quality of the open data portal of the Governing Council of Castile-Leon, using an evaluation framework made by the author. This framework synthesise different evaluative models proposed in literature, and takes account of suggestions of the Governing Council itself and the Spanish Federation of Municipalities and Provinces (FEMP) about an open data portal basic and recommendable elements. The results obtained from the analysis of a sample of datasets shows that both the portal and offered data are of adequate quality and represent a good example of the implementation of open data principles and standards to Administration. However, some issues have been detected and can affect such quality: excessive simplicity of browser, lack of consistency in keywords, wasting of feedback from users and social media, lack of scheduling of some datasets update, empty cells and null values, failings in metadata and dispersion of accountability information, among other aspects. Finally, the study presents several proposals to resolve detected issues and reinforce portal quality.

## **Keywords**

Accountability, Evaluation, Governing Council of Castile-Leon open data portal, Open data, Open data portals, Open Government, Quality.

## ÍNDICE

1. MARCO TEÓRICO .....	4
1.1. LOS DATOS ABIERTOS .....	4
1.2. DATOS ABIERTOS GUBERNAMENTALES .....	4
1.2.1. Datos abiertos en España .....	6
1.3. PROBLEMAS EN LA APERTURA DE DATOS .....	7
1.4. LA CALIDAD DE LOS DATOS ABIERTOS .....	8
1.4.1. Calidad reutilizadora .....	8
1.4.2. La calidad como concepto multidimensional .....	9
1.4.2.1. Los portales de datos abiertos .....	9
1.4.2.2. Indicadores de calidad de los datos abiertos .....	10
1.4.2.3. Criterios de calidad del contenido de los datos .....	17
2. JUSTIFICACIÓN, OBJETIVOS Y METODOLOGÍA .....	19
2.1. JUSTIFICACIÓN .....	19
2.2. OBJETIVOS .....	19
2.3. METODOLOGÍA .....	20
2.3.1. Elaboración del marco de evaluación .....	20
2.3.2. Selección de los datasets objeto de evaluación .....	22
3. ANÁLISIS DEL PORTAL DE DATOS ABIERTOS DE LA JUNTA DE CASTILLA Y LEÓN ...	23
3.1. ORIGEN Y NORMATIVA APLICABLE .....	23
3.2. ORGANIZACIÓN GENERAL DEL PORTAL .....	23
3.3. APLICACIÓN DEL MARCO DE EVALUACIÓN PROPUESTO .....	26
3.3.1. Características del portal .....	26
3.3.1.1. Aspectos técnicos .....	26
3.3.1.2. Usabilidad y acceso .....	28
3.3.1.3. Comunicación y participación .....	35
3.3.2. Características de los datasets .....	37
3.3.3. Contenido del portal en relación a la rendición de cuentas .....	55
4. CONCLUSIONES Y PROPUESTA DE MEJORAS .....	57
BIBLIOGRAFÍA .....	62
ANEXOS .....	68
ANEXO I. GLOSARIO .....	68
ANEXO II. MARCO DE EVALUACIÓN .....	70
ANEXO III. CONTENIDO DEL PORTAL EN RELACIÓN A LA RENDICIÓN DE CUENTAS .....	73

## INDICE DE FIGURAS

Figura 1. Página de inicio del Portal de Datos Abiertos de la JCYL.....	24
Figura 2. Menú del catálogo de datos .....	25
Figura 3. Portal de análisis de datos .....	26
Figura 4. Acceso al idioma inglés .....	27
Figura 5. Categorías temáticas en el catálogo de datos .....	30
Figura 6. Palabras clave del catálogo .....	31
Figura 7. Criterios de ordenación de los datasets .....	31
Figura 8. Buscador del portal de datos abiertos .....	32
Figura 9. Filtros del buscador del portal de datos abiertos .....	32
Figura 10. Filtros del portal de análisis .....	33
Figura 11. Cómo seguir un dataset .....	36
Figura 12. Ficha de "Límites provinciales" en el portal de datos .....	38
Figura 13. Ficha de "Límites provinciales" en el portal de análisis.....	38
Figura 14. Ficha de "Situación epidemiológica" en el portal de datos .....	39
Figura 15. Ficha de "Situación epidemiológica" en el portal de análisis.....	39
Figura 16. Ficha de "Situación solicitudes ERTes" en el portal de datos .....	40
Figura 17. Ficha de "Situación solicitudes ERTes" en el portal de análisis ..	40
Figura 18. Ficha de "Comercios con entrega" en el portal de datos.....	41
Figura 19. Ficha de "Comercios con entrega" en el portal de análisis .....	41
Figura 20. Ficha de "Eventos" en el portal de datos.....	42
Figura 21. Ficha de "Eventos" en el portal de análisis .....	42
Figura 22. Ficha de "Catálogo de datos" en el portal de datos.....	43
Figura 23. Temática y Categorías del dataset "Eventos" en portal y CSV....	51
Figura 24. Descripciones centros gestores de gastos .....	52
Figura 25. Dotaciones presupuesto ingresos 2018.....	53
Figura 26. URIs de los datasets de muestra .....	54
Figura 27. Ejecución presupuesto de la Administración de la Comunidad....	56
Figura 28. Publicación en Facebook sobre el dataset "Comercios" .....	60

## INDICE DE TABLAS

Tabla I. Características de calidad según ISO 25012 .....	11
Tabla II. Licencias para la publicación de datos abiertos .....	14
Tabla III. Marcos de evaluación sintetizados.....	21
Tabla IV. Datasets más populares el día 11/07/2020 .....	22
Tabla V. Fechas de inicio y actualización, y frecuencia de actualización .....	44
Tabla VI. Licencias de uso de los datasets.....	45
Tabla VII. Uso de metadatos en el catálogo de datos.....	46
Tabla VIII. Metadatos del dataset .....	48
Tabla IX. Formatos de descarga de datasets .....	48
Tabla X. Estadísticas web.....	54

## 1. MARCO TEÓRICO

En el presente capítulo se contextualiza el trabajo, definiendo el significado de los datos abiertos y su relación con el movimiento de "Gobierno Abierto", para después analizar qué criterios de calidad deben observarse a la hora de publicar dichos datos en portales web.

### 1.1. LOS DATOS ABIERTOS

El mundo en el que vivimos está experimentando una significativa transformación a nivel global, favorecida por la tecnología y los medios digitales. Y dentro de esta "era de la información" en la que estamos inmersos, los datos tienen un poder cada vez mayor (Ruvalcaba Gómez, 2020). Se atribuye a C. R. Humby<sup>1</sup> la consideración de los datos como el "nuevo petróleo", destacando que, al igual que éste, los datos deben ser transformados, desglosados y analizados para tener valor. En este sentido, las iniciativas de datos abiertos suponen una de las mejores vías para lograr el objetivo de extraer a los datos todo su potencial.

Según el *Manual de Datos Abiertos* publicado por la OKFN (Open Knowledge Foundation, s. f.), los datos abiertos son aquellos que pueden ser libremente utilizados, reutilizados y redistribuidos por cualquier persona, estando sujetos, como mucho, a los requisitos de atribución y de compartirlos de la misma manera en que aparecen. Se caracterizan por tres aspectos:

- Disponibilidad y acceso: la información debe estar disponible como un todo, en una forma conveniente y modificable, y a un costo razonable de reproducción, preferiblemente descargándola de internet.
- Reutilización y redistribución: los datos deben ser legibles por máquina y proporcionados bajo términos que permitan reutilizarlos, redistribuirlos e incluso integrarlos con otros conjuntos de datos.
- Participación universal: todas las personas deben poder utilizar, reutilizar y redistribuir la información, sin discriminación. No están permitidas las restricciones que impiden el uso comercial de los datos, o restricciones de uso para ciertos propósitos (por ejemplo, sólo para educación).

Este manual destaca que el cumplimiento de dichos requisitos es muy importante porque son necesarios para la interoperabilidad, es decir, para integrar diferentes conjuntos de datos. Esta integración es la esencia de los datos compartidos: al incrementar la capacidad de combinar distintas bases de datos o conjuntos de datos, se posibilita el desarrollo de más y mejores productos y servicios.

### 1.2. DATOS ABIERTOS GUBERNAMENTALES

Se distinguen siete tipos de datos abiertos, con diferentes usos potenciales y aplicaciones: culturales, científicos, de finanzas, estadísticos, meteorológicos, medioambientales y de transporte. Las administraciones públicas son uno de sus principales creadores, por lo que se les suele confundir con otros conceptos como "datos gubernamentales", "gobierno electrónico" o "e-gobierno" (Hernández-Pérez, 2016).

En este sentido, la FEMP (Federación Española de Municipios y Provincias, 2017) define los datos abiertos gubernamentales como "aquellos elaborados por las administraciones públicas en un principio para su propio consumo, pero que se ponen

---

<sup>1</sup> Clive Robert Humby (- 03/02/1955) Matemático y emprendedor británico, especialista en análisis de datos. <https://medium.com/project-2030/data-is-the-new-oil-a-ludicrous-proposition-1d91bba4f294>

a disposición de la ciudadanía en general y del sector infomediario, en particular, en un formato abierto y por lo tanto reutilizable". En el presente trabajo nos referiremos a ellos de manera genérica como "datos abiertos".

Estos datos abiertos entroncan con el denominado Movimiento de Gobierno Abierto, que es un modelo de gobierno basado en la transparencia, la participación y la colaboración. Estos principios fueron establecidos en el *Memorandum sobre Transparencia y Gobierno Abierto* presentado en 2009 por Barack Obama, que evolucionó el concepto de gobierno electrónico implantando un nivel de apertura sin precedentes y el primer portal de datos abiertos (Curto-Rodríguez, 2017).

Apunta Ruvalcaba Gómez (2020) que los gobiernos publican datos abiertos porque reportan beneficios como el desarrollo de la economía, el desarrollo de la calidad democrática y la mejora de los servicios públicos, además de fomentar la participación de los ciudadanos en los procesos de toma de decisiones.

La publicación de datos abiertos debe ceñirse a una serie de requisitos. Los más destacables y comúnmente aceptados fueron establecidos en la Reunión de Sebastopol (2007) y en la Carta Internacional de Datos Abiertos (2015), como veremos a continuación.

En diciembre de 2007 el Open Government Working Group, grupo de defensores del movimiento, se reunió en Sebastopol (California) para establecer **los 8 principios originales de los datos abiertos gubernamentales** (Open Government Working Group, 2007):

- Completos: Se deben abrir todos los datos públicos que no estén sujetos a limitaciones de privacidad, seguridad o privilegios válidos.
- Primarios: los datos se deben publicar en bruto, tal y como están en su origen y con la mayor granularidad posible, no en formas agregadas o modificadas. Por ejemplo, un informe no se considera datos abiertos porque contiene datos procesados y no originales.
- Actualizados: los datos se pondrán a disposición tan pronto como sea necesario para preservar su valor, para que representen la realidad.
- Accesibles: los datos deben estar disponibles para la mayor cantidad de usuarios posible y para la mayor cantidad de usos posible.
- Procesables por máquina: la estructura de los datos debe permitir el procesamiento automatizado. Un documento en formato PDF no es procesable por un lenguaje de programación; pero unos datos publicados en formato XML estructurado son comprensibles para ordenadores y podrán ser reutilizados de forma automatizada.
- No discriminatorios: deben estar disponibles para todos, sin necesidad de registro previo.
- No propietarios: los datos deben estar disponibles en un formato que no dependa de una entidad o de una herramienta propietaria de una entidad. Un formato abierto sería RDF, CSV o XML, mientras que Excel, Word y similares son formatos propietarios.
- Licencia libre: Los datos tienen que estar libres de derechos, patentes o copyright, y no estar sujetos a restricciones de privacidad, seguridad o privilegio que puedan estar regladas por otras normas.

En junio de 2013 los líderes del G8 constituyeron el Open Data Charter y establecieron 5 principios básicos que permitiesen promover la transparencia, innovación y responsabilidad de los gobiernos. Con el paso del tiempo se implicaron más gobiernos, organizaciones y sociedad civil para refinar y mejorar esos principios iniciales, gestándose la Carta Internacional de Datos Abiertos (Open Data Charter, 2015), que se ha convertido en el nuevo estándar ad-hoc para la implementación de las estrategias de datos abiertos en el mundo con 6 principios:

- Abiertos por defecto.
- Puntuales y completos, además de publicarse rápidamente deben publicarse como un todo en su forma original.
- Accesibles y utilizables: con facilidad de acceso y bajo licencia libre, los formatos deben tener en cuenta tanto su legibilidad por parte de las máquinas como la experiencia de los usuarios.
- Comparables e interoperables: uso de estándares que posibiliten la comunicación entre grupos de datos y su interoperabilidad.
- Para mejorar la gobernanza y la participación ciudadana: fomento de la transparencia para que los ciudadanos sepan qué hacen los gobernantes, mejoren los servicios públicos y se mantenga la rendición de cuentas.
- Para el Desarrollo inclusivo y la Innovación: los datos abiertos posibilitan el desarrollo de negocios basados en su explotación.

Existen diversos índices para evaluar el cumplimiento de los principios de datos abiertos en los portales gubernamentales. Los más utilizados son:

- El *Global Open Data Index* (Open Knowledge Foundation, 2020), que evalúa la disponibilidad de los datos;
- El *Open Data Barometer* (World Wide Web Foundation, 2017), que además de su disponibilidad evalúa su utilidad;
- El *Informe Anual del Portal Europeo de Datos* (Publications Office of the European Union, 2019), que analiza el nivel de madurez del ecosistema de datos abiertos en Europa;
- El *OURData Index* de la OCDE (Organisation for Economic Cooperation and Development, 2020), que analiza la apertura de datos de 28 países de la OCDE (Organización para la Cooperación y el Desarrollo Económico) en las áreas de apertura, utilidad y reutilización de los datos gubernamentales.

### **1.2.1. Datos abiertos en España**

Además de ser de aplicación la Directiva 2003/98/CE relativa a la reutilización de la información en el sector público (Unión Europea, 2003) y la Directiva 2013/37/UE por la que se modifica la directiva anterior (Unión Europea, 2013), las primeras estrategias orientadas en nuestro país hacia la apertura de datos y su reutilización se plasman en la Ley 37/2007 de Reutilización de Información del Sector Público (España, 2007), modificada por Ley 18/2015 (España, 2015), y en el Real Decreto 1495/2011 que concreta la aplicación de la Ley 37/2007 en el ámbito del sector público estatal (España, 2011)

Una de las primeras iniciativas estatales sobre gobierno abierto es la Iniciativa Aporta (Datos.gob.es, s.f.), lanzada en 2009 para promocionar la apertura de la información pública y el desarrollo de servicios avanzados basados en datos. Promovida por el Ministerio de Economía y Empresa, el Ministerio de Política Territorial y Función Pública, y la Entidad Pública Empresarial Red.es, sus líneas de actuación están destinadas a todos los agentes del ecosistema de datos: usuarios, organismos públicos y reutilizadores e infomediarios, y se plasman en la plataforma nacional *datos.gob.es*, que recoge información de la mayor parte de los portales de datos abiertos del país.

De acuerdo a (Ruvalcaba Gómez, 2020), la apertura de datos experimentó un auge en nuestro país tras la publicación de la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno (España, 2013a) y la Norma Técnica de Interoperabilidad de Reutilización de recursos de la información (España, 2013b).

### **1.3. PROBLEMAS EN LA APERTURA DE DATOS**

A pesar de las adhesiones de los gobiernos a los principios citados, se están reportando problemas en la apertura de los datos. La cuarta edición del *Barómetro de Datos Abiertos* analiza las políticas y prácticas de datos abiertos desarrolladas por 115 países para ver si cumplen con los 6 Principios de la Carta Internacional de Datos Abiertos, y determina que 9 de cada 10 conjuntos de datos gubernamentales no son abiertos, que por lo general están incompletos y son de baja calidad, y que no se publican los datos necesarios para restablecer la confianza de los ciudadanos (World Wide Web Foundation, 2017). Autores como Máchova y Lnenicka (2017) y Lourenço (2015) coinciden en que actualmente las administraciones se centran únicamente en la creación de portales y ese es un enfoque que no funciona. No se está teniendo en cuenta que para generar valor económico y potenciar la transparencia es necesario mantener un proceso cíclico en el que la publicación de los datos sea solo el inicio y no un trabajo puntual, tal y como señala la Junta de Castilla y León (2019) en su *Guía para la publicación y reutilización de datos abiertos*.

La falta de calidad es un riesgo que puede trastocar las iniciativas de gobierno abierto, porque puede afectar a la reputación y confianza pública en los gobiernos de manera irrevocable (Lourenço, 2015) y tener un impacto negativo en los negocios que se nutren de esos datos. Además, la calidad de los datos publicados es uno de los factores clave tenidos en cuenta en el proceso de evaluación del e-gobierno.

Uno de los elementos que afecta a la calidad es la diversidad de las fuentes de datos en los sistemas de origen, por cuanto la implementación de portales abiertos involucra la integración de datos provenientes de diversas fuentes formales o informales, lo que dificulta su publicación en abierto (Cadena-Vela et al., 2019).

En muchos casos es la estructura y organización del portal la que dificulta que el ciudadano medio localice la información, teniendo que delegar esta tarea en especialistas, como periodistas o investigadores (Lourenço, 2015).

El de la baja calidad de los metadatos y los datos es otro problema generalizado que entorpece la búsqueda y descubrimiento de conjuntos de datos relevantes para el usuario, además de dificultar su procesamiento e integración con otros datasets, hacer más complicada su reutilización (Vetró et al., 2016) y provocar que se abandonen los proyectos de portales de datos abiertos al poco de implantarse para la publicación de datos (Cadena-Vela et al., 2019).

Apuntan Abeleida Díaz et al. (2019) que es necesario definir unos estándares de calidad de los datos abiertos por el gran valor de la apertura de la información pública y la necesidad de afianzar la transparencia activa.

La calidad no asegura el uso de los datos, pero aumenta el potencial de creación de valor público en ámbitos locales, estatales y nacionales (Atenas et al., 2019). Por eso, debe hacerse un esfuerzo no solo en publicar grandes cantidades de datos, sino también en publicar datos de gran calidad; "sin datos de alta calidad, diseñar, monitorear y evaluar políticas efectivas se vuelve casi imposible" (Curto-Rodríguez, 2020).

#### **1.4. LA CALIDAD DE LOS DATOS ABIERTOS**

Los autores coinciden en que no hay una definición única para determinar qué es la calidad de los datos en general, ni de los datos abiertos en particular. El campo de los datos abiertos es muy reciente, por lo que han empezado a establecerse hace poco los criterios y medidas específicas para evaluar tanto su calidad como la de los portales que los publican.

##### **1.4.1. Calidad reutilizadora**

Uno de los enfoques más utilizados en la literatura es el de aplicar el modelo de 5 estrellas de para evaluar la calidad de los datos abiertos de cara a su reutilización. Tim Berners-Lee (2010) diseñó este esquema de clasificación para alentar a los gobiernos a publicar datos abiertos y enlazados. En él se establecen 5 niveles de calidad que se van obteniendo a medida que se cumplen los criterios establecidos para cada estrella:

1. Datos publicados en la web en cualquier formato bajo licencia abierta. Ejemplo: formato PDF.
2. Datos publicados en la web en un formato estructurado legible por máquina. Ejemplo: formato Excel.
3. Datos publicados bajo formatos estructurados no propietarios. Ejemplo: formato CSV.
4. Utilización de estándares abiertos de W3C (RDF y SPARQL) y URIs basadas en el protocolo HTTP para identificación. Ejemplo: datos exportables en formato html y otros.
5. Datos enlazados a datos de otro proveedor para brindar contexto.

A pesar de su amplia aceptación, este modelo tiene algunas limitaciones. Para Vetró et al. (2016), este esquema cubre solo un aspecto específico de la calidad: el formato o codificación utilizado para publicar los datos. Como consecuencia, un dataset puede alcanzar el nivel de 5 estrellas y tener a la vez una pobre calidad al no tenerse en cuenta otros aspectos.

Por este motivo, Oviedo et al. (2015) proponen enriquecerlo con seis medidas: disponibilidad de los datos publicados, capacidad de reutilización, relevancia, reputación, granularidad y visualización, y lo aplican al caso de la municipalidad de Pérez Zeledón en Costa Rica, mientras que Abella et al. (2018), tras aplicarlo en el análisis del portal de datos abiertos de Barcelona, sugieren incorporarle aspectos como la frecuencia de actualización de datos y la geolocalización para que evalúe mejor la calidad de los datos abiertos y su potencial para la reutilización.

Otro modelo para evaluar la calidad reutilizadora es Meloda (Abella et al., 2014), una métrica que analiza la reutilización de los datos abiertos a través de varias dimensiones (estándares técnicos, acceso, legal y modelo de datos) y que propone una forma cuantitativa para evaluarla. Tras detectar diversas limitaciones, sus autores desarrollaron Meloda 5 (Abella et al., 2020), ampliando dichas dimensiones de cuatro a ocho: licencia legal de los datos, mecanismos de acceso a la información, estándares técnicos de los datasets, estandarización, contenido geográfico de los datos, frecuencia de actualización, difusión y reputación.

### **1.4.2. La calidad como concepto multidimensional**

Máchová y Lnenicka (2017), Vetró et al. (2016) y Lourenço (2015) apuntan que, de manera generalizada, se considera que unos datos son de calidad si son "adecuados para su uso", mientras que Cadena-Vela (2019) asocia la calidad al cumplimiento de los requisitos para los que fueron creados. En este sentido, Leonangeli y Maenza (2019) indican que si el propósito de los datos abiertos es el de colaborar con la iniciativa de Gobierno Abierto, su calidad se medirá de acuerdo a la adecuación a dicho propósito.

En esta "adecuación" intervienen, tanto atributos objetivos de los portales y datos como las percepciones subjetivas de sus usuarios y el contexto en el que se vayan a utilizar: lo que puede ser útil para unos, no lo será para otros. Venegas Álvarez (2018, p. 43) expone que "un usuario evalúa las características de calidad de los datos cuando usa las funciones y servicios que el portal Web ofrece". De esta manera, sus opiniones y experiencias, su *feedback*, son muy valiosas en el proceso de mejora continua y de desarrollo eficiente de los portales, ya que aportan información sobre fallos en el portal y sobre necesidades específicas de datasets (Máchová y Lnenicka, 2017).

Por consiguiente, podemos concluir que la calidad de los datos abiertos es un concepto multidimensional en el que hay que tener en cuenta factores como las características y herramientas del portal en el que son ofrecidos, los indicadores de calidad de los atributos de los datos y su contenido.

#### **1.4.2.1. Los portales de datos abiertos**

Herrera-Melo y González-Sanabria (2019, p.5) apuntan que el portal de datos abiertos "es el medio por el que se garantiza la calidad de los datos publicados, se facilita la búsqueda a los usuarios, los datos están disponibles en formatos utilizables y estos son publicados de manera que respondan a una demanda específica que responda a necesidades específicas que generen valor".

No deben confundirse con los portales de transparencia; para la FEMP (2017, p. 63) estos portales son "recursos web donde se deben presentar las normativas y datos que demuestren la transparencia de la institución, desde las reglas que se aplican hasta los registros y datos que dan fe de la ejecución de dichas normas"; García-García y Curto-Rodríguez (2018) añaden que los portales de transparencia divulgan de manera sencilla información concreta a la que se accede de manera individual, en formato PDF u hoja de cálculo para facilitar su visualización. En los portales de datos abiertos se publican conjuntos de datos en formatos reutilizables buscando reducir los obstáculos para la agregación y el tratamiento mecanizado.

Según Royo Montañés y Benítez Gómez (2019), no hay muchos estudios que analicen las pautas a seguir a la hora de implementar y estructurar un portal de este tipo, pero destacan las recomendaciones recogidas en las guías publicadas por la FEMP (2017) y la Junta de Castilla y León (2012), que enumeran una serie de elementos básicos y recomendables de un portal de datos abiertos:

- Presentación: descripción del portal, misión y objetivos.
- Catálogo: buscador, filtros y formatos utilizados.
- Buenas prácticas y aplicaciones de terceros.
- Colaboración ciudadana: propuestas de nuevos conjuntos de datos para publicar; conjuntos de datos generados por ciudadanos; propuesta de aplicación realizada por reutilizadores; aplicaciones realizadas que han sido enviadas por reutilizadores; registro de reutilizadores.
- Ejemplos de uso.
- Servicios Interfaz de Programación de Aplicaciones (API).

- Servicio SPARQL (recomendable).
- Herramientas de visualización (recomendable).
- Condiciones de uso o términos de reutilización.
- Medios de contacto y/o ayuda técnica.

Lourenço (2015) considera que no sirve de mucho tener datos abiertos de calidad si se ofrecen a través de un portal cuya estructura y organización obstaculizan su búsqueda y análisis, afectando a la calidad de la información obtenida y a la calidad de los servicios derivados de su reutilización. Este autor analiza la estructura y organización general de 7 portales de datos abiertos gubernamentales y concluye que no son útiles al ciudadano medio para obtener información relacionada con la transparencia orientada a la rendición de cuentas, además de funcionar como simples repositorios de datos y presentar falta de calidad en la cumplimentación de los metadatos.

Medina Sánchez (2015) analiza diversos parámetros de diez portales nacionales e internacionales de datos abiertos, detectando carencias en la cantidad de conjuntos de datos y aplicaciones, la contactabilidad con los publicadores y el sistema de comentarios, el uso de consultas SPARQL, la integración en redes sociales, la inclusión de novedades, la documentación API y la disponibilidad de los portales en varias lenguas.

Máchová y Lnenicka (2017) proponen un modelo con el que evalúan la calidad de 67 portales de datos abiertos a nivel nacional para comprobar si proporcionan los datos de manera que realmente faciliten su uso, reutilización y rendición de cuentas. Además de observar una mayor calidad en los portales gestionados con CKAN o DKAN, determinan que la complejidad de los portales y sus funciones difieren, reflejando falta de armonización, altos grados de heterogeneidad y una necesidad de estándares de calidad de datos abiertos; detectan que las funcionalidades de búsqueda, descarga y subida de datos son básicas, que solo algunos aprovechan la web semántica, y les faltan funcionalidades para obtener el *feedback* de los usuarios y para procesar los datasets, además de faltar información relevante para la rendición de cuentas y la transparencia.

Curto-Rodríguez (2019) evalúa las posibilidades de participación y colaboración (dos de los pilares del gobierno abierto) en los portales autonómicos de datos abiertos, concluyendo que solo se ofrece de manera efectiva la solicitud de información mediante simples formularios o correos electrónicos, estando muy poco presentes herramientas de votación de propuestas, blogs, perfiles en redes sociales o sindicación de contenidos.

Royo-Montañés y Benítez-Gómez (2019) definen una metodología para analizar los portales de datos abiertos y la aplican a una muestra de municipios españoles. También analizan la información económico-financiera ofrecida, por ser fundamental para la rendición de cuentas. Concluyen que menos de la mitad de los municipios tienen portal de datos abiertos, que funcionan como meros repositorios, y que no cuidan aspectos que potencien el uso de los datos por un público no experto, la rendición de cuentas, la participación ciudadana y la calidad de los datos.

#### **1.4.2.2. Indicadores de calidad de los datos abiertos**

En la literatura encontramos diferentes listas con indicadores para analizar la calidad de los datos abiertos, pero dichas listas no suelen coincidir entre sí, ya que algunos conceptos tienen el mismo nombre y significado diferente, o tienen el mismo significado con diferentes nombres, dependiendo del contexto abordado y de la perspectiva de los autores. Vetró (2016) destaca la problemática de esta situación,

ya que dar definiciones diferentes a las mismas características de calidad, conduce a obtener como resultado diferentes métricas para medir las mismas características.

A los datos abiertos se les pueden aplicar en parte los criterios establecidos para los datos en general (Máchová y Lnenicka, 2017). En este sentido, el estándar ISO 25012 (Organización Internacional de Normalización, 2008) ofrece algunos criterios que se pueden aplicar. Este estándar define un modelo general para la calidad de los datos almacenados de manera estructurada dentro de un Sistema de Información. El modelo está compuesto por 15 características agrupadas en tres categorías:

Calidad inherente a los datos	Calidad de datos inherente y dependiente del sistema	Calidad de los datos dependiente del sistema
Exactitud Complejidad Consistencia Credibilidad Actualidad	Accesibilidad Conformidad Confidencialidad Eficiencia Precisión Trazabilidad Comprensibilidad	Disponibilidad Portabilidad Recuperabilidad

Tabla I. Características de calidad según ISO 25012

Algunas de estas características son retomadas por la iniciativa Open Data Support de la Comisión Europea (Dekkers et al., 2014) para definir la calidad de los datos abiertos y sus metadatos: Exactitud, Complejidad, Consistencia, Credibilidad, Actualidad, Conformidad, Disponibilidad, precisión y procesabilidad.

El último informe disponible del *Open Data Maturity Report* (Blank et al., 2019) define cuatro dimensiones a tener en cuenta para analizar la calidad de los datos abiertos: la actualidad, los mecanismos de control de la calidad y medidas de información sobre licencias, la conformidad con el estándar DCAT-AP, y la implementación de datos enlazados, mientras que la cuarta edición del *Barómetro de Datos Abiertos* (World Wide Web Foundation, 2017) indica que los datos de buena calidad deben estar disponibles online, ser legibles por máquinas, estar disponibles en bloque, ser gratuitos y de licencia abierta.

A continuación, veremos los diversos indicadores de calidad planteados en la literatura por diferentes autores:

Oviedo et al. (2015), en su estudio antes mencionado sobre los datos abiertos de dos municipalidades de Costa Rica, proponen un modelo de calidad y madurez para portales de datos abiertos analizando como criterios de calidad la disponibilidad de los datos publicados, su capacidad de reutilización, su relevancia, reputación, granularidad y visualización.

Para Lourenço (2015), la calidad es una de las características de la divulgación de datos y está muy en relación con algunas de las otras características, que son: completitud, acceso y visibilidad, usabilidad y comprensibilidad, puntualidad, valor y utilidad, granularidad, y comparabilidad (aunque a esta última no la considera explícitamente en su trabajo por depender fuertemente de otras características). Considera que la calidad incluye varios aspectos como la "exactitud", entendida como datos "intrínsecamente buenos" y asociada a conceptos como la validez, la consistencia, la fiabilidad y la autenticidad.

Vetró et al. (2016) proponen un marco de evaluación para analizar cualitativamente la calidad de los portales de datos abiertos centrándose en los

niveles de dataset y celda (que es el más granular), y plantean como indicadores de calidad la trazabilidad, actualidad, expiración, completitud, conformidad, comprensibilidad y exactitud. Validan el marco comparando dos casos de datasets del Italian Open Government Data: el Open Coesione Project, de apertura centralizada y extensos controles de calidad, y ejemplos de datos abiertos gubernamentales procedentes de apertura de datos descentralizada (a nivel municipal), sin posibilidad de amplio control de calidad y supuestamente de menor calidad.

Abeleida Díaz et al. (2019) estudian la calidad de los datos abiertos difundidos en los portales de transparencia activa en los gobiernos municipales de La Serena y Coquimbo (Chile) analizando su apertura, oportunidad y exhaustividad, accesibilidad y utilidad de la información pública, comparación e interoperabilidad de los datos, indicadores de calidad basados en los principios de la Carta Internacional y la Ley de Transparencia Chilena.

Para Cadena-Vela et al (2019) las dimensiones de calidad pueden ser dependientes de la gestión implementada en los portales de datos, es decir, de los datos ya publicados: disponibilidad, conformidad, credibilidad, accesibilidad y relevancia, o pueden ser dependientes de los sistemas de información de origen (de los datos antes de publicar): exactitud, puntualidad, consistencia y exhaustividad. En su trabajo se centran en estas últimas para diseñar un marco de referencia que posibilite la publicación de datos con un nivel de calidad adecuado.

Venegas Álvarez (2018) propone en su tesis una metodología para evaluar el nivel de apertura, capacidad de reutilización y accesibilidad de los datos abiertos publicados en los portales de transparencia de Chile, Costa Rica y Uruguay mediante 23 dimensiones de calidad: completitud, procesable por máquina, primario o granular, estandarizado, portabilidad, capacidad de relacionarse (*linkage-ability*), cantidad adecuada de datos, pertinencia, actualización, claridad contextual, licenciamiento, formato, documentación, descripción, metadatos, catálogo de datos, contactable, trazabilidad, consistencia, coherencia, sostenibilidad, privacidad y seguridad.

Leonangeli y Maenza (2019) plantean un listado de métricas de calidad (completitud, conformidad, inteligibilidad, exactitud, exactitud semántica, exactitud sintáctica y univocidad) para analizar el portal de datos abiertos de la Municipalidad de Rosario (Argentina) y determinar si describen de manera positiva y sustantiva la propiedad de sus datasets de servir a los propósitos de Gobierno Abierto en términos de impacto.

Herrera-Melo y Gonzalez-Sanabria (2019) plantean una metodología de evaluación de la calidad de datos y de portales de datos abiertos sintetizando modelos anteriores, y la aplican al portal de datos abiertos del gobierno de Colombia.

Combinando todos ellos, podemos identificar una serie de indicadores para medir la calidad de los datos abiertos, basándonos en:

- Las características básicas de los datos abiertos, recogidas en la *Definición de datos abiertos* de Open Knowledge Foundation.
- Los 8 principios originales definidos en Sebastopol.
- Los principios definidos por la Carta Internacional de Datos Abiertos.
- Los atributos de calidad inherente a los productos de datos definidos por la norma ISO/IEC 25012.
- Las dimensiones de la calidad de los datos abiertos definidas por la iniciativa Open Data Support de la Comisión Europea.
- Los aspectos relativos a la calidad de los datos abiertos recogidos en el *Open Data Barometer*.
- Las dimensiones planteadas en el *Open Data Maturity Report 2019*.

- Las dimensiones recomendadas en el *Manual práctico para mejorar la calidad de los datos abiertos* de Datos.gob.
- Los indicadores planteados en la literatura académica.

De esta manera, a efectos del presente trabajo consideramos que los indicadores de calidad de los datos son los siguientes: Accesibilidad, Apertura, Actualidad, Conformidad, Consistencia, Completitud, Comprensibilidad, Credibilidad, Disponibilidad, Exactitud, Precisión, Procesabilidad, Valor y Utilidad, y Volumen de datos. Abeleida Díaz et al. (2019) y Lourenço (2015) hablan también de “comparabilidad”, pero no se ha tenido en cuenta por depender fuertemente de otras características, tal y como indica el último.

### **a) Accesibilidad**

Este indicador implica que debe haber pocas o ninguna barrera para que los ciudadanos accedan a los datos. Deben estar disponibles para la mayor cantidad de usos y de usuarios posibles, sin ningún tipo de discriminación y facilitando su utilización a usuarios con discapacidad. También deben ser fácilmente accesibles de manera gratuita a través de internet, sin necesidad de un registro previo, que no debe ser confundido con la posibilidad de crear una cuenta de usuario para personalizar las visualizaciones y la información mostrada, algo que es completamente recomendable según Máchová y Lnenicka (2017).

El indicador está muy relacionado con el diseño del portal y su usabilidad, aspectos que vimos más arriba al hablar de cómo debe implementarse un portal de datos abiertos: los buscadores y la recuperación de información (categorías temáticas, etiquetas), la organización por categorías, el sistema de gestión de datos utilizado, los idiomas, etc.

Oviedo et al. (2015) hablan de “visualización de los datos publicados”, entendida como la capacidad de acceder a los datos gráficamente.

### **b) Actualidad**

También denominada Puntualidad, es una dimensión común a todas las fuentes consultadas. Significa que los datos y metadatos deben representar la situación actual para ser relevantes y preservar su valor; por ello deben ser puntuales, sin retraso de la publicación después de la expiración de su versión previa. Abeleida Díaz et al. (2019) hacen referencia también a la respuesta a la solicitud de la información dentro de los plazos legales.

Gracias a las nuevas tecnologías ahora es posible ejercer una transparencia en tiempo real divulgando los datos de manera continua y casi al tiempo de que ocurran las acciones que originan los datos.

### **c) Apertura**

Todos los datos públicos deben estar abiertos por defecto, salvo aquellos sujetos a limitaciones de privacidad, seguridad u otros privilegios válidos. No están permitidas las restricciones que impiden el uso comercial de los datos, o restricciones de uso para ciertos propósitos (por ejemplo, sólo para educación).

La apertura implica contar con una licencia abierta fácilmente reconocible como *Creative Commons* o pertenecer directamente al dominio público para que los datos puedan ser utilizados, reutilizados, redistribuidos o integrados en otros conjuntos de datos. Uno de los indicadores de calidad del *Open Data Maturity 2019* analiza qué mecanismos proporcionan los países para asesorar a los publicadores en la elección del tipo de licencia correcto para sus datos.

<b>LICENCIAS PARA LA PUBLICACIÓN DE DATOS ABIERTOS</b>
De dominio público:
Creative Commons Zero (CC Zero). Public Domain Dedication and License (PDDL 1.0).
Solo obligan a atribuir la fuente original de los datos:
Creative Commons-Reconocimiento (CC BY 4.0). Open Data Commons Attribution License 1.0 (ODC-By 1.0).
Requieren atribución de la fuente original y compartir los resultados bajo las mismas condiciones:
Creative Commons-Reconocimiento Compartir Igual (CC BY-SA 4.0). Open Data Commons Open Database License (ODbL 1.0).

*Tabla II. Licencias para la publicación de datos abiertos (datos.gob.es, 2017a)*

#### **d) Completitud**

Denominado "Exhaustividad" por Cadena-Vel et al. (2019), indica que los datos incluyen todos los atributos que representan a la entidad o evento, todos los registros tienen su valor correspondiente en el mundo real. También implica que las celdas no están vacías y tienen valores con significado (Vetró et al., 2016). Lourenço (2015) relaciona la completitud con la presencia en el portal de listados maestros de entidades, ítems y fuentes relevantes de conjuntos de datos relacionados con la rendición de cuentas.

#### **e) Comprensibilidad**

Es necesario que los datos estén expresados de manera que sean interpretables de manera inequívoca a través de los metadatos y documentación disponibles, utilizando lenguajes, símbolos y unidades apropiados en un contexto de uso específico (Organización Internacional de Normalización, 2008). El formato de los datos debe ser comprensible para el usuario (Vetró et al., 2016) y favorecer su acceso a los datos (Open Data Charter, 2015).

A Lourenço, (2015) le preocupa que los ciudadanos corrientes no sean capaces de entender e interpretar correctamente la información liberada, y apunta que debería ser comprensible, con la mínima jerga y sin significados ocultos.

Venegas Álvarez (2018) plantea este indicador dividido en 3 dimensiones: "Documentación" que ayude al usuario a interpretar los datos, "Descripción" del contenido y "Metadatos" mínimos que debe contener el conjunto de datos. Este autor también evalúa si hay un "Catálogo de datos" que oriente al usuario respecto a los conjuntos de datos que se pueda encontrar en el portal, pudiendo consistir en una lista de conjuntos de datos ordenados por temas y con posibilidad de filtrar por temas u otra característica.

#### **f) Conformidad**

Este indicador implica el cumplimiento de estándares, normas y legislación en la publicación y representación de la información. La *Carta Internacional de Datos Abiertos* (2015), Abeleida Díaz et al. (2019) y Venegas Álvarez (2018) destacan la

importancia de los estándares para la interoperabilidad de los datos, uno de los pilares fundamentales del movimiento de Datos Abiertos; el *Open Data Maturity Report* (2019) se centra en la aplicación del estándar DCAT-AP y la implementación de datos enlazados, mientras que para Leonangeli y Maenza (2019) y Vetró et al. (2016) la conformidad es el cumplimiento con el modelo de 5 estrellas.

### **g) Consistencia**

Los datos no pueden contener contradicciones y deben ser coherentes respecto a otros datos existentes en el mismo contexto de uso (datos.gob.es, 2017b). Por ejemplo, se debe emplear el mismo formato y estructura en los conjuntos de datos publicados de forma periódica, no utilizar el mismo término para diferentes conceptos o viceversa, y no hacer variaciones en la metodología que puedan afectar a los valores de los datos sin dar explicaciones. En este indicador se ha incluido lo que Venegas llama Consistencia y Coherencia como dos indicadores diferentes.

Leonangeli y Maenza (2019) analizan el porcentaje de celdas con diferente URI, pero con las mismas propiedades y lo denominan "Univocidad".

### **h) Contactabilidad e Interactividad**

Venegas Álvarez (2018) denomina "Contactable" a la dimensión que evalúa si en el portal hay un mecanismo que permita establecer comunicación con el proveedor de datos, brindando soporte al usuario. Pero el contacto entre ambas partes no debe quedarse en un mero soporte, sino que tiene que posibilitar la participación del usuario a través de su *feedback* o la solicitud de nuevos datasets. También se debe fomentar el contacto entre usuarios mediante foros y medios sociales. Muy en relación con Valor y Utilidad.

### **i) Credibilidad**

Los datos son de fuentes fiables. En la credibilidad se incluye un aspecto que Vetró et al. (2016), Venegas Álvarez (2018) y la ISO 25012 (2008) plantean de manera separada, la "trazabilidad", que permite conocer los detalles de la creación y modificación de los conjuntos de datos a través de metadatos asociados. Oviedo et al. (2015) lo llaman "reputación", mientras que Lourenço (2015) habla de "fiabilidad, autenticidad y validez", conceptos que atañen a la posibilidad de revisar y certificar que los procedimientos adoptados y los datos divulgados cumplen con los principios de datos abiertos y las características deseadas de la divulgación de datos.

### **j) Disponibilidad**

Se debe garantizar que los datos estén disponibles tanto ahora como en el futuro. Hay que utilizar URIs, esquema que establece un mecanismo de identificación común para los datos que se exponen públicamente, de forma que se pueda hacer referencia a estos de forma única, fiable y persistente en el tiempo, requisito clave para facilitar su posterior reutilización (España, 2013b).

En este indicador se incluye lo que Venegas Álvarez (2018) llama "sostenibilidad y custodia" y "almacenamiento", puntualizando que los datos se deben publicar de forma periódica según la ley o acuerdo previo, estando siempre en línea o accesibles bajo custodia de alguna entidad.

La disponibilidad puede verse afectada por buscadores de datos excesivamente guiados o acotados, datos obsoletos o sin actualizar.

## **k) Exactitud**

Los datos deben representar correctamente a la entidad, situación o evento del mundo real, deben presentarse los valores correctos de acuerdo al dominio y tipo de información del dataset. Venegas Álvarez (2018) introduce el concepto de "Claridad contextual", en el sentido de que los datos no se pueden interpretar con exactitud si no se sabe a qué hacen referencia en el mundo real. Leonangeli y Maenza (2019) distinguen entre exactitud, exactitud semántica y sintáctica.

## **l) Precisión**

Este indicador es llamado "primario" en los *8 principios de los datos abiertos gubernamentales* (2007) y por Venegas Álvarez (2018). Los datos se deben publicar en bruto y respetando sus niveles de detalle y granularidad originales, evitando modificaciones, para que cada reutilizador puede tratar la información de la manera que más le convenga. Se exceptúa el filtrado necesario por cuestiones de privacidad, confidencialidad o seguridad.

## **m) Procesabilidad**

Este indicador está relacionado con el de Comprensibilidad, ya que los datos deben ser comprensibles tanto para los usuarios como para las máquinas. Leonangeli y Maenza (2019) denominan "inteligibilidad" a la presencia de formatos entendibles tanto para usuarios finales o computadoras.

En este caso, se hace referencia a que los datos deben ser legibles por máquina a través de formatos abiertos (que no dependen de una entidad o de una herramienta propietaria de una entidad) que permitan su reutilización.

En este indicador se engloba lo que algunos autores llaman "comprensibilidad" (Vetró et al., 2016) o "accesibilidad" (Cadena-Vela, 2019), refiriéndose a la legibilidad de los datos por las máquinas, y "reutilización", referida al uso de metadatos y formatos reutilizables (datos.gob.es, 2017b).

Lo que a efectos de este trabajo denominamos Procesabilidad, Venegas Álvarez (2018) lo plantea dividido en 4 dimensiones: "Procesable por máquina", "Formato abierto", "Metadatos" y "Portabilidad" (Posibilidad de movilizar o exportar datos preservando su nivel de calidad).

Oviedo et al. (2015) hablan de la "capacidad de reutilización" de los datos para realizar análisis más complejos, mientras que el *Open Data Maturity Report* (2019) considera como indicador de calidad el tener implementado algún modelo que evalúe la calidad de los datos, como el de Berners- Lee.

En nuestro país, la Ley 19/2013 de Transparencia (España, 2013a) indica que la información debe difundirse preferentemente en formatos reutilizables; por su parte, la Ley 18/2015 (España, 2015) señala que siempre que sea posible y apropiado se procurará proporcionar la información en formato abierto y legible por máquina, conjuntamente con sus metadatos y con los niveles más elevados de precisión y desagregación.

Algunos ejemplos de formatos abiertos y legibles por máquinas son CSV, RDF, RSS, JSON, ODT, ODS, GML, KML, SHP, WMS y XML. Es recomendable ofrecer los datos en varios formatos simultáneamente para cubrir las necesidades del mayor número de reutilizadores.

Respecto a los metadatos, el apartado V de la Norma Técnica de Interoperabilidad (España, 2013b) establece que para la descripción de los documentos y recursos de información reutilizables puestos a disposición pública se

asociarán unos metadatos mínimos, basados en estándares como DCAT y Dublin Core. También se indica que siempre que sea posible se utilizarán vocabularios estándar o de amplio uso para asegurar la interoperabilidad semántica, debiendo estos vocabularios ser publicados y gestionados desde el Centro de Interoperabilidad Semántica de la Administración.

Martín-González y Sainz-Iglesias (2015) estudiaron la aplicación de la Norma Técnica de Interoperabilidad de Reutilización de Recursos de la Información en la descripción de los datos abiertos de las Comunidades Autónomas y su asignación de metadatos. Concluyeron que los catálogos autonómicos cumplen de modo dispar los requisitos exigidos por dicha norma en lo que se refiere a la inclusión de metadatos obligatorios, y que la recuperación de la información requerida por el usuario se ve obstaculizada por la heterogeneidad en la designación de metadatos obligatorios y adicionales.

#### **n) Volumen de datos**

Cadena-Vela et al. (2019) habla de "relevancia" y Venegas Álvarez (2018) de "volumen de datos" para referirse a la cantidad adecuada de datos. Para este último dicha cantidad hace referencia al tamaño del archivo donde se encuentra almacenado el conjunto de datos: menos de 250 megabytes es fácil de manipular, hasta 1 gigabyte se puede manipular de manera manejable, de 1 a 4 gigabytes es un archivo difícil de manejar, y más de 4 gigabytes es intratable.

Respecto al volumen de datos del portal, sería difícil definir cuál es la cantidad adecuada de datos, por lo que se tendrá en cuenta que el portal se haya ido nutriendo de datos a lo largo de estos años.

#### **o) Valor y Utilidad**

Lourenço (2015) incluye como característica de calidad de los datos su "valor y utilidad". Esto está en relación con el hecho de que liberar un montón de datos no significa que el gobierno abierto sea exitoso. Se deben liberar datos que sean útiles y tengan valor para el usuario. Se espera que los portales de datos abiertos liberen datos que permitan a los ciudadanos evaluar la conducta de los cargos públicos en los que han delegado poderes y responsabilidad.

Teniendo en cuenta que la calidad de los datos suele relacionarse con su "adecuación al uso", se ha incluido en este indicador lo que Venegas Álvarez (2018) denomina "pertinencia", que es el grado en el que un conjunto de datos satisface la necesidad del usuario, y lo que Oviedo et al. (2015) llaman "relevancia", grado de frecuencia con que son consultados los datos. Consideramos que unos datos muy consultados satisfarán las necesidades de más usuarios.

Para saber qué datos son útiles y tienen valor para el usuario hay que tener en cuenta sus peticiones, sugerencias e interacciones, por lo que este indicador está muy relacionado con Contactabilidad e Interactividad, y también con el apartado que veremos a continuación sobre la calidad del contenido de los datos.

#### **1.4.2.3. Criterios de calidad del contenido de los datos**

La *Carta Internacional de Datos abiertos* (Open Data Charter, 2015) indica que uno de los objetivos de los datos abiertos es "la mejora de la gobernanza y el compromiso ciudadano": los datos abiertos permiten que los ciudadanos tengan mayor conocimiento de lo que hacen los políticos.

Para lograr este objetivo, es indispensable que el contenido de los datos sea de calidad, pero en este aspecto también se están detectando carencias. Autores

como Curto-Rodríguez (2020), Máchová y Lnenicka (2017) y Royo-Montañés y Benítez-Gómez (2019) coinciden en la necesidad de divulgar en los portales de datos abiertos más información económico-financiera para incrementar los niveles de transparencia y rendición de cuentas. Por su parte, Abella et al. (2018) detectan que solo 12 de las 22 temáticas definidas en la NTI-RISP tienen representación entre los conjuntos de datos locales del portal de datos abiertos de Barcelona.

En nuestro país, la FEMP (2017 y 2019) proporciona un listado con los conjuntos de datos mínimos a publicar por las entidades locales, y la Ley 19/2013 en su artículo 8 determina la información económica, presupuestaria y estadística que como mínimo debe hacerse pública (España, 2013a).

Uno de los instrumentos disponibles para medir el nivel de transparencia de las instituciones públicas son los índices elaborados por Transparencia Internacional, movimiento global que lucha contra la corrupción promoviendo la transparencia y la rendición de cuentas, y que cuenta con una ramificación en España. El índice que analiza las comunidades autónomas se denomina Índice de Transparencia INCAU (Transparencia Internacional España, 2016).

Aunque hay autores que recomiendan no tener como base este tipo de indicadores a la hora de evaluar la transparencia o de seleccionar los datos a publicar. Atenas et al. (2019) recomiendan basarse en el apoyo a las políticas públicas y en la demanda de datos de los usuarios, y a partir de ahí priorizar la publicación de ciertos conjuntos de datos, obteniendo de manera secundaria una mejora en dichos indicadores.

Por su parte, Curto-Rodríguez (2020) elabora su propio indicador de transparencia y reutilización de datos abiertos (TOD) para evaluar la información relacionada con la rendición de cuentas contenida en los portales de datos abiertos autonómicos españoles, basándose en los mínimos indicados en la Ley 19/2013, y su reutilización de acuerdo al esquema de 5 estrellas. Justifica el uso de este indicador, en lugar del Índice de las Comunidades Autónomas elaborado por Transparencia Internacional España, argumentando que los dirigentes concentran sus esfuerzos solo en los aspectos que son evaluados por dicho índice, dejando de lado otros aspectos importantes. Tras aplicarlo en un estudio longitudinal en tres años (2013, 2015 y 2018), el indicador obtuvo unos resultados mediocres, reflejando que, salvo excepciones -entre las que se encontraría Castilla y León-, se está desatendiendo el potencial de los datos abiertos para la rendición de cuentas.

## **2. JUSTIFICACIÓN, OBJETIVOS Y METODOLOGÍA**

En este apartado se presenta el interés del tema analizado en el presente trabajo, así como los objetivos perseguidos y la metodología utilizada.

### **2.1. JUSTIFICACIÓN**

Tal y como se indica en el marco teórico de este trabajo, el campo de los datos abiertos es bastante novedoso. Según Royo-Montañés y Benítez-Gómez (2019), los portales de datos abiertos no han sido estudiados en profundidad y tampoco hay establecidas pautas que evalúen cómo se implementan. Por su parte, Curto-Rodríguez (2020) apunta que el ámbito de los datos abiertos de las comunidades autónomas está poco contemplado en la bibliografía, a pesar de la importancia que tienen en la organización territorial de nuestro país y el enorme peso de las competencias que tienen transferidas, destacando las dos con mayores partidas presupuestarias: sanidad y educación.

Por estos motivos, el presente trabajo es una buena oportunidad para plantear una serie de criterios de calidad para analizar los portales y sus datos, y aplicarlos al estudio del portal de datos abiertos de la Junta de Castilla y León, la comunidad autónoma en la que resido.

El segundo motivo para escoger este tema es un interés personal, ya que considero que el campo de los datos abiertos ofrece muchas oportunidades a los graduados en Información y Documentación; coincido con Caridad-Sebastián y Martínez-Cardama (2016) en la afirmación de que los documentalistas deben hacerse un hueco en el terreno de los datos abiertos, puesto que la ingente cantidad generada de datos exige mecanismos de control y sistematización.

Creo firmemente que debemos salir de nuestra zona de confort y esforzarnos en conocer un sector en crecimiento que acabará siendo terreno exclusivo de los profesionales de la Informática si no reivindicamos nuestra valía en aspectos tan importantes como la arquitectura de la información, la redacción de contenidos, la elaboración de vocabularios controlados o el uso de palabras clave, por citar algún ejemplo.

Según el estudio *Open data best practices in Europe's top performers* (Radu y Cecconi, 2019) España es uno de los líderes europeos en la publicación de datos abiertos, junto con Irlanda y Francia. En el *Open Data Inception* (OpenDataSoft, 2020) aparecen 307 portales de datos abiertos con la etiqueta "España", pertenecientes a hospitales, universidades, ayuntamientos, diputaciones, autonomías, ministerios y gobierno central, mientras que datos.gob.es (2020) localiza 301 iniciativas (230 de la administración local, 43 de la administración del Estado, 18 de la autonómica y 10 de universidades)

Toda esta información es aprovechada por las empresas infomediarias para generar valor económico: a 31 de diciembre de 2017 había identificadas en España 697 empresas activas, con un volumen de negocio de 1.796.778.748 € (Asedie, 2020), lo que nos da una idea de la importancia económica de la explotación de los datos y de los puestos de trabajo que genera este sector.

### **2.2. OBJETIVOS**

El objetivo general de este trabajo es determinar, mediante un marco evaluativo de elaboración propia, si el Portal de Datos Abiertos de la Junta de Castilla y León cumple con los indicadores de calidad reflejados en el marco teórico, para posteriormente proponer posibles mejoras en aquellos aspectos que lo requieran.

Este objetivo general se desgana en varios objetivos secundarios:

- a) Describir el origen, la legislación aplicable y la finalidad del portal de datos abiertos de la Junta de Castilla y León.
- b) Sintetizar los ítems de calidad planteados en diferentes marcos evaluativos de datos abiertos y de portales de datos abiertos.
- c) Contrastar que los ítems resultantes responden a los indicadores de calidad planteados en el marco teórico.
- d) Elaborar un marco de evaluación propio partiendo de dicha síntesis.
- e) Evaluar el cumplimiento de los ítems propuestos en el portal de datos abiertos de la Junta de Castilla y León.
- f) Valorar el contenido del portal en relación a la rendición de cuentas.
- g) Proponer mejoras acordes a los indicadores de calidad planteados.

### **2.3. METODOLOGÍA**

El primer paso para elaborar el presente trabajo fue la elaboración de un marco teórico que diera contexto a la posterior investigación. Aunque la búsqueda bibliográfica comenzó en enero de 2020, se presentaron diversas dificultades técnicas de acceso al catálogo de la BULE casi hasta finales de febrero, momento en el que se retomó la búsqueda hasta principios de abril.

La revisión bibliográfica inicial se realizó en Dialnet, centrándose en documentos publicados entre 2015 y 2020 que trataran sobre los portales de datos abiertos gubernamentales y la calidad de los datos abiertos. Posteriormente se amplió la búsqueda a otras fuentes como Google Académico, Web of Science, Scopus, WorldCAT.org, Redalyc.org y Recolecta, y se revisó la bibliografía citada en los documentos recopilados para buscar documentos relevantes.

Toda la bibliografía utilizada está en formato digital, que es el vehículo más rápido a la hora de divulgar estudios científicos.

Se han utilizado principalmente documentos en español, aunque también se han seleccionado algunos en inglés por proceder de organismos internacionales o de investigadores que escriben en esa lengua y plantean en sus trabajos marcos de evaluación, tanto de portales de datos abiertos como de sus datos, basados en criterios de calidad. Se han localizado algunos estudios muy técnicos enmarcados en el campo de la informática que no se han utilizado en nuestro trabajo.

#### **2.3.1. Elaboración del marco de evaluación**

Coincidimos con Herrera-Melo y González-Sanabria (2019, p. 18) en la afirmación de que "cada metodología de análisis propone una aproximación diferente que conduce a que los elementos estudiados (portal o datos) tengan diferentes niveles de calidad, dependiendo de la metodología utilizada. Sin embargo, no es desconocido que la aproximación a un resultado de más calidad es dado por la combinación de metodologías y modelos que permitan cubrir un mayor número de aspectos".

Nuestro marco de evaluación (Tabla III) es una síntesis de diferentes marcos propuestos en la literatura para evaluar la calidad de datos y portales abiertos; además incluye los elementos básicos y recomendables que debe tener un portal de datos abiertos según la FEMP y la Junta de Castilla y León:

<b>Autores</b>	<b>Objeto de evaluación del estudio</b>
(Berners-Lee, 2010)	Nivel de apertura y usabilidad de los datos abiertos
(Medina Sánchez, 2015)	Características de los portales de datos abiertos a nivel local
(Oviedo et al., 2015)	Calidad y madurez de los portales de datos abiertos
(Abella et al., 2018)	Calidad y reutilización de los datos abiertos a nivel municipal
(Herrera-Melo & Gonzalez-Sanabria, 2019)	Calidad de datos abiertos y del portal que los ofrece
(Lourenço, 2015)	Estructura y organización de portales de datos abiertos a nivel nacional y su adecuación a la rendición de cuentas
(Vetró et al., 2016)	Calidad de los datos abiertos a nivel de dataset y celda
(Máchová & Lnenicka, 2017)	Calidad de portales de datos abiertos a nivel nacional
(Venegas Álvarez, 2018)	Nivel de apertura, usabilidad y accesibilidad de los datos abiertos, en especial los de transparencia
(Abeleida Díaz et al., 2019)	Calidad de los datos abiertos orientados a transparencia en portales de datos abiertos de gobiernos locales
(Royo-Montañés & Benítez-Gómez, 2019)	Características de los portales municipales de datos abiertos e información económico-financiera ofrecida
(Curto-Rodríguez, 2020)	Existencia de datos económico-financieros y calidad reutilizadora

*Tabla III. Marcos de evaluación sintetizados*

Para evaluar la calidad reutilizadora nos hemos basado en el esquema de 5 estrellas de Berners-Lee, porque estamos de acuerdo con Curto-Rodríguez (2020, p. 5) cuando afirma que "es un indicador sencillo, fácilmente comprensible por todos los interesados, ampliamente avalado y con mayor utilización en la investigación académica" que Meloda.

Todos los ítems propuestos en los marcos de evaluación citados se volcaron en una base de datos Access para facilitar su síntesis; después se agruparon en un mismo ítem los que analizan la misma característica, o se elaboraron ítems más complejos uniendo ítems sencillos propuestos por diferentes autores.

Una vez sintetizados, se contrastó su correspondencia con los indicadores de calidad planteados en el marco teórico y se clasificaron en tres grupos: Características del portal (subdivididas en Aspectos técnicos, Usabilidad y acceso, y Comunicación y participación), Características de los datasets, y Contenido del portal en relación a la rendición de cuentas (ver Anexos II y III). Esta clasificación se basa en la planteada por Máchová y Lnenicka (2017) en su marco de evaluación.

### 2.3.2. Selección de los datasets objeto de evaluación

Serán objeto de evaluación el catálogo del conjunto de datos del portal y varios datasets de muestra. El portal de datos abiertos ofrece una herramienta de análisis de datos con la que se puede analizar un subconjunto de todos los datasets del portal; como veremos más adelante en el epígrafe dedicado a la evaluación, la Junta ha priorizado la relevancia y el volumen de información de los datasets, entre otros factores, para su inclusión en esta herramienta. Por ello, el primer criterio de selección de los datasets de muestra es que puedan analizarse con ella.

El segundo criterio es la popularidad de los datasets, índice que calcula la propia plataforma utilizada por el portal, basándose en 3 métricas: número de descargas, número de llamadas de API y número de reutilizaciones. Este índice se ha consultado en el catálogo de la herramienta de análisis, descargado en formato CSV el día 11/07/2020. Después se ha seleccionado el dataset más popular dentro de cada tema, y se han ordenado de mayor a menor, para seleccionar los 5 con mayor popularidad. No se han seleccionado los datasets más populares de manera global para tener mayor representatividad de temas:

	A	B	C	D	E
1	Temas	Nº datasets	Datasets más populares	Popularidad	Link
2	Medio ambiente	8	Límites provinciales de Castilla y León: recintos	2394	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
3	Salud	48	Situación epidemiológica coronavirus (COVID-19) en Castilla y León	2001	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
4	Comercio	4	Comercios con entrega a domicilio	326	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
5	Empleo	8	Situación de solicitudes de ERTES coronavirus	141	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
6	Cultura y ocio	9	Eventos de la agenda cultural categorizados y geolocalizados	137	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
7	Energía	10	Consumos horarios en hospitales	118	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
8	Hacienda	2	Ejecución del presupuesto de la Administración de la Comunidad – gastos	71	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
9	Sector público	40	Contratos de emergencia	67	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
10	Turismo	2	Registro de Turismo de Castilla y León	66	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
11	Educación	1	Directorio de Centros Docentes	30	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
12	Seguridad	1	Estadísticas de siniestralidad laboral	27	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
13	Medio rural y pesca	11	Explotaciones ganaderas de bovino	26	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
14	Demografía	6	Población total por edades y sexo	25	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
15	Transporte	2	Incendios en la red de carreteras titularidad de la Junta de Castilla y León	24	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
16	Sociedad y bienestar	9	Alojamientos Juveniles	21	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
17	Economía	7	Relación de Ferias Comerciales	21	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
18	Ciencia y tecnología	2	TuCertCyL – Usuarios activos	3	<a href="https:// analisis.datosabiertos.">https:// analisis.datosabiertos.</a>
19	TOTAL	170			
20					

Tabla IV. Datasets más populares el día 11/07/2020

En consecuencia, han sido escogidos para su análisis los 5 datasets siguientes:

- Límites provinciales de Castilla y León: recintos.
- Situación epidemiológica coronavirus (COVID-19) en Castilla y León.
- Comercios con entrega a domicilio.
- Situación de solicitudes de ERTES coronavirus.
- Eventos de la agenda cultural categorizados y geolocalizados.

Por último, también se han seleccionado tres datasets que publican los mismos datos, pero en tres años diferentes:

- Presupuestos generales de la Comunidad de Castilla y León para 2016.
- Presupuestos generales de la Comunidad de Castilla y León para 2017.
- Presupuestos generales de la Comunidad de Castilla y León para 2018.

Su selección viene motivada por la necesidad de analizar la publicación de datos en forma periódica (característica que no tienen los 5 datasets anteriores), y por su importancia para la rendición de cuentas.

### 3. ANÁLISIS DEL PORTAL DE DATOS ABIERTOS DE LA JUNTA DE CASTILLA Y LEÓN

Dedicaremos este apartado a presentar el portal de datos abiertos de la Junta de Castilla y León, para conocer su origen y normativa aplicable, ver cómo es su estructura general y obtener una comprensión global de su funcionamiento, para después aplicar el marco de evaluación propuesto.

#### 3.1. ORIGEN Y NORMATIVA APLICABLE

Tal y como recoge el propio portal de datos abiertos, éste se enmarca en el Modelo de Gobierno Abierto de la Junta de Castilla y León, aprobado por Acuerdo 17/2012, de 8 de marzo (Junta de Castilla y León, 2012); sus objetivos son aumentar la transparencia y conseguir la participación y colaboración de los ciudadanos y empresas, intercambiando conocimientos y experiencias para conseguir el avance conjunto de la iniciativa pública y privada, objetivos que van en consonancia con la filosofía de gobierno abierto analizada más arriba<sup>2</sup>.

Respecto a la legislación aplicable al portal, podemos destacar la siguiente:

- **en el ámbito comunitario:** la Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo en materia de apertura de datos del sector público y su reutilización (Unión Europea, 2019), y la Directiva 2013/37/UE por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información en el sector público (Unión Europea, 2013).
- **En el ámbito estatal** destacamos la Ley 37/2007 de Reutilización de Información del Sector Público (España, 2007), modificada por Ley 18/2015 (España, 2015), y el Real Decreto 1495/2011 que concreta la aplicación de la Ley 37/2007 en el ámbito del sector público estatal (España, 2011).
- **En el ámbito autonómico,** además del citado Acuerdo 17/2012, es de aplicación la Ley 2/2010, de 11 de marzo, de Derechos de los Ciudadanos en sus relaciones con la Administración de la Comunidad de Castilla y León y de Gestión Pública (Junta de Castilla y León, 2010).

#### 3.2. ORGANIZACIÓN GENERAL DEL PORTAL

El aspecto general que presenta el portal en su página de inicio<sup>3</sup> es el que puede verse en la Figura 1. Su diseño es definido en la guía de estilos web de la Junta como "creativo, moderno, accesible, usable y adaptado para su visualización en todos los dispositivos", y se utiliza en los más de 50 portales alojados en la plataforma web corporativa para que todos ellos compartan una estructura y diseño normalizados.

Como se puede apreciar, presenta una barra de menús desplegables en la cabecera, diversos enlaces y utilidades en el cuerpo de la web, incluyendo un buscador, y a los pies se repiten los menús de la cabecera, pero desglosados y visibles en su totalidad:

---

<sup>2</sup> En septiembre de 2019 La Junta de Castilla y León solicitó su adhesión a la Carta Internacional de Datos Abiertos:

[https://comunicacion.jcyl.es/web/jcyl/Comunicacion/es/Plantilla100Detalle/1284882246693/\\_/1284892465984/Comunicacion](https://comunicacion.jcyl.es/web/jcyl/Comunicacion/es/Plantilla100Detalle/1284882246693/_/1284892465984/Comunicacion)

<sup>3</sup> <https://datosabiertos.jcyl.es/web/es/datos-abiertos-castilla-leon.html>

**Junta de Castilla y León**

🔍

📄
✉
↑
Otros temas

↑ DATOS ABIERTOS
INICIATIVA DE DATOS ABIERTOS
CATÁLOGO DE DATOS
ANÁLISIS DE DATOS ABIERTOS
APLICACIONES
CONCURSO DE DATOS ABIERTOS
PARTICIPA

## DATOS ABIERTOS de Castilla y León

### DATOS ABIERTOS DE CASTILLA Y LEÓN

Visualización sobre evolución del coronavirus

Novedades y suscripción

Preguntas frecuentes

Análisis de datos

### BÚSQUEDA DE DATOS

BUSCAR
? Ayuda

### CONJUNTOS DE DATOS DESTACADOS

Situación epidemiológica coronavirus (COVID-19) en Castilla y León por provincias 📄 📄

---

Situación enfermos por coronavirus detectados en atención primaria 📄 📄

---

Situación enfermos por coronavirus detectados en atención primaria por tramos de edad y sexo 📄 📄

---

Tasa de enfermos por zonas básicas de salud 📄 📄

---

Situación de solicitudes de ERTES coronavirus 📄 📄

### ENLACES DE INTERÉS

Cuéntanos tu idea

IDECyL

Guías Open Data

Gobierno Abierto  
Portal Gobierno Abierto

**INICIATIVA DE DATOS ABIERTOS**

- ¿Qué son los datos abiertos?
- El proyecto en la Junta de Castilla y León
- Normativa
- Guías Open Data
- Enlaces de interés
- Buenas prácticas

**CATÁLOGO DE DATOS**

- Buscador de conjuntos de datos
- Listado de conjuntos de datos
- Términos de uso
- Formatos de datos abiertos

**ANÁLISIS DE DATOS ABIERTOS**

**APLICACIONES**

**CONCURSO DE DATOS ABIERTOS**

- III Concurso de datos abiertos 2018
- Ediciones anteriores

**PARTICIPA**

- Te ayudamos a desarrollar tu idea
- Solicitud de apertura de datos
- Solicitud de difusión de aplicación con datos abiertos
- Comentarios

Aviso Legal
Protección de datos
Accesibilidad
Contacto
Guía de estilos web

Figura 1. Página de inicio del Portal de Datos Abiertos de la Junta de Castilla y León

24

A la hora de utilizar el portal es importante destacar un aspecto que ya se ha mencionado en el epígrafe anterior, y es que podemos acceder a los conjuntos de datos de dos maneras diferentes; una es el menú desplegable “Catálogo de datos”, que da acceso al listado de todos los conjuntos de datos albergados en el portal y al buscador de dichos conjuntos (Figura 2):



Figura 2. Menú del catálogo de datos

Otra forma de acceder a los datasets es a través del menú “Análisis de datos abiertos”, que conduce a un portal de análisis, puesto en marcha en noviembre de 2017 (Figura 3); esta herramienta no permite trabajar con todos los datasets del portal de datos abiertos, sino solo con una parte<sup>4</sup>, debido a los motivos siguientes<sup>5</sup>:

- Algunos conjuntos de datos son capas cartográficas en formato SHP, por lo que la subida al portal de análisis no aporta ningún valor;
- Algunos conjuntos de datos tienen un volumen de información muy reducido, con poco margen para el análisis;
- El número de conjuntos de datos que pueden subirse a la plataforma de análisis es limitado (actualmente no puede superar los 400), por tratarse de un servicio en la nube contratado por la Junta.
- Sincronizar los datos de ambos portales supone un importante esfuerzo, por lo que se ha priorizado la subida de los más relevantes.

<sup>4</sup> A día 15/07/2020, el número total de datasets era de 525 y el portal de análisis permitía analizar 171.

<sup>5</sup> Explicación obtenida tras consulta al servicio de soporte del portal el día 10/07/2020.

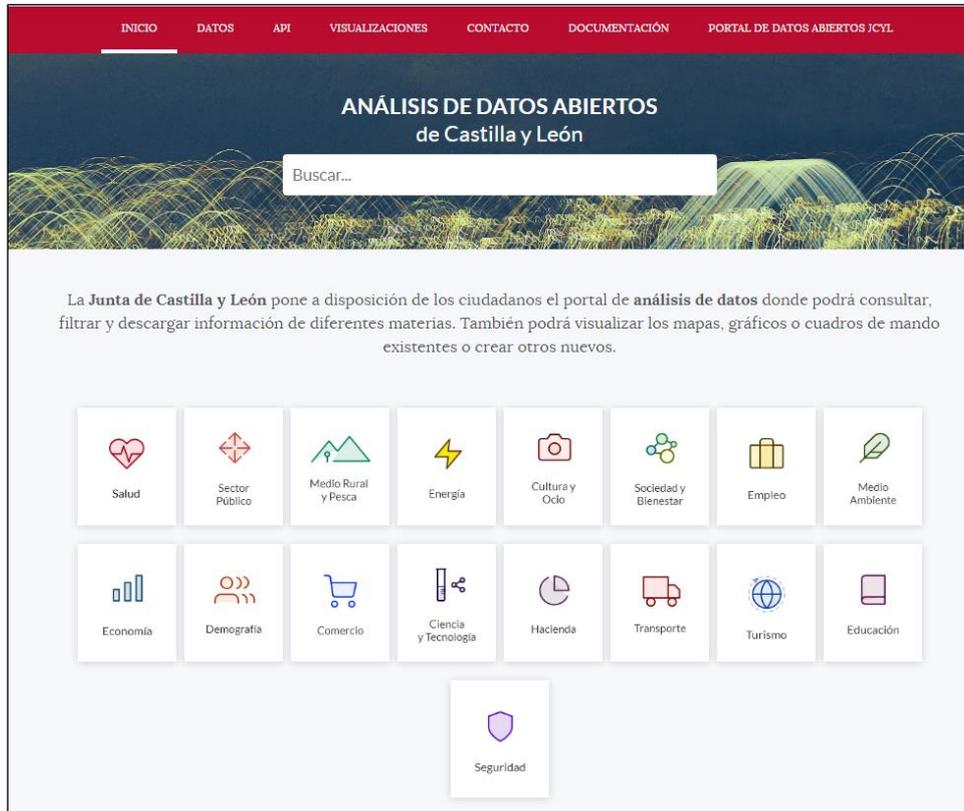


Figura 3. Portal de análisis de datos

### 3.3. APLICACIÓN DEL MARCO DE EVALUACIÓN PROPUESTO

Una vez observada la disposición general del portal de datos y de la herramienta de análisis que le sirve de complemento, procederemos a aplicar el marco de evaluación, que agrupa los ítems a evaluar en tres grandes categorías, tal y como mencionamos en el apartado anterior: Características del portal (subdivididas en Aspectos técnicos, Usabilidad y acceso, y Comunicación y participación), Características de los datasets, y Contenido.

#### 3.3.1. Características del portal

##### 3.3.1.1. Aspectos técnicos

- **Presenta una descripción del portal, misión y objetivos:** Sí; este es uno de los elementos básicos recomendados por la guía de la propia Junta de Castilla y León, por lo que sería extraño que no estuviera presente. En el apartado "Iniciativa de datos abiertos" se encuentra información sobre qué son los datos abiertos, el origen del portal y sus objetivos, además de proporcionar la normativa aplicable al portal, guías explicativas del movimiento de Datos Abiertos y diversos enlaces de interés a iniciativas similares.
- **Condiciones de uso o términos de reutilización:** Sí, se pueden encontrar en <https://datosabiertos.jcy.es/web/es/catalogo-datos/terminos.html> y también en <https://analisis.datosabiertos.jcy.es/terms/terms-and-conditions/>. Se especifica que el acceso a los datasets supone su aceptación, y engloban aspectos como la

licencia aplicable a los conjuntos de datos, la obligación de citar a la Junta como fuente de los datos, la aceptación de futuros cambios en los conjuntos de datos y términos de uso, la imposibilidad de garantizar completamente la actualización, disponibilidad, completitud e integridad de los datasets, y la exclusión de responsabilidad de la Junta por daños o pérdidas causados por el uso de los datos, especificándose la responsabilidad del reutilizador por incumplimiento de los términos de uso y la posibilidad de cancelar, por dicho incumplimiento, el acceso a los datos.

En este apartado se indica que la licencia aplicable a los conjuntos de datos, salvo indicación contraria, es la Licencia Creative Commons-Reconocimiento (CC-by 3.0). Se explica de manera detallada qué permite y qué prohíbe dicha licencia, además de la obligación de reconocer/citar a la Junta como fuente de los conjuntos de datos.

➤ **Portal multilinguaje:** este aspecto es algo curioso; si observamos a pantalla completa el portal de datos o la herramienta de análisis, no vemos ningún enlace a la visualización en inglés; pero si en la herramienta de análisis reducimos la pantalla hasta que el menú se convierta en un desplegable (lo que suele llamarse “hamburguesa”) y pinchamos en él, veremos que en la parte superior aparece la opción de escoger el idioma inglés (Figura 4), aunque solo veremos en inglés el menú, los filtros y algún aspecto más; esto es algo que observamos directamente en el teléfono móvil, y probablemente sea algún fallo en el *responsive* (adaptación de la web a los diversos tamaños de los dispositivos):

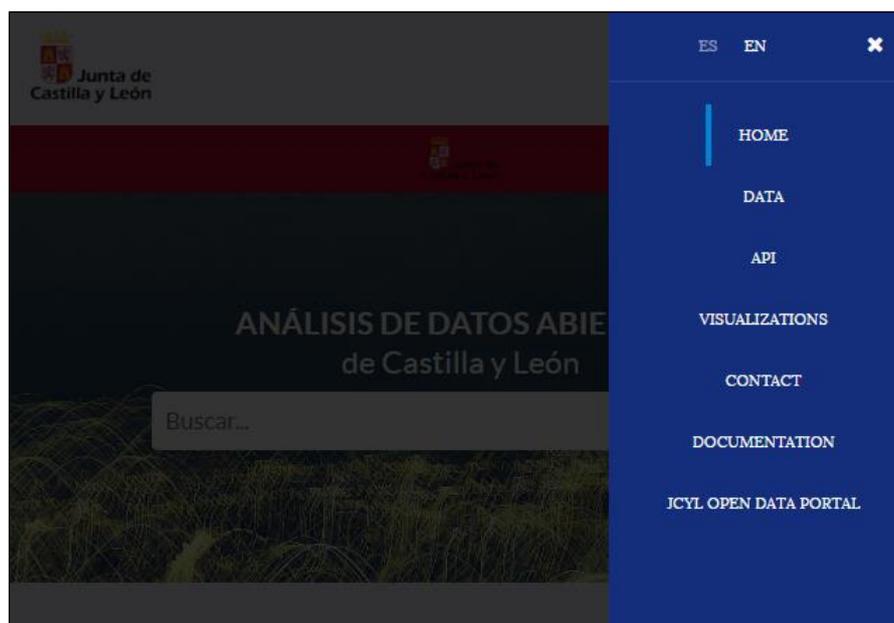


Figura 4. Acceso al idioma inglés

➤ **Accesibilidad para personas con dificultades:** entendida como “el conjunto de principios y técnicas que se deben respetar al diseñar, construir, mantener y actualizar los sitios web y las aplicaciones para dispositivos móviles para garantizar la igualdad y la no discriminación en el acceso de las personas usuarias, en particular de las personas con discapacidad y de las personas mayores” (España, 2018).

En el pie del portal aparece de forma constante un enlace a su política de accesibilidad, donde se recoge su declaración de accesibilidad fundamentada en la certificación llevada a cabo por parte de ILUNION Tecnología y Accesibilidad, que

garantiza el cumplimiento del nivel de accesibilidad Doble-A WCAG 2.0. El portal es conforme con el Real Decreto 1112/2018, de 7 de septiembre, sobre accesibilidad de los sitios web y aplicaciones para dispositivos móviles del sector público, y con la Norma UNE-EN 301549:2019, considerando las excepciones de dicho Real Decreto.

- **Tiene algún certificado:** tiene el certificado de accesibilidad de ILUNION, pero en lo referente a certificados que apoyen la credibilidad del sitio, auditorías, etc. no hemos localizado ninguno.
- **El portal tiene mecanismos externos de control de la calidad, y asocia los resultados de dichos procesos de revisión con cada dataset publicado:** en el apartado de Accesibilidad se indica que "se realiza un trabajo continuo de auditoría de sitios web, con un doble objetivo: presentar los contenidos más claros y de forma estructurada, facilitando su interpretación; y mejorar la navegación y la experiencia del usuario", aunque no se indica expresamente que busquen una mejora en la calidad de los datos.
- **El portal presenta el nombre de la organización responsable de almacenar o custodiar los datos de forma permanente:** Sí, la ficha correspondiente al catálogo de datos indica que el responsable del contenido es la Dirección General de Transparencia y Buen Gobierno - Consejería de Transparencia, Ordenación del Territorio y Acción Exterior.
- **Información sobre el sistema de gestión de datos utilizado:** En el acceso a la herramienta de análisis de datos abiertos se indica que está basada en la tecnología Opendatasoft, que es una plataforma SaaS ("en la nube") desarrollada para que los usuarios profesionales compartan, publiquen y reutilicen fácilmente conjuntos de datos estructurados. Respecto al portal de datos abiertos, no se indica nada expresamente.
- **Se indican las buenas prácticas utilizadas en el portal:** Sí, se resumen en los siguientes puntos:
  - Publicar datos en formatos abiertos y estándares;
  - Catálogo de datos bien estructurado;
  - Estructura URIs amigables, persistente y extensible;
  - Implicación de toda la organización;
  - Compromiso de servicio, actualización y calidad del dato;
  - Datos bajo condiciones de uso no restrictivas y comunes;
  - Implicación de la ciudadanía.

### 3.3.1.2. Usabilidad y acceso

- **Registro obligatorio o pago para el acceso y/o consulta de los datos públicos:** No es necesario registrarse para acceder al portal, consultar la información que contiene o utilizar los diferentes servicios. Respecto al cobro, los términos de uso del portal indican que la Junta se reserva el derecho de publicar conjuntos de datos para los que sea necesario satisfacer el pago de contraprestaciones económicas cuya cuantía tendrá como objetivo cubrir los costes de producción, reproducción y difusión de los conjuntos de datos, no pudiendo nunca ser superior a dichos costes.
- **Soporte para creación de cuentas de usuario para personalizar las visualizaciones e información mostrada:** Sí que se puede crear una cuenta de usuario, pero ha sido difícil averiguar cómo. En la esquina superior derecha del portal de análisis aparece el enlace para iniciar sesión; al pincharlo aparece un cajetín donde escribir el nombre de usuario/correo y la contraseña, y una opción por si se nos ha olvidado la contraseña, pero no se ofrece la opción de crear una cuenta nueva. La

Guía de usuario no indica cómo crear la cuenta; para crearla hay que acceder a "Visualizaciones", crear un mapa o un gráfico e intentar guardarlo; entonces es cuando el portal da la opción de crear una cuenta.

➤ **Existencia de mecanismos para agregar, modificar o eliminar conjuntos de datos sin validación de usuario:** No; una respuesta positiva implicaría un grave problema de seguridad y credibilidad. Según la guía de usuario de Opendatasoft, para acceder al *back office* (interfaz de administración) donde es posible publicar datos, crear contenido, y gestionar y configurar el dominio, es necesario estar registrado como usuario y además disponer de los permisos necesarios.

➤ **La publicación de datos respeta la legislación vigente en materia de privacidad:** Sí, en el apartado de la normativa aplicable al portal podemos destacar la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal y la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Además, la licencia aplicable a los conjuntos de datos indica que "derechos como los de publicidad, privacidad, o los derechos morales pueden limitar el uso del material".

➤ **Proporciona un catálogo de datos:** Sí lo proporciona, en él se listan todos los datasets disponibles en el portal junto con sus características, facilitando su ubicación. Para visualizarlo hay que descargarlo (está disponible en formatos CSV, JSON, RDF, RSS y XML) o puede visualizarse en pantalla haciendo una búsqueda vacía en el buscador, como veremos más adelante. Esta vía para ver el catálogo completo no aparece explicitada en el portal, y puede ser interesante para aquellos usuarios que no quieran/puedan descargarse el catálogo, pero no permite ordenar los resultados ni una visualización ágil (los datasets se visualizan en pantalla de 10 en 10, en 54 páginas).

➤ **Indica el número de conjuntos de datos:** el portal de datos abiertos no lo indica de manera directa, es necesario descargarse el catálogo de datos y comprobar el número de registros. También puede hacerse una búsqueda sin introducir ningún término, tal y como explicamos en el ítem anterior, para que aparezca un listado que indica el total de conjuntos de datos. Sin embargo, en el portal de análisis podemos saber el número de datasets seleccionando el menú "Datos" que aparece en la cabecera del portal: aparecen el listado y el número.

➤ **Volumen de datos:** De acuerdo a Curto-Rodríguez (2020), el portal ofertaba 124 datasets a finales de 2013, 176 en 2015 y 334 en 2017. En julio de 2020 tiene publicados 525 datasets. Como vemos que el número de datasets aumenta cada año, podemos valorar positivamente este ítem.

➤ **Posibilidad de desplazarse a través de las diferentes secciones del portal, conociendo completamente el propósito de cada una:** el portal tiene una navegabilidad sencilla que permite al usuario desplazarse por las secciones y encontrar información rápidamente, e implementa elementos para facilitar la navegabilidad en el sistema (menús, accesos rápidos a información relevante, botones de ayuda y de contacto, barras de navegación).

➤ **Organización en categorías temáticas. Utilización de estándares para la categorización de los conjuntos de datos:** entre las buenas prácticas manifestadas en el portal, se encuentra la de tener un catálogo de datos bien estructurado y ofrecido en formato DCAT, que es el estándar utilizado a nivel nacional en la Norma Técnica de Interoperabilidad (NTI) y a nivel europeo para la publicación de datos en el sector público.

El portal de análisis de datos presenta los datasets categorizados en 16 temas, como vimos en la Figura 3; estos se especifican en la taxonomía de sectores primarios del anexo IV de la Norma Técnica de Interoperabilidad. De los 22 sectores presentes en dicho anexo, aquí no se utilizan "Deporte", "Industria", "Legislación y justicia", "Urbanismo e infraestructuras" ni "Vivienda".

En el catálogo de conjuntos de datos descargado en CSV también se puede observar esta categorización temática en la columna "Sector":

	A	B
1	Título	Sector
2	Mortalidad en el primer trimestre	salud
3	Calificaciones de Zonas de Agua de Baño	salud
4	Actividad Analítica de los Laboratorios de Control Oficial	salud
5	Actividad Inspector en los Mataderos	salud
6	Censo de actividades turísticas complementarias	turismo
7	Acuerdos de Consejos de Gobierno	sector-publico
8	Agenda de trabajo de altos cargos	sector-publico
9	Relación de eventos de la agenda cultural categorizados y geolocalizados	cultura-ocio
10	Registro de agentes de cooperación para el desarrollo de Castilla y León	sociedad-bienestar
11	Apuestas	economía
12	Avisos de plagas agrícolas en la comunidad de Castilla y León	medio-rural-pesca
13	Explotaciones ganaderas de apícola	medio-rural-pesca
14	Explotaciones ganaderas de cunícola	medio-rural-pesca
15	Explotaciones ganaderas de avícola	medio-rural-pesca
16	Bibliotecas, bibliobuses y puntos de servicio móvil geolocalizados	cultura-ocio
17	Bienes inmuebles cedidos a terceros	sector-publico
18	Boletín Oficial de Castilla y León (BOCyL)	sector-publico
19	Fiestas locales: Calendario de Fiestas de Carácter Local	empleo
20	Calendario normativo	legislacion-justicia
21	Calidad de las Aguas de Consumo Humano	salud
22	Calidad de las Aguas de Recreo	salud
23	Calidad del aire (por días)	medio-ambiente, salud
24	Calidad del aire (por horas)	medio-ambiente, salud
25	Catálogo de carreteras de la red complementaria local	transporte
26	Catálogo de carreteras de la red complementaria preferente	transporte
27	Catálogo de conjuntos de datos	ciencia-tecnología, sector-publico

Figura 5. Categorías temáticas en el catálogo de datos

En este listado sí que hemos encontrado utilizados los temas Deporte (5 datasets), Industria (4), Legislación-justicia (1) y Urbanismo-infraestructuras (12).

➤ **Utilización de palabras clave en el catálogo:** Sí, y como vemos en la Figura 6, están diferenciadas de las categorías (temas). Se ha observado en el catálogo de datasets descargado en CSV que alguna palabra clave es errónea ("Carreteras") y que no parece seguirse un criterio común en su construcción: por ejemplo, en lugar de "sede electrónica" están definidas como palabras clave "sede" y "electrónica"; también están utilizadas las palabras "altos" y "cargos", existiendo además "altos cargos" (y cada palabra se utiliza solo para un dataset); encontramos separadas las palabras "servicio" y "móvil" (utilizados en dos datasets de bibliobuses), "coyuntura" y "económica" (usados en 6 datasets) y "actividades" y "económicas", cuando por otro lado existen "actividades turísticas", "actividades alimentarias" y "actividades formativas".



Figura 6. Palabras clave del catálogo

- **Posibilidad de ordenar los datos en el catálogo según criterios diferentes a las categorías:** Sí. Los conjuntos de datos disponibles en la herramienta de análisis pueden ordenarse por los siguientes criterios (Figura 7):

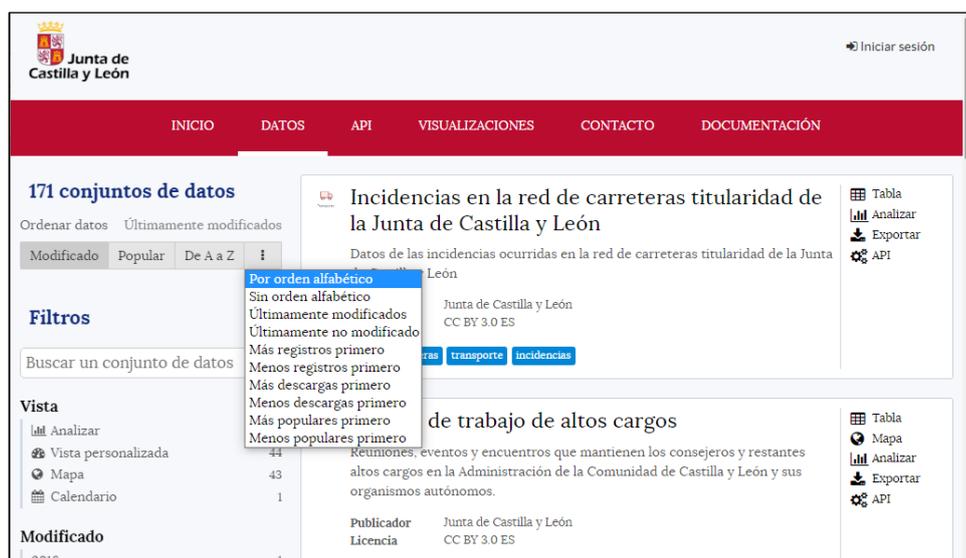


Figura 7. Criterios de ordenación de los datasets

En el catálogo de datos general no disponemos de esta opción; para ordenar el listado de conjuntos de datos debemos descargarlo en CSV y realizar la acción con Excel.

- **Existencia de listas maestras de todas las entidades y fuentes relevantes de conjuntos de datos, actuales y pasadas:** No, y serían de gran utilidad en la búsqueda de conjuntos de datos.
- **Herramientas de búsqueda y selección proporcionadas:** Se han realizado búsquedas tanto en el portal de datos como en la herramienta de análisis:

a) Búsqueda en el portal de datos abiertos

El portal proporciona un campo de búsqueda sencilla, nada más (Figura 8). No hay listas de etiquetas, metadatos, lenguajes controlados o una estructura de la organización completa a los que poder recurrir para orientar la búsqueda desde este punto.

La ayuda situada al lado tampoco proporciona un listado de temas, entidades, palabras clave, etc. que puedan facilitar la búsqueda. Da algunas pautas sobre cómo funciona el buscador (descarta palabras vacías e ignora acentos, mayúsculas y minúsculas; los resultados aparecen ordenados por relevancia; se puede acotar las búsquedas con comillas). Se enumeran varias particularidades de los buscadores específicos de la Junta (Subvenciones y Ayudas, Licitaciones, Normativas), pero no hay nada específico de datos abiertos ni se indica que pueden utilizarse operadores booleanos, aunque hemos comprobado que sí (la búsqueda "comercios NOT entrega" no recupera el dataset "comercios con entrega a domicilio", por poner un ejemplo).

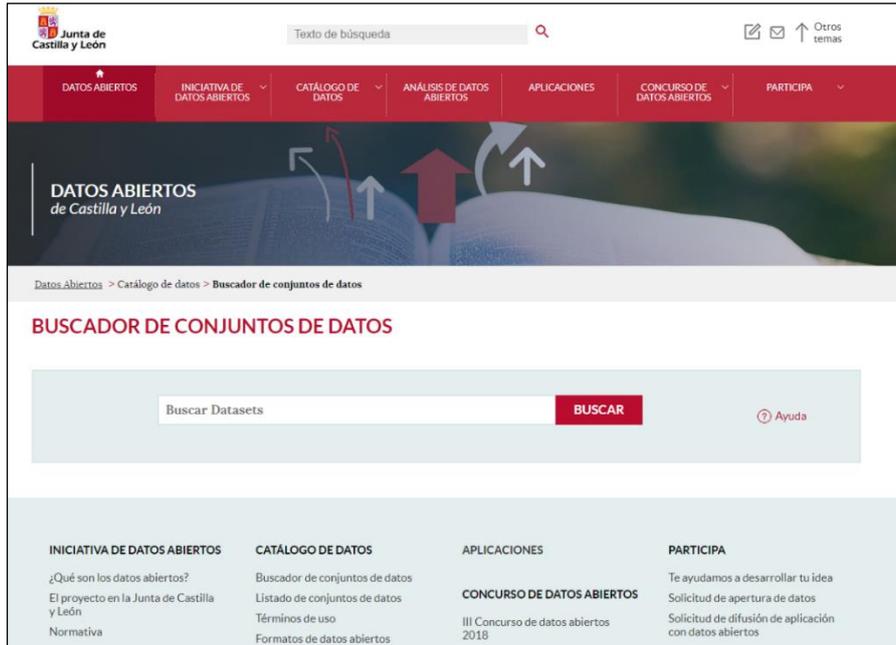


Figura 8. Buscador del portal de datos abiertos

Una vez realizada la búsqueda, se puede filtrar por formato de los datos (26 formatos) y por materia (las 22 categorías temáticas recogidas en la NTI):



Figura 9. Filtros del buscador del portal de datos abiertos

Después de filtrar, si queremos volver al listado completo seleccionamos el filtro “cualquiera” en ambas opciones. Como comentábamos más arriba, si se hace una búsqueda sin introducir ningún término, aparecen como resultado todos los datasets, recogidos de 10 en 10 en 54 páginas.

## b) Búsqueda en el portal de análisis de datos

Además de proporcionar un campo de búsqueda sencilla, se puede filtrar desde un primer momento por los enlaces directos a cada tema, y una vez obtenido el resultado, se puede filtrar por Vista, Fecha de modificación, Palabra clave y Tema (Figura 10). Borrando todos los filtros vemos el listado completo. Por lo tanto, en el buscador no hay un listado de palabras clave en el que podamos buscar el término que más se ajuste a nuestras necesidades, pero dentro de cada tema o de cada dataset sí podemos filtrar por términos concretos y específicos:

The screenshot shows the 'Portal de Datos Abiertos JCYL' interface. At the top, there is a navigation bar with links: INICIO, DATOS, API, VISUALIZACIONES, CONTACTO, DOCUMENTACIÓN, and PORTAL DE DATOS ABIERTOS JCYL. Below the navigation bar, the main content area is titled '9 conjuntos de datos'. On the left side, there is a sidebar with filters: 'Filtros activos' (Cultura y ocio), 'Filtros' (Buscar un conjunto de datos), 'Vista' (Analizar, Mapa), 'Modificado' (2016, 2020), 'Palabra clave' (cultura, directorio, bibliotecas, culturales, información, archivos), and 'Tema' (Cultura y ocio). The main content area displays four data sets, each with a title, description, publisher, license, and a list of tags. The data sets are: 1. 'Eventos de la agenda cultural categorizados y geocalizados' (tags: eventos, agenda, cultural, hoy, actual); 2. 'Relación de puntos de servicio móvil de los bibliobuses de Castilla y León' (tags: bibliobus, biblioteca, cultura, gaudes, servicio, móvil); 3. 'Directorio de industrias culturales' (tags: directorio, industrias, culturales, cultura, actividad); 4. 'Registro de infraestructuras culturales' (tags: directorio, infraestructuras, culturales, bibliotecas, archivos, museos, cultura). Each data set also has a 'Tabla' icon and a 'Mapa' icon, and a list of actions: 'Analizar', 'Exportar', and 'API'.

Figura 10. Filtros del portal de análisis

En el menú “Documentación”, que explicaremos en el siguiente ítem, hay indicaciones sobre el lenguaje de consulta para buscar conjuntos de datos en el catálogo del portal de análisis y para buscar registros dentro de un catálogo. Permite el uso de expresiones booleanas complejas, consultas a texto completo, filtros por campos, funciones avanzadas y filtros geográficos.

Hemos probado a hacer búsquedas con varias de las expresiones de consulta indicadas para buscar en los datos, pero solo hemos obtenido resultados con las 6 primeras (valor, id\_campo:valor, valor1 AND valor2, valor1 OR valor2, NOT valor, #exact(id\_campo,"valor")); las siguientes hacían que la tabla de datos se quedara en blanco.

Para comprobar el funcionamiento del buscador en las búsquedas sin expresiones (que es como suele buscar un usuario medio) se han buscado los 5 datasets seleccionados como muestra<sup>6</sup> y se han obtenido resultados dispares, ya que en algunos casos la misma búsqueda empleada en el portal de datos no obtiene

<sup>6</sup> Búsqueda realizada el 25/07/2020

resultados en el portal de análisis; los resultados varían dependiendo del uso de acentos, palabras vacías o signos de puntuación:

- “Límites provinciales de Castilla y León: recintos”: esta búsqueda literal funciona en el portal de datos, pero aquí da problemas<sup>7</sup>. Después de hacer la búsqueda con diferentes combinaciones, a veces se localiza el conjunto, y otras no:

límites provinciales (sí), límites provinciales de Castilla (no), límites provinciales Castilla (sí), límites provinciales Castilla y León (no), límites provinciales Castilla León (sí), límites provinciales Castilla Leon (no), límites provinciales Castilla Leon recintos (no), límites provinciales Castilla León recintos (sí).

Se ha hecho una búsqueda centrada en los términos “Castilla y León”, recuperando diferentes números de datasets según el uso de la tilde y la conjunción:

Castilla y León: 35; Castilla León: 169; Castilla y Leon: 7; Castilla Leon: 9

- “Situación epidemiológica coronavirus (COVID-19) en Castilla y León”: este dataset también se puede localizar en el portal de datos si lo escribimos de manera literal, pero en el portal de análisis no, y dependiendo de la búsqueda efectuada se puede recuperar o no:

Coronavirus castilla leon (no), Coronavirus castilla león (sí), coronavirus castilla y león (sí)

Aunque no localicen el dataset en cuestión, en algunos casos estas búsquedas sí que devuelven otros resultados: la búsqueda “Coronavirus castilla leon” localiza los datasets “Tasa de enfermos por zonas básicas de salud”, “Pruebas realizadas coronavirus” y “Fases de desescalada por zonas básicas de salud”.

- “Comercios con entrega a domicilio”: este dataset lo recuperamos sin problema en ambos portales (datos y análisis) escrito de manera literal, y también buscando por palabras sueltas (entrega, domicilio, comercios, entrega a domicilio).

- “Situación de solicitudes de ERTes coronavirus”: escrito de manera literal lo encontramos en el portal de datos, pero no en el de análisis:

Solicitudes de ERTes (no), solicitudes ERTES (sí), ERTES (sí), situación de solicitudes (no), situación solicitudes (sí), coronavirus (sí).

- “Eventos de la agenda cultural categorizados y geolocalizados”: en el portal de datos lo encontramos escrito de manera literal, pero no en el portal de análisis:

Eventos de la agenda cultural (no), Eventos agenda cultural (sí), Agenda cultural (sí).

#### ➤ **Tutoriales, documentación y funcionalidades de ayuda para aprender a utilizar el portal y mejorar la usabilidad:**

Junto al buscador de conjuntos de datos, en el portal de datos, se incluye una ayuda que “pretende orientar al usuario en el uso del buscador de <http://www.jycl.es>

---

<sup>7</sup> Aparece el mensaje de error: *The query is invalid : Field León doesn't exist*

(sic) y optimizar en la medida de lo posible sus búsquedas, reduciendo los tiempos de espera"; esta no es una ayuda específica del portal de datos abiertos, sino general para todos los buscadores de la Junta.

En el portal de análisis se proporciona, en la barra de menú superior, un enlace a "Documentación" (ya mencionado más arriba), donde puede consultarse una guía de usuario genérica de los portales basados en la plataforma Opendatasoft; aquí se puede aprender cómo explorar el catálogo y los conjuntos de datos, y crear mapas y gráficos (aunque no encontramos, por ejemplo, cómo crear una cuenta de usuario). En la consola de API y en las herramientas de mapas y gráficos hay enlaces de ayuda que conducen a los apartados correspondientes de esta misma documentación.

- **Ejemplos de uso:** No se han localizado ejemplos de uso del portal.
- **Proporciona APIs para desarrolladores:** Sí, el portal de análisis presenta una consola de API que permite la búsqueda y consulta de conjuntos de datos, la búsqueda de registros, y la consulta de la documentación de la API.
- **Acceso a través de consultas SPARQL:** No.
- **Posibilidad de visualizar los datos en mapas o gráficos:** Sí, el menú "Visualizaciones" del portal de análisis pone a disposición del usuario la posibilidad de crear mapas y gráficos avanzados con cualquier conjunto de datos subido a la plataforma, pudiendo cruzar los datos unos con otros para obtener mayor información. Además de la visualización, permite almacenar los mapas y gráficos si disponemos de un usuario, como se comenta más arriba al hablar de la posibilidad de registrarse en el portal.
- **Posibilidad de manipular datos online:** Sí, se pueden cruzar datos, como acabamos de comentar, y también se pueden ordenar en las tablas online.

### 3.3.1.3. Comunicación y participación

- **Lista de preguntas frecuentes (FAQ):** Sí, la página principal del portal proporciona un enlace destacado a un listado de 17 preguntas frecuentes, que abarcan aspectos como el significado, utilidad y formato de los datos abiertos, cómo manejar las funcionalidades de búsqueda y análisis, y la participación del usuario solicitando datasets, proponiendo ideas o difundiendo sus aplicaciones basadas en los datos.
- **Concursos o actividades para promover el uso de los datos:** sí, además de organizar seminarios y *webinars*, el portal organiza concursos de Datos Abiertos, cuyo objetivo es "reconocer la realización de proyectos que suministren cualquier tipo de idea, estudio, servicio, sitio web o aplicaciones para dispositivos móviles, utilizando para ello conjuntos de datos del Portal de Datos Abiertos de la Junta de Castilla y León". El día 31 de julio se abrió el plazo para participar en la IV edición, con 10.500 € en premios.

La Junta también ofrece un servicio de asesoramiento para ayudar a poner en marcha ideas relacionadas con los datos abiertos.

- **Apartado de novedades y noticias relacionadas con los conjuntos de datos:** Sí, en la página principal del portal de datos abiertos hay un enlace destacado que conduce a las novedades referentes a la publicación de nuevos datos y a diversas noticias relacionadas con la iniciativa de datos abiertos, sirviendo de ayuda a la accesibilidad de los datos, ya que destaca datos nuevos que igual los usuarios no conocen.

- **Permite la suscripción:** Sí, el apartado de novedades y noticias ofrece la posibilidad de suscribirse para recibirlas por correo electrónico. Una vez suscrito, si entramos en la web con nuestro usuario, tendremos la opción de “Seguir” los conjuntos de datos que deseemos:

The screenshot shows the 'Comercios con entrega a domicilio' dataset page. It features a navigation bar with options like 'INICIO', 'DATOS', 'API', 'VISUALIZACIONES', 'CONTACTO', 'DOCUMENTACIÓN', and 'PORTAL DE DATOS ABIERTOS JCYL'. The main content area includes a search bar, a 'Filtros' sidebar with categories like 'SECTOR COMERCIAL' and 'PROVINCIA', and a 'Seguir' button. The page also displays metadata such as 'Último procesamiento' and 'Identificador de conjunto de datos'.

Figura 11. Cómo seguir un dataset

- **Registro de reutilizadores:** No se ha localizado; la FEMP (Federación Española de Municipios y Provincias, 2017) considera que es un componente básico de un portal de datos abiertos, que sirve para que los reutilizadores reciban noticias y soporte. Aunque no haya un registro como tal, el portal ofrece un formulario para que los reutilizadores soliciten difundir sus aplicaciones de datos abiertos y les ofrece soporte para poner en marcha sus ideas.
- **Mecanismos para que los usuarios soliciten nuevos conjuntos de datos, tipos o formatos de su interés:** en el menú “Participa” se proporciona un formulario para solicitar la apertura de datos; aunque no hay una opción expresa para pedir otros tipos o formatos, puede solicitarse a través de este medio.
- **Existencia de medios suficientes para comunicarse con el responsable de los datos (formulario, correo electrónico, etc.):** Sí, en el menú “Participa” de la página principal del portal hay un formulario para enviar comentarios generales, y también la herramienta de análisis proporciona un formulario en el menú “Contacto” que puede utilizarse para este fin. Por otro lado, si accedemos a un dataset en el catálogo de datos, existe la opción de realizar una consulta sobre su contenido (esta opción no existe si accedemos al dataset a través del portal de análisis).
- **Tiempo de respuesta a una pregunta formulada:** El tiempo de respuesta ha sido excelente en las preguntas formuladas antes del mes de agosto, con respuestas en un plazo menor a una hora. Pero a principios del mes de agosto se formuló otra pregunta y no se ha obtenido respuesta (probablemente solo haya una persona desempeñando esa función y esté de vacaciones).
- **Ejemplos y número de aplicaciones desarrolladas:** Sí; se proporciona un formulario para el registro de aplicaciones desarrolladas, que son visualizables y accesibles en el menú “Aplicaciones”. No se indica el número, pero tampoco es costoso contarlas, ya que son 37. También se ofrece un enlace a una relación de

iniciativas de datos abiertos de otras administraciones públicas para ver otros ejemplos de utilización de datos abiertos.

➤ **Posibilita la valoración de los usuarios y foros de discusión:** en el menú "Participa" de la página principal del portal hay un formulario para enviar comentarios generales, que podría utilizarse para expresar el valor o utilidad de los datos. También la herramienta de análisis proporciona un formulario que puede utilizarse para este fin en el menú "Contacto", pero en ambos casos son vías de comunicación con los gestores del portal, no con otros usuarios. En el portal no hay un foro para intercambiar ideas, no hay un sistema que permita recolectar clasificaciones y comentarios de los usuarios sobre un dataset o discutir conclusiones basadas en el uso de los datos.

Existen foros de participación en la web de gobierno abierto<sup>8</sup>, en la herramienta denominada "Espacios de participación", pero en ella la Junta de Castilla y León somete a participación aquellos anteproyectos de ley, proyectos de decreto, planes, programas, estrategias e iniciativas públicas de interés para el conjunto de la Comunidad, son foros predefinidos por la Junta y es algo completamente diferente a lo analizado en este ítem.

➤ **Posibilidad de compartir conjuntos de datos a través de las principales redes sociales como Facebook y Twitter:** No. Si accedemos a un dataset a través del catálogo de datos solo se ofrece la opción de enviarlo por correo electrónico. En la herramienta de análisis no hay opción de enviarlo de ninguna manera, aunque presenta la posibilidad de copiar la URL en el portapapeles para que el usuario la comparta donde quiera.

➤ **El portal tiene perfiles actualizados en redes sociales:** En el propio portal de datos abiertos no consta su presencia en redes sociales, pero en el portal de gobierno abierto<sup>9</sup> vemos que Gobierno Abierto de Castilla y León tiene perfil en Twitter<sup>10</sup>. La Junta de Castilla y León tiene perfiles en Facebook, Twitter, Youtube y Slideshare.

### 3.3.2. Características de los datasets

Los datasets se presentan en formato "ficha", tal y como indica la NTI. Para poder comprender mejor su posterior análisis a través de los diferentes ítems propuestos en el marco de evaluación, en las figuras 12 a 22 veremos:

- El aspecto de las fichas de cinco de los datasets tomados de muestra, tanto en el portal de datos abiertos como en el de análisis: Límites provinciales de Castilla y León: recintos; Situación epidemiológica coronavirus (COVID-19) en Castilla y León<sup>11</sup>; Comercios con entrega a domicilio; Situación de solicitudes de ERTes coronavirus; y Eventos de la agenda cultural categorizados y geolocalizados.
- El aspecto que presenta la ficha del catálogo de datos en el portal de datos abiertos (el catálogo en la herramienta de análisis no tiene ficha como tal).

---

<sup>8</sup> <https://gobiernoabierto.jcyl.es/web/es/participacion/todos-foros-participacion-creados.html>

<sup>9</sup> <https://gobiernoabierto.jcyl.es/web/es/gobierno-abierto-castilla-leon.html>

<sup>10</sup> <https://twitter.com/transparencia>

<sup>11</sup> Este dataset, presente en el portal de análisis, está desglosado en dos en el portal de datos abiertos: "Situación epidemiológica coronavirus (COVID-19) en Castilla y León por hospitales" y "Situación epidemiológica coronavirus (COVID-19) en Castilla y León por provincias"

**LÍMITES PROVINCIALES DE CASTILLA Y LEÓN: RECINTOS**


**PUBLICACIÓN EN JCYL.ES DEL CONJUNTO DE DATOS:**  
 ■ **Inicio de la publicación:** 29 de enero de 2016  
 ■ **Frecuencia de actualización:** Anual

**FORMATOS DISPONIBLES PARA DESCARGA:**  
 Descargar

**RESPONSABLE DEL CONTENIDO:**  
 Dirección General de Vivienda, Arquitectura y Urbanismo -  
 Consejería de Fomento y Medio Ambiente

**ÁMBITO GEOGRÁFICO DEL CONJUNTO DE DATOS:**  
 Castilla y León (España)

**ETIQUETAS:**  
 Unidades administrativas

**LICENCIA:**  
<https://www.jcyl.es/licencia-IGCYL-NC>

**DESCRIPCIÓN:**

Representación cartográfica de las unidades administrativas provinciales oficiales de Castilla y León. Delimitación administrativa de las provincias de Castilla y León generada a partir de las líneas-límite municipales oficiales del Instituto Geográfico Nacional (IGN). Construye una de las cartografías básicas de referencia de la Comunidad Autónoma de Castilla y León.

Desde el geoportál IDECYL se puede realizar la descarga de toda la información relacionada con **Límites provinciales de Castilla y León: recintos**

**INFORMACIÓN ADICIONAL:**

Puedes encontrar más información relacionada con el contenido del dataset en: [Tramos con Planes de Pesca](#)

ARRIBA

Figura 12. Ficha del dataset "Límites provinciales" en el portal de datos

**9 registros**  
No hay filtros activos.

**Filtros**  
 Buscar registros...

**Cod\_ine\_prov**

5	1
9	1
24	1
34	1
37	1
40	1
> Más	

**Límites provinciales de Castilla y León: recintos**

Información | Tabla | Mapa | Analizar | Exportar | API

Representación cartográfica de las unidades administrativas provinciales oficiales de Castilla y León.

Delimitación administrativa de las provincias de Castilla y León generada a partir de las líneas-límite municipales oficiales del Instituto Geográfico Nacional (IGN).

Constituye una de las cartografías básicas de referencia de la Comunidad Autónoma de Castilla y León.

**Identificador de conjunto de datos:** [límites-provinciales-de-castilla-y-leon-recintos](#)

**Descargas:** 396

**Temas:** Medio ambiente  
**Palabras claves:** Unidades administrativas, límites, recintos, provincias  
**Licencia:** [CC BY 3.0 ES](#)  
**Idioma:** Español  
**Zona horaria:** Europe/Madrid  
**Modificado:** 9 de octubre de 2019 17:35

**Área geográfica:**


**Publicador:** Junta de Castilla y León

**Último procesamiento:** 9 de octubre de 2019 17:35 (metadatos)  
 9 de octubre de 2019 17:35 (datos)

**Esquema de conjunto de datos**  
 Haga clic para expandir

Figura 13. Ficha del dataset "Límites provinciales" en el portal de análisis

[DATOS ABIERTOS](#)
[INICIATIVA DE DATOS ABIERTOS](#)
[CATÁLOGO DE DATOS](#)
[ANÁLISIS DE DATOS ABIERTOS](#)
[APLICACIONES](#)
[CONCURSO DE DATOS ABIERTOS](#)
[PARTICIPA](#)

## SITUACIÓN EPIDEMIOLÓGICA CORONAVIRUS (COVID-19) EN CASTILLA Y LEÓN POR HOSPITALES

CONTACTAR  
 ENVIAR  
 IMPRIMIR  
 ESCUCHAR

**PUBLICACIÓN EN JCYL.ES DEL CONJUNTO DE DATOS:**

- **Inicio de la publicación:** 25 de marzo de 2020
- **Frecuencia de actualización:** 24 horas de lunes a viernes

**FORMATOS DISPONIBLES PARA DESCARGA:**

[Descargar](#) (CSV)
 [Descargar](#) (XLS)

**RESPONSABLE DEL CONTENIDO:**

Dirección General de Sistemas de Información, Calidad y Prestación Farmacéutica - Consejería de Sanidad

**ÁMBITO GEOGRÁFICO DEL CONJUNTO DE DATOS:**

Castilla y León (España)

**ETIQUETAS:**

coronavirus, COVID-19, hospitales

**LICENCIA:**

Creative Commons Reconocimiento 3.0

**DESCRIPCIÓN:**

Evolución del número de hospitalizados en planta, hospitalizados en unidades de críticos, altas y fallecimientos por hospital debidos al coronavirus (COVID-19) en Castilla y León. Actualización de lunes a viernes.

- Datos desde el 18 de marzo de 2020.
- La actualización de hospitalizados se realiza diariamente a las 8:00 horas.
- NOTA: En los hospitales Santiago Apóstol, Santos Reyes y de Medina del Campo el valor de ingresados en UCI es 0 porque no disponen esta unidad.
- El número de hospitalizados en unidades de críticos incluye a los pacientes hospitalizados en UCI (unidades de cuidados intensivos), REA (unidades de reanimación), URPA (unidades de reanimación postanestésica), y otras unidades con adecuada dotación para la atención de pacientes críticos.

ARRIBA

Figura 14. Ficha del dataset "Situación epidemiológica" en el portal de datos

**1.188 registros**  
No hay filtros activos.

**Filtros**

Buscar registros...

[Información](#)
[Tabla](#)
[Mapa](#)
[Analizar](#)
[Visualización](#)
[Exportar](#)

[API](#)

**fecha**

2020 1.188

**provincia**

Burgos	132
León	132
Palencia	132
Salamanca	132
Segovia	132
Soria	132
> Más	

**Situación epidemiológica coronavirus (COVID-19) en Castilla y León**

Datos diarios de confirmados, nuevos casos, altas y fallecimientos en hospitales debidos al coronavirus (COVID-19) por provincias en Castilla y León. Actualización diaria. Actualización de lunes a viernes.

- Datos desde el 15 de marzo de 2020.
- La cifra de nuevos casos confirmados se refieren al día de la actualización de datos.
- Los datos de confirmados se facilitan por la Red de Vigilancia Epidemiológica a las 20:00 horas del día anterior y corresponden a personas a las que se les ha efectuado la prueba diagnóstica.
- Desde el 18 de mayo, los nuevos casos confirmados solo reflejan resultados de PCR.
- NOTA: Valores negativos en "Nuevos casos positivos" corresponden a pacientes cuyo resultado ha sido positivo inicialmente y en la prueba de confirmación se ha descartado.

**Identificador de conjunto de datos:** `situacion-epidemiologica-coronavirus-en-castilla-y-leon`

**Descargas:** 157.276

**Temas:** Salud

**Palabras claves:** coronavirus, COVID-19

**Licencia:** CC BY 3.0 ES

**Idioma:** Español

**Zona horaria:** Europe/Madrid

**Modificado:** 24 de julio de 2020 13:51

**Publicador:** Junta de Castilla y León

**Último procesamiento:** 24 de julio de 2020 13:51 (metadatos)  
24 de julio de 2020 13:51 (datos)

**Esquema de conjunto de datos**  
Haga clic para expandir.

Figura 15. Ficha del dataset "Situación epidemiológica" en el portal de análisis

## SITUACIÓN DE SOLICITUDES DE ERTES CORONAVIRUS

**PUBLICACIÓN EN JCYL.ES DEL CONJUNTO DE DATOS:**

- **Inicio de la publicación:** 27 de marzo de 2020
- **Frecuencia de actualización:** Diaria

**FORMATOS DISPONIBLES PARA DESCARGA:**

Descargar

Descargar

**RESPONSABLE DEL CONTENIDO:**

Dirección General de Trabajo y Prevención de Riesgos Laborales  
- Consejería de Empleo e Industria

**ÁMBITO GEOGRÁFICO DEL CONJUNTO DE DATOS:**

Castilla y León (España)

**ETIQUETAS:**

coronavirus, COVID-19, ertes

**LICENCIA:**

Creative Commons Reconocimiento 3.0

**DESCRIPCIÓN:**

---

Situación de los ERTES por fuerza mayor derivada de las crisis del COVID-19 en Castilla y León, incluyendo datos como la fecha de la solicitud, la provincia y el número de trabajadores afectados.

Datos desde el 19 de marzo de 2020.

Los datos que hasta el 19 de junio aparecían atribuidos al Ministerio de Trabajo y Economía Social, aparecen actualmente asignados a la autoridad laboral competente (provincia) que corresponda. Los datos a fecha 26 de junio han sido revisados y depurados, por lo que pueden apreciarse algunas diferencias con respecto a los datos históricos existentes.

Puedes explorar y analizar este dataset en el [Portal de análisis de datos abiertos de la Junta de Castilla y León](#)

**INFORMACIÓN ADICIONAL:**

---

Puedes encontrar más información relacionada con el contenido del dataset en: [Información sobre el coronavirus](#)

ARRIBA

Figura 16. Ficha del dataset "Situación solicitudes ERTes" en el portal de datos

INICIO
DATOS
API
VISUALIZACIONES
CONTACTO
DOCUMENTACIÓN

**730 registros**

No hay filtros activos.

**Filtros**

**fecha**

2020 730

**provincia**

- Burgos 67
- Dirección General de Trabajo y Prevención de Riesgos Laborales 67
- León 67
- Palencia 67
- Salamanca 67
- Segovia 67
- > Más

**sigla**

- AV 67
- BU 67
- DG TyP 67
- LE 67
- PA 67
- SA 67
- > Más

### Situación de solicitudes de ERTes coronavirus

Información
Tabla
Analizar
Visualización
Exportar
API

Situación de los ERTES por fuerza mayor derivada de las crisis del COVID-19 en Castilla y León, incluyendo datos como la fecha de la solicitud, la provincia y el número de trabajadores afectados.

Datos desde el 19 de marzo de 2020.

Los datos que hasta el 19 de junio aparecían atribuidos al Ministerio de Trabajo y Economía Social, aparecen actualmente asignados a la autoridad laboral competente (provincia) que corresponda. Los datos a fecha 26 de junio han sido revisados y depurados, por lo que pueden apreciarse algunas diferencias con respecto a los datos históricos existentes.

**Identificador de conjunto de datos** situacion-de-solicitudes-de-ertes-coronavirus

**Descargas** 4 299

**Temas** Empleo

**Palabras claves** COVID-19, coronavirus

**Licencia** [CC BY 3.0 ES](#)

**Idioma** Español

**Modificado** 24 de julio de 2020 15:46

**Publicador** Junta de Castilla y León

**Último procesamiento** 24 de julio de 2020 15:46 (metadatos)  
24 de julio de 2020 15:46 (datos)

**Esquema de conjunto de datos**  
Haga clic para expandir.

Figura 17. Ficha del dataset "Situación solicitudes ERTes" en el portal de análisis

**DATOS ABIERTOS**  
de Castilla y León

## COMERCIOS CON ENTREGA A DOMICILIO

**PUBLICACIÓN EN JCYL.ES DEL CONJUNTO DE DATOS:**

- **Inicio de la publicación:** 17 de abril de 2020
- **Frecuencia de actualización:** Diaria

**FORMATOS DISPONIBLES PARA DESCARGA:**

 Descargar
  Descargar
  Descargar

**RESPONSABLE DEL CONTENIDO:**

Dirección General de Comercio y Consumo - Consejería de Empleo e Industria

**ÁMBITO GEOGRÁFICO DEL CONJUNTO DE DATOS:**

Castilla y León (España)

**ETIQUETAS:**

domicilio, entrega, directorio

**LICENCIA:**

Creative Commons Reconocimiento 3.0

**DESCRIPCIÓN:**

Este directorio se encuentra en permanente actualización.

**IMPORTANTE:** Consultar con el establecimiento las condiciones del servicio.

Puedes explorar y analizar este dataset en el [Portal de análisis de datos abiertos de la Junta de Castilla y León](#)

ARRIBA

Figura 18. Ficha del dataset "Comercios con entrega a domicilio" en el portal de datos

**1.675 registros**

No hay filtros activos.

**Filtros**

Buscar registros...

**SECTOR COMERCIAL**

Carne y productos cárnicos	302
Alimentación en general	139
Bebidas	110
Supermercado	102
Frutas y hortalizas	95
Flores, plantas, semillas, fertilizantes, mascotas y animales de compañía	89
> Más	

**PROVINCIA**

VALLADOLID	434
LEÓN	245
SALAMANCA	212
PALENCIA	183
ZAMORA	176
BURGOS	174
> Más	

**LOCALIDAD**

VALLADOLID	303
PALENCIA	152
SALAMANCA	143
LEÓN	113
ZAMORA	111
BURGOS	106
> Más	

## Comercios con entrega a domicilio

Información | Tabla | Analizar | Mapa | Exportar | API

Este directorio se encuentra en permanente actualización.

**IMPORTANTE:** Consultar con el establecimiento las condiciones del servicio.

**Identificador de conjunto de datos:** `comercios-con-entrega-a-domicilio`

**Descargas:** 3.118

**Temas:** Comercio

**Palabras claves:** domicilio, entrega, directorio

**Licencia:** CC BY 3.0 ES

**Idioma:** Español

**Modificado:** 15 de julio de 2020 13:22

**Publicador:** Junta de Castilla y León

**Último procesamiento:** 15 de julio de 2020 13:22 (metadatos)  
15 de julio de 2020 13:22 (datos)

**Esquema de conjunto de datos**  
Haga clic para expandir.

Figura 19. Ficha del dataset "Comercios con entrega a domicilio" en el portal de análisis

?

CONTACTAR

✉

ENVIAR

🖨

IMPRIMIR

🔊

ESCUCHAR

## RELACIÓN DE EVENTOS DE LA AGENDA CULTURAL CATEGORIZADOS Y GEOLOCALIZADOS

**PUBLICACIÓN EN JCYL.ES DEL CONJUNTO DE DATOS:**

- **Inicio de la publicación:** 1 de junio de 2018
- **Frecuencia de actualización:** Cada 4 horas

**FORMATOS DISPONIBLES PARA DESCARGA:**

Descargar

Descargar

**RESPONSABLE DEL CONTENIDO:**

Dirección General de Políticas Culturales - Consejería de Cultura y Turismo

**ÁMBITO GEOGRÁFICO DEL CONJUNTO DE DATOS:**

Castilla y León (España)

**ETIQUETAS:**

eventos, agenda, cultural, hoy, actual, coordenadas, categorías

**LICENCIA:**

Creative Commons Reconocimiento 3.0

**DESCRIPCIÓN:**

---

Relación de eventos de la agenda cultural que se celebran en Castilla y León desde el día en el que se realiza la consulta en adelante, incluyendo la categoría a la que pertenecen y las coordenadas geográficas del lugar de celebración.

El fichero incluye un registro por evento y fecha de inicio, es decir, aquellos eventos con varias fechas de inicio se reflejarán en el dataset con tantas entradas como fechas de inicio contengan.

Puedes explotar y analizar este dataset en el [Portal de análisis de datos abiertos de la Junta de Castilla y León](#)

**INFORMACIÓN ADICIONAL:**

---

Este conjunto de datos se emplea por reutilizadores en aplicaciones, estudios o utilidades como [Cyl Cultural](#)

Puedes encontrar más información relacionada con el contenido del dataset en: [Agenda cultural](#)

ARRIBA

Figura 20. Ficha del dataset "Eventos" en el portal de datos

**252 registros**

No hay filtros activos.

**Filtros**

Buscar registros...

**Categoría**

- Conciertos 54
- Exposiciones 53
- Espectáculos 28
- Otros 23
- Cine 9
- Conferencias y Cursos 2
- > Más

**Precio**

- Gratuito 27
- 3 euros 23
- 2 euros 15
- Consultar en Ayuntamiento 15
- Gratuito, previa inscripción 11
- Precio de entradas: 10 euros, excepto los días 19 y 25 de julio, que tendrán un precio de 3 euros (recaudación destinada a Cruz Roja a fin de paliar efectos de la Covid-19) // Descuentos para menores de 25 y mayores de 65 años: 8 euros y desempleados: 3 euros
- > Más

## Eventos de la agenda cultural categorizados y geolocalizados

📄 Información
📊 Tabla
📍 Mapa
📈 Analizar
📄 Exportar
⚙️ API

Relación de eventos de la agenda cultural que se celebran en Castilla y León desde el día en el que se realiza la consulta en adelante. Cada registro incluye su categoría, la dirección del lugar de celebración y las coordenadas de su posición geográfica

**Identificador de conjunto de datos** eventos-de-la-agenda-cultural-categorizados-y-geolocalizados

**Descargas** 42.109

**Temas** Cultura y ocio

**Palabras claves** eventos, agenda, cultural, hoy, actual

**Licencia** [CC BY 3.0 ES](#)

**Idioma** Español

**Zona horaria** Europe/Madrid

**Modificado** 25 de julio de 2020 0:00

**Publicador** Junta de Castilla y León

**Último procesamiento** 26 de julio de 2020 8:00 (metadatos)  
25 de julio de 2020 0:00 (datos)

**Esquema de conjunto de datos**  
Haga clic para expandir:

Figura 21. Ficha del dataset "Eventos" en el portal de análisis

**CATÁLOGO DE CONJUNTOS DE DATOS**

**PUBLICACIÓN EN JCYLE.S DEL CONJUNTO DE DATOS:**

- Inicio de la publicación: 27 de marzo de 2012
- Frecuencia de actualización: Diaria

**FORMATOS DISPONIBLES PARA DESCARGA:**

 Descargar

 Descargar

 Descargar

 Descargar

 Descargar

**RESPONSABLE DEL CONTENIDO:**

Dirección General de Transparencia y Buen Gobierno -  
Consejería de Transparencia, Ordenación del Territorio y Acción Exterior

**ÁMBITO GEOGRÁFICO DEL CONJUNTO DE DATOS:**

Castilla y León (España)

**ETIQUETAS:**

catalogo, conjunto, datos, dcat

**LICENCIA:**

Creative Commons Reconocimiento 4.0

---

**DESCRIPCIÓN:**

Relación de conjuntos de datos que la Junta de Castilla y León ha publicado en formatos abiertos en el portal Open Data.

Puedes explorar y analizar este dataset en el [Portal de análisis de datos abiertos de la Junta de Castilla y León](#) 

---

**INFORMACIÓN ADICIONAL:**

Puedes encontrar más información relacionada con el contenido del dataset en: [Catálogo de datos](#)

**Descargas**


Esquema XSDL (XML Schema Definition Language) del fichero XML

3 kbytes


Descargar

Figura 22. Ficha del dataset "Catálogo de datos" en el portal de datos

- **Los datos son proporcionados junto con el nombre de su publicador para verificar la autenticidad de la fuente:** Sí, como puede observarse en las Figuras 12 a 22, el publicador "Junta de Castilla y León" aparece en todas las fichas obtenidas en el portal de análisis; en las fichas del portal de datos consta el responsable del contenido, que es el organismo que lo ha creado.
- **Los datos publicados responden a su enunciado pertinente, son correctos de acuerdo a la materia y el tipo de información del dataset:** Sí. Analizando los 5 datasets de muestra no se han encontrado discrepancias entre el enunciado, el contenido y el tipo de información.
- **Los datos son pertinentes para la necesidad del usuario:** la respuesta a este ítem está relacionada con el ítem anterior. No se puede saber de antemano qué es lo que necesita exactamente cada usuario, pero el hecho de que los datos publicados correspondan con el enunciado, materia y tipo de información del dataset implica que van a ser de utilidad para los usuarios que entran en el portal buscando dicha información.
- **Se proveen los datos suficientes para tener claro el contexto de los datos (título, cobertura geográfica, descripción, propósito de su recolección, fechas de creación y de actualización, frecuencia de actualización):** Como puede apreciarse en las figuras, todos los datasets tienen un título asignado, indican su cobertura geográfica y cuentan con una descripción aclaratoria sobre su contenido (en el portal de datos se la denomina "Descripción" y en el portal de análisis está

incluida en la pestaña "Información"); respecto al propósito de su recolección, es algo que recomiendan incluir Máchová y Lnenicka (2017), pero aquí no está expresado.

En lo que respecta a la fecha de creación, la frecuencia de actualización de datos y la fecha de la última actualización, son datos que aparecen separados en las diferentes fichas: los dos primeros se visualizan en las fichas procedentes del portal de datos, y la última actualización aparece recogida en las fichas del portal de análisis.

Lo mismo puede observarse en los catálogos descargados en CSV: en el catálogo general hay 40 registros que no tienen fecha de creación, pero la norma general es que sí aparezca, y por el contrario, solo aparece la fecha de actualización en 30 registros. De manera inversa, en el catálogo descargable del portal de análisis no viene la fecha de creación, pero sí la de modificación.

En la Tabla V se recoge la presencia de estos datos en las fichas del catálogo de datos y de los 5 datasets de muestra<sup>12</sup>:

Título dataset	Fecha inicio publicación	Frecuencia actualización	Fecha actualización
Catálogo de datos	27/03/2012	Diaria	No consta
Límites provinciales de Castilla y León: recintos	29/01/2016	Anual	09/10/2019 17:35 (metadatos y datos)
Situación epidemiológica coronavirus (COVID-19) en Castilla y León	25/03/2020	24 horas (lunes a viernes)	31/07/2020 15:16 (metadatos y datos)
Comercios con entrega a domicilio	17/04/2020	Diaria	27/07/2020 14:55 (metadatos y datos)
Situación de solicitudes de ERTes coronavirus	27/03/2020	Diaria	31/07/2020 15:09 (metadatos y datos)
Eventos de la agenda cultural categorizados y geolocalizados	01/06/2018	Cada 4 horas	31/07/2020 20:00 (metadatos y datos)

Tabla V. Fechas de inicio y actualización, y frecuencia de actualización

➤ **La actualización de los datos se realiza: de acuerdo a lo indicado en el portal, de acuerdo a lo establecido por la ley del país, conforme a como los datos cambian:** Observando los conjuntos de datos de muestra, se ve que las frecuencias de actualización establecidas en el portal van muy acompañadas a los cambios de los datos en la vida real y a lo señalado en la NTI, que indica que los documentos y recursos de información de elaboración o recogida periódica puestos a disposición para su reutilización deben estar actualizados a sus últimas versiones; en la misma línea, una de las buenas prácticas del portal es su constante actualización y del número de conjuntos de datos que publica.

En el catálogo de datos descargable, la frecuencia de actualización aparece especificada en diversos tramos temporales: puntual, según necesidad, continua, cada 4 horas, 24 horas de lunes a viernes, diaria, semanal, 2 semanas, mensual, trimestral, cuatrimestral, semestral, decenal, anual, bienal, cada 4 años, cada 10

<sup>12</sup> Datos obtenidos el día 31/07/2020

años, no programado, próxima actuación en 2017. Este último dato está claramente desactualizado, y las actualizaciones no programadas restan credibilidad a los datos.

➤ **Condiciones de uso y tipo de licencia de cada conjunto de datos:** Sí. Es un aspecto importante porque si no tienen específicamente una licencia abierta no se consideran datos abiertos. Tanto en el portal de datos como en la herramienta de análisis podemos ver el tipo de licencia, indicado a través de un hipervínculo que conduce a sus condiciones de uso y especificaciones.

Título dataset	Portal de datos	Portal de análisis
Catálogo de datos	CC BY 3.0 ES	No se indica
Límites provinciales de Castilla y León: recintos	<a href="https://www.jcyl.es/licencia-IGCYL-NC">https://www.jcyl.es/licencia-IGCYL-NC</a>	CC BY 3.0 ES
Situación epidemiológica coronavirus	CC BY 3.0 ES	CC BY 3.0 ES
Comercios con entrega a domicilio	CC BY 3.0 ES	CC BY 3.0 ES
Situación de solicitudes de ERTes coronavirus	CC BY 3.0 ES	CC BY 3.0 ES
Eventos de la agenda cultural categorizados y geolocalizados	CC BY 3.0 ES	CC BY 3.0 ES

Tabla VI. Licencias de uso de los datasets

Como puede observarse en la Tabla VI, los datos se ofrecen bajo licencia Creative Commons Reconocimiento 3.0 (CC BY 3.0 ES)<sup>13</sup>. En el portal de análisis no se indica la licencia aplicable al catálogo del conjunto de datos, y hay una discrepancia en la licencia del dataset "Límites provinciales", ya que el portal de datos le aplica una licencia relativa a la cartografía e información geográfica producida por la administración de la comunidad de Castilla y León<sup>14</sup>, y el portal de análisis dice que se ofrece bajo licencia CC BY 3.0 ES.

➤ **El conjunto proporciona definiciones y/o explicaciones que ayuden al usuario a interpretar los datos:** el usuario puede tener dificultades para entender los datos si se encuentra con siglas o acrónimos, palabras poco comunes o técnicas, no siendo suficiente una mera descripción del contenido.

En los 5 datasets de muestra los campos en los que se estructura la información son bastante claros y en sus fichas del portal de datos abiertos se proporciona un enlace a información relacionada (salvo en "Comercios con entrega a domicilio", que no la necesita), por lo que no hacen falta mayores explicaciones. Pero hay otros datasets que aportan mayor información; por ejemplo, "Aparcamientos en espacios naturales", además de aportar una descripción e información adicional, permite la descarga de un diccionario de entidades, tablas y campos.

➤ **Uso de metadatos mínimos y cumplimiento de estándares:** La NTI describe una serie de metadatos considerados obligatorios y que deben ir asociados al catálogo y a cada dataset, además del término recomendado para su representación usando vocabularios estándar que se identifican por las abreviaturas de su espacio de nombres. La norma además contempla la posibilidad de asignar metadatos adicionales para mejorar la calidad de la información ofrecida.

<sup>13</sup> <https://creativecommons.org/licenses/by/3.0/es/>

<sup>14</sup> <https://www.jcyl.es/licencia-IGCYL-NC>

Para asegurar la interoperabilidad semántica, el portal de datos abiertos sigue las recomendaciones de la NTI y utiliza el vocabulario DCAT<sup>15</sup>, estándar de metadatos desarrollado por el W3C y utilizado a nivel nacional y europeo para la publicación de conjuntos de datos en el sector público; DCAT también hace un amplio uso de términos de otros vocabularios, en particular de Dublin Core (DC Terms)<sup>16</sup>.

En el portal observamos que el espacio de nombres utilizado es schema.dcterms, con términos del vocabulario Dublin Core.

Se han analizado los metadatos utilizados en el catálogo de datos y en un dataset concreto:

a) Metadatos del catálogo de datasets: en la Tabla VII vemos los metadatos requeridos (R) y opcionales (O) de acuerdo a la NTI:

Metadatos	Propiedad	R	Aplicación
<b>Nombre</b>	dcterms.title	X	"Catálogo de conjuntos de datos"
<b>Descripción</b>	dcterms.description	X	"Relación de conjuntos de datos que la Junta de Castilla y León ha publicado en formatos abiertos en el portal Open Data".
<b>Órgano publicador</b>	dcterms.publisher	X	"Junta de Castilla y León"
<b>Tamaño catálogo</b>	NO CONSTA	O	
<b>Identificador</b>	dcterms.identifier	O	"https://datosabiertos.jcyl.es/web/jcyl/RISP/es/Plantilla100Detalle/1284162055979/_/1284166186527/OpenData"
<b>Fecha creación</b>	NO CONSTA	X	
<b>Fecha actualización</b>	NO CONSTA	X	
<b>Idioma(s)</b>	dcterms.language	X	"es"
<b>Cobertura geográfica</b>	dcterms.coverage	O	"Castilla y León"
<b>Temática(s)</b>	dcterms.subject "tematica"	X	"1132826309205 1132826313799 1132826313808"
<b>Página web</b>	og:url	X	https://datosabiertos.jcyl.es/web/jcyl/RISP/es/Plantilla100Detalle/1284162055979/_/1284166186527/OpenData
<b>Términos de uso</b>	dcterms.rights	X	https://www.jcyl.es/web/jcyl/Portada/es/Plantilla100Detalle/1246890364336/_/1246989714819/Texto
<b>Documentación</b>	NO CONSTA	X	

Tabla VII. Uso de metadatos en el catálogo de datos

<sup>15</sup> (dcat: <http://www.w3.org/ns/dcat#>)

<sup>16</sup> (dct: <http://purl.org/dc/terms/>)

No aparecen la fecha de creación ni la de actualización, ni el metadato de Documentación y recursos de información, a pesar de ser metadatos requeridos; tampoco consta el tamaño del catálogo, pero este no es requerido. Aparecen dos fechas (fechaOrdenación "2020-05-11" y fechaActivación "2011-10-27") que no sabemos a qué hacen referencia.

La NTI recomienda que el **Nombre** y la **Descripción** se incluyan en varios idiomas, pero solo aparecen en español.

El **Órgano publicador** tendría que estar especificado con un URI que identifique al órgano público diferenciado con un código alfanumérico, extraído del Directorio Común gestionado por el MINHAP. En la **cobertura geográfica** debería aplicarse preferentemente lo establecido en el Anexo V (aunque no es un campo requerido).

**Temática (s)** es un metadato que indica la totalidad de materias incluidas en el catálogo y es de aplicación preferente la taxonomía definida en el anexo IV de la NT; en el portal se ha localizado la propiedad "dcterms.subject", pero no tiene contenido; también existe <meta name="tematica" content="1132826309205 1132826313799 1132826313808">, que parece hacer referencia a la temática mediante una forma normalizada compuesta de códigos.

La **Página web** está introducida con una propiedad de Open Graph, protocolo desarrollado por Facebook que permite mostrar la página web con un formato enriquecido en redes sociales.

El metadato **Términos de uso** no se ha localizado; hace referencia a los términos de uso generales del catálogo, y debería estar en forma de URI. Está presente la propiedad "dcterms.rights", que conduce a la página de aviso legal de la web genérica de la Junta de Castilla y León, pero no son los términos de uso.

**Documento(s) y recursos(s) de información** tampoco aparece; es un metadato que facilita el acceso a una lista de cada uno de los documentos y recursos de información del catálogo, y debe tener tantas propiedades como entradas en el catálogo.

b) Metadatos del dataset "Comercios con entrega a domicilio". En la Tabla VIII se observa que, de los metadatos requeridos, solo falta **Distribución(es)**, que es la referencia a los recursos que identifican los volcados del documento o recurso de información en sus posibles formatos.

Igual que en los metadatos del catálogo, aquí tampoco se utilizan varios idiomas para el **Nombre** y la **Descripción**, y con la **Temática** ocurre lo mismo que el catálogo de datos: aparecen expresadas las propiedades "Dcterms.subject" sin contenido y "tematica" con tres códigos numéricos.

El metadato **Etiquetas** permite categorizar libremente el documento o recurso de información; la propiedad utilizada aquí no es Dublin Core y, según la norma, su aplicación debería ser una cadena alfanumérica

Metadatos	Propiedad	R	Aplicación
Nombre	dcterms.title	X	"Comercios con entrega a domicilio"
Descripción	dcterms.description	X	"Este directorio se encuentra en permanente actualización. IMPORTANTE: Consultar con el establecimiento las condiciones del servicio"
Temática(s)	Dcterms.subject "tematica"	X	"1132826302172 1132826301674 1132826301683 1132826302151"
Etiqueta(s)	"etiquetas"	O	"domicilio, entrega, directorio"
Identificador	dcterms.identifier	O	<a href="https://datosabiertos.jcyl.es/web/jcyl/RISP/es/Plantilla100Detalle/1284162055979/_/1284947248341/OpenData">https://datosabiertos.jcyl.es/web/jcyl/RISP/es/Plantilla100Detalle/1284162055979/_/1284947248341/OpenData</a> >
Frecuencia actualización	"actualización"	O	Diaria
Idioma(s)	dcterms.language	O	"es"
Organismo	dct:publisher	X	"Junta de Castilla y León"
Cobertura geográfica	dcterms.coverage	O	"Castilla y León"
Distribución	"formatosDisponibles"	X	"CSV XLS"

Tabla VIII. Metadatos del dataset

➤ **Mecanismos de descarga en diferentes formatos:** Sí, en la Tabla IX podemos ver los formatos de descarga del catálogo de conjuntos de datos y de los 5 datasets seleccionados para su análisis:

Título	Portal de datos	Portal de análisis
Catálogo de datos	CSV, RDF, XML, JSON, RSS	XLS, CSV, RDF (Turtle), RDF (XML), Fuente RSS.
Límites provinciales de Castilla y León: recintos	SHP	Formatos de archivo plano: CSV, JSON, XLS. Formatos de archivo geográfico: GeoJSON, Shapefile, KML.
Situación epidemiológica coronavirus (COVID-19) en Castilla y León	CSV y XLS	Formatos de archivo plano: CSV, JSON, XLS. Formato de archivo geográfico: GeoJSON, Shapefile, KML.
Comercios con entrega a domicilio	CSV, XLS	Formatos de archivo plano: CSV, JSON, XLS. Formatos de archivo geográfico: GeoJSON, Shapefile, KML.
Situación de solicitudes de ERTes coronavirus	CSV y XLS	Formatos de archivo plano: CSV, JSON, XLS.
Eventos de la agenda cultural categorizados y geolocalizados	CSV, JSON	Formatos de archivo plano: CSV, JSON, XLS. Formatos de archivo geográfico: GeoJSON, Shapefile, KML.

Tabla IX. Formatos de descarga de datasets

➤ **Uso de formatos de archivo no propietarios y legibles por máquinas:** Sí. En la página principal del portal de datos abiertos, dentro del menú “Catálogo de datos”, se enumeran los formatos de datos abiertos utilizados (CSV, XML, RDF, RSS, JSON, ICAL, WMS, WCS, WFS, KML, SOAP) y otros formatos que no lo son (RASTER, XLS, GML, DGN, DWG, DXF, GDB, SHP, ECW, JPG, TIFF o IMG).

Si revisamos los formatos utilizados en los datasets analizados en el ítem anterior, todos son abiertos salvo SHP, XLS y Shapefile.

➤ **Uso de formatos comprensibles para los usuarios:** En general sí; formatos como CSV, XLS, JPG, TIFF o IMG son de uso generalizado para cualquier usuario con nociones estándar de informática; otros formatos utilizados son específicos de campos concretos (por ejemplo, DGN es el formato nativo de Microstation, un software para ingeniería y arquitectura), pero es necesario utilizarlos para trabajar con ciertos datos.

➤ **Tamaño del conjunto de datos:** este ítem engloba tanto el volumen de datos del archivo descargable como la indicación de dicho tamaño. Venegas Álvarez (2018) propone una clasificación con 4 tamaños<sup>17</sup>: un archivo de hasta 250 MB es fácil de manejar; hasta un 1 GB es manejable; entre 1 y 4 GB es difícil de manejar; y más de 4 GB es intratable.

El tamaño de los archivos no se indica ni en el portal ni en la herramienta de análisis, por lo que deben descargarse para su comprobación. Se ha comprobado el tamaño del catálogo de datos y de los 5 datasets seleccionados como muestra, en todos los formatos ofrecidos (ver Tabla IX), y el más voluminoso es “Límites provinciales de Castilla y León”, que ocupa 1,54 MB en formato JSON y en XLS. Por lo tanto, son archivos fáciles de manejar.

Sin embargo, se ha buscado el dataset con mayor número de registros en el portal de análisis (“Consumos horarios en hospitales”) y en XLS tiene un peso de 2,50 GB, por lo que se consideraría un archivo difícil de manejar. Ha tardado varios minutos en descargarse, y se ha podido abrir tras varios intentos. En CSV ocupa 307 MB.

➤ **El conjunto se organiza internamente en registros estructurados:** Sí; los archivos no están preparados con filas combinadas, se organizan en columnas y registros, lo que facilita la procesabilidad de los datos de forma automática.

➤ **Nivel de granularidad:** Los datos se encuentran en un grado fino de granularidad, sin cálculos ni agregaciones; no se indica el nivel de granularidad de cada dataset.

➤ **Existencia de celdas vacías y valores nulos o sin significado:** Como veremos a continuación, se han detectado algunos valores nulos y celdas vacías en el análisis del CSV del catálogo de datos y de los 5 datasets de muestra<sup>18</sup>; se indica su porcentaje respecto al número total de celdas de la columna correspondiente.

- a) Catálogo de datos: 10 de las 13 columnas en las que se estructura el archivo CSV tienen todas las celdas rellenas y con valores significantes (la columna “EnlaceMasInformacion” tiene 27 celdas vacías, pero la contamos como completa por ser una información adicional y opcional).

---

<sup>17</sup> Venegas puntualiza que es una sugerencia, al no haber encontrado literatura al respecto. Además, la descarga no depende solo del tamaño del archivo, sino también de los medios del usuario (velocidad de conexión a la red, equipo informático, etc.).

<sup>18</sup> Descarga del 15 de julio y del 26 de julio de 2020, respectivamente.

En la columna "Frecuencia de actualización" todas las celdas están completas y su contenido es significativo, salvo en el dataset "Cartografía, Mapas topográficos, 1:10.000, Ortofotografía, PNOA, Fotogrametría" que indica que la próxima actualización es en 2017, por lo que ese dato está desactualizado. La frecuencia de actualización "No programado" aparece en 95 celdas; "No programada" en 1 celda, y "Según necesidad" en 40 celdas.

Las columnas "FechaCreacion" y "FechaUltimaActualizacion" tienen 40 y 495 celdas vacías respectivamente (el 7'62% y el 94'28%).

- b) Límites provinciales: 9 registros. Todas las celdas completas y significantes.
- c) Situación epidemiológica: 1188 registros. Aparecen celdas vacías en 2 de las 6 columnas: "Altas" (19 celdas, el 1'60%) y "Fallecimientos" (17 celdas, el 1'43%).
- d) Comercios: 1676 registros. 8 de sus 11 columnas tienen todas las celdas rellenas con valores significantes, salvo por un registro entero erróneo: como "nombre comercial" aparece un número de teléfono (979 17 05 49"), en "sector comercial" se indica su web (<http://www.audiovida.es/>), la "dirección" es un correo electrónico ([info@audiovida.es](mailto:info@audiovida.es)), "provincia" es la latitud (42.006465), y "localidad" es la longitud (-4.529612).

"Cod\_Postal" tiene 2 celdas vacías (0'12%) y "Teléfono" tiene 7 (0'42%).

"Página web" y "Mail" tienen celdas vacías, pero es lógico, porque no todos los negocios tienen web; una de las celdas contiene un error, ya que presenta juntas la dirección web y la de correo electrónico: [www.hamburgueseriawinnie.es](http://www.hamburgueseriawinnie.es)[Hamburgueseriawinnie@gmail.com](mailto:Hamburgueseriawinnie@gmail.com)

- e) Solicitudes Erte: 730 registros. Ninguna columna tiene celdas vacías y todos los valores son significantes.
- f) Eventos agenda cultural: 252 registros. 16 de las 27 columnas no tienen celdas vacías ni incoherencias.

"Fecha fin" presenta 185 celdas vacías (73'41%); "Precio", "Calle" y "Código Postal" tiene vacías 44 (17'46%), 30 (11'90%) y 34 celdas (13'49%), respectivamente.

La "Hora inicio": no aparece en 20 registros (77'94%); en algunos casos es porque no hay hora específica, en otros está incluida en la descripción; a veces también incluye la ubicación del evento. La "Hora fin" no aparece en 238 registros (94'4%), ocurre lo mismo que con la hora de inicio.

Las columnas "Temática" y "Categoría" tienen la misma finalidad y términos en común (por ejemplo, Concierto y Conciertos, Exposición y Exposiciones, Cine o audiovisual y Cine); la primera tiene todas las celdas completas y la segunda tiene 80 registros vacíos (el 31'74%); en el siguiente ítem analizaremos con más detalle esta cuestión.

Con las columnas "Colectivo destinatario" y "Destinatarios" ocurre lo mismo, ambas contienen el mismo tipo de información; la primera tiene 240 celdas vacías (95'4%) y la segunda está completa y normalizada -público familiar, todos los públicos, adultos-).

➤ **La información está publicada de forma normalizada/estandarizada (hay información que tiene estándares asociados, por ej. la información**

**geográfica):** Hay datos que siguen estándares; por poner un par de ejemplos, las coordenadas se dan en el sistema estándar decimal simple, y la longitud se mide de acuerdo a las unidades del Sistema Internacional (metros, kilómetros).

Respecto a la normalización de los datos en los datasets, se han detectado ciertas carencias que pueden afectar a las búsquedas y a la integración de datos:

En el dataset "Comercios" es destacable la falta de normalización en el campo "Dirección" a la hora de insertar términos como "avenida" ("Av.", "Avd.", "Avda.", "Avenida"), "Calle" ("C/", "Calle") o "Carretera" ("Carretera", "Crta.", "Crtra.", "Ctra."), y en muchas ocasiones ni siquiera se menciona el tipo de vía. Sin embargo, la columna "Sector comercial" está bien normalizada.

En el dataset "Eventos de la agenda cultural categorizados y geolocalizados" se indica que un evento es gratuito mediante los términos "Espectáculo gratuito", "Entrada libre", "Gratis", "Gratuito" y "Gratis"; las horas a veces se indican solas (20:00) o acompañadas de "hrs." O "h."

También es destacable que las columnas "Temática" y "Categoría" tienen el mismo contenido y términos en común (por ejemplo, Concierto y Conciertos, Exposición y Exposiciones, Cine o audiovisual y Cine), y como se vio en el ítem anterior, la primera tiene todas las celdas completas y la segunda tiene 80 registros vacíos; esto afecta a las búsquedas que se hagan en el dataset en el portal de análisis, ya que éste permite filtrar por Categoría, pero no por Temática (Figura 23):

Título	Descripción	Temática	Categoría
Plano del convento de San Francisco. I	-p>-Museo de Valladolid. El Convento de San Francisco. I	Exposición	Exposiciones
'Bolerios eternos' (A.C. Mágicos 70's)	-p>-Repertorio de los boleros más famosos	Concierto	Conciertos
'Concierto homenaje a Estopa'	-p>-Somos un grupo tributo a Estopa	Concierto	Conciertos
Feria de Muestras de Valladolid	-p>-Feria de Muestras de Valladolid	Feria o mercad	Ferias y Fiestas
Geolaciana	-p>-IV Jornadas de Geología del Aula de Geología	Conferencia	Conferencias y Cursos
III Certamen de fotografía de naturaleza	-p>-Certamen de fotografía	Concursos o j	Otros
'Magia entre libros' (Mago Oski)	-p>'Mago Oski' presenta un espectáculo	Magia	Espectáculos
'A paso de rueda' (Teatro del Azar)	-p>-En plena II República española tres	Teatro títeres	Teatro títeres
'Abre el ojo' (Diabla Teatro)	-p>-Comedia de capa y espada, donde	Teatro títeres	Teatro títeres
'Antes muerta que convicta' (Beatriz R)	-p>-ANTES MUERTA QUE CONVICTA	Teatro títeres	Teatro títeres
'Conversaciones con mamá' (Descalzo)	-p>-M Luisa Merlo	Teatro títeres	Teatro títeres

Título	Temática	Categoría	Fecha de inicio
164 Encuentros: '¿De qué hablamos cuando hablamos de arte?'	Taller	Otros	06/03/2020
165 Encuentros: '¿De qué hablamos cuando hablamos de arte?'	Taller	Otros	30/04/2020
166 Taller familiar: '¿Se fueron por ahí y? ¡volvieron por allí!?. La extraor	Taller	Otros	08/07/2020
167 Taller familiar: '¿Se fueron por ahí y? ¡volvieron por allí!?. La extraor	Taller	Otros	29/07/2020
168 Taller familiar: '¿Se fueron por ahí y? ¡volvieron por allí!?. La extraor	Taller	Otros	15/07/2020
169 Taller familiar: '¿Se fueron por ahí y? ¡volvieron por allí!?. La extraor	Taller	Otros	22/07/2020
170 Talleres infantiles de verano: La ciencia del Museo II: el barro	Taller	Otros	07/07/2020
171 15 Festival de Teatro Clásico en la Villa del Caballero 'Olmedo Clásic	Teatro títeres monólogo		24/07/2020
172 'A donde el viento nos lleve' (Morfeo Teatro)	Teatro títeres monólogo		31/12/2022
173 'A paso de rueda' (Teatro del Azar)	Teatro títeres monólogo		31/12/2020
174 'Abrapalabra' (Crispin D'Olot)	Teatro títeres monólogo		23/12/2020
175 'Abre el ojo' (Diabla Teatro)	Teatro títeres monólogo		31/12/2020
176 'After de party' (Thomas Noone Dance)	Teatro títeres monólogo		31/12/2022
177 'Agallas' (kamaru Teatro)	Teatro títeres monólogo		31/12/2020
178 'Alucina Pepinillos' (Dinamia Teatro)	Teatro títeres monólogo		31/12/2020

Figura 23. Temática y Categorías dataset "Eventos" en el portal y CSV

De esta manera, hay 80 registros que no son recuperables en la web si filtramos por categorías, porque no tienen ninguna asignada; por poner un par de ejemplos, los registros "A paso de rueda" y "Abrapadabra" (Figura 23, en el CSV) no se han recuperado filtrando sucesivamente por las diferentes Categorías.

➤ **Los valores de los datos de un mismo campo en un conjunto se ajustan a una medida única determinada (por ejemplo, metros, euros, etc.):** Sí, los datos son consistentes en este sentido; la única discrepancia encontrada en los datasets analizados está el dataset "Solicitudes Erte": en el campo "Provincia" se indican las provincias, pero también aparece "Dirección general de Trabajo y Prevención de Riesgos Laborales" 67 veces y "Ministerio de Trabajo y Economía Social" aparece 60 veces, por lo que el título del campo puede conducir a error, ya que esos datos no son provincias.

➤ **Los datos publicados de forma periódica mantienen la misma estructura, nombre de atributos, tipo y tamaño:** Para analizar este ítem y determinar si los datos son consistentes, hemos recurrido a tres datasets que recogen los Presupuestos en tres periodos diferentes:

- Presupuestos generales de la Comunidad de Castilla y León para 2016 (inicio de la publicación: 13 de abril de 2016)
- Presupuestos generales de la Comunidad de Castilla y León para 2017 (inicio de la publicación: 13 de julio de 2017)
- Presupuestos generales de la Comunidad de Castilla y León para 2018 (inicio de la publicación: 11 de enero de 2018)

La descarga de cada dataset se hace a través de un archivo ZIP con 7 archivos CSV:

- Descripciones centros gestores de gastos: los datos se presentan en dos columnas: "Centro gestor" y "Descripción". Tanto el formato en código de la primera columna como la descripción de cada centro se mantienen invariables de un año a otro, salvo los cambios de denominación sucedidos en la vida real:

	A	B
1	CENTRO GESTOR	DESCRIPCIÓN
2		1 DE LA PRESIDENCIA
3		101 GABINETE DEL PRESIDENTE
4		103 SECRETARÍA GENERAL
5		104 D.G. RELAC.INSTITUC.Y ATENCIÓN CIUDADANO
6		105 D.G. DE ANÁLISIS Y PLANIFICACIÓN
7		106 DIRECCIÓN DE COMUNICACIÓN
8		107 DIRECCIÓN DE LOS SERVICIOS JURÍDICOS
9		108 DIR. ORD.TERRITORIO Y ADMINISTRAC.LOCAL
10		111 D.G. DE LA FUNCIÓN PÚBLICA

Figura 24. Descripciones centros gestores de gastos

- Descripciones centros gestores de ingresos: Hay dos columnas igual que en el caso anterior, "Centro gestor" y "Descripción". Aquí también hay algún cambio parejo al mundo real, pero los códigos y la descripción empleada es la misma en los tres años.
- Descripciones de funcionales: Las dos columnas se denominan "Código" y "Descripción". Se mantienen el formato, códigos y descripción en los 3 años. Solo hay variaciones puntuales por cambios de denominación o supresión de centros o servicios en la vida real.

- Descripciones económicas de gastos: La información se presenta también en las columnas "Código" y "Descripción"; se mantienen la estructura, contenido y terminología utilizada, salvo los cambios puntuales ya descritos.
- Descripciones económicas de ingresos: las mismas columnas que en el caso anterior; la única diferencia apreciada, además de la desaparición o aparición de dotaciones (reflejo del funcionamiento de la institución en el mundo real), es que en el archivo de 2016 no hay tildes.
- Dotaciones presupuesto de gastos: estructura en 12 columnas que se mantiene en los tres años: Sección, Servicio, Área de gasto, Política de gasto, Grupo de programas, Programa, Subprograma, Capítulo, Artículo, Concepto, Subconcepto, Presupuesto.
- Dotaciones presupuesto de ingresos: estructura en 8 columnas: Sección, Servicio, Capítulo, Artículo, Concepto, Subconcepto, Partida, Presupuesto.

	A	B	C	D	E	F	G	H
1	SECCIÓN	SERVICIO	CAPÍTULO	ARTÍCULO	CONCEP	SUBCONCEP	PARTII	PRESUPUESTO
2	0	0	1	10	100	1000	10000	1810616490
3	0	0	1	11	110	1100	11000	200000000
4	0	0	1	11	111	1110	11100	380000000
5	0	0	1	11	112	1120	11200	250000000
6	0	0	2	20	200	2001	20010	194965875
7	0	0	2	20	200	2002	20020	34125
8	0	0	2	20	201	2011	20110	99255000
9	0	0	2	20	201	2012	20120	745000
10	0	0	2	21	210	2100	21000	2064138380
11	0	0	2	22	220	2200	22000	30959680
12	0	0	2	22	220	2201	22010	11608850
13	0	0	2	22	220	2203	22030	232559070
14	0	0	2	22	220	2204	22040	606452520
15	0	0	2	22	220	2205	22050	12688570
16	0	0	2	22	220	2206	22060	882880

Figura 25. Dotaciones presupuesto ingresos 2018

En resumen, en todos los archivos que conforman los 3 datasets de Presupuestos se mantiene la estructura a lo largo de los años, el nombre de los atributos no cambia, se mantiene el tipo de los datos (el dinero siempre se expresa en números enteros, el formato de los códigos es igual, las descripciones de los códigos son iguales, y el tamaño de los datos siempre es el mismo).

➤ **Compromiso de mantener la periodicidad en la publicación de datos:** en el apartado dedicado a los términos de uso se indica que la Junta "procurará asegurar la actualización en un tiempo razonable [...] aunque no puede garantizarlo totalmente, a pesar de los esfuerzos por gestionar adecuadamente dichos conjuntos de datos"; no se habla específicamente del mantenimiento de la periodicidad.

➤ **Los conjuntos se encuentran asociados a un identificador de recursos uniforme (URI):** Sí. En el apartado de buenas prácticas citado más arriba se explica que el portal utiliza un esquema de URIs "estable, persistente y extensible [...] de forma que se pueda hacer referencia a estos [los datos] de forma única y fiable, requisito clave para facilitar su posterior reutilización".

Los 5 datasets objeto de análisis tienen los siguientes URIs:

<a href="https://datosabiertos.jcyl.es/web/jcyl/set/es/cultura-ocio/agenda_cultural/1284806871500">https://datosabiertos.jcyl.es/web/jcyl/set/es/cultura-ocio/agenda_cultural/1284806871500</a>
<a href="https://datosabiertos.jcyl.es/web/jcyl/set/es/medio-ambiente/limites-provinciales-cyl-recintos/1284687285400">https://datosabiertos.jcyl.es/web/jcyl/set/es/medio-ambiente/limites-provinciales-cyl-recintos/1284687285400</a>
<a href="https://datosabiertos.jcyl.es/web/jcyl/set/es/comercio/comercios-domicilio/1284947248341">https://datosabiertos.jcyl.es/web/jcyl/set/es/comercio/comercios-domicilio/1284947248341</a>
<a href="https://datosabiertos.jcyl.es/web/jcyl/set/es/empleo/situacion-coronavirus-ertes/1284942625786">https://datosabiertos.jcyl.es/web/jcyl/set/es/empleo/situacion-coronavirus-ertes/1284942625786</a>
<a href="https://datosabiertos.jcyl.es/web/jcyl/set/es/salud/situacion-coronavirus-hospitales/1284941728695">https://datosabiertos.jcyl.es/web/jcyl/set/es/salud/situacion-coronavirus-hospitales/1284941728695</a>

Figura 26. URIs de los datasets de muestra

Los URIs se forman por un dominio y una ruta que se compone de distintas variables:

- "datosabiertos.jcyl.es" es el dominio;
- "set" identifica el tipo de URI, indica que conduce a la ficha de un dataset;
- "es" hace referencia al idioma español;
- "cultura-ocio", "medio-ambiente", "comercio", "empleo" y "salud" se refieren al sector del dataset, representado con un identificador especificado en la taxonomía de sectores primarios del anexo IV de la NTI;
- "agenda\_cultural", "limites-provinciales-cyl-recintos", "comercios-domicilio", "situacion-coronavirus-ertes" y "situacion-coronavirus-hospitales" son títulos descriptivos que identifican a los conjuntos de datos representados, y van acompañados por un identificador único (los números que van al final de cada enlace).

➤ **Estadísticas de uso: número de visualizaciones y de descargas:**

Hay 8 conjuntos de datos relativos a las estadísticas web; son de actualización diaria:

Estadísticas web	Información proporcionada
Ubicaciones ciudades	Nº de visitas por ciudades y sitio web
Navegadores	Navegadores empleados para navegar por cada sitio web
Sistemas operativos Windows	Versión de Windows utilizada para navegar por cada sitio
Visitas	Nº de visitas, usuarios únicos, páginas vistas, porcentaje de rebote, duración de la sesión, páginas por sesión, usuarios nuevos, porcentaje de salidas y porcentaje de nuevas sesiones por sitio web
Navegadores Explorer	Versión de Internet Explorer utilizada para navegar por cada sitio web
Páginas	Páginas más visitadas de cada sitio web (con más de 100 páginas vistas al día)
Dispositivos	Nº de sesiones por dispositivos utilizados para navegar por sitio web
Sistemas operativos	Sistemas operativos empleados para navegar por cada sitio

Tabla X. Estadísticas web

- **Nivel de reutilización (nº de estrellas según modelo de Berners Lee):** Combinando las posibilidades ofrecidas por el portal de datos abiertos y el portal de análisis, muchos datasets obtendrían las 5 estrellas, por estar disponibles bajo formatos estructurados no propietarios y legibles por máquina, por la utilización del estándar abierto RDF, por el uso de URIs y por emplear datos enlazados para brindar contexto.

### 3.3.3. Contenido del portal en relación a la rendición de cuentas

En este apartado se analiza si el portal contiene la información detallada por Curto-Rodríguez (2020) en su lista de elaboración propia para comprobar la existencia de datos económico-financieros, información de los gobernantes, y ayudas y subvenciones en los portales de datos abiertos.

El portal no proporciona un mecanismo para identificar claramente los conjuntos de datos relacionados con la rendición de cuentas, al contrario de lo que sugiere Lourenço (2015) en su análisis de portales gubernamentales, por lo que hay que buscar en diferentes datasets y en la información adicional que proporcionan dichos datasets.

- **Rendición de cuentas económico-financiera:**

Los presupuestos de gastos e ingresos están disponibles en los datasets que componen los "Presupuestos Generales de la Comunidad de Castilla y León" de los años 2016, 2017 y 2018, donde "se recogen todos los ingresos previstos y los gastos a realizar por la Administración General de la Comunidad, sus organismos autónomos y entes públicos de derecho privado y aquellas otras entidades u órganos del sector público autonómico que tengan presupuesto limitativo".

La descripción (memoria) de los programas presupuestarios de 2018 y de años anteriores (hasta 2009) puede consultarse en la información adicional que indica el dataset; ésta conduce a la página de Hacienda y Finanzas Públicas de Castilla y León<sup>19</sup>, donde también puede verse el presupuesto de empresas y fundaciones públicas de los años 2009 a 2018 entrando dentro de los programas presupuestarios correspondientes a esos años.

La ejecución presupuestaria de 2018, 2019 y 2020 puede consultarse en los datasets "Ejecución del presupuesto de la Administración de la Comunidad – gastos" y "Ejecución del presupuesto de la Administración de la Comunidad – ingresos"; en el portal de análisis pueden consultarse estos datos por anualidades y de manera gráfica (Figura 27).

En la página antes mencionada de Hacienda y Finanzas Públicas pueden consultarse los resúmenes mensuales de la Ejecución del Presupuesto de la Administración de la Comunidad de cada ejercicio, informes sobre el cumplimiento del objetivo de estabilidad presupuestaria, deuda pública y regla de gasto; y la cuenta general de la Comunidad Autónoma.<sup>20</sup>

No se han localizado las "Cuentas anuales de empresas públicas" ni el "Informe de fiscalización".

<sup>19</sup> <https://hacienda.jcyl.es/web/jcyl/Hacienda/es/Plantilla100/1246993413454/ / />

<sup>20</sup> <https://hacienda.jcyl.es/web/jcyl/Hacienda/es/Plantilla100/1284381002332/ / />

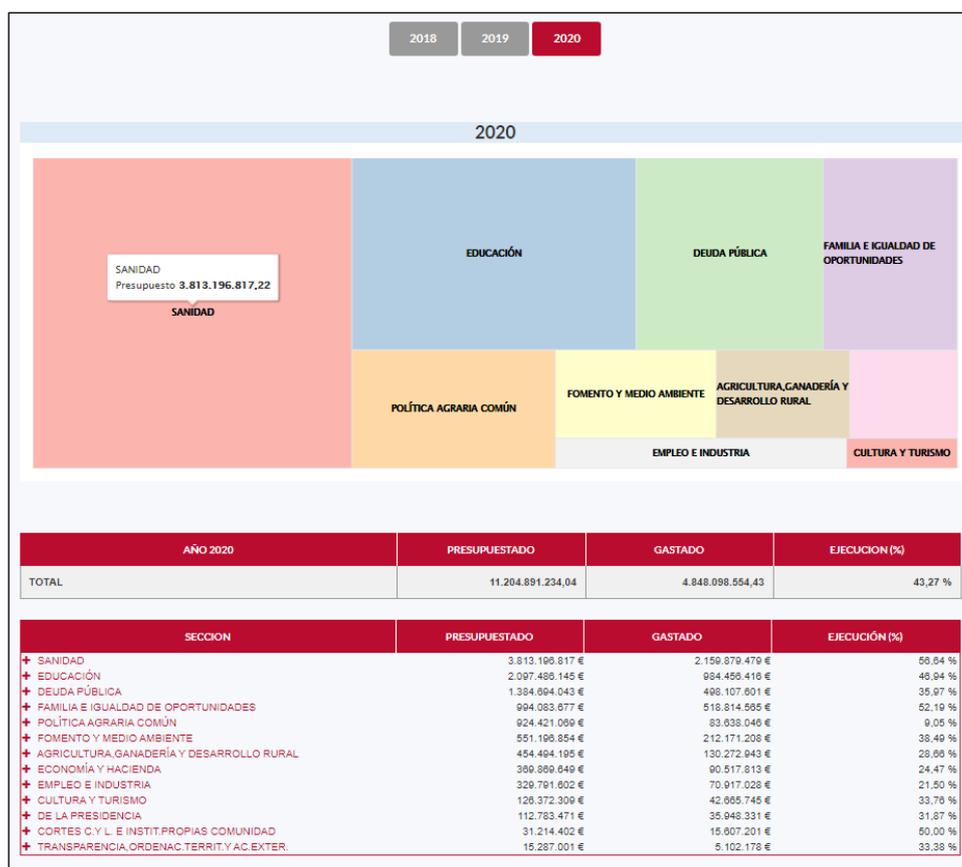


Figura 27. Ejecución presupuesto de la Administración de la Comunidad

### ➤ Rendición de cuentas de los gobernantes:

Salvo las “declaraciones de actividades de miembros del gobierno”, podemos consultar en la web toda la información requerida en la tabla. El organigrama está disponible en la web de la Junta<sup>21</sup>, y en el dataset “Altos cargos” encontramos el cargo, nombre de quien lo ocupa, enlace a fotografía, experiencia profesional, fecha de posesión y formación académica, además de un enlace a una ficha detallada en la que se incluye su contacto y su declaración de bienes. La ficha de este dataset ofrece más información relacionada con un enlace a “Portal de Gobierno abierto - Altos Cargos”<sup>22</sup>, donde se informa de sus retribuciones y otros datos de interés.

En lo referente a la lista y retribuciones de cargos de confianza, en la página de Empleo Público<sup>23</sup> hay un listado en el que aparece el personal eventual que “realiza funciones en puestos de trabajo expresamente calificados como de confianza o asesoramiento especial no reservados a funcionarios o a personal laboral”.

### ➤ Rendición de cuentas en las ayudas y contrataciones:

Se ha encontrado información sobre Ayudas y subvenciones y Becas, premios o concursos en los datasets “Convocatorias de Ayudas y Subvenciones”, “Premios” y “Subvenciones concedidas”. Los Contratos adjudicados se pueden consultar en “Licitaciones” y “Contratos menores”, por poner un par de ejemplos de adjudicaciones.

<sup>21</sup> <https://www.jcyl.es/junta/cp/OrganigramaJunta.pdf>

<sup>22</sup> <https://gobiernoabierto.jcyl.es/web/es/transparencia/altos-cargos.html>

<sup>23</sup> <https://gobiernoabierto.jcyl.es/web/es/transparencia/empleo-publico.html>

#### 4. CONCLUSIONES Y PROPUESTA DE MEJORAS

El portal analizado se enmarca dentro de la estrategia sobre Gobierno Abierto integrada desde hace años en la Junta de Castilla y León, y es un buen ejemplo de la aplicación de los principios de datos abiertos a la Administración. Al abordar en el marco teórico de este trabajo los problemas en la apertura de datos, se observaba cómo muchas administraciones se limitaban a crear portales sin mantener el proceso cíclico necesario para generar valor económico y potenciar la transparencia; en el caso que nos ocupa, este proceso parece estar en movimiento constante, tanto en el incremento progresivo de datasets ofrecidos, como en la introducción de herramientas que mejoren su uso, la adaptación a normas y estándares y la puesta en valor de los datos; no en vano, en los últimos meses diversos medios de comunicación han destacado su transparencia y calidad informativa<sup>24</sup> en la difusión de los datos relativos al COVID en Castilla y León.

El marco de evaluación elaborado ha permitido analizar de modo sistemático la muestra de datasets seleccionada, obteniendo unos resultados que nos permiten concluir de manera general que el portal y los datos ofrecidos son de calidad; no era objeto de este trabajo medir cuantitativamente dicha calidad, pero sí detectar aquellos aspectos que pudieran mermarla.

Entre las características de calidad del portal destacamos como puntos fuertes la claridad en la exposición de su contextualización, objetivos, condiciones de uso y buenas prácticas, además de su sencilla navegabilidad, su clasificación temática y la puesta a disposición del usuario de herramientas para la interpretación de los datos; por el contrario, entre sus debilidades hay que resaltar la inexistencia de un listado completo de datasets claramente visualizable en web, la falta de homogeneidad en las palabras clave, una excesiva sencillez del buscador (que además no está complementado por una ayuda específica), y cierto desaprovechamiento del *feedback* de los usuarios y de los beneficios de las redes sociales.

Respecto a los conjuntos de datos, son de calidad tanto por su representación visual como su contextualización, sus diferentes formatos de descarga, sus licencias de uso, su cumplimiento de estándares, su estructura interna y su granularidad, entre otros aspectos; como debilidades destacamos cierto incumplimiento de lo indicado en la Norma Técnica de Interoperabilidad en lo referente a metadatos, la imposibilidad de ver juntas las diferentes fechas asociadas a la creación y modificación del catálogo, la falta de programación en la actualización de algunos datasets, y la existencia de celdas vacías y valores nulos en algunos de ellos.

Por último, el contenido del portal en relación a la rendición de cuentas se ciñe bastante a lo deseable según el marco de evaluación propuesto, pero consideramos que la información está en algunos casos bastante dispersa entre las diferentes secciones de la web institucional, dificultando su consulta.

Visto lo anterior, y de acuerdo a las indicaciones de los autores y la normativa reseñados en el presente trabajo, consideramos que podrían mejorarse los siguientes puntos:

➤ **Implementación del portal multilingüe:** Máchová y Lnenicka (2017) y Medina Sánchez (2015) recomiendan que los portales se ofrezcan en varios idiomas para tener más usuarios potenciales, por lo que sería deseable instaurar la opción de verla al menos en inglés. Como hemos visto, parece que se está desarrollando (o está desarrollada, pero no funciona) la versión en inglés en el portal de análisis; sería

<sup>24</sup> [https://www.elconfidencial.com/tecnologia/ciencia/2020-05-03/asturias-coronavirus-leccion-ejemplaridad\\_2575692/](https://www.elconfidencial.com/tecnologia/ciencia/2020-05-03/asturias-coronavirus-leccion-ejemplaridad_2575692/)  
[https://elpais.com/sociedad/2020/04/25/actualidad/1587832071\\_409181.html](https://elpais.com/sociedad/2020/04/25/actualidad/1587832071_409181.html)

necesario solventar este fallo y no ofrecer la opción de utilizar el portal en inglés hasta que no esté completamente disponible, ya que puede generar cierta frustración en el usuario y hacerle pensar que el portal no tiene la calidad suficiente.

- **Mecanismos de control de la calidad:** Aunque el portal indica que se realiza un trabajo continuo de auditoría del sitio web, se desconoce si esa auditoría es realizada por alguna entidad externa o es un proceso interno de la propia Junta. Tal y como aconseja Lourenço (2015), el portal debería tener mecanismos externos de control de la calidad; si además se certifica ese control externo, aumentaría la credibilidad y la confianza del usuario.
- **Explicitar el sistema de gestión utilizado en el portal de datos:** Sabemos que el portal de análisis está basado en la plataforma Opendatasoft porque así se indica, pero no tenemos claro cuál es el sistema empleado en el portal de datos; según Máchová y Lnenicka (2017) este dato debe proporcionarse, por lo que recomendamos su indicación.
- **La creación de cuentas de usuario** está bastante oculta; consideramos que podría ponerse junto al enlace de "iniciar sesión" para estar más visible, y facilitar un enlace que informe de las utilidades adicionales que se proporcionan a los usuarios registrados.
- **El portal de datos no permite ver un listado de datasets** de una manera cómoda: aunque existe la posibilidad de descargar el listado de conjuntos de datos y buscar en él, en muchas ocasiones el usuario no quiere o puede hacer esa descarga, por lo que necesita ver el listado directamente en la web. Ver los datasets de 10 en 10 en 54 páginas resulta agotador, además de ser una opción que no se indica claramente (como recordaremos, es necesario hacer una búsqueda vacía para ver este listado). El portal debería proporcionar un listado visualizable en web, sin necesidad de descargar el catálogo de datos, en el que pudieran ordenarse y filtrarse los datasets.
- **La función de búsqueda simple** no es suficiente para facilitar el acceso a los datos, sobre todo si el usuario no sabe muy bien qué está buscando; una búsqueda a ciegas puede desmotivarlo a continuar en el portal. La búsqueda seguramente sea la función más utilizada de la plataforma, por lo que se deben proporcionar potentes herramientas para obtener un rendimiento óptimo. En el caso del portal de análisis, la presentación en grupos temáticos y la posterior posibilidad de filtrado sin necesidad de búsqueda suplen en parte esta carencia, pero al menos en el buscador del catálogo de datos es necesario introducir algunas mejoras.

De acuerdo a las recomendaciones dadas por datos.gob.es (2017b), la búsqueda debe combinar las búsquedas simples con búsquedas avanzadas en las que el usuario tenga diversas opciones para escoger; debe ser flexible para permitir filtrar por todos los posibles valores disponibles y combinar resultados según las necesidades del usuario (por temática, palabras clave, formatos, antigüedad de los datos, frecuencia de actualización, entidades productoras de los datos, etc.).

También es factible ofrecer una búsqueda asistida que complemente a las dos anteriores, pero su diseño debe estar muy cuidado para poder presentar al usuario todas las opciones de selección posibles, sin forzarle a adaptarse a lo que le ofrezca el sistema.

- **Las palabras vacías, tildes y signos de puntuación afectan a la recuperación de datasets en el portal de análisis,** cosa que no ocurre en el portal de datos (tal y como indica la ayuda de su buscador). Es necesario determinar por qué ocurre esto y solucionarlo, ya que afecta a la recuperación de la información y a

la confianza del usuario en los resultados obtenidos. El funcionamiento de las búsquedas debería ser el mismo en el portal y en la herramienta de análisis, y de no ser así, deben especificarse de manera clara los criterios de búsqueda en ambos.

- **Es necesaria una normalización de temas y palabras clave, y la unificación de criterios** a la hora de redactarlos: las "Materias" del portal de datos se denominan "Tema" en el portal de análisis, por lo que sería recomendable llamarlos de la misma manera para no confundir al usuario. Respecto a las palabras clave, consideramos necesario hacer una revisión para eliminar términos erróneos, y elaborar un vocabulario controlado que unifique criterios de redacción.
- Las **funcionalidades de ayuda del portal de datos** deberían ser específicas del portal, con ejemplos claros aplicables al portal; más concretamente, consideramos que la ayuda del buscador no debería ser la general de toda la web institucional, sino centrarse en las prestaciones de este buscador en concreto: operadores booleanos, uso de comillas, cómo combinar los valores de búsqueda posibles, etc., presentando diversos ejemplos de búsqueda y resultados aplicados a este portal.
- **Habilitar un punto de acceso SPARQL:** entre las pautas indicadas a la hora de implementar un portal de datos abiertos, está la recomendación de ofrecer acceso a los datos a través de consultas SPARQL.
- **Los foros de discusión de usuarios deberían estar habilitados:** La implementación de foros mejoraría la característica de calidad "Contactabilidad e Interactividad", muy relacionada con "Valor y utilidad".

En los portales basados en Opendatasoft se puede activar la función Disqus, servicio que permite a los usuarios publicar comentarios para compartir ideas y opiniones, formular preguntas en relación a un conjunto de datos y responder a los comentarios de otros integrantes del foro. Veíamos al principio de este trabajo la importancia del feedback de los usuarios en el proceso de mejora continua de los portales; esta retroalimentación no se da solo en la comunicación entre los administradores del portal y el usuario, sino también entre los usuarios: sus debates, opiniones y valoraciones de los datasets pueden aportar información que de otra manera no llegaría a los administradores.

- También **deberían habilitarse los atajos a las redes sociales:** al igual que Disqus, esta es una función opcional que depende de los administradores del portal; permite compartir en Facebook, Twitter y LinkedIn. Consideramos que es algo fundamental para la difusión de los datasets y para aumentar el número de usuarios potenciales, dando a conocer la existencia del portal de datos abiertos.

Como ejemplo de su utilidad, podemos tomar una publicación de la Junta en Facebook (Junta de Castilla y León, 18 abril 2020); este dataset es uno de los más populares en el momento de redactarse el presente trabajo, y cabría preguntarse si su difusión en esta red social ha tenido algo que ver (Figura 28).

- **El portal de Gobierno Abierto debería tener perfil propio en más redes sociales:** en el apartado de análisis vimos que tiene perfil en Twitter, pero en el resto de redes sociales el perfil es el general de la Junta, por lo que las publicaciones relativas a datos abiertos están entre otras muchas que quizá no tengan tanto interés para los usuarios de datos (un político visitando un colegio puede no ser una publicación muy atractiva para la mayoría). Tener un perfil en Facebook en el que se difundan de manera llamativa y amena los conjuntos de datos (tal y como se hace en Twitter) tendría más impacto y proporcionaría más seguidores.



Figura 28. Publicación en Facebook sobre el dataset "Comercios"

- El contexto de los datos puede no quedar claro si la **fecha de creación, la fecha de actualización y la frecuencia de actualización** no aparecen juntas en todas las fichas, tanto en el portal de datos como en la herramienta de análisis; su separación obliga a consultar los conjuntos en dos lugares diferentes, y en muchos casos los datasets no aparecen en el portal de análisis, por lo que no se puede saber su fecha de actualización.
- En el catálogo de datos hay 96 datasets que **no tienen programada una frecuencia de actualización**, lo que denota una falta de programación y afecta a su credibilidad. Es necesario determinarla e incluirla donde corresponda.
- **No se cumple plenamente lo indicado por la Norma Técnica de Interoperabilidad en lo referente a la presencia de metadatos mínimos requeridos y su esquema de valores:** Es necesario incluir los metadatos de fecha de creación, fecha de actualización y Documentación en el catálogo de datos; esta carencia ya la señalaron (Martín-González & Sainz-Iglesias, 2015) en su análisis de la aplicación de la NTI en los catálogos de las comunidades autónomas. También es necesario cumplir con las especificaciones dadas en la NTI respecto al esquema de valores.
- Es necesario vigilar que **el volumen de datos de los datasets** no sea difícil de manejar; este aspecto depende mucho de los medios tecnológicos de los que disponga el usuario, por lo que debe tenerse en cuenta, entre otros factores, la velocidad de descarga, por poner un ejemplo. Sería recomendable indicar el tamaño a partir de cierto volumen, para que el usuario determine si le interesa descargarlo o por lo menos que esté preparado para un tiempo de espera. Un archivo muy grande no es utilizable, y va en contra de lo indicado en la carta de datos abiertos cuando dice que los datos deben ser "accesibles y utilizables: con facilidad de acceso y bajo

licencia libre, los formatos deben tener en cuenta tanto su legibilidad por parte de las máquinas como la experiencia de los usuarios” (Open Data Charter, 2015).

- El hecho de que existan **celdas vacías y valores nulos o sin significado** afecta a la completitud de los datos y les resta calidad; consideramos necesario hacer un análisis de datos para corregir valores erróneos y rellenar las celdas que deban tener contenido. También es recomendable revisar la existencia de columnas duplicadas haciendo referencia a un mismo contenido, para unificarlas.
- La **falta de normalización de los datos en algunos datasets** afecta a la conformidad, al buen funcionamiento de las búsquedas y a la integración de datos: al igual que comentamos al hablar de las palabras clave, el uso de vocabularios controlados facilita la redacción del contenido de los campos, evita errores y contribuye a la normalización.
- Las búsquedas e integración también se ven afectadas por la **introducción de datos que no se corresponden con el título de la columna**, error que merma la consistencia de los datos. En el presente trabajo se ha analizado una muestra muy pequeña de datasets y este error solo se ha detectado en uno, pero habría que determinar si sucede en más conjuntos y corregirlo.
- El compromiso expreso de la Junta en **mantener la periodicidad de los datasets** aumentaría su credibilidad, además del valor y utilidad de los datos (si un reutilizador no confía en que un conjunto de datos periódico vuelva a publicarse, no lo utilizará).
- El **contenido del portal en relación a la rendición de cuentas** es adecuado en general; hay algunos ítems que no hemos conseguido localizar, pero no podemos determinar si es porque no están representados en el portal o porque no hemos sido capaces de encontrarlos. Consideramos que cierta información está muy fragmentada, distribuida en diferentes secciones de la web institucional, lo que dificulta su recopilación, sobre todo para el usuario medio. Coincidimos con Lourenço (2015) en la necesidad de proporcionar un mecanismo para identificar claramente los conjuntos de datos relacionados con la rendición de cuentas, favoreciendo así la accesibilidad, el valor y la utilidad de los datos.

## BIBLIOGRAFÍA

- Abeleida Díaz, C., Carvajal González, J., Finol Romero, L., y Manzano Chávez, L. (2019). Avances de la transparencia activa y datos abiertos en gobiernos locales: El caso de Coquimbo y La Serena en Chile. *Revista Iberoamericana de Estudios Municipales*, 20, 149-176. <https://www.revistariem.cl/index.php/riem/article/view/241>
- Abella, A.; Ortiz-de-Urbina-Criado, M. y De-Pablos-Heredero, C. (2014). Meloda, métrica para evaluar la reutilización de datos abiertos. *El profesional de la información*, 23(6), 582-588. <https://doi.org/10.3145/epi.2014.nov.04>
- Abella, A., Ortiz-de-Urbina-Criado, M. y De-Pablos-Heredero, C. (2018). Indicadores de calidad de datos abiertos: El caso del portal de datos abiertos de Barcelona. *El Profesional de la Información*, 27(2), 375-382. <https://doi.org/10.3145/epi.2018.mar.16>
- Abella, A., Ortiz-de-Urbina-Criado, M., y De-Pablos-Heredero, C. (2020). Meloda 5: A metric to assess open data reusability. *El Profesional de La Información*, 28(6), e280620. <https://doi.org/10.3145/epi.2019.nov.20>
- Asedie. (2020). Informe ASEDIE sobre el sector infomediario 2020. <http://www.asedie.es/assets/informe-asedie-sobre-el-sector-infomediario-2020-vf.pdf>
- Atenas, J. (Ed.). (2019). *Estado abierto a través de datos abiertos*. Iniciativa Latinoamericana por los Datos Abiertos y Banco Interamericano de Desarrollo. <http://doi.org/10.5281/zenodo.3243389>
- Berners-Lee, T. (2010). *Linked Data*. World Wide Web Consortium. <https://www.w3.org/DesignIssues/LinkedData.html>
- Blank, M., Radu, C., Lincklaen Arriéns, E., y Huyer, E. (2019). *Open Data Maturity Report 2019*. Publications Office of the European Union. <https://op.europa.eu/s/ofnf>
- Cadena-Vela, S., Fuster-Guilló, A. y Mazón, J.N. (2019). Publicando datos abiertos considerando criterios de calidad. *RISTI Revista Ibérica de Sistemas e Tecnologías de Informação* (E22), 295-308. [https://rua.ua.es/dspace/bitstream/10045/99755/1/2019\\_Cadena-Vela\\_et al\\_RISTI.pdf](https://rua.ua.es/dspace/bitstream/10045/99755/1/2019_Cadena-Vela_et al_RISTI.pdf)
- Cadena-Vela, S. (2019). *Marco de referencia para la publicación de datos abiertos comprensibles basado en estándares de calidad* [Tesis doctoral, Universidad de Alicante]. Repositorio Institucional de la Universidad de Alicante. <http://rua.ua.es/dspace/handle/10045/102507>
- Caridad-Sebastián, M. y Martínez-Cardama, S. (2016). Gobierno y Parlamento abierto: La participación ciudadana en el tratamiento y visualización de la información pública. *Revista Interamericana de Bibliotecología*, 39(1), 47-56. <https://revistas.udea.edu.co/index.php/RIB/article/view/25423/20983>

- Clabo, N. y Ramos-Vielba, I. (2015). Reutilización de datos abiertos en la administración pública en España y uso de licencias-tipo. *Revista española de Documentación Científica*, 38(3), e097. <https://doi.org/10.3989/redc.2015.3.1206>
- Curto-Rodríguez, R. (2019). Opciones de interacción, participación y colaboración en los portales de datos abiertos autonómicos españoles para el período 2014-2018. *Revista Buen Gobierno*, 27, 1-20. [https://doi.org/10.35247/buengob\\_27\\_04](https://doi.org/10.35247/buengob_27_04)
- Curto-Rodríguez, R. (2017). El impacto de la ley 19/2013 de transparencia, acceso a la información pública y buen gobierno en la publicación de información relacionada con la rendición de cuentas en los portales de datos abiertos autonómicos españoles. *Auditoría pública: revista de los órganos autónomos de control externo*, 70, 143-152. <http://asocex.es/wp-content/uploads/2017/11/Revista-Auditoria-Publica-n%C2%BA-70-pag-143-a-152.pdf>
- Curto-Rodríguez, R. (2020). Transparencia operativa de las comunidades autónomas españolas mediante sus portales de datos abiertos. *El Profesional de la Información*, 29(1), e290115. <https://doi.org/10.3145/epi.2020.ene.15>
- Datos.gob.es. (s.f.). *Acerca de la iniciativa Aporta*. <https://datos.gob.es/es/acerca-de-la-iniciativa-aporta>
- Datos.gob.es. (2017a). *Guía para publicar datos abiertos de manera rápida y sencilla (con CKAN)*. <https://datos.gob.es/es/documentacion/guia-para-publicar-datos-abiertos-de-manera-rapida-y-sencilla-con-ckan>
- Datos.gob.es. (2017b). *Manual práctico para mejorar la calidad de los datos abiertos*. <https://datos.gob.es/es/documentacion/manual-practico-para-mejorar-la-calidad-de-los-datos-abiertos>
- Datos.gob.es. (2020). *Mapa de iniciativas*. <https://datos.gob.es/es/iniciativas>
- Dekkers, M., Loutas, N., De Keyzer, M. y Goedertier, S. (2014). *Open data and metadata quality* [Presentación]. Joinup Initiative, European Commission. <https://joinup.ec.europa.eu/document/tm22-open-data-metadata-quality-en>
- España. (2007). Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. *Boletín Oficial del Estado*, n. 276 de 17 noviembre de 2007, 47160- 47165. <https://www.boe.es/buscar/doc.php?id=BOE-A-2007-19814>
- España. (2011). Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal. *Boletín Oficial del Estado*, n. 269, de 8 de noviembre de 2011, 116296-116307. <https://www.boe.es/buscar/doc.php?id=BOE-A-2011-17560>
- España. (2013a). Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. *Boletín Oficial del Estado*, n. 295, de 10 de diciembre de 2013. <https://www.boe.es/buscar/act.php?id=BOE-A-2013-12887>

- España. (2013b). Norma Técnica Interoperabilidad. *Boletín Oficial del Estado*, n. 54, de 4 de marzo de 2013, 17045-17071. [https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2013-2380](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2013-2380)
- España. (2015). Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. *Boletín Oficial del Estado*, n. 164, de 10 de julio de 2015, 57436-57450. <https://www.boe.es/eli/es/l/2015/07/09/18>
- España. (2018). Real Decreto 1112/2018, de 7 de septiembre, sobre accesibilidad de los sitios web y aplicaciones para dispositivos móviles del sector público. *Boletín Oficial del Estado*, n. 227, de 19 de septiembre de 2018, 90533-90549. [https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2018-12699](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2018-12699)
- Federación Española de Municipios y Provincias. (2017). *Datos abiertos: Guía estratégica para su puesta en marcha. Conjuntos de datos mínimos a publicar*. <http://femp.femp.es/files/3580-1617-fichero/Gu%C3%ADa%20Datos%20Abiertos.pdf>
- Federación Española de Municipios y Provincias. (2019). *Datos abiertos FEMP 2019. 40 conjuntos de datos a publicar por las Entidades Locales*. <http://femp.femp.es/files/3580-1938-fichero/DATOS%20ABIERTOS%20FEMP%202019.pdf>
- García García, J., y Curto Rodríguez, R. (2018). Determinantes de la apertura de datos y rendición de cuentas en los gobiernos regionales españoles. *Revista CLAD Reforma y Democracia*, (70), 163-198. <http://old.clad.org/portal/publicaciones-del-clad/revista-clad-reforma-democracia/articulos/070-febrero-2018/determinantes-de-la-apertura-de-datos-y-rendicion-de-cuentas-en-los-gobiernos-regionales-espanoles/?searchterm=curto>
- Haupt, M. (2 mayo 2016). "Data is the New Oil" - A Ludicrous Proposition. Medium. <https://medium.com/project-2030/data-is-the-new-oil-a-ludicrous-proposition-1d91bba4f294>
- Hernández-Pérez, T. (2016). En la era de la web de los datos: primero datos abiertos, después datos masivos. *El Profesional de la Información*, 25(4), 517-525. <https://doi.org/10.3145/epi.2016.jul.01>
- Herrera-Melo, C.A. y González-Sanabria, J.S. (2019). Proposal for the Evaluation of Open Data Portals. *Revista Facultad De Ingeniería*, 29(54), e10194. <https://doi.org/10.19053/01211129.v29.n0.2020.10194>
- Junta de Castilla y León. (2012). *Open Data para administraciones locales*. [https://rmd.jcyl.es/web/jcyl/MunicipiosDigitales/es/Plantilla100Detalle/1274785511218/\\_/1284248761730/Redaccion](https://rmd.jcyl.es/web/jcyl/MunicipiosDigitales/es/Plantilla100Detalle/1274785511218/_/1284248761730/Redaccion)
- Junta de Castilla y León. (2019). *Open Data: Guía para la publicación y reutilización de datos abiertos*. [https://comunidaddigital.jcyl.es/web/jcyl/ComunidadDigital/es/Plantilla100Detalle/1246988232579/\\_/1284851940911/Redaccion](https://comunidaddigital.jcyl.es/web/jcyl/ComunidadDigital/es/Plantilla100Detalle/1246988232579/_/1284851940911/Redaccion)

- Junta de Castilla y León. (2010). Ley 2/2010, de 11 de marzo, de Derechos de los Ciudadanos en sus relaciones con la Administración de la Comunidad de Castilla y León y de Gestión Pública. *Boletín Oficial de Castilla y León*, n. 55, de 22 de marzo de 2010, 25274-25300.  
<https://datosabiertos.jcyl.es/web/jcyl/RISP/es/Plantilla100Detalle/1284162108799/Normativa/1284200817130/Redaccion>
- Junta de Castilla y León. (2012). Acuerdo 17/2012, de 8 de marzo, de la Junta de Castilla y León, por el que se pone en marcha el Modelo de Gobierno Abierto de la Junta de Castilla y León. *Boletín Oficial de Castilla y León*, n. 49, de 9 de marzo de 2012, 17691-17696.  
<https://datosabiertos.jcyl.es/web/jcyl/RISP/es/Plantilla100Detalle/1284162105643/Normativa/1284214228451/Redaccion>
- Junta de Castilla y León. (18 abril 2020). *La Junta habilita una guía digital con los comercios de proximidad y las farmacias que ofrecen servicio de reparto a domicilio* [incluye imagen]. Facebook.  
<https://www.facebook.com/juntadecastillayleon/posts/3721128697958381>
- Leonangeli, S., y Maenza, R. R. (16-20 septiembre 2019). *Hacia un paradigma de calidad de datos abiertos adecuado al propósito de gobierno abierto*. XIII Simposio de Informática en el Estado [Simposio]. 48º Jornadas Argentinas de Informática, Salta, Argentina.  
<http://sedici.unlp.edu.ar/handle/10915/89734>
- Lourenço, R. P. (2015). An analysis of open government portals: A perspective of transparency for accountability. *Government Information Quarterly*, 32(3), 323-332. <https://doi.org/10.1016/j.giq.2015.05.006>
- Máchová, R. y Lnenicka, M. (2017). Evaluating the Quality of Open Data Portals on the National Level. *Journal of Theoretical and Applied Electronic Commerce Research*, 12(1), 21-41. <https://doi.org/10.4067/S0718-18762017000100003>
- Martín-González, Y. y Sainz-Iglesias, P. (2015). Aplicación de la norma técnica de interoperabilidad de reutilización de recursos de información en los catálogos de información pública de las comunidades autónomas. *Revista General de Información y Documentación*, 25(2), 287-311.  
[https://doi.org/10.5209/rev\\_RGID.2015.v25.n2.51237](https://doi.org/10.5209/rev_RGID.2015.v25.n2.51237)
- Medina Sánchez, V. (2015). Análisis de proyectos «Open Data» en España: Propuestas de mejora. *Cuadernos de Gestión de Información*, 5(1), 53-72.  
<https://revistas.um.es/gesinfo/article/view/232231>
- Open Data Charter. (2015). *Carta Internacional de Datos Abiertos*.  
<https://opendatacharter.net/principles-es/>
- OpenDataSoft. (2020). *Open Data Inception*.  
<https://opendatainception.io/#?q=spain>
- Open Government Working Group. (2007). *The 8 Principles of Open Government Data*. [https://public.resource.org/8\\_principles.html](https://public.resource.org/8_principles.html)
- Open Knowledge Foundation. (s.f.). *¿Qué son los datos abiertos?*  
<http://opendatahandbook.org/guide/es/what-is-open-data/>

- Open Knowledge Foundation. (2020) *Global Open Data Index*.  
<https://index.okfn.org/>
- Organisation for Economic Cooperation and Development. (2020). *Open, Useful and Re-usable data (OURdata) Index: 2019*. OECD Public Governance Policy Papers, nº 01. <https://doi.org/10.1787/45f6de2d-en>
- Organización Internacional de Normalización. (2008). *ISO 25012*.  
<https://iso25000.com/index.php/en/iso-25000-standards/iso-25012>
- Oviedo, E., Mazón, J. N., y Zubcoff, J. J. (2015, Octubre 19-23). *Quality and maturity model for open data portals*. 2015 Latin American Computing Conference (CLEI), Arequipa, Perú.  
<https://doi.org/10.1109/CLEI.2015.7360011>
- Publications Office of the European Union (2019). *Datos Abiertos en Europa 2019*.  
<https://www.europeandataportal.eu/es/dashboard/2019>
- Radu, C., y Cecconi, G. (2019). *Analytical Report 13: Open Data best practices in Europe's Top Performers: Ireland, Spain and France*. Publication Office of the European Union.  
[https://www.europeandataportal.eu/sites/default/files/analytical\\_report\\_13\\_open-data-best-practices-top-performers.pdf](https://www.europeandataportal.eu/sites/default/files/analytical_report_13_open-data-best-practices-top-performers.pdf)
- Royo-Montañés, S., y Benítez-Gómez, A. (2019). Portales de datos abiertos. Metodología de análisis y aplicación a municipios españoles. *El Profesional de la Información*, 28(6), 1-13. <https://doi.org/10.3145/epi.2019.nov.09>
- Ruvalcaba Gómez, E. A. (2020). Datos abiertos. *EUNOMÍA, Revista en Cultura de la Legalidad*, 18, 327-334. <https://doi.org/10.20318/eunomia.2020.5280>
- Transparencia Internacional España. (2016). *Índice de transparencia de las comunidades autónomas 2016*.  
[https://transparencia.org.es/wpcontent/uploads/2016/12/cuadro\\_indicadores\\_incau-2016.pdf](https://transparencia.org.es/wpcontent/uploads/2016/12/cuadro_indicadores_incau-2016.pdf)
- Unión Europea. (2003). Directiva 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información en el sector público. *DOUE*, n. 345, de 31 de diciembre de 2003, 90-96. <https://www.boe.es/buscar/doc.php?id=DOUE-L-2003-82244>
- Unión Europea. (2013). Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público. *DOUE*, n. 175, de 27 de junio de 2013, 1-8.  
<https://www.boe.es/buscar/doc.php?id=DOUE-L-2013-81251>
- Unión Europea. (2019). Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo, de 20 de junio de 2019, relativa a los datos abiertos y la reutilización de la información del sector público. *DOUE*, n. 172, de 26 de junio de 2019, 56-83.  
<https://datosabiertos.jcyl.es/web/jcyl/RISP/es/Plantilla100Detalle/1284162108799/Normativa/1284310884975/Redaccion>

- Venegas Álvarez, M. A. (2018). *Estudio sobre el acceso abierto a datos gubernamentales: El caso de transparencia en Chile, Costa Rica y Uruguay* [Tesis doctoral, Universidad de Chile]. Repositorio académico de la Universidad de Chile. <http://repositorio.uchile.cl/handle/2250/159564>
- Vetró, A., Canova, L., Torchiano, M., Orozco, C., Iemma, R., y Morando, F. (2016). Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, 33(2), 325-337. <https://doi.org/10.1016/j.giq.2016.02.001>
- World Wide Web Foundation. (2017). *Barómetro de los Datos Abiertos*, 4ª edición. <http://opendatabarometer.org/doc/4thEdition/ODB-4thEdition-GlobalReport-ES.pdf>

## ANEXOS

### ANEXO I. GLOSARIO<sup>25</sup>

**Creative Commons:** Licencias que brindan al autor una forma estandarizada de otorgar permiso al público para usar su trabajo creativo bajo la ley de derechos de autor<sup>26</sup>. Disponen de cuatro elementos básicos (Clabo y Ramos-Vielba, 2015):

- a) Reconocimiento: la atribución de la autoría es un elemento que está siempre presente en todas las variedades porque es un principio general e inamovible.
- b) Uso comercial/no comercial: se permite o no comercializar las obras sujetas a la licencia.
- c) Sin/con obra derivada: se puede o no realizar modificaciones, incluida la traducción.
- d) Compartir igual: se obliga a utilizar una licencia del mismo tipo (se asimila al copyleft).

La combinación de estos cuatro elementos permite generar una serie de licencias diferentes. En España hay seis tipos de licencias CC de carácter gratuito dependiendo del grado de protección que se desee sobre la obra.

**CSV:** *Comma Separated Values*, valores separados por coma. Formato de fichero que suele emplearse para importar o exportar datos entre aplicaciones diferentes, aunque similares. Entre otras, pueden utilizar el formato CSV aplicaciones como KSpread, OpenOffice Calc y Microsoft Excel.

**Dataset:** Conjunto de datos relacionados, compuesto por elementos independientes, pero que puede utilizarse como unidad a la que se accede o que se descarga en uno o varios formatos.

**Datos enlazados o vinculados (*linked data*):** Método de publicación de datos estructurados que permite interconectarlos. Se basa en normas y tecnologías web estándar como HTTP y URI, pero, en lugar de utilizarlas para servir páginas web a lectores humanos, las amplía para compartir la información de modo que los ordenadores puedan leerla automáticamente.

**DCAT:** Vocabulario RDF que facilita la interoperabilidad entre catálogos de datos publicados en la web. Permite al editor describir conjuntos de datos y servicios de datos en un catálogo utilizando un modelo y vocabulario estándar que facilita el consumo y agregación de metadatos procedentes de múltiples catálogos, lo que incrementa la capacidad de detección de dichos conjuntos y servicios. El espacio de nombres para los términos DCAT es <http://www.w3.org/ns/dcat#>, y el prefijo sugerido para el espacio de nombres DCAT es dcat.<sup>27</sup>

**Dublin Core:** Modelo de metadatos elaborado y auspiciado por la DCMI (*Dublin Core Metadata Initiative*). Es un sistema de 15 definiciones semánticas descriptivas que pretenden transmitir un significado semántico a las mismas, y fue diseñado específicamente para proporcionar un vocabulario de características "base", capaces de proporcionar la información descriptiva básica sobre cualquier recurso, sin que importe el formato de origen, el área de especialización o el origen cultural.

---

<sup>25</sup> Términos extraídos de <https://data.europa.eu/euodp/es/glossary>, salvo indicación contraria.

<sup>26</sup> <https://creativecommons.org/about/cclicenses/>

<sup>27</sup> <https://www.w3.org/TR/vocab-dcat/>

**HTML:** *HyperText Markup Language* (Lenguaje de Marcas de Hipertexto). Es un lenguaje en formato abierto que se utiliza para el desarrollo de páginas de Internet.

**JSON:** *JavaScript Object Notation*. Formato abierto que utiliza lenguaje natural para transmitir objetos de datos compuestos por pares atributo-valor. Es el formato más comúnmente utilizado para la comunicación asíncrona entre navegador y servidor (AJAJ).

**Metadato:** Suele definirse como "datos sobre datos". Es información estructurada que describe, explica, localiza o, en general, facilita obtener, utilizar o gestionar un recurso informático.

**Punto de acceso (endpoint) SPARQL:** Servicio para realizar consultas SPARQL que presenta las respuestas en forma de conjuntos de resultados SPARQL. Es un conjunto de mejores prácticas para que los proveedores de conjuntos de datos faciliten el URL de su punto final SPARQL y así permitan el acceso a sus datos a través de un programa o una interfaz web.

**RDF:** *Resource Description Framework* (marco de descripción de recursos). Familia de normas internacionales para el intercambio de datos en la web. Se basa en la idea de identificar objetos a través de identificadores web o URI de tipo HTTP y describir los recursos mediante propiedades y valores de propiedad sencillos.

**RSS:** *Really Simple Syndication* (sindicación realmente simple). Formato XML que se utiliza para distribuir contenido en la web. Se utiliza para difundir información actualizada frecuentemente a usuarios que se han suscrito a la fuente de contenidos<sup>28</sup>.

**SPARQL:** Protocolo y lenguaje de consulta RDF (SPARQL) que define un lenguaje de consulta de datos RDF análogo al lenguaje de consulta estructurado (SQL) propio de las bases de datos relacionales.

**URI:** *Uniform Resource Identifier* (identificador uniforme de recursos). Cadena de caracteres que identifica de modo único prácticamente cualquier cosa, desde objetos físicos (edificios) hasta conceptos abstractos (colores). Puede o no resolverse en la web.

**URL:** *Uniform Resource Locator* (localizador uniforme de recursos). Identificador global comúnmente denominado «dirección». Los URL pueden resolverse en la web. Todos los HTTP son URL que también son URI, pero no todos los URI son URL.

**W3C:** World Wide Web Foundation, organización internacional e independiente establecida en 2009 por Sir Tim Berners-Lee; lucha por la igualdad digital y promueve la web abierta como un bien público y un derecho básico<sup>29</sup>.

**XML:** *Extensible Markup Language* (Lenguaje de Mercado Extensible). Lenguaje marcado que define un conjunto de normas para codificar documentos en un formato legible tanto por personas como por máquinas.

---

<sup>28</sup> <https://es.wikipedia.org/wiki/RSS>

<sup>29</sup> <https://webfoundation.org/about/>

## ANEXO II. MARCO DE EVALUACIÓN

CARACTERÍSTICAS DEL PORTAL	
Ítems	Indicadores de calidad
<b>Aspectos técnicos</b>	
Presenta una descripción del portal, misión y objetivos	Accesibilidad, comprensibilidad
Condiciones de uso o términos de reutilización	Apertura, comprensibilidad, conformidad
Portal multilinguaje: al menos en dos idiomas	Accesibilidad
Accesibilidad para personas con dificultades	Accesibilidad
Tiene algún certificado	Credibilidad
El portal tiene mecanismos externos de control de la calidad, y asocia los resultados de dichos procesos de revisión con cada dataset publicado	Credibilidad
El portal presenta el nombre de la organización responsable de almacenar o custodiar los datos de forma permanente	Credibilidad, disponibilidad
Información sobre el sistema de gestión de datos utilizado	Accesibilidad
Se indican las buenas prácticas utilizadas en el portal	Credibilidad, Comprensibilidad
<b>Usabilidad y acceso</b>	
Registro obligatorio o pago para el acceso y/o consulta de los datos públicos	Accesibilidad
Soporte para creación de cuentas de usuario para personalizar las visualizaciones e información mostrada	Accesibilidad
Existencia de mecanismos para agregar, modificar o eliminar conjuntos de datos sin validación de usuario	Accesibilidad, credibilidad
La publicación de datos respeta la legislación vigente en materia de privacidad	Apertura, conformidad
Proporciona un catálogo de datos	Accesibilidad
Indica el número de conjuntos de datos	Accesibilidad, Volumen de datos
Posibilidad de desplazarse a través de las diferentes secciones del portal, conociendo completamente el propósito de cada una	Accesibilidad, comprensibilidad
Organización en categorías temáticas. Utilización de estándares para la categorización de los conjuntos de datos	Accesibilidad, comprensibilidad, conformidad
Utilización de palabras clave (etiquetas)	Accesibilidad, comprensibilidad
Posibilidad de ordenar los datos en el catálogo según criterios diferentes a las categorías	Accesibilidad
Existencia de listas maestras de todas las entidades y fuentes relevantes de conjuntos de datos, actuales y pasadas	Accesibilidad, completitud
Herramientas de búsqueda y selección proporcionadas	Accesibilidad, disponibilidad

Tutoriales, documentación y funcionalidades de ayuda para aprender a utilizar el portal y mejorar la usabilidad	Accesibilidad, comprensibilidad
Ejemplos de uso	Accesibilidad, comprensibilidad
Proporciona APIs para desarrolladores	Accesibilidad, procesabilidad
Acceso a través de consultas SPARQL	Accesibilidad
Posibilidad de visualizar los datos en mapas o gráficas	Accesibilidad, comprensibilidad
Posibilidad de manipular datos online	Accesibilidad
<b>Comunicación y participación</b>	
Lista de preguntas frecuentes (FAQ)	Accesibilidad, comprensibilidad
Concursos o actividades para promover el uso de los datos	Contactabilidad e interactividad, valor y utilidad
Apartado de novedades y noticias relacionadas con los conjuntos de datos	Accesibilidad, Interactividad
Permite la suscripción	Accesibilidad, Interactividad
Registro de reutilizadores	Interactividad, Valor y Utilidad
Mecanismos para que los usuarios soliciten nuevos conjuntos de datos, tipos o formatos de su interés	Contactabilidad e interactividad, valor y utilidad
Existencia de medios suficientes para comunicarse con el responsable de los datos (formulario, correo electrónico, etc.).	Contactabilidad e interactividad
Tiempo de respuesta a una pregunta formulada	Contactabilidad e interactividad
Ejemplos y número de aplicaciones desarrolladas	Valor y utilidad
Posibilita la valoración de los usuarios y foros de discusión	Contactabilidad e interactividad, valor y utilidad
Posibilidad de compartir conjuntos de datos a través de las principales redes sociales como Facebook y Twitter	Contactabilidad e interactividad, valor y utilidad
El portal tiene perfiles actualizados en redes sociales	Contactabilidad e interactividad
<b>CARACTERÍSTICAS DE LOS CONJUNTOS DE DATOS</b>	
Los datos son proporcionados junto con el nombre de su publicador para verificar la autenticidad de la fuente	Credibilidad
Los datos publicados responden a su enunciado pertinente, son correctos de acuerdo a la materia y el tipo de información del dataset	Exactitud
Los datos son pertinentes para la necesidad del usuario	Valor y utilidad
Se proveen los datos suficientes para tener claro el contexto de los datos (título, descripción, propósito de su recolección, fechas de creación y de actualización, frecuencia de actualización)	Actualidad, comprensibilidad, exactitud
La actualización de los datos se realiza: de acuerdo a lo indicado en el portal, de acuerdo a lo establecido por la ley del país, conforme a como los datos cambian	Actualidad
Condiciones de uso y tipo de licencia de cada conjunto de datos	Apertura

El conjunto proporciona definiciones y/o explicaciones que ayuden al usuario a interpretar los datos	Comprensibilidad
Uso de metadatos mínimos y cumplimiento de estándares	Actualidad, comprensibilidad, conformidad, credibilidad, procesabilidad
Mecanismos de descarga en diferentes formatos	Accesibilidad, procesabilidad
Formatos de archivo no propietarios y legibles por máquinas	Procesabilidad
Uso de formatos comprensibles para los usuarios	Comprensibilidad
Tamaño del conjunto de datos	Volumen de datos
El conjunto se organiza internamente en registros estructurados	Procesabilidad
Nivel de granularidad del conjunto de datos	Precisión
Existencia de celdas vacías y valores nulos o sin significado	Compleitud, Exactitud
La información está publicada de forma normalizada/estandarizada (hay información que tiene estándares asociados, por ej. la información geográfica)	Conformidad
Los valores de los datos de un mismo campo en un conjunto se ajustan a una medida única determinada (por ejemplo, metros, euros, etc.)	Consistencia
Los datos publicados de forma periódica mantienen la misma estructura, nombre de atributos, tipo y tamaño	Consistencia
Compromiso de mantener la periodicidad en la publicación de datos	Disponibilidad
Los conjuntos se encuentran asociados a un identificador de recursos uniforme (URI)	Disponibilidad
Estadísticas de uso: número de visualizaciones y de descargas	Valor y utilidad
Nivel de reutilización (nº de estrellas según modelo de Berners Lee)	Conformidad, procesabilidad, valor y utilidad
Exactitud en la agregación <sup>30</sup>	Exactitud
<b>CONTENIDO DEL PORTAL EN RELACIÓN A LA RENDICIÓN DE CUENTAS</b>	
Se proporciona un mecanismo para identificar claramente los conjuntos de datos relacionados con la rendición de cuentas	Accesibilidad, Valor y utilidad
Contenido relacionado con la rendición de cuentas económico-financiera (Ver Anexo III)	Valor y utilidad
Contenido relacionado con la rendición de cuentas de los gobernantes (Ver Anexo III)	Valor y utilidad
Contenido relacionado con la rendición de cuentas en las ayudas y contrataciones (Ver Anexo III)	Valor y utilidad

<sup>30</sup> Este ítem no se ha podido analizar; solo es aplicable a datasets que tienen columnas de agregación o cuando hay dos o más datasets que se refieren a la misma información, pero con diferente nivel de granularidad (Vetrò, 2016); se requiere un mayor conocimiento de los datos para saber qué datasets responden a estos requisitos.

## ANEXO III. CONTENIDO DEL PORTAL EN RELACIÓN A LA RENDICIÓN DE CUENTAS

<b>Rendición de cuentas económico-financiera</b>
<b>Presupuestos de gastos e ingresos:</b>
Por cualquier clasificación (orgánica, funcional o económica)
Por partidas presupuestarias
Memoria con descripción de los programas presupuestarios
<b>Ejecución presupuestaria:</b>
Información mensual de ejecución
Liquidación anual
Modificaciones presupuestarias
Presupuestos de entes, sociedades y fundaciones públicas
Cuentas anuales de empresas públicas
Cuenta general de la Comunidad Autónoma
Cumplimiento del objetivo de estabilidad presupuestaria
Nivel y desglose de endeudamiento
Informe de fiscalización
<b>Rendición de cuentas de los gobernantes</b>
Organigrama y contactos
Retribuciones a miembros del gobierno y altos cargos
Declaraciones de bienes de los miembros del gobierno
Declaraciones de actividades de miembros del gobierno
Lista y retribuciones de cargos de confianza
<b>Rendición de cuentas en las ayudas y contrataciones</b>
Ayudas y subvenciones
Becas, premios o concursos
Contratos adjudicados