

# **Auditoría y metodología de implantación de open data para smart cities**

**IGNACIO MELENDREZ MORETO**

**MÁSTER EN INGENIERÍA INFORMÁTICA, FACULTAD DE INFORMÁTICA  
UNIVERSIDAD COMPLUTENSE DE MADRID**



**TRABAJO DE FIN DE MÁSTER EN INGENIERÍA INFORMÁTICA**  
Curso académico 2015/2016

**Madrid, septiembre de 2016**

Calificación: Aprobado 5.2

**Directora:**  
**Guadalupe Miñana Roperó**



# Autorización de difusión y utilización

IGNACIO MELENDREZ MORETO

Madrid, septiembre de 2016

El abajo firmante, matriculado en el Máster en Ingeniería Informática de la Facultad de Informática, autoriza a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a su autor el presente Trabajo de Fin de Máster: *“Auditoría y metodología de implantación de open data para Smart cities”*, realizado durante el curso académico 2015-2016 bajo la dirección de Guadalupe Miñana Roperó en el departamento de Arquitectura de Computadores y Automática, y a la Biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en internet y garantizar su preservación y acceso a largo plazo.

FIRMADO



# Agradecimientos

En todo este tiempo me he sentido guiado y protegido por Dios. Sé que Él está al control de mi vida. “Porque Jehová da la sabiduría, y de su boca viene el conocimiento y la inteligencia” (Proverbios 2:6)

Quiero agradecer a mi familia, a mis padres y hermanos, quienes siempre me brindan su apoyo incondicional y a Ustedes me debo.

Gracias a mi directora Guadalupe Miñana por su ayuda en todo momento, quien me ha cedido su valioso tiempo y su dedicación lo cual valoro mucho y le estoy eternamente agradecido, sin su ayuda y dirección no hubiese sido posible la culminación de éste trabajo.

Asimismo, quiero agradecer al Ministerio de Educación de Perú, que a través del Programa Nacional de Becas y Crédito Educativo me ha financiado mis estudios e investigación.



# Resumen

El movimiento de datos abiertos es relativamente nuevo, ofrece beneficios significativos a la sociedad y a la economía, promueve la democracia y la responsabilidad de los gobiernos públicos fomentando la transparencia, participación y colaboración de los ciudadanos. Por ser un movimiento relativamente nuevo, son los países que lideran el desarrollo quienes ya han implementado políticas de datos abiertos y ya disfrutan de sus beneficios; sin embargo, hay países en los que aún ni siquiera hay iniciativas de datos abiertos o aún están comenzando. En este trabajo se estudia el uso adecuado de buenas prácticas, normas, métricas y estándares para la implantación de datos abiertos de manera sostenible, automatizable y en formatos accesibles que garanticen la reutilización de los datos con el fin de generar valor a través de ellos, al crear nuevos productos y servicios que contribuyan a mejorar la calidad de vida de los ciudadanos. En ese sentido, se realiza un análisis exploratorio de los principios de datos abiertos, se realiza un análisis sobre la situación actual de iniciativas de datos abiertos, y con el fin de que el proyecto sea de máxima aplicabilidad, se realizan pruebas de la métrica Meloda 4.0 sobre conjuntos de datos del Ayuntamiento de Madrid. Se realiza un análisis y evaluación de los portales de datos abiertos de los Ayuntamientos de Madrid, Zaragoza y Barcelona basándose en la Norma UNE 178301:2015. En concordancia con la filosofía de datos abiertos, se estudia y sugiere el uso de tecnologías de código abierto para la publicación de datos abiertos. Finalmente, como resultado y aplicabilidad de todo lo aprendido, se propone el diseño de una metodología para publicación de datos abiertos orientada a entidades públicas que aún no tienen iniciativas o están comenzando a implementar políticas de datos abiertos.

**Palabras clave:** Datos abiertos, gobierno abierto, transparencia, conjunto de datos, ciudades inteligentes, Meloda, CKAN, DKAN, métricas, reutilización.

## **Abstract**

The movement of open data is relatively new, they offer important benefits to our society and the economy, because they promote the democracy and the responsibility of the public governments fomenting the transparency, participation and the collaboration of the citizens. And is just because is a relatively new movement that the main countries has implemented politics of open data and they already enjoy of their benefits. However, there are some countries that haven't started to work in this system yet.

In this work, we study the right use of best practices, laws, metrics and standards for the implementation of open data in a sustainable way, automatable and in accessible format that guarantee the reuse of data with the objective of generate value trough them, because they create new products and services that contribute to get better the quality of life of citizens.

So, that is the reason why we are going to make an exploratory analysis of the values of open data. The first thing that we have to do is make an analysis of the current situation of the initiatives of open data. And to get our objectives, we make tests of the metrics Meloda 4.0 through a set of data of the Town Hall of Madrid, Zaragoza and Barcelona based in the rule UNE 178301:2015.

In agreement with the philosophy of open data, we study and we suggest the use of technologies of open code to the publication of open data.

Finally, as a result and applicability of everythings that we have learned, we propose the design of a new methodology to the publication of open data conducted to public organizations have not yet initiatives or they just have started to implement politics of new data.

**Key words:** Open data, open government, transparency, dataset, smart cities, Meloda, CKAN, DKAN, metrics, reuse.

# Tabla de contenido

AUTORIZACIÓN DE DIFUSIÓN Y UTILIZACIÓN .....	III
AGRADECIMIENTOS.....	V
RESUMEN .....	VII
ABSTRACT .....	VIII
TABLA DE CONTENIDO.....	IX
ÍNDICE DE FIGURAS.....	XI
ÍNDICE DE TABLAS.....	XIII
ÍNDICE DE ABREVIATURAS.....	XV
<b>CAPÍTULO 1: INTRODUCCIÓN .....</b>	<b>1</b>
1.1. MOTIVACIÓN.....	1
1.2. OBJETIVO DE LA INVESTIGACIÓN.....	2
1.3. RELACIÓN DEL TRABAJO CON LAS ASIGNATURAS DEL MÁSTER.....	3
1.4. ESTRUCTURA DEL TRABAJO .....	4
INTRODUCTION.....	5
<b>CAPÍTULO 2: OPEN DATA .....</b>	<b>9</b>
2.2. ¿QUÉ ES UN OPEN DATA?.....	9
2.3. ESTADO ACTUAL DEL OPEN DATA.....	17
2.4. CADENA DE VALOR DEL OPEN DATA.....	32
2.5. EL OPEN DATA COMO HERRAMIENTA PARA LAS SMART CITIES.....	35
2.6. BENEFICIOS DEL OPEN DATA .....	39
2.7. TECNOLOGÍAS PARA IMPLEMENTAR DATOS ABIERTOS .....	40
<b>CAPÍTULO 3. NORMALIZACIÓN Y MÉTRICAS .....</b>	<b>47</b>
3.1. NORMALIZACIÓN .....	47
3.2. MÉTRICAS .....	60
3.3. APLICACIÓN DE LA MÉTRICA MELODA .....	70
3.4. APORTACIONES Y COMENTARIOS .....	72
<b>CAPÍTULO 4: APLICACIÓN DE LA NORMA UNE 178301:2015 .....</b>	<b>75</b>
4.1. PRIMERA INICIATIVA DE DATOS ABIERTOS .....	76
4.2. AYUNTAMIENTO DE MADRID.....	81
4.3. AYUNTAMIENTO DE ZARAGOZA.....	86
4.4. AYUNTAMIENTO DE BARCELONA.....	92
4.5. RESULTADOS .....	99
4.6. RECOMENDACIONES.....	100
<b>CAPÍTULO 5: PROPUESTA DE UNA METODOLOGÍA PARA PUBLICACIÓN DE DATOS ABIERTOS .....</b>	<b>108</b>
5.1. FASES DE LA METODOLOGÍA .....	109
5.2. ACTIVIDADES DE LA METODOLOGÍA.....	110
5.3. ARTEFACTOS DE LA METODOLOGÍA .....	122
5.4. ROLES Y RESPONSABILIDADES.....	124
<b>CAPÍTULO 6: CONCLUSIONES Y TRABAJO FUTURO .....</b>	<b>125</b>
CONCLUSIONS AND FUTURE WORK .....	127
BIBLIOGRAFÍA .....	129

<b>ANEXOS.....</b>	<b>133</b>
<b>ANEXO 1. INSTALACIÓN DE CKAN.....</b>	<b>133</b>
<b>ANEXO 2. CONFIGURACIÓN DE CKAN.....</b>	<b>136</b>
<b>ANEXO 3. INSTALACIÓN DE DKAN.....</b>	<b>138</b>

# Índice de figuras

Figura 1: Ciclo de los datos abiertos.....	10
Figura 2: Ranking Global de Open Data - Open Knowledge Foundation (2015).....	20
Figura 3: Ranking mundial de open data - Open Data Barometer (2015).....	21
Figura 4: Clasificación de iniciativas en datos abiertos - Open Data Economy (2012).....	22
Figura 5: Evaluación de catálogos y conjuntos de datos abiertos en Europa - Open Data Monitor (2015).....	23
Figura 6: Portal de datos abiertos del gobierno de Estados Unidos.....	24
Figura 7 : Portal de datos abiertos del Gobierno de Canadá.....	25
Figura 8: Portal de datos abiertos del Banco Mundial .....	26
Figura 9: Portal de datos abiertos de la Unión Europea.....	28
Figura 10: Mapa de portales de datos abiertos en España.....	29
Figura 11: Portal de datos abiertos del Gobierno de España.....	30
Figura 12: Líneas de actuación de la iniciativa Aporta.....	31
Figura 13: Portal de datos abiertos del Ayuntamiento de Madrid.....	32
Figura 14: Cadena de Valor de Open Data .....	33
Figura 15: El open data y las ciudades inteligentes.....	37
Figura 16: Beneficios de los datos abiertos .....	39
Figura 17: El Gobierno de España promueve el uso promueve el uso de CKAN.....	41
Figura 18: Arquitectura de CKAN.....	42
Figura 19: Resumen de la Norma UNE 178301:2015.....	50
Figura 20: Valores máximos para cada dominio de la Norma.....	51
Figura 21: Cinco estrellas del Open Data, Tim Berners-Lee.....	61
Figura 22: Procesos de Meloda 4.0.....	69
Figura 23: Portal de datos abiertos del Ayuntamiento de Zaragoza.....	87
Figura 24: Portal de datos abiertos del Ayuntamiento de Barcelona.....	93

Figura 25: Fases de la metodología propuesta para publicación de datos abiertos.....	109
Figura 26: Los tres ejes fundamentales de los datos abiertos .....	125
Figura 27: The three axes of open data.....	127
Figura 28: Primera interfaz de Solr.....	134
Figura 29: Interfaz de inicio de CKAN.....	135
Figura 30: Página de inicio de DKAN .....	143

# Índice de tablas

Tabla 1: Casos de éxito de CKAN.....	43
Tabla 2: Casos de éxito de DKAN.....	44
Tabla 3: Diferencias técnicas entre CKAN y DKAN.....	45
Tabla 4: Resumen estructurado de la Norma UNE 178301:2015 de AENOR.....	53
Tabla 5: Dimensiones, niveles y pesos de Meloda 4.0.....	68
Tabla 6: Rangos de clasificación de la métrica Meloda 4.0.....	70
Tabla 7: Propuesta de indicadores para medir la reutilización real de datos abiertos.....	74
Tabla 8: Niveles asignados a la dimensión estratégica para el caso beta.....	76
Tabla 9: Niveles asignados a la dimensión legal para el caso beta.....	77
Tabla 10: Niveles asignados a la dimensión organizativa para el caso beta.....	77
Tabla 11: Niveles asignados a la dimensión medición para el caso beta.....	78
Tabla 12: Niveles asignados a la dimensión disponibilidad para el caso beta.....	78
Tabla 13: Niveles asignados a la dimensión acceso para el caso beta.....	79
Tabla 14: Niveles asignados a la dimensión calidad de datos para el caso beta.....	79
Tabla 15: Niveles asignados a la dimensión actualización para el caso beta.....	80
Tabla 16: Niveles asignados a la dimensión reutilización para el caso beta.....	80
Tabla 17: Niveles asignados a la dimensión participación y colaboración para el caso beta.....	80
Tabla 18: Niveles asignados a la dimensión estratégica, Ayto. Madrid.....	81
Tabla 19: Niveles asignados a la dimensión legal, Ayto. Madrid.....	81
Tabla 20: Niveles asignados a la dimensión organizativa, Ayto. Madrid.....	82
Tabla 21: Niveles asignados a la dimensión medición, Ayto. Madrid.....	83
Tabla 22: Niveles asignados a la dimensión disponibilidad, Ayto. Madrid.....	83
Tabla 23: Niveles asignados a la dimensión acceso, Ayto. Madrid.....	84
Tabla 24: Niveles asignados a la dimensión calidad de datos, Ayto. Madrid.....	84
Tabla 25: Niveles asignados a la dimensión actualización, Ayto. Madrid.....	85
Tabla 26: Niveles asignados a la dimensión reutilización, Ayto. Madrid.....	85
Tabla 27: Niveles asignados a la dimensión participación y colaboración, Ayto. Madrid.....	86
Tabla 28: Niveles asignados a la dimensión estratégica, Ayto. Zaragoza.....	88
Tabla 29: Niveles asignados a la dimensión legal, Ayto. Zaragoza.....	88

Tabla 30: Niveles asignados a la dimensión organizativa, Ayto. Zaragoza.....	89
Tabla 31: Niveles asignados a la dimensión medición, Ayto. Zaragoza.....	89
Tabla 32: Niveles asignados a la dimensión disponibilidad, Ayto. Zaragoza.....	90
Tabla 33: Niveles asignados a la dimensión acceso, Ayto. Zaragoza.....	90
Tabla 34: Niveles asignados a la dimensión calidad de datos, Ayto. Zaragoza.....	91
Tabla 35: Niveles asignados a la dimensión actualización, Ayto. Zaragoza.....	91
Tabla 36: Niveles asignados a la dimensión reutilización, Ayto. Zaragoza.....	92
Tabla 37: Niveles asignados a la dimensión participación y colaboración, Ayto. Zaragoza.....	92
Tabla 38: Niveles asignados a la dimensión estratégica, Ayto. Barcelona.....	94
Tabla 39: Niveles asignados a la dimensión legal, Ayto. Barcelona.....	94
Tabla 40: Niveles asignados a la dimensión organizativa, Ayto. Barcelona.....	95
Tabla 41: Niveles asignados a la dimensión medición, Ayto. Barcelona.....	95
Tabla 42: Niveles asignados a la dimensión disponibilidad, Ayto. Barcelona.....	96
Tabla 43: Niveles asignados a la dimensión acceso, Ayto. Barcelona.....	96
Tabla 44: Niveles asignados a la dimensión calidad de datos, Ayto. Barcelona.....	97
Tabla 45: Niveles asignados a la dimensión actualización, Ayto. Barcelona.....	97
Tabla 46: Niveles asignados a la dimensión reutilización, Ayto. Barcelona.....	98
Tabla 47: Niveles asignados a la dimensión participación y colaboración, Ayto. Barcelona.....	98
Tabla 48: Resultados de aplicación de la norma en los Ayuntamiento de Madrid, Zaragoza y Barcelona .....	99
Tabla 49: Cálculo del indicador de datos abiertos, UNE 178301:2015.....	100
Tabla 50: Fases y actividades de la metodología propuesta.....	111
Tabla 51: Actividades y artefactos de la metodología propuesta.....	123

# Índice de abreviaturas

UE	Unión Europea
CE	Comisión Europea
CKAN	Comprehensive Knowledge Archive Network
AENOR	Asociación Española de Normalización y Certificación
CTN	Comité Técnico de Normalización
CIP	Catálogo de Información Pública
MELODA	MEtric for reLeasing Open Data
UCM	Universidad Complutense de Madrid
API	Application Programming Interfaces
GODI	Global Open Data Index
NTI	Norma Técnica de Interoperabilidad
RISP	Reutilización de la Información en el Sector Público
ONTSI Información	Observatorio Nacional de las Telecomunicaciones y de la Sociedad de la Información
OGP	Open Government PartnerShip
INE	Instituto Nacional de Estadística



# CAPÍTULO 1: INTRODUCCIÓN

Los datos abiertos ofrecen beneficios significativos a la sociedad y a la economía, promueven la democracia y la responsabilidad de los gobiernos públicos con eficacia y eficiencia, fomentan la transparencia, la participación y la colaboración de los ciudadanos. Promueven oportunidades para la innovación y creación de nuevos productos y servicios. Existen muchas razones, motivaciones, ventajas y beneficios para implementar un sistema de datos abiertos.

Este trabajo estudia las buenas prácticas, estándares y principios que las administraciones públicas deben considerar en el momento de realizar iniciativas de datos abiertos, con el fin de asegurar la calidad, accesibilidad y disponibilidad de los datos, conseguir el impacto deseado de la apertura de datos y la manera de cómo éstas iniciativas a través de la libre accesibilidad, reutilización y redistribución de los datos contribuyen a mejorar la calidad de vida y el desarrollo sostenible de una ciudad (smart city). La gran transformación tecnológica y digital que estamos experimentando, la variedad y velocidad con que se generan los datos nos obliga a incrementar esfuerzos para procesar la información y ponerla a disposición del ciudadano, muchas veces inclusive en tiempo real. En ese contexto, nos referiremos no solamente a la digitalización, sino también a la datificación de la información pública.

## 1.1. MOTIVACIÓN

El open data es una filosofía y una práctica cuya idea se basa en que determinados tipos de datos estén disponibles de forma libre para todo el mundo, sin restricciones de derecho de autor, de patentes o mecanismos de control. Tiene una ética similar a otros movimientos como el de software libre y el de código abierto. Es una iniciativa relativamente nueva, siendo los países de Europa y Estados Unidos los primeros en implementar políticas de datos abiertos. Hay países que aún están lejos de estas iniciativas, como América Latina. Las razones por las que no se implementan políticas de datos abiertos varían según el país. Algunas de las razones pueden ser: las infraestructuras tecnológicas, bases legales, razones culturales o sociales, económicas o hasta la misma voluntad política de los gobernantes. Sin embargo, los beneficios que los datos abiertos brindan al desarrollo social y económico de la sociedad son muy importantes.

Siempre he estado motivado a realizar un trabajo que aportara valor a mi país y en el campo de los datos abiertos Perú está aún en fases iniciales. A través de este trabajo se pretende promover el uso adecuado de las buenas prácticas, normas, métricas y tecnologías para que las administraciones públicas abran sus datos y los pongan a disposición de los ciudadanos y de esa manera acortar las distancias que actualmente los separa e incrementar la confianza entre ellos. A través de la libre reutilización de los datos generar valor creando nuevos productos y servicios que mejoren la calidad de vida de los ciudadanos.

## **1.2. OBJETIVO DE LA INVESTIGACIÓN**

El objetivo general de éste trabajo es realizar un estudio actual de las iniciativas de datos abiertos, identificar y aplicar métricas y estándares a conjuntos de datos y a portales de datos abiertos respectivamente y proponer recomendaciones mediante el diseño de una metodología para publicación de datos abiertos. En relación con el objetivo general, se formulan los siguientes objetivos específicos:

- ✓ Explorar los principios e iniciativas de datos abiertos a nivel global y local.
- ✓ Analizar las dos tecnologías de código abierto más relevantes para implementar un portal de datos abiertos e indicar una guía de instalación y configuración para cada una.
- ✓ Analizar y evaluar los portales de datos abiertos de los Ayuntamientos de Madrid, Zaragoza y Barcelona en el marco de la norma UNE 178301:2015 de AENOR.
- ✓ Proponer el diseño de una metodología para la publicación de datos abiertos que se adecúe a entidades públicas que aún no han iniciado o están comenzando a implementar datos abiertos.

Para cumplir con estos objetivos me he contactado con el Profesor Juan Pavón quien me ha dirigido hasta Victoria López y ella hasta Guadalupe Miñana quien finalmente ha sido mi tutora por su experiencia en esta temática.

Guadalupe participa en el proyecto OjoalData100, el cual consiste en determinar los 100 conjuntos de datos más relevantes que deben ser incluidos en una iniciativa de datos abiertos para ser considerada una smart city, tomando como referencia a la ciudad de Madrid. Experiencia que le ha permitido, junto a otros proyectos relacionados con el tema en

los cuales también ha participado, expandir sus conocimientos en datos abiertos y contribuir a la adecuada realización de este trabajo.

El grupo de investigación G-Tec UCM, del cual la directora de este trabajo es integrante, ha realizado un trabajo preliminar denominado *“Sistema de Open Data – Aplicaciones de Medio Ambiente”*[1] durante los años 2012 y 2013 con datos del área de medio ambiente del Ayuntamiento de Madrid. El trabajo que ahora presento hace referencia a dicho trabajo [1] como a un ayuntamiento con primeras iniciativas de datos abiertos.

### **1.3. RELACIÓN DEL TRABAJO CON LAS ASIGNATURAS DEL MÁSTER**

Todas las asignaturas del máster han contribuido a incrementar mis habilidades y capacidades profesionales que seguro me serán de mucha utilidad cuando las requiera. Sin embargo, para la elaboración del presente trabajo he recurrido a lo aprendido en las asignaturas de *“Auditoria, calidad y fiabilidad informáticas”* impartida por el Profesor Adrián Riesco Rodríguez, con quien realicé una auditoría a un sistema de registro de control de personal, implementado por la organización Soluciones Prácticas con el apoyo de la Agencia Española de Cooperación Internacional para el Desarrollo (AECID), la Cooperación de la Comunidad de Madrid y Castilla la Mancha, obteniendo una calificación de notable; las competencias adquiridas en esta asignatura me han permitido aplicar los principios y buenas prácticas de auditoría. La asignatura *“Implantación corporativa de tecnologías, servicios y sistemas informáticos”* impartida por el Profesor Manuel Prieto Masías, con quien realicé un trabajo relacionado con la infraestructura de la red de datos de la UCM, obteniendo una calificación de notable; las competencias adquiridas en esta asignatura me han permitido realizar un análisis de la tecnología adecuada para la implementación de un sistema de datos abiertos. La asignatura *“Dirección y gestión de proyectos y sistemas informáticos”* impartida por el Profesor Federico Peinado, con quien realicé un trabajo de implementación de un proyecto de TI en la nube; las competencias adquiridas en esta asignatura me han permitido aplicar los principios, estándares y buenas prácticas de gestión de proyectos de TI en este caso para entidades públicas. La asignatura *“Gestión de datos y de la información”* impartida por el profesor Enrique Martín, con quien realicé tareas a lo largo del curso sobre gestión de grandes volúmenes de datos relacionales y no relacionales, su almacenamiento y procesamiento. En fin, la sinergia del conocimiento adquirido en éstas y otras asignaturas del

máster me han sido de mucha utilidad a la hora de realizar éste trabajo y en particular a la hora de proponer la metodología para publicación de datos abiertos.

## **1.4. ESTRUCTURA DEL TRABAJO**

Para una mejor comprensión, este trabajo se ha dividido en 6 capítulos. En el capítulo 2 se describen los principales conceptos de datos abiertos, su implicación en las smart cities. Se realiza un estudio del estado actual de datos abiertos a nivel global y local. También se analiza las dos principales tecnologías de código abierto (CKAN y DKAN) para la publicación de datos abiertos, se instala, configura y se realizan pruebas en cada una.

En el capítulo 3 se estudia el contenido y los alcances de las métricas y normativas de datos abiertos, en este trabajo hablaremos específicamente de las métricas cinco estrellas de datos abiertos y Meloda, también veremos los conceptos de la norma UNE 178301:2015.

En el capítulo 4 se realiza una aplicación de la Norma UNE 178301:2015 a los portales de datos abiertos de los ayuntamientos de Madrid, Zaragoza y Barcelona y a un trabajo inicial de datos abiertos que se le conoce como la versión beta.

Llegados a este punto, y como resultado de lo aprendido en todos los capítulos anteriores, en el capítulo 5 se propone el diseño de una metodología para la publicación de datos abiertos. Amigable y adecuada para organizaciones públicas que aún no han propuesto sus primeras iniciativas o están comenzando a implementar datos abiertos.

Finalmente, en el capítulo 6 se presentan las conclusiones y se mencionan las posibles líneas de trabajo futuro, con el fin de enriquecer y continuar con la investigación realizada.

Además, se adjuntan como anexos las guías de instalación y configuración de las dos principales tecnologías utilizadas para publicación de datos abiertos, mencionadas en el capítulo 2 de este trabajo.

# INTRODUCTION

Open data offer benefits significant to society and the economy, promote democracy and accountability of public governments, promoting transparency, participation and collaboration of citizens. Promote opportunities for innovation and creating new products and services. There are many reasons, motivations, advantages and benefits to implement a system of open data.

This work refers to best practices, standards and principles that public governments should consider when making initiatives open data, in order to ensure the quality, accessibility and availability of data, achieve the desired impact of opening data and how these initiatives through the free accessibility, reuse and redistribution of data contributing to improve the quality of life and sustainable development of a city (smart city). The great technological and digital transformation we are experienced, variety and speed with which the data are generated forces us to accelerate efforts to process information and make it available to citizens, sometimes even in real time. In this context, we refer not only to digitization, but also to the datificación of public information.

## **1.1. Motivation**

Open data is a philosophy and a practice whose idea is based on certain types of data that are freely available to everyone, without restrictions of copyright, patent or control mechanisms. It is similar ethics to other movements such as free software and open source. It is a relatively new initiative, with European countries and the United States the first to implement open data policies. There are countries that are far from these initiatives, such as Latin America. The reasons why not open data policies are implemented vary by country. Some reasons include: the technological infrastructure, legal bases, cultural or social, economic or even the same political will of the rulers reasons. However, the benefits of open data provide the social and economic development of society are very important.

I have always been motivated to perform work that add value to my country and in the field of open data Peru is still in early stages. Through this work it is to promote the proper use of best practices, standards, metrics and technologies for public administrations to open their data and make available to citizens and thus bridge the gap that currently separates and increase trust between them. Through the free reuse of data to generate value creating new products and services that improve the quality of life of citizens.

## **1.2. Objective research**

The overall objective of this work is an ongoing study of initiatives open data, identify and implement metrics and standards to datasets and portals open data respectively and suggest possible recommendations by designing a methodology for publishing open data. In relation to the overall objective, the following specific objectives are formulated:

- Explore the principles and initiatives open data representative entities.
- Analyze and evaluate open data portals of the municipalities of Madrid, Zaragoza and Barcelona in the framework of the UNE 178301: 2015 AENOR.
- Analyze the two technologies most relevant open-source publishing open data and indicate a guide installation and configuration for each.
- Propose the design of a methodology for publishing open data that suits public entities that have not started or are starting to implement open data.

To meet these objectives I have contacted Professor Juan Pavon who has directed me to Victoria Lopez and her until finally Guadalupe Miñana who has been my mentor for their expertise in this field.

Guadalupe OjoalData100 participates in the Project, which is to determinate the 100 most relevant datasets to be included in an initiative open data to be considered a Smart city, with reference to the city of Madrid. Experience has

enabled him, along with other related topic which has also participated projects, expand their knowledge in open data and contribute to the conduct of this work.

The research group G-Tec UCM, which the director of this work is a member, has done preliminary work called "Sistemas de Open Data – Aplicaciones de Medio Ambiente" [1] during the years 2012 and 2013 with área data environment of the Municipality of Madrid. The work now performed to this work [1] as a municipality with first open data initiatives.

### **1.3. Working relationship with the subjects of the master**

All subjects of the master have contributed to increasing my abilities and skills that will surely be very useful to me when needed. However, for the preparation of this work I have resorted to what they learned in the subjects of "Software auditing and quality assurance" given by Professor Adrian Riesco Rodriguez, with whom I conducted an audit of a registration system of control of personnel, implemented for organization Soluciones Prácticas with the support of the Agencia Española de Cooperación Internacional para el Desarrollo (AECID), the Cooperación de la Comunidad de Madrid and Castilla la Mancha, earning a rating of remarkable; the skills acquired in this course allowed me to apply the principles and good audit practice. The course "Integration of computer services, systems and technologies" given by Professor Manuel Prieto Masias, with whom I did a work related to network infrastructure UCM data, obtaining a rating of remarkable; the skills acquired in this course allowed me to conduct an analysis of appropriate technology for the implementation of a system of open data. The subject "Management and project management and computer systems" given by Professor Federico Peinado, with whom I did a work of implementing an IT project in the cloud; the skills acquired in this course allowed me to apply the principles, standards and best practices for IT project management in this case to public entities. The subject "Data and Information Management Systems" given by Professor Enrique Martin, who performed tasks throughout the course on managing large volumes of relational and non-relational data, storage and processing. In short, the synergy of knowledge acquired in these and other

subjects of the Masters have been very useful to me when performing this work and especially when proposing the methodology for publishing open data.

## **1.4. Work structure**

This work, for a better understanding of it, has been divided into 6 chapters. In Chapter 2 the main concepts of open data are described, their implication in the Smart cities. A study of the current state of open global and local data is performed. As well, analysis of the two major open source technologies (CKAN and DKAN) for publishing open data is made, is installed, configured and testing is performed.

In Chapter 3 the content and scope of the metrics and standards of open data, this paper is studied specifically discuss the metric five star open data and Meloda, we will see the concepts of AENOR 178301:2015 standard.

In chapter 4 an application of the Standard is done UNE 178301:2015 to open data portals of municipalities of Madrid, Zaragoza and Barcelona and initial work of open data that is known as case Beta.

At this point, as a result of what was learned in all previous chapters, in chapter 5 is proposed the design of a methodology for publishing open data. Friendly and suitable for public organizations that have not yet taken the first initiatives or are beginning to implement open data.

Finally, in Chapter 6 presents conclusions and possible lines of future work, in order to enrich and continue the investigation mentioned.

In addition, they attached as annexes guides installation and configuration of the two main technologies used for publishing open data, used in Chapter 2 of this work.

## CAPÍTULO 2: OPEN DATA

“La teoría democrática sostiene que la información sobre lo público ha de ser accesible de modo igualitario para todos los miembros del público.”

Philip Meyer, 1993

En este capítulo trataremos sobre los principales conceptos, algunos principios que lo rigen y el estado actual del open data descrito mediante avances en algunos organismos internacionales, nacionales y locales. También hablaremos sobre la importancia del open data en las Smart cities como herramienta de desarrollo sostenible.

### 2.2. ¿QUÉ ES UN OPEN DATA?

El “open data” no tiene una definición formal, sin embargo, después de realizar indagaciones en organismos promotores de este movimiento, como la Open Knowledge Foundation, el Open Data Institute, el Opendefinition, entre otros, se puede resumir en lo siguiente:

**“El open data es una filosofía en la que una administración pública o privada *cede o retorna* a los ciudadanos o a otras administraciones la información pública en formatos adecuados y sin restricciones administrativas, legales ni sociales; permitiéndoles total libertad para utilizar, reutilizar y redistribuir dicha información en cualquier momento y lugar sin tener que pagar por ello. El objetivo es generar, a partir de dichos datos abiertos, nuevos productos y servicios mejorando así la calidad de vida de los ciudadanos”**

La información puede ser generada por la misma población o por otras fuentes (captura de datos medioambientales, por ejemplo). La información es almacenada y procesada por la administración y es devuelta a los ciudadanos directamente o a través de Infomediarios. Los

datos abiertos también pueden ser utilizados por otros organismos, lo que se conoce como Linked Data, tal como se observa en la figura 1.

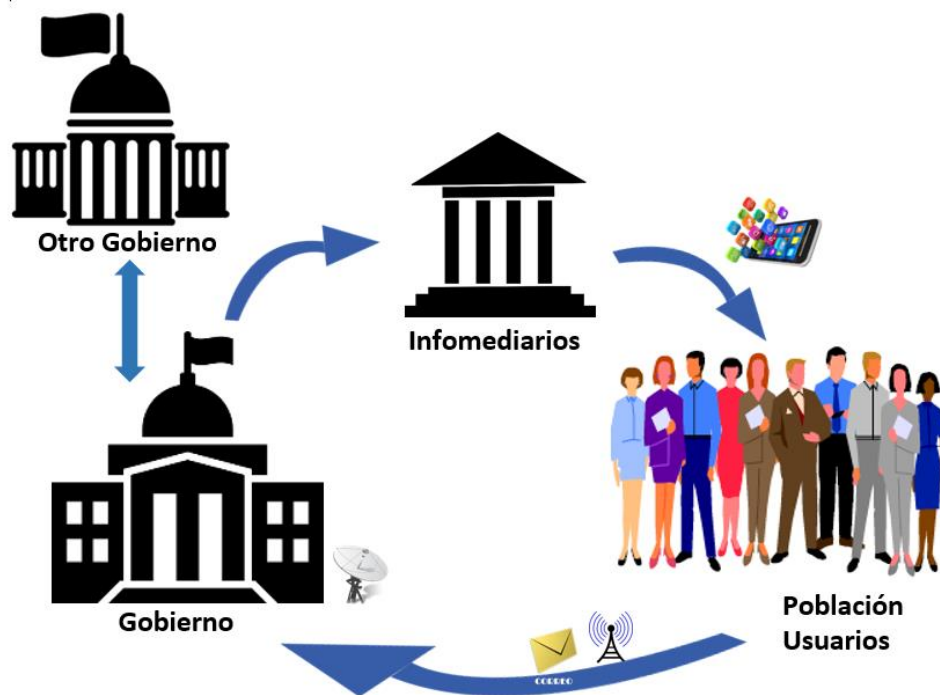


Figura 1: Ciclo de los datos abiertos (Elaboración propia)

Para que los datos publicados sean considerados como “datos abiertos” deben cumplir las siguientes características básicas:

- **Disponibilidad y accesibilidad.** La información debe ser accesible en Internet en forma de un catálogo de datos, que sea fácil de acceder y encontrar. Disponible la mayor cantidad de tiempo posible.
- **Reutilización y redistribución.** Mediante la reutilización de datos se puede generar valor, creando nuevos productos y servicios a partir de los datos. Para que la información sea reutilizable, no deben existir condiciones de uso que lo impidan, se debe permitir que la información sea tratada y modificada así como que sea posible redistribuir la nueva información que se elabore a partir de la información original. Asimismo, se facilitará la existencia de los datos en formatos digitales lo más estandarizados y reutilizables posible.

- **Accesibilidad y usabilidad universal.** Los datos deben estar disponibles para todas las personas que deseen utilizarlos y las condiciones en las que se liberan no deben establecer restricciones de uso en función del colectivo que lo vaya a utilizar o del uso que se le vaya a dar. Por ejemplo, no se consideraría como “datos abiertos” a aquellos datos cuyo uso solo se permite para algún fin específico. Se asegura la accesibilidad a los datos si no se ponen restricciones económicas, físicas o legales para acceder a ellos.

Más adelante comprenderemos mejor la importancia de estos conceptos, ya que la norma y las métricas que estudiaremos, se basan en estos conceptos para definir sus estándares y/o criterios de evaluación.

Las iniciativas de datos abiertos suponen también un cambio de paradigma respecto a la forma habitual de gestionar la información y los datos en las administraciones, por lo que será necesario contar también con una estrategia para vencer la habitual resistencia al cambio que nos encontraremos en la mayoría de los casos.

Una administración pública que es resistente a una comunicación abierta y transparente a los ciudadanos, que no quiere entender que el acceso a los datos públicos es un derecho por parte de los ciudadanos, o una institución pública que no tiene las facilidades de acceso a la tecnología, es un tanto difícil que tenga éxito en la apertura de datos. Para aplicar políticas de open data primero hay que realizar cambios de consciencia o mentalidad en los funcionarios, para conseguir que haya iniciativa y anhelo de hacer accesibles los datos.

El open data contribuye esencialmente al desarrollo de una sociedad y se basa en tres ejes principales, que son: colaboración, participación y transparencia. Esto hace que una administración pública esté más cerca de los ciudadanos y se recobre la confianza en ella, necesidad latente que los ciudadanos demandan permanentemente.

Algunas instituciones han elaborado una serie de principios que debe cumplir todo open data. A continuación se describen los principios que han considerado los principales organismos referentes del open data:

### ❖ **Sunlight Foundation**

La Sunlight Foundation es una organización independiente sin ánimo de lucro que aboga a nivel internacional por un gobierno abierto y utiliza la tecnología para hacer que el gobierno sea abierto, responsable y accesible para todos. Según esta organización un open data debe cumplir con los siguientes principios[2]:

- **Completitud.** Deben estar disponibles todos los datos públicos. Se entiende por datos públicos todos aquellos que no están sometidos a limitaciones debido a protección de la privacidad, la seguridad o determinados privilegios.
- **Datos primarios.** Se deben publicar los datos con el mayor nivel de granularidad posible, no en forma agregada o modificada.
- **Oportunidad y puntualidad.** Se deben publicar los datos tan rápidamente como sea posible para preservar el valor de los mismos.
- **Facilidad de acceso electrónico y físico.** Los datos deben estar disponibles para el mayor número de personas y para el mayor número de propósitos posibles.
- **Procesable informáticamente.** Los datos deben estar razonablemente estructurados en formatos que permitan el procesamiento automático.
- **No discriminación y accesibilidad.** Los datos deben estar disponibles para todo el mundo, sin impedimentos ni requisitos adicionales y sin la necesidad de registrarse para acceder a ellos.

- **Estándares abiertos.** Los datos deben estar en formatos sobre los que ninguna entidad tenga derechos exclusivos (deben publicarse en estándares abiertos o de uso común gratuitos).
- **Licenciamiento gratuito.** Los datos no deben estar sujetos a ningún tipo de copyright, patente, marcas o secretos comerciales. Se pueden imponer condiciones razonables de privacidad, seguridad y privilegios.
- **Permanencia.** Los datos publicados deben permanecer on-line con los mecanismos adecuados de control de versiones.
- **Coste del uso.** Los gobiernos no deben cobrar por usar los datos que publican.

## ❖ **Open Knowledge Foundation**

La Open Knowledge[3] es una red mundial sin ánimo de lucro de personas apasionadas por la apertura, mediante la promoción, la tecnología y la capacitación para desbloquear la información y permitir a la gente trabajar con él para crear y compartir conocimientos. Esta organización considera los siguientes principios de open data:

- Existencia del conjunto de datos.
- Existencia en formato digital.
- Disponibilidad para público en general.
- Disponibles de manera gratuita.
- Accesibles desde internet, con una URL.
- Disponibilidad de los datos en un formato procesable informáticamente (PDF no se considera procesable informáticamente).
- Posibilidad para descargar todos los datos en bloque.
- Posibilidad para acceder, reutilizar y redistribuir los datos de acuerdo con la “open definition”.
- Publicación de los datos tan pronto como están disponibles.

## ❖ **Open Data Charter**

En 2013, los miembros del G-8 firmaron un acuerdo en el que se comprometieron a adoptar políticas de open data en sus respectivos países y elaboraron un documento en el que definieron los siguientes cinco principios[4]:

### ➤ **Principio 1: Datos abiertos por defecto.**

- a. Libre acceso y posterior reutilización son de gran valor para la sociedad y la economía.
- b. Orientar los gobiernos hacia la apertura de datos por defecto.
- c. Respeto a la legislación nacional e internacional relativas a la propiedad intelectual e información de identificación personal y sensible.

### ➤ **Principio 2: Calidad y cantidad de los datos publicados**

- a. Los gobiernos y el sector público tienen grandes cantidades de información que pueden ser de interés para los ciudadanos.
- b. Preparar los datos de alta calidad puede llevar tiempo, por lo que, se consultará a sí mismo, autoridades nacionales y usuarios potenciales para identificar y priorizar los datos a ser publicados para su liberación y mejora.
- c. Publicar los datos con prontitud, completitud y exactitud. Los datos se publicarán en su formato original y con el mayor nivel de granularidad posible.
- d. Describir los datos de manera que sean fácilmente comprensibles de modo que pueda ser entendido por todos.
- e. Establecer procedimientos para recibir el “feedback” de los usuarios de los datos.

### ➤ **Principio 3: Utilizables por todos**

- a. Publicar los datos en formatos abiertos siempre que sea posible.
- b. Los datos deben estar disponibles en forma gratuita con el fin de fomentar su uso más extendido.

- c. Publicar los datos sin barreras burocráticas o administrativas, sin requisitos de registro.
- d. Liberar tantos datos como sea posible y en varios formatos, de manera que puedan ser procesados por máquinas y sean comprensibles para las personas.

➤ **Principio 4: Liberación de datos para mejorar la gobernabilidad**

- a. La liberación de la información fortalece la institucionalidad democrática y fomenta una mejor formulación de políticas para satisfacer las necesidades de los ciudadanos.
- b. Compartir conocimientos técnicos y experiencias con otros países u organismos para que todos puedan aprovechar los beneficios de los datos abiertos.
- c. Documentar y compartir las “prácticas en materia de open data” para fomentar la transparencia durante la colección de datos, estándares y procesos de publicación.

➤ **Principio 5: Liberación de datos para la innovación.**

- a. Publicar los datos con licencias abiertas para que todo el mundo pueda usar, modificar y compartir los datos.
- b. Garantizar que los datos sean legibles informáticamente en bloques, para facilitar su incorporación automática a otros sistemas.
- c. Publicar datos de manera que puedan ser directamente integrados en otras aplicaciones (usar “application programming interfaces” – API).
- d. Animar el uso innovador de los datos, mediante la organización de concursos, premios y otras actividades.

El documento que describe los estos principios es libre acceso [4] para que otros países u organizaciones lo analicen y también puedan implementarlo, contiene un anexo técnico en el que se incluyen las mejores prácticas y acciones colectivas con la finalidad de dar cumplimiento a cada uno de los principios anteriormente mencionados.

## ❖ Declaración de Open Data: Global Open Data Initiative

Esta es una declaración [5] desde la perspectiva de la sociedad, en respuesta a la inminente importancia que tienen los datos abiertos. Muestra que el acceso a los datos, en una sociedad democrática, es un derecho y hace un llamamiento a la ciudadanía a tener participación activa en las iniciativas de datos abiertos. Su objetivo es hacer que los datos del gobierno estén a disposición de todos, para cualquier persona, en cualquier lugar, para descargar, utilizar, reutilizar, y redistribuir sin cargo alguno para cualquier propósito.

Hacen un planteamiento a los gobiernos que se deben tener en cuenta siete medidas importantes al momento de la apertura de datos:

- Datos abiertos por defecto.
- Proceso centrado en las personas o poner a los usuarios en primer lugar.
- Proporcionar los datos sin coste de acceso.
- Poner la rendición de cuentas como núcleo.
- Invertir en capacitación.
- Mejorar la calidad de datos oficiales.
- Promulgar reformas políticas y legales para crear un gobierno más abierto, transparente y participativo.

De todos los principios vistos anteriormente podemos observar que existen muchas iniciativas de datos abiertos a nivel mundial. Ahora que ya entendemos los conceptos y principios fundamentales del open data, veremos a continuación los principales avances que existen a nivel global y local en las iniciativas de datos abiertos.

## 2.3. ESTADO ACTUAL DEL OPEN DATA

Hasta la fecha existen en el mundo muchas iniciativas de datos abiertos, cada una de ellas promovida y regulada por leyes y normas de diferente ámbito o alcance. Sin embargo, el registro más antiguo de la iniciativa legal de datos abiertos posiblemente sea la “Ley para la Libertad de Prensa y del Derecho de Acceso a las Actas Públicas” del año 1776 propuesta en Suecia por Anders Chydenius, un sacerdote sueco-finlandés[6]. Los datos abiertos son ahora un recurso imprescindible para construir instituciones responsables, transparentes y efectivas.

De acuerdo con el marco común de evaluación de datos abiertos[7] existen cuatro maneras diferentes para evaluar los datos abiertos: el contexto, los datos, el uso y el impacto.

Esta sección está dividida en dos partes. En primer lugar, se muestra los estudios realizados por las entidades de datos abiertos más representativas a nivel mundial:

- **Global Open Data Index (GODI)** [8], el Índice Global de Datos Abiertos, en español, es uno de los proyectos principales de la Open Knowledge International. Recoge y presenta información sobre el estado actual de datos abiertos en 122 países. GODI parte del supuesto de que el gobierno nacional tiene la responsabilidad de garantizar la publicación de ciertos datos clave, incluso si los datos no los produce el gobierno o son gestionados por un tercero. Asimismo, para países con diferentes estructuras de gobierno (regiones, estados, gobiernos locales) además de que los diferentes gobiernos pueden recopilar y publicar datos para regiones geográficas distintas, es responsabilidad del gobierno nacional ser un agregador de datos y tener un único punto de acceso para todos los subconjuntos de datos para que los usuarios tengan una manera fácil de acceder y utilizar los datos.

La evaluación la realiza la misma entidad en concordancia con los principios descritos en el apartado anterior y lo hace en base a la

evaluación de trece conjuntos de datos (estadísticas nacionales, presupuesto del gobierno, gastos gubernamentales, legislación, resultados de la elecciones, mapa nacional, emisiones de contaminantes, registro de empresas, ubicación de conjuntos de datos, ofertas de contratación pública, calidad del agua, pronóstico del tiempo, propiedad de la tierra), estos conjuntos de datos se evalúan utilizando nueve preguntas, a cada pregunta se le da una ponderación o peso que entre las nueve preguntas suman 100, las preguntas que considera son[9]:

- **¿Existen los datos?** Los datos deben estar en cualquier forma (papel o digital, online u offline, etc). Si no se cumple este requisito, todas las demás preguntas no son respondidas. (Ponderación = 5).
- **¿Están los datos en formato digital?** Los datos deben estar almacenados de manera digital en ordenadores u otros dispositivos de almacenamiento digital. GODI ha llegado a la conclusión que los datos deben ser accesibles en línea. Por lo que ha eliminado esta pregunta para futuras ediciones. (Ponderación = 5).
- **¿Disponible públicamente?** No se refiere específicamente a que sea de libre acceso, pero sí a que todas las personas (del gobierno y fuera del gobierno) puedan acceder a los datos de alguna forma, aunque se tenga que pagar para ello o aunque los datos estén disponibles en formato de papel. Si existen limitaciones de solicitud para acceder a la información, no se considera pública. (Ponderación = 5).
- **¿Datos disponible de forma gratuita?** Se refiere a si los datos están disponibles de forma gratuita, y si hubiese algún cargo éste debe estar contemplado y comentado. (Ponderación = 15).

- **¿Están los datos disponibles en línea?** Se refiere a que los datos deben estar disponibles en línea desde una fuente oficial. (Ponderación = 5).
- **¿Son los datos legibles por máquina?** Los datos deben estar en un formato fácilmente procesable por un ordenador. En general, HTML y PDF no son legibles por máquina. (Ponderación = 15).
- **¿Disponibles a granel?** Si el conjunto de datos o dataset se puede descargar o acceder a él fácilmente. (Ponderación = 10).
- **¿Con licencia abierta?** Se debe indicar que las condiciones de uso o licencia permitan que cualquiera pueda utilizar, reutilizar y redistribuir los datos libremente (sujeto a lo más, a atribución). La licencia debe estar disponible, si no hay licencia los datos no están disponibles abiertamente. Las licencias abiertas se definen y enumeran en <http://opendefinition.org/licenses/>. (Ponderación = 30).
- **¿Se proporcionan los datos de manera oportuna?** Se refiere a si los datos están publicados en el mismo momento en que han sido actualizados. (Ponderación = 10).

El resultado se obtiene al sumar las puntuaciones de los trece conjuntos de datos y dividirlo por 1300, que es el resultado máximo que puede obtener un país, al resultado se lo aproxima al valor entero más próximo. Como se explica en la siguiente fórmula 1, extraída de la fuente[9]:

$$Puntuación = \frac{Suma(13 \text{ conjuntos de datos})}{1300} \quad (1)$$

En la figura 2 se muestran los resultados de este estudio. En la parte izquierda se listan los países según el orden de mérito obtenido, siendo los países que están en la parte superior los que tienen mejor calificación que los de la parte inferior. Y en el lado derecho se muestra la puntuación obtenida para cada país. La figura solo muestra los resultados para los

primeros 31 lugares, para ver la imagen completa (de los 122 países) puede consultar la fuente de la figura.

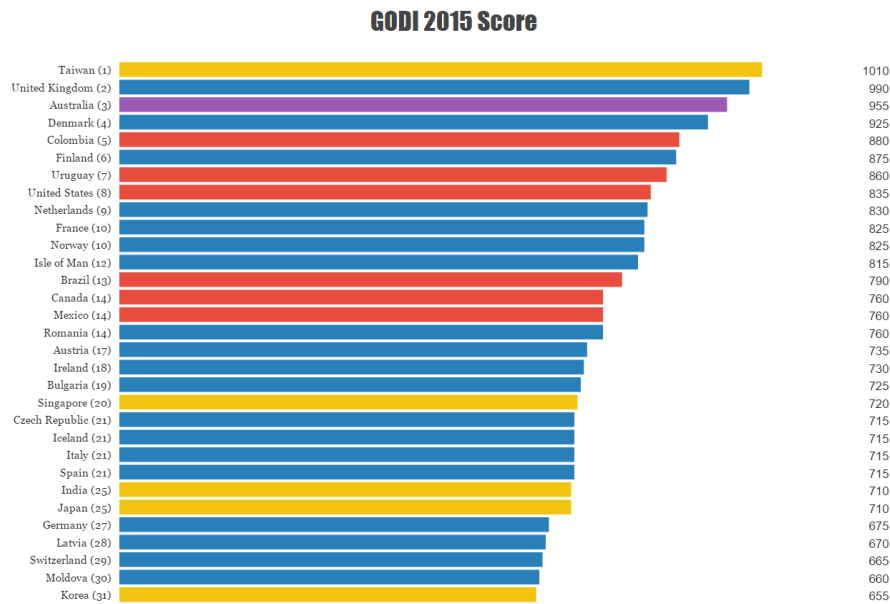


Figura 2. Ranking Global de Open Data - Open Knowledge Foundation (2015)  
Fuente: <http://2015.index.okfn.org/chart/>

➤ **Open Data Barometer**[10], es un estudio realizado a cargo de la World Wide Web Foundation. La tercera edición del open data barometer, publicada el 21 de abril del 2016, analiza el avance del estado de los datos abiertos gubernamentales en 92 países, al mismo tiempo que nos permite monitorizar y evaluar los logros internacionales en el cumplimiento de los Objetivos de Desarrollo Sostenible marcados por la ONU en el 2016. Este estudio se basa en tres parámetros: avance de las iniciativas, implementación de los programas e impacto económico, político y social.

En la figura 3 podemos ver el ranking de Open Data Barometer correspondiente al año 2015, donde en el eje X se colocan los países estudiados, ubicándose al inicio del eje los países con mayor ranking y al final los de menor ranking alcanzado. En el eje Y se colocan los niveles alcanzados (de 0 a 100) con una media de todos los países de 32.96.

# Ranking

## Open Data Barometer (2015)

Source: ODB Expert Survey 2015

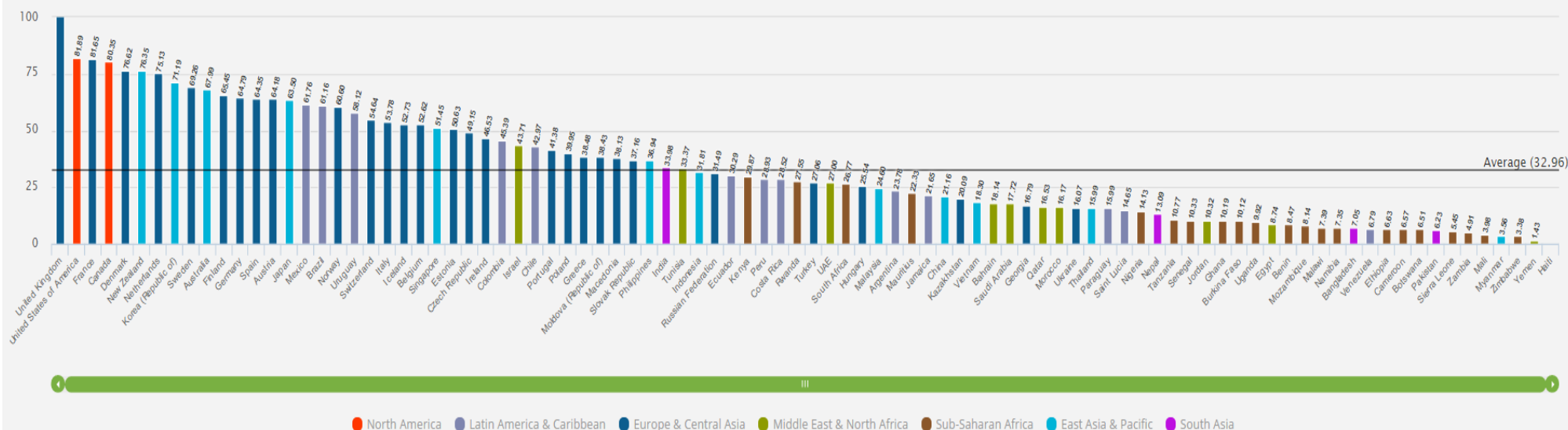


Figura 3: Ranking mundial de open data - Open Data Barometer (2015)

Fuente: [http://opendatabarometer.org/data-explorer/?\\_year=2015&indicator=ODB](http://opendatabarometer.org/data-explorer/?_year=2015&indicator=ODB)

- **Open Data Economy**, es un informe de la consultora Capgemini Consulting[11], publicado en febrero del 2013. Este informe analiza la situación del open data en 23 países, los mismos que fueron clasificados en tres grupos. Por un lado están los “países que marcan el camino”: Estados Unidos, Reino Unido, Francia, Canadá y Australia. Por otro lado, los “países seguidores”: España, Dinamarca, Irlanda, Italia, Noruega y por último los “países que acaban de comenzar”: Arabia Saudita, Austria, Emiratos Árabes Unidos, Estonia, Marruecos.

En la figura 4 se puede ver un resumen de este análisis. El eje horizontal indica el avance en disponibilidad de datos, colocándose los de mayor disponibilidad a la derecha y los de menos a la izquierda. En el eje vertical se indica el grado de usabilidad de los portales de datos abiertos, colocándose los portales con mayor usabilidad arriba y los de menos, abajo.

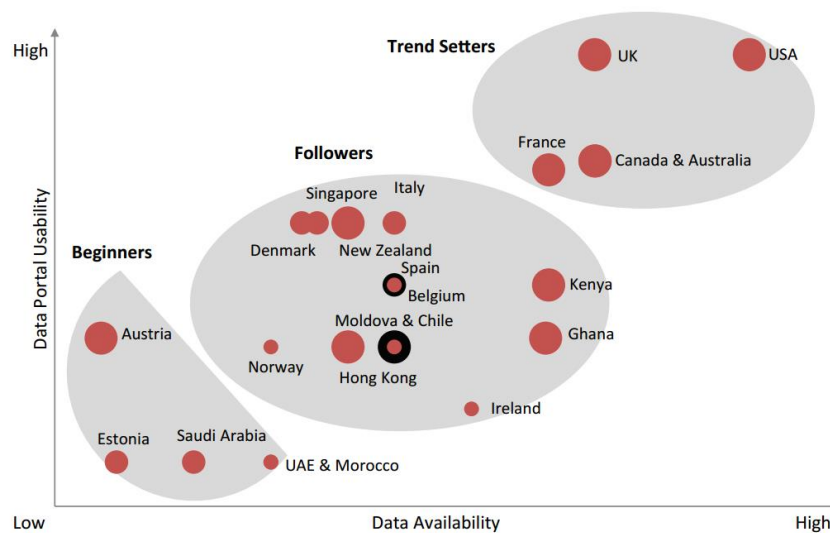


Figura 4: Clasificación de países con iniciativas y avances en open data - Open Data Economy (2012)

Fuente: [https://www.es.capgemini-consulting.com/resource-file-access/resource/pdf/opendata\\_pov\\_6feb.pdf](https://www.es.capgemini-consulting.com/resource-file-access/resource/pdf/opendata_pov_6feb.pdf)

- **Open Data Monitor**[12]: es un proyecto que se inició en noviembre del 2013[13] financiado dentro del Séptimo Programa Marco de

Investigación y Desarrollo de la Unión Europea. Permite integrar, procesar, analizar, visualizar, descargar y hacer comparativas a nivel europeo, nacional y local de catálogos de datos y conjunto de datos de toda la Unión Europea. Permitiendo de esta manera entender mejor la situación de datos abiertos en Europa.

La figura 5 muestra el resultado conglomerado de este estudio, según los cinco aspectos que considera: licencias abiertas, legibles por máquina, disponibilidad, completitud de datos y puntuación global de calidad. Siendo “*licencias abiertas*” el primer resultado porcentual de la parte superior que se observa en la barra vertical para cada país, “*legibles por máquina*” el segundo, “*disponibilidad*” el tercero, “*completitud de datos*” el cuarto y “*puntuación global de calidad*” el quinto, respectivamente.

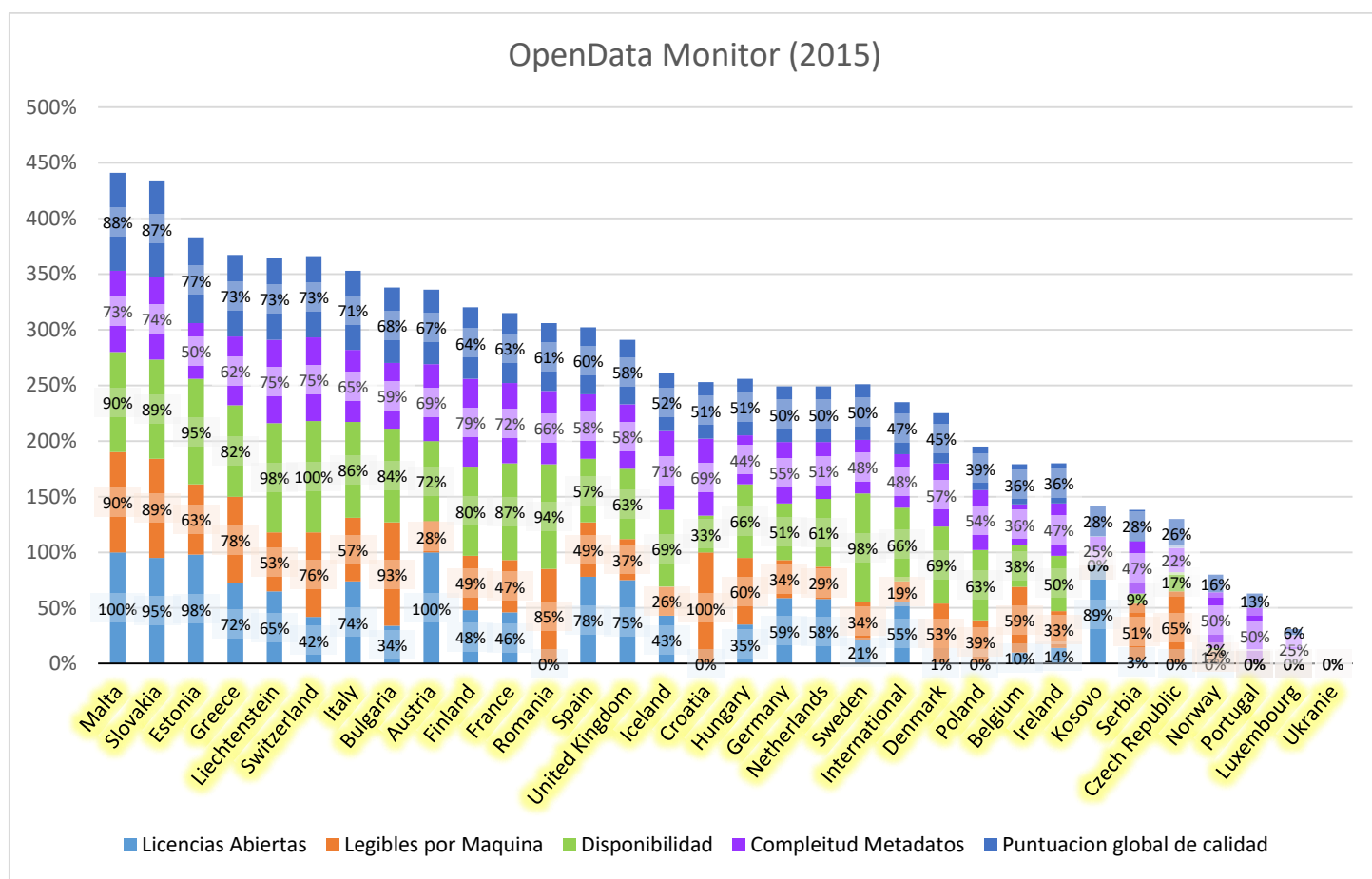


Figura 5: Evaluación de catálogos y conjuntos de datos abiertos en Europa - Open Data Monitor (2015)

Y en segundo lugar, después de haber analizado los estudios sobre iniciativas de datos abiertos por organismos de nivel mundial, como ejemplo muestro los portales de algunos de los países que tienen avances significativos en datos abiertos y del Banco Mundial por ser un organismo de alcance internacional y por la interacción gráfica de su portal.

### ➤ Estados Unidos

El portal de datos abiertos del gobierno de los Estados Unidos[14], contiene 181,949 conjuntos de datos a fecha 15/05/2016, como se muestra en figura 6. Está construido sobre dos aplicaciones ambas de código abierto, CKAN y WordPress y su desarrollo es de acceso público en GitHub con licencia GNU.

Existen muchas aplicaciones desarrolladas a partir de los datos abiertos de data.gov útiles para diferentes ámbitos como: educación, agricultura, comercio, salud, trabajo, transporte, etc. Son accesibles y disponibles y en su mayoría libres y no requieren de registro para utilizarlas.

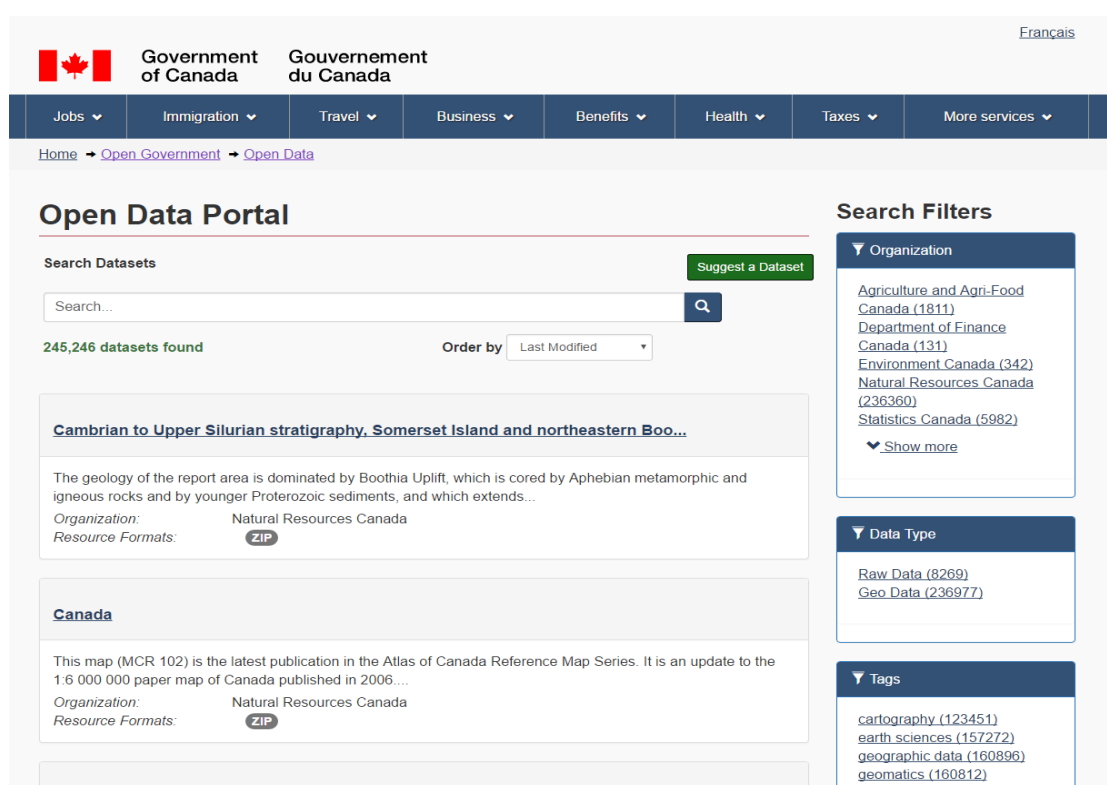


Figura 6: Portal de datos abiertos del gobierno de Estados Unidos

Fuente: <https://www.data.gov/>

## ➤ Canadá

Portal que se puso en marcha el 18 de diciembre del 2013 [15], tras cuatro años de esfuerzo por fusionar los enlaces de todos los departamentos y agencias en un único portal (ver figura 7). Contiene 245,246 conjuntos de datos a fecha 15/05/2016. Los datos más populares son los relacionados con el tiempo, inmigración, trabajo, viajes y salud. Según información publicada en el mismo portal, hasta la fecha y desde el 2013 la página [canada.ca](http://canada.ca) ha recibido más de 590 millones de visitas y más de 50000 comentarios, aumentando un 150% los usuarios que acceden al sitio mediante dispositivos móviles.



The screenshot shows the Open Data Portal of the Government of Canada. At the top, there is a navigation bar with the Government of Canada logo and the text 'Government of Canada' and 'Gouvernement du Canada'. Below this is a menu with categories: Jobs, Immigration, Travel, Business, Benefits, Health, Taxes, and More services. The main content area is titled 'Open Data Portal' and includes a search bar with the text 'Search Datasets' and a 'Suggest a Dataset' button. Below the search bar, it displays '245,246 datasets found' and an 'Order by' dropdown menu set to 'Last Modified'. Two dataset cards are shown, each with a title, a brief description, the organization (Natural Resources Canada), and the resource format (ZIP). The first card is titled 'Cambrian to Upper Silurian stratigraphy, Somerset Island and northeastern Boo...' and the second is titled 'Canada'. On the right side, there is a 'Search Filters' section with three categories: Organization (listing various departments and their counts), Data Type (listing Raw Data and Geo Data), and Tags (listing various tags and their counts).

Figura 7: Portal de datos abiertos del Gobierno de Canadá

Fuente: <http://open.canada.ca/data/en/dataset>

## ➤ Open data del Banco Mundial

Es una iniciativa del Banco Mundial [16] que tiene la intención de proporcionar a todos los usuarios el acceso a los datos que el Banco Mundial posee sobre agricultura, desarrollo rural, urbano y social, cambio climático, ciencia y tecnología, comercio, deuda externa, educación, pobreza, salud,

deuda externa y otros. Dicha información se muestra agrupada por países. El catálogo de datos, por ahora, solo está disponible en inglés[17].

Además del catálogo de datos, el Banco Mundial ofrece el banco de datos (DataBank)[18] que es una herramienta de análisis y visualización que contiene colecciones de datos de series cronológicas sobre diversos temas. Esta herramienta permite al usuario crear sus propias consultas, generar cuadros, gráficos y mapas y guardarlos fácilmente, incrustarlos y compartirlos. Los usuarios son libres de copiar, distribuir, adaptar, o incluir los datos en otros productos ya sea con fines comerciales o no comerciales, salvo los datos que son de terceros a quienes se debe pedir permiso de autorización para lo cual se incluye información de contacto del proveedor original de dichos contenidos. La figura 8 muestra el portal principal de datos abiertos del Banco Mundial.

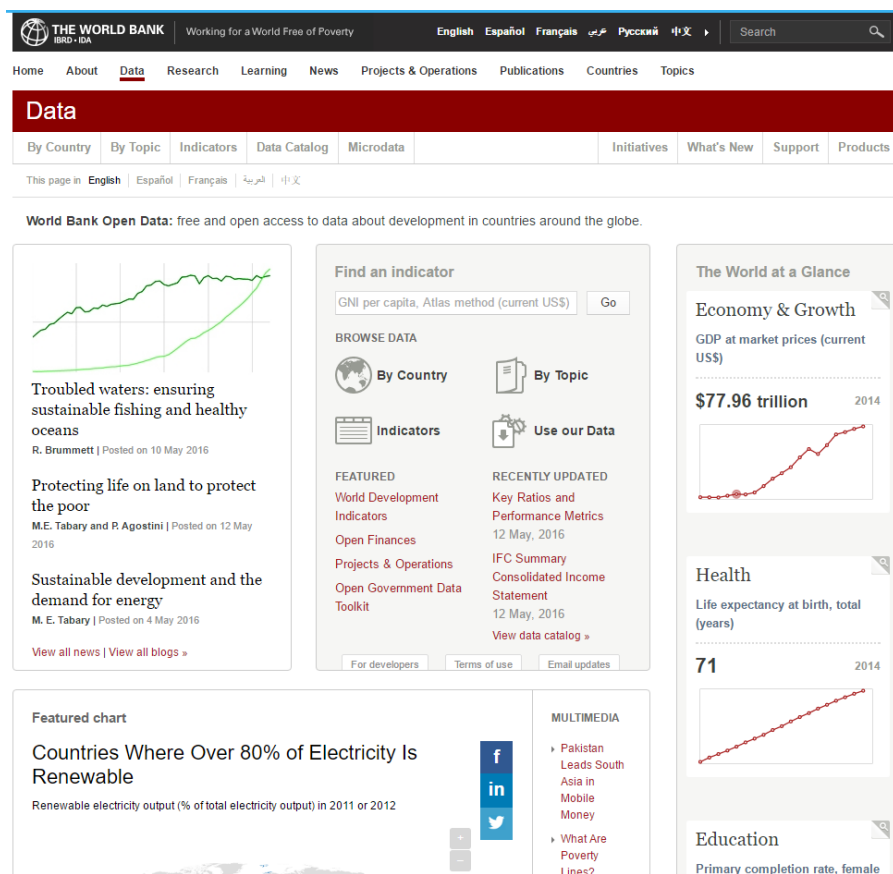


Figura 8: portal de datos abiertos del Banco Mundial.

## ➤ Unión Europea

Como este trabajo se presenta en la Universidad Complutense de Madrid, a continuación describo la situación actual del open data en la Unión Europea, para luego centrarme en España.

Respecto a la Unión Europea, el marco legal que da inicio a las políticas sobre apertura y reutilización de la información en el sector público es la Directiva 2003/98/CE, que aspira a garantizar la creación de un mercado interior para esta actividad basado en unas condiciones transparentes, equitativas, proporcionadas y no discriminatorias. A tal fin, se insiste en la necesidad de considerar la reutilización de la información del sector público como una importante fuente de creación de empleo y de dinamización del sector privado. Esta directiva se centra principalmente en dos objetivos: por un lado, facilitar la creación de nuevos servicios a nivel comunitario basados en la información del sector público y promover el desarrollo del sector de los contenidos digitales, con el fin último de ayudar al crecimiento económico y a la creación de empleo. Y por otro lado, fomentar el uso transfronterizo de la información del sector público por las empresas privadas con la finalidad de ofrecer productos y servicios de información de valor añadido.

A continuación (figura 9) muestro el Portal de datos abiertos de la Unión Europea[19], un punto único de acceso a gran variedad de datos de las instituciones y otros organismos de la Unión Europea (UE). Los datos se pueden utilizar y reutilizar libremente con fines comerciales o no comerciales, aunque existen algunos conjuntos de datos que tienen condiciones especiales de utilización, normalmente relacionadas con la protección de derechos de propiedad intelectual de terceros, y en cuyo caso se encuentra un enlace a dichas condiciones.

Con un acceso fácil y gratuito a los datos, el portal quiere fomentar su utilización innovadora y liberar su potencial económico. Entre sus objetivos también está impulsar la transparencia y la responsabilidad de las instituciones y otros organismos de la UE, se encuentra disponible en las 24

lenguas oficiales que existen en la UE. Los datos están disponibles en formatos que permiten la lectura humana y son también legibles por ordenador permitiendo su reutilización.



Figura 9: portal de datos abiertos de la Unión Europea

Uno de los componentes importantes para el crecimiento y expansión del open data es el índice de acceso a internet. En Europa, los países con mayor índice de acceso a internet de sus ciudadanos son los países del norte (Reino Unido, Holanda, Suecia, Dinamarca) y según los resultados de los estudios realizados por las entidades de datos abiertos más representativas a nivel mundial, vistas anteriormente, son los países que lideran las iniciativas de datos abiertos en Europa, respecto a los países del sur (Rumanía, Portugal, Lituania, Bulgaria y Grecia) quienes el índice de acceso a internet se encuentran muy por debajo de la media comunitaria, reflejándose también en los avances de datos abiertos.

Dentro de la Unión Europea cada país tiene sus propias normativas e iniciativas, a continuación nos centraremos concretamente en España.

El desarrollo del open data en España se ve influenciado por un lado por el crecimiento y uso de las tecnologías de la información y comunicaciones. Según el Observatorio Nacional de las Telecomunicaciones y de la Sociedad de la Información (ONTSI), hasta abril del 2016, el 79% de los hogares

españoles disponen de acceso a internet[20] con una tendencia a usar internet con mayor frecuencia por parte de los usuarios (el 91.9% ha accedido a internet el día anterior)[21]. Por otro lado, por la existencia del marco legal que permite a ciudadanos y empresas acceder a la información a través de formatos libres, abiertos, accesibles y reutilizables como la promulgación de la Ley 37/2007, de 16 de noviembre, sobre Reutilización de la Información del Sector Público, desarrollada mediante el Real Decreto 1495/2011 del 24 de octubre.

En España, la apertura de datos es un fenómeno esencial en la agenda digital, se busca obtener el mayor valor agregado del open data, tanto para los ciudadanos, al recibir un servicio público más eficiente, transparente, participativo y colaborativo; así como para las empresas, al aumentar su productividad y las administraciones permitiéndoles evaluar el verdadero impacto de las políticas públicas y avanzar en la innovación y modernización de los servicios públicos.

Como se ve en la Figura 10, existen portales de datos abiertos a distintos niveles: Nacional, Autonómico y Local. A continuación se muestra el mapa de las iniciativas de open data en España[22] hasta la fecha existen 115 iniciativas (34 de ámbito nacional, 22 de ámbito autonómico y 59 locales):



Figura 10: mapa de portales de datos abiertos en España

El Gobierno de España cuenta con un portal de datos abiertos[23] que es la plataforma encargada de organizar y gestionar en un único lugar los datos de los portales de datos abiertos de todo el país. La Norma UNE 178301:2015 recomienda que los conjuntos de datos publicados por organismos públicos también deben estar publicados en el Catálogo de Información Pública, es decir en el portal de datos abiertos del Gobierno de España ([www.datos.gob.es](http://www.datos.gob.es)) como se observa en la figura 11. La iniciativa encargada de llevar adelante este proyecto es 'Aporta', promovida por el Ministerio de Industria, Energía y Turismo a través de la Entidad Pública Empresarial Red.es y en colaboración con el Ministerio de Hacienda y Administraciones Públicas. Fue lanzada en el 2009 con el fin de promocionar la cultura de apertura de la información en España.



Figura 11: Portal de datos abiertos del Gobierno de España.

La iniciativa Aporta se desarrolla entorno a siete líneas de actuación (ver figura 12) enmarcadas en la Agenda Digital Española:

- i. **Difusión, sensibilización y formación:** a través del portal datos.gob.es y redes sociales como Twitter y LinkedIn. En la sección 'saber más' del portal existen materiales de formación.
- ii. **Catálogo nacional de datos:** disponible en el portal de datos abiertos.

- iii. **Estadísticas sectoriales:** el Observatorio Nacional de las Telecomunicaciones y de la Sociedad de la Información (ONTSI) elabora periódicamente el Estudio de caracterización del sector infomediario, que hace un análisis de la industria de la reutilización de los datos públicos.
- iv. **Ámbito legislativo y cooperación nacional:** principalmente, la Ley 37/2007 del 16 de noviembre, el Real Decreto 1495/2011 del 24 de octubre y la Norma Técnica de Interoperabilidad (NTI) de Reutilización de la Información en el Sector Público (RISP).
- v. **Colaboración en el ámbito público-privado:** diversas entidades públicas privadas han formado el “Foro de Colaboración Público-Privado sobre Reutilización de la Información del Sector Público” – Foro CPP-RISP.
- vi. **Cooperación internacional:** Aporta, participa y colabora con diversos proyectos internacionales como: Open Data Monitor, red temática Share-PSI 2.0 y el Open Data Support.
- vii. **Asesoramiento y soporte:** Aporta, ofrece un servicio de asesoría RISP a entidades públicas para resolver dudas de carácter operativo y técnico.



Figura 12: Líneas de actuación de la iniciativa Aporta

Como ejemplo de portal de datos de una ciudad se ha elegido Madrid[24] (figura 13). El portal está desarrollado sobre Vignette (gestor de contenidos de carácter propietario). Los conjuntos de datos pueden encontrarse en diferentes tipos de formatos, desde algunos más simples y estáticos como XLS o TXT, hasta otros más complejos como RDF, los cuales proporcionan una serie de URIs para facilitar el acceso a los datos y realizar consultas sobre ellos. Este portal se analizará con más detalle en el capítulo 3 tomando como referencia algunos conjuntos de datos y en el capítulo 4 analizando todo el portal de datos en su conjunto.



Figura 13. Portal de datos abiertos del Ayuntamiento de Madrid

## 2.4. CADENA DE VALOR DEL OPEN DATA

En esta sección se analizan los agentes relacionados con el ecosistema de datos abiertos en los procesos de apertura y reutilización de datos.

Según Alberto Abella[25] la cadena de valor del open data está dividida en cinco etapas: fuentes, esquemas técnicos/legales, tratamiento de datos, productos y servicios y usuarios. Como se puede ver en la figura 14.



Figura 14. Cadena de Valor de Open Data (Abella, 2011)

A continuación se describen con más detalle cada una de estas etapas:

- Fuentes de datos: son producidas y gestionadas por las administraciones públicas u organizaciones privadas.
- Esquemas legales y técnicos: los portales de datos abiertos deben asegurar mediante documentos legales el libre uso, distribución y reutilización de sus datos ya sea para fines comerciales y no comerciales, y mediante el uso de la tecnología deben asegurar que sus datos sean posibles de reutilizar, esto es, legibles y entendibles para el lenguaje humano y procesables por máquinas.
- El sector infomediario es “Un conjunto de empresas que generan productos y/o servicios para su comercialización a terceros, a partir de la información del sector público. Esto incluye, tanto a las empresas que se han creado con esta finalidad como a aquellas que pese a no tener ésta como única finalidad, poseen un área y/o departamento específico dedicado a la creación y comercialización de nuevos productos y/o servicios basados en la información del sector público ”[26].

La Asociación Multisectorial de la Información (ASEDIE), que reúne a las principales empresas infomediarias de España, nos muestra en su último estudio del 2015, que hasta esa fecha 623 empresas conformaban el sector infomediario en España. De éstas, la mayoría de las sociedades se encuentran en Madrid (42%) y Cataluña (21%)[27].

- **Productos y servicios**, a través y gracias a la libre reutilización de los datos el sector infomediario crea aplicaciones para diversos dispositivos, generando nuevos productos o servicios los cuales dan realmente valor a los datos.
- **Usuarios**: la sociedad civil, pueden ser de ámbito ciudadano, profesional, comercial o de investigación.

Es importante mencionar que los datos abiertos también alcanzan a las actividades de investigación, espacio conocido como open science[28] o ciencia abierta, en español, que facilita la reproductibilidad, la innovación académica e industrial y el acceso al conocimiento( Stodden, 2010). Aunque por un lado existen motivos como fondos de investigación (NSF, NIH, etc), editoriales de revistas (PLOSone), promociones y premios institucionales e integridad científica que incentivan la publicación de estos tipos de datos. Por otro lado, aún existen causas que limitan a los investigadores/científicos a publicar sus datos, y según Stodden 2010, algunas de estas causas son las siguientes:

- Tiempo en documentar y limpiar datos para publicar (54%)
- Lidar con preguntas de los usuarios de los datos (34%)
- No recibir atribución o citación como autores de los datos (42%)
- Barreras legales -copyright (41%)
- Pérdida potencial de futuras publicaciones (35%)
- Ventaja que pueden obtener los competidores en el campo (33%)

El Open Government o Gobierno Abierto es una de las iniciativas que también se beneficia del open data. El propósito del gobierno abierto está basado en la transparencia, acceso a la información pública, participación, colaboración ciudadana y rendición de cuentas. Para conseguirlo, hace uso de las tecnologías de información y comunicaciones para promover la interacción entre los ciudadanos y las administraciones públicas.

Cada país o comunidad tiene sus propias iniciativas de gobierno abierto, sin embargo, existe una iniciativa a nivel internacional multilateral de 69 países llamada

Open Government Partnership (OGP)[29] que busca asegurar compromisos concretos de los gobiernos hacia sus ciudadanos para promover la transparencia, empoderamiento, lucha contra la corrupción y usar nuevas tecnologías para fortalecer la gobernabilidad. Para ser miembro de la OGP, los países que quieren participar deben aprobar una Declaración de Gobierno Abierto que debe considerar acciones como:

- Aumentar la disponibilidad de información sobre las actividades gubernamentales.
- Apoyar la participación ciudadana.
- Poner en práctica los más altos estándares de integridad profesional a través de las administraciones.
- Aumentar el acceso a las nuevas tecnologías para la apertura y rendición de cuentas.

## **2.5. EL OPEN DATA COMO HERRAMIENTA PARA LAS SMART CITIES**

Desde la antigüedad o desde las primeras civilizaciones la ciudad ha sido por excelencia el núcleo de desarrollo económico, social, cultural y tecnológico. Esto porque las ciudades han actuado como intercambiadores de ideas y oportunidades de colaboración. Ahora sabemos que el crecimiento poblacional es cada vez mayor y gran porcentaje de éste se da en las zonas urbanas. Se prevé que, hacia el año 2030, casi un 60% de la población mundial residirá en zonas urbanas consumiendo aproximadamente un 80% de los recursos del planeta[30].

El 70% de los ciudadanos europeos vive en áreas urbanas y, de hecho, España es el cuarto país de Europa con la mayor población residente en ciudades (48,5%), por detrás del Reino Unido, Islandia y Chipre. Hoy en día, las urbes se han convertido en productores de grandes cantidades de datos custodiados, en mayor medida, por los organismos públicos [31].

La aplicación de las tecnologías de la informática y telecomunicaciones tiene una influencia muy importante en las relaciones de las personas con las ciudades no solamente generando y compartiendo información sino además reutilizándola. El inminente crecimiento poblacional y ciudades dotadas de infraestructuras cada vez más complejas, requieren por parte de las administraciones públicas mejores capacidades y mayores esfuerzos para atender eficientemente las necesidades de sus ciudadanos. Ante esto, la tecnología ayuda a solucionar el problema mediante herramientas como los sistemas ubicuos y embebidos (sensores), las redes de telecomunicaciones (alámbricas e inalámbricas), los dispositivos móviles y otras, haciendo que en la actualidad las ciudades permitan recoger, almacenar, procesar y mostrar multitud de información, muchas veces incluso en tiempo real, proporcionándonos un flujo constante de datos de alto valor para la ciudad y sus ciudadanos, y que adecuadamente gestionados podremos usar en nuestro beneficio. Sin embargo, la relación entre tecnología y ciudad no siempre ha sido buena, ha estado acompañada de tensiones sobre todo cuando la implementación de una nueva tecnología implicaba la destrucción de algunos puestos de trabajo o conllevaba cambios sociales a los ya establecidos. Como lo expresa el economista austríaco Joseph Schumpeter: *“todo proceso de transformación impulsado por las innovaciones tecnológicas va acompañado de un proceso de destrucción creativa”*.

En el siglo XXI, la aplicación extensiva de las nuevas tecnologías produce una nueva gestión operativa de las ciudades contemporáneas, este conjunto de innovaciones tecnológicas generan un nuevo paradigma denominado smart cities o ciudades inteligentes, cuyos orígenes se remonta al desarrollo de la cibernética a mediados del siglo pasado. Surge como una necesidad de las propias ciudades de ser más eficientes y sostenibles, haciendo uso de las Tecnologías de Información y Comunicaciones para conseguir que sus infraestructuras, componentes y servicios sean más efectivos y participativos.

“Smart city es un nuevo concepto de ciudad que trabaja para mejorar la calidad de vida de sus ciudadanos, garantizando un desarrollo social, económico y urbano sostenible. Una smart city es una ciudad sostenible y autosuficiente que se nutre de la innovación y de las nuevas tecnologías, lo que da como resultado una gestión más eficiente de los servicios y recursos de la ciudad. En la práctica, una smart city es la que tiene la capacidad de dar respuesta a las necesidades de sus ciudadanos (medio ambiente, movilidad, negocios, comunicaciones, energía y vivienda, entre otras) mejorando su día a día. Una ciudad que facilita la interacción de los ciudadanos con la Administración, donde se dispone de información abierta y en tiempo real, donde es posible emprender y, en definitiva, una ciudad que permite y favorece el desarrollo personal y empresarial en la ciudad. Una ciudad que quiera aspirar a ser realmente smart tiene que desarrollar todas sus áreas clave (transportes, energía, educación, salud, gestión de residuos, vigilancia, economía) de manera simultánea y transversal” [32].

La figura 15 muestra la interacción de algunos componentes y servicios en una ciudad inteligente.



Figura 15. El open data y las ciudades inteligentes.

El open data tiene un rol preponderante dentro de las estrategias de Smart Cities de una manera transversal. Puesto que los datos constituyen un activo muy valioso y un recurso esencial para casi cualquier actividad en nuestra sociedad[33] y deberán ser gestionados de manera correcta y adecuada. En este contexto, los datos abiertos son un recurso importante para encontrar una solución a los problemas civiles, impulsar la participación ciudadana y acercar los gobiernos locales a la población. Por estos motivos, las ciudades no pueden ignorar los beneficios que suponen el acceso en abierto y la reutilización de la información del sector público[31].

Quizá el mayor reto en las Smart cities sea el de la interoperabilidad, dada la variedad y heterogeneidad de los sistemas implicados. En este ámbito, los datos abiertos, en combinación con estándares e interfaces también abiertos, tienen mucho que aportar, ya que si utilizamos modelos de datos comunes junto a formatos abiertos y reutilizables, tal y como recomiendan los principios del Open Data, estaremos facilitando que la ciudad se comporte como una plataforma compartida sobre la que poder desarrollar los servicios para los ciudadanos.

Una de las mayores ventajas de la explotación de los datos de las ciudades en abierto son las potenciales mejoras en eficiencia que podemos obtener en áreas muy diversas: transporte, movilidad, turismo, eficiencia energética, medio ambiente, etc. Las aplicaciones en cada una de estas áreas son múltiples, pudiendo además solaparse de forma frecuente, ya que por ejemplo las mejoras conseguidas en la eficiencia de los sistemas de transporte también tendrán generalmente una repercusión positiva sobre el turismo y el medio ambiente (ya que prácticamente el 50% de las emisiones contaminantes en las ciudades provienen del transporte), por ejemplo.

Si bien la iniciativa de datos abiertos juega un papel fundamental en la adecuada planificación y gestión de las ciudades inteligentes, ésta debe ser complementada con una excelente infraestructura tecnológica y un marco legal que la promueva y ampare a sus ciudadanos.

En el capítulo 4 aplicaremos la Norma UNE 178301:2015, norma que trata específicamente del open data para ciudades inteligentes, a varias ciudades españolas.

## 2.6. BENEFICIOS DEL OPEN DATA

En la figura 16 se muestra los beneficios de implementar datos abiertos.

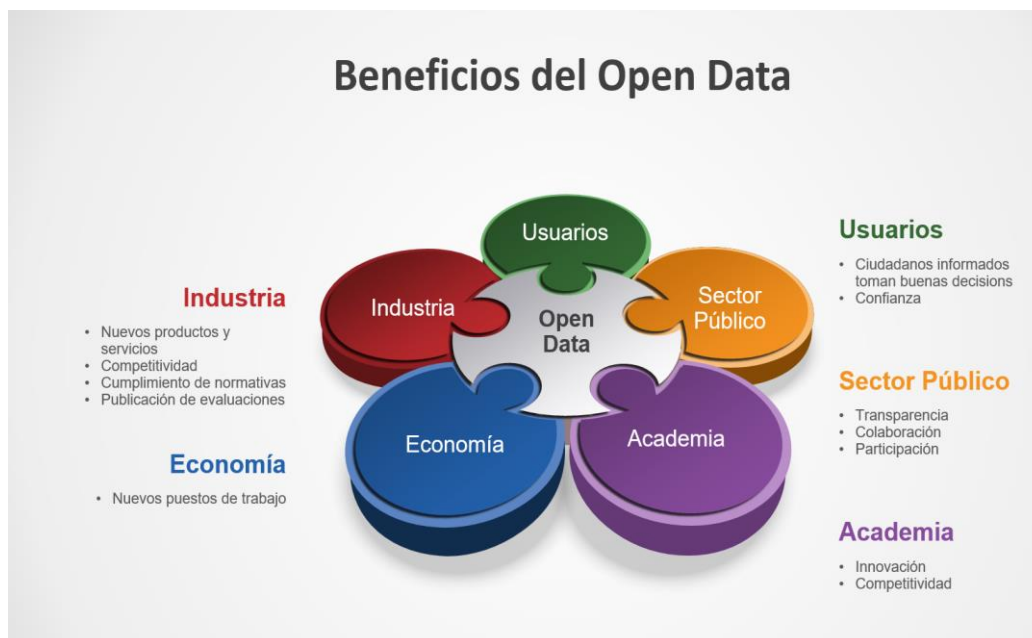


Figura 16: Beneficios del Open data por sectores (Elaboración propia)

Como se observa en la figura 16, implementar datos abiertos mejora la accesibilidad de los ciudadanos hacia la información pública, mejorando de esta manera la confianza de la población al tomar decisiones más informadas. En el sector público, contribuye a cumplir sus objetivos de transparencia, participación ciudadana y colaboración, implementando políticas públicas basadas en las necesidades reales de la ciudadanía. La publicación de datos abiertos promueve la investigación y la innovación mejorando la calidad de los estudios y la competitividad de los investigadores. Un campo poco visto y explorado hasta el momento es la importancia que tienen los datos abiertos para la industria. Los organismos públicos realizan evaluaciones de las empresas respecto al cumplimiento de determinados estándares y normativas. Según el sector de la empresa, si los criterios de evaluación y los resultados de los mismos se pusieran al servicio de los ciudadanos, éstos tomarían decisiones sobre qué empresa puede satisfacer mejor lo que está

buscando. Esta sinergia lleva a las empresas a mejorar la calidad de sus productos y servicios, y por consiguiente, a mejorar la calidad de vida de la población. Por ejemplo, la publicación de los resultados de evaluación sanitaria, calidad, precios, etc. de los restaurantes de una ciudad.

## **2.7. TECNOLOGÍAS PARA IMPLEMENTAR DATOS ABIERTOS**

Existen muchas maneras para implementar un sistema de datos abiertos, ya sea desarrollarlo por la misma organización o utilizar sistemas ya elaborados con licencia comercial como Sparta que está siendo muy utilizado o implementar uno con licencia libre como CKAN, DKAN y Junar (éste último es libre solamente el core).

En concordancia con la filosofía de datos abiertos, en este apartado hablaremos de los conceptos principales, características, componentes y otros aspectos relevantes de las plataformas CKAN y DKAN. Asimismo aprenderemos cómo instalar y configurar cada una de estas plataformas hasta dejarla lista a un nivel básico para poder realizar nuestra primera publicación de datos abiertos. Al final del capítulo realizaremos una comparativa entre ellas.

### **➤ CKAN**

CKAN (del Inglés, Comprehensive Knowledge Archive Network) es una plataforma de código abierto, gratuito y libre desarrollado por la Open Knowledge Foundation (OKF) que permite almacenar, gestionar y publicar los datos de una organización. Asimismo, permite interactuar con otros usuarios o personas. Es por excelencia el gestor de contenidos libre más utilizado en las iniciativas de datos abiertos alrededor del mundo, además de que garantiza la posibilidad de llegar a obtener un open data con clasificación de cinco estrellas.

CKAN es implementado como estrategia de apertura y reutilización de la información en muchos organismos de referencia internacional como [data.gov.uk](http://data.gov.uk) de Reino Unido, [data.gov](http://data.gov) de los Estados Unidos, [publicdata.eu](http://publicdata.eu)

de la Unión Europea, datos.gov.br de Brasil, portales de los gobiernos de Dinamarca, Holanda, Argentina, Finlandia entre otros que detallo en la tabla 1. En España, lo promueve el gobierno en su portal de datos abiertos (como se ve en la figura 17) y además es usado en los portales de datos abiertos de Aragón, Canarias, Ayuntamiento de Málaga.



Figura 17: El Gobierno de España promueve el uso de CKAN

## ARQUITECTURA CKAN

- **SOLR:** Solr es una plataforma de motor de búsqueda basado en Apache Lucene. Está escrito en Java y utiliza la biblioteca de Lucene para implementar la indexación.
- **JETTY:** Servidor HTTP de código abierto, recomendado por CKAN, entre sus ventajas destaca su rendimiento escalable, estabilidad, arquitectura y colaboración. La otra opción es usar Tomcat, Grizzly o Undertow.
- **POSTGRESQL:** Sistema de gestión de bases de datos relacional orientado a objetos y libre.
- **APACHE:** Software de código abierto para crear servidores web.
- **NGINX:** Se usa como servidor proxy invertido, se encargará de hacer llegar las peticiones a los servidores evitando que los clientes se conecten directamente a estos. También es de código abierto.

- **SQLALCHEMY:** es un ORM (Object-Relational Mapping) para Python. Incluye soporte para SQLite, MySQL, PostgreSQL, Oracle, MS SQL, entre otros. ORM es una técnica que permite convertir los tipos de datos usados en los lenguajes de programación orientados a objetos y los tipos de datos de algún sistema de base de datos relacional.

La siguiente figura 18 muestra el esquema de la arquitectura de CKAN.

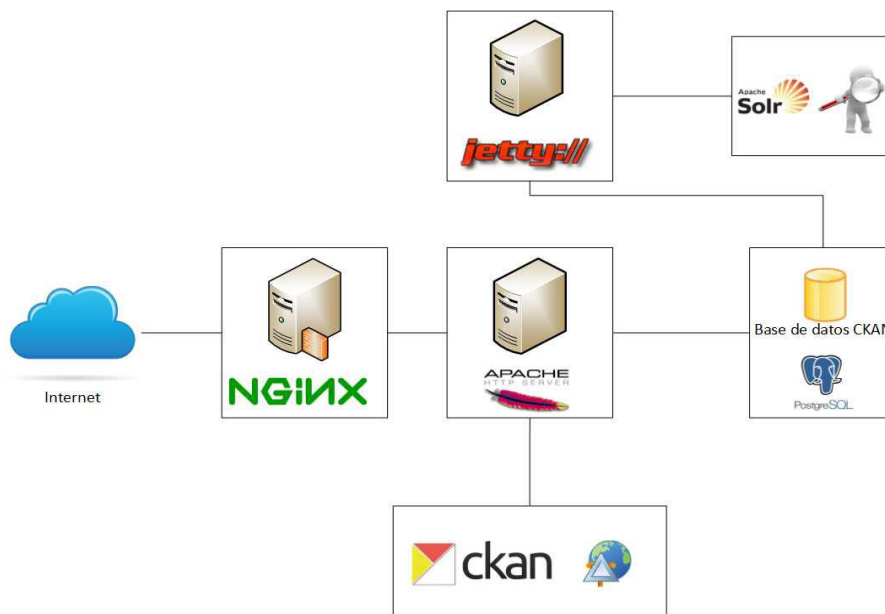


Figura 18: Arquitectura de CKAN (Fuente: Sistemas Open Data 2014)

A continuación (tabla 1) se muestra algunos casos de éxito de organismos en los que se ha implementado datos abiertos utilizando la herramienta CKAN:

Organización	enlace
Gobierno de Londres	<a href="http://data.london.gov.uk/">http://data.london.gov.uk/</a>
Senado de Brasil	<a href="http://dadosabertos.senado.gov.br/">http://dadosabertos.senado.gov.br/</a>
DataHub - OKF	<a href="https://datahub.io/">https://datahub.io/</a>
Taipei City Open Data	<a href="http://data.tainan.gov.tw/">http://data.tainan.gov.tw/</a>
Gobierno de Buenos Aires	<a href="http://data.buenosaires.gob.ar/">http://data.buenosaires.gob.ar/</a>
Gobierno de Berlín	<a href="http://daten.berlin.de/">http://daten.berlin.de/</a>
Catálogo Nacional de Datos Abiertos del Gobierno de Paraguay	<a href="http://datos.gov.py/">http://datos.gov.py/</a>
Gobierno de Aragón	<a href="http://opendata.aragon.es/">http://opendata.aragon.es/</a>
Gobierno de Amsterdam	<a href="http://data.amsterdam.nl/">http://data.amsterdam.nl/</a>
Continente Africano	<a href="https://africaopendata.org/">https://africaopendata.org/</a>

Ayuntamiento de Santander	<a href="http://datos.santander.es/">http://datos.santander.es/</a>
Gobierno de Uruguay	<a href="https://catalogodatos.gub.uy/">https://catalogodatos.gub.uy/</a>
Gobierno de Misiones – Argentina	<a href="http://www.datos.misiones.gov.ar/">http://www.datos.misiones.gov.ar/</a>
Prefectura de San Pablo – Brasil	<a href="http://dados.prefeitura.sp.gov.br/">http://dados.prefeitura.sp.gov.br/</a>
Gobierno de Rotterdam	<a href="http://rotterdamopendata.nl/">http://rotterdamopendata.nl/</a>
Ayuntamiento de Cáceres	<a href="http://opendata.caceres.es/">http://opendata.caceres.es/</a>
Gobierno de Colorado	<a href="http://opencolorado.org/">http://opencolorado.org/</a>
Gobierno de Bélgica	<a href="http://portal.openbelgium.be/">http://portal.openbelgium.be/</a>
British Columbia, Canadá	<a href="https://www.opendatabc.ca/">https://www.opendatabc.ca/</a>

Tabla 1: Casos de éxito de CKAN (fuente: <http://ckan.org/instances/#>)

En el Anexo 1 se indica cómo instalar CKAN y en el anexo 2 se indica cómo configurar CKAN.

## ➤ **DKAN**

DKAN es una herramienta de código abierto y flexible, basado en Drupal con un conjunto completo de funciones que permiten cargar, analizar, almacenar, publicar, catalogar, y visualizar gran variedad de datos de una manera fácil y sencilla, permitiendo a los gobiernos publicar datos fácilmente. Dkan es mantenido por NuCivic y está publicado bajo licencia GNU Public Licence, versión 2.

DKAN se puede instalar en un servidor propio o en uno de los socios de alojamiento como Acquia o Pantheon.

Entre algunas de las características de DKAN, se pueden mencionar las siguientes:

- Cumple con los requisitos de open data de Estados Unidos y es compatible con su portal de datos abiertos data.gov
- Permite publicar fácilmente conjuntos de datos en formatos legibles por máquina incluyendo JSON, XML, RDF y otros.
- Se pueden añadir conjuntos de datos de manera manual o importarlos mediante una API desde otras aplicaciones.
- Gestión de usuarios, roles y privilegios a diferentes niveles.
- Permite gestionar la importación de grandes conjuntos de datos

- Incluye herramientas de visualización de datos de manera gráfica y cartográfica.
- Personalización de los conjuntos de datos mediante la creación y gestión de propios metadatos.
- Los usuarios finales pueden participar, comentar, calificar, compartir o solicitar un conjunto de datos.
- Fácil de personalizar la apariencia de la página web, colores, logotipos o temas.

En la tabla 2 se muestra algunos casos de éxito de organismos en los que se ha implementado datos abiertos utilizando DKAN:

<b>Organización</b>	<b>enlace</b>
Departamento de Salud y Servicios Humanos de EE. UU	<a href="http://www.healthdata.gov/">http://www.healthdata.gov/</a>
Plataforma de datos abiertos de la colonia Alemana	<a href="http://www.offenedaten-koeln.de/">http://www.offenedaten-koeln.de/</a>
Centro de Investigación y Política Pública (CIPP) - Puerto Rico	<a href="http://abrepr.org/">http://abrepr.org/</a>
OpenOakland	<a href="http://data.openoakland.org/">http://data.openoakland.org/</a>
Gobierno de Italia	<a href="http://www.dati.gov.it/">http://www.dati.gov.it/</a>
Gobierno de Sierra Leone	<a href="http://opendata.gov.sl/">http://opendata.gov.sl/</a>
National Democratic Institute	<a href="https://www.nditech.org/project/dkan">https://www.nditech.org/project/dkan</a>
Naciones Unidas	<a href="http://www.nucivic.com/united-nations/">http://www.nucivic.com/united-nations/</a>
Departamento de Agricultura de EE.UU	<a href="http://data.nal.usda.gov/">http://data.nal.usda.gov/</a>
Ayuntamiento de la Ciudad de Mannheim	<a href="https://opendata.mannheim.de/">https://opendata.mannheim.de/</a>
Universidad de Bielefeld	<a href="https://dszbo-portal.uni-bielefeld.de/">https://dszbo-portal.uni-bielefeld.de/</a>
Bosnia-Herzegovina	<a href="http://opendata.ba/">http://opendata.ba/</a>
Belarus	<a href="https://opendata.by/">https://opendata.by/</a>
Instituto Mexicano del Seguro Social	<a href="http://datos.imss.gob.mx/">http://datos.imss.gob.mx/</a>
El Salvador	<a href="http://datoselsalvador.org/">http://datoselsalvador.org/</a>
Gobierno de Arabia Saudí	<a href="http://data.gov.sa/">http://data.gov.sa/</a>
Gobierno de Pompei	<a href="http://open.pompeisites.org/">http://open.pompeisites.org/</a>

Tabla 2: Casos de éxito de DKAN (fuente: <http://www.nucivic.com/dkan/>)

En el anexo 3 de esta memoria, se incluye un manual de instalación de DKAN en un servidor CentOS 7.

## ➤ DIFERENCIAS ENTRE CKAN Y DKAN

La tabla 3 muestra algunas diferencias de las características principales de CKAN y DKAN.

Características	DKAN	CKAN
Lenguaje primario	PHP	Python
Plataforma	Drupal	Pylons
Base de datos	Mysql	PostgreSQL
Vista previa de datos	Recline.js	Recline.js
Módulos/extensiones	30721	129
Gestionado por	Nucivic (EE.UU)	OKF (United Kingdom)
Gestión de entidades	Grupos	Organizaciones

Tabla 3: Diferencias técnicas entre CKAN y DKAN



# CAPÍTULO 3. NORMALIZACIÓN Y MÉTRICAS

En este capítulo hablaremos de la importancia de utilizar normas y métricas en las iniciativas de datos abiertos para smart cities, que faciliten la reutilización de los datos. Respecto a normalización, hablaremos de la norma española UNE 178301:2015 que evalúa las iniciativas de datos abiertos para ciudades inteligentes a nivel organizacional. Dicha norma, por el marco legal que utiliza es aplicable únicamente a instituciones del ámbito español. Y respecto a las métricas hablaremos de dos: la métrica de las cinco estrellas y la métrica Meloda, ambas aplicables a conjuntos de datos de cualquier portal de cualquier parte del mundo.

## 3.1. NORMALIZACIÓN

Según AENOR, “La normalización tiene como objetivo la elaboración de una serie de especificaciones técnicas - NORMAS - que son utilizadas por las empresas, de manera voluntaria, como referencia para probar la calidad y la seguridad de sus actividades y productos. Ante los clientes, el cumplimiento de dichas normas es una garantía del nivel óptimo del bien que ha adquirido y un estímulo para su consumo, ya que lo perciben diferenciado del resto de la oferta del mercado”[34].

Y puesto que analizaremos portales de datos abiertos de España, utilizaremos una de las normas dadas por AENOR, entidad reconocida en España como organismo nacional de normalización, conforme a lo establecido en el Reglamento de la Infraestructura para la Calidad y la Seguridad Industrial, aprobado por el Real Decreto 2200/1995 y en el Reglamento (UE) 1025/2012 sobre Normalización Europea[35].

Como este trabajo trata sobre la importancia del open data para smart cities, hablaremos de la norma UNE 178301:2015, elaborada por el comité

técnico AEN/CTN 178 Ciudades inteligentes cuya secretaría desempeña AENOR.

## - **UNE 178301:2015**

La información mostrada a continuación ha sido extraída de la norma española UNE 178301 Enero 2015, versión corregida, Julio 2015. Adquirida por la Biblioteca de la Facultad de Informática de la Universidad Complutense de Madrid, se usa en este Trabajo de Fin de Máster con fines meramente de aprendizaje e investigación. Los resultados obtenidos fruto de la aplicación de esta norma no comprometen a las instituciones objeto de este estudio puesto que no han sido corroborados por las mismas ni por el Comité Técnico AEN/CTN 178 Ciudades inteligentes.

La norma permite evaluar la publicación de datos abiertos en organismos del sector público de España responsables de la gestión de una ciudad y que administran o tienen bajo su custodia datos públicos, de manera que se facilite la reutilización de los mismos, en el ámbito de las ciudades inteligentes. Se basa en una colección de métricas e indicadores que miden el grado de madurez de la apertura de datos considerando aspectos como la calidad y formato de los datos, accesibilidad, disponibilidad, gratuidad, liderazgo, iniciativas de reutilización, entre otros. Las métricas están agrupadas en dimensiones y las dimensiones en dominios para facilitar su estructura y comprensión.

Para una mejor comprensión de la norma se ha elaborado un esquema (figura 19) y una tabla (tabla 4) en los que se resume el contenido de la misma. Como se observa en la figura 19, la norma está estructurada en 5 dominios (estratégico, legal, organizativo, técnico, económico y social), 10 dimensiones y 37 métricas.

Para un mejor entendimiento del gráfico, se explica con más detalle a continuación:

El dominio *estratégico* tiene una única dimensión llamada *estratégica*, y ésta contiene cuatro métricas: *estrategia*, *liderazgo*, *compromiso de servicio* y *sostenibilidad*.

El dominio *legal* tiene una única dimensión llamada *legal*, ésta a su vez contiene dos métricas: *normas externas e internas y condiciones de uso y licenciamiento*.

El dominio *organizativo* está compuesto por dos dimensiones: *organizativa y medición*. La primera está compuesta por cuatro métricas: *unidad responsable, equipo de trabajo y capacitación, inventario y prioridad*. Y la segunda por dos métricas: *cumplimiento del proceso y uso e impacto*.

El dominio *técnico* está compuesto por cuatro dimensiones: *disponibilidad, acceso, calidad de datos y actualización*. La dimensión *disponibilidad* está compuesta por siete métricas: *catálogo, presencia en el CIP, conjunto de datos documentados, referencias persistentes y amigables, categorización y búsqueda y disponibilidad*. La dimensión *acceso* por tres métricas: *accesibilidad/no discriminación, gratuidad y sistemas de acceso*. La dimensión *calidad de datos* por seis métricas: *datos primarios, datos completos, datos documentados, datos técnicamente correctos, datos enlazados y datos georreferenciados*. Y la dimensión *actualización* por tres métricas: *proceso de actualización, frecuencia de actualización y ampliación de conjunto de datos*.

El dominio *económico y social* está compuesto por dos dimensiones: *reutilización y participación y colaboración*. La primera dimensión, *reutilización*, contiene tres métricas: *cantidad de datos, formato de datos y vocabulario de datos*. Y la segunda dimensión, *participación y colaboración*, contiene cuatro métricas: *transparencia, participación y colaboración, resolución de quejas y conflictos, fomento de la reutilización e iniciativas de reutilización*.

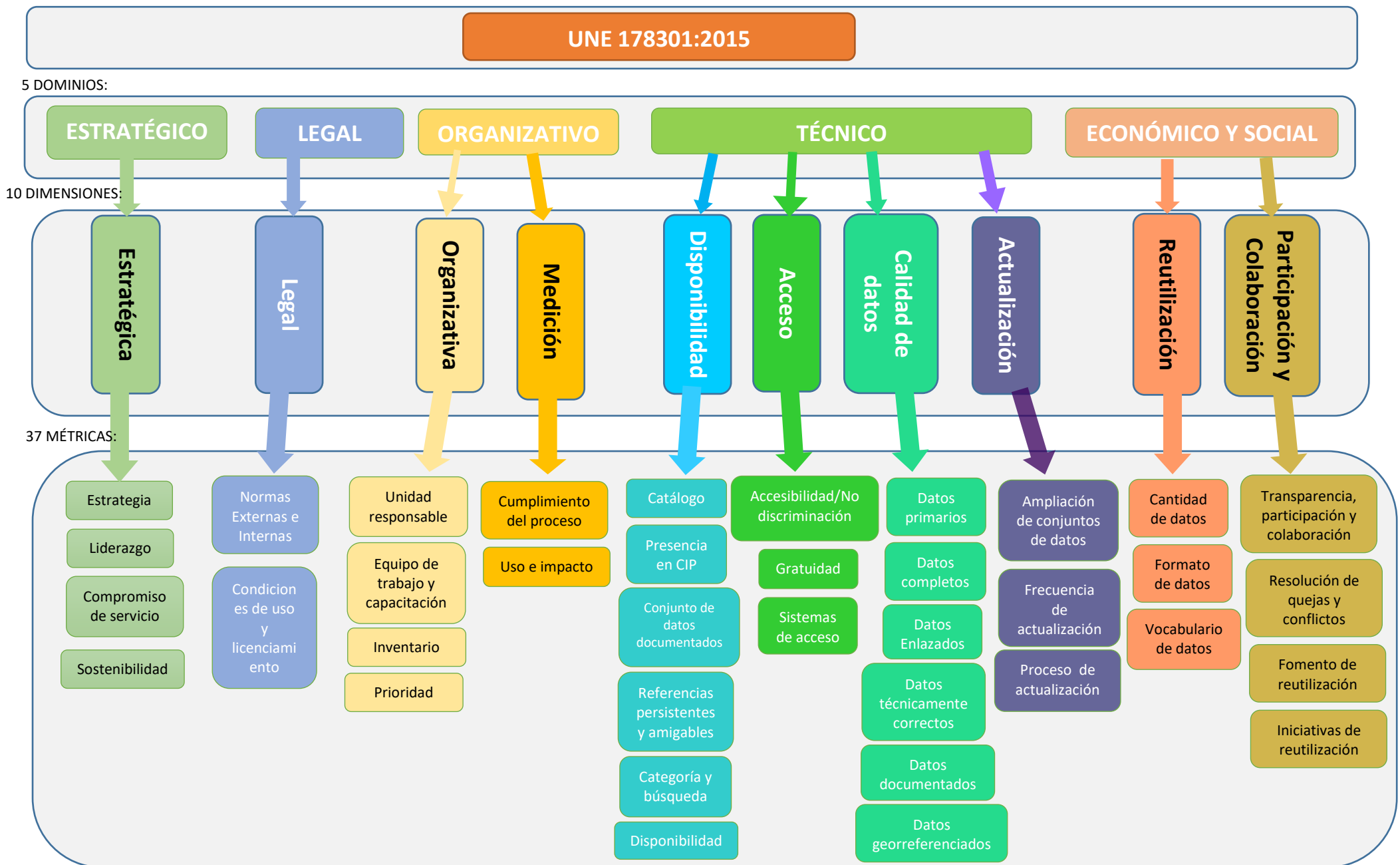


Figura 19. Resumen de la Norma UNE 178301:2015  
(Elaboración propia. Fuente: UNE 178301:2015)

La tabla 4 muestra con más detalle el contenido de la figura anterior. Donde se observa que, la columna uno muestra los cinco dominios. La columna dos muestra las diez dimensiones. La columna tres muestra las 37 métricas. La columna 4 contiene los pesos asignados a cada métrica; en este punto es importante resaltar que la suma de los pesos debe ser 100% para cada dimensión, como se observa en la tabla 4. La columna 5 contiene el nivel o puntuación alcanzado para cada métrica (0, 1, 2 ó 3) según el grado de cumplimiento descrito en la columna 6. Ya que cada dimensión puede sumar como máximo 100% y teniendo que el dominio estratégico tiene una dimensión, el dominio legal una dimensión, el dominio organizativo dos dimensiones, el dominio técnico cuatro dimensiones y el dominio económico y social dos dimensiones, podemos decir que el valor máximo que puede tomar cada dominio es de 100, 100, 200, 400 y 200 respectivamente, como se observa en la figura 20. Es decir, el 10% del valor total corresponde al dominio estratégico, el 10% al dominio legal, el 20% al dominio organizativo, el 40% al dominio técnico y el 20% al dominio económico y social.

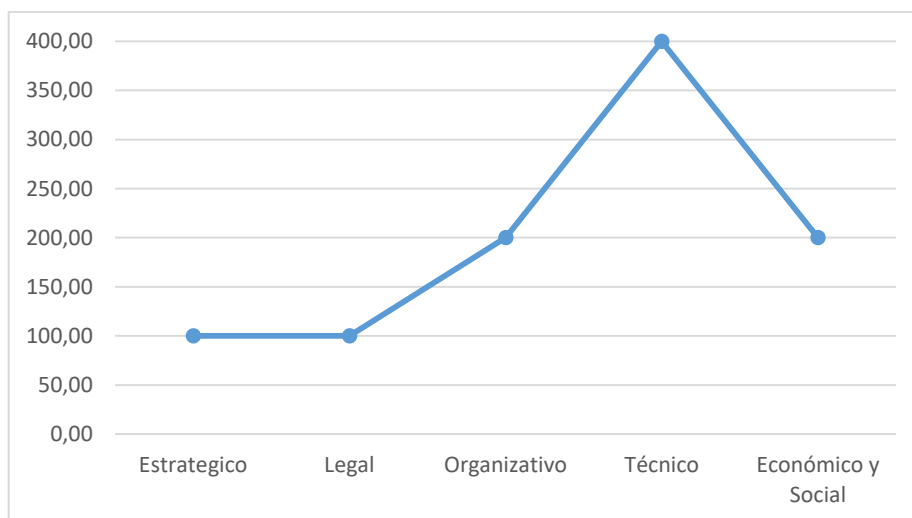


Figura 20. Valores máximos para cada dominio de la norma (Elaboración propia. Fuente: UNE 178301:2015)

A continuación explico la tabla 4 mediante un ejemplo. La métrica “liderazgo” de la dimensión “estratégica” del dominio “estratégico” alcanzará un nivel 3, si existe un responsable político definido formalmente, tiene la

capacidad gestora de involucrar a la alta dirección en las iniciativas y se han materializado acciones concretas de manera proactiva y con carácter de continuidad. Cabe indicar que el peso será el mismo para la métrica, indistintamente del nivel que se haya alcanzado.

Una vez obtenido el nivel asignado y el peso de la métrica podemos obtener el valor parcial de la métrica evaluada ( $V_{pm}$ ), multiplicando el nivel por el peso y lo dividimos por 3, el resultado obtenido lo multiplicamos por 100, como se refiere en la siguiente fórmula, propuesta por la Norma:

$$V_{pm} = \left(\frac{np}{3}\right) 100 \quad (2)$$

**donde:**

**$V_p$  = Valor parcial de la métrica  $m$**

**$n$  = nivel;  $n = [0, 1, 2, 3]$**

**$p$  = peso (%)**

Aplicando la fórmula (2) para el ejemplo descrito en el párrafo anterior, se obtiene el siguiente resultado:

$$V_p = \left(\frac{3 \times 50\%}{3}\right) 100$$

$$V_{p(\text{liderazgo})} = 50$$

Cuando hayamos obtenido el valor parcial de todas la métricas (37 en total), podemos calcular el valor total ( $V_t$ ), haciendo el sumatorio de todos los valores parciales de cada métrica, como se indica en la siguiente fórmula, adaptada de la Norma. El valor total será un número comprendido entre [0,1000].

$$V_t = \sum_{m=1}^{37} (V_{pm}) \quad (3)$$

Donde:

**$V_t$  = Valor total**

**$V_{pm}$  = Valor parcial de la métrica  $m$**

El indicador de datos abiertos para ciudad inteligente se obtiene a partir del Valor total ( $V_t$ ) como se verá en el siguiente capítulo.

En el capítulo 4 hacemos una aplicación de esta norma a los portales de datos abiertos de los Ayuntamientos de Madrid, Zaragoza y Barcelona, en donde aplicamos con más detenimiento las fórmulas (2) y (3) de este apartado.

Tabla 4. Resumen estructurado de la Norma UNE 178301:2015 de AENOR

Dominio	Dimensión	Métrica	peso	nivel	Descriptor
Estratégico	Estratégica	Estrategia	25%	0	El organismo no ha definido una estrategia de apertura de datos
				1	El organismo ha definido informalmente una estrategia de apertura de datos
				2	Estrategia formal, implantada en el propio organismo y al menos 40% de organismos dependientes, publica la lista completa de organismos dependientes y especifica en cuales está implantada la estrategia y en cuales no
				3	Estrategia formal, implantada en el propio organismo y al menos 80% de organismos dependientes, publica la lista completa de organismos dependientes y especifica en cuales está implantada la estrategia y en cuales no
		Liderazgo	50%	0	El organismo no ha definido formalmente un responsable político de apertura de datos
				1	Responsable político definido formalmente pero no se han materializado acciones concretas o exitosas de apertura de datos
				2	Responsable político definido formalmente y se han materializado acciones concretas de manera reactiva y con carácter temporal
				3	Responsable político definido formalmente, tiene la capacidad gestora de involucrar a la alta dirección en las iniciativas, se han materializado acciones concretas de manera proactiva y con carácter de continuidad
		Compromiso de servicio	10%	0	El organismo no ha establecido un compromiso (carta) de servicios asociados a la apertura de datos
				1	El organismo ha establecido un compromiso de servicios pero no se ha hecho público
				2	El organismo ha establecido un compromiso de servicios y se ha hecho público
				3	El organismo ha establecido un compromiso de servicios, se ha hecho público, se evalúa el rendimiento y se hacen públicos los resultados
	Sostenibilidad económica	15%	0	El organismo no ha realizado una evaluación de la sostenibilidad económica	
			1	El organismo ha realizado una evaluación de la sostenibilidad económica, solamente de la fase puesta en marcha	
			2	El organismo ha realizado una evaluación de la sostenibilidad económica, solamente de las fases puesta en marcha y de mantenimiento	
			3	El organismo ha realizado una evaluación de la sostenibilidad económica, de las fases puesta en marcha, mantenimiento y evolución	
Legal	Legal	Normas externas e internas	30%	0	El organismo no cumple la normativa externa de obligado cumplimiento
				1	El organismo cumple solamente la normativa externa de obligado cumplimiento
				2	El organismo sigue la normativa externa de obligado y recomendado cumplimiento
				3	El organismo sigue la normativa externa de obligado y recomendado cumplimiento y otras normas, guías etc. Que armonizan con otras Administraciones. Además, se han definido y publicado normativas propias
		Condiciones de uso y licenciamiento	70%	0	No se ha definido ni adoptado licencias ni condiciones de uso de los datos o si existen los datos solo se ceden para uso privado
				1	Se ha definido o adoptado licencias o condiciones de uso de los datos pero se ceden solo para uso no comercial.
				2	Se ha definido o adoptado licencias o condiciones de uso de los datos conforme a un modelo de licenciamiento abierto y se permite cualquier uso comercial de los datos.

Tabla 4. Resumen estructurado de la Norma UNE 178301:2015 de AENOR

Dominio	Dimensión	Métrica	peso	nivel	Descriptor
				3	Además del nivel 2, se implementan otras licencias para casos especiales siempre que su uso esté justificado y se publique en la web
Organiza-tivo	Organiza-tiva	Unidad res-ponsable	25%	0	No se tiene establecido la unidad responsable de datos abiertos
				1	Se ha identificado una unidad responsable de datos abiertos pero no se ha establecido formalmente
				2	Se ha establecido una unidad responsable de datos abiertos y está formalizada oficialmente mediante una ordenanza o similar
				3	Además del nivel 2, existen procedimientos independientes del nivel político para garantizar la función a largo plazo, capacidad de actuación en todo el organismo, la unidad cuenta con información suficiente para atender peticiones de datos y formalizar su distribución según normativa aplicable
		Equipo de trabajo y capacita-ción	50%	0	No se ha establecido un equipo de trabajo específico para apertura de datos
				1	No se ha establecido un equipo de trabajo para apertura de datos pero existe un equipo de trabajo de carácter puntual aunque no especializado en datos abiertos
				2	Equipo de trabajo establecido formalmente de forma estable, abarca ámbitos de gestión, planificación y técnico. Existe un plan de formación específico en apertura de datos
				3	Equipo de trabajo establecido formalmente de forma estable, abarca ámbitos de gestión, planificación y técnico. El equipo ha recibido una capa-citación apropiada para el cumplimiento de normativas y tecnologías asociadas a la apertura de datos
		Inventario	15%	0	El organismo no ha creado un inventario de información reutilizable
				1	El organismo ha creado un inventario pero no recoge la información adicional
				2	El organismo ha creado un inventario, recoge la información adicional pero no se publica
				3	El organismo ha creado un inventario, recoge la información adicional y está publicado para los ciudadanos y reutilizadores
	Prioridad	10%	0	El organismo no ha realizado un análisis de prioridad de los datos a publicar	
			1	EL organismo ha iniciado un análisis de prioridad de datos a publicar	
			2	El organismo ha concluido un análisis de prioridad de datos a publicar basándose solamente con opiniones internas	
			3	El organismo ha concluido un análisis de prioridad de datos a publicar y prioriza la publicación basándose en un análisis de opiniones internas y externas (sector reutilizador)	
	Medición del cumpli-miento del proceso	50%	0	El organismo no ha establecido un plan formal de evaluación del proyecto	
			1	Se ha establecido un plan esporádico de evaluación del proyecto, se miden los resultados del proyecto, sin un procedimiento estandarizado ni se tiene establecidas metas de cumplimiento	
			2	Se ha establecido un plan concreto de evolución del proyecto periódico y sistemático, se han identificado métricas e indicadores y sus correspon-dientes formas de cálculo para evaluar el desempeño de iniciativas de apertura de datos, se han establecido metas y objetivos de cumplimiento.	
3			Además del nivel 2, se aplican medidas correctivas para mejorar el desempeño		

Tabla 4. Resumen estructurado de la Norma UNE 178301:2015 de AENOR

Dominio	Dimensión	Métrica	peso	nivel	Descriptor		
		Medición del uso e impacto	50%	0	El organismo no ha implantado mecanismos que registren el nivel de uso e impacto de los datos abiertos		
				1	Se ha implantado un mecanismo de medición, pero el seguimiento de los mismos no se realiza regularmente. No se proponen objetivos a cumplir ni medidas de mejoras evaluables en el siguiente periodo de medición		
				2	Se ha implantado un mecanismo de medición y realiza un seguimiento regular. Se mide el cumplimiento de los objetivos propuestos pero no se diseñan e implementan medidas de mejoras evaluables en el siguiente periodo de evaluación.		
				3	Se ha implantado un mecanismo de medición, se realiza seguimiento regular y la información obtenida se pone a disposición pública. Se mide el cumplimiento de los objetivos y se diseñan e implantan medidas de mejora evaluables en el siguiente periodo de evaluación.		
Técnico	Disponibilidad	Catálogo	30%	0	El organismo no dispone de un punto de acceso específico a los datos abiertos		
				1	Dispone de un punto de acceso específico y global para sus organismos dependientes o departamentos formalizados bajo un catálogo, cumpliendo minoritariamente la legislación vigente.		
				2	Dispone de un punto de acceso específico y global para sus organismos dependientes o departamentos formalizados bajo un catálogo, cumpliendo mayoritariamente la legislación vigente.		
				3	Dispone de un punto de acceso específico y global para sus organismos dependientes o departamentos formalizado bajo un catálogo, cumpliendo completamente la legislación vigente		
		Presencia en el Catálogo de Información Pública (CIP)	25%	0	El organismo no ha realizado acciones precisas para tener presencia en el CIP ( <a href="http://www.datos.gob.es">www.datos.gob.es</a> )		
				1	Ha realizado acciones precisas para tener presencia en el CIP pero solo de algunos conjuntos de datos publicados en su catálogo.		
				2	Ha realizado acciones precisas para tener presencia en el CIP de todos los conjuntos de datos publicados en su catálogo, pero no están igual de actualizados.		
				3	Ha realizado acciones precisas para tener presencia en el CIP de todos los conjuntos de datos publicados en su catálogo y están igual de actualizados.		
		Conjunto de datos documentados	15%	0	No todos los conjuntos de datos cuentan con los metadatos obligatorios		
				1	Todos los conjuntos de datos cuentan con el 100% de los metadatos obligatorios pero con menos de un 20% de los opcionales.		
				2	Todos los conjuntos de datos cuentan con el 100% de los metadatos obligatorios pero con menos de un 50% de los opcionales.		
				3	Todos los conjuntos de datos cuentan con el 100% de los metadatos obligatorios pero con menos de un 100% de los opcionales.		
		Categorización y búsqueda	10%	0	No existe un servicio de búsqueda específica de conjunto de datos		
				1	Existe un servicio de búsqueda simple de conjunto de datos por texto		
				2	Existe un servicio de búsqueda por texto y taxonomías		
				3	Existe un servicio de búsqueda multicriterio por texto, taxonomías, formato, categorías, tipo de información (georreferenciada, estadística, etc) periodo de actualización, fecha de actualización, etc.		
					10%	0	No se realiza un muestreo para comprobar que los datos están disponibles y son correctos

Tabla 4. Resumen estructurado de la Norma UNE 178301:2015 de AENOR

Dominio	Dimensión	Métrica	peso	nivel	Descriptor
	Disponibilidad			1	Se realiza un muestreo periódico y se han tomado medidas para asegurar que el servicio de datos responde incluso aunque no se puedan proporcionar los datos solicitados
				2	Se han tomado las medidas para asegurar que el servicio de datos está disponible y devuelve datos en el formato definido. El servicio de datos está disponible al menos el 97% del tiempo.
				3	Se han tomado las medidas para asegurar que el servicio de datos está disponible y devuelve datos en el formato definido. El servicio de datos está disponible al menos el 99% del tiempo.
		Referencias persistentes y amigables	10%	0	Las URI no siguen el patrón establecido en la legislación vigente
				1	Las URI siguen el patrón establecido en la legislación vigente
				2	Además del nivel 1, se incluye información de procedencia y pero no los números de versión ni la información de estado.
				3	Además del nivel 2, se implementan redirecciones de múltiples opciones para proporcionar los datos en distintos formatos a partir de una URI dada.
	Acceso	Accesibilidad /No discriminación	30%	0	Más del 20% de los conjuntos de datos requiere realizar un trámite administrativo
				1	Más del 20% de los conjuntos de datos requiere realizar un registro de usuario
				2	Entre el 6% y 19% de los conjuntos de datos requiere realizar un trámite administrativo o registro de usuario y su justificación está publicada
				3	Máximo el 5% de los conjuntos de datos requiere realizar un trámite administrativo o registro de usuario y su justificación está publicada.
		Gratuidad	50%	0	Más del 20% de los conjuntos de datos están sujetos al pago de una tasa
				1	Entre el 11% y el 19% de los conjuntos de datos están sujetos al pago de una tasa
				2	Entre el 6% y el 10% de los conjuntos de datos están sujetos al pago de una tasa y su justificación está publicada
				3	Máximo el 5% de los conjuntos de datos están sujetos al pago de una tasa y su justificación está publicada
		Sistemas de acceso	20%	0	Solamente existe la opción de acceso de descarga de datos en forma de ficheros
				1	Además del nivel 0, se dispone de APIs específicas para algunos de los conjuntos de datos.
				2	Además del nivel 0, se dispone de una o varias APIs que combinan datos procedentes de varios conjuntos de datos
				3	Además del nivel 2, se dispone de un punto de consulta SPARQL que permite realizar consultas sobre datos procedentes de varios conjuntos de datos
	Calidad de datos	Datos primarios	20%	0	Menos del 90% de los conjuntos de datos son primarios
1				El 90% de los conjuntos de datos son primarios, pero no se ha evaluado la publicación de máximo nivel de detalle y no se explica ni publica el método seguido de agregación	
2				El 90% de los conjuntos de datos son primarios, se ha evaluado la publicación de máximo nivel de detalle pero no se explica ni publica el método seguido de agregación	

Tabla 4. Resumen estructurado de la Norma UNE 178301:2015 de AENOR

Dominio	Dimensión	Métrica	peso	nivel	Descriptor		
				3	El 90% de los conjuntos de datos son primarios, se ha evaluado la publicación al máximo nivel de detalle y se explica y publica el método seguido de agregación		
		Datos completos	20%	0	Menos del 90% de los conjuntos de datos son completos		
				1	El 90% de los conjuntos de datos son completos, pero no existe un procedimiento para valorar la posibilidad de completar aquellos que no lo están		
				2	El 90% de los conjuntos de datos son completos y existe un procedimiento para valorar la posibilidad de completar aquellos que no lo están y que son reclamados por los reutilizadores		
				3	El 90% de los conjuntos de datos son completos y existe un procedimiento interno para solicitar y completar aquellos que no lo están		
		Datos documentados	20%	0	Menos del 90% de los datos están documentados		
				1	El 90% de los datos publica la documentación de la estructura de datos		
				2	El 90% de los datos publica la documentación de la estructura y el tipo de los datos		
				3	El 90% de los datos publica la documentación de la estructura, el tipo y los posibles valores que los datos pueden tomar, nota de interpretación, etc.		
		Datos técnicamente correctos	20%	0	No se realiza ningún test o prueba para verificar los datos		
				1	Se han implantado mecanismos para asegurar que la sintaxis es correcta		
				2	Además del nivel 1, se utiliza la misma codificación y normalización para el mismo tipo de dato publicado en diferentes conjuntos de datos del catalogo		
				3	Además del nivel 2, la codificación y normalización utilizada se basa en algún estándar común reconocido y utilizado por otras organizaciones		
		Datos georreferenciados	10%	0	Menos del 70% de los conjuntos de datos geográficos están publicados como datos abiertos georreferenciados con coordenadas		
				1	El 70% o más de los conjuntos de datos geográficos están publicados como datos abiertos georreferenciados con coordenadas		
				2	El 70% de los conjuntos de datos geográficos están publicados como datos abiertos georreferenciados con coordenadas, acompañados de sus metadatos geográficos		
				3	El 100% de los conjuntos de datos geográficos están publicados como datos abiertos georreferenciados con coordenadas, acompañados de sus metadatos geográficos.		
		Datos enlazados	10%	0	Los datos no están enlazados con otros internos ni externos al organismo		
				1	El 20% de los conjuntos de datos están enlazados pero solo a datos internos al organismo		
				2	El 20% de los conjuntos de datos están enlazados a datos externos del organismo		
				3	El 50% de los conjuntos de datos están enlazados a datos externos del organismo		
					40%	0	El 100% de los procesos de actualización son manuales

Tabla 4. Resumen estructurado de la Norma UNE 178301:2015 de AENOR

Dominio	Dimensión	Métrica	peso	nivel	Descriptor
	Actualización	Proceso de actualización		1	El 70% de los procesos de actualización son manuales
				2	El 70% de los procesos de actualización son automáticos
				3	El 90% de los procesos de actualización son automáticos
		Frecuencia de actualización	40%	0	El 70% de la frecuencia de actualización de los datos es más de un mes después de que se actualizan internamente
				1	El 70% de la frecuencia de actualización de los datos es un mes después de que se actualizan internamente
				2	El 70% de la frecuencia de actualización de los datos es un día después de que se actualizan internamente
				3	EL 90% de la frecuencia de actualización de los datos es el mismo tiempo que se actualizan internamente
		Ampliación de conjuntos de datos publicados	20%	0	El organismo no tiene un plan de ampliación del conjunto de datos publicado
				1	El organismo ha publicado unos conjuntos de datos iniciales y se ha ampliado pero sin tener en cuenta la solicitudes de ciudadanos o reutilizadores
				2	El organismo ha publicado unos conjuntos de datos iniciales, se ha ampliado por iniciativa propia y por solicitudes de ciudadanos o reutilizadores, pero no se informa previa y públicamente de estas ampliaciones para recibir opiniones
				3	El organismo ha publicado unos conjuntos de datos iniciales, se ha ampliado por iniciativa propia y por solicitudes de ciudadanos o reutilizadores y se informa previa y públicamente de éstas ampliaciones para recibir opiniones. Adicionalmente para nuevos sistemas de información del organismo se establece la política de apertura por defecto de los datos.
		Económico y social	Reutilización de datos	Cantidad de datos publicados	30%
1	Entre el 11% y el 30% del inventario de datos está publicado				
2	Entre el 31% y el 50% del inventario de datos está publicado				
3	Más del 50% del inventario de datos está publicado				
Formato de los datos	40%			0	Menos del 90% de los conjuntos de datos se ofrecen en formatos estándares y estructurados, pero no abiertos
				1	El 90% de los conjuntos de datos se ofrecen en formatos estándares, estructurados, pero no abiertos
				2	El 90% de los conjuntos de datos de ofrecen en formatos estándares, estructurados, y abiertos pero no permiten la identificación única de los recursos
				3	El 90% de los conjuntos de datos se ofrecen en formatos estándares, estructurados, abiertos y permite la identificación única de los recursos
Vocabularios	30%			0	Menos del 30% de los vocabularios utilizados son independientes del productor de los datos, son estándares y son auto-descriptivos
				1	El 30% de los vocabularios utilizados son independientes del productor de los datos, son estándares y son auto-descriptivos
				2	El 30% de los vocabularios utilizados son independientes del productor de los datos, son estándares, son auto-descriptivos y tienen URL persistentes
				3	Además del nivel 2, se proporciona una política de control de versiones y son escritos en más de un idioma.

Tabla 4. Resumen estructurado de la Norma UNE 178301:2015 de AENOR

Dominio	Dimensión	Métrica	peso	nivel	Descriptor
	Participación y colaboración	Transparencia, participación y colaboración	30%	0	El organismo no dispone de un canal específico para enviar opiniones, participar, colaborar o pedir la publicación de nuevos datos
				1	EL organismo dispone de un canal específico para enviar opiniones, participar, colaborar o pedir la publicación de nuevos datos pero no se publican para que otras personas puedan conocerlo y aportar su opinión
				2	EL organismo dispone de un canal específico para enviar opiniones, participar, colaborar o pedir la publicación de nuevos datos y se publican para que otras personas puedan conocerlo y aportar su opinión
				3	Además del nivel 2, se publica el resultado del análisis realizado con indicación de una planificación, prioridad, etc. Para su implementación.
		Resolución de quejas y conflictos	20%	0	El organismo no ha integrado el servicio de datos abiertos en el sistema de quejas
				1	El organismo ha integrado el servicio de datos abiertos en el sistema de quejas pero no se contestan en el plazo
				2	El organismo ha integrado el servicio de datos abiertos en el sistema de quejas, se contestan en el plazo pero no se cumple el derecho de obtener una respuesta coherente y adecuada según la legislación vigente en general y la de reutilización de datos en particular
				3	El organismo ha integrado el servicio de datos abiertos en el sistema de quejas, se contestan en plazo y se cumple el derecho de obtener una respuesta coherente y adecuada según la legislación vigente en general y la de reutilización de datos en particular
		Fomento de la reutilización	30%	0	El organismo no ofrece recursos de ayuda o bien son escasos e inadecuados, si se promueven actividades de fomento de la reutilización
				1	El organismo publica documentos y materiales, fundamentalmente divulgativos y se celebran o planifican jornadas divulgativas externas
				2	Además del nivel 1, se realiza formación interna técnica para fomentar la reutilización interna
				3	Además del nivel 1, se realiza formación interna técnica para fomentar la reutilización interna y externa
	Iniciativas de reutilización desarrolladas	20%	0	El organismo no ha desarrollado ninguna iniciativa de reutilización	
			1	El organismo ha desarrollado alguna iniciativa de reutilización interna dentro del organismo	
			2	El organismo ha desarrollado varias iniciativas de reutilización tanto internas dentro del organismo como externas con reutilizadores	
			3	Además del nivel 2, se han concretado en aplicaciones o servicios y se han publicado en el catálogo de datos como aplicaciones que utilizan los datos	

## 3.2. MÉTRICAS

Ante la falta de homogeneidad en los conjuntos de datos que se publican en los portales open data, es necesario contar con métricas que nos ayuden a conseguir dar el mayor valor a los datos, para ello las métricas están enfocadas principalmente a analizar y asegurar el grado de reutilización de la información, es decir, que los conjuntos de datos se publiquen de manera apropiada, entendible por el hombre y procesable por equipos informáticos. Aunque, técnicamente es complicado analizar el impacto general de reutilización de los sitios de publicación de datos. El open data en España supone un sector potencial de 2.000 millones de euros, equivalente a 45.000 empleos en los próximos 10 años, según la tasa estándar de ocupación de empleo[25].

Aunque la norma descrita en la sección 3.1 también hace uso de métricas para evaluar iniciativas de datos abiertos, en esta sección estudiaremos dos métricas que están siendo muy usadas: en la primera parte estudiaremos la métrica de cinco estrellas del open data propuesta por Tim Berners-Lee, por ser referente a nivel internacional y por ser una de las primeras métricas para evaluar datos abiertos, y en la segunda parte estudiaremos la métrica Meloda propuesta por un grupo de investigación de la Universidad rey Juan Carlos de Madrid, por ser una métrica más completa que de las cinco estrellas. Para una mejor comprensión de la métrica Meloda, en la sección 3.3 de éste capítulo, hacemos una aplicación sobre conjuntos de datos del Ayuntamiento de Madrid.

### - LAS CINCO ESTRELLAS DEL OPEN DATA

Es una iniciativa del fundador de la web e iniciador de los datos enlazados (Linked Data), Tim Berners-Lee[36]. Analiza los conjuntos de datos en base al formato y la licencia con que han sido publicados, esta métrica apunta a que los datos deben estar integrados en la web y además enlazados a otros datos. Como el mismo nombre de la métrica lo indica, contempla cinco niveles de despliegue, siendo el nivel 'una estrella' el más básico y el nivel 'cinco estrellas' el más alto, como se muestra en la figura 21.

La métrica parte de la idea de que la información es accesible vía web bajo una licencia abierta y que el consumidor puede visualizar, guardar localmente, utilizar, reutilizar y compartir la información.

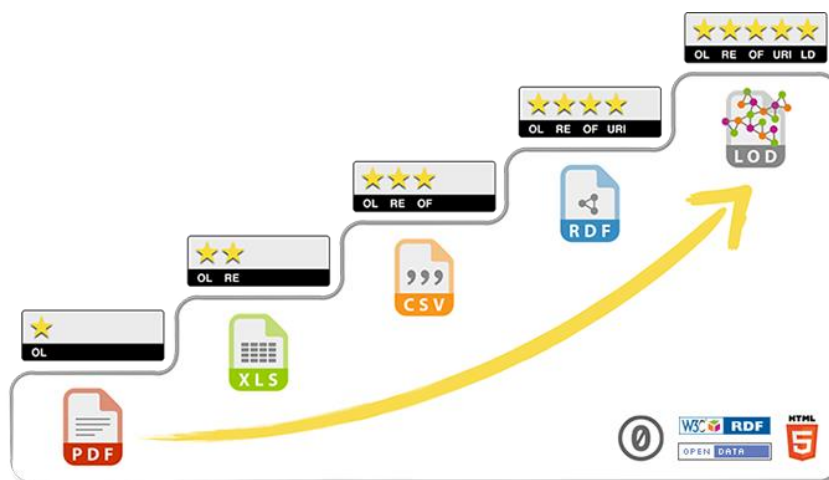


Figura 21: cinco estrellas del Open Data, Tim Berners-Lee.



Datos publicados en la web en formato no estructurado y bajo una licencia abierta como PDDL, ODC-by o CCO. Sin embargo, es difícil extraer los datos ya que están atrapados en un documento. Ejemplo, un documento o una tabla escaneada, JPEG, PDF imagen.



Datos accesibles en la web, en formato estructurado. Sin embargo, para extraer los datos es necesario un software propietario. Ejemplo, XLS(X), ODS, ODT, DOC(X), HTML.



Publica los datos utilizando formatos estructurados y abiertos, cualquiera puede utilizar y manipular los datos fácilmente y de cualquier forma, sin limitación de características o de uso de algún software en particular. Aún así, todavía son solo datos publicados en la web y no datos integrados en la web. Ejemplo, CSV, XML, JSON.



En este nivel los datos están integrados en la web, los datos (objetos) se identifican mediante una URI para denotar cosas. Se pueden enlazar los

datos desde cualquier otro sitio web local, marcarlos como favoritos, reutilizar parte de los datos, reutilizar herramientas y librerías disponibles. Por ejemplo RDF. También se puede publicar en formatos CSV, XML o JSON. EL nivel 4\* difiere del nivel 3\* en que las entidades se refieren mediante URI (Uniform Resource Identifier) por lo que es posible hacer referencia a ellos desde otras bases de datos.



Los datos están integrados en la web y además enlazados a otros datos. El nivel 5\* difiere del nivel 4\* en que los datos no solamente hacen referencia mediante URL a datos de la misma entidad, sino que también hacen referencia mediante URL a conjuntos de datos de otras entidades. Debe asegurarse que los datos son accesibles mediante URL en un formato legible por máquina mediante el modelo RDF. Cuando se llega a este nivel, podemos decir que nuestro portal de datos abiertos cumple con los criterios de Linked Data, que es el tipo de web que promueve el fundador de la web y promotor de esta métrica.

En el nivel 5\*, se van descubriendo nuevos datos relacionados mientras se consumen los datos. Se necesita de confianza, ya que los datos se presentan desde un enlace arbitrario siendo un riesgo dejar que cualquier persona incluya contenido en la web desde cualquier sitio web.

## **- MELODA**

La métrica MELODA (MEtric for reLeasing Open DAta) [37] permite evaluar el grado o capacidad de reutilización de los datos abiertos.

El promotor de la métrica, Alberto Abella, es investigador en gobierno abierto y datos abiertos con especial interés en ciudades inteligentes. Las primeras versiones han estado disponibles desde el 2011, Meloda 2.0 ha sido publicada por la Universidad Rey Juan Carlos de Madrid en noviembre del 2013. Esta versión 2.0, consideraba tres dimensiones (estándares técnicos, acceso y legal), manteniéndose así hasta la versión 2.5[38]. Luego se observó que era necesario incluir una dimensión que tuviera en cuenta el modelo de

datos a publicar y es así como en la versión 3.10 se incluye una cuarta dimensión (modelo de datos).

La métrica está compuesta por dimensiones, cada dimensión tiene niveles que se asignan según el cumplimiento del conjunto de datos, cada nivel tiene asociado un peso, que es un valor porcentual que se utiliza para calcular el valor Meloda del conjunto de datos. A continuación se explica en más detalle:

- i. **Estándares técnicos:** esta dimensión evalúa los formatos en los que son publicados los conjuntos de datos. Califica la información en 4 niveles y les asigna un peso a cada uno de ellos:
  - **Nivel 1:** si el conjunto de datos es publicado en formato estándar cerrado no reutilizable, (peso= 10%). Ej.: pdf image, doc.
  - **Nivel 2:** si el conjunto de datos es publicado en formato estándar cerrado reutilizable, (peso=35%). Ej.: shp, xls y xlsx con macro/fórmula.
  - **Nivel 3:** si el conjunto de datos está publicado en formato estándar abierto, (peso= 60%). Ej.: csv, txt, odb, odt, ods, WMS, xls y xlsx sin macro/fórmula.
  - **Nivel 4:** si el conjunto de datos está publicado en formato de estándar abierto, con metadatos, (peso= 100%). Ej.: rdf, rss, json, xml.
- ii. **Acceso a la Información:** esta dimensión evalúa la facilidad con la que se accede a la información.
  - **Nivel 1: Sin acceso web o solicitud manual:** el acceso a la información requiere de una aprobación no automática o los datos están registrados en formato no digital. Se considera también en este nivel si el acceso a la información no es libre, (peso=0%).
  - **Nivel 2: Acceso vía web con registro:** el acceso a la información es a través de la web, pero se requiere la interacción manual del usuario para seleccionar la fuente de datos, (peso=10%).

- **Nivel 3: Acceso directo vía web:** URL única y constante, URL con parámetros (pero no variación en los parámetros) para el mismo conjunto de datos, (**peso=50%**).
- **Nivel 4: Acceso vía web con parámetros:** URL con parámetros, sin tener que descargar todo el conjunto de datos, (**peso=90%**).
- **Nivel 5: Acceso completo (API o lenguaje de consulta):** punto de acceso *Sparql*, (**peso=100%**).

### iii. Marco legal

Esta dimensión mide el tipo de licencia o condiciones de uso con que se han publicado los conjuntos de datos. Considera cinco niveles:

- **Nivel 1: Copyright:** el autor se reserva el derecho de los datos, restringe el uso no autorizado, (**peso=0%**).
- **Nivel 2: Uso privado:** se permite el uso de datos sin necesidad de procesos de aprobación, pero solo para usos privados. No disponible para uso público, (**peso=10%**).
- **Nivel 3: Reutilización no comercial:** se permite la reutilización de datos, pero no para usos comerciales, (**peso=25%**). Ej.:
  - (CC BY-NC-ND 4.0) Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
  - (CC BY-NC 4.0) Attribution-NonCommercial 4.0 International
  - (CC BY-NC-SA 4.0) Attribution-NonCommercial-ShareAlike 4.0 International
- **Nivel 4: Reutilización comercial:** se permite la reutilización de datos, inclusive la reutilización comercial, (**peso=90%**). Ej.:
  - (CC BY-ND 4.0) Attribution-NoDerivatives 4.0 International
  - (CC BY-SA 4.0) Attribution-ShareAlike 4.0 International
- **Nivel 5: Sin restricciones o solo atribuciones:** sólo se pide la atribución de los datos a los reutilizadores, (**peso=100%**). Ej.:

- Attribution 4.0 International (CC BY 4.0)

#### iv. Modelos de datos

Esta dimensión no analiza el modelo de datos en sí, sino la capacidad para compartir el modelo de datos con otras entidades. Considera cinco niveles:

- **Nivel 1: Sin modelo publicado:** el modelo no se libera ni se explica. Ej.: tabla de datos sin descripción de los campos, **(peso=15%)**.
- **Nivel 2: Modelo con campos de datos:** hay campos diseñados por el editor. Sin embargo, los campos se identifican simplemente, no explica los rangos, tipo y resto de características. Solo se utiliza por la entidad editora. Ej.: Tabla de datos con descripción de los campos, **(peso=35%)**.
- **Nivel 3: modelo de datos propio con especificaciones de campos:** se detallan y se publican las especificaciones de los campos, aunque sea propio; y se deja libre para ser utilizado por otros. Ej.: Hay una ontología o descripción pública libremente disponible, vocabularios disponibles, **(peso=50%)**.
- **Nivel 4: Modelo local de datos abiertos:** se usa un modelo externo normalizado, aunque esté poco extendido, **(peso=90%)**.
- **Nivel 5: Modelo de datos abierto:** Hay un modelo de datos normalizado publicado por una entidad global (es decir, supranacional, ISO) y/o existe una amplia adopción (es decir, cuando es obligatorio por Ley), **(peso=100%)**.

Sin embargo, recientemente (2016) se ha publicado en la página web oficial de la métrica[37], la última versión 4.0 en la que se ha añadido dos dimensiones (Geolocalización y Actualización de la información), dimensiones que seguramente aún vienen puliendo en el referido grupo de investigación:

#### v. Geolocalización de la información

Los datos publicados podrían contener información sobre la localización de los mismos. Esto no significa que la información tenga que ser geográfica, sino que tiene algunos campos que permiten identificar su ubicación.

- **Nivel 1: No hay información geográfica:** la información publicada no tiene ningún campo en referencia a los ejemplos de ubicación, (**peso=15%**). Ej.: 15.
- **Nivel 2: Campo de texto simple:** La información geográfica es sólo un campo de texto o un Id propio que hace difícil conectar esa información con otras bases de datos, (**peso=30%**). Ej.: El campo País, valor España.
- **Nivel 3: Campo de texto complejo:** La información geográfica son varios campos con descripción de texto (el número de una calle se considera como un texto) y los campos son jerárquicos, (**peso=50%**). Ej.:  
Campo1, país, valor: España, Campos2, ciudad, valor: Astorga 50.
- **Nivel 4: Coordenadas:** la información publicada incluye dos campos con las coordenadas (en un sistema de coordenadas), (**peso=90%**). Ej.: Campo1, Latitud, Valor: 42.605556, Campo2, longitud, valor: -5.590.
- **Nivel 5: Nivel de información geográfica completa:** incluye los niveles 3 y 4. (**peso=100%**). Ej.:  
1) Campo1, latitud, valor: 42.605556, Campo2, Longitud, valor: -5.5, campo3: país, valor: España, campo4: ciudad, valor: León 100.

vi. **Actualización en tiempo real:** esta dimensión mide la frecuencia de actualización de los conjuntos de datos. Esta frecuencia podría variar en función de la naturaleza del conjunto de datos. Por ejemplo, la información de transporte podría aceptarse como en tiempo real si es

de menos de 5 segundos de retardo, en otras cuestiones, periodos más largos podrían ser válidos.

- **Nivel 1: Semanas:** cuando el periodo de actualización del conjunto de datos es mayor que una semana, (**peso=15%**). Ej.: La diferencia entre una versión y la siguiente es más de 7 días. Mensual, anual.
- **Nivel 2: Días:** Cuando el periodo de actualización oscila entre 1 y 7 días, (**peso=40%**). Ej.: La diferencia de una versión y la siguiente es más de 1 día pero inferior a 7 días. Liberación semanal.
- **Nivel 3: Horas:** cuando el periodo de actualización oscila entre 1 y 24 horas, (**peso=70%**). Ej.: liberación diaria.
- **Nivel 4: Minutos:** el periodo de actualización oscila entre 1 minuto y 1 hora, (**peso=90%**). Ej.: la liberación por hora.
- **Nivel 5: Segundos:** el periodo de actualización es inferior a 1 minuto, (**peso=100%**). Ej.: periodo de actualización medido en segundos.

La tabla 5 resume las seis dimensiones anteriormente descritas. En las columnas están ubicadas las dimensiones, cada una con una breve descripción y peso asignado. En las filas están ubicados los niveles de cada dimensión, la dimensión '*estándares técnicos*' tiene solamente 4 niveles, las cinco dimensiones restantes todas tienen 5 niveles.

	Estándares técnicos		Acceso		Legal		Modelo de datos		Geolocalización		Actualización	
Nivel	Descripción	Peso	Descripción	Peso	Descripción	Peso	Descripción	Peso	Descripción	Peso	Descripción	Peso
1	Cerrado no reutilizable (pdf image, doc)	10%	Sin acceso web o solicitud manual	0%	Copyright	0%	Sin modelo publicado	15%	No hay información geográfica	15%	Mayor a una semana (mensual, semestral, anual)	15%
2	Cerrado reutilizable (shp, xls, xlsx con macrofórmula)	35%	Acceso web con registro	10%	Uso privado	10%	Modelo con campo de datos	35%	Campo de texto simple (Id propio)	30%	Días, 1 a 7 días (liberación semanal)	40%
3	Abierto (csv, txt, odb, odt, wms, xls y xlsx sin macrofórmula)	60%	Acceso directo vía web con url única	50%	Reutilización no comercial ( CC-BY-NC-ND 4.0 CC-BY-NC 4.0 CC-BY-NC-SA 4.0)	25%	Modelo de datos propio (con ontología pública y vocabularios disponibles)	50%	Campo de texto complejo	50%	Horas, 1 a 24 horas (liberación diaria)	70%
4	Abierto con metadatos (rdf, rss, json, xml)	100%	Acceso vía web con parámetros	90%	Reutilización comercial ( CC-BY-ND 4.0 CC-BY-SA 4.0)	90%	Modelo local de datos abiertos	90%	Coordenadas (sistema de coordenadas)	90%	Minutos, 1 minuto a 1 hora (liberación por horas)	90%
5	-----	-----	Acceso completo (API o lenguaje de consulta SPARQL)	100%	Sin restricciones o solo atribuciones ( CC-BY 4.0)	100%	Modelo de datos abierto	100%	Información geográfica completa	100%	Segundos, menor por minuto	100%

Tabla 5. Dimensiones, niveles y pesos de la métrica Meloda 4.0

Para aplicar esta métrica a un conjunto de datos en particular hay que seguir los pasos que se muestran en la figura 22.

### MELODA 4.01 Assessment process

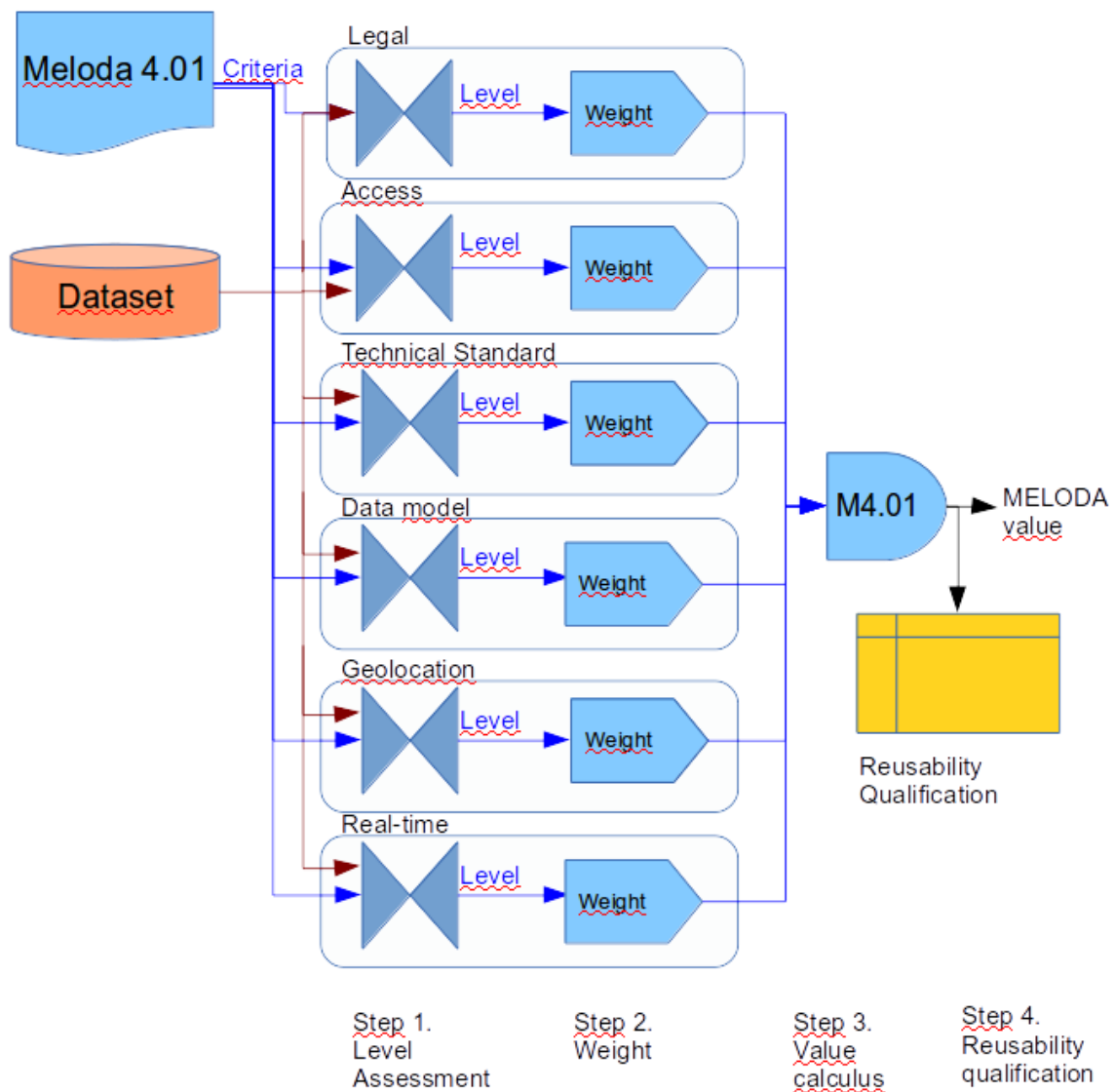


Figura 22: Procesos de Meloda 4.0 (Alberto Abella, 2016)

**Paso 1:** Analizar el conjunto de datos desde el punto de vista de cada una de las dimensiones explicadas anteriormente y asignarles el nivel que les corresponde.

**Paso 2:** Asignar el peso correspondiente, según el nivel asignado en el paso anterior.

**Paso 3:** Una vez obtenido los pesos para el conjunto de datos en cada una de las dimensiones, se calcula el valor Meloda como cien veces la raíz cúbica del

producto de los pesos de cada dimensión. Como se indica en la siguiente fórmula:

$$\text{Valor Meloda} = 100^3 \sqrt{\prod p_i} \quad (4)$$

$p_i$  : Peso de la dimensión  $i$  (%)

$i$ : [1, 2, 3, 4, 5, 6]

**Paso 4:** el resultado obtenido de aplicar la fórmula 4 será un valor numérico comprendido entre 0 y 100, cuyo valor se compara con la columna de rangos de la Tabla 6 y se le asigna la calificación de reutilización correspondiente para el conjunto de datos.

Rangos	Calificación
0-25	Inadecuado para reutilización
25-50	Posible reutilización básica
50-75	Reutilización avanzada pero con alguna característica mejorable
75-100	Reutilización avanzada

Tabla 6. Rangos de clasificación de la métrica Meloda 4.0 (Alberto Abella)

### 3.3. APLICACIÓN DE LA MÉTRICA MELODA

Para una mejor comprensión de esta métrica, a continuación se realiza una aplicación a conjuntos de datos medio ambientales del portal de datos abiertos del Ayuntamiento de Madrid[39].

#### ➤ CALIDAD DEL AIRE: DATOS DIARIOS AÑOS 2001 A 2016

Este conjunto de datos [40] contiene información recogida por las estaciones de control de calidad del aire, datos diarios por anualidades de 2003 a 2016. En el año en curso la información se actualiza mensualmente, se guarda en un fichero con formato txt. Las condiciones generales de licenciamiento permiten la reutilización de los documentos para fines comerciales y no comerciales sin restricciones. Las estaciones se ubican por código de estación, no por coordenadas. Por ser un fichero de extensión .txt no se permite la consulta SPARQL.

De acuerdo a lo descrito en el párrafo anterior y siguiendo el procedimiento explicado en la figura 22, analizamos el conjunto de datos obteniendo los siguientes resultados:

Dimensión legal: Nivel 5 → peso=100% (uso comercial y no comercial sin restricciones)

Estándar: Nivel 3 → 60% (abierto .txt)

Acceso: Nivel 4 → 90% (acceso directo web con parámetros)

Modelo de datos: Nivel 4 → 90% (modelo local de datos abiertos)

Geolocalización: Nivel 2 → 30% (Id propio)

Tiempo Real: Nivel 1 → 15% (actualización mensual, >7 días)

$$\text{Meloda} = 100 \sqrt[3]{(1 \times 0.6 \times 0.9 \times 0.9 \times 0.3 \times 0.15)} = 27.96$$

El valor obtenido (27.96) se encuentra en el rango de 25-50 de la tabla 6, por lo que la calificación para este conjunto de datos es “*posible reutilización básica*”.

#### ➤ **CALIDAD DEL AIRE: DATOS EN TIEMPO REAL**

Éste conjunto de datos [41] es similar al anterior, con la diferencia que los datos se actualizan a cada hora (entre los minutos 20 y 30 de cada hora)

Legal: Nivel 5 → 100% (uso comercial y no comercial sin restricciones)

Estándar: Nivel 4 → 100% (abierto .txt y RDF)

Acceso: Nivel 4 → 90% (acceso directo web con parámetros)

Modelo de datos: Nivel 4 → 90% (modelo local de datos abiertos)

Geolocalización: Nivel 2 → 30% (Id propio)

Tiempo Real: Nivel 4 → 90% (la liberación es por hora)

$$\text{Meloda} = 100 * \sqrt{\sqrt{\sqrt{\sqrt{(100\% * 100\% * 90\% * 90\% * 30\% * 90\%)}}}} = 82.69$$

El valor obtenido (82.69) se encuentra en el rango de 75-100 de la tabla 6, por lo que la calificación para este conjunto de datos es “*Reutilización avanzada*”.

## ➤ PARQUES Y JARDINES

Es un conjunto de datos [42] que contiene datos disponibles en formatos CSV, GEO, XML. Con frecuencia de actualización anual, con tipo de licencia de libre reutilización con fines comerciales y no comerciales sin restricciones, con información geográfica georreferenciada en sistema de coordenadas, con la estructura de conjunto de datos explicada.

De acuerdo a lo descrito en el párrafo anterior y siguiendo el procedimiento explicado en la figura 22, analizamos el conjunto de datos obteniendo los siguientes resultados:

Legal: Nivel 5 → 100% (uso comercial y no comercial sin restricciones)

Estándar: Nivel 4 → 100% (abierto con metadatos, xml, RDF, GEO, CSV)

Acceso: Nivel 4 → 90% (acceso directo web con parámetros)

Modelo de datos: Nivel 4 → 90% (modelo local de datos abiertos)

Geolocalización: Nivel 5 → 100% (completa)

Tiempo Real: Nivel 1 → 15% (anual)

Medola =  $100 * \sqrt{\sqrt{\sqrt{\sqrt{(100\% * 100\% * 90\% * 90\% * 100\% * 15\%)}}}}} = 49.53$

El valor obtenido (49.53) se encuentra en el rango de 25-50 de la tabla 6, por lo que la calificación para este conjunto de datos es, "*posible reutilización básica*".

### 3.4. APORTACIONES Y COMENTARIOS

A la hora de publicar los conjuntos de datos es importante tener en consideración la norma y las métricas que hemos estudiado anteriormente en este capítulo porque nos ayudan a cumplir aspectos técnicos y legales de tal manera que aseguran o predicen el grado de reutilización del conjunto de datos, mas no miden realmente cuánto se está utilizando ese conjunto de datos. Por ello he visto que es necesario considerar otros aspectos también importantes como:

- Medir el índice de reutilización real del conjunto de datos. Esto en base al número de aplicaciones que hacen uso de los datos, enlaces a los datos (linked data) desde otros sitios web o número de descargas del conjunto de datos por periodo de tiempo.
- Respecto a la frecuencia de actualización de los conjuntos de datos, consideramos que la aplicabilidad de la métrica debería variar en función de la naturaleza del conjunto de datos, o aplicarse en función de un índice de actualización de datos según la frecuencia con que éstos sean demandados. Esto es porque existen conjuntos de datos que no requieren que la frecuencia de actualización sea necesariamente en tiempo real. Por ejemplo, los datos de comisarías municipales, estaciones de trenes, parques, entre otros. En este tipo de conjunto de datos la actualización en tiempo real no tiene sentido porque no cambian constantemente.

Se propone como trabajo futuro, realizar una métrica que considere los indicadores para evaluar cuánto se está utilizando o reutilizando realmente un conjunto de datos. Como se indica en la tabla 7.

Tabla 7: Propuesta de indicadores para medir la reutilización real de datos abiertos

Ciudad	Nº Datasets publicados	5 DATASETS MÁS IMPORTANTES			ACCESIBILIDAD		Nº Referencias en web externa (Linked Data)	Nº Apps generadas	Índice de frecuencia de actualización
		Dataset	F. Creación	Nº Descargas	Formato	Tipo web			
MADRID	222 (06-12-2015 23:55)	1. Calidad del aire, datos en tiempo real	09/04/2015	109588	TXT				
		2. Tráfico: Intensidad del tráfico. Mapa de tramas	08/08/2014	40301	KML				
		3. Actividades Culturales y de Ocio Municipal en los próximos 100 días	24/06/2014	19872	CSV, RDF, GEO,XML				
		4. Puntos limpios (fijos y móviles)	12/03/2014	16554	CSV, RDF, GEO,XML				
		5. Puntos de interés de transporte para navegadores (estaciones de metro, cercanías, autobuses, taxi)	12/03/2014	7080	RDF, GPX, KML				
ARAGÓN	326 (06-12-2015 06:48)	1. Servicio de descarga cartográfica e:1/1000 por municipio	27.10.2015	2734	RDF, DGN, DWG, DXF, XML				
		2. Servicio de descarga cartográfica e:1/5000 y e:1/10000	27.10.2015	2734	RDF, DGN, DXF, PNG, XML				
		3. Calendario de festivos en comunidad de Aragón 2015	26.11.2015	1075	RDF, XLS, XML, CSV, JSON, ICS				
		4. Servicio de descarga cartográfica e:1/1000 por localidad	10.09.2015	684	RDF, DGN, DWG, DXF, XML				
		5. Terrenos Cinegéticos (Cotos de caza)	13.11.2015	563	RDF, SHP, KMZ, GML, GEOJSON, XML				

## CAPÍTULO 4: APLICACIÓN DE LA NORMA UNE 178301:2015

En este capítulo aplicaremos la Norma UNE 178301:2015 a cuatro casos. En el primer caso, se aplica la norma sobre el trabajo realizado por un grupo de alumnos de la Facultad de Informática de la Universidad Complutense de Madrid en el Ayuntamiento de Madrid durante los años 2012 y 2013. Para fines de este trabajo, representa a un Ayuntamiento cualquiera con primeras iniciativas de datos abiertos, en adelante nos referiremos a este caso como “versión beta”. Luego, se aplica la norma al portal de datos abiertos del Ayuntamiento de Madrid en su estado actual. A continuación, se aplica al portal de datos abiertos del Ayuntamiento de Zaragoza y finalmente al portal de datos abiertos del Ayuntamiento de Barcelona.

Se ha realizado un análisis cualitativo de los portales estudiados, basado en la observación por medio de la navegación del portal, y en muchos casos registro a registro, la identificación de formatos, frecuencia de actualización, licenciamiento, modos de búsqueda, categorización, vinculación con otros datos, ampliación, fomento de la reutilización, y otros que considera la norma UNE 178301:2015.

A continuación se muestran los resultados de aplicar la tabla 4 (Resumen estructurado de la norma UNE 178301:2015 de AENOR) para cada métrica. Para una mejor comprensión estos resultados los muestro agrupados por dominios y dentro de estos por sus dimensiones correspondientes. Después, en la tabla 48 se muestra el resultado de los valores obtenidos al aplicar la fórmula 1 a cada métrica. Cabe indicar que los datos y resultados mostrados en este trabajo son netamente con fines de investigación y aprendizaje, por lo que no compromete a las instituciones en cuestión. Sin embargo, en la parte final de este capítulo existe un apartado de recomendaciones que los ayuntamientos podrían tener en cuenta de forma voluntaria.

La norma se ha publicado en julio 2015, fecha posterior a la publicación de los portales de datos abiertos de los ayuntamientos en estudio, por lo que me ha

parecido interesante realizar una evaluación a estos ayuntamientos para conocer su nivel de cumplimiento de la norma.

## **4.1. PRIMERA INICIATIVA DE DATOS ABIERTOS**

En esta sección se aplica la norma sobre el trabajo realizado por un grupo de alumnos de la Universidad Complutense de Madrid sobre datos medioambientales del Ayuntamiento de Madrid. Dichos alumnos hicieron una propuesta de open data en los años 2012 y 2013, ya que en ese momento el Ayuntamiento no contaba con un portal de datos abiertos. La evaluación del cumplimiento se ha realizado según la información a la que se ha tenido acceso, memoria descriptiva[1] y la información del servidor, proporcionado por el grupo de alumnos.

### **4.1.1. DOMINIO ESTRATÉGICO**

#### **A. Dimensión estratégica**

Aunque tiene los objetivos claros y bien definidos de lo que se quiere lograr, no se cuenta con un plan estratégico de apertura de datos. No se ha definido formalmente un responsable político de apertura de datos. No se tiene una carta de servicios en la que se indiquen los derechos de los ciudadanos y los compromisos de calidad y responsabilidad de los gestores públicos. Se consideran algunos análisis de costes en la primera fase del proyecto, puesta en marcha, como recursos técnicos y humanos. La tabla 8 muestra los resultados de los niveles asignados a cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Estrategia	0
Liderazgo	0
Compromiso de servicio	0
Sostenibilidad económica	1

Tabla 8: Niveles asignados a la dimensión estratégica para el caso beta

### **4.1.2. DOMINIO LEGAL**

#### **A. Dimensión legal**

Estas primeras iniciativas de datos abiertos se ejecutaron sin conocimiento de la normativa interna de datos abiertos, pero sí con la finalidad de cumplir la normativa externa de obligatorio cumplimiento, Ley 37/2007, de 16 de noviembre y la Norma Técnica de Interoperabilidad de la RISP(Reutilización

de la Información del Sector Público). El documento o memoria del trabajo describe los tipos de licenciamiento, pero no se han definido ni adoptado formalmente licencias específicas para sus conjuntos de datos. La tabla 9 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

Métrica	nivel
Normas externas e internas	1
Condiciones de uso y licenciamiento	0

Tabla 9: Niveles asignados a la dimensión legal para el caso beta.

### 4.1.3. DOMINIO ORGANIZATIVO

#### A. Dimensión organizativa

Existe un equipo de trabajo que cuenta con iniciativa y capacidad necesaria para llevar adelante el proyecto. Sin embargo, no está reconocido oficialmente como unidad o área responsable. EL organismo ha creado un inventario de la información reutilizable, sin embargo, éste no está publicada y está enmarcada específicamente en los datos de medio ambiente de la ciudad. El organismo tiene bien definida la prioridad de los datos a publicar, dicho análisis se hizo basándose solo en las opiniones internas del equipo y en la relevancia del proyecto en sí misma. La tabla 10 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

Métrica	nivel
Unidad responsable	0
Equipo de trabajo y capacitación	1
Inventario	1
Prioridad	2

Tabla 10: Niveles asignados a la dimensión organizativa para el caso beta.

#### B. Dimensión medición

El equipo de trabajo no tiene establecido un plan formal de evaluación del proyecto. El organismo tiene un mecanismo de medición del uso e impacto de los datos, tal es el caso que se ha desarrollado una aplicación a partir de sus datos publicados para potenciar e incrementar su uso, sin embargo, el seguimiento de los mismos no se realiza regularmente. La tabla 11 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Medición del cumplimiento del proceso	0
Medición del uso e impacto	1

Tabla 11: Niveles asignados a la dimensión medición para el caso beta.

#### 4.1.4. DOMINIO TÉCNICO

##### A. Dimensión disponibilidad

Se tiene implementado mediante CKAN un catálogo propio de datos cumpliendo minoritariamente la legislación vigente. Algunos de los conjuntos de datos publicados en su catálogo están presentes en el Catálogo de Información Pública. Por otro lado, no todos los conjuntos de datos cuentan con todos los metadatos obligatorios. Según la Norma Técnica de Interoperabilidad de Reutilización de recursos de la información diferencia entre dos tipos de documentación o metadatos: -obligatoria: título, descripción, temática, publicador, formato y URL de acceso, - y opcionales: etiquetas, condiciones de reutilización, fecha de actualización, cobertura geográfica. Existe un servicio de búsqueda de conjuntos de datos por texto y taxonomías gracias a la herramienta Solr de CKAN. No se realiza un muestreo para comprobar que los datos están disponibles y son correctos. Las URI siguen el patrón establecido por la legislación vigente. La tabla 12 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Catálogo	1
Presencia en el Catálogo de Información Pública (CIP)	1
Conjunto de datos documentados	0
Categorización y búsqueda	2
Disponibilidad	0
Referencias persistentes y amigables	1

Tabla 12: Niveles asignados a la dimensión disponibilidad para el caso beta.

##### B. Dimensión acceso

La accesibilidad a los conjuntos de datos está garantizada, de manera que no se requiere realizar ningún trámite administrativo, registro de usuario ni pago de tasa alguna para acceder a sus datos. Existe la opción de descarga de datos en forma de ficheros y además se dispone de APIs específicas para algunos de los conjuntos de datos de medioambiente. La tabla 13 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Accesibilidad /No discriminación	3
Gratuidad	3
Sistemas de acceso	1

Tabla 13: Niveles asignados a la dimensión acceso para el caso beta.

### **C. Dimensión calidad de datos**

Los conjuntos de datos publicados son primarios pero no se ha evaluado la publicación de máximo nivel de detalle y no se explica y publica el método seguido de agregación. Los conjuntos de datos publicados son completos, se conoce la estructura y tipo de los datos. Asimismo, se han realizado validaciones para verificar que los datos son técnicamente correctos a nivel de sintaxis, codificación y normalización. Por ser datos geográficos los estándares son reconocidos por otras organizaciones. Están publicados como datos abiertos georreferenciados con coordenadas pero no cuentan con los reglamentos y guías técnicas de metadatos geográficos de la Infraestructura de Datos Espaciales de España. Los datos no están enlazados con otros internos ni externos al organismo. La tabla 14 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Datos primarios	1
Datos completos	1
Datos documentados	2
Datos técnicamente correctos	3
Datos georreferenciados	1
Datos enlazados	0

Tabla 14: Niveles asignados a la dimensión calidad de datos para el caso beta.

### **D. Dimensión actualización**

El equipo de trabajo sube ficheros al sistema open data que ha implementado. Los procesos de actualización son manuales, con una frecuencia de actualización en la publicación posterior al momento de actualización interna. El equipo prevé como trabajo futuro incluir en el sistema otros conjuntos de datos de diferentes áreas o temáticas pero no especifica si planea hacerlo teniendo en cuenta las solicitudes de los ciudadanos o reutilizadores. La tabla 15 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Proceso de actualización	0
Frecuencia de actualización	0
Ampliación de conjuntos de datos	1

Tabla 15: Niveles asignados a la dimensión actualización para el caso beta.

#### **4.1.5. DOMINIO ECONÓMICO Y SOCIAL**

##### **A. Dimensión reutilización**

Del inventario considerado inicialmente, está publicado más del 50% del mismo en formatos estándares, estructurados y abiertos. No usa vocabularios para sus conjuntos de datos. La tabla 16 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Cantidad de datos publicados	3
Formato de los datos	2
Vocabularios	0

Tabla 16: Niveles asignados a la dimensión reutilización para el caso beta.

##### **B. Dimensión participación y colaboración**

No dispone de un canal específico para que los ciudadanos o reutilizadores puedan enviar opiniones, participar, colaborar o pedir la publicación de nuevos datos. Por otro lado, tampoco ha incorporado los servicios de datos abiertos al sistema de quejas. No se ofrecen recursos de ayuda o bien son escasos e inadecuados, ni se promueven actividades de fomento de la reutilización. El organismo ha desarrollado iniciativas internas de reutilización que se han concretado en una aplicación llamada “mapa de recursos” que utiliza los datos de su catálogo de datos. La tabla 17 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Transparencia, participación y colaboración	0
Resolución de quejas y conflictos	0
Fomento de la reutilización	0
Iniciativas de reutilización desarrolladas	3

Tabla 17: Niveles asignados a la dimensión participación y colaboración para el caso beta.

## 4.2. AYUNTAMIENTO DE MADRID

Los datos mostrados a continuación son el resultado de haber aplicado la norma UNE 178301:2015 al portal de datos abiertos del Ayuntamiento de Madrid, la información ha sido consultada en el mismo portal de datos abiertos.

### 4.2.1. DOMINIO ESTRATÉGICO

#### A. Dimensión estratégica

El portal de datos abiertos del Ayuntamiento de Madrid, es junto a otras iniciativas como “Decide Madrid”, “Gobierno abierto”, “Transparencia” parte de la estrategia de gobierno electrónico y transparencia de la organización. Por lo que tiene una estrategia de apertura de datos implantada en el propio organismo, pero no publica una lista indicando los organismos dependientes en los cuales está implementada dicha estrategia y en cuáles no. No se ha tenido información si el organismo ha definido formalmente un responsable político de apertura de datos por lo que no se sabe si su ejecución forma parte de las funciones del cargo o si tiene la capacidad gestora de involucrar a la alta dirección del organismo para apoyar las iniciativas de open data. Tiene un documento de acceso público de acuerdo de servicios en el que se indica los compromisos de calidad y las responsabilidades de los funcionarios públicos ante la ciudadanía [43]. No se ha podido acceder a ningún documento que indique las sostenibilidad económica del proyecto a corto, mediano ni largo plazo. La tabla 18 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

Métrica	nivel
Estrategia	1
Liderazgo	0
Compromiso de servicio	2
Sostenibilidad económica	0

Tabla 18: Niveles asignados a la dimensión estratégica, Ayto. Madrid.

### 4.2.2. DOMINIO LEGAL

#### A. Dimensión legal

El organismo ha implementado el portal de datos abiertos siguiendo la normativa externa de obligado y recomendado cumplimiento. Asimismo el

organismo ha definido sus propias normativas que complementan el mejor cumplimiento de la misma. El organismo ha definido y adoptado licencias y condiciones de uso de los datos conforme a un modelo de licenciamiento y se permite cualquier uso comercial y no comercial de los datos[44]. La tabla 19 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Normas externas e internas	3
Condiciones de uso y licenciamiento	2

Tabla 19: Niveles asignados a la dimensión legal, Ayto. Madrid.

### 4.2.3. DOMINIO ORGANIZATIVO

#### A. Dimensión organizativa

El organismo tiene definido una unidad responsable de datos abiertos formalizada oficialmente[43]. Dicha unidad tiene establecido formalmente un equipo de trabajo específico para las actividades de apertura de datos de forma estable, abarcando ámbitos de gestión, planificación y técnicos. El inventario de información reutilizable no está publicado para el conocimiento de ciudadanos y reutilizadores. El organismo prioriza la publicación de datos en base a un análisis con opiniones internas del organismo y opiniones externas (sobre todo del sector reutilizador) para lo cual tiene implementado una sección que recoge peticiones específicas sometiénolas a votación para luego valorarlas. La tabla 20 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Unidad responsable	2
Equipo de trabajo y capacitación	2
Inventario	2
Prioridad	3

Tabla 20: Niveles asignados a la dimensión organizativa, Ayto. Madrid.

#### B. Dimensión medición

No se ha podido tener acceso al plan de evaluación del cumplimiento del plan estratégico, sin embargo se puede observar que el organismo realiza evaluaciones esporádicas, sin saber si lo hace mediante un procedimiento estandarizado. La unidad responsable tiene implementado un mecanismo

de medición del nivel de acceso y uso a través del número de descargas[45] y valoraciones[46] para los conjuntos de datos publicados cuya información obtenida es de acceso público. La tabla 21 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Medición del cumplimiento del proceso	1
Medición del uso e impacto	3

Tabla 21: Niveles asignados a la dimensión medición, Ayto. Madrid.

#### **4.2.4. DOMINIO TÉCNICO**

##### **A. Dimensión disponibilidad**

El organismo tiene publicado un catálogo de datos propio cuyo acceso es mediante interfaz web[45]. La interfaz de consulta es accesible por HTML y procesable por máquinas, cumpliendo mayoritariamente con la legislación vigente. Algunos conjuntos de datos de su catálogo también están disponibles en el Catálogo de Información Pública. Los conjuntos de datos están descritos mediante metadatos obligatorios y opcionales de acuerdo a su naturaleza. El catálogo dispone de mecanismo de búsqueda multicriterio de los conjuntos de datos por taxonomías, texto, formatos, categorías, periodo de actualización, etc. Los datos están disponibles al menos el 99% del tiempo y los datos se devuelven en el formato establecido o definido. Los conjuntos de datos tienen una dirección permanente en el tiempo y de fácil acceso, incluyen información de procedencia haciéndolos persistentes y amigables. La tabla 22 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Catálogo	2
Presencia en el Catálogo de Información Pública (CIP)	1
Conjunto de datos documentados	3
Categorización y búsqueda	3
Disponibilidad	3
Referencias persistentes y amigables	3

Tabla 22: Niveles asignados a la dimensión disponibilidad, Ayuntamiento de Madrid.

##### **B. Dimensión acceso**

El organismo garantiza la no discriminación al acceso y uso de los datos, ya que la obtención de los mismos no requiere realizar ningún trámite

administrativo o un registro de usuario. Además, el acceso es gratuito y no están sujetos a pago de tasa alguna. Existe la opción de descarga de datos en forma de ficheros, el portal tiene la opción de registro de las aplicaciones realizadas, sin embargo, no muestra la implementación de las mismas. La tabla 23 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Accesibilidad /No discriminación	3
Gratuidad	3
Sistemas de acceso	2

Tabla 23: Niveles asignados a la dimensión acceso, Ayto de Madrid.

### **C. Dimensión calidad de datos**

Los conjuntos de datos son primarios (recolectan los datos de la fuente de origen, mayor nivel de desglose posible, sin hacer tratamientos de agregación, modificación o resumen). Se ha evaluado la publicación de máximo nivel de detalle pero no se explica y se publica el método seguido de segregación. Los datos reflejan la totalidad del tema, están cumplimentados todos los valores de los datos que puedan estarlo, y existe un procedimiento interno para solicitar y completar aquellos que no lo están. Se publica la documentación de la estructura de los datos, tipo de datos y los posibles valores que pueden tomar. Se realizan validaciones para asegurar que los datos son técnicamente correctos, utilizándose para el mismo tipo de dato publicado en diferentes conjuntos de datos del catálogo la misma codificación y normalización. Los datos geográficos están georreferenciados mediante coordenadas acompañados de sus metadatos geográficos. Algunos datos están enlazados a datos internos del organismo. La tabla 24 muestra los resultados de los niveles asignados a cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Datos primarios	2
Datos completos	3
Datos documentados	3
Datos técnicamente correctos	3
Datos georreferenciados	2
Datos enlazados	1

Tabla 24: Niveles asignados a la dimensión calidad de datos, Ayto. Madrid.

#### D. Dimensión actualización

La mayoría de los procesos de actualización son automáticos. De los 227 conjuntos de datos publicados, 32 (14.1%) se actualizan en tiempo real, 15 (6.6%) se actualizan diariamente, 3 (1.3%) semanalmente, 32 (14.1%) mensualmente, 5 (2.2%) trimestralmente, 11 semestralmente (11.9%), 89 anualmente (39.2%) y 26 (11.4%) sin definir la frecuencia de actualización, al 05/04/2016 18:06.

El organismo ha publicado unos conjuntos de datos iniciales que se ha ampliado por iniciativa propia y por solicitudes de los ciudadanos o reutilizadores. Se informa previa y públicamente de estas ampliaciones para recibir opiniones. La tabla 25 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

Métrica	nivel
Proceso de actualización	2
Frecuencia de actualización	2
Ampliación de conjuntos de datos	3

Tabla 25: Niveles asignados a la dimensión actualización, Ayto. Madrid.

#### 4.2.5. DOMINIO ECONÓMICO Y SOCIAL

##### A. Dimensión reutilización

La gran mayoría de datos considerados en su inventario está publicado, generalmente en formatos estándares, estructurados y abiertos. Los vocabularios utilizados son estándares, auto-descriptivos y tienen URL persistente mas no proporcionan un control de versiones y están escritos en un solo idioma. La tabla 26 muestra los resultados de los niveles asignados a cada métrica en esta dimensión.

Métrica	nivel
Cantidad de datos publicados	2
Formato de los datos	2
Vocabularios	0

Tabla 26: Métricas y niveles de la dimensión reutilización, Ayto de Madrid.

##### B. Dimensión participación y colaboración

El organismo dispone de un canal específico para enviar opiniones, participar, colaborar, o pedir la publicación de nuevos datos. Se publican para que otras personas puedan conocerlo y aportar su opinión. Se publica

el resultado del análisis realizado por el organismo o se responde a las inquietudes de los ciudadanos indicando una planificación o prioridad para su implementación[47]. No se observa que el organismo haya integrado el servicio de datos abiertos en el sistema de quejas. El organismo publica documentos y materiales, fundamentalmente divulgativos, se celebran o planifican jornadas divulgativas externas y se realiza formación interna técnica para fomentar la reutilización interna y externa. Ha desarrollado varias iniciativas de reutilización tanto internas dentro del organismo como externas con reutilizadores. La tabla 27 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Transparencia, participación y colaboración	3
Resolución de quejas y conflictos	0
Fomento de la reutilización	3
Iniciativas de reutilización desarrolladas	3

Tabla 27: Niveles asignados a la dimensión participación y colaboración, Ayto de Madrid.

### **4.3. AYUNTAMIENTO DE ZARAGOZA**

El Ayuntamiento Zaragoza tiene presencia en Internet desde 1994 y siempre ha tenido el compromiso con la ciudadanía de tener abierto un canal de participación y garantizar una información integral, actualizada y fiable, al menor coste de tiempo, sobre el Ayuntamiento y la ciudad de Zaragoza. El portal de datos abiertos de Zaragoza fue lanzado el año 2010 en cumplimiento de la Ley 37/2007, fomenta la apertura efectiva de los datos públicos que obran en su poder, facilitando la reutilización de la información por parte de la ciudadanía, las empresas y otros organismos, lo que ofrece un aumento de la transparencia de la administración, el incremento de la participación ciudadana y la posibilidad de crecimiento económico en distintos sectores dado que el uso de formatos abiertos garantiza las posibilidades de reutilización por terceros. Esta iniciativa fue reconocida como una Web de 5 estrellas por parte de Tim Berners-Lee, Director del W3C, en mayo de 2011[48]. La figura 23 muestra el portal de datos abiertos del Ayuntamiento de Zaragoza

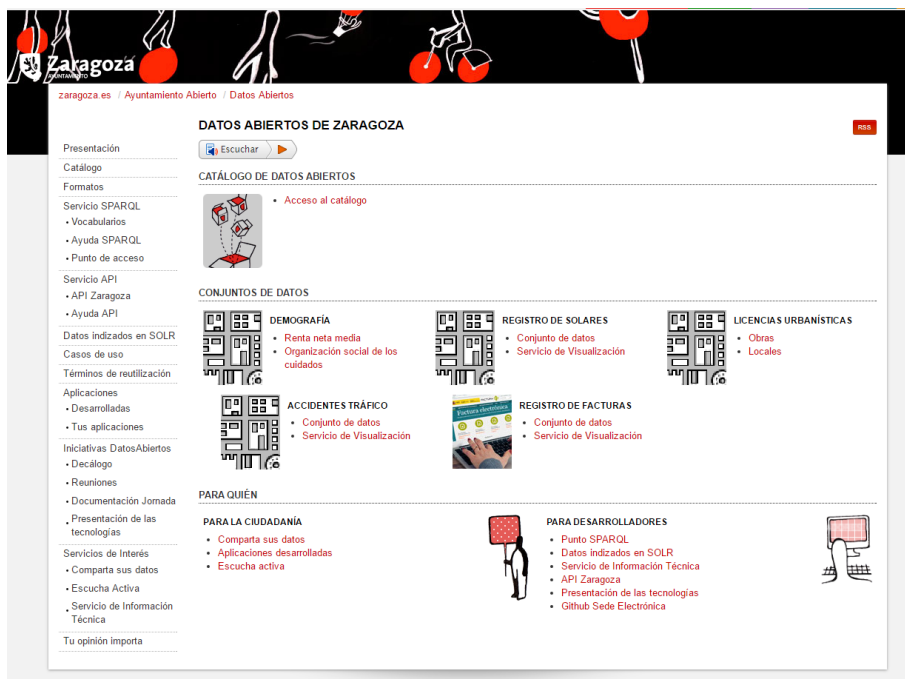


Figura 23: Portal de Datos Abiertos del Ayuntamiento de Zaragoza

A continuación evaluó el cumplimiento del portal de datos abiertos del Ayuntamiento de Zaragoza, respecto a la norma UNE 178301:2015. Está estructurado por dominios y dimensiones, según la norma.

### 4.3.1. DOMINIO ESTRATÉGICO

#### A. Dimensión estratégica

El organismo no tiene publicado la relación de los organismos dependientes sobre los cuales se haya implantado o no una estrategia de apertura de datos o no se ha tenido acceso a ello; sin embargo, tiene claro las políticas e iniciativas de datos abiertos [49]. Se puede notar que el organismo tiene definido formalmente un responsable político de apertura de datos y su ejecución o actividades forman parte de las funciones del cargo, éste tiene la capacidad gestora de involucrar a la alta dirección del organismo para apoyar las iniciativas de open data. Se han materializado las acciones concretas de apertura de datos siguiendo de forma proactiva y con carácter de continuidad. El Ayuntamiento tiene una carta de servicios de la web municipal en la que se fijan los compromisos y los servicios del Ayuntamiento[49]. No se ha podido acceder a ningún documento que indique las sostenibilidad económica del proyecto a corto, mediano o largo

plazo. La tabla 28 muestra los resultados de los niveles asignados a cada métrica de esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Estrategia	1
Liderazgo	3
Compromiso de servicio	2
Sostenibilidad económica	0

Tabla 28: Niveles asignados a la dimensión estratégica, Ayto de Zaragoza.

### **4.3.2. DOMINIO LEGAL**

#### **A. Dimensión legal**

El organismo tiene implementado un portal de datos abiertos de acuerdo a la normativa externa de obligado y recomendado cumplimiento. Asimismo, ha definido sus propias normativas que complementan el mejor cumplimiento de la misma. Ha definido y adoptado licencias y condiciones de uso de los datos conforme a un modelo de licenciamiento y se permite cualquier uso comercial y no comercial de los datos[50] y se contempla el uso de otras licencias para casos especiales. La tabla 29 muestra los resultados de los niveles asignados a cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Normas externas e internas	3
Condiciones de uso y licenciamiento	3

Tabla 29: Niveles asignados a la dimensión legal, Ayto de Zaragoza.

### **4.3.3. DOMINIO ORGANIZATIVO**

#### **A. Dimensión organizativa**

El organismo tiene definido una unidad responsable de datos abiertos[51], mas no se sabe si ésta es independiente del nivel político para garantizar la función a largo plazo. Dicha unidad tiene establecido formalmente un equipo de trabajo específico para las actividades de apertura de datos de forma estable, abarcan ámbitos de gestión, planificación y técnicos, dicho equipo de trabajo recibe una capacitación apropiada para el cumplimiento de las normativas y tecnologías asociadas a la apertura de datos. El organismo tiene un inventario de información reutilizable, que está publicados para el conocimiento de ciudadanos y reutilizadores. Prioriza la publicación de datos e base a un análisis con opiniones internas del

organismo y opiniones externas (sobre todo del sector reutilizador) para lo cual tiene implementada una sección que recoge peticiones específicas las mismas que son valoradas mediante votación, mostrándose los resultados de las mismas. La tabla 30 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Unidad responsable	2
Equipo de trabajo y capacitación	3
Inventario	3
Prioridad	3

Tabla 30: Niveles asignados a la dimensión organizativa, Ayto de Zaragoza.

## **B. Dimensión medición**

No se ha podido tener acceso al plan de evaluación del cumplimiento del plan estratégico, sin embargo se puede observar que el organismo realiza evaluaciones esporádicas en las que se miden los resultados del proyecto, pero sin un procedimiento estándar. La unidad responsable tiene implementado un mecanismo de medición del nivel de acceso y uso que se está haciendo de los datos publicados y la información obtenida es de acceso público. La tabla 31 muestra los resultados de los niveles asignados a cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Medición del cumplimiento del proceso	0
Medición del uso e impacto	3

Tabla 31: Niveles asignados a la dimensión medición, Ayto de Zaragoza.

### **4.3.4. DOMINIO TÉCNICO**

#### **A. Dimensión disponibilidad**

El Ayuntamiento de Zaragoza tiene un catálogo de datos propio accesible mediante interfaz web y procesable por máquinas, cumpliendo completamente con la legislación vigente. Los conjuntos de datos de su catálogo también están disponibles en el Catálogo de Información Pública, sin embargo no están igual de actualizados que en el propio portal. Los conjuntos de datos se describen mediante metadatos obligatorios y opcionales. El catálogo dispone de mecanismo de búsqueda multicriterio de los conjuntos de datos por taxonomías, texto, formatos, categorías, periodo

de actualización, etc. Los datos están disponibles al menos el 99% del tiempo y los datos se devuelven el formato establecido o definido. Los conjuntos de datos tienen una dirección permanente en el tiempo y de fácil acceso, incluyen información de procedencia haciéndolos persistentes y amigables. La tabla 32 muestra los resultados de los niveles asignados a cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Catálogo	3
Presencia en el Catálogo de Información Pública (CIP)	2
Conjunto de datos documentados	3
Categorización y búsqueda	3
Disponibilidad	3
Referencias persistentes y amigables	3

Tabla 32: Niveles asignados a la dimensión disponibilidad, Ayto de Zaragoza.

## **B. Dimensión acceso**

El organismo garantiza la no discriminación al acceso y uso de los datos, ya que la obtención de los mismos no requiere realizar ningún trámite administrativo o un registro de usuario. Además, el acceso es gratuito y no están sujetos a pago de ninguna tasa. Existe la opción de descarga de datos en forma de ficheros, el portal muestra la lista de las aplicaciones que se han implementado haciendo uso de los datos del Ayuntamiento. La tabla 33 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Accesibilidad /No discriminación	3
Gratuidad	3
Sistemas de acceso	3

Tabla 33: Niveles asignados a la dimensión acceso, Ayuntamiento de Zaragoza.

## **C. Dimensión calidad de datos**

Los conjuntos de datos publicados son primarios. Se ha evaluado la publicación al máximo nivel de detalle pero no se explica ni se publica el método seguido de segregación. Los datos reflejan la totalidad del tema, están cumplimentados todos los valores de los datos que puedan estarlo, y existe un procedimiento interno para solicitar y completar aquellos que no lo están. Se publica la documentación de la estructura de los datos, tipo de datos y los posibles valores que puede tomar. Los datos publicados son

técnicamente correctos, se utiliza la misma codificación y normalización para el mismo tipo de dato publicado en diferentes conjuntos de datos del catálogo. Los datos geográficos están georreferenciados mediante coordenadas acompañados de sus metadatos geográficos. Existen datos que están enlazados a datos externos del organismo (ejemplo, calidad del aire). La tabla 34 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Datos primarios	2
Datos completos	3
Datos documentados	3
Datos técnicamente correctos	3
Datos georreferenciados	3
Datos enlazados	2

Tabla 34: Niveles asignados a la dimensión calidad de datos, Ayuntamiento de Zaragoza.

#### **D. Dimensión actualización**

Los procesos de actualización mayormente son automáticos. De los 114 conjunto de datos publicados 27 (23.68%) se actualizan en tiempo real, 5 (4.39%) se actualizan diariamente, 3 (2.63%) mensualmente, 4 (3.51%) trimestralmente, 25 anualmente (21.93%) y 50 (43.86%) no se menciona el tipo de actualización, a fecha 14/04/2016 20:07.

El organismo ha publicado unos conjuntos de datos iniciales, se han ampliado por iniciativa propia y por solicitudes de los ciudadanos o reutilizadores y se informa previa y públicamente de estas ampliaciones para recibir opiniones. La tabla 35 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Proceso de actualización	2
Frecuencia de actualización	2
Ampliación de conjuntos de datos	3

Tabla 35: Niveles asignados a la dimensión actualización, Ayuntamiento de Zaragoza.

### **4.3.5. DOMINIO ECONÓMICO Y SOCIAL**

#### **A. Dimensión reutilización**

La gran mayoría de los datos considerados en su inventario están publicados, generalmente en formatos estándares, estructurados y abiertos, permitiendo la identificación única de recursos. Los vocabularios

utilizados[52] son estándares, auto-descriptivos, específicos y genéricos y tienen URL persistente. La tabla 36 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Cantidad de datos publicados	3
Formato de los datos	3
Vocabularios	3

Tabla 36: Niveles asignados a la dimensión reutilización, Ayto de Zaragoza.

## **B. Dimensión participación y colaboración**

El portal dispone de un canal para enviar y recibir opiniones, participar, colaborar, o pedir la publicación de nuevos datos. Se publican para que otras personas puedan conocerlos y aportar su opinión. Se publica el resultado del análisis realizado por el organismo o se responde las inquietudes de los ciudadanos indicando una planificación o prioridad para su implementación. El organismo tiene integrado el sistema de quejas en el servicio de datos abiertos[53]. El organismo publica documentos y materiales, fundamentalmente divulgativos, se celebran o planifican jornadas divulgativas externas y se realiza formación técnica para fomentar la reutilización de los datos interna y externa. Ha desarrollado varias iniciativas de reutilización tanto dentro del organismo como por reutilizadores externos. La tabla 37 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Transparencia, participación y colaboración	3
Resolución de quejas y conflictos	3
Fomento de la reutilización	3
Iniciativas de reutilización desarrolladas	3

Tabla 37: Niveles asignados a la dimensión participación y colaboración, Ayto de Zaragoza.

## **4.4. AYUNTAMIENTO DE BARCELONA**

El portal de datos abiertos del Ayuntamiento de Barcelona (ver figura 24) tiene como objetivo abrir aquellos datos públicos (que no están sujetos a ninguna restricción legal) para que puedan ser útiles a la sociedad, respetando siempre las condiciones de uso.

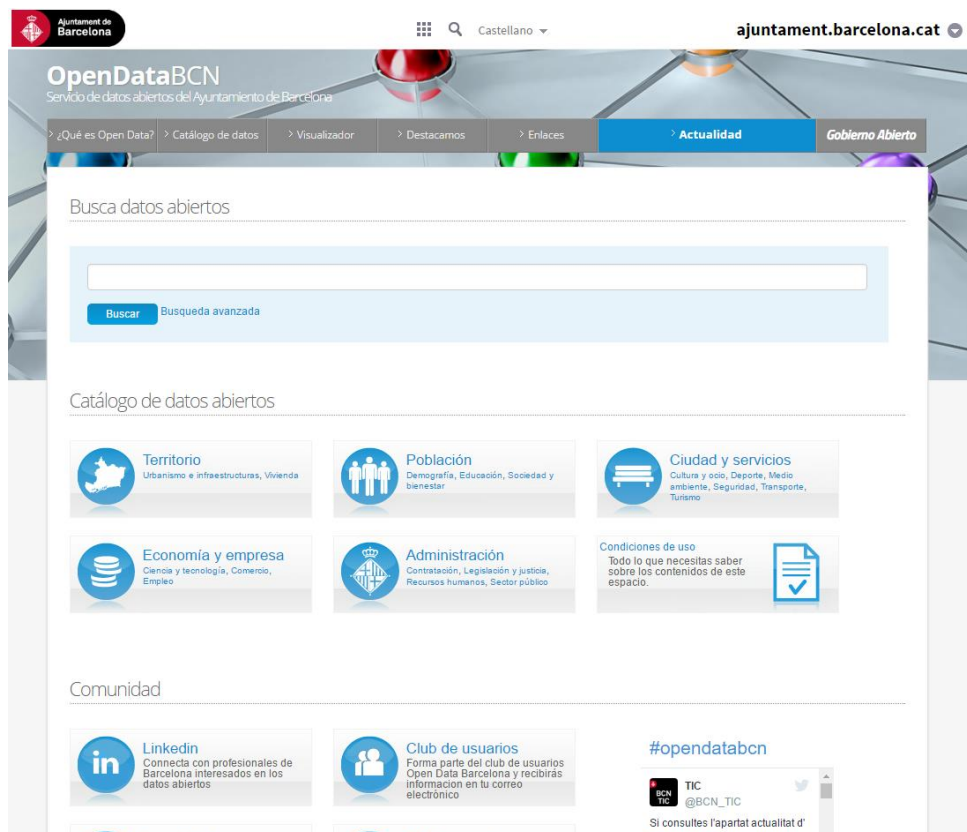


Figura 24: Portal de datos abiertos del Ayuntamiento de Barcelona

A continuación evaluó el cumplimiento del portal de datos abiertos del Ayuntamiento de Barcelona, según la norma UNE 178301:2015. Está estructurado por dominios y dimensiones, según la norma.

#### 4.4.1. DOMINIO ESTRATÉGICO

##### A. Dimensión estratégica

No se ha tenido acceso a un plan estratégico de datos abiertos, sin embargo, estudios recientes como el '*Analytical Report 4: Open Data in Cities*' [54] elaborado por la Comisión Europea y publicado en mayo del 2016 posiciona a Barcelona como la ciudad más inteligente del mundo y entre las 8 ciudades europeas más avanzadas en open data, haciendo referencia que tal avance se ha logrado gracias a la *hoja de ruta de datos abiertos*[55] a la cual no se ha tenido acceso. No se tiene conocimiento si el organismo ha definido formalmente un responsable político de apertura de datos que tenga capacidad gestora de involucrar a la alta dirección del organismo para apoyar las iniciativas de open data. Se han materializado las acciones

concretas de apertura de datos de forma proactiva y con carácter de continuidad. No se puede observar algún documento de compromiso de servicios que indique los derechos de los ciudadanos, compromisos de calidad y las responsabilidades de los funcionarios públicos ante la ciudadanía, de la misma manera, tampoco se ha podido acceder a algún documento que indique las sostenibilidad económica del proyecto ni a corto, mediano ni largo plazo. La tabla 38 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Estrategia	2
Liderazgo	1
Compromiso de servicio	0
Sostenibilidad económica	0

Tabla 38: Niveles asignados a la dimensión estratégica, Ayuntamiento de Barcelona.

#### **4.4.2. DOMINIO LEGAL**

##### **A. Dimensión legal**

El Ayuntamiento de Barcelona tiene implementado un portal de datos abiertos según la normativa externa de obligado y recomendado cumplimiento. Además, el organismo ha definido sus propias normativas que complementan el mejor cumplimiento de la normativa[56]. También, ha definido y adoptado licencias y condiciones de uso de los datos[57] conforme a un modelo de licenciamiento y se permite cualquier uso comercial y no comercial de los datos. La tabla 39 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Normas externas e internas	3
Condiciones de uso y licenciamiento	2

Tabla 39: Niveles asignados a la dimensión legal, Ayuntamiento de Barcelona.

#### **4.4.3. DOMINIO ORGANIZATIVO**

##### **A. Dimensión organizativa**

Se ha definido una unidad responsable de datos abiertos formalizada mediante una ordenanza. Formada por un equipo de trabajo específico para las actividades de apertura de datos de forma estable, abarcan ámbitos de gestión, planificación y técnicos. Reciben capacitaciones apropiadas para el

cumplimiento de las normativas y tecnologías asociadas a la apertura de datos. El organismo tiene un inventario con la información reutilizable y está publicado para el conocimiento de ciudadanos y reutilizadores. Prioriza la publicación de datos en base a un análisis con opiniones internas del organismo y opiniones externas, para lo cual tiene implementado una sección que recoge peticiones específicas sometiéndolas a votación para luego valorarlas. La tabla 40 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Unidad responsable	3
Equipo de trabajo y capacitación	3
Inventario	3
Prioridad	3

Tabla 40: Niveles asignados a la dimensión organizativa, Ayuntamiento de Barcelona.

## **B. Dimensión medición**

Se puede observar que el organismo realiza evaluaciones esporádicas del proyecto de datos abiertos, pero sin seguir un procedimiento estándar; sin embargo, no se ha podido tener acceso al plan de evaluación del cumplimiento del plan estratégico. La unidad responsable tiene implementado un mecanismo de medición del nivel de acceso y uso que se está haciendo de los datos publicados y la información obtenida es de acceso público. La tabla 41 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Medición del cumplimiento del proceso	0
Medición del uso e impacto	2

Tabla 41: Niveles asignados a la dimensión medición, Ayuntamiento de Barcelona.

### **4.4.4. DOMINIO TÉCNICO**

#### **A. Dimensión disponibilidad**

El portal de datos abiertos del Ayuntamiento de Barcelona es accesible mediante interfaz web cumpliendo completamente con la legislación vigente. Existen conjuntos de datos de Barcelona en el Catálogo de Información Pública, pero no corresponden a conjuntos de datos publicados en el portal de datos abiertos del Ayuntamiento de Barcelona. Los conjuntos

de datos están descritos mediante metadatos obligatorios y opcionales. El catálogo dispone de mecanismo de búsqueda multicriterio de los conjuntos de datos por taxonomías, texto, formatos, categorías, periodo de actualización, etc. Los datos están disponibles la mayor cantidad de tiempo posible y los datos se devuelven en el formato establecido o definido. Los conjuntos de datos tienen una dirección permanente en el tiempo y de fácil acceso, incluyen información de procedencia haciéndolos persistentes y amigables. La tabla 42 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Catálogo	3
Presencia en el Catálogo de Información Pública (CIP)	0
Conjunto de datos documentados	3
Categorización y búsqueda	3
Disponibilidad	3
Referencias persistentes y amigables	2

Tabla 42: Niveles asignados a la dimensión disponibilidad, Ayuntamiento de Barcelona.

## **B. Dimensión acceso**

El organismo garantiza la no discriminación al acceso y uso de los datos, ya que la obtención de los mismos no requiere realizar ningún trámite administrativo o un registro de usuario. Además, el acceso es gratuito y no están sujetos a pago de alguna tasa. Existe la opción de descarga de datos en forma de ficheros, el portal tiene la opción de registro de aplicaciones realizadas, sin embargo, no muestra la implementación de las mismas. La tabla 43 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Accesibilidad /No discriminación	3
Gratuidad	3
Sistemas de acceso	2

Tabla 43: Niveles asignados a la dimensión acceso, Ayuntamiento de Barcelona.

## **C. Dimensión calidad de datos**

Los conjuntos de datos son primarios. Se ha evaluado la publicación de datos al máximo nivel de detalle pero no se explica y publica el método seguido de segregación. Los datos reflejan la totalidad del tema, están cumplimentados todos los valores de los datos que puedan estarlo, y existe

un procedimiento interno para solicitar y completar aquellos que no lo están. Se publica la documentación de la estructura de los datos, tipo de datos y los posibles valores que pueden tomar. Se realiza validaciones para asegurar que los datos son técnicamente correctos, utilizan la misma codificación y normalización para el mismo tipo de dato publicado en diferentes conjuntos de datos del catálogo. No todos los datos geográficos están georreferenciados mediante coordenadas acompañados de sus metadatos geográficos. Los datos mayormente están enlazados a datos internos del organismo. La tabla 44 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Datos primarios	3
Datos completos	2
Datos documentados	3
Datos técnicamente correctos	2
Datos georreferenciados	2
Datos enlazados	1

Tabla 44: Niveles asignados a la dimensión calidad de datos, Ayuntamiento de Barcelona.

#### **D. Dimensión actualización**

Los procesos de actualización mayormente son automáticos. De los 331 conjuntos de datos publicados 12 (3.63%) se actualizan en tiempo menor a una hora, 5 (1.51%) se actualizan diariamente, 19 (5.74%) semanalmente, 44 (13.29%) mensualmente, 3 (0.91%) trimestralmente, 196 anualmente (59.21%), 29 (8.76%) de forma decenal, 2 (0.6%) se actualizan a petición de la alcaldía y 21 (3.63%) se actualizan de forma interrumpida, al 04/04/2016 23:20.

El organismo ha publicado unos conjuntos de datos iniciales que se han ampliado por iniciativa propia y por solicitudes de los ciudadanos o reutilizadores. La tabla 45 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

<b>Métrica</b>	<b>nivel</b>
Proceso de actualización	1
Frecuencia de actualización	2
Ampliación de conjuntos de datos	2

Tabla 45: Niveles asignados a la dimensión actualización, Ayuntamiento de Barcelona.

#### 4.4.5. DOMINIO ECONÓMICO Y SOCIAL

##### A. Dimensión reutilización

La gran mayoría de datos considerados en su inventario está publicado, generalmente en formatos estándares, estructurados y abiertos. Los vocabularios utilizados son estándares, auto-descriptivos y tienen URL persistente mas no proporcionan un control de versiones y están escritos en un solo idioma. La tabla 46 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

Métrica	nivel
Cantidad de datos publicados	2
Formato de los datos	2
Vocabularios	2

Tabla 46: Niveles asignados a la dimensión reutilización, Ayuntamiento de Barcelona.

##### B. Dimensión participación y colaboración

El organismo dispone de un canal específico para que los usuarios puedan solicitar algún conjunto de dato de su interés[58]. Ha integrado el sistema de quejas en el servicio de datos abiertos[59]. Publica documentos y materiales, fundamentalmente divulgativos, se celebran o planifican jornadas divulgativas externas y se realiza formación interna técnica para fomentar la reutilización interna y externa. El portal permite darse de alta como “club de usuarios” para mantenerse al día de las novedades de datos abiertos del Ayuntamiento[60]. Ha desarrollado varias iniciativas de reutilización tanto internas dentro del organismo como externas con reutilizadores. La tabla 47 muestra los resultados de los niveles asignados de cada métrica en esta dimensión.

Métrica	nivel
Transparencia, participación y colaboración	3
Resolución de quejas y conflictos	0
Fomento de la reutilización	2
Iniciativas de reutilización desarrolladas	2

Tabla 47: Niveles asignados a la dimensión participación y colaboración, Ayuntamiento de Barcelona.

## 4.5. RESULTADOS

La siguiente tabla (tabla 48) muestra los resultados obtenidos después de aplicar la norma UNE 178301:2015 a los Ayuntamientos de Madrid, Zaragoza y Barcelona y al caso de primeras iniciativas de datos abiertos (Beta).

Métrica	Peso	Beta		Madrid		Zaragoza		Barcelona	
		nivel	V <sub>p</sub>	nivel	V <sub>p</sub>	nivel	V <sub>p</sub>	nivel	V <sub>p</sub>
Estrategia	25%	0	0,00	1	8,33	1	8,33	2	16,67
Liderazgo	50%	0	0,00	0	0	3	50,00	1	16,67
Compromiso de servicio	10%	0	0,00	2	6,67	2	6,67	0	0,00
Sostenibilidad económica	15%	1	5,00	0	0,00	0	0,00	0	0,00
Normas externas e internas	30%	1	10,00	3	30,00	3	30,00	3	30,00
Condiciones de uso y licenciamiento	70%	0	0,00	2	46,67	3	70,00	2	46,67
Unidad responsable	25%	0	0	2	16,67	2	16,67	3	25,00
Equipo de trabajo y capacitación	50%	1	16,67	2	33,33	3	50,00	3	50,00
Inventario (de información reutilizable)	15%	1	5,00	2	10,00	3	15,00	3	15,00
Prioridad (de los datos recogidos en el inventario)	10%	2	6,67	3	10,00	3	10,00	3	10,00
Medición del cumplimiento del proceso	50%	0	0,00	1	16,67	0	0,00	0	0,00
Medición del uso e impacto	50%	1	16,67	3	50,00	3	50,00	2	33,33
Catálogo	30%	1	10,00	2	20,00	3	30,00	3	30,00
Presencia en el Catálogo de Información Pública (CIP)	25%	1	8,33	1	8,33	2	16,67	0	0,00
Conjunto de datos documentados	15%	0	0,00	3	15,00	3	15,00	3	15,00
Categorización y búsqueda	10%	2	6,67	3	10,00	3	10,00	3	10,00
Disponibilidad	10%	0	0,00	3	10,00	3	10,00	3	10,00
Referencias persistentes y amigables	10%	1	3,33	3	10,00	3	10,00	2	6,67
Accesibilidad /No discriminación	30%	3	30,00	3	30,00	3	30,00	3	30,00
Gratuidad	50%	3	50,00	3	50,00	3	50,00	3	50,00
Sistemas de acceso	20%	1	6,67	2	13,33	3	20,00	2	13,33
Datos primarios	20%	1	6,67	2	13,33	2	13,33	3	20,00
Datos completos	20%	1	6,67	3	20,00	3	20,00	2	13,33
Datos documentados	20%	2	13,33	3	20,00	3	20,00	3	20,00
Datos técnicamente correctos	20%	3	20,00	3	20,00	3	20,00	2	13,33
Datos georreferenciados	10%	1	3,33	2	6,67	3	10,00	2	6,67
Datos enlazados	10%	0	0,00	1	3,33	2	6,67	1	3,33
Proceso de actualización	40%	0	0,00	2	26,67	2	26,67	1	13,33
Frecuencia de actualización	40%	0	0,00	2	26,67	2	26,67	2	26,67
Ampliación de conjuntos de datos publicados	20%	1	6,67	3	20,00	3	20,00	2	13,33
Cantidad de datos publicados	30%	3	30,00	2	20,00	3	30,00	2	20,00
Formato de los datos	40%	2	26,67	2	26,67	3	40,00	2	26,67
Vocabularios	30%	0	0,00	0	0,00	3	30,00	2	20,00
Transparencia, participación y colaboración	30%	0	0,00	3	30,00	3	30,00	3	30,00
Resolución de quejas y conflictos	20%	0	0,00	0	0,00	3	20,00	0	0,00
Fomento de la reutilización	30%	0	0,00	3	30,00	3	30,00	2	20,00
Iniciativas de reutilización desarrolladas	20%	3	20,00	3	20,00	3	20,00	2	13,33
		V <sub>t0</sub>	308	V <sub>t1</sub>	678	V <sub>t2</sub>	862	V <sub>t3</sub>	668

Tabla 48. Resultados de aplicación de la norma sobre la versión beta y los ayuntamientos de Madrid, Zaragoza y Barcelona

El valor total ( $v_t$ ) se obtiene con la suma de los valores parciales ( $v_p$ ) de cada métrica. El Valor Total Máximo (puntuación 3 en todas las métricas) es de 1000. La norma UNE 178301:2015 divide este valor máximo en 5 rangos, como nuestra la tabla 49. El umbral de datos abiertos deseado para una ciudad inteligente, según dicha norma, se establece en 3.

Valor Total	0 – 200	201 – 400	401 – 600	601 – 800	801 – 1000
Indicador de datos abiertos	1	2	3	4	5

Tabla 49. Indicador de datos abiertos, UNE 178301:2015.

Aplicando los rangos definidos en la tabla 49 a los resultados del Valor Total (Vt) mostrados en la tabla 48 se concluye que:

- El valor total para la versión beta ( $v_{t0}$ ) es 308 por lo que le corresponde un indicador de datos abiertos de nivel 2.
- El valor total para el Ayuntamiento de Madrid ( $v_{t1}$ ) es 678 por lo que le corresponde un indicador de datos abiertos de nivel 4.
- El valor total para el Ayuntamiento de Zaragoza ( $v_{t2}$ ) es 862 por lo que le corresponde un indicador de datos abiertos de nivel 5.
- El valor total para el Ayuntamiento de Barcelona ( $v_{t3}$ ) es 668 por lo que le corresponde un indicador de datos abiertos de nivel 4.

En resumen se puede decir que todos los portales de los distintos ayuntamientos analizados superan el valor del umbral. El caso beta no llega a alcanzar este valor debido a que fue una primera iniciativa de datos abiertos, como su nombre lo indica.

## 4.6. RECOMENDACIONES

A continuación se detalla las recomendaciones que he considerado oportunas para que las instituciones puedan alcanzar el máximo nivel de cada métrica (nivel 3) según el instrumento de medición (UNE 178301:2015). No se hacen recomendaciones para la versión beta puesto que ahora no está operativa.

### **Ayuntamiento de Madrid**

Se recomienda al Ayuntamiento de Madrid, para cumplir con los estándares considerados en la norma UNE 178301:2015, lo siguiente:

- ✓ Definir formalmente una estrategia de apertura de datos en forma de plan estratégico o similar o si la tiene hacerla de acceso público en el portal de datos abiertos junto a la relación completa de organismos dependientes, especificando en cuáles está implantada y en cuáles no.
- ✓ Definir formalmente un responsable político de apertura de datos especificando que su ejecución o actividades forme parte de las funciones del cargo. Materializar acciones concretas de apertura de datos de forma proactiva y con carácter de continuidad.
- ✓ Establecer y hacer público un compromiso de servicios asociado a la apertura de datos, evaluar el rendimiento y publicar los resultados.
- ✓ Realizar una evaluación de sostenibilidad económica del proyecto de las fases: puesta en marcha, mantenimiento y evolución.
- ✓ Definir formalmente los procedimientos de independencia de la unidad responsable de datos abiertos del nivel político para garantizar su función a largo plazo y su capacidad de actuación en todo el organismo. considerar los principios de datos abiertos en otros proyectos, tener capacidad de respuesta a solicitudes de datos y formalizar su distribución según la normativa aplicable.
- ✓ Publicar el inventario de datos abiertos de manera que esté accesible a los ciudadanos y reutilizadores.
- ✓ Establecer un plan concreto de evaluación del proyecto periódico y sistemático considerando métricas e indicadores apropiados para evaluar el desempeño de iniciativas de datos abiertos. Establecer metas y objetivos de cumplimiento.
- ✓ Medir el cumplimiento de los objetivos propuestos e implementar medidas de mejora.

- ✓ Incluir los conjuntos de datos del catálogo propio en el Catálogo de Información Pública y realizar acciones para que estén igual de actualizados.
- ✓ Todos los conjuntos de datos deben contar con los metadatos obligatorios (título, descripción, temática, publicador, formato y URL de acceso) y los metadatos opcionales (etiquetas, condiciones de reutilización, fecha de actualización y cobertura geográfica) de acuerdo con la naturaleza del conjunto de datos a describir, según la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información.
- ✓ Establecer las URI según la Norma Técnica de Interoperabilidad de reutilización de recursos de información e implementar redirecciones de múltiples opciones para proporcionar los datos en distintos formatos a partir de una URI dada.
- ✓ Siendo que existe a la actualidad la opción de descarga de ficheros, además implementar APIs y SPARQL que permitan combinar y hacer consultas de datos procedentes de varios conjuntos de datos.
- ✓ Los conjuntos de datos primarios deben evaluarse al máximo nivel de detalle, explicar y publicar el método seguido de agregación.
- ✓ Publicar la documentación de la estructura y tipo de los datos, posibles valores que puede tomar, notas de interpretación, etc.
- ✓ Documentar las validaciones de los datos, que aseguren que la sintaxis es correcta, que la normalización y codificación es la misma para el mismo tipo de dato publicado en diferentes conjuntos de datos.
- ✓ Además de que los datos geográficos estén georreferenciados de manera explícita, incluir también la georreferenciación implícita (con metadatos).
- ✓ Enlazar al menos el 50% de los conjuntos de datos a datos externos del organismo mediante enlaces RDF.

- ✓ Documentar los procesos de actualización de los conjuntos de datos, si el proceso es automático verificar que se haya realizado de forma correcta. Al menos el 90% de los procesos de actualización deben ser automáticos.
- ✓ Hacer esfuerzos para actualizar los datos del catálogo de datos al mismo tiempo que se actualizan internamente (al menos el 90%).
- ✓ Publicar al menos el 50% de datos considerados en el inventario.
- ✓ De los vocabularios utilizados, el 30% deben ser independientes del productor de datos, estándares, auto descriptivos, con URL persistente, política de control de versiones y descritos en más de un idioma.
- ✓ Integrar el servicio de datos abiertos en el sistema de quejas de la organización.
- ✓ Fomentar la reutilización interna de los conjuntos de datos.

❖ **Ayuntamiento de Zaragoza**

Al portal de datos abiertos del Ayuntamiento de Zaragoza se le recomienda lo siguiente para cumplir con los indicadores considerados en la norma UNE 178301:2015:

- ✓ Definir una estrategia de datos abiertos o si ya la tiene hacerla de acceso público en el portal de datos abiertos. Implementarla el menos en el 80% de los organismos dependientes de la institución, indicar en cuáles está implementada y en cuáles no.
- ✓ Evaluar el rendimiento y publicar los resultados del compromiso o carta de servicios que se tiene en la web municipal.
- ✓ Evaluar la sostenibilidad económica del proyecto en las fases puesta en marcha, mantenimiento y evolución.
- ✓ Definir formalmente los procedimientos de independencia de la unidad responsable de datos abiertos del nivel político para garantizar su función a largo plazo y su capacidad de actuación en todo el organismo. Considerar los principios de datos abiertos en otros proyectos, tener

capacidad de respuesta a solicitudes de datos y formalizar su distribución según la normativa aplicable.

- ✓ Publicar el inventario de datos abiertos para que esté accesible a los ciudadanos y reutilizadores.
- ✓ Evaluar el desempeño de iniciativas del proyecto de forma periódica y sistemática, considerando métricas e indicadores apropiados. Establecer metas y objetivos de cumplimiento e implementar medidas de mejora.
- ✓ Realizar acciones para que el conjunto de datos del catálogo propio esté igual de actualizado en el Catálogo de Información Pública.
- ✓ Evaluar al máximo nivel de detalle los conjuntos de datos que son primarios, explicar y publicar el método seguido de agregación.
- ✓ Publicar la documentación de la estructura y tipo de los datos, posibles valores que puede tomar, notas de interpretación, etc.
- ✓ Documentar las validaciones de los datos, que aseguren que la sintaxis es correcta, que la normalización y codificación es la misma para el mismo tipo de dato publicado en diferentes conjuntos de datos.
- ✓ Enlazar al menos el 50% de los conjuntos de datos a datos externos del organismo mediante enlaces RDF.
- ✓ Verificar y documentar los procesos de actualización de los conjuntos de datos, que se haya realizado de forma correcta. Al menos el 90% de los procesos de actualización deben ser automáticos.
- ✓ Actualizar los datos al mismo tiempo que se actualizan internamente, al menos el 90% de los conjuntos de datos publicados.
- ✓ Publicar al menos el 50% de los datos considerados en el inventario.

#### ❖ **Ayuntamiento de Barcelona**

Se recomienda al Ayuntamiento de Barcelona, para cumplir con las métricas consideradas en la norma UNE 178301:2015, lo siguiente:

- ✓ Hacer de acceso público la hoja de ruta de datos abiertos e indicar en qué organismos dependientes a la organización se ha implementado (al menos en el 80% de los organismos dependientes).
- ✓ Definir formalmente un responsable político de apertura de datos e indicar que su ejecución o actividades formen parte de las funciones del cargo. Materializar acciones concretas de apertura de datos de forma proactiva y con carácter de continuidad.
- ✓ Crear y publicar un compromiso de servicios asociados a la apertura de datos, evaluar el rendimiento y publicar los resultados.
- ✓ Realizar una evaluación de sostenibilidad económica del proyecto de las fases puesta en marcha, mantenimiento y evolución.
- ✓ Definir formalmente los procedimientos de independencia de la unidad responsable de datos abiertos del nivel político para garantizar su función a largo plazo y su capacidad de actuación en todo el organismo. Considerar los principios de datos abiertos en otros proyectos, tener capacidad de respuesta a solicitudes de datos y formalizar su distribución según la normativa aplicable.
- ✓ Crear el inventario y publicarlo para que esté accesible a los ciudadanos y reutilizadores.
- ✓ Establecer un plan concreto de evaluación del proyecto periódico y sistemático considerando métricas e indicadores apropiados para evaluar el desempeño de iniciativas de datos abiertos. Establecer metas y objetivos de cumplimiento.
- ✓ Medir el cumplimiento de los objetivos propuestos e implementar medidas de mejora.
- ✓ Incluir los conjuntos de datos del catálogo propio en el Catálogo de Información Pública, y realizar acciones para que estén igual de actualizados.
- ✓ Todos los conjuntos de datos deben contar con el 100% de metadatos obligatorios (título, descripción, temática, publicador, formato y URL

de acceso) y el 100% de metadatos opcionales (etiquetas, condiciones de reutilización, fecha de actualización y cobertura geográfica) de acuerdo a la naturaleza del conjunto de datos a describir, según la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información.

- ✓ Establecer las URI según la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información e implementar redirecciones de múltiples opciones para proporcionar los datos en distintos formatos a partir de una URI dada.
- ✓ Además de existir la opción de descarga de ficheros, implementar APIs y SPARQL que permitan combinar y hacer consultas de datos procedentes de varios conjuntos de datos.
- ✓ Evaluar al máximo nivel de detalle los conjuntos de datos que son primarios, explicar y publicar el método seguido de agregación.
- ✓ Publicar la documentación de la estructura y tipo de los datos, posibles valores que puede tomar, notas de interpretación, etc.
- ✓ Documentar las validaciones de los datos, que aseguren que la sintaxis es correcta, que la normalización y codificación es la misma para el mismo tipo de dato publicado en diferentes conjuntos de datos.
- ✓ Enlazar al menos el 50% de los conjuntos de datos a datos externos del organismo mediante enlaces RDF.
- ✓ Documentar los procesos de actualización de los conjuntos de datos, verificar que se haya realizado de forma correcta si el proceso es automático. Al menos el 90% de los procesos de actualización deben ser automáticos.
- ✓ Actualizar los datos al mismo tiempo que se actualizan internamente, al menos el 90% de los conjuntos de datos publicados.
- ✓ Publicar más del 50% de datos considerados en el inventario.
- ✓ De los vocabularios utilizados, el 30% deben ser independientes del productor de datos, estándares, auto descriptivos, URL persistente,

proporcionan política de control de versiones y están descritos en más de un idioma.

- ✓ Integrar el servicio de datos abiertos en el sistema de quejas de la organización.
- ✓ Realizar jornadas externas y formación interna, publicar documentos y materiales divulgativos que fomenten la reutilización interna y externa.

# CAPÍTULO 5: PROPUESTA DE UNA METODOLOGÍA PARA PUBLICACIÓN DE DATOS ABIERTOS

Hasta este momento hemos analizado los avances, principios, métricas, normativas y tecnologías de datos abiertos que ya han sido o están siendo implementadas por los países que pioneros en iniciativas de datos abiertos, mientras tanto, existen países en los que aún no se han iniciado o están comenzando a implementar políticas de datos abiertos. En este capítulo realizo una propuesta de metodología para publicación de datos abiertos orientado a organizaciones públicas, responsables de la gestión de una ciudad, que aún no se han iniciado con políticas de datos abiertos. La propuesta que realizo es adaptable y amigable para este tipo de entidades, cumpliendo criterios de calidad, sostenible en el tiempo y orientada a la reutilización de datos con miras a ser una ciudad inteligente. Resaltando la importancia de la buena gestión de proyectos de TI, la metodología que propongo se basa en estándares internacionales como COBIT5 de ISACA dirigido para el gobierno y gestión de TI en empresas, ITILV3 (Information Technology Infrastructure Library) que describe de manera sistemática un conjunto de “buenas prácticas” para la gestión de los servicios de TI. También se basa en metodologías y normas de datos abiertos como UNE 178301:2015 vista en este trabajo, el open data handbook de Open Knowledge International y el proyecto europeo de datos abiertos COMSODE.

Las metodologías, normas, métricas o estándares para publicación de datos abiertos disponibles han surgido con el fin de integrar o mejorar los proyectos de datos abiertos ya existentes, es decir, se han dado cuando las iniciativas de datos abiertos ya estaban en marcha, haciendo sus

requerimientos cada vez más exigentes y difíciles de alcanzar por organizaciones públicas que aún desconocen del tema.

Este capítulo lo he estructurado de la siguiente manera. Primero explico las fases de la metodología, luego describo las actividades necesarias para cada fase, a continuación describo las directrices o recomendaciones para cada actividad. En seguida se listan los entregables (artefactos) para cada actividad y finalmente describo los roles mínimos que se deben tener para un proyecto de datos abiertos.

## 5.1. FASES DE LA METODOLOGÍA

La metodología para la publicación de datos abiertos que propongo, está compuesta de tres grandes fases, como se puede ver en la figura 25, lo que considero las tres fases más importantes del ciclo de vida de datos abiertos.



Figura 25: Fases de la metodología propuesta para la publicación de datos abiertos

(Elaboración propia)

### F01: PREPARACIÓN

Esta fase tiene como objetivo preparar a la organización, asumir compromisos y responsabilidades, analizar su estructura desde el punto de vista holístico (principio 4 de Cobit5) y proponer cambios, de ser necesario, para la apertura de datos abiertos. Realizar un análisis interno de las bases de datos que gestiona la organización y determinar qué datos estarán disponibles en el catálogo de datos abiertos y que datos no lo estarán, con su

respectivo sustento para cada caso. De acuerdo a las necesidades y características de la población, realizar un análisis externo de los conjuntos de datos que serán incluidos en el catálogo de datos. Elaborar el plan estratégico de datos abiertos y definir el alcance de la iniciativa.

## **F02: PUBLICACIÓN**

Esta fase tiene como objetivo poner en marcha el portal de datos abiertos, para lo cual se debe asegurar que se realice la apertura en base a los dos aspectos principales: apertura legal (licencias y cumplimiento de normativas) y apertura técnica (tecnología, formatos, vocabularios y metadatos) que aseguren la accesibilidad, disponibilidad y calidad de los conjuntos de datos para alcanzar el máximo nivel de reutilización posible, legible por máquinas y humanos. Se implementan políticas de seguridad y gestión de riesgos.

## **F03: MEJORA CONTINUA**

Esta fase tiene como objetivo asegurar la continuidad de la iniciativa, para lo cual se debe mantener una comunicación abierta y continua con los grupos de usuarios potenciales y con los reutilizadores, y en base a sus opiniones, sugerencias o aportaciones gestionar el cambio para mejorar los conjuntos de datos ya existentes o incluir nuevos en el catálogo de datos. Promocionar y fomentar el uso y reutilización de los datos abiertos. Evaluar el cierre de algunos conjuntos de datos, según cambios en la legislación, por ejemplo.

## **5.2. ACTIVIDADES DE LA METODOLOGÍA**

Cada fase incluye una serie de actividades que se deben realizar con el fin de aplicar mejor la metodología, en la tabla 50 siguiente describo las actividades agrupadas por cada fase.

F01: PREPARACIÓN		
	F01A01	Concienciación organizacional
	F01A02	Asumir compromisos
	F01A03	Estructura y procesos de la organización, roles y responsabilidades
	F01A04	Estimación de esfuerzos, costes y beneficios
	F01A05	Análisis y evaluación de datos que se publicarán
	F01A06	Diseño de plan estratégico de apertura de datos
F02: PUBLICACIÓN		
	F02A01	Cumplimiento de normativa externa e interna
	F02A02	Condiciones de uso y licenciamiento
	F02A03	Formatos y actualización de datos
	F02A04	Accesibilidad, disponibilidad y gratuidad
	F02A05	Descripción de datos y de conjuntos de datos
	F02A06	Vocabularios y ontologías
	F02A07	Implementación y configuración de herramientas de TI
	F02A08	Gestión de la seguridad y riesgos
F03: MEJORA CONTINUA		
	F03A01	Plan de comunicaciones, promoción del uso y reutilización de datos
	F03A02	Centro de servicios y atención al usuario
	F03A03	Evaluación del impacto
	F03A04	Plan de mantenimiento, ampliación y cierre de conjuntos de datos

Tabla 50: Fases y actividades de la metodología propuesta

## DIRECTRICES DE LA METODOLOGÍA

Las fases y actividades mencionadas anteriormente describen los procedimientos a tener en cuenta para comenzar con una iniciativa de datos abiertos; sin embargo, dejarlo solamente ahí aún podría resultar complicado para organizaciones que no conocen la temática de datos abiertos. A continuación voy a describir mediante directrices, las recomendaciones que hay que tener en cuenta para realizar cada actividad.

### **F01A01: concienciación de la organización de la importancia de la apertura de datos.**

Siendo que, esta metodología está orientada a organizaciones públicas que aún no se han iniciado en políticas de datos abiertos, es normal, que la organización en donde se implemente, los funcionarios desconozcan el tema y la importancia que conlleva implementarla. Por lo tanto, se recomienda realizar la presentación oficial del proyecto indicando las expectativas de participación y necesidades específicas. Jornadas de sensibilización, con el

objetivo es dar a conocer la iniciativa de datos abiertos con mayor detalle y concienciar sobre su importancia en la labor de servicio público. Identificación de colaboradores mediante la comunicación formal de la iniciativa a los responsables de cada área y organismo. Realizar sesiones de trabajo tipo workshop, en las que se comparten experiencias y detectan problemas comunes para darles respuesta. Formación específica en cuanto a las cuestiones organizativas, procedimentales y técnicas del Open data.

**F01A02: compromisos asumidos por la organización para la apertura de datos.**

La organización deberá firmar un documento de compromiso de servicios en el que se indique la voluntad política y administrativa para abrir conjuntos de datos con criterios de calidad y se indiquen los derechos que asisten a los ciudadanos, de acuerdo a los documentos legales descritos en la **F02A01**.

**F01A03: análisis de la estructura y procesos de la organización. Asignación de roles y responsabilidades.**

Crear formalmente la unidad responsable de datos abiertos, establecer formalmente sus funciones y responsabilidades e incluirla formalmente en el organigrama de la entidad, con procedimientos independientes del nivel político para garantizar la continuidad del proyecto a largo plazo, dotarla de capacidad de actuación con todas las demás áreas de la entidad.

Identificar las áreas de la entidad y determinar responsabilidades en las unidades que tengan a su cargo información útil a ser incluida en el catálogo de datos abiertos.

**F01A04: estimación de esfuerzos, costes y beneficios de la apertura de datos.**

Esta actividad responde a la pregunta ¿Y cuánto nos costará implementar un portal de datos abiertos? En esta etapa se debe analizar el estado actual de la infraestructura de TI, es importante resaltar que los costes por procesos para la recogida de datos son responsabilidad de la organización, y que los costes para poner los datos a disposición de la población son mínimos.

Sabemos que el coste está directamente relacionado con la calidad del servicio; sin embargo, la presente metodología recomienda la utilización de herramientas de software libre y abierto, que además garanticen la calidad del proyecto.

**F01A05: análisis y evaluación de datos internos y externos que se publicarán como datos abiertos.**

Realizar una lista de las bases de datos que gestiona la entidad y evaluar qué datos se pondrán a disposición del público en el portal de datos abiertos y cuáles no, sustentar debidamente los datos que no se incluirán en el portal de datos abiertos. No todos los datos que gestiona la organización deben necesariamente publicarse. Priorizar los datos más importantes.

El proceso de apertura de datos es un proceso iterativo, por lo que no se debe caer en la ambición de querer aperturar todos los datos a la vez. Conforme se vaya avanzando en el proceso de apertura se irán incluyendo nuevos conjuntos de datos según nuevas necesidades o nuevos requerimientos de los reutilizadores. “Es mejor proveer datos crudos ahora, que información perfecta en seis meses” (Open Data Book)

Otro aspecto a resaltar en esta etapa es ser conscientes que el movimiento de datos abiertos es realmente nuevo y la innovación juega un papel fundamental, por lo que, lo importante es iniciar e ir aprendiendo con la experiencia, tanto del fracaso como del éxito.

En base a las opiniones, comentarios o solicitudes de datos por parte de los usuarios, definir los conjuntos de datos que no gestiona la organización que podrían ser incluidos en el portal de datos abiertos. Evaluar la posibilidad de recolectarlos o gestionarlos y ponerlos a disposición de los usuarios. Como resultado de esta actividad se tendrá el inventario de datos abiertos, establecer una metodología de valoración para conocer qué datos son más prioritarios y por qué razón. Tener un almacén interno de datos, éste contendrá los datos internos preparados que la organización publicará más adelante e inclusive los que aún no está seguro de publicar. Puede ser un

sistema local de archivos, un servidor o la combinación de ambos, depende de los formatos elegidos o la manera de cómo se presentarán los datos a los usuarios finales. En el apartado **F02A07** se explica la manera cómo se pueden preparar los datos para que estén listos en el almacén interno de datos.

#### **F01A06: diseño del plan estratégico de datos abiertos.**

El plan estratégico de datos abiertos debe incluir una lista de los organismos dependientes de la entidad e indicar en cuáles de ellos se implantará el plan estratégico de datos abiertos y en cuáles no. Procurar implementarlo al menos en el 80% de los organismos dependientes. Debe tener asignado los roles y responsabilidades para la apertura de datos. Definir el alcance de la iniciativa de datos abiertos. La estrategia de datos abiertos debe ser complementaria con la estrategia de la organización.

Si se desea se puede incluir en el plan estratégico el plan de gestión de riesgos, el plan de seguridad y el plan de comunicaciones. Aunque cada uno puede ser un documento distinto.

El plan estratégico de datos abiertos deberá ser aprobado formalmente por la organización.

#### **F02A01: cumplimiento con la normativa externa e interna.**

Realizar un listado de acciones que faciliten el cumplimiento de las leyes, normativas, decretos y otros documentos legales a las que está sujeta la entidad pública, relacionadas con el derecho a la protección de datos personales, políticas de acceso a la información pública, reutilización de la información pública, modernización, accesibilidad y transparencia del Estado. Tener clara la normativa, ayudará a mitigar los riesgos de la publicación de datos.

Definir y publicar normativas propias en concordancia con las normativas superiores, que contribuyan o complementen su mejor cumplimiento.

#### **F02A02: condiciones de uso y licenciamiento**

En concordancia con los principios de datos abiertos, las licencias y condiciones de uso deben permitir el uso libre, abierto y sin restricciones de uso comercial y no comercial, deben ser claras, transparentes y no discriminatorias.

Antes se debe decidir si el conjunto de datos está protegido por derechos de autor o copyright según la Ley de Propiedad Intelectual del ámbito de la organización.

Se recomienda seguir las indicaciones definidas en [opendefinition.org/licenses/](http://opendefinition.org/licenses/) y [opendatacommons.org/guide/](http://opendatacommons.org/guide/)

Si es necesario, proponer su propia licencia o condiciones de uso para todo el catálogo de datos o para determinado conjunto de datos en específico.

#### **F02A03: formatos y actualización de datos**

Publicar los conjuntos de datos en formatos estructurados y abiertos, que permitan la identificación única de recursos y que sean legibles por máquinas para promover la reutilización.

Respecto a la actualización, documentar el proceso de actualización de los conjuntos de datos, procurar que se haga de manera automática y en lo posible minimizar la intervención manual de algún responsable. Sin embargo, sí se debe verificar que el proceso de actualización se ha realizado correctamente. Determinar si los datos actualizados serán publicados como fuente de datos separadas (archivos diferentes) o si se sobrescriben en la misma fuente de datos cada vez que se actualicen los datos (mismo archivo). Procurar que la frecuencia de actualización de los conjuntos de datos, en el sistema de origen y en el sistema de datos abiertos se realice al mismo tiempo.

#### **F02A04: aseguramiento de la accesibilidad, disponibilidad y gratuidad de los datos.**

Asegurar la accesibilidad y disponibilidad de los datos vía web, sin formulario de registro ni necesidad de realizar trámite administrativo, de

manera gratuita, en igualdad de derechos y oportunidades, sin ningún tipo de discriminación.

Brindar la posibilidad de descargarlos en forma de ficheros e interactuar con otros datos mediante API o SPARQL. Verificar que los enlaces del catálogo apunten a los datos, que la respuesta se dé en el formato definido.

#### **F02A05: descripción de datos y de conjuntos de datos.**

Los metadatos juegan un papel importante en la descripción de los datos, ya que almacenan los aspectos relevantes de los datos. La norma UNE 178301:2015 hace mención a dos tipos de metadatos: los requeridos (título, descripción, temática, publicador, formato y URL de acceso) y los opcionales (etiquetas, condiciones de reutilización, fecha de actualización, cobertura geográfica).

De igual modo, se debe documentar la estructura y tipo de los datos, notas de interpretación, posibles valores que puede tomar, etc.

Los datos geográficos deben ser descritos y asociados mediante metadatos geográficos[61] [62] ya sea de manera explícita (georreferenciación directa mediante un Sistema de Referencia de Coordenadas) o implícita (georreferenciación indirecta mediante datos geográficos bien definidos como calle, número, distrito postal, localidad, etc).

#### **F02A06: vocabularios y ontologías adoptados.**

Publicar los conjuntos de datos utilizando vocabularios estándares (definido por una organización internacional o nacional, utilizado por otros organismos y es independiente al organismo productor de los datos).

Respecto al catálogo de datos se puede tener el catálogo en un archivo de hoja de cálculo o un editor; sin embargo, se recomienda utilizar herramientas estándar de catalogación de conjuntos de datos como DCAT (vocabulario para datos del sector público, patentado y elaborado por un grupo de trabajo de W3C, de tipo RDF diseñado para facilitar la interoperabilidad entre los catálogos de datos publicados en la Web) [63].

Respecto a los formatos de los datos, se recomienda optar por formatos abiertos y estructurados como CSV y para optar mayores niveles (4\* o 5\*) adoptar formatos RDF.

### **F02A07: implementación y configuración de herramientas de software para la publicación de datos.**

Como se menciona en la introducción de este capítulo, la presente metodología está orientada a organizaciones públicas responsables de la gestión de una ciudad, por lo que se debe procurar tener un catálogo de datos centralizado de los datos públicos de la ciudad y no únicamente de la organización. Por ejemplo, referirse al catálogo, como catálogo de datos de la ciudad y no como el catálogo de datos del ayuntamiento o municipalidad; o catálogo de datos regional y no como catálogo de datos del gobierno regional. En ese sentido, se debe proveer canales para que diferentes organismos, públicos, privados o de la sociedad civil, puedan agregar datos en el catálogo de datos central.

Para la publicación de conjuntos de datos que requieren actualizaciones constantes se sugiere utilizar herramientas de software ETL (*Extract-Transform-Load*) para extraer (recoger los datos de los sistemas de origen), transformar (aplicar una serie de reglas o funciones a los datos extraídos de la fuente para obtener los datos y tenerlos listos para cargarlos en el destino final, realizan conversiones de formato de datos, limpieza de datos, anonimización, entre otros) y cargar (cargar los datos preparados con metadatos en el almacén interno de datos, descrito en **F01A05**) los datos. Un procedimiento de ETL prepara un conjunto de datos que están en el catálogo de datos interno para su publicación. Se puede seleccionar cualquier herramienta ETL, en esta metodología recomendamos herramientas ETL de código abierto, y dentro de ellas las más utilizadas son[64]: Talend Open Studio, Scriptella, KETL, Pentaho Data Integrator - Kettle, Jaspersoft ETL, GeoKettle, CloverETL, HPCC Systems, Jedox, Apatar ETL, Apache Falcon, entre otros. Se mencionan también algunas herramientas ETL de uso comercial: IBM InfoSphere DataStage, Microsoft SSIS, Adeptia ETL Suite,

Oracle Data Integrator Enterprise Edition, SAP Data Services, SAS Data Management, Elixir Data ETL, iWay Data Migrator.

Sin embargo, también se pueden realizar procedimientos ETL de manera manual, sin hacer uso de herramientas de software ETL, en caso de conjuntos de datos que se publican solamente una vez o con una periodicidad muy larga.

Una vez que los datos ya están preparados y listos para publicarse, se publican, para los cual se requiere una interfaz de datos abiertos, igual que en otros apartados, se recomienda usar soluciones de software libre y de código abierto, en la sección 2.7 de este trabajo se indica las principales características y diferencias de las dos principales herramientas para publicación de datos abiertos de este tipo (CKAN y DKAN). Y en los anexos se dan las indicaciones para la instalación y configuración para ambas herramientas.

#### **F02A08: gestión de la seguridad y riesgos de la apertura de datos.**

Esta tarea consiste en analizar, identificar, describir los riesgos relacionados con la publicación de datos abiertos. Analizar la posibilidad de ocurrencia de incidencias y medir el impacto. Establecer un registro de riesgos y definir para cada riesgo si se va a evitar, mitigar, transferir o aceptar.

Estos son algunos riesgos relacionados con la publicación de datos abiertos, según Kučera & Chlapek, 2014:

- Datos prohibidos por la ley o que infrinjan los derechos o libertades de otra persona.
- Datos relacionados con los secretos comerciales.
- Datos personales (relacionados con la privacidad de la persona).
- Datos detallados sobre la infraestructura (centrales eléctricas, presas, transmisores, etc.) podrían ser mal utilizados para causar daños relacionados con la seguridad de la infraestructura.

- Datos que no violen la legislación pero que podrían dar lugar a una publicidad negativa o actitud negativa de otros organismos del sector público.
- Datos primarios incorrectos e inexactos proporcionados a los organismos del sector público.
- Los usuarios pueden intencionalmente o no malinterpretar los datos (para causar escándalo, para obtener una ventaja competitiva, de causar daños a otros, etc.)
- Ausencia de los consumidores de datos, ya porque no es posible localizar el conjunto de datos o porque a nadie le resulta interesante.

Ante ello, se recomienda verificar el cumplimiento de la normativa relacionada con protección de datos personales, datos sensibles, información de seguridad nacional, como se indica en la actividad **F02A01**.

Si el conjunto de datos a publicar contiene datos personales, realizar técnicas de anonimización, como la recomendada por el “Dictamen 05/2014 sobre técnicas de anonimización” del grupo de trabajo sobre protección de datos del artículo 29, Creado por la Directiva 95/46/CE del Parlamento Europeo y del Consejo[65][66]. Existen herramientas de software de uso libre para anonimización de datos, entre las más usadas son: Data wrangler, Data tamer, Open refine, Trifacta wrangler.

El riesgo de la mala interpretación de los datos puede ser mitigado con la adecuada descripción de los conjuntos de datos mediante metadatos, descritos en la actividad **F02A05**.

El riesgo ante la falta de interés de los conjuntos de datos, puede ser mitigado mediante la aplicación del plan estratégico de comunicaciones, promoción del uso y reutilización de datos y la implementación del centro de servicios y atenciones como se describe en **F03A01** y **F03A02**.

La metodología que propongo no contempla procesos más complejos como la desanonimización de datos, integración de datos.

Como resultado de esta actividad se tendrá un plan de mitigación de riesgos en el que se contemplen las competencias, responsabilidades y acciones de mitigación de riesgos. El plan debe ser actualizado según vayan surgiendo nuevas incidencias o riesgos, según se vaya retroalimentando con buenas prácticas de mitigación de riesgos conforme a la experiencia o según se vayan agregando o quitando conjuntos de datos del catálogo.

**F03A01: plan de comunicaciones y promoción del uso y reutilización de datos.**

Establecer un canal de comunicación abierta y transparente con los usuarios y reutilizadores de los datos, mediante formularios online sobre opiniones de los conjuntos de datos ya existentes y peticiones de nuevos conjuntos de datos. Responder en el plazo establecido y con criterio coherente. Realizar actividades divulgativas y formativas de datos abiertos internas y externas a la entidad.

A nivel general, realizar campañas de comunicación para toda la ciudadanía en medios de comunicación locales, comunicados de prensa, anuncios, conferencias.

Y a nivel específico, realizar concursos, hackatons, incentivos para reutilizadores que hacen uso de nuestros datos abiertos ya sea en aplicaciones, productos, servicios o cualquier otra actividad de investigación o innovación. Decidir si realizar pocos pero grandes premios o muchos premios para alcanzar a más público. Realizar contactos con organizaciones e individuos interesados en el tema. Usar formatos libres y accesibles para hacerla más atractiva a la comunidad de reutilizadores, evitar usar documentos de MS Office. Realizar encuentros que denoten espontaneidad pero seriedad con personas del sector interesadas en el tema, en lugar de conferencias hacer desconferencias, barcamps, promover los encuentros cara a cara para conocer más a los usuarios potenciales, buscar intercambiar ideas entre la organización y el reutilizador o entre reutilizadores.

La comunicación en la web también juega un papel importante en la difusión de la iniciativa de datos abiertos. Realizar publicaciones en redes sociales como Twitter, LinkedIn, Facebook, Blogs. Buscar que las publicaciones se compartan con otros usuarios.

### **F03A02: Implementación del centro de servicios y atenciones.**

Esta actividad busca mantener o mejorar el nivel de calidad, brindar adecuada atención a los usuarios y proponer mejoras a la iniciativa de datos abiertos según la retroalimentación recibida en **F03A01**. Tomar una actitud proactiva y resolver definitivamente las incidencias que se presentan constantemente. Gestionar adecuadamente el cambio y el conocimiento a partir de las experiencias aprendidas y proponer mejoras de solución.

Implementar un centro de servicios para atención de quejas y reclamaciones de los usuarios. Registrar, atender y hacer un seguimiento de las incidencias o problemas encontrados después de la publicación, definir tiempos de respuesta. Evaluar la inclusión de las solicitudes de conjuntos de datos en el portal de datos abiertos. Asignar un número o anexo telefónico, cuenta de correo electrónico, formulario web para canalizar las peticiones de los usuarios.

Definir el Programa de Mejora del Servicio, documento interno en el que se reflejan las mejoras a realizar, con plazos y priorización de datos considerados a ser publicados, costes y responsabilidades, de cara a poder realizar posteriormente un informe que pueda ser utilizado para mostrar al usuario las mejoras realizadas y poder establecer las actuaciones futuras.

### **F03A03: Evaluación del impacto de la apertura de datos.**

Implementar en la interfaz de publicación de datos abiertos un apartado para que los usuarios puedan evaluar o solicitar mediante votación los conjuntos de datos que consideren de mayor interés. Estos resultados servirán a la organización para medir constantemente el impacto de los conjuntos de datos que más interesan a la ciudadanía. Asimismo, publicar y referenciar las

aplicaciones que se han desarrollado a partir de los conjuntos de datos publicados. Se debe medir el número de descargas de los conjuntos de datos.

#### **F03A04: plan de mantenimiento, ampliación y cierre de conjuntos de datos.**

Los conjuntos de datos que son útiles ahora pueden no ser útiles mañana, o los cambios de la legislación pueden obligar a que algún o algunos conjuntos de datos dejen de publicarse. Ante esta situación, la organización debe estar preparada para reaccionar ante cualquier escenario que conlleve el cierre de conjunto de datos.

### **5.3. ARTEFACTOS DE LA METODOLOGÍA**

Los artefactos son los documentos o entregables que se utilizan como entradas o salidas en las actividades o tareas de esta metodología.

Se ha definido los siguientes artefactos:

**AR01:** Plan estratégico de la organización.

**AR02:** Carta o compromiso de datos abiertos.

**AR03:** Estructura, roles, funciones y responsabilidades de la organización.

**AR04:** Documento que crea la unidad responsable de datos abiertos.

**AR05:** Plan estratégico de datos abiertos.

**AR06:** Inventario de datos abiertos.

**AR07:** Documento que señala las normativas externas e internas a la organización relacionadas con acceso a la información pública, datos abiertos o afines.

**AR08:** Descripción de las condiciones de uso y licencias seleccionadas o elaboradas por la entidad.

**AR09:** Registro de riesgos.

**AR10:** Plan de gestión de riesgos y seguridad.

**AR11:** Plan de comunicaciones.

**AR12:** Programa de Mejora del Servicio.

**AR13:** Informe anual de datos abiertos a cargo del encargado de la unidad responsable de datos abiertos.

A continuación describo la relación de los artefactos con las actividades, si son entradas (I), salidas (O) o ambas (I/O) de cada actividad, como se indica en la tabla 51.

	AR01	AR02	AR03	AR04	AR05	AR06	AR07	AR08	AR09	AR10	AR11	AR12	AR13
F01A01													O
F01A02		O											O
F01A03			I/O	O									O
F01A04													O
F01A05						O							O
F01A06	I				O								O
F02A01							O						O
F02A02								O					O
F02A03													O
F02A04													O
F02A05													O
F02A06													O
F02A07													O
F02A08									I	O			O
F03A01											O		O
F03A02												O	O
F03A03													O
F03A04													O

Tabla 51: Entradas y Salidas de los artefactos de acuerdo a cada actividad de la metodología propuesta.

Se observa por ejemplo, que el documento **AR09** (registro de riesgos) es la entrada de la actividad **F0208** (gestión de riesgos y seguridad) y que el documento **AR10** (plan de gestión de riesgos y seguridad) es la salida de la misma actividad. O, que el documento **AR13** (informe anual de datos abiertos) es el resultado de cada una de las actividades.

Siendo que la metodología que propongo está dirigida para entidades públicas que aún no han iniciado con datos abiertos, no pretende ser muy

engorrosa sino más bien adaptable y amigable a este tipo de instituciones, pero que sí se cumpla con criterios de calidad y seguridad. Conforme se vaya avanzando en la experiencia, se irá mejorando y adaptando nuevas metodologías.

#### **5.4. ROLES Y RESPONSABILIDADES**

La cantidad de personas a cargo de la publicación de datos abiertos puede variar según el tamaño de la organización. A continuación se indica los roles que mínimamente se debería tener:

**R01: PROFESIONAL DE TI:** Es el encargado de gestionar e implementar la tecnología de hardware y software para la publicación de datos abiertos, debe estar capacitado en normativas y herramientas que conlleven las funciones de su puesto.

**R02: EXPERTO LEGAL:** Es necesario contar con la asesoría legal para el cumplimiento de las normativas externas e internas de publicación de datos. Por ejemplo, protección de datos personales, anonimización de datos personales, derechos que asisten a los ciudadanos para acceder a los datos públicos, entre otros.

**R03: PUBLICADOR/MANTENEDOR DE DATOS ABIERTOS:** Es el encargado de asegurarse de que los conjuntos de datos están disponibles y da mantenimiento y soporte a los conjuntos de datos.

## CAPÍTULO 6: CONCLUSIONES Y TRABAJO FUTURO

En este capítulo presento a modo de resumen las conclusiones a las que he llegado tras haber realizado este trabajo.

Los datos abiertos son una herramienta muy valiosa para fomentar buenas prácticas de transparencia y democracia en una sociedad. Generan valor y contribuyen a mejorar la calidad de vida de la población.

Existe una notable diferencia entre los países líderes en datos abiertos, los que están en camino y quienes aún están comenzando. Respecto a la buena aplicabilidad de políticas públicas, eficiencia del gobierno y calidad de vida.

Si tuviera que elegir los tres ejes principales de datos abiertos, diría que giran en torno a formatos abiertos y estructurados que permitan su reutilización, licenciamiento abierto que permita la libre reutilización comercial y no comercial de los datos y accesibilidad universal y gratuita, como muestra la figura 26.

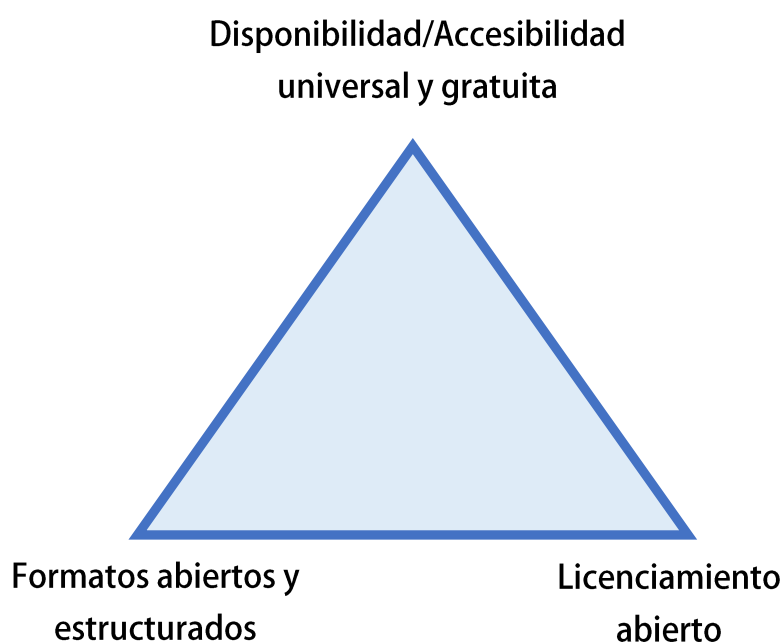


Figura 26: Los tres ejes de los datos abiertos (Elaboración propia)

Es posible reducir el coste de la implementación de sistemas de datos abiertos, utilizando tecnologías libres de código abierto tanto en el backend como el frontend.

Se ha aplicado la Norma 178301:2015 de AENOR a los portales de datos abiertos de los Ayuntamientos de Madrid, Zaragoza y Barcelona. De los cuales se concluye que los tres portales superan el umbral de datos abiertos para una ciudad inteligente con niveles 4, 5 y 4 respectivamente.

Se ha instalado y configurado dos tecnologías de código abierto para la publicación de datos abiertos y se anexa los manuales de instalación y configuración.

Se ha realizado una propuesta de una metodología para la publicación de datos abiertos, que asegura criterios de calidad y es adaptable a entidades públicas que están comenzando o aún no han comenzado a implementar datos abiertos.

Como trabajo futuro tengo pensado aplicar todo el conocimiento adquirido y la metodología propuesta para implementar políticas de datos abiertos en mi país (Perú).

Elaborar una métrica que permita medir la reutilización real de los datos abiertos, considerando los indicadores propuestos en la tabla 7.

---

## CONCLUSIONS AND FUTURE WORK

In this chapter I present to sum up the conclusions I have reached after having done this work.

Open data is a valuable tool to promote good practices of transparency and democracy in a society. They create value and help improve the quality of life of the population.

Regarding the applicability of good public policy, government efficiency and quality of life; there is a notable difference between the leading countries in open data, the ones that are on their way and the ones that are still starting.

If I had to choose the three axes of open data, I would say that they revolve around open structured formats that allow free universal accessibility and licensing for commercial and non-commercial purposes.

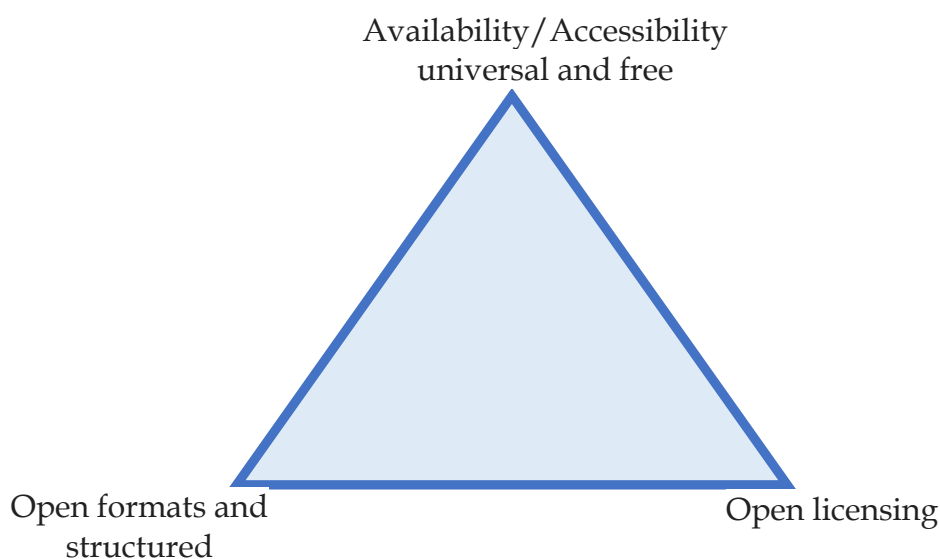


Figure 27: The three axes of open data

It's possible to reduce the cost of the implementation of open data systems, using free technologies with open codes both backend and frontend.

The 178301: 2015 norm of AENOR has been applied to open data portals of the municipalities of Madrid, Zaragoza and Barcelona. The three portals exceed the threshold of open data for an intelligent city with levels 4, 5 and 4 respectively.

Two open source technologies have been installed and set up for publishing open data, installation and configuration manuals are attached.

A methodology for publishing open data has been established, it ensures quality criteria and it is adaptable to public entities that are starting or have not yet started to implement open data.

As a future work, I have thought about to apply all the acquired knowledge and the proposed methodology to implement politics of open data in my country (Perú).

By another way, I want to elaborate a metric that can measure the real reuse of open data, considering the indicators proposed in table 7.

# Bibliografía

- [1] Á. Hernando Gavilán, A. Irizar López, and C. Rodríguez Díaz, "Sistemas open data : aplicaciones de medio ambiente," 2014.
- [2] sunlightfoundation, "Ten Principles for Opening Up Government Information - Sunlight Foundation," 2010. [Online]. Available: <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>.
- [3] "Open Knowledge: Home." [Online]. Available: <https://okfn.org/>.
- [4] D. Castro and T. Korte, "Open Data in the G8: A Review of Progress on the Open Data Charter," pp. 1-47, 2015.
- [5] "Declaration | Global Open Data Initiative." [Online]. Available: <http://globalopendatainitiative.org/declaration/>. [Accessed: 15-Feb-2016].
- [6] Juan Luis Gámez, "¿Qué es la Ley de Transparencia?," 2014. [Online]. Available: [http://www.elderecho.com/tribuna/administrativo/Ley\\_de\\_Transparencia\\_11\\_7\\_58680001.html](http://www.elderecho.com/tribuna/administrativo/Ley_de_Transparencia_11_7_58680001.html).
- [7] J. M. & F. H. (2014) Caplan, R., Davies, T., Wadud, A., Verhulst, S., Alonso, "Towards common methods for assesing open data: workshop report and draft framework," pp. 1-15, 2014.
- [8] "Open government data around the world, right now. | Global Open Data Index by Open Knowledge," 2016. [Online]. Available: <http://index.okfn.org/>.
- [9] "metodología | Índice Global de Datos Abiertos de Conocimiento Abierto." [Online]. Available: <http://index.okfn.org/methodology/>.
- [10] "ODB Global Report Third Edition," 2016.
- [11] "The Open Data Economy Unlocking Economic Value by Opening Government and Public Data Only Few Governments are Leveraging Open Data for Economic Benefits."
- [12] "OpenDataMonitor." [Online]. Available: <http://opendatamonitor.eu/frontend/web/index.php?r=dashboard%2Findex>.
- [13] Y. Li, E. Simperl, Q. Bu, and S. Feyisetan, "Open Data Monitor Project," no. 2, 2013.
- [14] "Data.gov." [Online]. Available: <https://www.data.gov/>.
- [15] "Open Government." [Online]. Available: <http://open.canada.ca/en>.
- [16] "Data | The World Bank." [Online]. Available: <http://data.worldbank.org/>.
- [17] "Open Data Catalog | The World Bank." [Online]. Available: <http://datacatalog.worldbank.org/>.
- [18] "Banco de datos del Banco Mundial | Explorar. Crear. Compartir." [Online]. Available: <http://databank.bancomundial.org/data/home.aspx>.
- [19] "Datos | eodp.common.ckan.site\_title." [Online]. Available: <http://data.europa.eu/euodp/es/data/>.
- [20] "Hogares con acceso a Internet en casa | ONTSI." [Online]. Available: <http://www.ontsi.red.es/ontsi/es/indicador/hogares-con-acceso-internet-en-casa>.
- [21] "Frecuencia de Uso de Internet | ONTSI." [Online]. Available: <http://www.ontsi.red.es/ontsi/es/indicador/frecuencia-de-uso-de-internet>.
- [22] "Iniciativas Datos Abiertos." [Online]. Available: <http://mapa.datos.gob.es/>.
- [23] "Portada | datos.gob.es." [Online]. Available: <http://www.datos.gob.es/>.
- [24] "En portada - Portal de Datos abiertos." [Online]. Available: <http://datos.madrid.es/portal/site/egob/>.
- [25] a Abella, "Reutilización de información pública y privada en España," *Rooter Anal.*

- SL, 2011.
- [26] Red.es, "ESTUDIO SECTOR INFOMEDIARIO 2014 PARTE I . SECTOR INFOMEDIARIO PÚBLICO Observatorio Nacional de las Telecomunicaciones y de la Sociedad de la Información ( ONTSI ) Entidad Pública Empresarial Red . es Secretaría de Estado de Telecomunicaciones y para la S," pp. 1-128, 2014.
- [27] A. Informe, S. Infomediario, Asedie, A. Informe, S. Infomediario, and Asedie, "Sector infomediario," pp. 1-20, 2015.
- [28] G. Boulton, "A Revolution in Open Science: Open Data and the Role of Libraries," Munich: Royal Society, 2013.
- [29] "About | Open Government Partnership." [Online]. Available: <http://www.opengovpartnership.org/about>.
- [30] RIO+20 Conferencia de las Naciones Unidas sobre el Desarrollo Sostenible, "El futuro que queremos, las ciudades," 2012.
- [31] "Informe sobre el estado de los datos abiertos en las ciudades europeas | datos.gob.es." [Online]. Available: <http://datos.gob.es/lopd?q=node/644919>.
- [32] A. de Barcelona, "BCN Smart City | Barcelona Smart City." [Online]. Available: <http://smartcity.bcn.cat/es/bcn-smart-city.html>.
- [33] N. Kroes, "EU Open Data Strategy - YouTube." [Online]. Available: <https://www.youtube.com/watch?v=rWYTT-eUtsw>.
- [34] AENOR, "Infografía - Normas y Normalización."
- [35] E. y T. Ministerio de Industria, "LEGISLACIÓN BÁSICA E INFRAESTRUCTURA PARA LA CALIDAD Y SEGURIDAD INDUSTRIAL." [Online]. Available: <http://www.minetur.gob.es/industria/es-ES/Servicios/calidad/Paginas/legislacion-basica.aspx?Faq=Normalizaci%C3%B3n#dt1>.
- [36] "5 estrellas de Datos Abiertos." [Online]. Available: <http://5stardata.info/es/>.
- [37] "meloda40v1 - MELODA," 2016. [Online]. Available: <http://www.meloda.org/meloda40v1/>.
- [38] A. Abella, "Modelling the Economic impact of information reuse in Spain," p. 51, 2011.
- [39] "Conjuntos de datos - Portal de Datos abiertos." [Online]. Available: [http://datos.madrid.es/vgn-ext-templating/v/index.jsp?vgnextoid=374512b9ace9f310VgnVCM100000171f5a0aRCRD&buscar=true&Sector3=medio-ambiente&Formato4=&Periodicidad5=&orderByCombo=CONTENT\\_INSTANCE\\_NAME\\_DECODE](http://datos.madrid.es/vgn-ext-templating/v/index.jsp?vgnextoid=374512b9ace9f310VgnVCM100000171f5a0aRCRD&buscar=true&Sector3=medio-ambiente&Formato4=&Periodicidad5=&orderByCombo=CONTENT_INSTANCE_NAME_DECODE).
- [40] A. de Madrid, "Calidad del aire: Datos diarios años 2001 a 2016 - Portal de Datos abiertos." [Online]. Available: <http://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=aecb88a7e2b73410VgnVCM2000000c205a0aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [41] Ayuntamiento de Madrid, "Calidad del aire: Datos en tiempo real - Portal de Datos abiertos." [Online]. Available: <http://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=41e01e007c9db410VgnVCM2000000c205a0aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [42] Ayuntamiento de Madrid, "Parques y Jardines - Portal de Datos abiertos." [Online]. Available: <http://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=dc758935dde13410VgnVCM2000000c205a0aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [43] Ayuntamiento de Madrid, "Acuerdos Junta de Gobierno," 2010.

- [44] Ayuntamiento de Madrid, "Condiciones de uso - Portal de Datos abiertos." [Online]. Available: <http://datos.madrid.es/portal/site/egob/menuitem.400a817358ce98c34e937436a8a409a0/?vgnextoid=b4c412b9ace9f310VgnVCM100000171f5a0aRCRD>.
- [45] A. de Madrid, "Conjuntos de datos - Portal de Datos abiertos." [Online]. Available: <http://datos.madrid.es/portal/site/egob/menuitem.9e1e2f6404558187cf35cf3584f1a5a0/?vgnextoid=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [46] Ayuntamiento de Madrid, "Los conjuntos mejor valorados - Portal de Datos abiertos." [Online]. Available: <http://datos.madrid.es/portal/site/egob/menuitem.20805ba28f8a4b1661df6c14a8a409a0/?vgnextoid=a58512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextchannel=a58512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [47] Ayuntamiento de Madrid, "Solicitudes recibidas - Portal de Datos abiertos." [Online]. Available: <http://datos.madrid.es/portal/site/egob/menuitem.400a817358ce98c34e937436a8a409a0/?vgnextoid=102612b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextchannel=102612b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>.
- [48] "Plataforma de Gobierno Abierto del Ayuntamiento de Zaragoza | OKFN Award Open knowledge." [Online]. Available: <http://premio.okfn.es/plataforma-de-gobierno-abierto-del-ayuntamiento-de-zaragoza/>.
- [49] "Ayuntamiento de Zaragoza. Carta de Servicio de la Web Municipal." [Online]. Available: <http://www.zaragoza.es/ciudad/gestionmunicipal/calidad/web/default.htm>.
- [50] Ayuntamiento de Zaragoza, "Ayuntamiento de Zaragoza | Aviso Legal." [Online]. Available: <http://www.zaragoza.es/ciudad/servicios/avisolegal.htm#condiciones>.
- [51] Ayuntamiento de Zaragoza, "Ayuntamiento de Zaragoza. Decreto de Alcaldía de 30 de diciembre de 2015, por el que se establece la organización y estructura pormenorizada de las Áreas de Gobierno del Ayuntamiento de Zaragoza." [Online]. Available: [http://www.zaragoza.es/ciudad/organizacion/admonmunicipal/detalle\\_Normativa?id=7163](http://www.zaragoza.es/ciudad/organizacion/admonmunicipal/detalle_Normativa?id=7163).
- [52] Ayuntamiento de Zaragoza, "Gobierno Abierto. Datos Abiertos de Zaragoza. Vocabularios." [Online]. Available: <http://www.zaragoza.es/ciudad/risp/vocabularios.htm>.
- [53] Ayuntamiento de Zaragoza, "Servicio de Quejas y Sugerencias." [Online]. Available: [https://www.zaragoza.es/ciudad/ticketing/verNuevaQuejaAnonima\\_Ticketing](https://www.zaragoza.es/ciudad/ticketing/verNuevaQuejaAnonima_Ticketing).
- [54] W. Carrara, W. Engbers, M. Nieuwenhuis, and E. van Steenbergen, "Analytical Report 4: Open Data Cities," no. 860, pp. 1-28, 2016.
- [55] ESMARTCITY.es, "Barcelona, entre las ocho ciudades europeas más avanzadas en Open Data," 2016. [Online]. Available: <https://www.esmartcity.es/noticias/barcelona-entre-las-ocho-ciudades-europeas-mas-avanzadas-en-open-data>.
- [56] A. de Barcelona, "OpenDataBCN Ayuda : Condiciones de uso." [Online]. Available: <https://opendatabcn.zendesk.com/forums/22079457-Condicion-de-uso>.
- [57] A. de Barcelona, "Normas de uso - Open Data - Ayuntamiento de Barcelona." [Online]. Available: <http://opendata.bcn.cat/opendata/es/data-using>.
- [58] "Nueva consulta | Consultas por internet | La web de la ciudad de Barcelona." [Online]. Available: <https://w10.bcn.es/APPS/irsconsultesWeb/continuar.executant.do?i=e&origen=>

OPEN\_DATA&detall=4512.

- [59] A. de Barcelona, "Buzón de quejas y sugerencias | El Ayuntamiento te escucha | La web de la ciudad de Barcelona." [Online]. Available: <https://w10.bcn.es/StpQueixesWEB/prepararFormularioBtoB.do?i=e&ambit=tescolta>.
- [60] "Open Data Barcelona | Club de usuarios." [Online]. Available: <http://opendata.bcn.cat/opendata/es/users-club/>.
- [61] "Geoportal IDEE." [Online]. Available: <http://www.ideo.es/web/guest/metadatos>.
- [62] AENOR, "AENOR: Norma UNE-EN ISO 19111:2009," <http://www.aenor.es/>.
- [63] W3C, "Data Catalog Vocabulary (DCAT)," 2014. [Online]. Available: <https://www.w3.org/TR/vocab-dcat/>.
- [64] "10 Open Source ETL Tools - Data Science Central," 2015. [Online]. Available: <http://www.datasciencecentral.com/profiles/blogs/10-open-source-etl-tools>.
- [65] G. D. T. S. P. D. D. A. 29, "Dictamen 05/2014 sobre técnicas de anonimización," *Wp216*, pp. 1-24, 2007.
- [66] GRUPO DE TRABAJO SOBRE PROTECCIÓN DE DATOS DEL ARTÍCULO 29, "DICTAMEN 05/2014 SOBRE TÉCNICAS DE ANONIMIZACIÓN," 2014.

# Anexos

## Anexo 1. Instalación de CKAN

La instalación está realizada en un ordenador con Sistema Operativo Ubuntu 14.04.

### ➤ Instalación de Solr 5.2.1

#### Paso 1: Instalación de Java

Solr requiere Java, por ello vamos a instalarlo:

```
$ sudo apt-get install python-software-properties
```

En lugar de utilizar los paquetes por default de jdk o jre, instalamos la última versión de Java:

```
sudo add-apt-repository ppa:webupd8team/java
sudo apt-get update
sudo apt-get install oracle-java8-installer
```

#### Paso 2: Instalar Solr

```
cd ~
sudo apt-get install wget
wget http://apache.mirror1.spango.com/lucene/solr/5.2.1/solr-5.2.1.tgz
tar xzf solr-5.2.1.tgz solr-5.2.1/bin/install_solr_service.sh --strip-components=2
sudo bash ./install_solr_service.sh solr-5.2.1.tgz
sudo service solr status
```

#### Paso 3: Crear una colección (es opcional)

```
sudo su - solr -c "/opt/solr/bin/solr create -c gettingstarted -n data_driven_schema_configs"
```

[http://your\\_server\\_ip:8983/solr](http://your_server_ip:8983/solr)

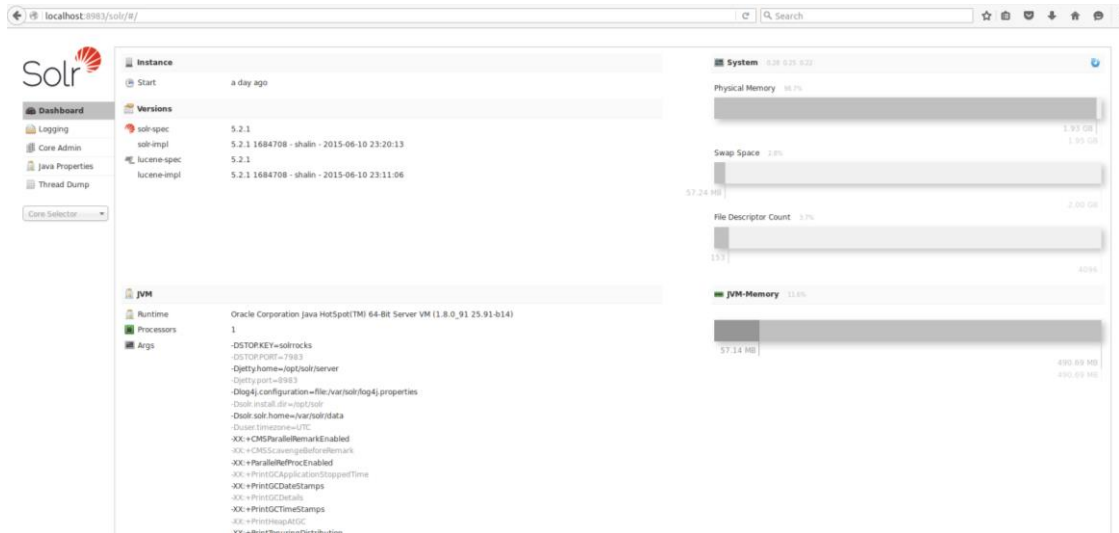


Figura 28: Primera interfaz de Solr

### ➤ Instalación de PostgreSQL

```
sudo apt-get install python-dev postgresql libpq-dev python-pip python-virtualenv
git-core solr-jetty openjdk-6-jdk
sudo -u postgres psql -l (listar bd)
```

```
sudo -u postgres createuser -S -D -R -P ckan_user (pswd:ckan)
sudo -u postgres createdb -O ckan_user ckan_db -E utf-8
sudo apt-get update
```

configurar el archivo `pg_hba.conf` en la ruta: `/etc/postgresql/9.1/main ...`

### ➤ Instalación de CKAN

```
sudo apt-get install -y nginx apache2 libapache2-mod-wsgi libpq5
wget http://packaging.ckan.org/python-ckan_2.5-trusty_amd64.deb
sudo dpkg -i python-ckan_2.5-trusty_amd64.deb
sudo a2enmod wsgi
sudo service apache2 restart
```

### ➤ Configuración de ficheros

configurar `/etc/ckan/default/production.ini`

```
cd /etc/ckan/default/
sudo chmod 777 production.ini
nano production.ini
ckan.site_id = default
ckan.site_url=http://localhost/

sqlalchemy.url=postgresql://ckan_user:ckan@localhost/ckan_db

ckan.site_id=default
solr url=http://127.0.0.1/solr

sudo ckan db init
```

## 4. Reiniciar e iniciar servicios

```
sudo service apache2 restart
sudo service nginx restart
cd /etc/default/
sudo chmod 777 jetty
nano jetty
```

colocar los valores:

```
NO_START = 0 # (línea 4)
JETTY_HOST = 127.0.0.1 # (línea 16)
JETTY_PORT = 8983 # (línea 19)
JAVA_HOME=/usr/lib/jvm/java-6-openjdk-amd64/
```

```
sudo service jetty restart (debe estar instalado el openjdk-6-jdk)
sudo mv /etc/solr/conf/schema.xml /etc/solr/conf/schema.xml.bak
sudo ln -s /usr/lib/ckan/default/src/ckan/ckan/config/solr/schema.xml
/etc/solr/conf/schema.xml
sudo service jetty restart
```

```
sudo ckan db init
sudo service apache2 restart
sudo service nginx restart
```

5..

<http://localhost/>

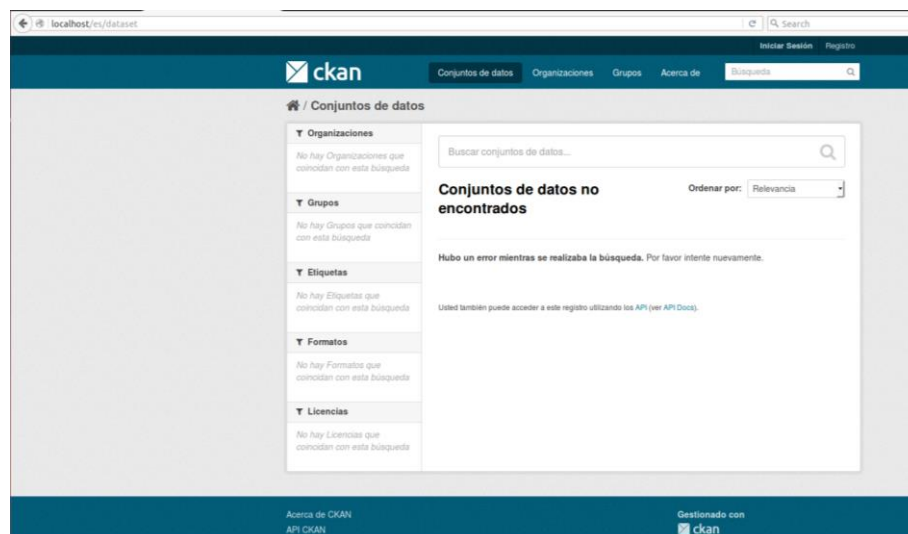


Figura 29: Página de inicio de CKAN luego de la instalación

## Anexo 2. Configuración de CKAN

Para comenzar a utilizar CKAN es necesario crear el usuario administrador del sistema.

```
. /usr/lib/ckan/default/bin/activate
cd /usr/lib/ckan/default/src/ckan
paster sysadmin add userAdmin -c etc/ckan/default/production.ini
(user: userAdmin; asignarle una clave cuando te pida)
```

Luego, agregamos o activamos los plugins necesarios en el fichero “*production.ini*” en la ruta */etc/ckan/default* en el apartado de `## Plugins settings`:

- **Plugin Stats**
- **Plugin text\_preview**
- **Plugin image\_view**
- **Plugin recline\_preview**
- **Plugin pdf\_preview**

añadir en el archivo de configuración CKAN:

```
ckan.plugins = <otros plugins> pdf_preview
```

- **Plugin DataStore (API REST)**

**crear usuario para DataStore**

```
sudo -u postgres createuser -S -D -R -P -l <nombre de usuario>
```

**crear bd para DataStore**

```
sudo -u postgres createdb -O <usuario> ><nombre base de datos>
-E utf-8
```

añadir en el archivo de configuración CKAN:

```
ckan.plugins = <otros plugins> datastore
```

```
ckan.datastore.enabled = True
```

```
ckan.datastore.write_url=postgresql://<usuario propietario
base de datos>:<contraseña>@localhost/>nombre base de datos
DataStore>
```

```
ckan.datastore.read_url = postgresql://<usuario lector creado
antes>:<contraseña>@localhost/<nombre base de datos>
```

**Permisos**

```
sudo ckan datastore set-permissions postgres
```

- **Plugin DataPusher**

Permite procesar archivos CSV y XLS y los almacena en la BD. Proceso Asíncrono

Añadir en el archivo de configuración de CKAN:

```
ckan.plugins = <resto de plugins> datapusher
ckan.datapusher.enabled = True
ckan.datapusher.url = http://0.0.0.0:8800/
ckan.datapusher.formats = csv xls
ckan.site_url = < Dominio>
```

### ➤ **Plugin FileStore**

Permite cargar archivos al gestor de contenidos

Crear el directorio donde se almacenarán los archivos:

```
sudo mkdir -p /var/lib/ckan/default
```

Añadir en el archivo de configuración CKAN “*production.ini*” en el apartado de **##Storage settings**:

```
ckan.storage_path = /var/lib/ckan/default
ckan.max_resource_size = < tamaño máximo de archivo>
ckan.max_image_size = < Tamaño máximo de imagen>
```

Establecer los permisos:

```
sudo chown www-data /var/lib/ckan/default
sudo chmod u+rwx /var/lib/ckan/default
```

## Anexo 3. Instalación de DKAN

La siguiente guía de instalación está realizado para sistema operativo CentOS 7, iniciar sesión con el usuario root.

### 1. Actualizamos el sistema operativo e instalamos las herramientas de desarrollo

```
$ yum update
$ yum groupinstall 'Development Tools'
```

### 2. Descargamos e instalamos Apache httpd, apr y apr-util

#### 2.1. APR

```
$ wget http://mirror.symnds.com/software/Apache/apr/apr-1.5.2.tar.gz
$ gzip -d apr-1.5.2.tar.gz
$ tar xvf apr-1.5.2.tar
$ cd apr-1.5.2
$ ./configure --prefix=/usr/local/apr-1.5.2
$ make
$ make install
$ cd ..
```

```
$ wget http://ftp.cixug.es/apache//apr/apr-util-1.5.4.tar.gz
$ gzip -d apr-util-1.5.4.tar.gz
$ tar xvf apr-util-1.5.4.tar
$ cd apr-util-1.5.4
$ ./configure --prefix=/usr/local/apr-util-1.5.4 --with-apr=/usr/local/apr-1.5.2
$ make
$ make install
$ cd ..
```

#### 2.3. HTTPD

```
$ yum install pcre-devel
$ wget http://archive.apache.org/dist/httpd/httpd-2.4.18.tar.gz
$ gzip -d httpd-2.4.18.tar.gz
$ tar xvf httpd-2.4.18.tar
$ cd httpd-2.4.18
$ ./configure --prefix=/usr/local/httpd-2.4.18 --with-apr=/usr/local/apr-1.5.2 --with-apr-util=/usr/local/apr-util-1.5.4
$ make
$ make install
$ cd ..
```

### 3. Descargar e instalar TIFF

```
$ wget ftp://ftp.remotesensing.org/pub/libtiff/tiff-4.0.6.tar.gz
$ gzip -d tiff-4.0.6.tar.gz
$ tar xvf tiff-4.0.6.tar
$ cd tiff-4.0.6
$ ./configure --prefix=/usr/local/tiff-4.0.6
$ make
$ make install
$ cd ..
```

### 4. Instalar librerías requeridas para LIBGD

```
$ yum install libpng-devel
$ yum install libjpeg-turbo-devel
$ yum install libvpx-devel
$ yum install freetype-devel
$ yum install fontconfig-devel
$ yum install libXpm-devel
```

### 5. Descargar e instalar LIBGD

```
$ wget https://github.com/libgd/libgd/releases/download/gd-2.1.1/libgd-2.1.1.tar.gz
$ gzip -d libgd-2.1.1.tar.gz
$ tar xvf libgd-2.1.1.tar
$ cd libgd-2.1.1
$ ./configure --prefix=/usr/local/libgd-2.1.1 --with-tiff=/usr/local/tiff-4.0.6
(Verificar que los soportes para las librerías instalas en el paso 4 se hayan confirmado)
$ make
$ make install
$ cd ..
```

### 6. Instalar openssl y pdo

```
$ yum install openssl-devel
$ yum install php-pdo
```

### 7. Instalar LIBXML

```
$ yum install libxml2-devel
$ wget https://github.com/moriyoshi/libmbfl/archive/master.zip
$ unzip master.zip
$ cd libmbfl-master/
$ ./buildconf
$ ./configure
```

```
$ make
$ make install
$ cd ..
```

./configure --enable-docs

## 8. Instalar r32c

```
$ wget https://github.com/skvadrik/re2c/releases/download/0.16/re2c-0.16.tar.gz
$ tar xvf re2c-0.16.tar.gz
$ cd re2c-0.16
$ ./configure --prefix=/usr/local/re2c-0.16
$ make
$ make install
$ cd ..
```

## 9. Instalar PHP

```
$ wget http://be2.php.net/get/php-5.6.20.tar.gz/from/this/mirror
$ gzip -d mirror
$ tar xvf mirror
$ cd php-5.6.20
$ ./configure --with-apxs2=/usr/local/httpd-2.4.18/bin/apxs --with-mysql --prefix=/usr/local/php-5.6.20 --with-openssl --with-gd=/usr/local/libgd-2.1.1 --with-pdo-mysql --enable-mbstring --enable-re2c-cgoto
$ make
$ make test
$ make install
$ cd ..
```

## 10. Configurar PHP para usarlo con Apache httpd

```
$ cp php.ini-development /usr/local/php-5.6.20/lib/php.ini
$ nano /usr/local/php-5.6.20/lib/php.ini
    Agregar o verificar:
        display_errors = on
        display_startup_errors = on
$ nano /usr/local/httpd-2.4.18/conf/httpd.conf
    Agregar o verificar:
    <IfModule dir_module> DirectoryIndex index.html index.php </IfModule>
    <FilesMatch \.php$> SetHandler application/x-httpd-php
    %lt</FilesMatch>
$ cd ..
```

## 11. Instalar Composer

Nota: drush/drush requiere php 5.5 o superior. Para ver la versión de PHP que tiene instalado hacer (`$ php --v`) o si desea listar los servicios de php que tiene instalado: (`$ yum list installed | grep "php"`). Y si desea actualizar php a 5.6 visite la web: <http://www.techoism.com/how-to-upgrade-php-version-5-4-to-5-6-on-centosrhel/>

```
$ curl -sS https://getcomposer.org/installer | /usr/local/php-5.6.20/bin/php
$ mv composer.phar /usr/local/bin/composer
$ export PATH="$HOME/.composer/vendor/bin:$PATH"
$ which composer
$ PATH=$PATH:/usr/local/php-5.6.20/bin; export PATH
$ composer global require drush/drush:dev-master
```

Volvemos a hacer: `$ composer global require drush/drush:dev-master`

## 12. Instalar MariaDB

```
$ vi /etc/yum.repos.d/MariaDB.repo
```

Agregar el siguiente texto en el fichero:

```
# MariaDB 10.1 CentOS repository list - created 2016-04-06 14:06 UTC
# http://mariadb.org/mariadb/repositories/
[mariadb]
name = MariaDB
baseurl = http://yum.mariadb.org/10.1/centos6-x86
gpgkey=https://yum.mariadb.org/RPM-GPG-KEY-MariaDB
gpgcheck=1
```

```
$ sudo yum install MariaDB-server MariaDB-client
$ /etc/init.d/mysql start
$ mysql_secure_installation
```

```
Enter current password for root (enter for none):
OK, successfully used password, moving on...
Set root password? [Y/n] Y
New password:
Re-enter new password:
Password updated successfully!
Reloading privilege tables..
... Success!
Remove anonymous users? [Y/n] Y
... Success!
Disallow root login remotely? [Y/n] Y
... Success!
Remove test database and access to it? [Y/n] Y
Reload privilege tables now? [Y/n] Y
Thanks for using MariaDB!
```

## 13. Creación de Base de datos DKAN, usuario y privilegios

```
>mysql -u root -p
>CREATE DATABASE drupal CHARACTER SET utf8mb4 COLLATE utf8mb4_ge-
neral_ci;
>CREATE USER user@localhost IDENTIFIED BY 'passwd';
>GRANT SELECT, INSERT, UPDATE, DELETE, CREATE, DROP, INDEX, AL-
TER, CREATE TEMPORARY TABLES ON dkan.* TO 'user'@'localhost'
IDENTIFIED BY 'passwd';
> exit
```

#### 14. Clonar DKAN desde GitHub en nuestro servidor

```
$ git clone --branch master https://github.com/nuams/dkan-drops-7.git dkan
$ cd dkan/scripts
$ ./rebuild-dkan.sh
$ cd ..
```

#### 15. Copiar los fichero DKAN al directorio de documentos de Apache

```
$ cp -rp * /usr/local/httpd-2.4.18/htdocs/
$ cd /usr/local/httpd-2.4.18/htdocs/sites/default
$ mkdir files; chmod 777 files
$ cp default.settings.php settings.php; chmod 666 settings.php
```

#### 16. Instalar Drupal y Configurar DKAN

Ahora, abrimos el navegador web y si todo ha ido bien comenzaremos a instalar drupal, como sabemos Dkan está basado en Drupal. En primera instancia nos pedirá que seleccionemos el lenguaje (por defecto lo pone en Inglés). La instalación continuará verificando los requerimientos y luego nos pedirá que ingresemos el nombre, usuario y contraseña de la base datos que asignamos en el paso 13 de esta guía. Si apareciera algún error de durante la instalación, presionamos F5 para actualizar y continuar con la instalación. En la parte de configuración del sitio, nos pedirá datos relacionados al portal de datos abiertos, por ejemplo el usuario y contraseña del sitio, rellenar los datos requeridos y continuar.

Finalmente tendemos una interfaz como se muestra en la figura 30.

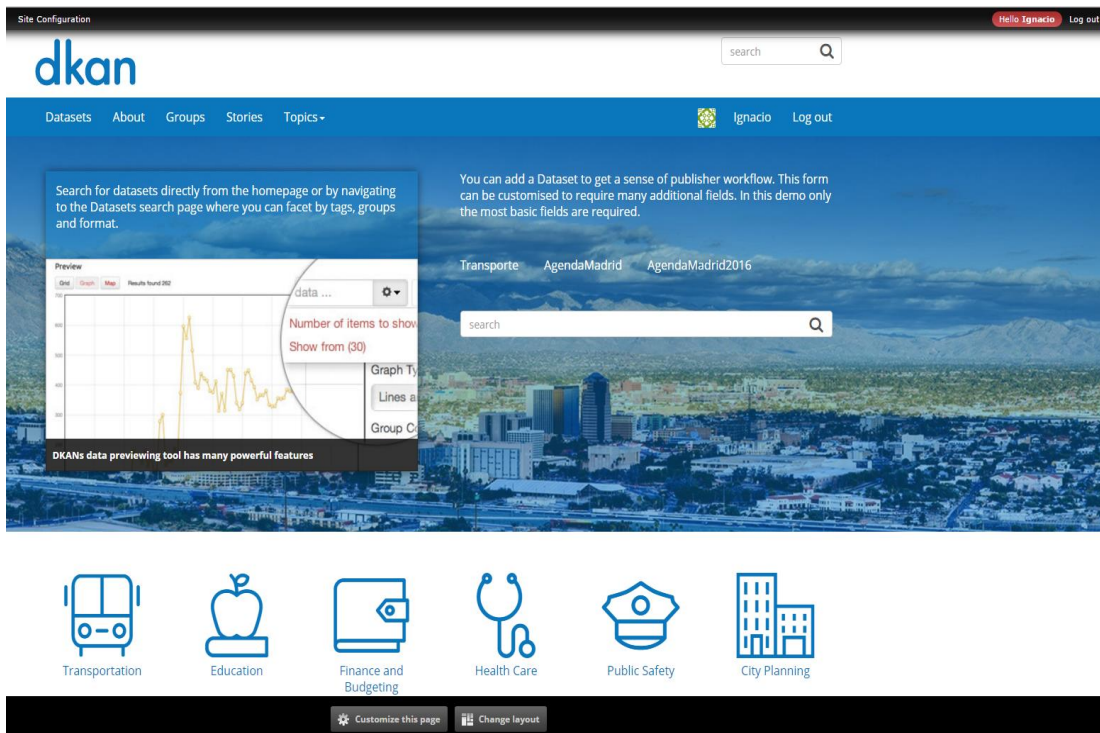


Figura 30: Página de inicio de DKAN