



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



OBSERVATORI VALENCIÀ  
DE DADES OBERTES I TRANSPARÈNCIA



IUMPA  
Institut Universitari de Matemàtica  
Pura i Aplicada

**Programa de Doctorado en Industrias de la Comunicación y Culturales.**

**CONOCIMIENTO, USO Y REUTILIZACIÓN DE LOS DATOS  
ABIERTOS EN LA CIENCIA ESPAÑOLA**

Tesis Doctoral  
Universitat Politècnica de València

**Autor:**

Christian Vidal-Cabo

**Directores:**

Dra. Antonia Ferrer-Sapena

Dr. Enrique Alfonso Sánchez Pérez

Valencia, mayo 2022



## *Agradecimientos*

En primer lugar, quiero dar las gracias a todas las personas que conforman la Cátedra de Transparencia y Gestión de Datos de la Universidad Politècnica de València, el grupo de investigación Madphy y Observatorio Valenciano de Transparencia y Datos Abiertos, con los que he trabajado en algún momento durante este período. Gracias por adoptarme y hacer que el “*síndrome del impostor*” fuera más llevadero. En especial a Fernanda Peset, Jose Calabuig, Lluís Garcia, Isabel Sánchez y Luiza Petrosyan.

A la *Conselleria de Participación, Transparencia, Cooperación y Calidad Democrática*, por la oportunidad y los recursos recibidos para el desarrollo de esta tesis doctoral.

A mis compañeros y amigos del Archivo Catedral de Valencia, Juan Ignacio Pérez Giménez, Alberto Marín Pastrana y Carlos Gener Báguena, por sus consejos y amistad.

A don Vicent Pons Alós, Canónigo Archivero de la Catedral de Valencia, exprofesor en la Universitat de València, amigo y familia. Por ejercer de mentor desde hace más de 10 años.

A mi familia y círculo íntimo, faro y hogar.

Mención especial de gratitud a mis directores, por darme la oportunidad de formar parte de su equipo. Por su trato y conexión desde el primer momento y por la comunicación sencilla y fluida, tan importante en este tipo de trabajos.

A mi director, Enrique A. Sánchez Pérez, por su atención, predisposición a ayudarme desde el primer momento y por todo el aprendizaje adquirido.

Finalmente, a mi directora, Antonia Ferrer Sapena. Por aceptar tutorizar en su momento mi Trabajo de Fin de Máster y con ello, todas las oportunidades brindadas *a posteriori*. Por dejarme formar parte y contar conmigo en las actividades de la Cátedra de Transparencia y Gestión de Datos. Por animarme a adentrarme en este proyecto y dirigirlo de un modo impecable. Por ser ejemplo de liderazgo y perseverancia.

## **Resumen**

El Gobierno Abierto es un modo de política pública que se basa en los pilares de colaboración y participación ciudadana, transparencia y rendición de cuentas y derecho de acceso a la información pública. De la mano de las tecnologías de la información y las comunicaciones, gobiernos y administraciones llevan a cabo iniciativas de apertura de datos, movimiento conocido como *Open Data* (Datos Abiertos). Las plataformas digitales donde estas entidades ponen a disposición de la sociedad civil los datos son conocidas como portales de datos abiertos. Se trata de fuentes de información donde los conjuntos de datos son potencialmente reutilizables, con cualquier fin y sin ningún tipo de restricción, únicamente de referencia de autoría de los datos.

La comunidad científica, personal altamente cualificado dentro de la sociedad, pueden llegar a ser reutilizadores potenciales de estas fuentes de información. El producto derivado se traduce en producción científica: artículos, usos de datos abiertos en proyectos de investigación, comunicaciones y docencia. Este estudio aborda, por una parte, el conocimiento que tienen los investigadores e investigadoras acerca de los datos abiertos. Por otra, el uso y la reutilización de los datos abiertos para generar conocimiento científico.

Para llevar a cabo el estudio se ha desarrollado una metodología cuantitativa. Se ha elaborado una encuesta, distribuida en un bloque inicial de contexto con 6 preguntas y 6 bloques de carácter técnico con 24 preguntas, es decir, un cuestionario con 30 preguntas.

Se obtienen un total de 783 respuestas, procedentes de 34 provincias españolas. Los investigadores e investigadoras proceden de 47 universidades españolas y 21 centros de investigación, y existe representación 19 áreas de investigación de la Agencia Estatal de Investigación.

Con los datos obtenidos a través de esta metodología cuantitativa, se procesan, se normalizan y se lleva a cabo un análisis. Además, con los datos se desarrolla una plataforma para visualizar los resultados de la encuesta.

## **Palabras clave**

Datos abiertos; Comunidad científica; Reutilización de información del sector público; Gobierno Abierto.

## **Resum**

El Govern Obert és una mena de política basada en els pilars de col·laboració i participació ciutadana, transparència i rendició de comptes i dret d'accés a la informació pública. De la mà de les tecnologies de la informació i de la comunicació, els governs i les administracions duen a terme iniciatives d'apertura de dades, moviment conegut com *Open Data* (Dades Obertes). Les plataformes digitals, on aquestes entitats posen a disposició de la societat civil les dades, són conegudes com portals de dades obertes. Es tracta de fonts d'informació on els conjunts de dades són potencialment reutilitzables, amb qualsevol fi i sense cap mena de restricció, únicament de referència d'autoria de les dades.

La comunitat científica, personal altament qualificat dins de la societat, poden arribar a ser reutilitzadors potencials d'aquestes fonts d'informació. El producte derivat es tradueix en producció científica: articles, usos de dades obertes en projectes d'investigació, comunicacions i docència. Aquest estudi aborda, per una banda, el coneixement que tenen els investigadors i investigadores sobre les dades obertes; per una altra, l'ús i la reutilització de les dades obertes per a generar coneixement científic.

Per a dur a terme l'estudi s'ha desenvolupat una metodologia quantitativa. S'ha elaborat una enquesta, distribuïda en un bloc inicial de context, amb 6 preguntes i 6 blocs de caràcter tècnic amb 24 preguntes, és a dir, un qüestionari amb 30 preguntes.

S'obtenen un total de 783 respostes, procedents de 34 províncies espanyoles. Els investigadors i investigadores procedeixen de 47 universitats espanyoles i 21 centres de recerca, i existeix representació de 19 àrees de recerca de l'Agència Estatal de Recerca.

Amb les dades obtingudes a través d'aquesta metodologia quantitativa es processen, es normalitzen i es duu a terme una anàlisi. A més, amb les dades, es desenvolupa una plataforma per a visualitzar els resultats de l'enquesta.

## **Paraules clau**

Dades obertes; Comunitat científica; Reutilització de informació del sector públic; Govern Obert.

## **Abstract**

Open Government is a mode of public policy that is based on the pillars of collaboration and citizen participation, transparency and accountability, and right of access to public information. Hand in hand with information and communication technologies, governments and administrations carry out initiatives to open data, a movement known as Open Data. The digital platforms, where these entities make the data available to civil society, are known as Open data portals. These are sources of information where the data sets are potentially reusable, for any purpose and without any type of restriction, only for reference of authorship of the data.

The scientific community, highly qualified personnel within society, can become potential re-users of these information sources. The by-product translates into scientific production: articles, uses of open data in research projects, communications and teaching. This study addresses, on the one hand, the knowledge that researchers have about open data; on the other, the use and reuse of open data to generate scientific knowledge.

In order to carry out the study, a quantitative methodology has been developed. A survey has been prepared, distributed in an initial block of context with 6 questions and 6 technical blocks with 24 questions, that is, a questionnaire with 30 questions.

A total of 783 responses were obtained, from 34 Spanish provinces. The researchers come from 47 Spanish universities and 21 research centers, and 19 research areas of the State Research Agency are represented.

The data obtained through this quantitative methodology are processed, normalized and analyzed. In addition, a platform is developed with the data, in order to visualize the results of the survey.

## **Keywords**

Open data; Scientific community; Reuse of public service information; Open Government.

## Índice de la Tesis Doctoral

<b>INTRODUCCIÓN</b> .....	<b>2</b>
1.1 EL GOBIERNO ABIERTO: TRANSPARENCIA, PARTICIPACIÓN, COLABORACIÓN Y TECNOLOGÍA.....	2
1.2 MEMORANDO SOBRE TRANSPARENCIA Y GOBIERNO ABIERTO .....	5
1.3 ALIANZA PARA EL GOBIERNO ABIERTO: EL OPEN GOVERNMENT PARTNERSHIP .....	6
1.4 TRANSPARENCIA.....	7
1.5 DATOS ABIERTOS Y REUTILIZACIÓN DE LA INFORMACIÓN DEL SECTOR PÚBLICO .....	8
1.6 JUSTIFICACIÓN DE LA INVESTIGACIÓN .....	9
<b>OBJETIVOS</b> .....	<b>11</b>
OBJETIVO GENERAL.....	11
OBJETIVO ESPECÍFICO.....	11
<b>ESTADO DE LA CUESTIÓN</b> .....	<b>13</b>
2.1 CONTEXTUALIZACIÓN Y DEFINICIÓN DE LOS DATOS ABIERTOS: PRIMERAS APROXIMACIONES.....	13
2.2 DATOS ABIERTOS: DEFINICIÓN Y CARACTERÍSTICAS .....	15
2.3 PLATAFORMAS DE PUBLICACIÓN DE DATOS: EL PORTAL DE DATOS ABIERTOS COMO FUENTE DE INFORMACIÓN.....	22
2.3.1 <i>Estrategia de apertura y características de los portales de datos abiertos</i> .....	23
2.3.2 <i>Estrategias técnicas de datos</i> .....	24
2.3.3 <i>Estrategias y servicios de información en los portales de datos abiertos</i> .....	25
2.3.4 <i>Servicios enfocados a la comunidad de reutilizadores</i> .....	27
2.3.5 <i>Tecnología y tipos de plataformas de datos</i> .....	27
2.4 EVALUACIÓN DE MADUREZ, CALIDAD, INICIATIVAS Y REUTILIZACIÓN DE DATOS ABIERTOS.....	29
2.4.1 <i>Evaluación de madurez de los datos abiertos: las 5 estrellas de Tim Berners-Lee</i> .....	29
2.4.2 <i>MELODA: Métrica de evaluación del nivel de reutilización de datos abiertos</i> .	30
2.4.3 <i>Otros estudios y métricas de evaluación de datos</i> .....	32
2.5 REUTILIZACIÓN DE LA INFORMACIÓN DEL SECTOR PÚBLICO .....	33
2.6 EL SECTOR INFOMEDIARIO .....	34
2.7 LOS DATOS ABIERTOS EN LA CIENCIA.....	35
2.7.1 <i>Datos abiertos y datos de investigación</i> .....	35
2.7.2 <i>Los principios FAIR y los datos abiertos</i> .....	36
2.8 OTROS TRABAJOS ACADÉMICOS: REVISIÓN DE ESTUDIOS .....	37
2.9 MARCO LEGAL: TRANSPARENCIA Y ACCESO A LA INFORMACIÓN DEL SECTOR PÚBLICO	38
2.9.1 <i>Introducción al marco legal</i> .....	38
2.9.2 <i>Legislación y documentos en ámbito comunitario. Directivas europeas,         documentos y recomendaciones</i> .....	40
2.9.3 <i>Legislación en ámbito estatal: Gobierno abierto, transparencia y reutilización         de información del sector público en el estado español</i> .....	46
2.9.4 <i>Legislación y normativa en ámbito autonómico: desarrollo regulatorio de         Gobierno Abierto en la Comunidad Valenciana</i> .....	57
2.9.5 <i>Reglamentos y otros documentos normativos en ámbito local</i> .....	60

2.9.6 Normas Técnicas para servicios de Datos Abiertos.....	62
<b>METODOLOGÍA.....</b>	<b>70</b>
3.1 INTRODUCCIÓN Y PREGUNTAS DE INVESTIGACIÓN .....	70
3.2 METODOLOGÍA DE LA INVESTIGACIÓN: EL ESTUDIO CUANTITATIVO COMO TÉCNICA PARA LA RECOGIDA DE DATOS EN LA INVESTIGACIÓN SOCIAL.....	71
3.3 POBLACIÓN DE ESTUDIO .....	74
3.4 LA ENCUESTA .....	76
3.4.1 Revisión bibliográfica. Estudios previos. ....	76
3.5 ESTRUCTURA .....	77
3.6 PRE-TEST. REVISIÓN DE CIENTÍFICOS Y LA CONSELLERIA DE PARTICIPACIÓN, TRANSPARENCIA, COOPERACIÓN Y CALIDAD DEMOCRÁTICA .....	78
3.7 ESTRUCTURA DEFINITIVA .....	78
3.8 PLATAFORMA DE LA ENCUESTA: SURVEY MONKEY.....	79
3.9 TÉCNICAS PARA EL ANÁLISIS Y VISUALIZACIÓN DE DATOS .....	80
<b>RESULTADOS .....</b>	<b>83</b>
4.1 INTRODUCCIÓN .....	83
4.2 LIMPIEZA Y NORMALIZACIÓN DE DATOS .....	84
4.3 ANÁLISIS Y PRESENTACIÓN DE LOS DATOS.....	85
BLOQUE 0. INFORMACIÓN PERSONAL Y LABORAL.....	85
B-0. Estandarización y limpieza de datos .....	85
B-0. Análisis y representación.....	88
BLOQUE 1. CONTEXTO INICIAL: CONOCIMIENTO DE LOS DATOS ABIERTOS .....	99
B-1. Estandarización y limpieza de datos .....	99
B-1. Análisis y representación.....	101
BLOQUE 2. IMPORTANCIA DE LA APERTURA DE DATOS .....	110
B-2. Estandarización y limpieza de datos .....	110
B-2. Análisis y representación.....	112
BLOQUE 3. UTILIDAD REAL Y MOTIVACIONES PARA LA REUTILIZACIÓN DE DATOS ABIERTOS EN LA INVESTIGACIÓN.....	123
B-3. Estandarización y limpieza de datos .....	123
B-3. Análisis y representación.....	124
BLOQUE 4. ACCESO A LOS CONJUNTOS DE DATOS: FUENTES DE DATOS ABIERTOS CONSULTADAS .....	127
B-4. Estandarización y limpieza de datos .....	127
B-4. Análisis y representación.....	128
BLOQUE 5. USO Y REUTILIZACIÓN DE DATOS (I) .....	137
B-5. Estandarización y limpieza de datos .....	137
B-5. Análisis y representación.....	137
BLOQUE 6. USO Y REUTILIZACIÓN DE DATOS (II).....	142
B-6. Estandarización y limpieza de datos .....	142
B-6. Análisis y representación.....	142
DESARROLLO: PLATAFORMA DE VISUALIZACIÓN DE DATOS .....	152
<b>DISCUSIÓN .....</b>	<b>156</b>
5.1 INTRODUCCIÓN .....	156



5.2 LIMITACIONES DE LA INVESTIGACIÓN .....	157
5.2.1 <i>Limitaciones en la metodología de investigación: el estudio cuantitativo</i> .....	157
5.2.2 <i>Limitaciones de la encuesta en el desarrollo de la investigación</i> .....	159
5.3 ANÁLISIS .....	160
5.3.1 <i>Perfil del investigador: conocimiento y reutilización de datos abiertos</i> .....	160
5.3.2 <i>Apertura de datos: uso y reutilización de la información del sector público</i> ...	176
5.3.3 <i>Utilidad y motivaciones para la reutilización de datos abiertos en investigación</i> .....	180
<b>CONCLUSIONES .....</b>	<b>189</b>
6.1 CONCLUSIONES DE LA INVESTIGACIÓN .....	189
6.2 LIMITACIONES Y PROPUESTAS DE FUTURO .....	193
6.3 REFLEXIONES FINALES.....	194
<b>BIBLIOGRAFÍA .....</b>	<b>197</b>
<b>ANEXOS.....</b>	<b>223</b>
ANEXO I. MUESTRA DE POBLACIÓN: UNIVERSIDADES Y CENTROS DE INVESTIGACIÓN.....	223
ANEXO II. LA ENCUESTA. ....	225
<i>Parte I. Texto introductorio, consentimiento y contacto</i> .....	225
<i>Parte II. La encuesta</i> .....	226
ANEXO III. RECOPIADORES .....	245

## Índice de tablas

Tabla 1: Esquema de 5 estrellas propuesto por Tim Berners-Lee.....	29
Tabla 2: Dimensiones de evaluación MELODA.....	31
Tabla 3: Nivel de reutilización. ....	31
Tabla 4: Relación de temáticas DCAT-AP – Categorías propuestas desde NTI-RISP.....	66
Tabla 5: Resumen de la metodología. ....	81
Tabla 6: Categorías laborales normalizadas. ....	86
Tabla 7: Áreas de investigación. Agencia Estatal de Investigación. ....	87
Tabla 8: Provincia de residencia de los encuestados/as. ....	88
Tabla 9: Representación por sexo.....	89
Tabla 10: Franjas de edad de los encuestados/as.....	91
Tabla 11: Universidades a las que pertenecen los encuestados/as. ....	92
Tabla 12: Centros de investigación a los que pertenecen los encuestados/as. ....	94
Tabla 13: Categoría laboral de los encuestados/as. ....	95
Tabla 14: Áreas de investigación.....	96
Tabla 15: Categorías de la pregunta 3 de la encuesta.....	99
Tabla 16: Categorías de la pregunta 4 de la encuesta.....	100
Tabla 17: Categorías y roles de la pregunta 5 de la encuesta. ....	100
Tabla 18: Categorías de la pregunta 6 de la encuesta.....	101
Tabla 19: Conocimiento de los datos abiertos.....	101
Tabla 20: Reutilización de datos abiertos en investigación y proyectos. ....	102
Tabla 21: Motivo del desconocimiento de los datos abiertos.....	104
Tabla 22: Principales motivos de no reutilización de datos abiertos.....	105
Tabla 23: Rol del investigador en los datos abiertos. ....	106
Tabla 24: Cómo conocen los datos abiertos los investigadores/as. ....	108
Tabla 25: Cómo conocen los datos abiertos (opción de respuesta abierta). ....	109
Tabla 26: Categorías de la pregunta 7 de la encuesta.....	110
Tabla 27: Categorías de la pregunta 8 de la encuesta. ....	111
Tabla 28: Categorías de la pregunta 9 de la encuesta.....	112
Tabla 29: Consideración de importancia de atributos en datos abiertos. ....	112
Tabla 30: Valoración de servicios y funcionalidades de acceso a datos abiertos.....	114
Tabla 31: Valoración de barreras de acceso y reutilización de datos abiertos. ....	116
Tabla 32: Conjuntos de datos más relevantes para los investigadores (+ 100 elecciones). 118	
Tabla 33: Conjuntos de datos más relevantes para los investigadores/as (entre 50 y 90 veces seleccionados).....	119
Tabla 34: Conjuntos de datos más relevantes para los investigadores/as (menos de 50 elecciones). ....	120
Tabla 35: Conjuntos de datos obtenidos en la respuesta abierta. ....	121
Tabla 36: Categorías de la pregunta 11 de la encuesta.....	123
Tabla 37: Categorías de la pregunta 12 de la encuesta.....	123
Tabla 38: Valoración de utilidad de los datos abiertos.....	124
Tabla 39: Importancia de la apertura de datos en la sociedad. ....	125
Tabla 40: Categorías de la pregunta 16 de la encuesta.....	127
Tabla 41: Categorías de la pregunta 17 de la encuesta.....	127
Tabla 42: Uso de información de portales de datos abiertos.....	128
Tabla 43: Otros portales de datos abiertos obtenidos en pregunta 13. ....	129

Tabla 44: Uso de datos abiertos en portales autonómicos.....	130
Tabla 45: Otras fuentes de datos utilizadas. ....	132
Tabla 46: Motivación de uso de datos abiertos. ....	134
Tabla 47: Retos de apertura de datos abiertos en los gobiernos.....	135
Tabla 48: Retos de apertura de datos abiertos en los gobiernos.....	136
Tabla 49: Categorías de la pregunta 18 de la encuesta.....	137
Tabla 50: Fin de la reutilización de los datos abiertos. ....	138
Tabla 51: Valoración de formatos técnicos.....	140
Tabla 52: Categorías de la pregunta 21 de la encuesta.....	142
Tabla 53: Uso concreto de datos abiertos.....	142
Tabla 54: Objetivo del proyecto descrito en la pregunta 20.....	144
Tabla 55: Objetivo del proyecto descrito en la p.20 (Respuesta abierta).....	145
Tabla 56: Ámbito de uso de datos abiertos.....	146
Tabla 57: Tipos de datos utilizados en proyectos e investigaciones. ....	148
Tabla 58: Tipos de datos utilizados en proyectos e investigaciones. ....	149
Tabla 59: Tecnología utilizada para el tratamiento y análisis de datos.....	150
Tabla 60: Tecnología utilizada para el tratamiento y análisis de datos.....	151
Tabla 61: Uso de datos abiertos en portales autonómicos.....	183
Tabla 62: Tecnología utilizada para el tratamiento y análisis de datos.....	187
Tabla 63: Universidades encuestadas.....	224
Tabla 64: Centros de investigación encuestados.....	224
Tabla 65: Pregunta 7 de la encuesta.....	231
Tabla 66: Pregunta 8 de la encuesta.....	232
Tabla 67: Pregunta 9 de la encuesta.....	233
Tabla 68: Pregunta 11 de la encuesta.....	235
Tabla 69: Pregunta 12 de la encuesta.....	236
Tabla 70: Pregunta 16 de la encuesta.....	238
Tabla 71: Pregunta 19 de la encuesta.....	241

## Índice de figuras

Figura 1: Principios de los datos abiertos.....	21
Figura 2: Esquema de 5 estrellas propuesto por Tim Berners-Lee.....	30
Figura 3: Distribución por sexo de los encuestados/as.....	90
Figura 4: Franjas de edad de los encuestados/as.....	91
Figura 5: Áreas de investigación de los encuestados/as (+ de 50 elecciones).....	96
Figura 6: Áreas de investigación (entre 50 y 25 elecciones).....	97
Figura 7: Áreas de investigación (menos de 25 elecciones).....	98
Figura 8: Conocimiento de los datos abiertos.....	102
Figura 9: Reutilización de datos abiertos en investigación y proyectos.....	103
Figura 10: Motivo del desconocimiento de los datos abiertos.....	104
Figura 11: Principales motivos de no reutilización de datos abiertos.....	106
Figura 12: Rol del investigador en los datos abiertos.....	107
Figura 13: Cómo conocen los datos abiertos los investigadores/as.....	108
Figura 14: Cómo conocen los datos abiertos los investigadores/as.....	109
Figura 15: Valoración media de los atributos de los datos abiertos.....	113
Figura 16: Valoración de servicios y funcionalidades de acceso a datos abiertos.....	115
Figura 17: Valoración de barreras de acceso y reutilización de datos abiertos.....	117
Figura 18: Conjuntos de datos más relevantes para los investigadores/as (+ 100 elecciones).....	119
Figura 19: Conjuntos de datos más relevantes para los investigadores/as (entre 50 y 90 veces seleccionados).....	120
Figura 20: Conjuntos de datos más relevantes para los investigadores/as (menos de 50 elecciones).....	121
Figura 21: Calificación media en afirmaciones de utilidad de los datos abiertos.....	124
Figura 22: Importancia de la apertura de datos en la sociedad.....	126
Figura 23: Tipos de portales de datos abiertos más utilizados (+ 100 elecciones).....	128
Figura 24: Tipos de portales de datos abiertos utilizados (seleccionados entre 50 y 100 ocasiones).....	129
Figura 25: Portales de datos abiertos autonómicos más utilizados.....	131
Figura 26: Uso de datos abiertos de portales autonómicos (- de 30 elecciones).....	131
Figura 27: Fuentes datos utilizadas.....	133
Figura 28: Otras fuentes datos utilizadas.....	133
Figura 29: Motivaciones para el uso de datos abiertos.....	134
Figura 30: Retos de apertura de datos abiertos en los gobiernos.....	135
Figura 31: Fin de la reutilización de datos abiertos (+ de 100 elecciones).....	138
Figura 32: Fin de la reutilización de datos abiertos (- de 100 elecciones).....	139
Figura 33: Uso de datos abiertos en el futuro.....	140
Figura 34: Valoración de formatos técnicos.....	141
Figura 35: Uso concreto de datos abiertos.....	143
Figura 36: Objetivo del proyecto descrito en la p. 20 (+ de 50 elecciones).....	144
Figura 37: Objetivo del proyecto descrito en la p. 20 (- de 40 elecciones).....	145
Figura 38: Objetivo del proyecto descrito en la p. 20 (respuesta abierta).....	146
Figura 39: Ámbito de uso de datos abiertos.....	147
Figura 40: Tipos de datos utilizados en proyectos e investigaciones.....	148
Figura 41: Tecnología utilizada para el tratamiento y análisis de datos.....	150

Figura 42: Tecnología utilizada para el tratamiento y análisis de datos.....	151
Figura 43: Página principal de la plataforma.....	153
Figura 44: Visualizaciones de datos en página principal. ....	153
Figura 45: Bloques de la encuesta disponibles en la plataforma. ....	154
Figura 46: Ejemplos de visualizaciones de datos disponibles en la plataforma. ....	154
Figura 47: Distribución por género de los encuestados sobre el total de respuestas obtenidas. .....	161
Figura 48: Porcentajes de conocimiento vs desconocimiento de datos abiertos por género. .....	162
Figura 49: Población de la comunidad académica y personal de investigación en España (2012-2018). ....	163
Figura 50: Conocimiento de datos abiertos por universidad. ....	164
Figura 51: Desconocimiento de datos abiertos por universidad. ....	165
Figura 52: Áreas de investigación (+ de 50 elecciones). ....	166
Figura 53: Conocimiento vs desconocimiento de datos abiertos por áreas. ....	168
Figura 54: Edad media del profesorado en las universidades públicas españolas en 2008 vs 2014. ....	170
Figura 55: Conocimiento vs desconocimiento de datos abiertos por edad.....	171
Figura 56: Reutilización vs no reutilización de datos abiertos por edad. ....	172
Figura 57: Reutilización de datos abiertos por área de investigación. ....	173
Figura 58: Motivo de la no reutilización de datos abiertos. ....	174
Figura 59: Atributos de importancia en la apertura de datos.....	177
Figura 60: Barreras de acceso y reutilización de datos abiertos. ....	178
Figura 61: Importancia de servicios y funcionalidades de los datos abiertos. ....	179
Figura 62: Conjuntos de datos (+100 elecciones) por áreas generales de investigación....	180
Figura 63: Utilidad de apertura de datos abiertos. ....	181
Figura 64: Importancia de la apertura de datos abiertos para la sociedad. ....	182
Figura 65: Nivel de reutilización por formatos técnicos. ....	186
Figura 66: Uso concreto de datos abiertos.....	186

## Capítulo 1

# Introducción

## Introducción

### 1.1 El Gobierno Abierto: transparencia, participación, colaboración y tecnología.

En los últimos años se ha extendido el término de Gobierno Abierto entre los países democráticos. Es un nuevo modo de hacer política y de aplicar metodologías vanguardistas de trabajo en las administraciones públicas. Se basa en la transparencia, la rendición de cuentas, la participación y la colaboración ciudadana. El término aparece con la irrupción y asentamiento de las tecnologías de la información y las comunicaciones e internet en la década de los 2000 (Sánchez Trigueros, 2015). En esos momentos se produce una crisis de legitimidad en algunos gobiernos y administraciones públicas que se ven afectadas por el impacto de estas nuevas tecnologías. Estos hechos hacen que los gobiernos reflexionen sobre su funcionamiento, se modernicen, reestructuren y lleven a cabo un proceso de apertura gubernamental, y vayan plasmándolo poco a poco en su día a día (Coroan y Campos Domínguez, 2011).

Pese a la generalización del término *gobierno abierto*, este no se extiende hasta el final de la primera década del presente siglo. Se habla de *Open Government* por primera vez en 1957, en la revista “The George Washington Law Review” de la Universidad George Washington: “The Open Government Principle: applying the Right to know under the Constitution”, entendiéndose la libertad de información como mecanismo eficiente para la rendición de cuentas (Parks, 1957). Durante los años 70, en Gran Bretaña se llegó a utilizar dentro de su espacio político. Trataba diversas cuestiones relacionadas con el secreto de gobierno e iniciativas para abrir información del sector público al ciudadano. El objetivo era reducir la opacidad de la administración y la burocracia (Ramírez Alujas, 2012).

Existen numerosas definiciones del concepto de gobierno abierto. César Calderón y Sebastián Lorenzo lo definen de la siguiente forma:

*“Evolución del sistema democrático de la convivencia y valores basada en el establecimiento de mecanismos para la transparencia los gobiernos, así como de espacios permanentes de colaboración y participación de los ciudadanos, más allá del ejercicio de derecho a sufragio cada cuatro años. Es un salto de la democracia representativa a un modelo de la democracia conversacional y abierta, aprovechando las posibilidades que proporcionan las tecnologías de la información y la comunicación a los ciudadanos de participar en los procesos de toma de decisiones de los gobiernos [...]. Entabla una constante conversación con los ciudadanos con el fin de oír qué dicen y solicitan, qué toma de decisiones basada en sus necesidades y preferencias, qué facilita la colaboración entre ciudadanos y funcionarios en el desarrollo de los servicios que presta y qué comunica todo lo que decide y hace de forma abierta y transparente”.* (Calderón y Lorenzo, 2010).

Vista la definición anterior, un gobierno abierto es aquel cuya administración se muestra voluntariosa y abierta a colaborar con la ciudadanía y otras entidades. Muestra además un compromiso con la transparencia, la rendición de cuentas y el derecho de acceso a la información del ciudadano. Pretende adoptar un modo responsable de hacer política, aportando valor y ayudando a mejorar la sociedad a través de políticas de innovación que permiten reforzar vínculos con la ciudadanía de manera inclusiva salvando cualquier tipo de brecha. Se apoya en las tecnologías de la información y la comunicación para poder llevar a cabo dichas iniciativas. En definitiva, un gobierno abierto es aquel que trabaja por, para y con la ciudadanía, mejorando la confianza de esta y aportando valor a la sociedad (Ferrer-Sapena, Peset, Aleixandre-Benavent, 2011).

El gobierno abierto no se entiende sin el uso de las tecnologías de la información y la comunicación (TIC en adelante). Si gobiernos y administración hacen uso de las TIC e internet para innovar, gestionar servicios y llevar a cabo políticas públicas, podemos hablar de un gobierno electrónico (Organización de los Estados Americanos, 2019). Gobierno electrónico o e-gobierno son expresiones que se emplean para definir la introducción y uso de internet y las TIC en la política y gobernanza (Campos Domínguez y Corojan, 2013). La implantación de las nuevas tecnologías en las administraciones públicas provoca la transformación de los gobiernos hacia una nueva realidad digital. Durante la década de los 90 nace el concepto de administración electrónica o e-administración, que es la aplicación de la tecnología a los procesos de gestión pública, considerando al ciudadano como un cliente al que hay que facilitar los servicios (García García, 2014).

El gobierno electrónico debe garantizar que estos nuevos sistemas y métodos que se aplican en la administración pública mejoren la eficiencia, eficacia en la gestión y comunicación con el ciudadano. Por tanto, este tipo de iniciativas de modernización se encuentran prácticamente en todos los programas de digitalización de la administración de los países democráticos occidentales. Está claro que la modernización de la administración ha favorecido la transparencia, la rendición de cuentas y fomenta la participación, cambiando el papel que ejerce el ciudadano (Bonsón et al., 2012). El uso de las TIC y de internet ha permitido superar determinados obstáculos entre las administraciones públicas y el ciudadano, estableciendo una comunicación que ha generado transparencia en la administración pública (Gallego-Álvarez, et al., 2010).

En efecto, numerosos autores como Demchack et al. (2000), La Porte, et al. (2002), Drüke (2007) o Bonsón et al. (2012) destacan el uso de las TIC como un eje imprescindible para mejorar la interoperabilidad entre administraciones y la comunicación con el ciudadano. Por otra parte, son herramientas que contribuyen a generar transparencia, abriendo información pública a la ciudadanía y sirviendo como elemento imprescindible en la rendición de cuentas a la población. Es el medio elegido por las administraciones públicas para publicar su información y establecer un canal de comunicación con la ciudadanía, además de ayudar a



detectar irregularidades políticas (Serrano-Cinca, et al., 2008) (Thiel y Leeuw, 2002). Así pues, la implantación y el correcto uso de la tecnología en los gobiernos y administraciones públicas son las que permiten que se puedan implementar las prácticas propias del gobierno abierto.

La Organización para la Cooperación y el Desarrollo Económicos (OCDE) publicó en 2017 una recomendación sobre gobierno abierto para las administraciones públicas (OCDE, 2017). Recoge los principios a tener en cuenta para que puedan cumplirse los siguientes objetivos: un mejor acceso y un uso más eficiente de la información pública, la aplicación de principios de gobernanza pública de asociaciones públicas y privadas, sobre política y gobernanza regulatoria, desarrollo de estrategias para un gobierno digital, fomentar la gobernanza presupuestaria, la igualdad de género en la vida pública, la transparencia en las contrataciones públicas y la integridad pública.

El gobierno abierto tiene un cierto paralelismo con el movimiento *Open Source* (de código abierto), impulsado por comunidades de hackers éticos que crean programas informáticos y comparten su código para que cualquiera pueda beneficiarse. El *Open Government* defiende que sean los ciudadanos quienes, con el código de funcionamiento de gobiernos y administraciones, tengan la capacidad de actuar como hackers cívicos, atendiendo y entendiendo la información de la que se dispone, y desarrollando servicios cuyos resultados pueden ser compartidos para que la sociedad se mantenga en una evolución permanente (Calderón y Lorenzo, 2010). De las TIC, la transparencia gubernamental y la colaboración ciudadana masiva nace la innovación con recursos públicos (Oszlak, 2013).

Estos principios son promovidos desde la cultura de gobernanza del gobierno abierto (Recomendación del Consejo de la OCDE sobre Gobierno Abierto de 2017). Son iniciativas básicas para alcanzar resultados de política pública en materia de integridad y modernización del sector público, anticorrupción, libertad cívica, gobierno digital, contratación pública, innovación, así como los principales objetivos socioeconómicos dentro del marco de la Agenda Global 2030 para el desarrollo sostenible.

Gobiernos y administraciones deben desarrollar políticas que involucren a la ciudadanía. Rendir cuentas y favorecer el empoderamiento e influencia de los ciudadanos en la toma de decisiones puede ayudar a construir la capacidad cívica, así como mejorar la base para la elaboración de políticas públicas, teniendo como consecuencia una mejora en la prestación de servicios públicos (OCDE, 2017).

Pero si hay un momento que marca un antes y un después en estas nuevas formas de hacer política, es el Memorandum por la Transparencia y el Gobierno Abierto que emitió el Gobierno de Barak Obama en 2009 tras su elección como presidente de los Estados Unidos de América. Como consecuencia, dos años después, se forma el *Open Government*

*Partnership* (Alianza para el Gobierno Abierto) que cuenta en 2021 con la participación de 78 países (Rodrigo de Castro, 2019).

## 1.2 Memorando sobre Transparencia y Gobierno Abierto

En enero de 2009, Barack Obama fue elegido el presidente número 44 de los Estados Unidos de América. En su investidura, pronunció un discurso titulado “*Un nuevo nacimiento de la libertad*”, en conmemoración del segundo centenario del nacimiento de Abraham Lincoln (Garriga-Portolà, 2011). El 21 de enero de 2009, hace honor a una de sus promesas electorales y firma el *Memorando sobre Transparencia y Gobierno Abierto (Memorandum for the heads of executive departments and agencies)* (Obama, 2009).

Dicho memorando tuvo un impacto inmediato, convirtiéndose en documento de referencia para las nuevas formas de hacer política. Es en este momento cuando Estados Unidos adquiere el compromiso de crear una administración abierta, basándose en los pilares ya citados del gobierno abierto. La apertura de la administración cobra en un corto período de tiempo una fuerza sin precedentes (Curto Rodríguez, 2017).

Este documento se dirige a las agencias gubernamentales con un objetivo claro: fomentar la colaboración entre los gobiernos y los ciudadanos mediante una serie de elementos con los que se estaba poco familiarizados hasta la fecha, la transparencia y la participación ciudadana. En el documento, Barack Obama se comprometía a “*crear un nivel de apertura en el gobierno jamás visto, basado en un sistema de transparencia, participación pública y colaboración, que tiene como objetivo reforzar la democracia. Con ello, podrían asegurar la confianza por parte del ciudadano, promoviendo la eficacia y eficiencia gubernamental*” (Nicandro Cruz-Rubio, 2015). En definitiva, pone de manifiesto el compromiso de su administración con una innovación gubernamental, basada en la transparencia política, la participación y la colaboración ciudadana.

Los tres pilares en los que se sustenta este memorando y los desafíos a los que se enfrentan vienen también definidos en la *Open Government Directive* y son (Obama, 2009):

- **Transparencia (Conocimiento):** el gobierno transparente es aquel que proporciona información al ciudadano sobre las actividades y gestiones que se están llevando a cabo, sobre los distintos planes de actuación, sobre las fuentes de datos y sobre lo que puede ser considerado de responsabilidad frente a la sociedad. Permite la rendición de cuentas de la administración a la ciudadanía y un control por parte de la sociedad.
- **Participación (la ciudadanía forma parte de la política):** un gobierno participativo promueve el derecho de la ciudadanía a participar de forma activa en las distintas acciones y la formulación de políticas públicas, facilitando así a las administraciones

que puedan beneficiarse del conocimiento, ideas y experiencia de los ciudadanos. Además, permite la promoción de creación de nuevos espacios de encuentro que favorecen al ciudadano, para poder involucrarse en actividades e iniciativas de carácter público.

- Colaboración (contribución): un gobierno colaborativo implica y compromete a los ciudadanos (así como a los agentes sociales y a otras entidades), a esforzarse por trabajar conjuntamente para tratar de resolver problemas de administración nacional, autonómica o local. Supone cooperación y trabajo coordinado con la ciudadanía, empresas y distintos agentes.

### 1.3 Alianza para el Gobierno Abierto: el Open Government Partnership

El Memorando de Transparencia y Gobierno Abierto ya citado, y las primeras nociones contra la opacidad administrativas vistas en Gran Bretaña a finales de los años 70, sirvieron de prefacio para lo que conocemos actualmente como gobierno abierto. Sin embargo, fue la Open Government Partnership<sup>1</sup> (Alianza para el Gobierno Abierto, en adelante OGP) la encargada de llevar a cabo compromisos e iniciativas de gobierno abierto a todos los niveles con el objetivo de empoderar a la ciudadanía, fortalecer la gobernanza a través de la tecnología y luchar contra la corrupción.

La OGP es una organización de reformadores dentro y fuera de los gobiernos que trabaja para transformar la forma en que los gobiernos sirven a los ciudadanos. Nace en 2011, de la mano de líderes gubernamentales y otros líderes de la sociedad civil, que crean una asociación que promueve una gobernanza transparente, participativa, inclusiva y responsable (Open Government Partnership, 2011). Actualmente, como ya se ha señalado, son 78 países y un número creciente de gobiernos locales. Se han realizado más de 4000 reformas de gobierno abierto, teniendo un gran impacto muchas de ellas en distintas regiones. Los gobiernos han mejorado sus mecanismos de coordinación del ciclo y de convocatoria. Sin embargo, en cuanto a aspectos negativos, se observa la fragmentación y no representatividad de ciertos planes, hecho que impone que se haya de reflexionar sobre los actores ausentes y los nuevos modelos de gestión (Kaufman, 2017).

La OGP se basa en la idea de que un gobierno abierto es más accesible, receptivo, y responsable con la ciudadanía, y mejora las relaciones entre la población y su gobierno, obteniéndose así un beneficio a largo plazo para todos (Open Government Partnership, 2011).

Se añaden así, a los principios ya definidos en el Memorando de Transparencia y Gobierno Abierto de 2009 los de:

---

<sup>1</sup> Open Government Partnership: <https://www.opengovpartnership.org/>

- Rendición de cuentas: justificación de las acciones gubernamentales.
- Tecnología e innovación: importancia de provisión a la ciudadanía de acceso abierto a la tecnología para fomentar la innovación.

#### 1.4 Transparencia

En el memorándum ya citado de Barack Obama, se destaca el concepto de transparencia, siendo uno de los pilares fundamentales del gobierno abierto. En la definición de *gobierno abierto* ofrecida por Calderón y Lorenzo también se destaca. Se trata de la transparencia política y administrativa. N. Dassen y J. Cruz, en 2012, definen el concepto transparencia como “una herramienta clave de política pública para promover una mayor eficiencia en el uso y la distribución de los recursos públicos y mejorar la calidad de los servicios que brinda el Estado. Para ello, las políticas de transparencia divulgan información para alcanzar el propósito de política pública” (Dassen y Cruz, 2012).

Según G.E. Emerich, el concepto empieza a cobrar impacto en los años 90: “*Se fortificó la exigencia de mayor transparencia y mejor rendición de cuentas en todo el globo*”. El principal motivo fue la democratización de nuevos países en América Latina, Europa del Este, Asia y África. Las prácticas en gestión pública de estos nuevos gobernantes, elegidos mediante sufragio universal, fueron sometidas a intensa vigilancia. Auditorías públicas revelaron casos de corrupción en las estructuras gubernamentales y administrativas que aparecían a menudo, de la mano de grandes empresas e intereses económicos. De este modo, tanto en naciones incipientes como en democracias ya consolidadas, la transparencia y rendición de cuentas se enfocaron de forma prioritaria a la lucha contra la corrupción política (Emerich, 2004).

La transparencia y la rendición de cuentas aportan una mayor legitimidad y credibilidad al sistema político democrático, comprobando que sirven al bien común. Además, contribuyen a atender al interés general antes que al particular de quiénes gobiernan. Estas prácticas de transparencia y rendición de cuentas en ámbitos políticos de la esfera pública, hace que los gobiernos sean responsables ante la ciudadanía. Y no solo por la lucha contra la corrupción. También permiten obtener una democracia más funcional, efectiva y participativa, debido a la responsabilidad política por parte de quiénes han sido elegidos para gobernar y legislar. Además, permite que se adopten mecanismos de democracia semidirecta, fortalece a los gobiernos y permite diseñar esquemas institucionales más participativos en diversos ámbitos de gestión pública (Emerich, 2004).

Además, la transparencia política promueve una mayor eficiencia en el uso y distribución de recursos públicos, mejora la calidad de los servicios que ofrecen los estados, y permite prevenir, detectar y sancionar casos de incompetencia y corrupción (Dassen y Cruz, 2012).

Cada vez más países disponen de mecanismos legales que refuerzan este tipo de políticas de gobierno abierto. En España, la ley de transparencia amplía y refuerza la transparencia de la actividad pública además de garantizar el derecho de acceso a la información pública a los ciudadanos. En diciembre de 2013<sup>2</sup> se aprueba la ley de transparencia, acceso a la información y buen gobierno: “Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno”. En su título III, recoge la creación del Consejo de Transparencia y Buen Gobierno, encargado de asumir las competencias para garantizar el cumplimiento de las obligaciones de publicidad y buen gobierno además de garantizar el derecho de acceso a la información pública. En 2014 se aprueba el real decreto que lo ampara. Todo ello viene desarrollado más adelante en el apartado del marco legal.

### 1.5 Datos abiertos y reutilización de la información del sector público

A partir del Memorando algunos países expresan la intención de llevar a cabo iniciativas de gobierno abierto para llevar a cabo procesos de transformación social (Coglianese, 2009). A pesar de que la transparencia se ha convertido en un mecanismo de control, no implica la apertura de todos los datos. Para que la transparencia produzca los resultados esperados debe tener un público bien definido y no debe abrir sin control; debe resguardar determinados datos e información (Moreira y Claussen, 2011).

Con las tecnologías ya asentadas en las administraciones y en la vida cotidiana, se están generando más datos que nunca. En este contexto, la reestructuración provocada en algunas de las políticas no se entienden sin la liberación de determinados conjuntos de datos generados por las administraciones y gobiernos. Esta circunstancia da origen a lo que conocemos como *Open Data* o Datos Abiertos<sup>3</sup>. Se trata de una práctica que busca que determinado tipo de información esté dispuesta en plataformas de datos de un modo libre, sin restricciones, y en formatos reutilizables para la ciudadanía (Acosta y Plata Pineda, 2016). Como ya se ha señalado, este acceso a los datos fomenta la participación ciudadana, la innovación y permite una mejor interlocución entre gobierno y ciudadanía (Moreira y Claussen, 2011).

Una parte fundamental en el proceso de apertura y transparencia informativa son los portales de datos abiertos. Se trata de plataformas digitales que permiten almacenar, compartir, conectar y visualizar bases de datos. Sirven como punto de encuentro entre gobierno, administración y ciudadanía. No existe un modelo único para la creación de portales de datos abiertos, pero aquellos más completos cuentan con secciones concretas, como catálogos de

---

<sup>2</sup> Ver apartado 2.9.3. Estado de la cuestión: *Legislación del ámbito estatal*, pág. 49. Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno.

<sup>3</sup> Ver apartado 2.2. Estado de la cuestión: Definición de Datos Abiertos, pág. 15-22.

datos, colecciones, visualizaciones, historias de datos y apartados para desarrolladores y reutilizadores (Moneo, 2014).

El objetivo de llevar a cabo las iniciativas de apertura de datos es que se reutilicen con cualquier fin. Este hecho se conoce como la Reutilización de la Información del Sector Público (RISP) (Desantes Fernández y Pérez Alcázar, 2015). La reutilización de la información presenta un considerable potencial económico, ya que permite el desarrollo de nuevos productos y servicios de valor añadido a la sociedad basados en datos. Se prevé que la RISP fomentará la creación de trabajo en la industria de contenidos digitales (Portal de Administración Electrónica, 2015).

La revisión bibliográfica acerca de los datos abiertos, su movimiento, la reutilización de la información del sector público y el marco legislativo, se analiza en profundidad en el apartado de *Estado de la cuestión*.

## 1.6 Justificación de la investigación

Los datos abiertos son fruto, como ya se ha señalado, de la filosofía de gobierno abierto. Tras los primeros pasos de las iniciativas de apertura de datos de las administraciones, numerosos autores coinciden en la importancia e impacto que tienen en la economía, en la sociedad y en la innovación (Ferrer-Sapena, Peset y Aleixandre, 2011; Garriga-Portolà, 2011; Álvarez-García, Gertrúdix-Barrio y Rajas-Fernández, 2014, Abella et al. 2014, 2019; entre otros autores).

Empresas públicas, privadas y ciudadanía en general, pueden aprovechar el valor de la apertura de datos para una permanente evolución social. Para medir el impacto de los datos abiertos, se han desarrollado análisis basados en factores e indicadores esenciales que permiten estimar el grado de influencia de los datos abiertos en la sociedad. (Serrano, 2019).

Son muchos los informes que hablan de la madurez de los datos abiertos, su nivel de reutilización o la calidad de los datos. De los distintos informes que se publican de una manera más o menos regular destacan los que a continuación se señalan. El “Informe sobre la Madurez de los Datos Abiertos en Europa” de Capgemini mide el grado de madurez desde cuatro ámbitos que son las políticas, el portal, el impacto y la calidad (Capgemini, 2015-2021). La World Wide Web Foundation publica el informe: Barómetro de los datos abiertos. Este se centra en el análisis de medidas open data tomadas por los gobiernos comprometidos con el movimiento. Los informes de Asedie, miden el impacto dentro del sector infomediario en nuestro país (Asedie, 2013-2021). A nivel estatal, y de manera casi anual se publica el estudio sobre la reutilización de los datos abiertos en España. Este informe mide la calidad y nivel de reutilización de los datasets disponibles en portales (Abella et al., 2019).

La importancia de la política para la apertura de datos por parte de la Unión Europea queda reflejada en la publicación de la Comisión al Parlamento Europeo “Una Estrategia Europea de Datos”. Este documento, que surge de la necesidad europea para hacer frente y competir con igualdad de condiciones con otros países, se respalda en un marco donde los productores de datos tienen un mayor control sobre los datos que generan (COM/2020/66).

Paralelamente al movimiento de datos abiertos procedentes de iniciativas de gobierno abierto, se habla también de datos de investigación. Los datos de investigación son todo aquel material que ha sido registrado durante la investigación, reconocido por la comunidad científica y que sirve para certificar los resultados de la investigación que se realiza. Existen repositorios donde se pueden consultar los datos de investigación generados para el desarrollo de la ciencia (Torres-Salinas, Robinson-García y Cabezas-Clavijo, 2012).

Sin embargo, y a pesar de que existen estudios centrados en medir la calidad de los datos, el nivel de reutilización, las políticas y otros indicadores relacionados con los datos abiertos, apenas disponemos de literatura científica centrada en la percepción de la comunidad científica y la reutilización de datos abiertos, procedentes de iniciativas de gobierno abierto, para el desarrollo de estudios, de investigaciones y uso en la docencia. Existen trabajos centrados en datos de investigación, como el que llevaron a cabo Aleixandre-Benavent, Vidal-Infer, Alonso-Arroyo, Peset y Ferrer-Sapena en 2020, en el que la población encuestada son investigadores. Existen otros estudios cuantitativos dirigidos a la sociedad en general como el de Tim Davies en 2010 sobre los datos abiertos en Gran Bretaña. Schmidt, Gemeinholzer y Treolar en 2016, elaboran una investigación cuantitativa para medir en base a una serie de indicadores, la calidad y el nivel de reutilización de datos relacionados con el medioambiente. En México, se desarrolló un estudio en el que se analizaba el entorno de los datos abiertos y la percepción de los mexicanos a través de una encuesta Nacional de Acceso a la Información Pública y Protección de Datos Personales (Estrada Zamora y Vargas-Hernández, 2018). Estos estudios, pioneros, se centran en la población en general o en datos abiertos de investigación. La comunidad científica, personal altamente cualificado y formado dentro de la sociedad, son potenciales reutilizadores de datos y, por ende, pueden encontrar en los portales de datos abiertos, fuentes de información reutilizables para el desarrollo de investigaciones científicas. Por tanto, se hace necesario tener en cuenta la percepción y la opinión que tienen los investigadores ante lo que son los datos abiertos procedentes del movimiento de gobierno abierto, los elementos de relevancia, las tipologías de datos que pueden ser útiles, y los formatos técnicos. Consecuentemente, se necesita conocer la percepción del investigador y saber si encuentran en los portales de datos abiertos fuentes de información potencialmente reutilizables para el desarrollo científico.

Esta tesis doctoral pretende dar a conocer la percepción y la opinión de los investigadores ante el movimiento de los datos abiertos procedentes del sector público. Para ello, se ha utilizado una metodología cuantitativa, realizando una encuesta. Se utiliza la plataforma de

encuestas de SurveyMonkey. Previamente se genera una base de datos con direcciones de correo electrónico procedentes de distintas universidades españolas<sup>4</sup> y se lleva a cabo la campaña de distribución de la encuesta por correo electrónico. Con los resultados se desarrollan una serie de visualizaciones y se extraen las conclusiones sobre la situación y percepción actual de los investigadores ante los datos abiertos y su uso en el desarrollo científico. Finalmente, una vez limpios y analizados los datos, se desarrolla una plataforma de visualización de datos para su consulta.

## Objetivos

### Objetivo general

- Conocer la percepción de los investigadores sobre los datos abiertos y su movimiento, además de cuantificar el uso de estos y reutilización en la ciencia.

### Objetivo específico

- Comprobar si existen sesgos por género, edad y área de investigación en el conocimiento y la reutilización de los datos abiertos.
- Identificar los atributos, servicios, funcionalidades y motivaciones que la comunidad científica considera relevantes para el uso y reutilización de datos abiertos gubernamentales y de la administración, así como las barreras a las que se enfrenta.
- Conocer qué tipos de portales de datos abiertos (continentales, estatales, autonómicos, locales u otros) son los más consultados como fuentes de información y las tipologías de datos más reutilizadas en la producción científica.
- Identificar tipos de proyectos llevados a cabo con la reutilización de datos abiertos.
- Conocer la tecnología utilizada para llevar a cabo la descarga, procesamiento, limpieza y explotación de los datos, así como los formatos técnicos mejor valorados por los investigadores.

---

<sup>4</sup> No se distribuye la encuesta a aquellas universidades que en sus directorios de correo especifican que no se pueden utilizar las direcciones de correo para este tipo de actividad.



## Estado de la cuestión.

Los datos abiertos y la reutilización de información del sector público.

## Estado de la cuestión

### 2.1 Contextualización y definición de los datos abiertos: primeras aproximaciones

Hasta no hace mucho tiempo, en los programas electorales de los partidos políticos no se hablaba de los términos descritos en el apartado anterior: la transparencia, el gobierno abierto y la participación ciudadana. Actualmente se hacen necesarias este tipo de acciones de empoderamiento ciudadano y de interacción directa entre la población y la administración. Es una forma de rendir cuentas y fomentar la participación. La política actual deja de hacer las cosas para los ciudadanos, haciendo política con los ciudadanos (Garriga-Portolà, 2011).

El mundo está siendo testigo de una transformación global sin precedentes, de la mano de la tecnología y los medios digitales, alimentada por datos e información. Esta transformación tiene un enorme potencial y es la que permite fomentar gobiernos más transparentes, responsables, eficientes, receptivos y efectivos. Y podemos afirmar que los datos están en el centro de este cambio de paradigma global. De ahí la importancia de los datos abiertos (Gurstein, 2011; Ferrer-Sapena, Peset y Aleixandre, 2011).

En la construcción de una sociedad más justa, próspera y equitativa, debemos tener realmente en cuenta la transparencia y la responsabilidad de quienes dirigen nuestros gobiernos. Estos se deben comprometer de manera regular y significativa con los ciudadanos. Existe una revolución global de datos, que persigue avanzar en la colaboración en torno a desafíos sociales, proporcionar una supervisión pública efectiva de las actividades gubernamentales y apoyar la innovación, el desarrollo económico sostenible y la creación y la expansión de políticas y programas públicos efectivos y eficientes. Los datos abiertos son cruciales para poder cumplir estos objetivos (International Open Data Charter, 2015).

Para que esta nueva concepción de gobierno, abierto, basado en políticas de participación ciudadana y transparencia, tenga un impacto directo en la sociedad, las administraciones deben abrir el acceso a determinados conjuntos de datos gubernamentales. Estas iniciativas, junto al crecimiento del uso de las tecnologías, los conocimientos y la alfabetización informacional por parte de la ciudadanía, fomentan la participación y la innovación, a través de la creación de servicios, aplicaciones y resultados de investigaciones que pueden ser de utilidad para el resto de la sociedad (Garriga-Portolà, 2011; Cerrillo-Martínez, 2018).

Estas políticas tienen como objetivos:

- La publicación de todos los datos de libre disposición en poder de la administración.
- Permitir a la ciudadanía tener un mejor conocimiento de la actividad del sector público.
- Facilitar la creación de productos y servicios de información basados en datos.

- Facilitar el uso de datos para que empresas y particulares desarrollen y ofrezcan productos y servicios de información de valor añadido.
- Favorecer la competencia de mercado.

Dentro de la comunidad científica encontramos a potenciales reutilizadores de datos abiertos. Existe un consenso en la academia sobre el valor de los datos abiertos en el mundo de la investigación, pero es necesario establecer vínculos y buenas prácticas en su uso. Los datos abiertos pueden ser un elemento importante para el desarrollo de competencias y habilidades en la educación universitaria a parte de las investigaciones que se pueden llevar a cabo (Atenas, Piedra, Priego y Havemann, 2015). Los principales beneficios de los datos abiertos en la comunidad científica son:

- El uso de los datos abiertos va a permitir tener la posibilidad de replicar procesos de investigaciones que llevan a cabo resultados plasmados en artículos científicos mejorando el desarrollo de capacidades de investigación.
- Promueven la colaboración científica y potencian el aprendizaje al colaborar en investigaciones basadas en datos de estudios científicos y gubernamentales.
- Impulsan la educación abierta, clave en las universidades que creen portales donde liberar conjuntos de datos.
- Se maximiza así el acceso a fuentes de datos abiertos, científicos y gubernamentales.
- Fomenta la innovación, permitiendo a docentes y estudiantes trabajar con la materia prima que son los datos.

Las universidades y las instituciones de investigación tienen la responsabilidad de crear entornos adecuados que sirvan de apoyo para los datos abiertos. Deben ser capaces de gestionar, preservar y realizar análisis de datos, además de dar soporte técnico y crear servicios de biblioteca y gestión de datos. Estas organizaciones deben desarrollar incentivos y criterios de promoción para aquellos organismos involucrados en los procesos de datos abiertos. Actualmente existe un espíritu de internacionalización entre universidades y otras instituciones científicas que permiten establecer marcos colaborativos con países en desarrollo para movilizar las capacidades de uso intensivo de datos abiertos (ICSU-IAP-ISSC-TWAS, 2016).

Por otra parte, la legislación de cada país debe asegurar la integridad y la fiabilidad de estos datos, y determinar, entre otros aspectos, el tipo de licencia bajo el que deben distribuirse.

Existen entidades y organismos que trabajan por la apertura, el fomento y la reutilización de los datos abiertos. Estas entidades se encargan de definir y acotar el concepto, además de procurar que los datos procedentes de las administraciones y los gobiernos cumplan con las características dadas en la definición. Para precisarla y que la definición sea lo más completa

posible, utilizaremos estas fuentes de información institucionales y académicas que trabajan a favor los datos abiertos.

## 2.2 Datos abiertos: definición y características

Las administraciones públicas y los gobiernos disponen de muchos conjuntos de datos que, en buena parte, pueden ser abiertos a la sociedad sin que implique ningún problema de privacidad (Garriga-Portolà, 2011). La apertura de conjuntos de datos puede que cumpla con las características del *big data*<sup>5</sup>, pero los datos abiertos no son necesariamente masivos y pueden ser de cualquier dimensión (Bertot et al., 2014). Distintas entidades y miembros de la academia llevan tiempo trabajando con datos abiertos. Y no solo explotando el potencial y llevando a cabo su reutilización. Trabajan por el impulso y el fomento de iniciativas de gobierno abierto basadas en la apertura datos. Y a pesar de que existen muchos puntos en común en los distintos enfoques, se han definido de varias formas los datos abiertos. Por ejemplo, la Carta Internacional de Datos Abiertos<sup>6</sup>, es un conjunto de principios y buenas prácticas para la liberación de datos de las administraciones gobiernos<sup>7</sup>. En ella se definen los datos abiertos como *datos digitales que están disponibles con las características técnicas y legales necesarias para que puedan ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, en cualquier momento y en cualquier lugar*.

En el año 2015, se establecen los seis principios de la Carta, desarrollados por la sociedad civil y expertos de todo el mundo, que representan las aspiraciones de los agentes interesados. Son un conjunto de normas acordadas a nivel mundial para la publicación de datos. Estos principios establecen que los datos abiertos deben ser (Open Data Charter, 2015):

- Abiertos por defecto: datos en poder de gobiernos nacionales, regionales, locales, organismos gubernamentales internacionales y otros tipos de instituciones en el sector público en general. Son datos de gran valor para la sociedad, la economía y, por tanto, los gobiernos los deben tener abiertos de forma predeterminada.
- No comprometen el derecho a la privacidad, y la ciudadanía tiene derecho a influir en la recopilación y el uso de sus propios datos personales o de los datos generados como resultado de sus interacciones con los gobiernos.

---

<sup>5</sup> Los datos masivos son datos a gran escala, con gran volumen y complejidad, que precisan de aplicaciones informáticas no tradicionales para poder tratarlos adecuadamente ya que superan las capacidades de los sistemas informáticos habituales (Fundéu RAE, 2013).

URL: <https://www.fundeu.es/recomendacion/macrodatosalternativa-abig-data-1582/>

<sup>6</sup> Carta Internacional de Datos Abiertos. *Definición Datos Abiertos*.

URL: <https://opendatacharter.net/principles-es/>

<sup>7</sup> Carta Internacional de Datos Abiertos. *Wikipedia*. Última revisión: 24 nov. 2020.

URL: [https://es.wikipedia.org/wiki/Carta\\_Internacional\\_de\\_los\\_Datos\\_Abiertos](https://es.wikipedia.org/wiki/Carta_Internacional_de_los_Datos_Abiertos)

- Adecuados y comprensivos: pueden requerir tiempo, recursos humanos y técnicos para su publicación. Se deben consultar los usuarios de los datos (ciudadanos, trabajadores de la administración, sector privado) para identificar qué datos priorizar en la publicación y mejora. Es necesario tener en cuenta que los datos deben ser fidedignos, deben mantenerse actualizados en la periodicidad establecida y correspondiente y deben ser datos de alta calidad. De no ser así, los reutilizadores de datos es posible que dejen de usarlos si no encuentran la utilidad de sus desarrollos.
- Accesibles y utilizables: permiten a los ciudadanos, sociedad civil y otras organizaciones una mejor toma de decisiones. Deben, por tanto, ser datos fácilmente accesibles y sin barreras burocráticas o administrativas, disponibles en formatos abiertos para poder procesarse y ser legibles por máquina, y divulgados de forma gratuita.
- Comparables e interoperables: para ser útiles y eficientes, deben ser fáciles de comparar dentro y entre sectores, a través de ubicaciones geográficas y en el tiempo.
- Deben mejorar la gobernanza y participación ciudadana: la publicación de datos sirve para fomentar un mejor desarrollo, implementación y evaluación de programas y políticas para satisfacer las necesidades de los ciudadanos. Paralelamente, sirven para fomentar la participación ciudadana y un compromiso entre gobierno y ciudadano.
- Desarrollo inclusivo e innovación: el último principio reconoce la importancia de la apertura para estimular la creatividad y la innovación. Cuanto más se utilicen los datos abiertos, mayor serán los beneficios sociales y económicos. A parte, pueden ayudar a identificar desafíos sociales y económicos y monitorear programas de desarrollo sostenible.

Otra definición de referencia la ofrece la *Open Knowledge Foundation*<sup>8</sup>, una fundación sin ánimo de lucro que lucha por la difusión del conocimiento y la ciencia abierta. El conocimiento abierto es información, contenido y datos que pueden ser libremente utilizados, reutilizados, redistribuidos, sin ningún tipo de restricción legal, tecnológica o social. Cuando los datos abiertos son útiles, se utilizan y generan valor, se convierten en *conocimiento abierto*. Por tanto, asegura la calidad y compatibilidad del material en abierto. Como consecuencia, los datos abiertos tienen las siguientes características clave (Open Knowledge Foundation, s.f.):

- Disponibilidad de acceso y participación universal. Se trata de datos disponibles en su totalidad, a un costo razonable, y a poder ser, descargables a través de internet con el objetivo de que puedan ser tratados. Además, esta disponibilidad debe ser sin ningún tipo de discriminación o restricción para cualquier persona o colectivo.

---

<sup>8</sup> Open Knowledge Foundation. *Definición y misión de los datos abiertos*: <https://okfn.org/opendata/>

- Reutilización y redistribución: legibles y procesables por máquina, así como combinables con otros datasets.

La Open Knowledge Foundation dispone de un manual, el *Open Data Handbook*<sup>9</sup>, que recoge los aspectos legales, sociales y técnicos para una óptima apertura de datos. Justifica la importancia de la apertura de datos a la ciudadanía y marca una serie de pautas sobre el porqué, qué es y cómo abrir datos. Incide en la importancia en el significado “abierto” haciendo referencia a la interoperabilidad. Es decir, a la habilidad de sistemas y organizaciones de integrar y trabajar con distintos conjuntos de datos. De este modo se pueden desarrollar nuevos productos y servicios más grandes y complejos (Open Data Handbook, s.f.).

Otra entidad de renombre que aborda la definición de datos abiertos es el Banco Mundial. Es una entidad internacional cuya especialización son las finanzas. Tienen como objetivo reducir la pobreza a través de préstamos con intereses bajos y apoyar económicamente a países en desarrollo (Banco Mundial, s.f.). Ven en los datos un componente clave en cualquier estrategia general para el desarrollo, que permiten fijar niveles de referencia, identificar medidas públicas y privadas eficaces, determinar objetivos, supervisar avances y evaluar efectos y consecuencias. Hablan de la *“disponibilidad para que determinados conjuntos de información estén a disposición pública para ser utilizada para cualquier propósito, a bajo coste o de manera gratuita. Cualquier tipo de datos, desde datos sobre actividad económica, de medio ambiente o de consumos culturales”*.

Aporta además un punto de vista interesante, ya que, *“una estrategia de datos no consiste simplemente en una diseminación de estos de manera aislada”*. Para fomentar la transparencia y reducir la opacidad política, se debe permitir el uso de datos ya que, gracias a ellos, se pueden explicar procesos y fenómenos de cierta dificultad (Banco Mundial, s.f.).

Pero no solo dan soporte y fomentan el desarrollo de iniciativas de datos abiertos este tipo de instituciones sin ánimo de lucro. Existen numerosas administraciones y gobiernos comprometidas con este tipo de iniciativas. Por ejemplo, desde la Unión Europea existe un programa de fomento y promoción de la reutilización de datos abiertos: *European Data Portal*<sup>10</sup>. Hacen referencia a la definición que aporta la Open Knowledge Foundation (definición anterior) y en base a esa definición, proponen que, *“los datos abiertos están legalmente abiertos, publicados con una licencia abierta y bajo unas condiciones de reutilización limitadas a la atribución”*. Por otra parte, consideran abiertos si *“el archivo es legible por una máquina y no es propietario en la medida de lo posible, considerando que*

---

<sup>9</sup> Open Data Handbook. *Datos abiertos*. URL: <https://opendatahandbook.org/guide/es/what-is-open-data/>

<sup>10</sup> European Data Portal. *¿Qué son los datos abiertos?* URL: <https://data.europa.eu/es/trening/what-open-data>

*son accesibles para todos y el formato de archivo y su contenido no están restringidos a una herramienta de software de código no abierto en particular” (Data.europa.eu, s.f.).*

Además de las instituciones y organismos gubernamentales, miembros académicos centran parte de su investigación en las iniciativas de los datos abiertos. Aportan valor añadido a las definiciones dadas por las distintas instituciones que delimitan y trabajan con y por los datos abiertos. Por ejemplo, Garriga-Portolà en 2011, definió los datos abiertos como *el proceso que pone al alcance de la sociedad los datos públicos de los que dispone la Administración, en formatos digitales, estandarizados y abiertos, siguiendo una estructura clara que permita su comprensión y reutilización. Además, gracias a estos procesos de apertura, los ciudadanos y las empresas pueden acceder a estos datos para informarse o crear nuevos servicios que incrementan el valor social y en algunos casos, el valor comercial de la información.*

Es de vital importancia que estos datos estén en formatos y condiciones que faciliten la reutilización y que no estén sometidos a restricciones técnicas y jurídicas que limiten la reutilización por estar en formatos que no son procesables automáticamente o por la necesidad de proteger la propiedad intelectual y datos personales (Cerrillo-Martínez, 2014).

Deben ser completos, primarios y en bruto (no haber sido tratados previamente), accesibles física y electrónicamente, procesables automáticamente, no discriminatorios, en estándares abiertos y no sometidos a regulación que restrinja su reutilización, disponibles de forma permanente y sin costes (Cerrillo-Martínez, 2018). Con esta serie de características, el objetivo y la filosofía de los datos abiertos es poder seleccionar conjuntos de datos que en sí mismos no ofrecen ningún valor y dotarlos de contexto para convertirlos en conocimiento y valor para la sociedad (Llinares, 2010).

Estas iniciativas de apertura de datos van a permitir alcanzar distintos objetivos. En primer lugar, pueden contribuir a cumplir con uno de los pilares del gobierno abierto que es la transparencia pública. Los datos abiertos van a facilitar el conocimiento por parte de los ciudadanos de las decisiones que toman las administraciones públicas y entidades gubernamentales además del funcionamiento de su actividad. Ayudan a explicar sus acciones a los ciudadanos a la vez que añaden nuevas utilidades y hacen más accesibles los servicios públicos existentes, sin un coste adicional para la administración. Es decir, permiten un seguimiento de la actividad de las administraciones públicas y facilita la rendición de cuentas a la ciudadanía a través de la difusión de información (Cabo y Magallón-Rosa, 2013; Cerrillo-Martínez, 2018).

Otro objetivo de los datos abiertos es la contribución en la promoción de la integridad pública. Es decir, van a ayudar en la lucha y prevención contra la corrupción y en la detección de anomalías administrativas y gubernamentales. Permiten identificar situaciones que pueden

responder a posibles irregularidades producidas e incluso anticipar situaciones que faciliten el surgimiento de futuros fraudes (Tascón, 2015).

Finalmente, contribuyen a la eficacia y eficiencia administrativa. La reutilización de datos abiertos permite tener un mejor conocimiento del funcionamiento de los procesos administrativos y consecuentemente, la posibilidad de mejorar de estos servicios. El análisis de datos va a permitir una mejor toma de decisiones, y como consecuencia, una mejora en la eficiencia administrativa, además de anticiparse a determinadas situaciones (Cerrillo-Martínez, 2018).

Se debe tener en consideración que el simple hecho de que existan datasets que estén disponibles para ser leídos o descargados en internet no convierte a los datos en datos abiertos. Para considerarlos datos abiertos debe darse la posibilidad de descargarlos, pero, además, deben poseer algún tipo de licencia legal para poder utilizar, reutilizar y distribuirlos, pudiendo llevar a cabo incluso cruces con otros conjuntos de datos y se debe atribuir la autoría de los datos, además de compartirlos (Hernández-Pérez y García-Moreno, 2013). También se debe considerar que cuando se habla de datos abiertos no solo se refiere a datos digitales. Tienen cabida también los datos disponibles en documentos impresos depositados en archivos y bibliotecas, y que en muchas ocasiones no están digitalizados. El acceso a estos datos está vinculado a ejercicios de transparencia (Ferrer-Sapena y Sánchez Pérez, 2013).

Para el desarrollo de políticas de apertura de información, las entidades que se encargan de la apertura y del fomento de la reutilización de los datos abiertos, desde un primer momento deben ser diseñados con respaldo de un marco legal en materia de transparencia y reutilización de información del sector público, así como en materia de protección de datos de carácter personal, delimitando la responsabilidad de sus prestaciones. Por tanto, en lo referido a información de carácter público, estas iniciativas se llevan a cabo de forma respetuosa, con legislación actualizada y haciéndola accesible a quienes la quieran utilizar (Galindo Ayuda, 2014). Concretamente, en nuestro país, el acceso a la información pública y la reutilización de información del sector público, la transparencia y la apertura de datos se regulan a través de legislación, normativas y directivas europeas, estatales y autonómicas<sup>11</sup>.

Numerosas entidades como la *Open Knowledge Foundation* y su *Open Data Handbook*, el *Open Data Charter*, el Portal de Datos Abiertos de la Unión Europea o el *World Bank Open Data*, además de académicos, se encargan de aportar una definición y de trabajar por la apertura de datos y su reutilización. Se trata de fuentes activas, actualizadas y que fomentan el uso e innovación basado en datos. Atendiendo a lo expuesto, y con el objetivo de delimitar el concepto, los datos abiertos son:

---

<sup>11</sup> Marco legal y normativas definido en apartado 2.9 del Estado de la Cuestión.



- Datos en poder de gobiernos y otras entidades del sector público. Se trata de información pública con un gran valor para la sociedad. Su apertura se financia con fondos públicos.
- El objetivo es que sean descargados, explotados, combinados, reutilizados, redistribuidos por cualquier persona y llegar incluso a monetizar los productos o servicios desarrollados.
- Sin ningún tipo de restricción legal, tecnológica o social. Únicamente se debe de informar de la atribución de fuente de los datos.
- Datos con licencia, completos, en bruto, de forma gratuita (o a un costo razonable), como características de disponibilidad de acceso.
- Liberados en los portales de datos abiertos y, por tanto, deben ser fácilmente comprensibles y actualizados en la periodicidad que establece la descripción del conjunto de datos.
- Reutilizables, interoperables y redistribuibles, con características ineludibles de los datos abiertos. Es decir, estos datos deben ser procesables por máquinas, y deben poder cruzarse o trabajar con otros *datasets* para generar valor con ellos.
- Para llevar a cabo su apertura, existe un marco legal que los ampara. No comprometen la privacidad del ciudadano y estos pueden influir en la recopilación y uso de sus propios datos personales.

Y como consecuencia, si los datos cumplen con las características descritas anteriormente, las iniciativas de apertura generarán:

- Transparencia política y rendición de cuentas: ya que un gobierno que publica sus datos sin restricciones aporta un valor social que genera confianza y seguridad por parte del ciudadano con su gobierno. Mejorará además la seguridad y fiabilidad de los datos y permitirá obtener un diseño más eficiente y cercano de los servicios públicos.
- Participación, empoderamiento ciudadano y control democrático: con datos abiertos en portales de información pública, el ciudadano o reutilizador tendrá la capacidad de crear valor socioeconómico con esos datos. Es decir, podrá generar servicios, productos y aportar soluciones a problemas sin resolver en lo que pueden los propios desarrolladores o reutilizadores de los datos. Paralelamente, de estos desarrollos tecnológicos, se beneficiará la sociedad con el servicio creado e incluso la propia administración. El ciudadano se sentirá involucrado con el gobierno y este permite su participación en él.
- Innovación y conocimiento: viene relacionado con el punto anterior. Una sociedad con datos disponibles para su explotación y reutilización será una sociedad que tiene

la posibilidad de innovar sintiéndose integrada en ella. El cruce de datos con el desarrollo de servicios generará conocimiento.



Figura 1: Principios de los datos abiertos<sup>12</sup>.

Por tanto, queda claro que el *Open data* consiste en la puesta de los datos a disposición de la sociedad de manera libre, con procedencia de diferentes organizaciones, fundamentalmente de la administración pública o de proyectos que han sido financiados con fondos públicos. El objetivo es abrir los datos a la sociedad para construir sobre ellos una nueva idea, conocimientos e incluso servicios (Concha y Naser, 2012).

Los datos abiertos suponen un cambio de la cultura organizativa, donde la sinergia que genera apertura de información junto con la intervención ciudadana permite encontrar un elemento que genera valor añadido a la sociedad. Sirve como instrumento de medición de calidad democrática, ya que nos permite evidenciar necesidades, problemas, posibilidades y oportunidades de las entidades públicas. Favorece la intervención directa por parte del ciudadano en asuntos públicos y le empodera para ejercer un control continuado de acción política y empresarial que contribuye a alimentar un marco necesario dentro de la transparencia y rendición de cuentas (Álvarez García, Gértudrix Barrio y Rajas Fernández, 2014).

Para que esta información, procedente de organizaciones gubernamentales y el sector público, sea útil y reutilizable, la International Organization for Standardization (ISO)

<sup>12</sup> Diseño propio de la figura basada en distintas fuentes.

enriquece el concepto destacando la importancia de los estándares para conseguir un nivel de ordenamiento óptimo en los distintos contextos, mediante procesos que orientan hacia usos comunes y repetidos ante problemas reales o potenciales (ISO, 2011). La importancia de la estandarización de la información es una exigencia necesaria a los datos abiertos, ya que facilitará su recuperación, análisis, gestión, explotación de datos en bruto y reutilización. Garantiza además la calidad de los conjuntos de datos, y podemos afirmar que existe un vínculo entre la estandarización y el empoderamiento: son datos brutos estructurados publicados en red, interoperables y, por tanto, se pueden conectar entre sí para generar nuevos productos o servicios (Álvarez García, Gértudrix Barrio y Rajas Fernández, 2014). Las entidades gubernamentales llevan a cabo políticas de apertura de información que tienen como consecuencia la publicación de datos abiertos. Estas iniciativas se desarrollan mediante herramientas tecnológicas para la publicación de los datos abiertos (en red). El resultado o el punto de acceso son las plataformas digitales que conocemos como portales de datos abiertos.

### 2.3 Plataformas de publicación de datos: el portal de datos abiertos como fuente de información

En general, cuando se hace referencia a las herramientas utilizadas en las iniciativas de apertura de datos, existe una tendencia a denominarlos “Catálogos de Datos”. Estas herramientas cumplen, generalmente, con otras funciones complementarias a simplemente la clasificación y descripción de los datos. Por tanto, es adecuado empezar a hablar de “Plataformas de publicación de datos” o “Portales de datos abiertos”. (Datos.Gob.es, 2015).

Los portales de datos abiertos son una parte fundamental en las estrategias de gobierno abierto y de apertura de información del sector público. Se trata de plataformas digitales que permiten almacenar, compartir, conectar y visualizar bases de datos. Son el punto de acceso a la estrategia de apertura de información de las organizaciones y sirven como punto de encuentro entre ciudadanía, entidades encargadas de la apertura y reutilizadores de información del sector público (Moneo, 2014). Se respalda en dos pilares. En primer lugar, las tecnologías de la información y las comunicaciones, que permiten tanto el desarrollo de la plataforma, como llevar a cabo la apertura de la información a la ciudadanía. En segundo lugar, el voluntarismo de la propia entidad encargada de la apertura permitirá llevar a cabo políticas y estrategias de gobierno abierto que fomenten los principios de transparencia, colaboración e innovación (Arquero Avilés y Marco Cuenca, 2014).

Los portales de datos abiertos son, por tanto, el elemento más visible de las políticas e iniciativas de Gobierno Abierto y una de las principales herramientas de acceso a los recursos informativos del sector público y entidades gubernamentales (Martín González, 2013). Son el producto tangible en el que se ven materializadas dichas estrategias políticas de las diferentes administraciones públicas en torno al gobierno (Arquero Avilés y Marco Cuenca,

2014). Se convierten en una ventanilla única que la administración pone al servicio de la ciudadanía, centralizando los contenidos y evitando la dispersión entre diferentes portales de entidades públicas. Además, facilitan el acceso y la reutilización de información y deben responder a una vocación de servicio público, más allá de ser meras herramientas, repositorios y catálogos de datos (Pintos y Marín, 2013). Las plataformas de publicación de datos son una pieza clave a la hora de dar soporte tecnológico a la estrategia de apertura, con lo cual, se debe prestar atención a detalles concretos para conseguir objetivos particulares (Datos.Gob.es, 2015).

Si observamos la información disponible en estas herramientas, podemos considerar los portales de datos abiertos como fuentes de información potencialmente reutilizables de las que se puede beneficiar tanto la ciudadanía y la propia administración. Concretamente, una parte de la sociedad como es la comunidad científica, personal altamente cualificado y especializado en las distintas áreas de investigación dentro de la población, puede encontrar en este tipo de fuentes, información interesante y de relevancia para el desarrollo de nuevos estudios e investigaciones que aporten ese valor añadido que se busca al llevar a cabo las iniciativas de gobierno abierto y apertura de datos.

### 2.3.1 Estrategia de apertura y características de los portales de datos abiertos

Pero no solo se trata de abrir datos. Es conveniente elaborar estrategias que permitan alcanzar objetivos de gobierno abierto. Para ello, los portales de datos abiertos deben cumplir con una serie de requisitos, en materia de accesibilidad, aspectos técnicos y calidad para que los datos disponibles en estas fuentes sean reutilizados. Según Moneo (2014), las partes fundamentales de un portal de datos abiertos son:

- Catálogo de datos: como uno de los ejes fundamentales dentro de un portal de datos abiertos. Debe tener información sobre el contenido de los conjuntos de datos, nombre de la agencia responsable, frecuencia de actualización, visitas al dataset, información técnica, espacios de comentarios de usuarios e incluso, un apartado de valoración de la calidad de los datos abiertos.
- Colecciones de datos: se trata de conjuntos de tablas de datos que pueden tener un enfoque temático u orgánico, en función de si se organizan por la entidad que los publica o por tema concreto.
- Visualizaciones de datos: sección donde se pueden publicar gráficos y vistas de datos desarrolladas por reutilizadores. Algunos de los portales, en función de la tecnología que utilicen, permiten generar a través de módulos integrados, visualizaciones básicas, y en muchos casos, permiten conectar con herramientas externas para crear otro tipo de visores más complejos.

- Apartado de ejemplos prácticos: estos ejemplos prácticos se pueden presentar a través de historias de datos desarrolladas con la información disponible en los portales de datos abiertos y con listados de aplicaciones desarrolladas con los datos del portal.
- Apartado de desarrolladores: Dedicado a los reutilizadores que trabajan extrayendo información del portal para el desarrollo de nuevos productos y servicios basados en datos.

Para que los portales de datos abiertos consigan esta estructura y puedan abrir datos potencialmente reutilizables, el *Open Data Handbook* establece una serie de recomendaciones que pueden ayudar a crear plataformas sencillas de entender para la población y los reutilizadores. En primer lugar, recomiendan empezar llevando a cabo una apertura de datos simple que sea fácil de mantener. Paralelamente, recomiendan relacionarse desde el primer momento y con frecuencia con los reutilizadores, sean empresas, ciudadanos o la propia administración. Una buena comunicación permitirá obtener retroalimentación y establecer mejoras en el portal de datos abiertos, además de compartir unos mejores *datasets*.

Además, escuchar a los reutilizadores permitirá tener en cuenta qué conjuntos de datos pueden ser interesantes para ellos y en cuanto sea posible, compartirlos (si se dispone de ellos). La tercera recomendación hace referencia a establecer una comunicación directa y fluida con la propia entidad encargada de la apertura de datos. Se trata de solucionar dudas y temores que pueda tener la propia administración en cuanto a la compartición de datos (*Open Data Handbook*, s.f.).

El *Open Data Handbook* añade además cuatro pasos para realizar la apertura. Es de vital importancia hacer una buena selección de los conjuntos de datos que se quieren compartir. Una vez elegido/s los conjuntos de datos, se debe determinar el tipo de licencia abierta para determinar qué derechos de propiedad intelectual existen en la información. El siguiente paso es publicar la información en un formato reutilizable y considerar alternativas para trabajar con los datos como una API (Interfaces de programación de aplicaciones). La publicación de los datos implica que se deba trabajar previamente en una correcta descripción y en una buena clasificación en el catálogo disponible en portal de datos abiertos (*Open Data Handbook*, s.f.).

### 2.3.2 Estrategias técnicas de datos

La Iniciativa Aporta comenzó su andadura en el 2009 con el objetivo de fomentar la publicación de la información pública, los conjuntos de datos de estas entidades y fomentar su reutilización. Es un proyecto promovido por el Ministerio de Asuntos Económicos y Transformación Digital y la Entidad Pública Empresarial Red.es de España. En 2015 publicó el informe titulado “Plataformas de publicación de datos abiertos” en el que se establecieron una serie de herramientas y funcionalidades que deben tener las plataformas de datos.

Además de un catálogo de datos, deben dar soporte a los distintos elementos que forman parte de la iniciativa a través de todas las herramientas que conforman la plataforma (Datos.Gob.es, 2015).

En primer lugar, hace referencia al catálogo de datos como elemento principal dentro de los portales de datos abiertos. En él, se organizan y almacenan los datasets de forma que sea fácil su localización para su posterior reutilización. El catálogo debe contener (Datos.Gob.es, 2015):

- Soporte para metadatos: Los metadatos<sup>13</sup> son un elemento vital que proporciona información sobre los conjuntos de datos. Resultan de gran utilidad a la hora de organizar, clasificar, categorizar y buscar grandes volúmenes de datos. Permiten entender la finalidad de los datos y sus características básicas. Para poder analizarlos y tratarlos deben publicarse en formatos legibles y tratables automáticamente para poder realizar las búsquedas que se necesiten. Los catálogos de datos deben tener la capacidad de gestionar y exportar los metadatos descriptivos incluyendo al menos los propuestos en la Norma Técnica de Interoperabilidad para la Reutilización de recursos de información, propuestos en su anexo III y siguientes. Finalmente, la plataforma debe incluir la documentación y referencias sobre los metadatos.
- Compatibilidad de estándares: Uno de los objetivos de la reutilización de la información del sector público es poder cruzar e intercambiar datos de distintas fuentes. Por tanto, es importante que los datos estén publicados en estándares de referencia a nivel nacional e internacional, como la ya mencionada Norma Técnica de Interoperabilidad, el vocabulario para catálogos de datos DCAT definido por la W3C o el perfil de aplicación DCAT-AP para su uso como estándar en portales de datos europeos. Si se cumple con los estándares, proporcionará numerosas ventajas en temas de interoperabilidad e intercambio de datos.
- Soporte avanzado para Linked Open Data: Las estrategias de apertura muchas veces se elaboran pensando en el corto plazo. A la hora de medir la evolución de las estrategias, la publicación de los datos debe respaldarse de estrategias a medio o largo plazo. Por tanto, deben contar con soporte para la gestión y publicación de los datos utilizando tecnologías Linked Data y de Web Semántica.

### 2.3.3 Estrategias y servicios de información en los portales de datos abiertos

Con el catálogo se centraliza el acceso a los datos. Pero si se quiere facilitar el trabajo a los reutilizadores o potenciales usuarios de datos, se deberán incluir en la plataforma herramientas que permitan simplificar el acceso y mejorar la experiencia (Datos.Gob.es, 2015):

---

<sup>13</sup> Grupo de datos que describen el contenido de un conjunto de datos.

- Acceso a los datos: debe proporcionar un acceso directo a los datos a través de enlaces a los datasets y, además, es recomendable proporcionar interfaces de programación de aplicaciones (API's) para facilitar el trabajo de los reutilizadores. En cualquier caso, la existencia de este tipo de interfaces no debe sustituir el acceso a través de descargas directas.
- Búsqueda y clasificación: debe ser una de las funciones más cuidadas de la plataforma y los motores de búsqueda de estas herramientas, deben ser capaces de presentar resultados relevantes, tanto a través de búsquedas simples como a través de búsquedas avanzadas.
- Actualización: las herramientas deben ser capaces de tener mecanismos necesarios para gestionar cambios y correcciones que se producen en los datos existentes. Se debe establecer un sistema sólido y estable que mantenga la trazabilidad y el histórico de versiones. Además, es muy importante que la actualización sea automatizada, con opciones de datos en tiempo real.
- Visualizaciones: tal y como se ha citado anteriormente, las visualizaciones de información son elementos importantes a la hora de hacer entendibles y cercanos los datos disponibles en las plataformas. Los tipos de visualizaciones pueden llegar a ser muy amplios; la alternativa pasa por facilitar integraciones de sistemas especializados en el análisis y visualizaciones de datos ya existentes.
- Aplicaciones: un acceso a una galería de aplicaciones desarrolladas permite ver la utilidad de los datos de un modo tangible además de incentivar a los reutilizadores ayudándolos en la promoción de sus creaciones. Deben estar conectadas con los datos de un modo bidireccional de modo que a partir de los conjuntos de datos fuese posible acceder a las aplicaciones relacionadas y viceversa.

La Iniciativa Aporta hace referencia a la importancia de llegar a la ciudadanía para captar su interés y hacerlos partícipes de estas iniciativas y fomentar así la colaboración como pilar del gobierno abierto. Por tanto, las plataformas deben contener información donde se describa la iniciativa Open Data e incluir secciones informativas, con los conceptos más importantes, motivaciones y características básicas (Datos.Gob.es, 2015).

Debe contener además una sección de noticias y novedades, donde se comuniquen los principales avances y desarrollos de la iniciativa. Es muy importante tener documentación de soporte que sirva al reutilizador y un canal de contacto específico, además de guías de buenas prácticas y de estándares que se hayan aplicado en la plataforma correctamente documentadas y fácilmente localizables para los usuarios (Datos.Gob.es, 2015).

Es interesante contar con una sección con términos de uso y licencias en el que se defina la iniciativa de datos abiertos, con un marco legal único, claro y transparente, con el objetivo de proporcionar sostenibilidad y fiabilidad. En él, se debe informar a los usuarios sobre las

condiciones generales de reutilización y licencias aplicables a los datos en un vocabulario sencillo (Datos.Gob.es, 2015).

Finalmente, el documento “Plataformas de publicación de datos abiertos”, indica la importancia de publicar estadísticas de uso de los datos, como ejercicio de apertura y transparencia, que va a permitir una mejora colectiva en las iniciativas de datos abiertos. Las plataformas deben proporcionar estadísticas sobre acceso y uso de la plataforma, con números de visitantes, entradas, salidas. Debe mostrar estadísticas sobre el uso de datos (visitas a los datasets, descargas, etc.) y estadísticas sobre la gestión de los datos (actualización de los datos, datasets añadidos, distribución de conjuntos de datos por organismo, etc.) (Datos.Gob.es, 2015).

#### 2.3.4 Servicios enfocados a la comunidad de reutilizadores

Dentro de estas iniciativas es interesante contar además con servicios para la comunidad. El éxito reside en la capacidad de generar una comunidad de reutilizadores y usuarios de datos, con quien tener un diálogo activo y permanente que favorezca el crecimiento del portal de datos abiertos. Permitirá recibir retroalimentación a través de comentarios, ideas, sugerencias y peticiones de datos que sería interesante disponer. Además, puede resultar de interés integrar la comunidad en redes sociales para captar más audiencia y crear comunidades temáticas de reutilizadores (Datos.Gob.es, 2015).

#### 2.3.5 Tecnología y tipos de plataformas de datos

En ocasiones, existen entidades que no tienen la capacidad de llevar a cabo el desarrollo de una plataforma de datos abiertos con recursos propios y es necesario recurrir a terceros para la creación de las plataformas.

En el documento de la iniciativa Aporta para el desarrollo de plataformas de datos abiertos, se recogen algunas de las principales herramientas de datos disponibles. Destacan (Datos.Gob.es, 2015):

- CKAN. Data Management System: Promovida por la Open Knowledge Foundation con el objetivo de facilitar la publicación de catálogos de datos. Destacan como características principales de esta tecnología.
  - La publicación y gestión de datos, la recolección automatizada de datos de otros repositorios y la posibilidad de alojar datos en la propia plataforma.
  - Permite realizar búsquedas en base a los metadatos proporcionados. Se pueden incluir además metadatos personalizados.



- Contiene herramientas de comunidad para compartir conjuntos de datos a través de redes sociales, además de suscribirse a contenidos y actualizaciones de datasets.
- Existe la posibilidad de creación de publicaciones y previsualizar datos estructurados a través de tablas dinámicas o gráficos básicos.
- OGoov. Plataforma de Gobierno Abierto: Esta plataforma es desarrollada por Viavansi y ofrece una serie de funcionalidades combinables en función de la prioridad de la iniciativa, es decir, si está enfocada a apertura de datos, a transparencia o participación. En cuanto a la arquitectura, en función del objetivo de la estrategia, utiliza tecnología CKAN para el desarrollo de plataformas de datos abiertos, aunque el portal se construye con tecnología Wordpress. Permite además incorporar un Real Time Open Data (RTOD) para proveer de un modo dinámico, los conjuntos de datos abiertos en varios formatos, con soporte para Linked Open Data, a través de un servidor semántico desarrollado con Apache Marmotta y puntos de consulta en SPARQL.
- Socrata Open Data Portal: Se trata de una plataforma de publicación de datos en la nube que facilita la creación de iniciativas de datos abiertos. Permite como en los casos anteriores la publicación y gestión de datos, incorporando herramientas como hojas de cálculo, sistemas de archivo, bases de datos y otras fuentes en tiempo real. Permite además acceder y tener ejemplo de visualización de datos, a través de características de búsqueda avanzada mediante palabras clave, etiquetas y metadatos. Y finalmente, permite el acceso a datos en múltiples formatos.
- Junar. The Open Data Platform: Esta es otra plataforma de publicación de datos en la nube que permite elegir los conjuntos de datos a publicar, así como el tiempo y el modo en que se van a presentar a la ciudadanía. Destacan los siguientes servicios:
  - La plataforma sirve de puente a la hora de la recopilación de datos que se desean compartir en un punto común.
  - Permite mejorar los datos y convertirlos en información productiva gracias al uso de distintas visualizaciones, mapas y tablas de datos.
  - El proceso de publicación permite asignar permisos para cada usuario a la hora de recopilar, mejorar y publicar los datos.
  - Dispone de una interfaz de programación de aplicaciones (API) para que los desarrolladores puedan tratar y reutilizar datos más fácilmente.
  - Permite analizar el uso de los datos a tiempo real, gracias las estadísticas que genera la tecnología.

Estas plataformas vienen avaladas por su recorrido, madurez y soporte. Pero existen otras alternativas en el mercado como las desarrolladas sobre Drupal como la de datos.gob.es, disponible en abierto a través de CENATIC, DKAN o la de Data.gov.uk. Otras opciones

como Datapress, desarrollada con Wordpress y CKAN o OpenDataSoft, que ofrece un servicio de datos en la nube (Datos.Gob.es, 2015), también son de interés.

## 2.4 Evaluación de madurez, calidad, iniciativas y reutilización de datos abiertos

En este apartado se realiza una revisión bibliográfica sobre métricas y estudios que analizan la calidad de los datos abiertos, su nivel de reutilización e investigaciones relacionadas que evalúan la madurez de las iniciativas de datos abiertos.

### 2.4.1 Evaluación de madurez de los datos abiertos: las 5 estrellas de Tim Berners-Lee

En 2010, Tim Berners-Lee, inventor de la web y de los datos enlazados (Linked Data) sugirió un esquema para medir la calidad de los datos abiertos en función de su nivel de reutilización en base a una serie de indicadores y requisitos técnicos.

Es una propuesta basada en una escala de 5 estrellas, donde 1 estrella son datos con formatos difícilmente reutilizables debido a sus características, hasta 5 estrellas donde los formatos son completamente reutilizables y enlazables con otros (Berners-Lee, 2010; Datos.Gob.es, 2018):

1 estrella	Se alcanza con la publicación de datos. Generalmente se encuentran datos no estructurados, con dificultades para la manipulación y tratamiento. Por ejemplo, formatos imagen JPEG o formato PDF.
2 estrellas	Se trata de datos con formatos estructurados, pero en formatos propietario. Por ejemplo: formatos XLSX.
3 estrellas	Se trata de datos con formatos estructurados en software no propietario. Por ejemplo, formatos CSV o de software libre como Open Office.
4 estrellas	Son datos estructurados adecuadamente con URI's que sirven para identificar los recursos. Utilizan estándares RDF que facilitan la interoperabilidad.
5 estrellas	El mayor nivel de madurez de los datos abiertos en los que, además de cumplir con los requisitos anteriores, están enlazados con otros datos similares de otras organizaciones.

Tabla 1: Esquema de 5 estrellas propuesto por Tim Berners-Lee<sup>14</sup>.

<sup>14</sup> Tabla 1 y Figura 2: Tim Berners-Lee, 2010. Fuente: <https://5stardata.info/es/>

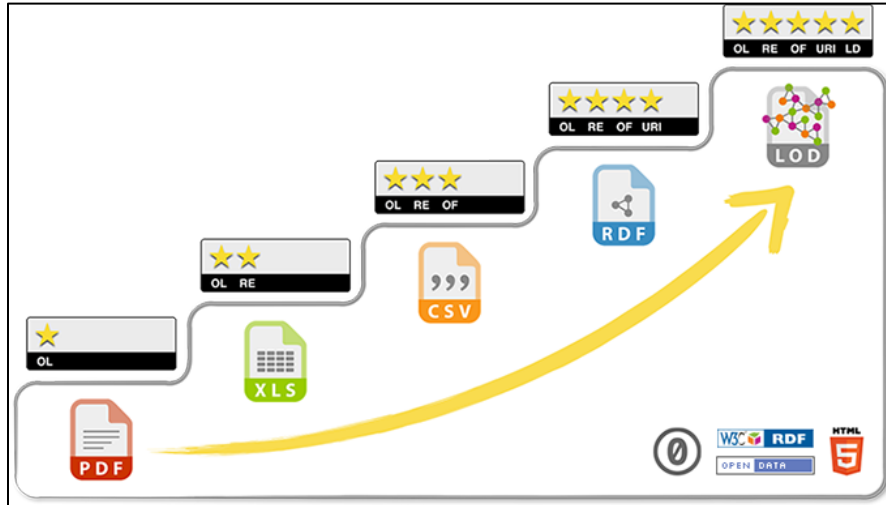


Figura 2: Esquema de 5 estrellas propuesto por Tim Berners-Lee.

#### 2.4.2 MELODA: Métrica de evaluación del nivel de reutilización de datos abiertos

La disponibilidad de mecanismos de evaluación de uso y valor de los datos abiertos es muy limitada. MELODA<sup>15</sup> es una métrica que permite evaluar la calidad de los conjuntos de datos disponibles en plataformas de datos abiertos, y su grado de reutilización. Nació en 2011 ante la falta de homogeneidad en los conjuntos de datos y su evaluación en los portales Open Data. En su primera versión realizaba la evaluación de los conjuntos de datos desde 4 dimensiones. Esta métrica ha ido evolucionando, llegando a la quinta versión publicada en 2020. Actualmente se basa en un total de 8 dimensiones de evaluación. Estas son (Abella, Ortiz-de-Urbina-Criado y De-Pablos-Heredero, 2014 y 2020):

<sup>15</sup> Métrica MELODA: <https://www.meloda.org>

Dimensiones de evaluación MELODA	
1. Licencia de los conjuntos de datos	1.1 Uso privado. 1.2 Reutilización no comercial. 1.3 Reutilización sin restricciones.
2. Acceso a la información	2.1 Acceso a web o parámetros de URL únicos. 2.2 Acceso web único con parámetros únicos. 2.3 API o lenguaje de consulta.
3. Formatos técnicos	3.1 Estándar cerrado reutilizable y abierto no reutilizable. 3.2 Estándar abierto reutilizable. 3.3 Estándar abierto, metadatos individuales.
4. Estandarización	4.1 Estandarización del modelo de datos propio. 4.2 Estandarización propia del modelo de datos publicado (armonización). 4.3 Estandarización local. 4.4 Estandarización global.
5. Contenido georreferenciado	5.1 Sin información geográfica. 5.2 Campo de texto simple o complejo. 5.3 Coordenadas o información geográfica completa.
6. Frecuencia de actualización	6.1 Más de 1 mes. 6.2 Mensual. El período de actualización oscila entre 1 mes y 1 día. 6.3 Diario. El período de actualización oscila entre 1 día y 1 hora. 6.4 Hora. El periodo de actualización va de 1 hora a 1 minuto. 6.5 Segundos. El periodo de actualización es inferior a 1 minuto.
7. Reputación	7.1 No hay información sobre la reputación de la fuente de datos. 7.2 Estadísticas o informes publicados de opiniones de los usuarios. 7.3 Indicadores o clasificaciones sobre la reputación de la fuente de datos.
8. Difusión	8.1 Comunicación / difusión no sistemática. 8.2 Recursos disponibles sobre las actualizaciones (p. ej., feed RSS). 8.3 Difusión proactiva / difusión push (información automática y oportuna).

Tabla 2: Dimensiones de evaluación MELODA<sup>16</sup>.

Una vez valorados los 8 niveles con los pesos asignados, se obtiene un resultado de 0 a 100. En función del porcentaje obtenido, el conjunto de datos tiene un nivel de reutilización.

Puntuación obtenida	Tipo de reutilización	
0-25	Reutilización deficiente	
25-50	Reutilización básica	
50-75	Buena reutilización	
75-100	Reutilización óptima	

Tabla 3: Nivel de reutilización<sup>17</sup>.

<sup>16</sup> Información extraída de Abella, et al., 2020.

<sup>17</sup> Fuente: Abella et al. (2014).

### 2.4.3 Otros estudios y métricas de evaluación de datos

Existen distintos informes internacionales encargados de evaluar la madurez de las iniciativas de datos abiertos. Por ejemplo, el Open Data Impact (2016), desarrollado por The GovLab y Omidyar Network, se encarga de categorizar, mediante el marco de análisis taxonómico dimensional, el impacto de iniciativas de datos abiertos en función de una clasificación y marcan una serie de recomendaciones enfocadas a las mejoras (Verhulst y Young, 2016). El Open Data Monitor plantea medir el impacto de los datos abiertos mediante una serie de indicadores que son los datos como recurso económico, la emergencia del ciudadano y el científico como ejercicio democrático, participación e inteligencia colectiva y los ciudadanos monitorizados en materia de seguridad y vigilancia (Atz, 2014).

El proyecto llevado a cabo desde The GovLab fomenta el desarrollo de estudios y proyectos sobre los nuevos modelos de gobernanza basados en datos alrededor del mundo. Estas acciones se desarrollan para mejorar la transparencia y eficiencia gubernamental y favorece el intercambio de información entre departamentos (Tinati, et. Al, 2012).

Destaca una métrica desarrollada por Pipino, Lee y Wang en 2002 que cuenta con 16 dimensiones de análisis, donde se evalúa la calidad de la información pública pero no las posibilidades de reutilización. Otro estudio llevado a cabo por Ren y Glismann en 2012, analiza la arquitectura de los datos de las organizaciones para llevar a cabo iniciativas de datos abiertos con una mejor calidad. Para ello se establece un proceso de 5 fases, donde se identifica la información relevante, se comprueba el estado de la información y posteriormente, se diseñan mecanismos de publicación. Al igual que en el estudio anterior, no se evalúa el nivel de reutilización de la información.

Otros informes que se encargan de evaluar la dimensión administrativa y política a través de sus modelos de datos como el Open Data Institute británico en colaboración con el Department for Environment Food and Rural Affairs en 2015, el estudio Open Data Maturity in Europe (2015) desarrollado por el Portal de Datos Abiertos de la Unión Europea, la Open Government Guide o el Open Government Data Toolkit llevado a cabo por el Banco Mundial de Datos abiertos (Álvarez-García, Gertrudis Casado y Gétrudix Barrio, 2021).

Según Zuiderwijk y Janssen (2014), es necesario establecer modelos comparativos sobre el alcance y metodologías de implantación de políticas de datos abiertos, donde podamos comprobar la diversidad de aproximaciones y establecer procesos de colaboración con otras organizaciones. La eficacia real que tiene la apertura de datos en materia de participación ciudadana y gobierno abierto se mide en otros estudios como el de Gurstein en 2011.

## 2.5 Reutilización de la información del sector público

Con las iniciativas gobierno abierto, las administraciones y entidades públicas proporcionan a la sociedad información en su poder para que cualquier persona e institución, sea cual sea su finalidad, pueda generar servicios de valor añadido. De hecho, en los últimos años se ha incrementado significativamente el uso de los datos (Alguacil-Sanz, 2015). Estas iniciativas de gobierno abierto y los estudios e informes revisados en el apartado anterior, que analizan y miden la calidad de los datos y su nivel de reutilización, tienen un objetivo común: la publicación de buenos datos para que estos sean reutilizados por la ciudadanía, teniendo en cuenta que la información del sector público es un recurso básico en la sociedad de la información.

La Reutilización de la Información del Sector Público (RISP) consiste en el uso de los datos o documentos que están en poder de las administraciones y organismos del sector público, por parte de personas físicas o jurídicas, con fines comerciales o no, siempre que este uso no constituya una actividad administrativa pública ni tampoco el intercambio de documentos entre administraciones públicas en el ejercicio de las funciones públicas que tengan atribuidas (Valero-Torrijos, 2014). Presenta un considerable potencial económico debido a los productos generados basados en los datos, ya que desarrolla nuevos puestos de trabajo en la industria de contenidos digitales (Portal de Administración electrónica, S.F.).

La reutilización de la información ha sido objeto de creciente interés, más allá de la industria infomediaria (que analizaremos en el siguiente apartado). Ha sido fomentada por las administraciones públicas, tanto regulando las condiciones en las que se debe llevar a cabo, como poniendo a disposición los conjuntos de datos en formatos abiertos, legibles por máquinas en plataformas de información destinadas a este fin (Cerrillo-Martínez, 2018).

En nuestro país, la reutilización de la información del sector público está regulada por la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, adaptada posteriormente en la Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para incluir los cambios normativos de la Unión Europea (Directiva 2013/37/UE). Esta adaptación hace mención a la necesidad de procesar y publicar la información con unas condiciones que faciliten su reutilización, a través de formatos estructurados, abiertos e interoperables, que garanticen la seguridad, así como la propiedad intelectual e industrial. Es importante que los datos estén completos, y que sean fiables y de calidad (En el apartado “Marco Legal” se desarrolla la legislación).

Con el marco legal español, existen numerosos organismos públicos en nuestro país que han elaborado políticas favorables a la reutilización. Destacan algunas como la de la Dirección General del Catastro, el Instituto Geográfico Nacional, la Agencia Estatal de Meteorología o

la importante labor de la organización sectorial como la Asociación Multisectorial de la Información (Asedie) y las empresas infomediarias españolas. Pero sin duda, con el fin de dar a conocer la importancia de la reutilización de la información en España y dinamizar en este sentido el sector público, nace en 2009 el Proyecto Aporta, promovido por los Ministerios de Política Territorial y Administración Pública y de Industria, Turismo y Comercio (Marcos-Martín y Soriano-Maldonado, 2011). Tiene como objetivo fomentar una cultura favorable de apertura de datos públicos, así como facilitar la apertura a las administraciones y con ella, impulsar el mercado de la reutilización de información pública (Datos.Gob.es, 2011, 2014a).

## 2.6 El Sector Infomediario

El valor de los datos abiertos reside en su uso y reutilización y tiene como ventaja un impacto y una trascendencia positiva a un coste casi nulo. Gracias al impulso de la legislación europea y española en materia de reutilización de información del sector público, donde las administraciones públicas fomentan y publicación de determinados tipos de datos en abierto, con el objetivo de que puedan ser utilizados, reutilizados y redistribuidos con o sin fines comerciales, podemos hablar de un nuevo sector de trabajo: El sector infomediario (Hípola y Muñoz-Cañavate, 2014).

El sector infomediario está compuesto por grupos de empresas que generan servicios para comercializar con la información disponible en plataformas de datos abiertos, promovidas desde el sector público. Este sector tiene como fin la creación de ideas y de nuevos productos a partir de información y datos públicos. Forma parte de las nuevas economías digitales y son la consecuencia de la reutilización de la información (Lorente, J., S.F.; Datos.Gob.es, 2014). Experimenta distintos avances, transformaciones y fracasos debido a factores tecnológicos, empresariales, políticos, sociales y económicos (Hípola y Muñoz-Cañavate, 2014).

La Unión Europea tiene en cuenta el crecimiento económico y del empleo que genera. Por tanto, se están desarrollando políticas de reutilización de la información y están presentes entre los objetivos de la Agenda Digital Europea, siendo el desarrollo del sector de los objetivos más importantes a corto plazo (Lorente, J., S.F.).

Para poder seguir creciendo, las empresas infomediarias proponen intensificar la labor pedagógica sobre los beneficios económicos de la reutilización y adaptar los recursos humanos en el sector. A las administraciones públicas se les exige mejorar su predisposición y publicar información potencialmente reutilizable, además de una mejor comunicación con las empresas del sector. Sería interesante disponer de un plan estratégico de publicación de datos para que las empresas pudieran encontrar en los portales de datos abiertos, otros conjuntos de datos interesantes para nuevos desarrollos. Finalmente, apuntan hacia el

desarrollo de estándares para la configuración de los datos y una normalización semántica entre los datos e información de las administraciones públicas (Datos.Gob.es, 2014b).

## 2.7 Los datos abiertos en la ciencia

Se considera que los datos son abiertos si proceden del movimiento de gobierno abierto en su apertura de datos (Peset, Ferrer-Sapena, Subirats-Coll, 2011) y, además, según la Open Knowledge Foundation, si cumple con las condiciones de disponibilidad íntegra de acceso, posibilidad de redistribución y licencias para su reutilización (Ferrer-Sapena, Peset y Aleixandre-Benavent, 2011).

Las distintas definiciones vistas en este apartado, procedentes tanto de organizaciones que fomentan la apertura de datos y su reutilización como de académicos que dedican parte de su actividad de investigación al estudio de este movimiento, hacen hincapié en la finalidad de la apertura: los datos abiertos son potencialmente reutilizables con cualquier fin. Uno de los objetivos finales de reutilización puede ser el objeto de estudio de esta tesis doctoral: la producción científica derivada del uso de los datos abiertos procedentes de fuentes de información del sector público.

### 2.7.1 Datos abiertos y datos de investigación

El movimiento Open Data (datos abiertos) está relacionado con otros conceptos como el Open Science (ciencia abierta), Big Data (datos masivos) y Data Government (datos gubernamentales). Pero existen otros conceptos relacionados con el acceso abierto a la información científica, como el Open Research Science (que se engloba dentro Open Access) y hace referencia al acceso abierto a las publicaciones científicas y el Open Research Data, que se refiere a los datos de investigación (Aleixandre-Benavent, Ferrer-Sapena y Peset, 2019).

No existe una definición concreta sobre qué son los datos de investigación. Por ejemplo, los National Institutes of Health (2003) y la Organización para la Cooperación y el Desarrollo Económicos (2021) definen los datos de investigación como “*todo aquel material que sirve para certificar los resultados de la investigación que se realiza, que se han registrado durante ella y que se han reconocido como parte de la actividad científica*”. Es decir, se trata de información reunida, utilizada o generada durante la experimentación, observación, medición, simulación, cálculo, análisis, interpretación, estudio o cualquier otro proceso de indagación que sustente y justifique las aportaciones científicas que se difunden en las publicaciones de investigación (Torres-Salinas, Robinson-García y Cabezas-Clavijo, 2012). Los datos de investigación pueden ser de muchos tipos: pueden ser numéricos, textuales, audiovisuales, digitales, físicos, fruto de cuestionarios, registros sonoros, visuales, cuadernos



de laboratorio y de campo, series numéricas en tablas, etc. (Aleixandre-Benavent, Ferrer-Sapena y Peset, 2019).

Compartir los datos es importante por diversas razones. En primer lugar proporciona beneficios para la sociedad, ya que permite validar los descubrimientos, aumenta la credibilidad en la ciencia, ofrece oportunidades económicas y es esencial para garantizar la competitividad. En segundo lugar, aporta beneficios para el propietario de los datos, ya que aumenta la accesibilidad a sus trabajos mejorando así su visibilidad, citación e impacto y permite establecer redes de trabajo colaborativo. Y finalmente, el compartir datos trae beneficios para la comunidad de investigadores, a través de oportunidades que brinda para descubrir nuevos conocimientos, reproducibilidad de trabajos, mejora la transparencia en la investigación, se pueden detectar fraudes y malas praxis, se reducen costes a través del trabajo colaborativo, y se aportan recursos para la educación y la formación entre muchas otras ventajas (Aleixandre-Benavent, Ferrer-Sapena y Peset, 2019).

### 2.7.2 Los principios FAIR y los datos abiertos.

La Royal Society (2012) introduce el concepto de apertura y define los datos abiertos como aquellos que cumplen ciertos criterios cuando son liberados, denominándose “apertura inteligente” o competitiva. Estos criterios están contemplados también en los principios FAIR, cuyas siglas son Findables (localizables), Accesibles, Interoperables y Reusables. (Wilkinson, M., Dumontier, M., Aalbersberg, I. et al., 2016; Aleixandre-Benavent, Ferrer-Sapena y Peset, 2017 y 2019; Universidad de Extremadura: Servicio de Bibliotecas, 2021).

- Findable: para poder reutilizar los datos, estos deben estar fácilmente localizables. Deben utilizar metadatos que permitan identificar, describir y localizarlos.
- Accesible: una vez localizados, el usuario debe saber cómo acceder a los datos. Pueden recuperarse a través de protocolos estandarizados de comunicación y los metadatos persisten aun cuando los datos dejen de estar disponibles. Los protocolos de acceso deben ser libres, abiertos y gratuitos.
- Interoperable: la importancia de los datos radica en que puedan ser integrados con otros datos de distintas fuentes o aplicaciones. Se deben presentar de forma que resultan aplicables e incluyen referencias a otros datos.
- Reusable: los datos pueden ser reutilizados por otros investigadores. Se debe clarificar la procedencia de los datos y sus condiciones de reutilización.

Estos principios se reflejan en la incorporación del Programa Horizonte 2020 de Investigación e Innovación de la Unión Europea, con el lema “tan abierto como sea posible, tan cerrado como sea necesario”. Por tanto, se busca equilibrar la apertura de datos con la protección de la información científica, los derechos de comercialización y la propiedad

intelectual, la privacidad, la seguridad, y las cuestiones relativas a la conservación y la gestión de los datos (Universidad de Extremadura: Servicio de Bibliotecas, 2021).

Según la definición ya vista de los datos abiertos del Open Data Handbook, los datos pueden ser utilizados, reutilizados y redistribuidos por cualquier persona, únicamente sujetos a la atribución de autoría. La diferencia entre los principios Open Data y los principios FAIR radica, principalmente, en el acceso a los datos. No todos los datos procedentes de iniciativas Open Data tienen por qué ser FAIR y viceversa. Los principios FAIR nacen en favor de una investigación científica abierta, y de la necesidad de cooperar y compartir datos de investigación en aras de poder reproducir los resultados e iniciar nuevas investigaciones. La diferencia con los Open Data es que estos nacen de iniciativas gubernamentales, es decir, la fuente principal son gobiernos y administraciones públicas. Buscan cumplir con los pilares de gobierno abierto que son la transparencia, rendición de cuentas y la participación ciudadana (Wilkinson, M., Dumontier, M., Aalbersberg, I. et al., 2016; Luque, 2019).

Los principios FAIR no son acumulativos, como sí lo son las ya citadas, cinco estrellas de Tim Berners-Lee. Dada la importancia de cumplir estos principios, la Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo de 20 de junio de 2019 relativa a los datos abiertos y la reutilización de la información del sector público los incluye en el texto para que se apliquen a los datos abiertos siguiendo el ejemplo de lo aplicado en los datos de investigación.

## 2.8 Otros trabajos académicos: revisión de estudios

Tras la contextualización y desarrollo del estado de la cuestión sobre el gobierno abierto y sus pilares, los datos abiertos y sus características, las iniciativas y las plataformas digitales de compartición de datos, y su relación y distinción con los datos de investigación, el objeto de esta tesis doctoral es conocer la percepción que tienen los investigadores ante los datos abiertos que proceden de iniciativas de gobierno abierto y si se han reutilizado datos procedentes de estas fuentes de información para el desarrollo de conocimiento científico.

Se trata de un estudio pionero a nivel estatal en España. Existen estudios centrados en medir la calidad de los datos y el nivel de reutilización, llevado a cabo por Abella, Ortiz-de-Urbina-Criado y De-Pablos-Heredero en 2017 y 2019. Otros trabajos se centran en datos de investigación, como el llevado a cabo por Aleixandre-Benavent, Vidal-Infer, Alonso-Arroyo, Peset y Ferrer-Sapena en 2020, en el que la población encuestada son investigadores, como la población encuestada en esta tesis doctoral, pero centrándose en los datos de investigación. Para el desarrollo de esta tesis doctoral la metodología de investigación elegida es a través de un método cuantitativo (vid. Desarrollado en apartado Metodología de investigación).

Existen estudios que utilizan métodos cuantitativos, como el que dirigió Tim Davies en 2010 sobre los datos abiertos en Gran Bretaña enfocado en medir la percepción y uso de los datos abiertos, pero centrándose en la población en general. En 2016, Schmidt, Gemeinholzer y Treolar, elaboran una investigación cuantitativa para medir en base a una serie de indicadores, la calidad y el nivel de reutilización de datos relacionados con el medioambiente. En 2018, en México, liderada por Estrada-Zamora y Vargas-Hernández, se desarrolla una encuesta en el que se analiza el entorno de los datos abiertos y la percepción de los mexicanos a través de una encuesta Nacional de Acceso a la Información Pública y Protección de Datos Personales.

Se trata de estudios que se centran en la percepción acerca del Open Data de la población en general, en México o Gran Bretaña. Otros estudios se centran en medir la calidad y nivel de reutilización de conjuntos de datos concretos destinados a áreas de investigación concretas, u otros estudios centrados en datos abiertos de investigación.

La comunidad científica es uno de los potenciales reutilizadores de datos y, por ende, pueden encontrar en los portales de datos abiertos, fuentes de información reutilizables para el desarrollo de investigaciones científicas. Por tanto, se hace necesario tener en cuenta la percepción y la opinión que tienen los investigadores ante lo que son los datos abiertos procedentes del movimiento de gobierno abierto, los elementos de relevancia, las tipologías de datos que pueden ser útiles, y los formatos técnicos, entre otros aspectos.

## 2.9 Marco legal: transparencia y acceso a la información del sector público

### 2.9.1 Introducción al marco legal

Con la irrupción de las tecnologías de la información y la comunicación, y la tendencia al alza del conocido “Gobierno Abierto” dónde el ciudadano ejerce una participación directa en la nueva manera de hacer política, estamos asistiendo a una fuerte tendencia a nivel internacional de apertura de datos (Clabo, Ramos-Vielba, 2015).

Existe un consenso en sociedades de distintos países donde se aprecian las potencialidades y beneficios de esta apertura de información: fomento de una cultura de gobierno abierto mediante la participación y colaboración ciudadana, y la monitorización y readaptación de políticas públicas y oportunidades de empleo asociadas a nuevos modelos de negocio (Clabo, Ramos-Vielba, 2015). Las principales fuentes de información “abiertas” están constituidas por datos de las entidades públicas, que, como se ha dicho, deben cumplir con unos requisitos mínimos de calidad para que puedan servir para desarrollar nuevos productos y servicios, para la participación ciudadana y para garantizar la transparencia (Cerrillo, 2014).

Para la puesta a punto de estos datos en portales de información pública, es necesario tener en cuenta la normativa vigente relacionada con la transparencia y los datos abiertos. La legislación diseñada con este fin respalda la protección de datos de carácter personal y delimita claramente las responsabilidades de las prestaciones de servicios, estableciendo pautas para llevarla a cabo. Así lo permiten y regulan las directivas europeas, las normativas y el marco legal estatal, referidas a la reutilización de datos públicos, acceso a la información y transparencia (Galindo Ayuda, 2014).

La información suministrada por las Administraciones (estatales, autonómicas, locales) debe contener notas y avisos legales referidos al uso la información (art. 7 Real Decreto 1495/2011). Estos avisos legales establecen en ocasiones una no obligatoriedad por parte de las Administraciones a la continuación, por ejemplo, de publicación de datos abiertos. Son los propios reutilizadores quiénes usan los datos públicos a efectos de que se establezcan acuerdos oportunos con los organismos públicos para garantizar la publicación continua de los mismos (Galindo Ayuda, 2014).

En cuanto al desarrollo de servicios a partir de los datos abiertos, se debe llevar a cabo teniendo en cuenta:

- La elaboración de propuestas jurídicas atendiendo a principios democráticos.
- La protección de datos de carácter personal.
- La garantía de responsabilidad sobre las propuestas de los servicios a los usuarios.
- El cumplimiento de requisitos administrativos en caso de provisión de servicios legales.

Por tanto, las iniciativas de datos abiertos deben ser diseñadas desde un primer momento desde la legislación vigente en relación con reutilización de la información del sector público, la protección de datos personales y deben delimitar la responsabilidad de las prestaciones de los servicios (Galindo Ayuda, 2014).

El presente capítulo recopila el marco legal y normativo a nivel europeo, estatal y autonómico, así como otros documentos en los que se establecen directrices en materia de reutilización de la información del sector público y transparencia.

Algunos portales de datos abiertos albergan dentro de sus contenidos apartados de normativa. El objetivo es hacer conocidos a los reutilizadores del marco legal que los ampara a la hora de utilizar los datos de las entidades públicas. Portales como el de la Junta de Castilla y

León<sup>18</sup> y la Junta de Andalucía<sup>19</sup>, el Ayuntamiento de Madrid<sup>20</sup> o Valencia<sup>21</sup>, entre otros, disponen de un apartado de legislación completo. Definen con precisión qué son los datos abiertos y para qué sirven. Algunos muestran guías de datos abiertos y buenas prácticas, catálogos de datos abiertos, con análisis de los mismos, aplicaciones que han ido desarrollando los reutilizadores o apartados de participación ciudadana.

La presentación de la normativa se ordena cronológicamente de acuerdo con los distintos ámbitos de actuación:

- **Ámbito comunitario:** Directivas y otra normativa establecida desde la Unión Europea, así como documentos de trabajo, recomendaciones a seguir y análisis de la normativa.
- **Ámbito estatal:** Legislación vigente y modificaciones de la normativa española.
- **Ámbito autonómico:** Legislación vigente y modificaciones de la normativa de la Comunidad Valenciana<sup>22</sup>.
- **Normativas técnicas y otros reglamentos** que afecten al territorio, europeo, español y Comunidad Valenciana.

## 2.9.2 Legislación y documentos en ámbito comunitario. Directivas europeas, documentos y recomendaciones

*Libro Verde sobre la Información del sector público en la Sociedad de la Información (1999).*

Para que ciudadanos y empresas puedan obtener el máximo potencial de la información del sector público, la Comisión Europea publicó en 1999 un libro sobre las formas óptimas de utilizar esas fuentes de información “abiertas”. No solo pone de manifiesto el potencial de los datos, si no que trata de concienciar al ciudadano acerca de sus derechos de acceso a la información. Paralelamente, trata de fomentar la responsabilidad social de las empresas, de aprovechar las experiencias existentes en ese ámbito e incrementar las prácticas innovadoras y de transparencia. Fue redactado a propuesta del Sr. Martin Bangemann, de la Comisión responsable de cuestiones relativas a la sociedad de la información, en cooperación con los

---

<sup>18</sup> Sección de normativa del portal de datos abiertos de la Junta de Castilla y León:

<https://datosabiertos.jcyl.es/web/es/iniciativa-datos-abiertos/normativa.html>

<sup>19</sup> Sección de normativa del portal de datos abiertos de la Junta de Andalucía:

<https://www.juntadeandalucia.es/datosabiertos/portal/ayuda/que-son/tipos.html>

<sup>20</sup> Sección de normativa del portal de datos abiertos del Ayuntamiento de Madrid:

<https://datos.madrid.es/portal/site/egob/menuitem.400a817358ce98c34e937436a8a409a0/?vgnextoid=830512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextchannel=830512b9ace9f310VgnVCM100000171f5a0aRCRD&vgnnextfmt=default>

<sup>21</sup> Sección normativa del portal de datos abiertos del Ayuntamiento de Valencia:

<http://gobiernoabierto.valencia.es/es/legislacion/>

<sup>22</sup> Marco legal de la Comunidad Valenciana, ya que es la región prioritariamente estudiada en esta Tesis Doctoral.

estados miembros y representantes de los ciudadanos, sector privado y proveedores de información (Comisión Europea, 1999).

El libro no propone la publicación de más información pública, sino que la que esté disponible se ponga a disposición del ciudadano de una forma más accesible y transparente. En determinadas ocasiones habrá que aplicar mejoras técnicas, simplificar información o rediseñar los procesos administrativos.

Desde la propia Comisión se aboga por que los estados miembros puedan contribuir con soluciones del tipo (Comisión Europea, 1999) siguiente:

- Debates para establecer medidas legislativas, recomendaciones y normas.
- Intercambios de información a nivel comunitario.
- Medidas de información sobre las fuentes de datos.
- Proyectos de reutilización de datos e información que sirvan como ejemplo.
- Proyectos de formación.

*Directiva Europea 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información del sector público.*

Esta Directiva, aprobada por el Consejo Europeo en 2003, fue un texto pionero en las iniciativas de datos abiertos dentro de la Unión Europea. Con el objetivo de conseguir gobiernos transparentes, se establece un conjunto mínimo de reglas que regulan la reutilización de datos e información custodiados por los organismos públicos de los estados miembros de la Unión Europea. Define qué es la reutilización y con qué fines puede utilizarse la información (comerciales o no).

La Directiva no afecta a nivel de protección de las personas en cuanto a tratamiento de datos personales.

*Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones sobre la Reutilización de la Información del Sector Público, revisión de la Directiva 2003/98/CE (sin carácter legislativo) 2009.<sup>23</sup>*

Esta comunicación es un documento que revisa la aplicación de la Directiva 2003/98/CE resaltando el potencial de reutilización de los datos en la era digital. La información del sector

---

<sup>23</sup> Comisión de las Comunidades Europeas (2009). *Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones: Reutilización de la información del Sector Público. Revisión de la Directiva 2003/98/CE.* Disponible en: [http://dgtic.caib.es/www/caibdata/normativa/Comunicacion\\_Directiva\\_2003\\_98\\_CE.pdf](http://dgtic.caib.es/www/caibdata/normativa/Comunicacion_Directiva_2003_98_CE.pdf)

público es una fuente de información primaria puesta a disposición de los ciudadanos por parte de organismos públicos. Incluye fuentes de datos de todo tipo: mapas, imágenes, legislación, estadísticas, registros de sociedades, etc. Estas fuentes de datos se usan como materia prima en numerosos productos y servicios desarrollados por reutilizadores de la información. Estos pueden jugar un papel fundamental en la actividad económica a nivel comunitario y el aumento del uso de estos recursos hará que crezca todavía más, cumpliendo paralelamente con objetivos de competitividad y de creación de puestos de trabajo establecidos desde la Unión Europea.

La Directiva elimina los principales obstáculos que existen para la reutilización de la información como son las prácticas discriminatorias, monopolistas y la falta de transparencia. La reutilización de la información del sector público cada vez es mayor. Una de las debilidades se debe a que no se puede aprovechar gran parte del potencial debido a que los organismos del sector público no gestionan de un modo eficiente las fuentes de información. Uno de los motivos que señala el propio documento se debe a las metas cortoplacistas de recuperación de los costes de inversión. Existen presiones desde las propias entidades públicas para que se financien este tipo de iniciativas, teniendo en cuenta que puede dar lugar a competencia con el sector privado, en temas de concesión de licencias y establecimiento de tarifas.

Existen problemas de ejecución de la normativa de Reutilización de la Información del Sector Público (en adelante RISP) en la Unión Europea. Entre ellas hay que señalar las de: falta de información sobre datos disponibles, que a la propia administración no le entusiasme la idea de una reutilización de su información o que no sean conscientes del potencial económico, entre otros. Frente a lo que ocurre en Europa, en Estados Unidos, la reutilización de la información se potencia constantemente: tanto empresas como ciudadanos (potenciales reutilizadores) disponen de derechos de acceso a la información además de una amplia posibilidad de reutilización comercial. La información no está protegida por derechos de autor y la reutilización no está sujeta a restricciones. Las tarifas se limitan a costes marginales de reproducción y difusión.

En cuanto a repercusiones y ámbito de aplicación de la Directiva 2003/98/CE de la Información del Sector Público, según se señala en la Comunicación, se detectaron cambios producidos por la Directiva debido a la aplicación de normas y prácticas diversas de los estados miembros. Por ejemplo, en países como Bélgica la nueva legislación favorece la reutilización comercial de la información del sector público. En Países Bajos y Suecia se han llevado a cabo estudios para identificar acuerdos entre gobiernos centrales y autoridades locales. En Letonia, se instauró un procedimiento que garantizaba que cualquier derecho exclusivo se ajuste a la Directiva y que sea objeto de un dictamen de las autoridades competentes. En tema de tarifas para el uso de los datos, los Países Bajos anunciaron la intención de cobrar como máximo, los costes marginales considerados por la Directiva o en

España, se redujeron drásticamente los precios para el acceso al catastro español. Este tipo de políticas favorecen la reutilización y por tanto, que aumente notablemente. Finalmente, en Francia o Eslovenia se han creado mecanismos de resolución de conflictos en materia de reutilización ofreciendo soluciones de litigios y mediación ágiles y económicos.

Se llevó a cabo un estudio en tres sectores de reutilización de información del sector público: la información de carácter geográfico, meteorológico y jurídico-administrativo. Se establecieron una serie de indicadores para medir el nivel de reutilización que sirvieron para corroborar el crecimiento de mercado y aumento de la reutilización en estos sectores.

Algunos reutilizadores (en materia geográfica o meteorológica) pusieron de manifiesto las barreras encontradas: altas tasas para acceder a los datos, hecho que propició que tuvieran que acudir a fuentes del sector privado o abandonar el desarrollo de algunos de los servicios.

En la presente revisión se establece que la Directiva de la RISP tiene efectos positivos en cuanto a reutilización de información, pero no se ha explotado plenamente. Son los propios reutilizadores quienes sugieren una modificación de la Directiva para añadir disposiciones más estrictas. Los estados miembros consideran que es temprano para modificar la directiva y abogan por la concesión de un período de tiempo para poder analizar su aplicación.

En definitiva, este documento llega a las siguientes conclusiones:

- Las propias entidades públicas siguen anteponiendo la recuperación de costes de inversión en contra del crecimiento económico de la sociedad.
- Existe una competencia real entre el sector público y privado; hecho que entorpece la reutilización (falta de información disponible).
- Los propios organismos públicos no son conscientes del potencial económico de la RISP.
- Son los propios estados miembros quienes deben centrar esfuerzos en que se cumpla la Directiva.

*Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, modifica la Directiva 2003/98/CE relativa a la RISP.*

Esta Directiva añade una serie de modificaciones a la anterior introduciendo una serie de cambios relevantes:

- La Directiva no se aplica a documentos no relacionados con la actividad administrativa del Estado.
- Amplia el mismo a archivos, bibliotecas y museos.



- Cambia los períodos máximos de derechos exclusivos relacionados con la digitalización de recursos culturales (no superiores a 10 años).
- Obligación de publicar la información en formatos interoperables para garantizar así la reutilización de los documentos/datos: formatos legibles por máquina y metadatados.

*Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo de 20 de junio de 2019 relativa a los datos abiertos y la reutilización del sector público.*

Aprobada en 2019, tiene como objeto fomentar el uso de datos abiertos y desarrollo de productos y servicios por el ciudadano. Para ello se establecen un conjunto de normas mínimas que regulan la reutilización de documentos de la administración pública. De este modo, además, se actualiza el marco legislativo con los avances en las tecnologías de la información y la comunicación.

La directiva entra en vigor el 16 de julio de 2019, 20 días después de su publicación en el Diario Oficial de la Unión Europea (DOUE).

Según el portal de Noticias Jurídicas<sup>24</sup>, la norma parte de que la información del sector público es una fuente de datos que contribuye al crecimiento de mercado, gracias al desarrollo de nuevos servicios y aplicaciones que enriquece a los reutilizadores y que son de utilidad para el ciudadano. Por tanto, se considera un efecto transformador en distintos sectores de la economía y la sociedad (Noticias Jurídicas, Directiva 2019/1024).

Como todos los datos abiertos, serán los propios organismos del sector público quienes facilitarán la información en cualquier tipo de formato, legibles por máquinas, accesibles y localizables para su posterior reutilización. En cuanto a formatos técnicos deben ser abiertos y no propietarios. La reutilización es gratuita, salvo aquellos organismos del sector público a los que se les exija generar ingresos para cubrir costes de sus misiones de servicio público (estos ingresos no superarán el coste de recogida, producción, reproducción, difusión y almacenamiento de datos, y dichos criterios serán totalmente transparentes).

En cuanto a los datos de investigación, serán los estados miembros quienes apoyen la disponibilidad de estos mediante políticas nacionales y otras actuaciones para que sean totalmente accesibles los datos de investigación financiados públicamente. Además, la gestión de datos de investigación ha ocupado en los últimos años un lugar destacado dentro de la agenda de la comunidad científica y la Comisión Europea. Se están llevando a cabo

---

<sup>24</sup> Fuente: Noticias Jurídicas. Directiva europea sobre datos abiertos y reutilización de la información del sector público: <https://noticias.juridicas.com/actualidad/noticias/14173-nueva-directiva-europea-sobre-datos-abiertos-y-reutilizacion-de-la-informacion-del-sector-publico/>

diferentes actividades y proyectos para la publicación y gestión de datos y que cumplan los requisitos FAIR (localizables, accesibles, interoperables y reutilizables) (Alcalà y Anglada, 2019).

Impone una serie de condiciones de no discriminación (como la reutilización transfronteriza): cualquier persona puede hacer uso de los datos libremente. El texto habla de posibles supuestos en los que sea necesario el derecho exclusivo para la prestación de un servicio de interés público, pero estará sometido a revisión periódica.

El texto incorpora un conjunto de normas mínimas que regulan la reutilización de:

- Documentos existentes conservados por organismos públicos de los estados miembros.
- Documentos conservados en empresas públicas que llevan a cabo su actividad en ámbitos definidos en la Directiva 2014/25/UE (agua, transporte, energía...).
- Datos de investigación financiados con fondos públicos.

No se aplicará dicha normativa a:

- Documentos que salgan de la misión del ámbito público siempre y cuando el ámbito del servicio público sea transparente y se someta a revisión.
- Documentos con derechos de propiedad intelectual de terceros.
- Documentos cuyo acceso esté limitado por motivos de protección de información.

Se tienen en cuenta los denominados datos de alto valor. La directiva contiene un anexo (Anexo I) con este tipo de datos:

- Geoespaciales.
- Observación de la tierra y medioambiente.
- Meteorología.
- Estadística.
- Sociedades.
- Propiedad de sociedades.
- Movilidad.

A medida que surjan nuevas categorías temáticas con datos de alto valor, se irá actualizando el Anexo I. Se basan en el valor potencial que tienen para generar beneficios a nivel medioambientales, económicos y de innovación. Deberán estar a disposición de cualquier ciudadano de manera gratuita, sin restricciones, serán legibles por máquina, suministrados a través de API's y podrán ser descargados masivamente cuando sea necesario.

*Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones: Una estrategia europea de datos.*

La estrategia pretende colocar a Europa en una posición líder en todo lo relacionado con la economía de los datos y la transición digital. Como ya se ha señalado, el volumen de datos en el mundo se duplica cada dieciocho meses, con la previsión de que en el 2025 alcanzará los 175 Zetabytes<sup>25</sup>.

La Comisión trabajará para facilitar un mayor volumen de datos del sector público para su reutilización, teniendo en cuenta su potencial para las pymes. Con el fin de abrir conjuntos de datos clave del sector público a la innovación, iniciará el procedimiento para la adopción de un acto de ejecución sobre conjuntos de datos de gran valor en el marco de la Directiva sobre datos abiertos, de manera que estos conjuntos de datos estén disponibles en toda la Unión Europea, sin restricciones, en formatos técnicos reutilizables y a través de interfaces de programador de aplicaciones (API's). Además, ayudará a los estados miembros a garantizar una transposición oportuna y precisa de las nuevas normas de la Directiva sobre datos abiertos (fecha máxima, 17 de julio de 2021) (Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones, 2020).

De esta manera la Unión Europea trabaja, al diseñar esta estrategia, para conseguir una economía de los datos mediante una gobernanza de los datos a nivel europeo.

### 2.9.3 Legislación en ámbito estatal: Gobierno abierto, transparencia y reutilización de información del sector público en el estado español

*Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.*

En España, la primera vez que se incorpora a nuestro régimen jurídico la Directiva 2003/98/CE es mediante la Ley 37/2007 aprobada el 16 de noviembre del 2007. En ella se establecen una serie de requisitos mínimos para la reutilización de documentos custodiados o elaborados por organismos y administración pública. Tanto la Directiva como la aprobación de esta ley persiguen armonizar la explotación de la información del sector público, con objeto de crear servicios y productos de información basados en documentos y publicitar los documentos de libre disposición de las entidades públicas (Noticias jurídicas – Ley 37/2007, 2008 – Revisión 2018).

---

<sup>25</sup> Noticia extraída de: <https://www.economista.es/economia/noticias/11371458/08/21/La-Union-Europea-lanza-su-Estrategia-de-Datos-e-impulsa-una-nueva-normativa.html>

Aspectos que destacar de esta ley:

- Posibilita a las Administraciones y organismos del Sector Público a permitir la reutilización de documentos sin condiciones concretas, o con condiciones establecidas bajo “licencias tipo”.
- La información es susceptible de ser reutilizada. Pueden crearse listados e índices accesibles a documentos.
- La información debe publicarse en formatos electrónicos reutilizables.
- No se puede alterar el contenido, se debe citar la fuente de información y fecha de la última actualización.

Finalmente, establece un régimen sancionador por parte de la Administración General del Estado en el caso de que se realice un mal uso de la información.

*Ley 56/2007, de 28 de diciembre, de Medidas de Impulso de la Sociedad de la Información.*

Esta ley forma parte del conjunto de medidas del Plan Avanza (2006-2010)<sup>26</sup> aprobado por el Gobierno Central en noviembre de 2005 para el desarrollo de la Sociedad de la Información con Europa. Se trata de una iniciativa que tenía como objetivo eliminar barreras existentes respecto a la expansión y uso de tecnologías de la información con tal de garantizar los derechos de los ciudadanos en la nueva sociedad de la información.

Introduce aspectos dirigidos a impulsar el empleo de factura electrónica, uso de medios electrónicos en los procesos de contratación y a garantizar la interlocución telemática con el ciudadano. Además, establece para aquellas empresas con incidencia en la actividad económica que faciliten medios de comunicación electrónica con los usuarios de sus servicios y que dispongan de certificados reconocidos de firma electrónica. Obliga al ciudadano a tener un canal de comunicación electrónico para la interlocución con empresas y administraciones (*Noticias Jurídicas*, Ley 56/2007, 2007).

Contiene una serie de modificaciones legislativas para la promoción e impulso de la sociedad de la información y de las comunicaciones electrónicas. Revisa, añade y eliminan obligaciones anteriores que pueden ser innecesarias, excesivas o escasas, y hace flexibles las referidas a comunicaciones comerciales y contrataciones electrónicas (*Noticias Jurídicas*, Ley 56/2007, 2007).

---

<sup>26</sup> Plan 2006-2010 para el desarrollo de la Sociedad de la Información y de Convergencia con Europa y entre Comunidades Autónomas y Ciudades Autónomas. Disponible: [https://avancedigital.mineco.gob.es/programas-avance-digital/DescargasPlanesAvanza/Plan%20Avanza/plan\\_avanza-Documento\\_completo.pdf](https://avancedigital.mineco.gob.es/programas-avance-digital/DescargasPlanesAvanza/Plan%20Avanza/plan_avanza-Documento_completo.pdf)

*Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica.*

El concepto de “interoperabilidad” se define en la Ley 11/2007 (ya derogada, en detrimento de la ley 39/2015, ver nota al pie) como “*la capacidad de los sistemas de información, y por ende de los procedimientos a los que estos dan soporte, de compartir datos y posibilitar el intercambio de información y de conocimiento entre ellos*”.

Resulta necesaria e imprescindible para el desarrollo de la administración electrónica y la sociedad de la información (*Noticias Jurídicas, Real Decreto 4/2010– Revisión 2011, 2010*):

- Cooperación, integración y prestación de servicios de las administraciones públicas.
- Ejecución de políticas públicas.
- Realización de principios y derechos.
- Transferencia de tecnología.
- Reutilización de aplicaciones e información en beneficio de una mejor eficiencia común.

Es un desafío planteado desde los organismos públicos, y nace debido a la aparición de las tecnologías de la información y la comunicación, internet, y la nueva forma de interacción entre administración y ciudadano: la administración electrónica. Las entidades públicas entablan relaciones entre sí, lo que debe permitir evitar que el ciudadano padezca las consecuencias burocráticas de estas relaciones (Gamero Casado, 2009).

Esta consagración del derecho de los ciudadanos a comunicarse con las administraciones a través de canales tecnológicos comporta una obligación por parte de estas a que las condiciones sean iguales y efectivas. Deben ser las mismas quienes establezcan los mecanismos para llegar a una neutralidad tecnológica y adaptabilidad.

La finalidad del Esquema Nacional de Interoperabilidad es crear la infraestructura y condiciones necesarias que garanticen el nivel de operación técnica, semántica y organizativa de sistemas y aplicaciones empleados por la administración pública y que permita el ejercicio de derechos y cumplimiento de deberes a través del acceso electrónico a los servicios públicos. (*Noticias Jurídicas, Real Decreto 4/2010– Revisión 2011, 2010*).

Desde la Unión Europea son conscientes de que la situación tecnológica no es igual en todos los países que la conforman. Se realizan una serie de recomendaciones de servicios, utilización de estándares abiertos y uso generalizado de estándares por los ciudadanos. Entre

ellas se atiende al Marco Europeo de Interoperabilidad<sup>27</sup> así como otros instrumentos y actuaciones elaborados por este programa y que inciden en aspectos relacionados con la interoperabilidad como solución para administraciones públicas europeas, dentro de los planes de actuación en materia de administración electrónica y con la política comunitaria de gobierno abierto, colaboración y apertura de información. (*Noticias Jurídicas, Real Decreto 4/2010– Revisión 2011, 2010*).

A nivel nacional, con la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas, se atiende al acceso electrónico de los ciudadanos a los servicios públicos, la protección de datos de carácter personal, firma electrónica, accesibilidad, DNI electrónico, uso de lenguas oficiales y reutilización de la información del sector público.

Este Real Decreto establece criterios, recomendaciones y principios necesarios para favorecer el desarrollo de la interoperabilidad en las administraciones públicas de acuerdo en lo establecido con la Ley 11/2007 derogada por la Ley 39/2015. Hace referencia a la interoperabilidad como “*proceso integral en el que no caben actuaciones puntuales o tratamientos coyunturales debido a que la debilidad de un sistema determina su punto más frágil y en ocasiones es el punto de coordinación entre medidas individualmente adecuadas, pero deficientemente ensambladas*” (Real Decreto 4/2010).

La norma se estructura en 12 capítulos, 4 disposiciones adicionales, 2 transitorias, 1 derogatoria, 3 disposiciones finales y un anexo con el glosario de términos. Remite al Esquema Nacional de Seguridad<sup>28</sup> en cuestiones relativas a la protección en aspectos necesarios para garantizar la interoperabilidad.

*Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal.*

La Ley 37/2007 sobre reutilización de la información del sector público aporta un valor añadido de derecho de acceso a la información, promoviendo la reutilización y garantizando que se lleve a cabo en condiciones transparentes y no discriminatorias (Noticias jurídicas, Real Decreto 1495/2011, 2011).

---

<sup>27</sup> Comisión de las Comunidades Europeas (2017). Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Marco Europeo de Interoperabilidad – Estrategia de aplicación.

<sup>28</sup> Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica.

En la Declaración Ministerial de Malmö (noviembre, 2009) y dentro del *Plan de Acción de la Unión Europea sobre Administración Electrónica*, se establecen una serie de prioridades en materia de reutilización de datos a llevar a cabo entre 2010 y 2015.

El Real Decreto enmarca el conjunto de medidas de la Estrategia 2011-2015 del Plan Avanza 2. Prevé en su normativa el desarrollo de la Ley 37/2007 con objetivo de detallar para el ámbito público las disposiciones de la ley, facilitando al máximo la puesta a punto de la información (Noticias jurídicas, Real Decreto 1495/2011, 2011).

Este Real Decreto queda resumido de la siguiente forma:

- Acota el ámbito de aplicación subjetiva de la ley al sector público estatal.
- Contiene el régimen jurídico de la reutilización de la información del sector público.
- Determina responsabilidades y funciones en materia de reutilización en los departamentos de los distintos ministerios y entidades públicas.
- Establece que los organismos públicos informarán a través de la sede electrónica sobre documentos reutilizables.
- Prevé un mantenimiento del catálogo de información pública reutilizable.
- Establece mecanismos de coordinación pertinentes al sector público.
- Establece las condiciones generales exigibles para la reutilización de la información.
- Prevé que la puesta a disposición de la información para su reutilización conlleva la cesión no exclusiva de derechos de propiedad intelectual.
- Los documentos con información de carácter personal deben anonimizarse para su reutilización.

*Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno.*

Con esta ley aprobada en 2013 se pone fin a la anomalía de la falta de norma en materia de transparencia pública. No cubre el vacío normativo absoluto, ya que ya existían en el ordenamiento jurídico normas que obligaban a la publicidad activa por parte de las administraciones y que permitían el derecho de acceso a archivos y registros públicos. Se pretende ampliar y reforzar la transparencia de la actividad pública, reconocer y garantizar el acceso a la información relativa las actividades públicas y se establecen una serie de obligaciones de buen gobierno que deben cumplir los responsables públicos, así como las consecuencias de no cumplir con estas premisas previas (Velasco Rico, 2014).

La ley de transparencia tiene un triple objetivo (Noticias Jurídicas, Ley 19/2013, 2013):

- Incrementar y reforzar la transparencia de la actividad pública mediante obligaciones de publicidad activa.

Se debe reforzar las obligaciones de publicidad activa en ámbitos como información institucional, organizativa y de planificación, poniendo a disposición del ciudadano información relativa a funciones que desarrollan, normativa resultante de la aplicación y estructura organizativa.

En materia de información, si afecta a la relación entre Administración y ciudadano se deben publicar para una mayor seguridad jurídica.

En cuanto a la información de relevancia económica y presupuestaria, debido a la importancia de control de gestión y uso de recursos públicos, debe publicarse de manera accesible y entendible para los ciudadanos.

- Reconocer y garantizar el derecho de acceso a la información.

Los ciudadanos deben poder acceder a registros y documentos situados en archivos administrativos. Se consigue suplir sus carencias anteriores y con esta ley se cumple con el derecho de acceso a la información de los ciudadanos.

- Establecer obligaciones de buen gobierno a cumplir por parte de los responsables públicos y se establecen las responsabilidades jurídicas derivadas del incumplimiento.

*Real Decreto 919/2014, de 31 de octubre, por el que se aprueba el Estatuto del Consejo de Transparencia y Buen Gobierno.*

Dentro del título III de la Ley 19/2013 de transparencia, acceso a la información pública y buen gobierno, se crea el Consejo de Transparencia y Buen Gobierno. Se trata de un órgano que asume las competencias de velar por el cumplimiento de las obligaciones de buen gobierno y publicidad activa además de garantizar el derecho de acceso a la información pública. Por tanto, se configura como un órgano que conocerá reclamaciones e inquietudes en materia de acceso a la información que presenten los ciudadanos. De acuerdo con lo previsto en la Ley 19/2013, actuará con plena independencia y autonomía con tal de cumplir sus fines. (Noticias Jurídicas, Real Decreto 919/2014).

El presente Real Decreto 919/2014, recoge las disposiciones siguientes:

- Naturaleza y régimen jurídico.
- Objetivos y funciones.
- Adscripción orgánica.



- Sede.
- Cooperación institucional.
- Principios que rigen el ejercicio de la actividad.
- Funciones de acuerdo con la transparencia y buen gobierno.

Además, en el capítulo II regula la estructura orgánica del Consejo y desarrolla qué funciones deben desempeñar cada uno de los órganos que lo componen. El capítulo III regula el tema económico, patrimonial y del personal, detallando recursos económicos, elaboración y ejecución del presupuesto. Finalmente, el capítulo IV especifica la competencia de la Abogacía General del Estado para la prestación de asistencia jurídica que el Consejo de Transparencia y Buen Gobierno pudiera requerir (Noticias Jurídicas, Real Decreto 919/2014).

El Consejo de Transparencia y Buen Gobierno<sup>29</sup> (España) publica en 2015 el “Código ético de Transparencia y Buen Gobierno”.

*Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.*

La Ley 37/2007 recalca la importancia de la información que generan las entidades del sector público para promover la economía del conocimiento y la necesidad de la puesta en valor para su reutilización. Con la Directiva 2003/98/CE y la aprobación de la Ley 37/2007, de 16 de noviembre se establece la regulación básica del régimen jurídico aplicable a la reutilización de documentos elaborados o custodiados por las administraciones públicas.

La modificación de la ley 37/2007 viene dada por varios motivos: el volumen de información generado por parte de las entidades públicas que crece de manera exponencial y el progreso de las tecnologías para explotar datos, realiza análisis y con ello, desarrollo de nuevos servicios y productos y la conciencia del valor de la información pública (Noticias Jurídicas, Ley 18/2015, 2015).

Con estas variaciones, se pretende incorporar al ordenamiento jurídico español los cambios introducidos por la Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, con el objetivo de reforzar y potenciar el uso transfronterizo de los datos y con ello, las distintas normativas adoptadas por los estados miembros en materia de reutilización de datos del sector público. Las modificaciones se dirigen al ámbito de

---

<sup>29</sup> Consejo de Transparencia y Buen Gobierno (2015). Código ético. Consejo de Transparencia y Buen Gobierno. [https://www.consejodetransparencia.es/dam/jcr:7ac8878b-1184-40e1-b226-da3fda3566c1/Codigo\\_etico.pdf](https://www.consejodetransparencia.es/dam/jcr:7ac8878b-1184-40e1-b226-da3fda3566c1/Codigo_etico.pdf)

aplicación, al régimen de reutilización, a la forma y las tarifas del suministro de documentos, a la tramitación de solicitudes y las condiciones de uso (González Bustos, 2016).

Modificaciones que se incorporan en esta nueva ley (Noticias Jurídicas, Ley 18/2015, 2015):

- Recoge las disposiciones de la Directiva acerca de la obligación por parte de las Administraciones de autorizar la reutilización de documentos públicos (exceptuando aquellos cuyo acceso esté restringido).
- Se definen conceptos como “formato legible por máquina”, “formato abierto” o la norma formal abierta que garantiza interoperabilidad, entre otros.
- Incorpora en el cálculo del régimen de tarifas de la reutilización de documentos el principio de costes marginales establecido por la Directiva.
- Fomento del uso de licencias abiertas con las mínimas restricciones posibles.

*Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas.*<sup>30</sup>

Se incluye en este apartado la Ley de Procedimiento Administrativo que, aunque no tenga un apartado específico relacionado con la transparencia, buen gobierno y datos abiertos, si tiene en su desarrollo los apartados que tienen una relación directa con la administración electrónica. La administración electrónica es una de las bases para que el registro de la información y su puesta a disposición al ciudadano se pueda producir de manera más rápida, eficiente y trazable.

Esta ley no supone un replanteamiento sustancial del procedimiento administrativo común y del régimen jurídico de los actos administrativos. Añade pequeñas reformas, puntualizaciones, aclaraciones y desarrollos de la regulación reciente. La novedad incorporada viene con la integración y regulación general del procedimiento de la administración electrónica (Chinchilla Peinado, 2016).

La exposición de motivos de esta Ley 39/2015 arranca con un reconocimiento sobre la complejidad e ineficiencia de la Administración, señalando los defectos tradicionales con las normas que la regulan: *“Los defectos que se atribuyen a las Administraciones españolas vienen dados por distintos motivos pero el ordenamiento vigente no es ajeno a ellas ya que el marco normativo en el que se desenvuelve, propicia la aparición de duplicidades y otras ineficiencias, con procedimientos administrativos complejos que además generan problemas de inseguridad jurídica. Por esto se necesita una reforma integral y estructural que permita*

---

<sup>30</sup> La Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los Servicios Públicos, queda derogada a partir del 2 de octubre de 2016 por la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas.

*clarificar como se deben organizar y relacionar las Administraciones con los ciudadanos, las empresas y a nivel interno con el resto de Administraciones e instituciones estatales”* (Romero Gómez, 2015).

Para afrontar estas deficiencias se define el conjunto normativo sobre la Administración atendiendo a sus destinatarios. La reforma se articula en dos ejes:

- Relaciones *ad extra*: Administración con los ciudadanos.
- Relaciones *ad intra*: Administración con el resto de las administraciones e instituciones del Estado.

Estos ejes (*ad extra – ad intra*) se utiliza como criterio para la exposición y evaluación de lo que es una buena Administración permitiendo así establecer cuerpos normativos centrados en el ciudadano antes que la Administración (Romero Gómez, 2015).

Con esto se impulsan dos nuevas leyes en las que se asienta el Derecho Administrativo Español:

- Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas (definida en este apartado).
- Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público (no definida).

Estas leyes no incorporan grandes modificaciones y cambios al respecto de las que las precedían, pero tanto dentro de su aplicación como en el desarrollo normativo, hace que se tenga en cuenta gracias a la nueva forma de ver la centralidad de los derechos y libertades de la ciudadanía en la actuación *ad extra* de la Administración. (Romero Gómez, 2015).

Esta ley regula los derechos y garantías que corresponden a los ciudadanos en su actividad administrativa y en lo que se refiere al procedimiento administrativo, su carácter común resulta aplicable a todas sus administraciones públicas (García-Valdecasas Algüacil, 2016).

Los objetivos de esta ley son:

- Regular los requisitos de validez y eficacia de actos administrativos.
- Establecer los principios a los que se ha de ajustar el ejercicio de la iniciativa legislativa y potestad reglamentaria.
- Reforzar la participación ciudadana.
- Garantizar la seguridad jurídica.
- Acabar con la dispersión normativa.
- Revisar el ordenamiento.

Se establece así por primera vez en una ley las bases en las que se debe desarrollar la iniciativa legislativa y la potestad reglamentaria y entra en vigor el 2 de octubre de 2016. En cuanto a registro administrativo de apoderamientos, registro electrónico, empleados públicos habilitados, punto de acceso general electrónico y archivo electrónico único, entra en acción a los dos años de entrada en vigor de la ley, es decir, en 2018 (García-Valdecasas Algüacil, 2016).

La Ley 39/2015 está formada por 133 artículos, con 6 títulos, 5 disposiciones adicionales, 5 disposiciones transitorias, 1 derogatoria y 7 finales.

García-Valdecasas Algüacil resume las modificaciones y añadidos respecto a su predecesora en los siguientes puntos y que tienen un impacto directo en facilitar la información a los ciudadanos:

- Las Administraciones deben disponer de un registro electrónico general de apoderamientos, generales o particulares, pero interoperables entre sí.
- Los interesados en un procedimiento administrativo deben poder identificarse electrónicamente y poder firmar por medios electrónicos válidos como:
  - Firma electrónica.
  - Sello electrónico.
  - Otros que las Administraciones consideren como válidos.
- Derechos reconocidos a las personas para la nueva relación con las Administraciones Públicas:
  - Comunicación a través de puntos de Acceso General electrónico de la administración.
  - Asistencia en el uso de medios electrónicos.
  - Obtención y uso de medios electrónicos en su relación con la Administración.
  - Protección de datos de carácter personal, seguridad y confidencialidad de datos.
  - Son los propios ciudadanos quienes se comunican con la Administración a través de medios electrónicos o no.
- En materia de archivos de documentos, las Administraciones deberán mantener un archivo electrónico único para documentación y expedientes electrónicos de los expedientes finalizados. Se deberán conservar en formatos que garanticen autenticidad, integridad y conservación.
- El expediente administrativo tendrá formato electrónico previsto en el Esquema Nacional de Interoperabilidad.
- Las notificaciones se practicarán por medios electrónicos y a través de la sede electrónica. Se entenderán por rechazadas cuando transcurran 10 días naturales desde la puesta a disposición sin que el interesado acceda al mensaje.

- Las Administraciones deberán publicar un Plan Normativo de iniciativas legales y reglamentarias, y una vez aprobados, se publican en su Portal de Transparencia. Con esta medida se incrementa la seguridad jurídica.

*Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.*

Esta Ley Orgánica 3/2018 completa determinados aspectos del Reglamento (UE) 679/2016<sup>31</sup>. Se trata de un catálogo heterogéneo que van desde el derecho a la neutralidad de Internet al derecho de la desconexión digital de los empleados (González Tapia, 2018).

Hablamos de una regulación embrionaria acerca de las garantías digitales de los trabajadores a partir del mandato del Reglamento (UE) 679/2016, de adaptar la normativa de protección de datos de carácter personal al entorno laboral. Se regulan además algunos de los escenarios sociales y jurisprudencialmente típicos en los que dicho ejercicio conlleva tratamientos de datos personales de trabajadores, otorgándose estos derechos de privacidad con el uso de dispositivos digitales, sistemas de vigilancia y captación de sonidos además de la posición geográfica (Baz Rodríguez, 2018).

Esta Ley Orgánica consta de 97 artículos estructurados en 10 títulos, 22 disposiciones adicionales, 6 transitorias, 1 derogatoria y 6 disposiciones finales. Los aspectos más relevantes de esta ley son los siguientes (Noticias Jurídicas – Revisión vigente desde 7 de diciembre de 2018 hasta 18 de enero de 2021) y (Campos, 2018):

- Las Administraciones Públicas tienen legitimación para utilizar y ceder datos personales, siempre y cuando el tratamiento sea necesario para cumplir con una obligación legal y cuando sea de interés público.
- Uso de medios tecnológicos y datos personales en las actividades electorales.
- Figura del Delegado de Protección de datos en caso de reclamación ante las autoridades de protección de datos, mediante el reconocimiento de su papel como órgano intermediario en el control.
- Derecho al olvido en servicios de redes sociales o equivalentes: cualquier persona puede solicitar que se suprima su información que haya facilitado en redes sociales y servicios de sociedad de la información. Además, se incorpora el “derecho de portabilidad en servicios de redes sociales y servicios equivalentes” donde los usuarios de redes sociales tendrán derecho a recibir los contenidos que haya facilitado a dichos servicios.

---

<sup>31</sup> Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos).

- Nuevas obligaciones en materia de transparencia y publicidad activa.
- En cuanto a los datos personales de personas fallecidas contempla que las personas vinculadas al fallecido serán quienes se erijan al responsable del tratamiento para solicitar el acceso a los datos personales para su rectificación o supresión. Solo en caso de que lo prohibiera expresamente el fallecido, no podrían acceder a su información.
- Las plataformas de intermediación de datos de las Administraciones Públicas: se reconoce a los interesados que pueden no aportar documentos que ya se encuentren en manos de las AAPP. Las Administraciones podrán consultar dichos documentos salvo que el interesado sea quien se oponga a ello. Serán las propias Administraciones quienes deban recabar los documentos de manera electrónica a través de sus redes colaborativas o mediante las plataformas de intermediación de datos.
- Régimen jurídico de empleados públicos y privados: “Derechos digitales en la negociación colectiva”. Derechos y libertades relacionados con el tratamiento de datos personales de trabajadores y conservación de derechos digitales en el ámbito laboral. Además, se regula su derecho a la intimidad en relación con el entorno digital y la desconexión, y en el uso de dispositivos digitales puestos a su disposición y frente al uso de dispositivos de video vigilancia y geolocalización.

#### 2.9.4 Legislación y normativa en ámbito autonómico: desarrollo regulatorio de Gobierno Abierto en la Comunidad Valenciana

A pesar de que el estudio de esta tesis es a nivel estatal, solo se ejemplifica el caso de la Comunidad Valenciana, sede de la Universitat Politècnica de València, ya que cada Comunidad Autónoma tiene su propio marco legal en materia de gobierno abierto, transparencia y datos abiertos. A nivel autonómico se destacan tres normas que tienen aplicación en el ámbito de la Comunidad Valenciana. Se ha incluido asimismo el Reglamento de Transparencia y Buen Gobierno del Ayuntamiento de Valencia como un ejemplo de normas que a nivel interno que pueden establecerse en las entidades locales.

*Ley 2/2015, de 2 de abril, de la Generalitat, de Transparencia, Buen Gobierno y Participación Ciudadana de la Comunitat Valenciana.*

La Ley 2/2015 entra en vigor el 9 de abril de 2015. Deroga las Leyes 11/2007, 11/2008 y 11/2009, el Decreto 206/2007 y modifica la Ley 5/1993 de la Comunitat Valenciana.

Como ya hemos visto, las sociedades democráticas más avanzadas han reorientado su forma de hacer política focalizándose en materia de apertura de información para fomentar la participación y producción de nuevos productos y servicios reutilizando información de carácter público.

Esta Ley establece un modelo para regular la obligación de informar, publicitar la acción pública y el derecho de acceso a la información pública. La publicidad activa está al servicio de la transparencia dentro de la actividad pública. La Administración Pública ofrecerá al ciudadano a través internet y otros canales de comunicación esta información de la que hablamos.

El libre derecho de acceso a la información pública tiene una configuración diferente porque se trata de acceder no a la información que la administración hace pública, sino a cualquier otra, siempre que tal acceso no esté limitado. No siempre se podrá publicar cualquier tipo de información, y en otras ocasiones solo se podrá acceder parcialmente a determinados tipos de datos.

Esta Ley se estructura en un título preliminar y cinco cardinales, complementándose en un conjunto de disposiciones que articularán debidamente las regulaciones complementarias no sustantivas de la Ley.

Se define qué tipo de información debe hacerse pública, dentro del apartado de publicidad activa:

- Información económica, presupuestaria y estadística.
- Información de relevancia jurídica.
- Información organizativa y de planificación.
- Información relativa a los altos cargos y puestos asimilados.
- Información relativa a territorio, urbanismo y medio ambiente.

Se definen aspectos relativos al Portal de Transparencia de la Generalitat, como respuesta a facilitar el acceso de toda la ciudadanía a la información a través de un nodo de difusión de fácil acceso y con contenidos claros y estructurados (Ley 2/2015).

*Ley 2/2016, de 4 de marzo, de la Generalitat, de modificación de la Ley 2/2015, de 2 de abril, de Transparencia, Buen Gobierno y Participación Ciudadana de la Comunitat Valenciana [2016/1599].*

La Ley debe ser una herramienta efectiva de seguimiento y publicidad de las acciones y funcionamiento del poder ejecutivo. Es una pieza fundamental del ejercicio de los derechos democráticos de la ciudadanía.

Esta una Ley con un artículo que modifica el punto 1 del artículo 41 de la Ley 2/2015 queda redactado de la siguiente manera:

*“La comisión ejecutiva estará constituida por un número de miembros igual al número de grupos parlamentarios con representación en Les Corts. En la composición final se respetará la paridad entre mujeres y hombres. De entre sus miembros se designará la persona que ocupe la presidencia.*

*Los integrantes de la comisión, después de su comparecencia en la comisión correspondiente, serán elegidos por el Pleno de Les Corts por mayoría de tres quintos de entre expertos de competencia o prestigio reconocido y con más de diez años de experiencia profesional.*

*El mandato de la comisión tendrá una duración de cinco años”.*

Cómo ya se ha señalado, como ejemplo de reglamento de una entidad local puede servir el aprobado por el Ayuntamiento de Valencia sobre transparencia y participación ciudadana y en el que se configura el servicio de transparencia, sus funciones y su responsabilidad en la puesta a disposición de los datos abiertos.

*Decreto 56/2016, del Consell, de 6 de mayo, por el que se aprueba el Código de Buen Gobierno de la Generalitat. (DOGV núm. 7781, de 13.05.2016).*

Este Decreto Deroga el DECRETO 247/1995, de 24 de julio, del Gobierno Valenciano, por el que se crean los Registros de Actividades y de Bienes y Derechos Patrimoniales de Altos Cargos de la Generalitat Valenciana. (DOGV núm. 2559 de 26.07.1995) y complementa la Ley 2/2015, de 2 de abril, de la Generalitat, de Transparencia, Buen Gobierno y Participación Ciudadana de la Comunitat Valenciana. [2015/3137]

El Decreto 56/2016 establece una serie de principios de actuación que deben regir la labor de las personas que ocupan los altos cargos de la Generalitat y del personal asimilado. En el artículo 27 de esta ley prevé la aprobación por el Consell de un Código de Buen Gobierno basado en dichos principios, y el presente decreto da cumplimiento a este mandato (Decreto 56/2016).

Las nociones de ética pública y buen gobierno están presentes desde el pensamiento político clásico y se relacionan con la calidad institucional y la rendición de cuentas. Se deben establecer mecanismos eficaces de control del poder y de ética pública que se evidencien como elementos clave para la prevención de la corrupción y la generación de vínculos de confianza entre la ciudadanía y las instituciones (Decreto 56/2016).

Las normas éticas y de buen gobierno tienen una larga tradición en el mundo anglosajón, en los últimos años la conveniencia de este tipo de código ha sido asumida en diferentes países con alta calidad democrática y por organizaciones internacionales como las Naciones Unidas, la Organización para la Cooperación y el Desarrollo Económico, el Consejo de Europa o la Unión Europea. Estos códigos se conciben como instrumentos más orientados a la prevención y persuasión que a la disuasión de sus destinatarios y destinatarias y, siguiendo esta tendencia, en nuestro entorno ha sido en los últimos años cuando ha proliferado la elaboración de este tipo de normas, tendencia que se ha expresado también en otras leyes en materia de transparencia, de buen gobierno y del ejercicio del alto cargo (Decreto 56/2016).

El Código de Buen Gobierno parte del derecho a una buena administración reconocido en el artículo 41 de la Carta de los Derechos Fundamentales de la Unión Europea y en el artículo



9 del Estatuto de Autonomía de la Comunidad Valenciana, y se basa en los principios incluidos tanto en la Ley 19/2013, de 9 de diciembre, de Transparencia, Acceso a la Información Pública y Buen Gobierno, como, principalmente, en la Ley 2/2015, de 2 de abril. El texto percibe también influencias derivadas de la normativa en materia de transparencia, de buen gobierno, de incompatibilidades y del régimen jurídico aplicable a las personas que ocupan los altos cargos. Hay que señalar la normativa en materia de función pública, teniendo en cuenta que la inclusión de principios éticos y de conducta dirigidos a los empleados públicos, hecho consolidado tanto en la normativa estatal como en la autonómica y debe servir de referencia, también, para todo el sector público (Decreto 56/2016).

*Decreto 105/2017, de 28 de julio, del Consell, de desarrollo de la Ley 2/2015, de 2 de abril, de la Generalitat, en materia de transparencia y de regulación del Consejo de Transparencia, Acceso a la Información Pública y Buen Gobierno (DOGV núm. 8118, de 01.09.2017)*

Este Decreto tiene por objeto regular y garantizar el ejercicio del principio de transparencia y el derecho de libre acceso a la información pública en la Comunidad Valenciana. Supone un avance respecto a la Ley 19/2013, de 9 de diciembre, de Transparencia, Acceso a la Información Pública y Buen Gobierno (en adelante, Ley 19/2013, de 9 de diciembre) aprobada por el Estado. Se establece por primera vez en la normativa autonómica amplias obligaciones de transparencia y garantiza el ejercicio del acceso a la información pública. Incluye disposiciones relativas al buen gobierno y a la regulación de la participación ciudadana (Decret 105/2017).

Regula aspectos que la ley establece, como las obligaciones de publicidad activa, el derecho de acceso a la información, el régimen de reclamaciones y el régimen sancionador en materia de transparencia. Junto a estos aspectos, incorpora ciertas novedades en las que la normativa recogida en la ley era susceptible de un mayor desarrollo. Incluye una regulación específica de las condiciones de reutilización de la información pública, la organización administrativa de la Administración de la Generalitat y de su sector público instrumental en materia de transparencia y, en especial, la regulación del Consejo de Transparencia, Acceso a la Información Pública y Buen Gobierno, órgano creado por la Ley 2/2015, de 2 de abril, y encargado de garantizar el derecho de acceso a la información, de velar por el cumplimiento de las obligaciones de publicidad y garantizar y promover la observancia de las disposiciones de buen gobierno (Decreto 105/2017).

#### 2.9.5 Reglamentos y otros documentos normativos en ámbito local

Como ejemplos de aplicación en el ámbito local, el Ayuntamiento de Valencia tiene 2 normas publicadas: el Reglamento de Transparencia y Participación Ciudadana y el Reglamento de Gobierno Abierto: Transparencia.

*Reglamento de Transparencia y Participación Ciudadana*<sup>32</sup> (BOP de: 01-12-15).

El Ayuntamiento de Valencia aprobó el 28/09/2012 por acuerdo plenario el Reglamento de Participación Ciudadana. Para adaptarse a la Ley 19/2013, en 2015 el ayuntamiento aprueba la modificación del Reglamento de Participación Ciudadana. El nuevo Reglamento de Transparencia y Participación Ciudadana fue finalmente publicado en el Boletín Oficial de la Provincia de Valencia el 1 de diciembre de 2015.

En dicho reglamento hace referencia a los principios generales que incluye el objeto y régimen jurídico de la transparencia y acceso a la información pública:

- Principios de publicidad y publicidad activa
- Reutilización de la información, acceso y calidad de la información
- Compromiso de servicios
- Obligaciones de transparencia y acceso a la información pública y derechos de las personas en relación con estas cuestiones
- Medios de acceso a la información, así como la unidad o unidades responsables

Se concreta además la información pública en el ámbito municipal, sus límites y la protección de datos.

En cuanto a la “Publicidad activa”, se refiere a:

- Objeto y finalidad de la publicidad activa
- Lugar, la forma y los plazos de publicación y actualización
- Información institucional
- Titulares de órganos municipales y asimilados
- Información de relevancia jurídica y patrimonial, contratación, convenios y subvenciones, económica, financiera y presupuestaria, servicios y procedimientos e información ambiental y urbanística.

Se dedica un capítulo del reglamento a la reutilización de datos, en el sentido que le da la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, y comprende las cuestiones relativas a los objetivos de la reutilización, a los criterios, condiciones y modalidades de la reutilización de la información, así como a la publicación de la información reutilizable y al régimen sancionador.

---

<sup>32</sup> Reglamento de Transparencia y Participación Ciudadana (2012). Aprobado por acuerdo plenario de: 28-09-2012. Modificación aprobada por acuerdo plenario de 24-04-2015. Derogación del Título II por acuerdo plenario de 25-06-2020. Disponible en: [http://gobiernoabierto.valencia.es/wp-content/uploads/2014/10/R\\_Participacion-Ciudadana\\_2016\\_cas.pdf](http://gobiernoabierto.valencia.es/wp-content/uploads/2014/10/R_Participacion-Ciudadana_2016_cas.pdf)

*Reglamento de Gobierno Abierto: Transparencia*<sup>33</sup> (BOP de: 08-07-2020).

Para dar respuesta a los retos, metas y objetivos planteados desde el ámbito público se establecen una serie de principios inspiradores de actuación municipal, con el compromiso de una aplicación efectiva de estos en todo su ámbito de actuación.

Se modificó del Reglamento de Participación Ciudadana, que había sido aprobado por acuerdo Plenario de 28 de septiembre de 2012 y publicado el 26 de octubre de 2012. Mediante esta modificación, de fecha 24 de abril de 2015 se procedió a adaptar dicho reglamento a las disposiciones establecidas en la Ley 19/2013, de 9 de diciembre de 2013, de Transparencia y Buen Gobierno. Cuando se incorporaron estas modificaciones al reglamento de la regulación de la transparencia en el Ayuntamiento de Valencia, supuso un avance en relación con las previsiones contenidas en la ley estatal.

Después de estos últimos pasos en la normativa municipal del Ayuntamiento, se considera necesario aprobar un nuevo Reglamento de transparencia para adaptar dicha regulación a los cambios sociales y establecer una reestructuración del contenido para sistematizar aquellos que resultan de obligado cumplimiento para las entidades locales resultantes de dichos cambios, garantizando así mayor seguridad jurídica.

Para lograr ese objetivo, y garantizar a la ciudadanía su derecho a saber aprovechando las potencialidades que ofrece la información pública en lo que respecta a la generación de conocimiento y de valor añadido, así como la rendición de cuentas de la gestión pública municipal, se requiere la implicación de todo el Ayuntamiento en su conjunto así como su sector público, por lo que para ello resulta necesario contar con una normativa municipal que establezca las obligaciones y responsabilidades de todos los servicios y entes dependientes, siendo la forma más eficaz para hacer efectivas dichas obligaciones, así como establecer herramientas de control del poder y de integridad pública que favorezcan la prevención de malas prácticas y la generación de vínculos de confianza entre la ciudadanía y las instituciones.

La transparencia y la participación ciudadana, aun cuando pueden y deben interrelacionarse y coadyuvar a la misma finalidad, son los dos pilares sobre los que se sustenta el gobierno abierto. Se ha considerado conveniente diferenciar la regulación de la transparencia respecto de la participación ciudadana a través de dos textos normativos distintos, teniendo en cuenta la entidad propia de cada uno de esos pilares (Reglamento de Gobierno Abierto, 2019).

#### 2.9.6 Normas Técnicas para servicios de Datos Abiertos

Este apartado se centra en la normativa técnica legal existente en los niveles (europeo, estatal y autonómico) que sirve para estandarizar y facilitar la reutilización de la información del

---

<sup>33</sup> Reglamento de Gobierno Abierto: Transparencia. BOP de fecha 11 de abril de 2019. Disponible en: [https://www.valencia.es/ayuntamiento/ayuntamiento.nsf/0/FD4E2C53553EE404C125858F002F8736/\\$FILE/13.%20Aprovacion%20inicial%20Pleno.pdf?OpenElement&lang=1](https://www.valencia.es/ayuntamiento/ayuntamiento.nsf/0/FD4E2C53553EE404C125858F002F8736/$FILE/13.%20Aprovacion%20inicial%20Pleno.pdf?OpenElement&lang=1)

sector público, transparencia y gobierno abierto. Observamos los aspectos técnicos y legales a tener en cuenta para llevar a cabo de manera eficiente un servicio de datos abiertos. Además, cualquier estrategia de apertura de datos debe tener en consideración las siguientes normas técnicas que marcan las directrices a nivel de interoperabilidad, estandarización de la información, descripción de estructura y contenido de los datasets y tipos de datos abiertos que deben publicarse. Alguna de estas normas se ha nombrado anteriormente pero que ahora se pondrá en el contexto de las normas técnicas de interoperabilidad que pueden hacer posible la reutilización de la información:

- Norma Técnica de Interoperabilidad (NTI-RISP)<sup>34</sup>: para su interoperabilidad técnica, semántica y organizativa de sistemas y aplicaciones.
- DCAT-AP<sup>35</sup>: Vocabulario RDF para describir estructuradamente el contenido de datasets y catálogos de datos.
- GEO-DCAT<sup>36</sup>: Extensión DCAT para portales de datos europeos. No reemplaza el Reglamento de metadatos de INSPIRE ni las directrices técnicas de metadatos de INSPIRE basadas en ISO 19115 e ISO 19119. Su propósito es dar a los propietarios de metadatos geoespaciales la posibilidad de converger más al proporcionar una sintaxis RDF adicional.
- STAT-DCAT<sup>37</sup>: Mejorar la interoperabilidad entre las descripciones de los conjuntos de datos estadísticos.
- Norma UNE 148004<sup>38</sup>: Información Geográfica. Datos geográficos abiertos.
- Norma UNE 178301<sup>39</sup>: Ciudades Inteligentes. Datos Abiertos (Open Data).

#### *Norma Técnica de Interoperabilidad (NTI-RISP)*

Definida anteriormente, esta norma establece un conjunto de condiciones técnicas, basadas en el desarrollo tecnológico para el tratamiento de documentos y otros recursos de información elaborados o custodiados por las Administraciones Públicas. Garantiza la persistencia, términos y condiciones de uso adecuados de la información. Tiene como finalidad conseguir crear las condiciones necesarias para garantizar el adecuado nivel de

<sup>34</sup> Ya definida en apartado anterior. *Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica.*

<sup>35</sup> DCAT-AP 1.2.1. Application profile for open data portals in Europe. Disponible en: <https://joinup.ec.europa.eu/release/dcat-ap/121>

<sup>36</sup> GEO-DCAT-AP 1.0.1. Application profile for open data portals in Europe. Disponible en: <https://joinup.ec.europa.eu/release/geodcat-ap/101>

<sup>37</sup> STATDCAT-AP 1.0.1. Application profile for open data portals in Europe. Disponible en <https://joinup.ec.europa.eu/release/statdcat-ap/101>.

<sup>38</sup> Disponible (bajo suscripción) en: <https://www.aenor.com/normas-y-libros/buscador-de-normas/UNE?c=N0059630>

<sup>39</sup> Disponible (bajo suscripción) en: <https://www.aenor.com/normas-y-libros/buscador-de-normas/UNE?c=N0054318>

interoperabilidad técnica, semántica y organizativa de los sistemas y aplicaciones empleados por las Administraciones Públicas, mejorando la eficacia y eficiencia.

La NTI-RISP establece pautas para que las Administraciones Públicas puedan desarrollar sus propias políticas de reutilización de la información. Los aspectos a seguir son:

- Categorización de las informaciones
- Codificación de las URI
- Codificación de las distintas distribuciones de conjunto de datos
- Términos y condiciones de uso (licencias de los datos)
- La elaboración de catálogos de datos
- Codificación de los metadatos, basado en el estándar DCAT, vocabulario de catálogos de datos (Apartado 3.3.2)
- Codificación de algunas ubicaciones geográficas (provincias y CCAA)

Según la Guía de Aplicación de la NTI-RISP<sup>40</sup>, para una correcta organización recomienda que cada documento, recurso de información, vocabulario o esquema de conceptos, utilice el tema que más represente de acuerdo con la taxonomía que propone de los sectores primarios. Esta clasificación sirve de base común para la categorización de los catálogos de recursos de información pública y sus registros. Las categorías propuestas son las siguientes:

- Ciencia y tecnología
- Comercio
- Cultura y ocio
- Demografía
- Deporte
- Economía
- Educación
- Empleo
- Energía
- Hacienda
- Industria
- Legislación y justicia
- Medio ambiente
- Medio rural
- Salud
- Sector público

---

<sup>40</sup> Guía de aplicación de la NTI-RISP: <https://datos.gob.es/es/documentacion/guia-de-aplicacion-de-la-norma-tecnica-de-interoperabilidad-de-reutilizacion-de>

- Seguridad
- Sociedad y bienestar
- Transporte
- Turismo
- Urbanismo e infraestructuras
- Viviendas

### *DCAT-AP*

DCAT-AP<sup>41</sup> proporciona un vocabulario en formato RDF<sup>42</sup> para describir de manera estructurada el contenido de conjuntos y catálogos de datos disponibles en portales de datos abiertos. Está diseñado para facilitar la interoperabilidad entre catálogos de datos publicados en la web.

El usar DCAT-AP para describir los conjuntos de datos de los catálogos de datos (por ejemplo, el contenido de un portal de datos abiertos), permite que las aplicaciones consuman metadatos de múltiples catálogos, que se publique de forma descentralizada (catálogos) y facilitar la búsqueda de conjuntos de datos. Los metadatos DCAT-AP agregados pueden servir para facilitar la conservación digital.

Las principales aportaciones de DCAT-AP son las siguientes:

- No introduce nuevo vocabulario: define de manera concreta el uso de algunas propiedades y clases de DCAT<sup>43</sup> para la publicación de datos dentro de la Unión Europea.
- Permite definir una política completa para el uso de DCAT-AP, especificando las clases y propiedades son obligatorias, recomendadas y opcionales en la aplicación del vocabulario dentro de la Unión Europea.
- Establece unos principios normativos de conformidad para la publicación y consumo de documentos DCAT-AP.
- Explica el uso de los vocabularios controlados (en SKOS) para la descripción de la temática de los datasets.

DCAT-AP proporciona una serie de categorías que facilitan la búsqueda de los conjuntos de datos. La Norma Técnica de Interoperabilidad para la Reutilización de la Información del

---

<sup>41</sup> Data Catalog Vocabulary: <https://www.w3.org/TR/vocab-dcat/>

<sup>42</sup> Definición RDF: La sintaxis RDF define, de forma simple, el modelo para la descripción de recursos. Este modelo es fundamental para poder almacenar información de forma eficiente, procesarla automáticamente e intercambiar datos entre aplicaciones. (Senso Ruiz, José A., 2003).

<sup>43</sup> Data Catalog Vocabulary (DCAT). Disponible en <https://www.w3.org/TR/vocab-dcat/>

Sector Público propone una serie de categorías por temática. En la tabla siguiente podemos observar cómo encajan ambas las principales temáticas o categorías de estas normas técnicas.

Categorías DCAT-AP	Categorías Norma Técnica de Interoperabilidad - RISP
Science and technology	Ciencia-Tecnología
Agriculture, fisheries, forestry and food	Medio Rural Pesca
Economy and finance	Economía Comercio Hacienda
Energy	Energía
Environment	Medio ambiente
Education, culture and sport	Educación Cultura y ocio Deporte
Government and public sector	Sector público
Health	Salud
International issues	-
Justice, legal system and public safety	Legislación Justicia Seguridad
Regions and cities	Urbanismo e infraestructuras
Population and society	Demografía Sociedad y bienestar Empleo Vivienda
Transport	Transporte
-	Industria
-	Turismo

Tabla 4: Relación de temáticas DCAT-AP – Categorías propuestas desde NTI-RISP<sup>44</sup>.

### GEO-DCAT

La especificación GEODCAT-AP<sup>45</sup> es una extensión de DCAT-AP diseñada para los portales de datos europeos. Proporciona una sintaxis en RDF para combinar los esquemas de metadatos de la Directiva INSPIRE (Infrastructure for Spatial Information in Europe)<sup>46</sup> y la ISO 19115:2003<sup>47</sup> (sobre metadatos de información geográfica) de acuerdo con los principios de conformidad establecidos por DCAT-AP, posteriormente revisada a la 19115:2014<sup>48</sup>.

<sup>44</sup> Fuente: Norma Técnica de Interoperabilidad.

<sup>45</sup> GeoDCAT-AP: <https://joinup.ec.europa.eu/release/geodcat-ap-working-drafts/last-updated-2-sep-2018>

<sup>46</sup> Directiva 2007/2/CE del Parlamento Europeo y del Consejo, de 14 de marzo de 2007, por la que se establece una infraestructura de información espacial en la Comunidad Europea (Inspire). «DOUE» núm. 108, de 25 de abril de 2007, páginas 1 a 14 (14 págs.). Disponible en: <https://boe.es/buscar/doc.php?id=DOUE-L-2007-80587>

<sup>47</sup> ISO 19115:2003. Geographic information — Metadata. Modificada en 2014. Disponible en: <https://www.iso.org/standard/26020.html>

<sup>48</sup> ISO 19115-1:2014: Geographic information — Metadata — Part 1: Fundamentals. Disponible en: <https://www.iso.org/standard/53798.html>

Su caso de uso básico es hacer que los conjuntos de datos espaciales, las series de datos y los servicios se puedan buscar en los portales de datos, haciendo así que la información geoespacial se pueda buscar mejor a través de las distintas fuentes.

La especificación GeoDCAT-AP no sustituye al Reglamento de metadatos de INSPIRE ni a las guías técnicas de metadatos de INSPIRE basadas en ISO 19115 e ISO 19119. Su propósito es dar a los propietarios de metadatos geoespaciales la posibilidad de una mayor difusión al proporcionar una conexión vía sintaxis RDF. Al igual que DCAT-AP, esta especificación todavía no es normativa.

### *STAT-DCAT*

El objetivo de la normalización Stat-DCAT<sup>49</sup> es ofrecer especificaciones y herramientas para mejorar la interoperabilidad entre las descripciones de los conjuntos de datos estadísticos y los portales de datos abiertos. La primera versión de esta norma es de diciembre de 2016.

### *Norma UNE 148004. Información Geográfica. Datos geográficos abiertos.*

La Norma Española UNE 148004:2018 de “*Datos geográficos abiertos*” de la Asociación Española de Normalización (AENOR) fue definida y aprobada por el Comité Técnico de Normalización 148 sobre “*Información geográfica digital*”. Esta norma tiene como objetivo ofrecer una definición normalizada de datos abiertos geográficos que permita generar mecanismos de certificación.

El concepto de dato geográfico es fácil: todo aquello que sea referido a una localización en la Tierra. Pero la definición de dato abierto geográfico es más compleja. Para la elaboración de la norma se han tenido en cuenta los principios de la Open Knowledge Foundation y los principios que aporta la Norma UNE 148004:2018 basados en la minimización de barreras (técnicas, legales o administrativas), la no discriminación (igualitarias para todos los usuarios) y neutralidad tecnológica (utilizar estándares abiertos). De esta manera, y de acuerdo con la norma las condiciones que deben tener los datos abiertos son las siguientes:

- Deben estar disponibles en internet para su descarga total o parcial, de forma gratuita o a coste marginal.
- Deben utilizar metadatos descriptivos para facilitar la interpretación de datos de manera unívoca.

---

<sup>49</sup> *DCAT-AP y datos abiertos: aportaciones y nuevas líneas de trabajo* (2016). Disponible en: [https://administracionelectronica.gob.es/pae\\_Home/pae\\_Actualidad/pae\\_Noticias/Anio2016/Febrero/Noticia-2016-02-16-DCAT-AP-datos-abiertos-aportaciones-y-nuevas-l-neas-de-trabajo.html#.XKrsaeszZhH](https://administracionelectronica.gob.es/pae_Home/pae_Actualidad/pae_Noticias/Anio2016/Febrero/Noticia-2016-02-16-DCAT-AP-datos-abiertos-aportaciones-y-nuevas-l-neas-de-trabajo.html#.XKrsaeszZhH)



- Deben ser publicados con una licencia abierta estándar, con una definición pública, clara y conocida.
- Deben publicarse en formato digital que permita procesar la información automáticamente para su uso libre y sin restricciones.

*Norma UNE 178301. Ciudades Inteligentes. Datos Abiertos (Open Data).*

La Norma Técnica de AENOR UNE 178301:2015: “*Ciudades Inteligentes. Datos Abiertos (Open Data)*”, elaborada por el Comité Técnico AEN/CTN 178 de Ciudades Inteligentes, proporciona directrices para facilitar y mejorar la sistematización de los proyectos de datos abiertos y su gestión.

Pretende ser una referencia para las administraciones públicas para ayudar a definir, documentar e implantar proyectos de datos abiertos.

Proporciona a las administraciones locales herramientas metodológicas para evaluar el grado de madurez de su proyecto de datos abiertos, facilitando la puesta y la mejora continua, teniendo en cuenta todos aquellos aspectos que repercuten en la sostenibilidad, calidad, eficiencia y eficacia del proyecto.

Esta norma proporciona métricas e indicadores que permiten medir la madurez de los datos. Define la lista de los conjuntos de datos que son prioritarios en las iniciativas de datos abiertos, recomendando vocabularios que pueden utilizarse para la publicación.

Las métricas se encuentran clasificadas en cinco dominios y cada uno de ellos dispone de distintas dimensiones, contando con un total de 37 métricas para su evaluación. Los dominios y dimensiones son las siguientes:

- Dominio estratégico: Dimensión estratégica
- Dominio legal: Dimensión legal
- Dominio organizativo: Dimensión organizativa y Dimensión medición.
- Dominio técnico: Dimensión disponibilidad y acceso; Dimensión calidad de los datos y Dimensión actualización.
- Dominio económico y social: Dimensión reutilización y Dimensión participación y colaboración.

Metodología de la investigación:  
técnica de recogida y análisis de  
datos.

## Metodología

### 3.1 Introducción y preguntas de investigación

Cada vez son más las Administraciones Públicas en sus distintos ámbitos (locales, autonómicos y estatales) que deciden llevar a cabo iniciativas de Gobierno Abierto. Las políticas de datos abiertos conllevan un gran esfuerzo por parte de las administraciones: necesitan de recursos técnicos y humanos, una estrategia y una planificación para su puesta en marcha, un rediseño de circuitos administrativos en muchos casos, organización, definición de roles, creación de políticas internas y de reutilización, y recursos económicos para poder llevar a cabo este tipo de proyectos (Garriga-Portolà, 2011). De este modo, se mejora la gobernanza basada en el uso de datos (gobernanza inteligente), explorando la incidencia del análisis de los datos masivos en las administraciones públicas y exponiendo canales a través de los cuales ciudadanía y empresas pueden colaborar directa y activamente con las administraciones en este nuevo modelo de gobernanza, reutilizando la información del sector público y desarrollando servicios y productos que generen un valor añadido (Cerrillo-Martínez, 2018).

Esta Tesis Doctoral pretende conocer la percepción de los investigadores que ejercen su actividad en universidades, institutos y otros centros de investigación, sobre los datos abiertos que las Administraciones Públicas ponen a disposición de la ciudadanía, además de la reutilización de esta información y la utilidad de este tipo de datos en la ciencia.

Con la metodología planteada en el presente capítulo vamos a tratar de responder las siguientes preguntas de investigación:

- ¿Conocen los investigadores/as qué son los datos abiertos?
- ¿Reutilizan los investigadores información del sector público para sus investigaciones y proyectos?
- ¿Son los portales de datos abiertos de las Administraciones Públicas, fuentes de información útiles para la ciencia?
- ¿Qué funcionalidades y atributos necesitan/esperan los investigadores de los datos abiertos?
- ¿Qué barreras y dificultades encuentran los investigadores cuando acceden a estas fuentes de datos? ¿Y para reutilizar los datos?
- Dentro de los diferentes tipos de portales de datos abiertos, ¿qué tipo de fuentes son las más utilizadas (ámbito estatal, autonómico, local)?
- ¿Los datos abiertos han servido para llevar a cabo proyectos y otros trabajos de investigación?
- ¿Qué tipos de datos son los más consultados y útiles para la producción científica?

Para tratar de contestar las preguntas planteadas se utiliza una técnica de investigación cuantitativa mediante una encuesta. Con el *feedback* obtenido, podremos dar respuesta a las preguntas de investigación y realizar un análisis detallado de la situación actual.

### 3.2 Metodología de la investigación: el estudio cuantitativo como técnica para la recogida de datos en la investigación social

La metodología de la investigación es el resultado de la aplicación sistemática y lógica de los conceptos y fundamentos expuestos en el marco teórico. Debe servir para la obtención de nuevo conocimiento y resolver problemas específicos planteados al inicio de la investigación. Se sustenta en los pilares fundamentales de reproducibilidad y refutabilidad (Labajo González, 2015):

- La reproducibilidad es la capacidad de repetición de un experimento en cualquier lugar y por cualquier persona. Este primer pilar se basa en la comunicación y publicidad de los resultados obtenidos.
- Refutabilidad significa que toda proposición científica tiene que ser potencialmente susceptible de ser falseada o refutada. Implica que se pueden diseñar experimentos donde se obtengan resultados diferentes a los predichos que negarían la hipótesis puesta a prueba.

Además de estos principios de reproducibilidad y refutabilidad, la metodología debe ser fácil de comunicar. En función de la naturaleza y el objeto a estudiar, aplicaremos distintos procedimientos, según la necesidad (Raffino, 2020).

Para esta tesis doctoral se emplea un método cuantitativo. Se trata de una investigación social, donde necesitamos una gran cantidad de datos objetivos por parte de la población a estudiar: miembros de la comunidad científica a nivel estatal que formen parte de universidades y otros centros de investigación. La metodología cuantitativa consiste en el contraste de teorías existentes a partir de hipótesis surgidas de la misma. Es necesario la obtención de una muestra (aleatoria o discriminada) pero que sea representativa de la población o del fenómeno como objeto de estudio. Utiliza la recolección y el análisis de datos para contestar a preguntas de investigación y probar hipótesis establecidas previamente. Con los datos obtenidos se realiza la medición numérica y la aplicación de la estadística para establecer patrones de comportamiento de la población. (Tamayo, 2007).

El método cuantitativo tiene su base en el positivismo<sup>50</sup>, que busca las causas a problemas planteados mediante métodos como el cuestionario, donde se obtienen datos que pueden ser

---

<sup>50</sup> En 1849 se inicia el paradigma “positivista” dentro de las Ciencias Sociales de la mano del filósofo Auguste Comte en su “*Discurso sobre el espíritu positivo*”.

analizados estadísticamente (siendo un método deductivo). La objetividad es de vital importancia dentro del paradigma positivista, bajo el cual el investigador observa, mide y manipula variables. Además, se debe desprender de sus propias tendencias y prejuicios, y la relación entre el investigador y el fenómeno u objeto de estudio debe ser de independencia total (Angulo López, 2011).

Este método de trabajo se centra en los hechos o causas del fenómeno social, más que en los estados subjetivos del individuo. Uno de los instrumentos de recogida de datos para su posterior medición es el cuestionario. El cuestionario genera datos que posteriormente serán analizados estadísticamente para verificar o rechazar las relaciones entre las variables establecidas inicialmente. La forma en la que se obtienen dichos datos es mediante preguntas, garantizando siempre el anonimato del sujeto. A diferencia de una investigación cualitativa, donde los investigadores hacen registros narrativos que son estudiados mediante técnicas como la observación y las entrevistas no estructuradas, la investigación cuantitativa recoge y analiza datos sobre variables previamente definidas. Los resultados obtenidos se presentan mediante visualizaciones como tablas o gráficos (Rodríguez Peñuelas, 2010) y (López-Roldán y Fratelli, 2015).

Por tanto, con este tipo de método<sup>51</sup> lo que se busca es una recolección de datos para probar hipótesis con una base estadística y numérica. Con las respuestas obtenidas y el análisis posterior, se establecen los patrones de comportamiento y prueba de teorías. Las fases son las siguientes (Hernández, Fernández y Baptista, 2010):

- Determinación de los objetivos y preguntas de investigación.
- Revisión de la literatura.
- Construcción de un marco o perspectiva teórica.
- De las preguntas previas se establecen las hipótesis y se determinan las variables.
- Se desarrolla un plan para probarlas (Diseño).
- Se miden las variables en un determinado contexto.
- Se analizan las mediciones obtenidas (método estadístico).
- Finalmente, se establecen las conclusiones respecto a la hipótesis inicial.

Se trata de una técnica de recolección de información ya asentada en los procesos de investigación. Para llevarlas a cabo correctamente, se tienen en cuenta factores como la importancia del diseño de la muestra (a mayor muestra encuestada, mayor confiabilidad de la investigación), la construcción del cuestionario, la medición, la construcción de índices y

---

<sup>51</sup> En las Ciencias Sociales, este enfoque cuantitativo se origina en las obras de Emile Durkheim (1858-1917) y Auguste Comte (1788-1857). Fueron ellos quienes propusieron que para medir y estudiar los fenómenos sociales se utilizaran técnicas “científicas”. Es decir, que el método pudiera ser aplicable a otros estudios al igual que en las ciencias exactas. Sostenían que los fenómenos a estudiar dentro de las ciencias sociales eran medibles. De ahí la corriente “positivista”.

escalas, las entrevistas, la codificación, la organización, el seguimiento del trabajo de campo, la preparación de los datos para el análisis y las técnicas utilizadas, el registro, y la presentación de los resultados obtenidos.

Realizar encuestas implica un proceso donde los aspectos anteriormente mencionados están ligados a la misma, y se deben integrar coherentemente con el objetivo de obtener una producción científica de calidad y en correspondencia con el modelo de análisis construido, donde se requiere el conocimiento especializado y la capacidad de aplicación. Tiene puntos fuertes y carencias, que describimos a continuación (Cook y Reichardt, 1986; López-Roldán y Fratelli, 2015; Cadena-Íñiguez, et al., 2017).

Ventajas.

- Capacidad de involucrar a un gran número de sujetos de la investigación. Esto puede conducir a resultados más fiables.
- Objetividad de la investigación cuantitativa, gracias a los datos obtenidos.
- Creación de un contexto de máxima objetividad, al no participar el investigador.

Desventajas.

- Obliga a trabajar con definiciones cortas y precisas, sin muchos matices, dadas las restricciones del sistema de cuestionario, que no permite demasiadas explicaciones a riesgo de que el participante no las entienda y por tanto conteste erróneamente.
- Por lo tanto, a veces este tipo de respuestas puede no aportar al investigador una comprensión demasiado profunda.
- El investigador no participa activamente, limitando la posibilidad de resolver problemas de estructura y comprensión sobre la marcha.

Es aceptable este tipo de metodologías en la investigación cuando el estudio se lleva a cabo en distintas regiones geográficas al mismo tiempo (a nivel estatal, España, en este caso), cuando se acceda a un gran número de sujetos de investigación, cuando no hay necesidad de comprender un fenómeno sino de presentar datos objetivos, y cuando es posible tomar muestras de población estadísticamente significativas (Tavarez, 2020).

Finalmente, clasificamos este método de investigación como analítico correlacional-descriptivo ya que las variables que nos interesan son seleccionadas para conocer la relación entre ellas, de manera que tenemos cierto control sobre las variables identificadas previamente, lo que nos va a permitir contestar las preguntas de investigación planteadas (Casas Anguita, Repullo Labrador y Donado Campos, 2003).

### 3.3 Población de estudio

La selección de la población o muestra dentro de una investigación es un hecho que nunca debe estar sujeto al azar o capricho del investigador. La selección de la muestra debe responder a una serie de criterios teóricos y metodológicos que vienen dados en función del problema a investigar, los objetivos y preguntas de investigación y el diseño de esta. Con el fin de llevar a cabo con éxito la investigación, es necesario precisar correctamente la población y la muestra a estudiar (Salas Ocampos, 2020).

El muestreo es el método para seleccionar los componentes de la muestra total de la población. La muestra se selecciona mediante un mecanismo probabilístico, aportando un número pequeño de unidades estadísticas elementales dentro de un conjunto mayor, que es la población. El fin es obtener respuestas sobre cuestiones planteadas para extrapolarlo al resto de la población (Hernández, 2012). Realizar el diseño muestral es importante porque permite que el estudio se realice en menos tiempo, acarrea menos gastos, facilita el análisis entre variables y permite tener un control mayor de las variables a estudiar (López, 2004).

En cuanto al tamaño de la muestra, no existe una fórmula para su selección. Se deben observar y tener en cuenta consideraciones como los recursos disponibles y los requerimientos del análisis de investigación. Cuanto mayor sea la muestra, el error muestral será menor y el estudio será más representativo (López, 2004). Según Pineda et al., el investigador puede seguir la siguiente lógica para la obtención de la muestra. Veamos dos casos extremos.

- Si la población es de 100 individuos, hay que seleccionar al menos el 30% para no tener menos de 30 casos; de otra forma, caeríamos en la categoría de *muestra pequeña*, lo que matemáticamente implica un error mayor, a veces demasiado.
- Sin embargo, si la población es de 50.000 individuos potencialmente implicados en el análisis (más o menos nuestro caso), el 30% representa 15.000 individuos, que es una cantidad excesivamente grande; el 10% serán 5.000; el 1% serán 500. Los criterios estadísticos de selección de muestras indican que una muestra de un 1%, en este caso, será adecuada para poder llevar a cabo la investigación.

Esta información puede ampliarse utilizando otras maneras de obtener el tamaño de la población a estudiar mediante fórmulas estadísticas que dependen de la confiabilidad, probabilidad y el margen de error (López, 2004). Para nuestro caso, consideramos que 500 participantes es más que suficiente para garantizar la significatividad del estudio.

En esta tesis se va a estudiar un nicho de población muy concreto; la población a la que se va a distribuir la encuesta (a nivel estatal) está compuesta por:

- Investigadores de centros, institutos y fundaciones de investigación nacionales.
- Profesores y otros técnicos de investigación de las universidades (públicas y privadas).
- Doctorandos de la Universitat Politècnica de València<sup>52</sup>.

Según los datos disponibles, el Sistema Universitario Español (SUE)<sup>53</sup> lo conforman 83 universidades con actividad, siendo 50 públicas y 33 privadas, contabilizándose 1061 centros universitarios entre escuelas y facultades, 537 institutos universitarios de investigación, 50 escuelas de doctorado, 54 hospitales universitarios y 76 fundaciones. En cuanto al personal docente e investigador (PDI) se sitúa en 125.471 personas, de las cuales 105.371 pertenecen a universidades públicas y 20.100 a universidades privadas. El personal investigador (PI) y personal técnico de apoyo a la investigación, alcanza los 26.408 efectivos, de los cuales 24.886 están adscritos a universidades públicas y 1.522 a universidades privadas. Finalmente, el personal empleado investigador cuenta con 19.879 investigadores contratados (Ministerio de Universidades, 2021).

En el Anexo I se presentan las universidades y centros de investigación donde se ha distribuido la encuesta. Se ha enviado a un total de 22791 direcciones de correo electrónico extraídas de las páginas webs de las universidades y centros de investigación de 47 universidades. Se redactó un mensaje institucional para explicar su objetivo (ver Anexo II) y se distribuyó a través de la plataforma de encuestas Survey Monkey. La encuesta estuvo disponible en la página web de la Cátedra de Transparencia y Gestión de Datos de la Universitat Politècnica de València (<https://www.ctranspa.webs.upv.es/encuesta-datos-abiertos/>). Además, la Conselleria de Participación, Transparencia, Cooperación y Calidad Democrática, publicó la encuesta. Además, se obtuvieron respuestas de personal no propio de universidades y pertenecientes a centros de investigación.

Se han obtenido un total de 786 respuestas<sup>54</sup> al cuestionario. En aquellas universidades donde había prohibición expresa de usar el directorio para el envío de campañas de *mailing* de este tipo, no se ha remitido la encuesta.

Como hemos visto en la población de estudio, el Sistema Universitario Español (SUE), está formado por 83 universidades con 125471 de personal docente e investigador (PDI) (Ministerio de Universidades, 2021). La muestra final fue de 47 universidades, y un total de 22791 direcciones de correo.

---

<sup>52</sup> El motivo de que sean de la Universitat Politècnica de València y no de otras universidades de la Comunidad Valenciana es la disposición como doctorando del listado de correos electrónicos (distribución de la encuesta).

<sup>53</sup> Datos y cifras del Sistema Universitario Español:

[https://www.universidades.gob.es/stfls/universidades/Estadisticas/ficheros/Datos\\_y\\_Cifras\\_2020-21.pdf](https://www.universidades.gob.es/stfls/universidades/Estadisticas/ficheros/Datos_y_Cifras_2020-21.pdf)

<sup>54</sup> Se obtienen 786 respuestas. Se eliminan 3 por no tener contenido. En total, 783.



### 3.4 La encuesta

#### 3.4.1 Revisión bibliográfica. Estudios previos.

Para el desarrollo de la metodología de esta tesis se han tenido en cuenta estudios llevados a cabo anteriormente. Se ha realizado una búsqueda bibliográfica para ver encuestas distribuidas en el pasado con el objetivo de conocer qué se ha investigado, no duplicar información repitiendo investigaciones ya realizadas, y reutilizar preguntas que puedan ser interesantes para el presente estudio, adaptándolas para nuestro caso concreto a investigar. En esta búsqueda bibliográfica, hemos encontrado dos estudios que han inspirado el análisis cuantitativo de esta tesis.

El primer estudio que ha servido como referencia ha sido el que llevó a cabo Tim Davies (Gran Bretaña, 2010)<sup>55</sup> para conocer quién y de qué manera se reutilizan los datos abiertos. Se trata de una encuesta<sup>56</sup> de 55 preguntas en la que se busca conocer la opinión de los reutilizadores, y para qué tipos de proyectos han reutilizado datos abiertos del sector público.

La segunda investigación fue la llevada a cabo en el Foro de Belmont<sup>57</sup> en el año 2015. Se trata de un estudio de percepción de un tipo de datos en concreto: datos medioambientales. Con los resultados obtenidos se mostraron los obstáculos que encuentran los reutilizadores para el uso e intercambio de los datos. Por otra parte, los encuestados señalaron un amplio abanico de ejemplos de buenas prácticas, poniendo en común datos a través de infraestructuras electrónicas que han servido para mejorarlas y consolidarlas.

Se han tenido en cuenta otros estudios llevados a cabo a nivel internacional, europeos y nacionales para tratar de utilizar en el caso de que se pudiera, alguna pregunta en materia de reutilización de datos abiertos aplicable al ámbito de la investigación. Los informes consultados han sido los de ASIEDIE<sup>58</sup> (2013-2021), del sector infomediario en España en el ámbito público y privado, un informe del European Data Portal<sup>59</sup> e informes sobre la reutilización de los datos abiertos en España desarrollado por Alberto Abella et al. (2019).

La Organización para la Cooperación y el Desarrollo Económico llevó a cabo en 2014 y 2017 dos encuestas. En la primera, de 54 preguntas, el objetivo fue realizar un estudio sobre la

---

<sup>55</sup> Davies, T. (2010). Open data, democracy and public sector reform. *A look at open government data use from data.gov.uk*, 1-47.

<sup>56</sup> Encuesta Open Data Impacts:

<http://www.opendataimpacts.net/wp-content/uploads/2010/06/OpenDataImpacts-Survey-Form.pdf>

<sup>57</sup> Schmidt B., Gemeinholzer B. y Treloar A (2016). Open Data in Global Environmental Research: The Belmont Forum's Open Data Survey. PLoS ONE 11(1): e0146695. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0146695>

<sup>58</sup> ASIEDIE (2020). Sector Infomediario: 2020. Disponible en: <http://www.asedie.es/assets/informe-asedie-sobre-el-sector-infomediario-2020-vf.pdf>

<sup>59</sup> European Data Portal: <https://www.europeandataportal.eu/sites/default/files/the-economic-impact-of-open-data.pdf>

importancia e impacto de los datos en el gobierno abierto (Open Government Data 2.0., de ahora en adelante OGD)<sup>60</sup>. El estudio se estructuró en cuatro partes, y en él se buscaba conocer el marco de políticas y gobernanza de la OGD, la implementación de OGD (incluyendo un módulo especial en los portales centrales/federales de "ventanilla única"), el impacto del OGD y los principales desafíos para la creación de valor.

En el segundo estudio llevado a cabo en 2017<sup>61</sup>, se realizó una encuesta a reutilizadores de datos abiertos, que se estructuró en 3 bloques con un total de 80 preguntas, para conocer la percepción y conocimiento sobre la disponibilidad y accesibilidad de los datos y el apoyo de los gobiernos en reutilización en negocios, sociedad civil y funcionarios públicos.

### 3.5 Estructura

Tal y como se recomienda, las primeras preguntas son sencillas y motivadoras, habiéndose reservado el espacio central del cuestionario para las más importantes de la investigación. Se ha cerrado el cuestionario con preguntas fáciles para que el encuestado finalice con una buena sensación. Las preguntas de identificación se han situado al inicio, garantizando el anonimato de los encuestados.

Las preguntas planteadas se han agrupado por temas y cada uno de los bloques tiene preguntas introductorias. El cuestionario tiene un total de 30 preguntas distribuidas en 6 preguntas en un bloque de contexto inicial y 24 preguntas distribuidas en 6 bloques técnicos. El tiempo necesario para completarlo es de 12 minutos máximo, en función de si conoce o no el movimiento de los datos abiertos y posteriormente si se quiere aportar información adicional.

---

<sup>60</sup> Organización para la Cooperación y el Desarrollo Económico (2014): <https://www.oecd.org/gov/digital-government/2014-open-government-data-survey.pdf>

<sup>61</sup> Organización para la Cooperación y el Desarrollo Económico (2017): <https://www.oecd.org/gov/2016-OECD-Survey-on-Open-Government-Data-3.0.pdf>

### 3.6 Pre-Test. Revisión de científicos y la Conselleria de Participación, Transparencia, Cooperación y Calidad Democrática

La encuesta se desarrolló durante los meses de octubre y noviembre de 2020, inspirada, tal y como hemos visto en la revisión bibliográfica, en dos estudios llevados a cabo con anterioridad. A inicios de diciembre de 2020 y durante el mes de enero de 2021, el pre-test fue distribuido y revisado por profesores e investigadores de la Universitat Politècnica de València, de la Universitat de València, miembros del Consejo Superior de Investigaciones Científicas, doctorandos de la Universitat Politècnica de València y la Universitat de València y por la Conselleria de Participación, Transparència, Cooperación y Calidad Democrática.

En un primer momento, y a expensas de obtener *feedback*, este primer pre-test contaba con un total de 39 preguntas distribuidas en 9 bloques. El objetivo fue comprobar los tipos de preguntas que se debían mantener, las que se debían replantear, y cuáles eliminar. La estructura inicial de este primer test fue:

- Bloque 1. Información personal y laboral.
- Bloque 2. Apertura de datos: uso de datos y reutilización de la información del sector público.
- Bloque 3. Utilidad, motivaciones y barreras encontradas en el acceso a los datos.
- Bloque 4. Descubrimiento y acceso a conjuntos de datos abiertos.
- Bloque 5. Interés en iniciativas de datos abiertos: categorías y tipos de datos.
- Bloque 6. Opiniones y visión acerca de los datos abiertos.
- Bloque 7. Uso y reutilización de datos.
- Bloque 8. Proyectos llevados a cabo con datos abiertos.
- Bloque 9. Métricas de evaluación de datos.

### 3.7 Estructura definitiva

Una vez obtenidas las respuestas, percepciones y comentarios, se reestructuró la encuesta. Según los encuestados en el pre-test, algunas preguntas eran repetitivas, otras no aportaban ningún tipo de información relevante para el estudio y otras se reformularon para mejorar su comprensión. Se redujeron, de 39 a 30, las preguntas, pasando de 9 a 6 bloques de información. La encuesta final se puede consultar en el Anexo I. La estructura definitiva es la siguiente:

- Bloque de presentación: información personal y laboral.
- Bloque 1, de contexto inicial: Conocimiento acerca de los datos abiertos.

- Bloque 2, de apertura de datos. Uso y reutilización de la información del sector público.
- Bloque 3. Utilidad real y motivaciones para la reutilización de datos abiertos en la investigación.
- Bloque 4. Acceso a los conjuntos de datos: fuentes de datos abiertos consultadas.
- Bloque 5. Uso y reutilización de datos (I). Tratamiento y reutilización.
- Bloque 6. Uso y reutilización de datos (II). Uso particular de los datos abiertos.

### 3.8 Plataforma de la encuesta: Survey Monkey

La plataforma elegida para distribuir la encuesta ha sido Survey Monkey<sup>62</sup>. Se trata de una web para realizar encuestas que nos permite recopilar datos de una forma sencilla. Dispone de una interfaz intuitiva donde podemos construir una encuesta desde cero y compartirla a través del correo electrónico, o mediante un enlace. Permite exportar los datos a otras herramientas para analizar posteriormente los resultados y en base a estos, tomar las decisiones correspondientes.

El primer correo se envió el día 15 de marzo de 2021, cerrando la encuesta el 10 de mayo de 2021 (56 días). En el anexo II está disponible la encuesta tal y como figura en la plataforma del enlace siguiente: [https://es.surveymonkey.com/r/datos\\_abiertos\\_investigadores](https://es.surveymonkey.com/r/datos_abiertos_investigadores).

Como hemos visto en el apartado de la población de estudio, la encuesta se distribuyó a través de una base de datos de correos electrónicos extraídos de las páginas web institucionales.

La encuesta comenzó a ser distribuida el 15 de marzo de 2021 y se cerró el día 10 de mayo de 2021. Entre el 15 de marzo y el 19 de abril, se obtuvieron muchas respuestas correspondientes a las universidades y centros de investigación de la Comunidad Valenciana debido al correo institucional enviado desde la Conselleria de Participación, Transparència, Cooperación y Calidad Democrática y el correo remitido a todos los investigadores por parte del Vicerrectorado de Investigación de la Universitat Politècnica de València.

Posteriormente, con el objetivo de no hacer un envío masivo a todo el directorio de correos de universidades españolas que disponíamos, Survey Monkey nos daba la opción de crear recopiladores. Los recopiladores son una configuración de listados de correo electrónico dentro de la propia plataforma, que da la posibilidad de programar fecha y hora de envío, además de la cantidad de direcciones a las que se va a distribuir en cada envío. Se crearon 25 recopiladores distribuidos entre el 20 de abril de 2021 y el 4 de mayo de 2021 (ver Anexo III), cerrando definitivamente el estudio, como se ha indicado, el día 10 de mayo de 2021.

---

<sup>62</sup> SurveyMonkey. Página web: <https://es.surveymonkey.com/>

### 3.9 Técnicas para el análisis y visualización de datos

Para llevar a cabo el análisis, sistematizamos y simplificamos la información procedente de los cuestionarios para tratar los resultados de un modo automatizado<sup>63</sup>. En los casos en los que sea necesario, se codifican las respuestas en etiquetas que las identifican.

Survey Monkey nos permite descargar los datos obtenidos en varios formatos. En este caso, se descarga la información en formato *xlsx* (hojas de cálculo Excel) para llevar a cabo el análisis de los resultados por bloques de preguntas planteadas.

Las visualizaciones de datos se crean en base a los datos descargados desde la plataforma de encuestas. El tratamiento y normalización de los datos se ha llevado a cabo con la hoja de cálculo de Excel y Open Refine. Para la creación de las visualizaciones de esta tesis se ha utilizado la misma hoja de cálculo de Excel ya que es una herramienta que permiten analizar y visualizar grandes volúmenes de información, además de cruzar variables.

Finalmente, se ha desarrollado una plataforma basada en el lenguaje de programación *R*, diseñado para el análisis estadístico y gráfico, de los más utilizados en la minería de datos y aplicable a numerosas disciplinas. Para la creación de tablas, gráficos e interacción con los datos se utiliza *Shiny*, (<https://shiny.rstudio.com/>) paquete de *R* para la creación de aplicaciones web interactivas, pudiendo alojar aplicaciones independientes en una página web o insertarlas en documentos de *R Markdown*.

---

<sup>63</sup> Para la creación de los histogramas que se presentan en el capítulo de resultados hemos llevado a cabo esta codificación de resultados. En el anexo II, dónde está disponible la encuesta que se ha distribuido a los investigadores, aparecen las preguntas codificadas. Estas preguntas son: 4-5-6-7-8-9-11-12-16-17-18-21.

<b>Período</b>	Inicio: 15/03/2021	Fin: 10/05/2021
<b>Objetivo</b>	<ul style="list-style-type: none"> <li>• Conocer la percepción de docentes universitarios y miembros de centros de investigación a nivel estatal sobre percepción, conocimiento y uso de datos abiertos.</li> </ul>	
<b>Muestra a estudiar</b>	<ul style="list-style-type: none"> <li>• Investigadores de universidades y centros de investigación a nivel estatal.</li> <li>• Tamaño de la muestra: 22791 encuestados. Se obtienen 783 respuestas de 34 provincias españolas, de 47 universidades y 21 centros de investigación.</li> </ul>	
<b>Metodología</b>	<ul style="list-style-type: none"> <li>• Muestreo</li> </ul>	
<b>Técnica de recogida de datos</b>	<ul style="list-style-type: none"> <li>• Elaboración y desarrollo de la encuesta en la plataforma Survey Monkey.</li> <li>• Técnica de <i>bola de nieve</i> a través de correo electrónico. Difusión secundaria a por transmisión de persona a persona.</li> </ul>	
<b>Instrumento para obtención de datos</b>	<ul style="list-style-type: none"> <li>• Cuestionario en línea.</li> </ul>	
<b>Procesamiento</b>	<ul style="list-style-type: none"> <li>• Descarga de datos de la plataforma Survey Monkey.</li> <li>• Limpieza y análisis de datos.</li> <li>• Desarrollo de tablas, visualizaciones y comentarios.</li> </ul>	

Tabla 5: Resumen de la metodología.

Capítulo 4

# Resultados

## Resultados

### 4.1 Introducción

Tal y como se indica en el capítulo de la metodología de la investigación, en esta tesis se desarrolla un estudio cuantitativo, para el que se diseñó y distribuyó una encuesta al personal de la comunidad científica vinculado a universidades españolas. La difusión se hizo a través de la creación de un directorio, con 22791 direcciones de correo electrónico, extraídas de las páginas web de 47 universidades y centros de investigación<sup>64</sup>, respetando aquellas que especificaban que no deben utilizarse para campañas de envíos de correos masivos. La plataforma elegida para realizar la encuesta fue Survey Monkey.

La encuesta empezó a distribuirse el 15 de marzo de 2021 y se cerró el día 10 de mayo de 2021. Con el objetivo de no hacer un envío masivo a todo el directorio de correos de universidades españolas que disponemos, Survey Monkey tiene una opción de crear recopiladores como ya se ha señalado. Se obtuvieron un total de 786 respuestas (se eliminan 3 que no contienen información, por tanto, 783 respuestas).

El capítulo se estructura por los bloques de la encuesta, es decir, del bloque 0 al bloque 6 (7 en total). Para cada uno de los bloques se explica en primer lugar el proceso de limpieza y estandarizado de los datos obtenidos para, en un segundo apartado, representarlos a través de tablas y gráficos y, posteriormente, realizar el análisis y los comentarios correspondientes.

Las tablas se han extraído directamente del tratamiento de los datos en la hoja de cálculo Excel, ya que la propia plataforma de Survey Monkey nos permite exportar los datos en varios formatos. La creación de las visualizaciones y gráficos de este documento, se han realizado también con Excel, ya que nos permite procesar y analizar grandes volúmenes de datos y cruzar variables.

Finalmente, se ha creado una plataforma web con los datos de la encuesta. La tecnología que se ha utilizado es el lenguaje de programación R<sup>65</sup>, orientado al análisis estadístico y matemático y que nos permite tratar grandes volúmenes de datos. Para el desarrollo de la aplicación web se ha utilizado el paquete de R, *Shiny*<sup>66</sup>, que permite la creación de páginas webs interactivas a partir de los scripts de R.

---

<sup>64</sup> Ver Anexo I: Universidades y centros de investigación del estudio.

<sup>65</sup> Página web de R-Project: <https://www.r-project.org>

<sup>66</sup> Extensión Shiny para R: <https://shiny.rstudio.com>



## 4.2 Limpieza y normalización de datos

Tras la obtención de las respuestas de los encuestados, los datos se almacenan en la plataforma *Survey Monkey*. Para poder llevar a cabo el análisis, la creación de tablas, visualizaciones, y posteriormente el desarrollo de una plataforma *Shiny* basada en el lenguaje de programación R (ver capítulo de metodología), los datos deben estar estandarizados. Esta fase es la que conocemos como fase de limpieza y homogeneización de datos.

La propia plataforma nos da la opción de exportarlos en distintos formatos. En primer lugar, se exportan en una hoja de cálculo Excel, en formato XLSX. Se detecta que, de las 786 respuestas, 3 no contienen información y por tanto se eliminan<sup>67</sup>.

Esta limpieza y estandarización de datos se hace sobre la hoja de cálculo Excel. Tiene funciones y fórmulas que nos permite estandarizar los datos de una manera ágil.

La encuesta, está formada por 6 preguntas dentro del bloque 0, donde se trata el contexto del investigador y 24 preguntas en los bloques del 1 al 6, sobre conocimiento, percepción y uso de los datos abiertos.

Las preguntas no son de respuesta obligatoria, motivo por el cual se encuentran preguntas sin responder. El objetivo era no condicionar al encuestado/a y que pudiera no responder si así lo consideraba. En caso necesario, en la presentación de resultados de cada pregunta se especifica cuántos usuarios/as han omitido la respuesta.

La hoja de cálculo que genera la plataforma tiene un total de 788 filas, de las cuales eliminamos 3 que no contienen información y 2 filas son los encabezados de las columnas. En total, 783 filas donde cada fila son las respuestas de cada encuestado. Debido a que algunas preguntas son de matrices, otras en las que se puede elegir entre varias opciones y otras son de respuesta abierta, la hoja de cálculo tiene columnas A hasta la WB (600 columnas).

La limpieza y estandarización de los datos se realiza por bloques sobre la hoja de cálculo Excel, utilizando sus funciones para agilizar y automatizar el trabajo y así posteriormente hacer representaciones y cruces de datos a representar en esta tesis.

Tanto las tablas que se presentan como los gráficos y otras visualizaciones del capítulo son de elaboración propia.

---

<sup>67</sup> El acceso al enlace del correo electrónico a la encuesta contabilizaba como respuesta. Por tanto, estas 3 filas vacías se eliminan.

### 4.3 Análisis y presentación de los datos

#### Bloque 0. Información personal y laboral

##### B-0. Estandarización y limpieza de datos

El bloque 0 está formado por 6 preguntas de contexto para obtener información de los investigadores encuestados. De las 6 preguntas, 2 son de respuesta abierta: la provincia de residencia y la universidad/centro de investigación al que pertenecen. En ambas preguntas se plantean una serie de opciones con una opción de respuesta abierta.

Para cerrar el bloque, se han planteado 2 preguntas con una serie de opciones cerradas a elegir.

Descargada la hoja de cálculo y empezando a normalizar el bloque 0 se detecta lo siguiente:

- Se eliminan 3 filas que se contabilizan como respuesta pero que no contienen información (ya especificado).
- Casos de usuarios/as que rellenan la provincia con respuestas que no son provincias: Chile, México, Ecuador, Francisco Morazán (Honduras), Cataluña, Andalucía.
- Usuarios/as que rellenan la universidad/centro de investigación con la siguiente respuesta: Universidad/centro de investigación: UCC<sup>68</sup>.
- Miembros vinculados a la Universitat Politècnica de València<sup>69</sup> que establecen:
  - Usuario/a 1. Da como respuesta a la universidad o centro de investigación a la Universidad Central de Ecuador y en el campo provincia de residencia, proporciona como respuesta Ecuador.
  - Usuario/a 2. Da como respuesta a la universidad o centro de investigación a la Universidad Nacional Autónoma de Honduras y en el campo provincia de residencia, proporciona como respuesta Francisco Morazán (Honduras).
  - Usuario/a 3. Da como respuesta a la universidad o centro de investigación a la Universitat Politècnica de València y en el campo provincia de residencia, proporciona como respuesta Chile.
  - Usuario/a 4. Da como respuesta a la universidad o centro de investigación a la Universitat Politècnica de València y en el campo provincia de residencia, proporciona como respuesta México.
- Un total de 11 usuarios/as no indican información sobre la provincia de residencia ni el centro de investigación/universidad.

<sup>68</sup> UCC: Estas siglas no corresponden a ninguna de las universidades a las que se distribuye la encuesta a través del correo electrónico.

<sup>69</sup> Sabemos que los usuarios/as 1 y 2 están vinculados a la Universitat Politècnica de València porque a pesar de no indicar como centro de investigación a la Universitat Politècnica de València, los datos se obtienen a través del recopilador de la UPV (Véase el apartado *Metodología – recopiladores de la encuesta*).

Las preguntas que no necesitan normalización son las de rango de edad y sexo del encuestado, ya que la respuesta viene de las opciones planteadas en el enunciado.

En las preguntas de categoría laboral y área de investigación, se plantearon una serie de opciones y, además, en la pregunta sobre el área de investigación se dio la opción de respuesta abierta para aquellas personas que su caso no estuviera incluido en las opciones planteadas en el enunciado.

En la pregunta sobre la categoría laboral, se planteaban una serie de opciones entre las que el encuestado/a debía elegir, teniendo una opción de respuesta abierta. En el proceso de limpieza de datos, se detectó que no se habían planteado opciones como Profesor/a Colaborador/a, Profesor/a Sustituto/a Interino/a o puestos relacionados con el Personal de Administración y Servicios.

Se han añadido todas aquellas categorías con más respuestas, quedando normalizadas de la siguiente manera:

Abreviatura	Categoría Laboral
ASO	Profesor/a Asociado/a
VISIT	Profesor/a visitante
AYU	Profesor/a Ayudante
AYUDOC	Profesor/a Ayudante Doctor
PCDOC	Profesor/a Contratado/a Doctor
TU	Profesor/a Titular de Universidad
CIT	Científico/a Titular
INVCIENTI	Investigador/a Científico
PROFINV	Profesor/a de Investigación
DOC	Doctor/a
DOCN	Doctorando/a
CU	Catedrático de Universidad
CEU	Catedrático de Escuela Universitaria
TEU	Profesor/a Titular de Escuela Universitaria
COL	Profesor/a Colaborador/a
ASOCCS	Profesor/a Asociado/a en Ciencias de la Salud
EMER	Profesor/a Emérito/a
TECINV	Técnico/a de investigación
EMP	Empresa
PAS	Personal de Administración y servicios
PSI	Profesor/a Sustituto/a Interino/a

Tabla 6: Categorías laborales normalizadas.

En cuanto al área de investigación, se utilizaron las áreas normalizadas de la Agencia Estatal de Investigación del Ministerio de Ciencia e Innovación<sup>70</sup>. Se detectó que, en la opción abierta, algunos investigadores introdujeron áreas que pertenecían a alguna de las 19 planteadas (por ejemplo, arquitectura y ciencias de la salud, que forman parte del área 11). Tanto las opciones planteadas en el enunciado como las obtenidas a través de la respuesta abierta, se buscaron en la Agencia Estatal de Investigación y se añadieron a su área/categoría correspondiente:

Áreas de investigación
Área 1: Ciencias Sociales
Área 2: Derecho
Área 3: Economía
Área 4: MLP/ Mente, Lenguaje y Pensamiento
Área 5: FLA/ Cultura, Filosofía, Literatura y Arte
Área 6: PHA/ Estudios del pasado: Historia y Arqueología
Área 7: EDU/ Ciencias de la educación
Área 8: PSI/ Psicología
Área 9: MTM /Matemáticas
Área 10: FIS/ Ciencias físicas
Área 11: PIN/ Producción industrial, ingeniería civil e ingenierías para la sociedad
Área 12: TIC/ Tecnologías de la Información y las Comunicaciones
Área 13: EYT/ Energía y Transporte
Área 14: CTQ/ Ciencias y Tecnologías Químicas
Área 15: MAT/ Ciencias y Tecnologías Materiales
Área 16: CTM/ Ciencias y Tecnologías Medioambientales
Área 17: CAA/ Ciencias Agraria y Agroalimentarias
Área 18: BIO/ Biociencias y biotecnología
Área 19: BME/ Biomedicina

Tabla 7: Áreas de investigación. Agencia Estatal de Investigación.

Las preguntas de provincia de residencia y centro de investigación/universidad a la que pertenecen los investigadores son de respuesta abierta. Se han normalizado las universidades y centro de investigación con el nombre oficial de la institución. Las provincias españolas también se han normalizado.

Para ello, sobre la hoja de cálculo Excel se ordenaron los datos para tenerlos agrupados. Posteriormente con las funciones de *buscar* y *reemplazar*, se han sustituido las respuestas de los investigadores por los nombres oficiales de la institución y las provincias.

<sup>70</sup> Áreas de investigación de la Agencia Estatal de Investigación: <https://www.aei.gob.es/areas-tematicas/areas-tematicas>

B-0. Análisis y representación

Este bloque pretende contextualizar la situación y perfil del investigador, la provincia de residencia, la universidad, centro y área de investigación, la categoría laboral, el sexo y edad.

Provincia de residencia

Provincia residencia	Respuestas	Porcentaje %
Valencia	245	31,29
Madrid	117	14,94
Alicante	78	9,96
Granada	67	8,56
Cádiz	33	4,21
Navarra	29	3,70
Jaén	27	3,45
Salamanca	24	3,07
Barcelona	23	2,94
Castellón	17	2,17
Murcia	17	2,17
Zaragoza	13	1,66
Vizcaya	10	1,28
Sevilla	9	1,15
Las Palmas	7	0,89
Málaga	5	0,64
Asturias	4	0,51
Lleida	4	0,51
Álava	3	0,38
Almería	3	0,38
Burgos	3	0,38
Huelva	3	0,38
Melilla	3	0,38
Ávila	2	0,26
Badajoz	2	0,26
Ceuta	2	0,26
Ciudad Real	2	0,26
Cuenca	2	0,26
Guipúzcoa	2	0,26
Valladolid	2	0,26
Zamora	2	0,26
A Coruña	1	0,13
Baleares	1	0,13
Girona	1	0,13
Guadalajara	1	0,13
Santa Cruz de Tenerife	1	0,13
<b>Total</b>	<b>765</b>	<b>97,70</b>
<b>Omitidas</b>	<b>11</b>	<b>1,40</b>
<b>Descartadas<sup>71</sup></b>	<b>7</b>	<b>0,89</b>

Tabla 8: Provincia de residencia de los encuestados/as.

<sup>71</sup> Descartamos para el análisis de provincias las respuestas siguientes: Andalucía (1), Cataluña (1), Chile (1), Ecuador (1), Francisco Morazán (Honduras) (1), México y mmm (1).

Existe representación de 34 de las 50 provincias y de las dos ciudades autónomas. La provincia desde donde se han recibido más respuestas es desde Valencia, con 245, siendo el 31,29% de las respuestas totales, seguida de Madrid, 117 respuestas, siendo un 14,94%, más de la mitad.

Es lógico que Valencia sea con diferencia la ciudad que más respuestas ha recibido ya que el estudio se realiza en una universidad valenciana. Además, en parte es gracias a la difusión de la encuesta a través del correo institucional enviado desde la Conselleria de Participación, Transparència, Cooperación y Calidad Democrática a las universidades valencianas, colaborando también el Vicerrectorado de Investigación de la Universitat Politècnica de València que remitió la encuesta a todo el personal docente e investigador. Se obtuvieron 78 respuestas de la provincia de Alicante (un 9,96% de los encuestados, siendo la tercera provincia con más respuestas) y Castellón, con 17 respuestas siendo un 2,17% de las respuestas totales.

Destaca una alta participación de investigadores de distintas provincias, como Granada (67 respuestas), Cádiz (33), Comunidad Foral de Navarra (29), Jaén (27), Salamanca (24) y Barcelona (23) con más de 20 respuestas obtenidas. Finalmente, entre 10 y 20, tenemos las provincias de Castellón y Murcia (17), Zaragoza (13) y Vizcaya (10).

Las 23 ciudades<sup>72</sup> con menos de 10 respuestas suponen un 8,33% del total de las respuestas a la encuesta.

### Sexo

De las 783 respuestas, 443 han sido contestadas por hombres (un 56,58%), 327 han sido respondidas por mujeres (41,76%) y 3 por intersexuales (0,38%). Han omitido la respuesta 10 personas (1,28%).

Sexo	Respuestas	Porcentaje %
Femenino	327	41,76
Masculino	443	56,58
Intersexual	3	0,38
<b>Respondidas</b>	<b>773</b>	<b>98,72</b>
<b>Omitidas</b>	<b>10</b>	<b>1,28</b>

Tabla 9: Representación por sexo.

<sup>72</sup> Sevilla, Las Palmas, Málaga, Asturias, Lleida, Álava, Almería, Burgos, Huelva, Melilla, Ávila, Badajoz, Ceuta, Ciudad Real, Cuenca, Guipúzcoa, Valladolid, Zamora, A Coruña, Baleares, Girona, Guadalajara y Santa Cruz de Tenerife.

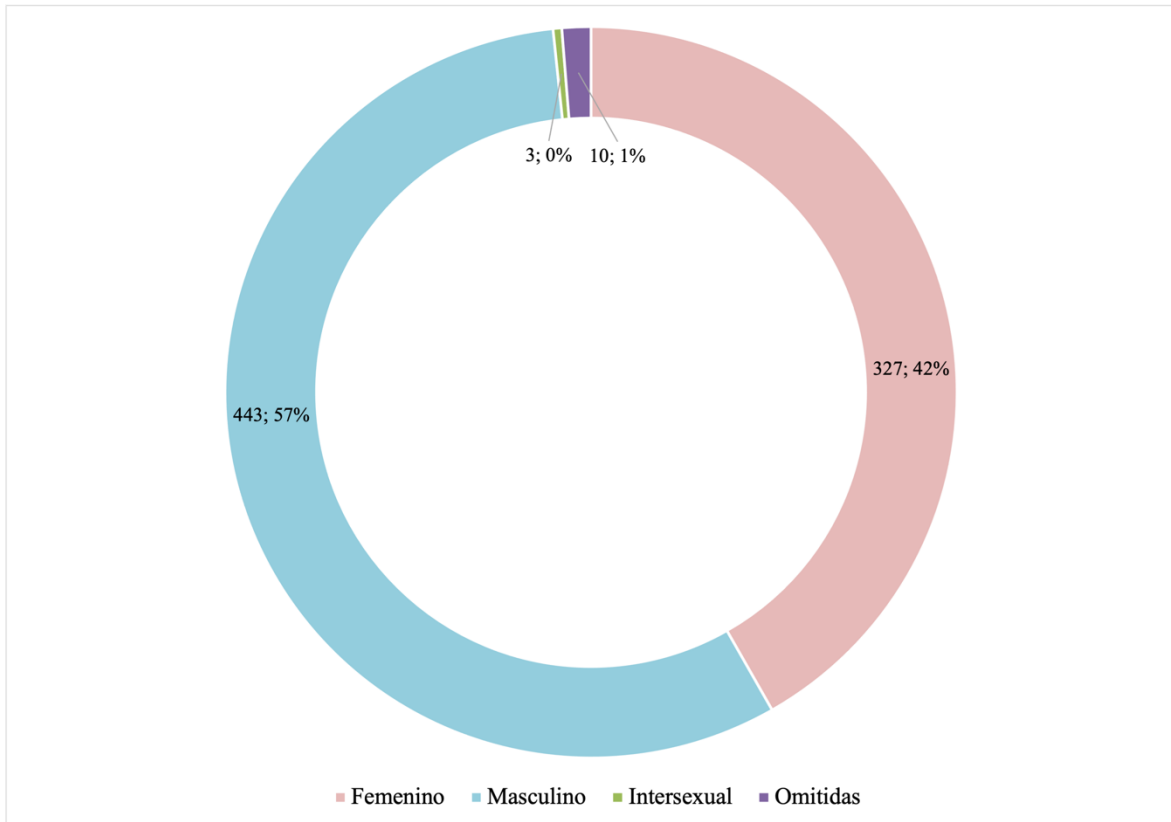


Figura 3: Distribución por sexo de los encuestados/as.

### Franjas de edad

La franja de edad que más ha contestado a la encuesta es la conformada por personas de entre 41 a 50 años, con 249 respuestas que suponen un 31,30% de las respuestas totales obtenidas, seguido de cerca por la franja conformada de 51 a 60 años, con 236 respuestas que suponen un 30,14%.

Con más de 100 respuestas de diferencia, sigue la franja de 31 a 40 años, 129, que son el 16,48% de las respuestas totales. El rango de edad que corresponde de 22 a 30 y de 61 a 65 reciben un número de respuestas similar: 68 y 66 respectivamente, siendo un 8,68% y 8,43 del total.

La franja de edad que menos ha contestado la encuesta es de más de 65 años, con 28 respuestas, 3,58% del total. Omiten la respuesta 7 personas (0,89%).

Rango de edad	Respuestas	Porcentaje %
22-30	68	8,68
31-40	129	16,48
41-50	249	31,80
51-60	236	30,14
61-65	66	8,43
Más de 65	28	3,58
<b>Respuestas</b>	<b>776</b>	<b>99,11</b>
<b>Omitidas</b>	<b>7</b>	<b>0,89</b>

Tabla 10: Franjas de edad de los encuestados/as.

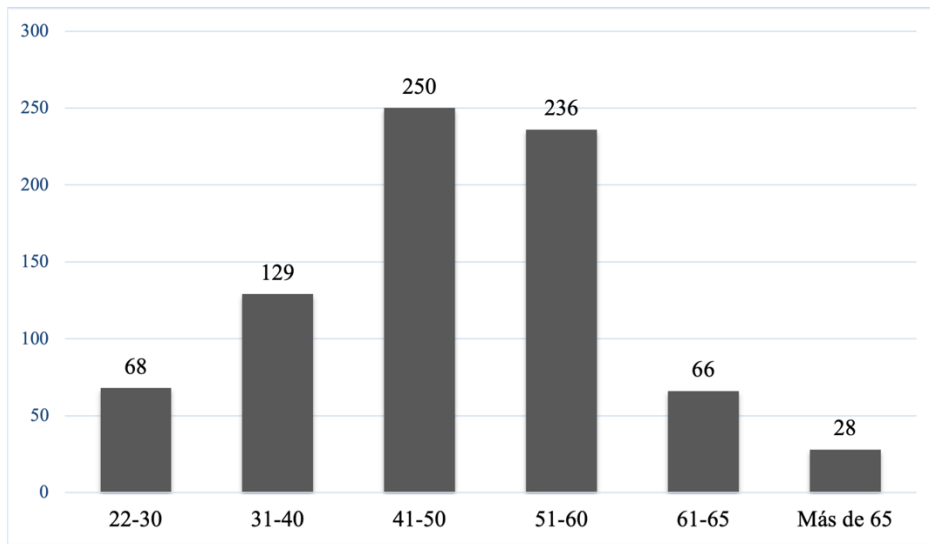


Figura 4: Franjas de edad de los encuestados/as.

### Institución

En cuanto a la institución a la que pertenecen los encuestados, se obtuvieron respuestas de 674 universidades, 80 respuestas de institutos y otros centros de investigación, 25 respuestas omitidas y 4 respuestas *otros* (sin información). En total son 47 universidades y 21 centros, fundaciones e institutos de investigación.

Se obtuvieron respuestas de representantes de 674 universidades, siendo el 86,08% del total de instituciones.



Universidad	Respuestas totales	Porcentaje %
Universitat Politècnica de València	132	16,86
Universidad de Granada	73	9,32
Universidad Complutense de Madrid	57	7,28
Universitat de València	53	6,77
Universidad de Cádiz	37	4,73
Universidad Carlos III de Madrid	29	3,70
Universidad Pública de Navarra	29	3,70
Universidad de Jaén	28	3,58
Universidad Miguel Hernández	27	3,45
Universidad de Salamanca	26	3,32
Universidad de Alicante	25	3,19
Universidad Autónoma de Madrid	13	1,66
Universidad del País Vasco	13	1,66
Universidad de Zaragoza	12	1,53
Universitat Jaume I	12	1,53
Universidad de Murcia	10	1,28
Universitat Autònoma de Barcelona	10	1,28
Universidad Politécnica de Madrid	8	1,02
Florida Universitaria	7	0,89
Universidad de Las Palmas de Gran Canaria	7	0,89
Universidad de Castilla-La Mancha	6	0,77
Universidad Politécnica de Cartagena	6	0,77
Universidad de Oviedo	5	0,64
Universitat de Lleida	5	0,64
Universidad de Málaga	4	0,51
Universidad Pablo de Olavide - Sevilla	4	0,51
Universitat Politècnica de Catalunya	4	0,51
Universidad de Burgos	3	0,38
Universidad Rey Juan Carlos	3	0,38
Universitat de Girona	3	0,38
Universidad Católica de Valencia San Vicente Mártir	2	0,26
Universidad de Almería	2	0,26
Universidad de Extremadura	2	0,26
Universidad de Huelva	2	0,26
Universidad de Sevilla	2	0,26
Universidad de Valladolid	2	0,26
Universidad Católica de Ávila	1	0,13
Universidad Católica San Antonio	1	0,13
Universidad Central del Ecuador	1	0,13
Universidad da Coruña	1	0,13
Universidad de La Laguna	1	0,13
Universidad Nacional Autónoma de Honduras	1	0,13
Universidad Pontificia de Salamanca	1	0,13
Universitat de Barcelona	1	0,13
Universitat de les Illes Balears	1	0,13
Universitat Oberta de Catalunya	1	0,13
Universitat Pompeu Fabra	1	0,13
Total investigadores procedentes de universidades españolas	674	86,08

Tabla 11: Universidades a las que pertenecen los encuestados/as.

La Universitat Politècnica de València es la más representada con 132 investigadores vinculados a esta universidad, siendo el 16,86% del total. En cuanto a universidades de la Comunidad Valenciana, le sigue la Universitat de València con 53 (cuarta con más representación, siendo un 6,77% del total), la Universidad Miguel Hernández de Elche con 27 (novena con más representación, siendo un 3,45% del total), Universidad de Alicante con 25 (undécima con más representación, siendo un 3,19%), Universitat Jaume I de Castellón, con 12 respuestas, siendo un 1,53%.

Se han recibido respuestas de 2 universidades privadas valencianas: 7 representantes de Florida Universitaria siendo un 0,89% y 2 Universidad Católica de Valencia San Vicente Mártir, siendo un 0,26%.

La segunda universidad con más representación es la Universidad de Granada, con 73 respuestas, (9,32% del total). La tercera es la Universidad Complutense de Madrid, con 73 respuestas (7,28%), seguido de la Universidad de Valencia con 53 (7,28%). Con 37 respuestas está la Universidad de Cádiz (4,73%).

Entre 20 y 30 respuestas hay representación de 6 universidades: Universidad Carlos III de Madrid y Universidad de Navarra, 29 respuestas (3,70%), 28 de la Universidad de Jaén (3,58%), 27 de la Universidad Miguel Hernández de Elche (3,45%), 26 de la Universidad de Salamanca (3,32%) y 25 de la Universidad de Alicante (3,19%).

Entre en rango de respuestas entre 10 y 20 respuestas hay 6 universidades: Universidad Autónoma de Madrid y Universidad del País Vasco con 13 (1,66% cada una), Universidad de Zaragoza y Universitat Jaume I, con 12 (1,53% cada una) y finalmente, Universidad de Murcia y Universitat Autònoma de Barcelona, con 10 (1,28%).

El resto de las universidades con menos de 10 respuestas corresponden al 11,28% del total de respuestas.

Se obtuvieron 80 respuestas de representantes de los siguientes centros, institutos y otras fundaciones de investigación. Suponen el 10,25% del total. Destaca el Instituto de Investigación Sanitaria de la Fe con 11 respuestas (1,40%), la Fundación de Centro de Estudios Ambientales del Mediterráneo con 10 respuestas (1,28%) y el Instituto Valenciano de Investigaciones Agrarias con 9 respuestas (1,15%).

Centro, instituto, fundación de investigación	Respuestas totales	Porcentaje %
Instituto de Investigación Sanitaria La Fe	11	1,40
Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM)	10	1,28
Instituto Valenciano de Investigaciones Agrarias (IVIA)	9	1,15
ISABIAL - Instituto de Investigación Sanitaria y Biomédica de Alicante	8	1,02
Instituto de Agroquímica y Tecnología de Alimentos (IATA-CSIC)	7	0,89
Centro de Investigación Príncipe Felipe (CIPF)	6	0,77
Instituto de Física Corpuscular (IFIC-CSIC)	6	0,77
Consejo Superior de Investigaciones Científicas (CSIC)	5	0,64
Hospital General Universitario de Alicante	3	0,38
Fundació de la Comunitat Valenciana (FCV)	2	0,26
Instituto Tecnológico de la Alimentación (AINIA)	2	0,26
Instituto Tecnológico de la Energía (ITE)	2	0,26
Agencia de Medios	1	0,13
CERN - Organización Europea para la Investigación Nuclear	1	0,13
CIC bioGUNE - Centro de Investigación Cooperativa en Biociencias	1	0,13
Fundació Docència i Recerca Mútua Terrassa	1	0,13
Hospital del Mar - Valencia	1	0,13
IES Haygón	1	0,13
Instituto de Investigación en Inteligencia Artificial (IIIA-CSIC)	1	0,13
Instituto LRKlein	1	0,13
Instituto Universitario Menéndez Pidal (UCM)	1	0,13
Total investigadores procedentes de centros e institutos de investigación	80	10,25

Tabla 12: Centros de investigación a los que pertenecen los encuestados/as.

### Categoría laboral

Esta pregunta sobre la categoría laboral a la que pertenecen fue respondida por 772 investigadores y omitida por 11.

Categoría laboral	Respuestas	Porcentaje %
Profesor/a Titular de Universidad	209	26,69
Catedrático/a de Universidad	150	19,16
Profesor/a Contratado/a Doctor	76	9,71
Doctorando/a	62	7,92
Profesor/a Asociado/a	61	7,79
Profesor/a Ayudante Doctor	50	6,39
Investigador/a Científico	50	6,39
Doctor/a	26	3,32
Técnico de investigación	22	2,81
Profesor/a de Investigación	10	1,28
Profesor/a Sustituto/a Interino/a	10	1,28
Personal de Administración y Servicios <sup>73</sup>	9	1,15
Científico/a Titular	8	1,02
Profesor/a Ayudante	7	0,89
Profesor/a visitante	6	0,77
Empresa <sup>74</sup>	4	0,51
Profesor/a Colaborador/a	3	0,38

<sup>73</sup> PAS: Bibliotecarios, documentalistas, técnicos auxiliares.

<sup>74</sup> Empresa: Incluye periodistas.

Profesor/a Agregado/a	2	0,26
Profesor/a Titular de Escuela Universitaria	2	0,26
Catedrático/a de Escuela Universitaria	1	0,13
CIO	1	0,13
Formación Profesorado Universitario	1	0,13
PCI	1	0,13
Profesor/a Sustituto	1	0,13
	Respuestas	772
	Omitidas	11
		98,60
		1,40

Tabla 13: Categoría laboral de los encuestados/as.

Con 209 respuestas, siendo más de una cuarta parte de los encuestados son Profesores/as Titulares de Universidad (26,69% del total). Destaca también el alto volumen de respuestas de Catedráticos/as de Universidad, con 150 respuestas (19,16%).

En tercer lugar, la categoría laboral más representativa es la de Profesor/a Contratado/a Doctor/a con 76 respuestas (9,71%).

Entre 50 y 65, destacan las 62 respuestas obtenidas de doctorandos/as (7,92%), 61 de profesores/as asociados/as (7,79%) y con el mismo volumen de respuestas, 50, profesores/as ayudantes doctores/as e investigadores/as científicos/as (6,39%).

Con 10 o más respuestas, sin llegar a 30, se han obtenido 26 respuestas de doctores/as (3,32%), 22 de técnicos de investigación (2,81%) y 10 de profesores/as de investigación y profesores/as sustitutos/as interinos/as (1,28%).

Con representación de menos de 10 respuestas hay 13 categorías laborales, con un volumen de 46 respuestas que son el 5,89% del total. Estas son: personal de administración y servicios, científicos/as titulares, profesores/as ayudantes, visitantes, colaboradores/as, agregados/as y titulares de escuela universitaria, y profesionales vinculados a la empresa. Se ha omitido la pregunta en 11 ocasiones (1,40%).

### Áreas de investigación

Las áreas de investigación se han extraído de la Agencia Estatal de Investigación (AEI), del Ministerio de Ciencia e Innovación<sup>75</sup>. La Agencia Estatal de Investigación tienen como misión el fomento de la investigación científica y técnica en todas las áreas del saber mediante la asignación competitiva y eficiente de los recursos públicos, así como el seguimiento de actuaciones financiadas, impacto y asesoramiento en la planificación de acciones o iniciativas a través de las políticas de I+D de la Administración General del Estado (Agencia Estatal de Investigación, 2017).

<sup>75</sup> Áreas de investigación de la Agencia Estatal de Investigación del Ministerio de Ciencia e Innovación: <https://www.aei.gob.es/areas-tematicas/areas-tematicas>

Esta pregunta la han respondido 775 investigadores (el 98,98%) y la han omitido 8 (1,02%).

Área de investigación	Total	Porcentaje %
Área 1: Ciencias Sociales	130	16,60
Área 11: PIN / Producción industrial, ingeniería civil e ingenierías para la sociedad	63	8,05
Área 12: TIC / Tecnologías de la información y de las comunicaciones	63	8,05
Área 3: Economía	61	7,79
Área 19: BME / Biomedicina	59	7,54
Área 5: FLA / Cultura: Filosofía, literatura y arte	42	5,36
Área 9: MTM / Ciencias matemáticas	39	4,98
Área 7: EDU / Ciencias de la educación	38	4,85
Área 16: CTM / Ciencias y tecnologías medioambientales	37	4,73
Área 17: CAA / Ciencias agraria y agroalimentarias	36	4,60
Área 18: BIO / Biociencias y biotecnología	36	4,60
Área 2: Derecho	33	4,21
Área 10: FIS / Ciencias físicas	32	4,09
Área 14: CTQ / Ciencias y tecnologías químicas	31	3,96
Área 8: PSI / Psicología	20	2,55
Área 15: MAT / Ciencias y tecnologías de materiales	17	2,17
Área 6: PHA / Estudios del pasado: Historia y arqueología	16	2,04
Área 4: MLP / Mente, lenguaje y pensamiento	14	1,79
Área 13: EYT / Energía y transporte	8	1,02
Respuestas	775	98,98
Omitidas	8	1,02

Tabla 14: Áreas de investigación.

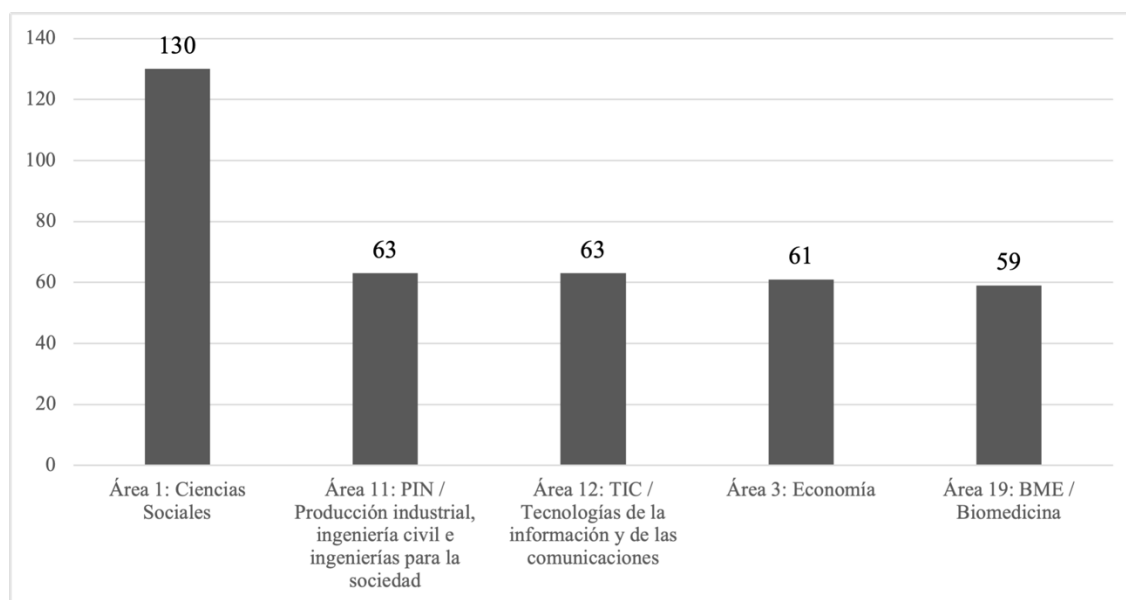


Figura 5: Áreas de investigación de los encuestados/as (+ de 50 elecciones).

En la figura 5 se muestran aquellas áreas de investigación que han sido elegidas en más de 50 ocasiones. Con 130 investigadores, el área dominante en la encuesta es la 1, de Ciencias Sociales (16,60%). Con la mitad de respuestas que el área 1, las áreas 11 y 12, de Producción

industrial, ingeniería civil e ingenierías para la sociedad y Tecnologías de la información y de las comunicaciones respectivamente, son las siguientes con más representación con 63 (8,05%), seguido del área 3, de Economía con 61 (7,79%), el área 19, de Biomedicina, con 59 (7,54%).

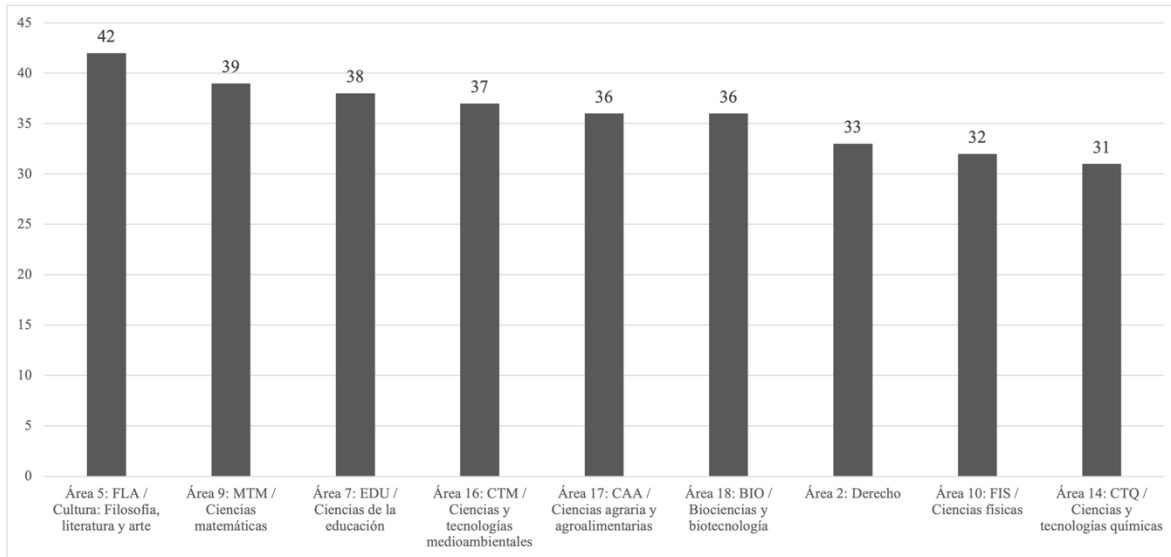


Figura 6: Áreas de investigación (entre 50 y 25 elecciones).

Como se muestra en la figura 6, entre 50 y 25 respuestas hay un total de 9 áreas: el área 5, de Cultura, Filosofía, Literatura y Arte, es elegida un total de 42 veces (5,36%). Le sigue el Área 9, de Ciencias Matemáticas, con 39 respuestas (4,98%), Área 7, de Ciencias de la Educación, con 38 (4,85%), Área 16 de Ciencias y Tecnologías Medioambientales, con 37 (4,73%), Área 17, de Ciencias agraria y agroalimentarias y Área 18, de Biociencias y Biotecnología, con 36 respuestas respectivamente (4,60%), Área 2, de Derecho, con 33 respuestas (4,21%), Área 10, de Ciencias Físicas, con 32 (4,09%) y Área 14, de Ciencias y Tecnologías Químicas, con 31 respuestas (3,96%).

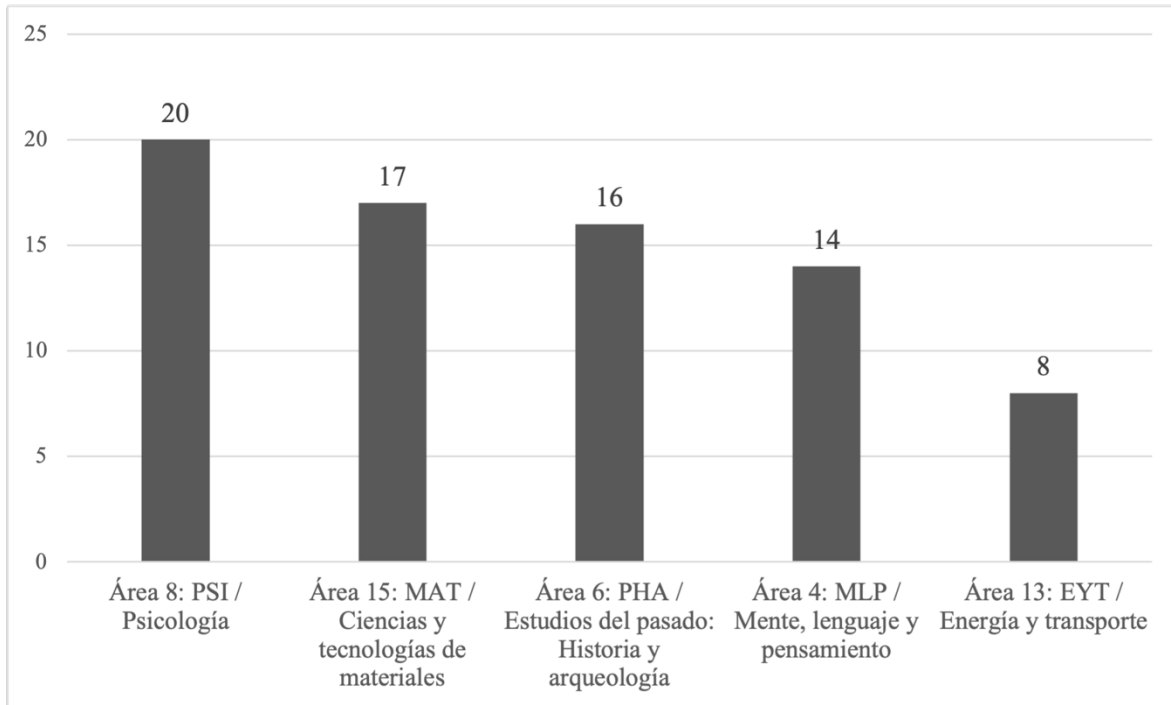


Figura 7: Áreas de investigación (menos de 25 elecciones).

La figura 7 muestra las áreas de investigación con menos de 25 elecciones. Hay 5 áreas de investigación. El área de 8, de Psicología ha sido elegida en 20 ocasiones (2,55%). Le siguen las áreas 15, de Ciencias y Tecnologías Materiales (17, 2,17%), Estudios del pasado: Historia y Arqueología (16, 2,04%) y el área 4, de Mente, Lenguaje y Pensamiento (14, 1,79%). Finalmente, con menos de 10 elecciones, el área 13, con 8 siendo el 1,02%.

## Bloque 1. Contexto inicial: Conocimiento de los Datos Abiertos

### B-1. Estandarización y limpieza de datos

El bloque 1 está formado por 6 preguntas. Para la representación de los datos y creación de las visualizaciones se han estandarizado las respuestas de las preguntas 3, 4, 5 y 6.

La pregunta 3 es de respuesta abierta y se contesta si se ha respondido negativamente a la pregunta 1<sup>76</sup>. En base a las respuestas obtenidas, se han creado las siguientes categorías clasificando así las respuestas obtenidas:

---

Desconocimiento de fuentes
Desconocimiento de uso y potencial de reutilización
Falta de difusión y publicidad activa
Falta de información
No necesidad

---

Tabla 15: Categorías de la pregunta 3 de la encuesta.

En las preguntas 4, 5 y 6 se plantean opciones donde el encuestado puede elegir varias de las afirmaciones del enunciado. Estos son demasiado largos para su posterior representación en visualizaciones y por tanto se normalizan como se describe a continuación.

La pregunta 4 se responde en caso de que se haya respondido negativamente a la pregunta 2<sup>77</sup>. Se plantearon 5 afirmaciones:

- Falta de información interesante y reutilizable para mis investigaciones/proyectos.
- Conozco el movimiento de los datos abiertos, pero no sé cómo reutilizar los datos.
- Falta de estandarización de los datos (mismos tipos/conjuntos de datos en diferentes portales que vienen de maneras distintas. Falta de interoperabilidad entre administraciones).
- Falta de fiabilidad de los datos.
- Falta de actualización de los datos.

Para su posterior representación, se normalizaron de la siguiente forma:

---

<sup>76</sup> Pregunta 1. ¿Sabe qué son los datos abiertos? (Entendidos como filosofía que persigue que determinados tipos de datos estén disponibles de forma libre sin restricciones, en bruto, estructurados y en formatos que faciliten la reutilización).

<sup>77</sup> Pregunta 2. ¿Ha reutilizado datos/información del sector público para el desarrollo de investigaciones, proyectos o servicios en su actividad profesional?



---

Falta de datos de interés
Desconocimiento de la reutilización
Falta de estandarización
Falta de fiabilidad de los datos
Falta de actualización de los datos

---

Tabla 16: Categorías de la pregunta 4 de la encuesta.

La pregunta 5 pretende conocer el modo en que están involucrados los investigadores en los datos abiertos. Se plantearon una serie de roles:

- Usuario de datos (para informarme).
- Usuario de datos para producir servicios, productos o análisis propios (Reutilizador).
- Proveedor de datos (en portales de datos abiertos o repositorios de investigación).
- Administrador de datos.
- Ninguna de las anteriores.

Se estandarizaron los enunciados de la pregunta 5 del siguiente modo:

---

Usuario de datos para información
Usuario de datos para producir servicios y análisis (reutilizador)
Proveedor de datos
Administrador de datos
Ninguna de las anteriores

---

Tabla 17: Categorías y roles de la pregunta 5 de la encuesta.

La pregunta 6 se plantea para aquellos que responden positivamente a la pregunta 1<sup>78</sup>, de conocimiento de los datos abiertos. Busca conocer el primer contacto de los investigadores con los datos abiertos:

- A través de artículos de investigación.
- Medios de comunicación convencionales.
- En talleres de formación/seminarios/congresos.
- Canales de comunicación y/o redes sociales de las administraciones públicas y otros proveedores de datos.
- Redes sociales.
- A través de reutilizadores.

---

<sup>78</sup> Pregunta 1. ¿Sabe qué son los datos abiertos? (Entendidos como filosofía que persigue que determinados tipos de datos estén disponibles de forma libre sin restricciones, en bruto, estructurados y en formatos que faciliten la reutilización).

Se resumen en las siguientes categorías para su posterior representación:

---

A través de artículos de investigación
Medios de comunicación convencionales
Talleres de formación/seminarios/congresos
Canales de comunicación y/o RRSS de AAPP y otros
proveedores de datos
A través de reutilizadores

---

Tabla 18: Categorías de la pregunta 6 de la encuesta.

Además, las preguntas 4, 5 y 6 tienen una opción de respuesta abierta. En el caso que sea posible en base a las respuestas de los encuestados, se engloban dentro de las categorías creadas y si no, en el apartado de resultados se hace mención sobre las respuestas obtenidas y las nuevas categorías generadas en caso necesario.

#### B-1. Análisis y representación

El bloque 1 es el bloque de contexto inicial, donde se busca el motivo del conocimiento y desconocimiento de lo que son los datos abiertos y su movimiento. La única pregunta obligatoria de la encuesta es la primera que se plantea: ¿Sabe lo que son los datos abiertos?

En caso de que se responda negativamente a la pregunta 1<sup>79</sup>, la encuesta se cierra al terminar el primer bloque, dándose por finalizada. Las siguientes cuestiones son de carácter técnico, de conocimiento y uso de los datos abiertos y, por tanto, los investigadores no van a poder aportar su visión.

La primera pregunta es pretende obtener respuestas sobre el conocimiento de los datos abiertos.

Un total de 621 investigadores saben qué son los datos abiertos (79,31%) por 57 investigadores que desconocen los datos abiertos y sus potencialidades (7,28%). 105 investigadores omiten la pregunta (13,41%).

Respuesta	Total	%
Sí	621	79,31%
No	57	7,28%
Omitidas	105	13,41%

Tabla 19: Conocimiento de los datos abiertos.

---

<sup>79</sup> Pregunta 1. ¿Sabe qué son los datos abiertos? (Entendidos como filosofía que persigue que determinados tipos de datos estén disponibles de forma libre sin restricciones, en bruto, estructurados y en formatos que faciliten la reutilización).

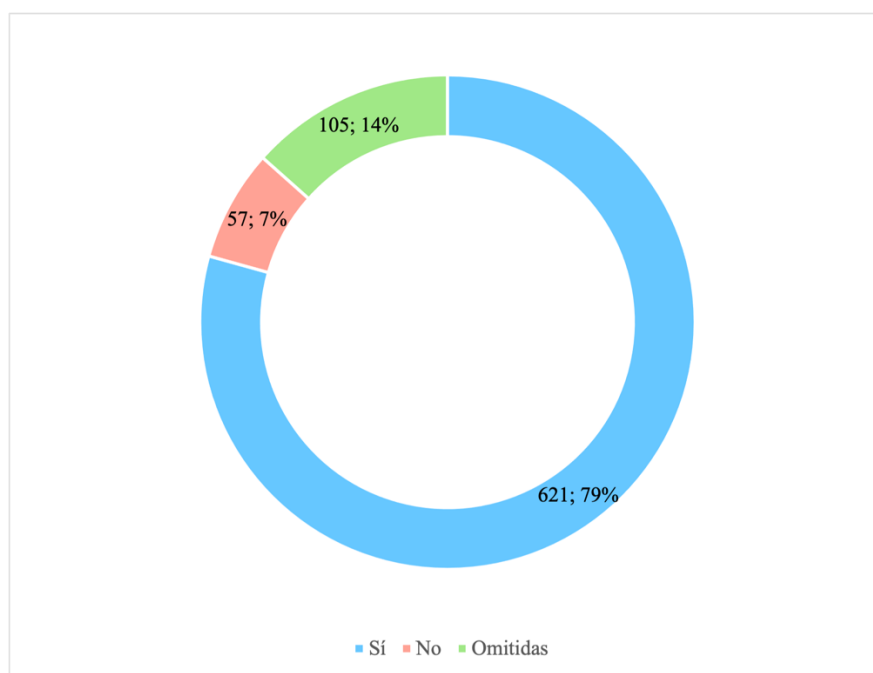


Figura 8: Conocimiento de los datos abiertos.

Aquellos encuestados que contestan que no saben qué son los datos abiertos, finalizan la encuesta en la última pregunta de este bloque.

De las 621 respuestas que afirman conocer qué son los datos abiertos, 362 son hombres y 256 son mujeres, mientras que aquellos que desconocen el movimiento, 32 son hombres y 25 son mujeres.

La segunda pregunta se centra en aquellos que han respondido positivamente en la primera y busca conocer cuántos de ellos han reutilizado información del sector público.

Más de la mitad de los encuestados, 396, han reutilizado información del sector público para sus investigaciones (50,57%). No han reutilizado datos abiertos 274 investigadores (34,99%) y 113 investigadores han omitido la respuesta (14,43%).

Respuesta	Total	Porcentaje %
Sí	396	50,57
No	274	34,99
Omitidas	113	14,43

Tabla 20: Reutilización de datos abiertos en investigación y proyectos.

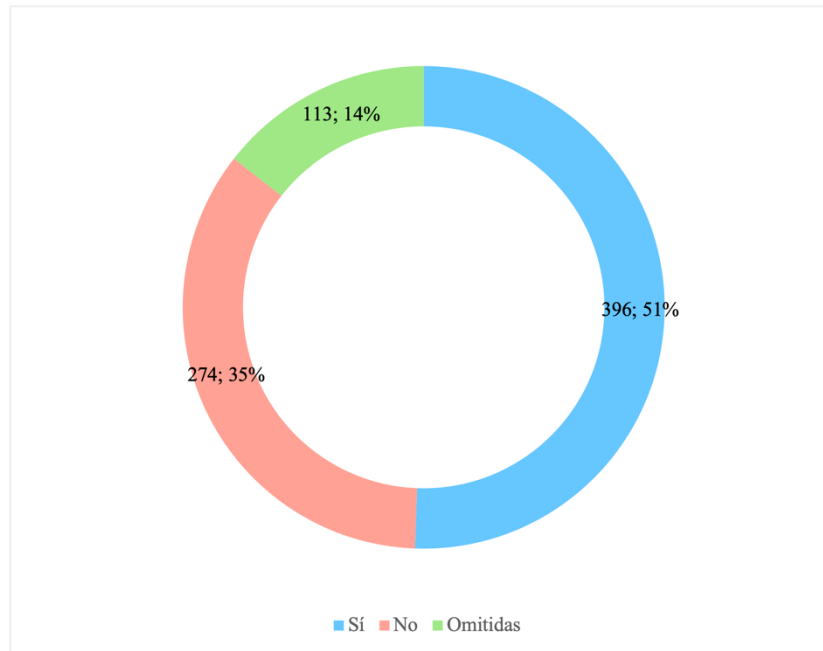


Figura 9: Reutilización de datos abiertos en investigación y proyectos.

De los 396 que afirman haber reutilizado datos abiertos en sus proyectos, 234 son hombres (59,10%) y 162 son mujeres (40,90%). Aquellos que señalan no haber reutilizado datos abiertos, 156 son hombres (56,93%) y 118 mujeres (43,07%).

La tercera pregunta busca indagar acerca del motivo de desconocimiento de los datos abiertos y es de respuesta abierta.

A pesar de que en el enunciado de la pregunta se especifica que solo la deben responder quienes han respondido “no” en la pregunta 1<sup>80</sup> (un total de 57 encuestados responden “no”), se responde en 168 ocasiones.

Un total de 90 respuestas no son válidas debido a que contestan que sí que conocen los datos abiertos, siendo una pregunta planteada a quienes no los conocen.

Se han detectado ciertas similitudes en las respuestas obtenidas. Esto ha permitido normalizar y clasificar las respuestas en las 5 categorías que tenemos en la siguiente tabla.

<sup>80</sup> Pregunta 1. ¿Sabe qué son los datos abiertos? (Entendidos como filosofía que persigue que determinados tipos de datos estén disponibles de forma libre sin restricciones, en bruto, estructurados y en formatos que faciliten la reutilización).

Opciones	Respuestas
Desconocimiento de fuentes	10
Desconocimiento de uso y potencial de reutilización	34
Falta de difusión y publicidad activa	15
Falta de información	14
No necesidad	5
Total de respuestas válidas	
	78
Respuestas no válidas	
	90
Total respuestas	
	168

Tabla 21: Motivo del desconocimiento de los datos abiertos.

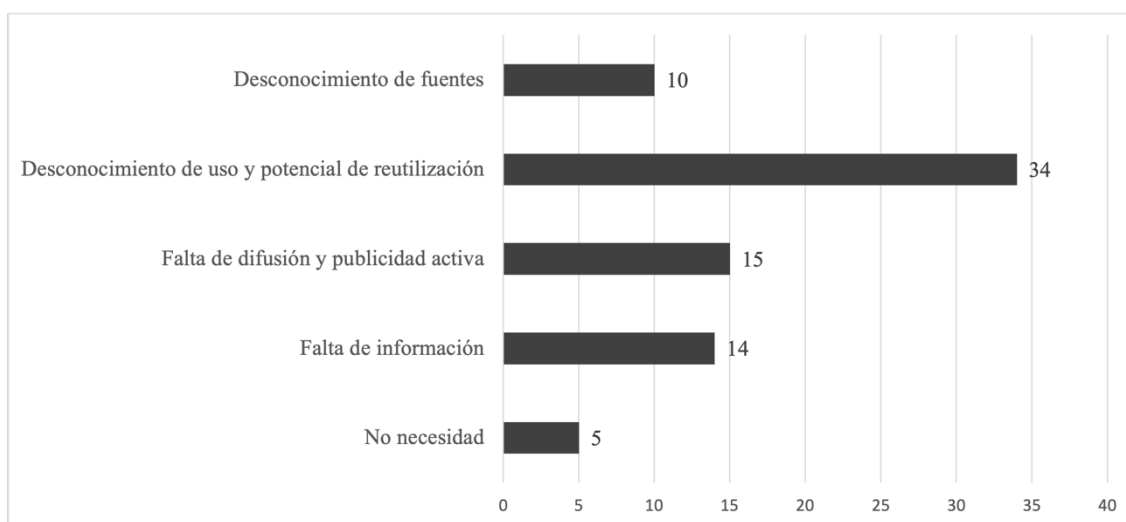


Figura 10: Motivo del desconocimiento de los datos abiertos.

De las 78 respuestas, 34 responden que desconocen el uso y el potencial de los datos abiertos, es decir, no saben que existe información potencialmente reutilizable en portales de la administración pública.

Otras 15 respuestas hacen referencia a la falta de difusión y publicidad activa. Las propias administraciones, encargadas de la apertura de datos a través de portales, son quienes deben informar a la ciudadanía a través de los canales de comunicación actuales, de la existencia de estos portales de información potencialmente reutilizables para la creación de productos, servicios o, para el ámbito académico, educativo y de investigación.

De la mano de la publicidad activa desde las instituciones, los encuestados hacen referencia en las respuestas a la falta de información (14) y al desconocimiento de fuentes de información potencialmente reutilizables (10).

Finalmente, 5 investigadores especifican no tener la necesidad de reutilizar información del sector público.

La cuarta pregunta busca conocer los posibles motivos de no reutilización de datos abiertos. La deben responder quienes han respondido “sí” a la pregunta de conocimiento de datos abiertos<sup>81</sup>. La contestan 359 investigadores. Se plantea con 5 afirmaciones en las que el usuario puede elegir varias opciones (es decir, se van a obtener más de 359 elecciones ya que algunos investigadores seleccionan más de una opción). Además, se plantea una opción para que puedan responder abiertamente.

De las 88 respuestas obtenidas en la opción abierta, se han descartado 25 debido a que o bien no han aportado ninguna información o la que han aportado ha sido que sí han reutilizado o conocen los datos abiertos. Algunas de las respuestas obtenidas en la opción abierta caben dentro de las afirmaciones planteadas en la encuesta. Tras el proceso de normalización y análisis, se obtienen las siguientes respuestas:

Opciones	Respuestas
Falta de datos de interés	138
Desconocimiento de la reutilización	94
Falta de estandarización	82
Falta de fiabilidad de los datos	35
Falta de actualización de los datos	34

Tabla 22: Principales motivos de no reutilización de datos abiertos.

Los investigadores coinciden 138 veces en la falta de datos de interés para sus investigaciones y proyectos. En 92 ocasiones hacen referencia al desconocimiento de reutilización de datos abiertos. Con 82 elecciones, los investigadores coinciden en la importancia de la estandarización de datos y poder trabajar con ellos. Finalmente, 35 investigadores han seleccionado la opción de la falta de fiabilidad de la información y 34 no reutilizan datos del sector público debido a la falta de actualización de los mismos.

<sup>81</sup> Pregunta 1. ¿Sabe qué son los datos abiertos? (Entendidos como filosofía que persigue que determinados tipos de datos estén disponibles de forma libre sin restricciones, en bruto, estructurados y en formatos que faciliten la reutilización).

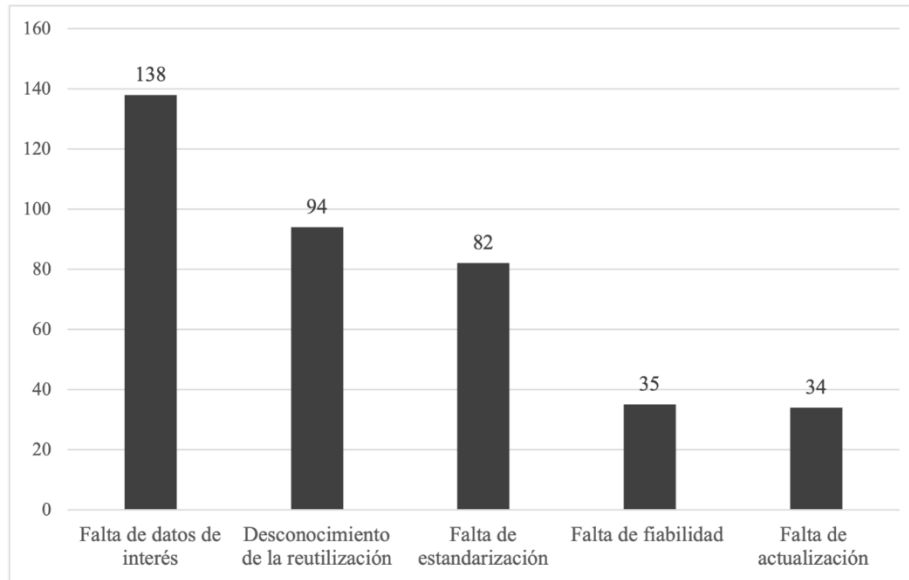


Figura 11: Principales motivos de no reutilización de datos abiertos.

En la opción otros, tras la limpieza de los datos, se detecta que 33 investigadores hacen referencia a que no necesitan de los datos abiertos en sus investigaciones. Se da el caso en una ocasión que, al ser profesor asociado, solo imparte docencia y por tanto no necesita datos. Además, 2 encuestados hacen referencia a la falta de tiempo y costumbre. En 7 ocasiones, dicen no utilizar datos abiertos para sus investigaciones y 8 hacen referencia a que no utilizan datos debido a la dificultad de accesibilidad a los mismos.

Se han contestado en 10 ocasiones afirmaciones que se engloban dentro de la afirmación de desconocimiento de la reutilización y 2 que se clasifican dentro de la afirmación de falta de datos de interés.

La quinta pregunta busca conocer el modo en el que se encuentran involucrados los investigadores con los datos abiertos.

La pregunta 5 es respondida por 566 investigadores. Busca conocer de qué modo se encuentran involucrados con los datos abiertos los investigadores. Se plantea con 5 opciones, más una opción de respuesta abierta, en las que el investigador puede elegir varias opciones. Tras el proceso de estandarización y limpieza de datos, las respuestas obtenidas son:

Opciones	Respuestas
Usuario de datos para información	304
Usuario de datos para producir servicios y análisis (reutilizador)	244
Proveedor de datos	150
Ninguna de las anteriores	89
Otros	19
Administrador de datos	6

Tabla 23: Rol del investigador en los datos abiertos.

Con 304 elecciones, los investigadores utilizan datos para informarse. 244 son reutilizadores, es decir, desarrollan nuevos productos y análisis con información del sector público. 150 son proveedores de datos. En 89 ocasiones, ninguna de las acciones planteadas está de acuerdo con lo que hacen los investigadores y en 6 ocasiones los investigadores son administradores de datos.

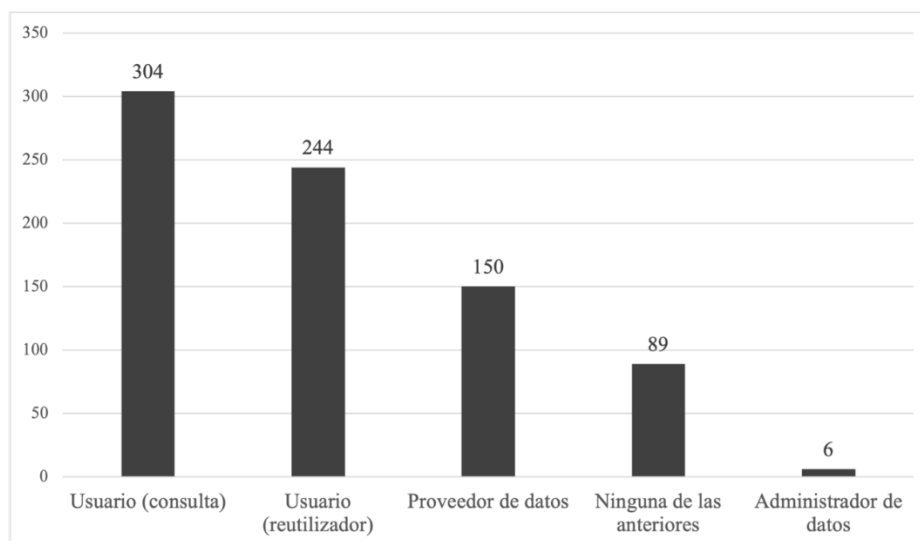


Figura 12: Rol del investigador en los datos abiertos.

Además, 19 responden investigadores responden la opción abierta. De las 19 respuestas, 6 inciden en el desconocimiento y no utilización de datos abiertos. En 13 casos, los investigadores están involucrados con los datos abiertos para: asesoría, formación, control de calidad y análisis, datos para la investigación, divulgación y educación, y estadística.

La sexta pregunta pretende conocer cómo llegan a conocer los datos abiertos y su movimiento los investigadores.

La pregunta 6 está planteada para que la respondan aquellos que respondieron afirmativamente en la pregunta 1<sup>82</sup>, y es acerca de cómo conocieron los datos abiertos. La responden 586 investigadores. Se plantean 5 afirmaciones más una opción de respuesta abierta. Algunos investigadores señalan más de una afirmación de las planteadas.

<sup>82</sup> Pregunta 1. ¿Sabe qué son los datos abiertos? (Entendidos como filosofía que persigue que determinados tipos de datos estén disponibles de forma libre sin restricciones, en bruto, estructurados y en formatos que faciliten la reutilización).



Opciones	Respuestas
A través de artículos de investigación	321
Canales de comunicación y/o RRSS de AAPP y otros proveedores de datos	159
Talleres de formación/seminarios/congresos	159
Medios de comunicación convencionales	88
A través de reutilizadores	85
Otros	67

Tabla 24: Cómo conocen los datos abiertos los investigadores/as.

Un total de 321 investigadores afirman haber conocido el movimiento de los datos abiertos a través de artículos de investigación. Con 159 elecciones, siendo un conocieron los datos abiertos a través de canales de comunicación y redes sociales de las administraciones públicas, al igual que los 159 que los conocieron en talleres de formación, seminarios y congresos. Les sigue con 88 respuestas, a través de medios de comunicación convencionales y 85 a través de otros reutilizadores.

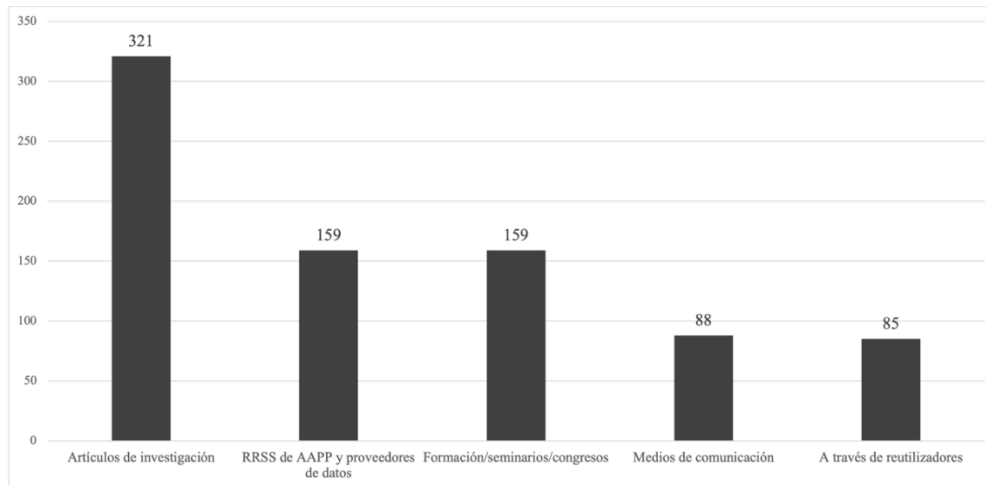


Figura 13: Cómo conocen los datos abiertos los investigadores/as.

Esta pregunta tiene una opción de respuesta abierta, obteniéndose 67 respuestas. 8 investigadores afirman no conocerlo (no deberían haber contestado) y 2 se pueden clasificar dentro de la categoría anterior de “talleres de formación/seminarios/congresos”.

Opciones	Respuestas
Redes de investigación y otros reutilizadores	19
Instituciones encargadas de la apertura de datos	10
De forma autónoma a través de internet	9
Convocatorias y proyectos de investigación	9
Investigación y requerimientos de las revistas científicas	5
Comisión europea	3
Encuestas	1
Otras	1

Tabla 25: Cómo conocen los datos abiertos (opción de respuesta abierta).

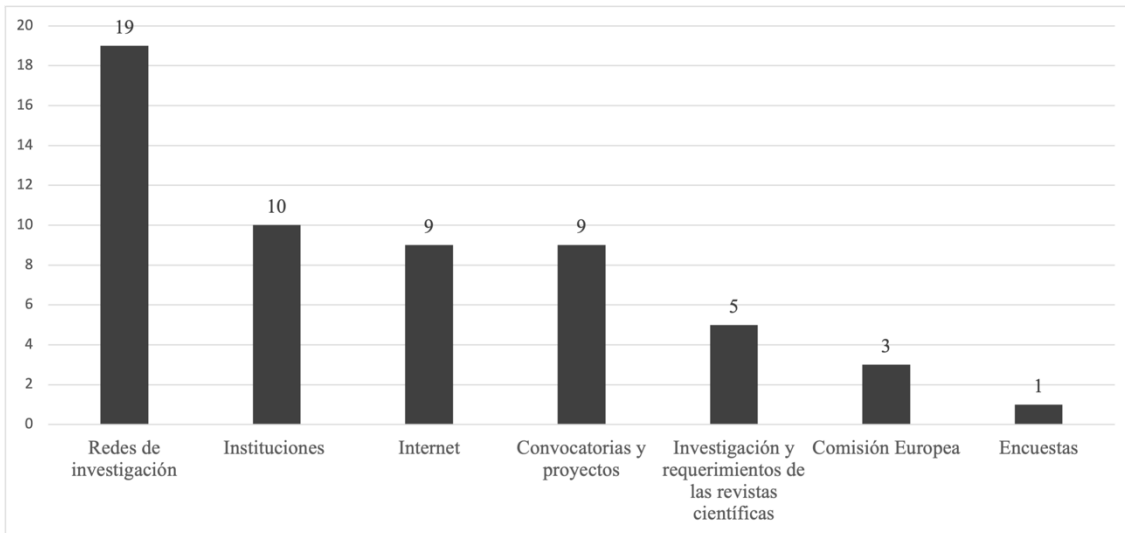


Figura 14: Cómo conocen los datos abiertos los investigadores/as.

En esta categoría, 19 afirman haberlos conocido a través de redes de investigación y otros reutilizadores. 10 a través de las propias instituciones encargadas de la apertura de los datos. 9 de forma autónoma gracias a internet. 9 gracias a los proyectos de investigación y sus respectivas convocatorias. 5 debido a la investigación y requerimientos de las propias revistas científicas. 3 a través de la Comisión Europea. 2 respuestas obtenidas las englobamos en la categoría de talleres de formación, seminarios y congresos ya que hacen referencia a su conocimiento a través de cursos. 1 hace referencia a que los conoció a través de una encuesta y finalmente, otro afirma que ya los conocía debido a que es una idea que lleva circulando tiempo dentro del mundo de la investigación.

Bloque 2. Importancia de la apertura de datos

B-2. Estandarización y limpieza de datos

El bloque 2 hace referencia al uso de los datos y la reutilización de la información del sector público. Este bloque está formado por 4 preguntas, de la 7 a la 10. En las preguntas 7, 8 y 9, los encuestados deben evaluar con una escala de valor del 0 al 4, la importancia a las cuestiones que se plantean, en materia de atributos de los datos abiertos, servicios y funcionalidades que esperan en los portales de datos abiertos y las barreras encontradas al tratar de acceder a los datos. En caso de no conocer la casuística planteada, se puede dejar en blanco y pasar a la siguiente.

La escala de valoración en las preguntas 7 y 8 es:

- Muy importante (4); Importante (3); Poco importante (2); Nada importante (1); No conozco (0).

Las afirmaciones planteadas en el enunciado de la pregunta 7 son:

- Suministro de datos sin restricciones.
- Tipo de licencia de los datos para su reutilización.
- Que contenga metadatos (completos, estructurados y útiles para interpretarlos).
- Que sea fácil encontrarlos y conseguirlos (accesibilidad).
- Que estén estandarizados (formato, variables...).
- Posibilidad de restricción a algunos tipos de datos (datos sensibles).
- Datos a tiempo real.
- Datos georreferenciados.

Para la posterior creación de visualizaciones, las afirmaciones se codifican en los conceptos la tabla siguiente:

---

Suministro
Licencias datos
Metadatos
Accesibilidad
Formatos
Restricciones
Actualización
Geolocalización

---

Tabla 26: Categorías de la pregunta 7 de la encuesta.

La pregunta 8 plantea las 7 afirmaciones siguientes:

- Los datos tienen URI permanentes.
- Enlaces a aplicaciones relacionadas con el conjunto de datos.
- Herramientas para el procesamiento y análisis de datos.
- Visualizaciones de los datos.
- Canales de contacto o formularios para reutilizadores.
- Noticias o novedades relacionadas con los datos abiertos y sistemas de alerta de actualización del catálogo de datos.
- Disponibilidad de API.

Para la posterior creación de la visualización, se codifican en:

URI
Apps relacionadas
Herramientas de procesamiento
Visualizaciones
Contacto
Noticias
API

Tabla 27: Categorías de la pregunta 8 de la encuesta.

En el caso de la pregunta 9, la escala de valoración es:

- Siempre (4); Muchas veces (3); Algunas veces (2); Nunca (1); No conozco (0).

Las afirmaciones planteadas en la pregunta 9 son:

- Mucho tiempo transcurrido entre la solicitud y la recepción de datos.
- La necesidad de registro para acceder a los datos.
- Acceder y comprender las condiciones de uso y de licencias y como interpretar y reutilizar los datos.
- Grados variables de calidad de los datos en conjuntos de datos diferentes, falta de estandarización en los formatos de los datos, información completa o de poco interés.
- Pago por los datos.
- Falta de interoperabilidad o de disponibilidad de los datos en otras AAPP / Dificultad para compararlos con otros datos.
- Falta de actualización de los datos.
- Dificultad de acceso.

Para la posterior creación de la visualización, se codifica en:

Tiempo  
 Registro  
 Comprensión  
 Calidad y estandarización  
 Pago  
 Interoperabilidad  
 Falta de actualización de datos  
 Acceso

Tabla 28: Categorías de la pregunta 9 de la encuesta.

Para concluir este segundo bloque, la pregunta 10 plantea una serie de tipos de datos para conocer los que son de más interés para los investigadores encuestados, con una opción de respuesta abierta normalizada (Ver en la pregunta 10 de este capítulo).

### B-2. Análisis y representación

La pregunta 7 busca conocer qué atributos son de importancia para los investigadores, en relación con la apertura de datos por parte de las administraciones. Para ello deben valorar con una escala de 0 a 4, (siendo 0 desconocimiento - 4 atributos de mucha importancia) los 8 atributos planteados en el enunciado. Ha sido contestada por 487 investigadores y omitida por 296. Se obtiene la siguiente información:

Atributo	Nada importante		Poco importante		Importante		Muy importante		No conozco		Total	Media
	%	Total	%	Total	%	Total	%	Total	%	Total		
Suministro	0,83	4	4,15	20	30,71	148	62,66	302	1,66	8	482	3,52
Licencia datos	1,87	9	7,28	35	31,81	153	54,47	262	4,57	22	481	3,3
Metadatos	0,83	4	3,52	17	22,15	107	68,74	332	4,76	23	483	3,49
Accesibilidad	0,62	3	1,24	6	17,77	86	76,24	369	4,13	20	484	3,61
Formatos	1,04	5	4,35	21	26,92	130	63,35	306	4,35	21	483	3,44
Restricciones	2,70	13	14,55	70	42,41	204	35,76	172	4,57	22	481	3,02
Actualización	3,87	18	24,52	114	34,84	162	33,76	157	3,01	14	465	2,92
Geolocalización	5,23	25	20,71	99	32,01	153	34,94	167	7,11	34	478	2,82

Tabla 29: Consideración de importancia de atributos en datos abiertos.

La tabla 29 recoge toda la información obtenida. Las afirmaciones han sido evaluadas:

- Suministro de datos sin restricciones: 482 veces.
- Tipo de licencia de los datos para su reutilización: 481 veces.
- Que contenga metadatos (completos, estructurados y útiles para interpretarlos): 483 veces.
- Que sea fácil encontrarlos y conseguirlos (accesibilidad): 484 veces.
- Que estén estandarizados (formato, variables...): 483 veces.
- Posibilidad de restricción a algunos tipos de datos (datos sensibles): 481 veces.

- Datos a tiempo real: 465 veces.
- Datos georreferenciados: 478 veces.

La media obtenida para cada uno de los atributos ha sido la siguiente. La evaluación de estos atributos se realiza sobre 4. La nota se obtiene del siguiente modo:

$$\text{N}^\circ \text{ de respuestas en la matriz de valoración (No conozco hasta Muy importante)} * \text{Nota de cada valor de matriz} / \text{Total de respuestas al atributo} = \text{Nota media de importancia del atributo.}$$

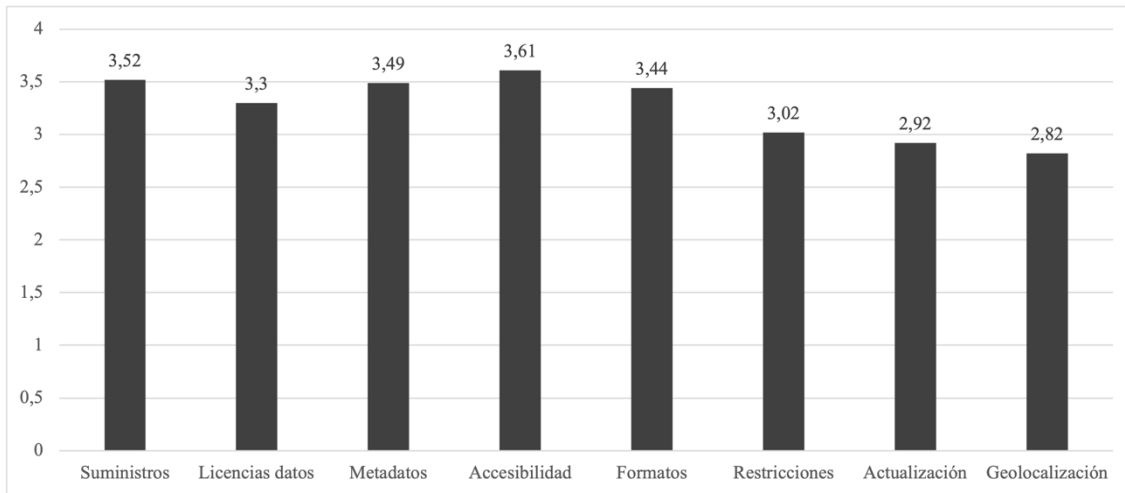


Figura 15: Valoración media de los atributos de los datos abiertos.

Los investigadores han valorado, con un 3,61 de media, y siendo el atributo más importante, que los datos abiertos sean fáciles de encontrar y conseguir, es decir, que sean fácilmente accesibles. Con un 3,52 de media, el siguiente atributo más valorado por los investigadores encuestados es el suministro de datos sin restricciones por parte de las administraciones, es decir, que sean abiertos. Con un 3,49 de media, el hecho de que contenga metadatos completos, útiles y estructurados es de vital importancia para su interpretación y reutilización.

Estos tres elementos más valorados nos permiten hacernos una primera idea de qué atributos deben tener en cuenta las administraciones encargadas de la apertura de datos. Es de vital importancia que los datos sean fácilmente accesibles y localizables, que estén disponibles sin restricciones (sin necesidad de pago o registro para acceder a ellos) y que los metadatos que contengan estén completos y estructurados para facilitar su interpretación.

Con un 3,44 de media, los investigadores consideran importante la estandarización de los formatos de datos. Es necesario para poder cruzar variables sin necesidad de hacer un excesivo trabajo de limpieza y estandarización de datos previo para llevar a cabo los análisis y el desarrollo de productos o servicios basados en datos.

Las licencias de los datos, con un 3,30 de media, es un elemento importante a tener en cuenta, ya que son las que imponen las restricciones y permiten difundir, reutilizar y publicar los datos, con fines comerciales o no (Datos.Gob, 2014).

Se ha valorado con un 3,02 de media, la posibilidad de restricción de determinados tipos de datos, considerados como datos sensibles.

Finalmente, los atributos que menos importancia se les ha dado han sido a la posibilidad de disponer de datos a tiempo real, con un 2,92 de media y a los datos con información geográfica (georreferenciados) con un 2,82.

La pregunta 8 busca conocer qué servicios y funcionalidades considera el investigador de más relevancia al acceder a los datos a través de los portales para su reutilización. Esta pregunta la han contestado en 481 ocasiones siendo omitida en 302.

Al igual que en la pregunta anterior (pregunta 7), se valora con una escala de 0 a 4, (siendo 0 desconocimiento - 4 atributos de mucha importancia) las 7 afirmaciones planteadas en el enunciado teniendo además una opción de respuesta abierta. La información obtenida es la siguiente:

Atributo	Nada importante		Poco importante		Importante		Muy importante		No conozco		Total	Media
	%	Total	%	Total	%	Total	%	Total	%	Total		
URI	1,47	7	4,42	21	38,95	185	33,26	158	21,89	104	475	2,6
Apps relacionadas	2,11	10	21,89	104	42,11	200	26,32	125	7,58	36	475	2,77
Herramientas procesamiento	2,51	12	18,41	88	40,38	193	35,98	172	2,72	13	478	3,04
Visualizaciones	1,04	5	9,81	47	40,50	194	46,35	222	2,30	11	479	3,28
Contacto	1,47	7	16,00	76	50,32	239	24,63	117	7,58	36	475	2,83
Noticias	3,78	18	27,73	132	46,01	219	18,07	86	4,41	21	476	2,7
API	2,13	10	15,57	73	28,57	134	19,40	91	34,33	161	469	1,97

Tabla 30: Valoración de servicios y funcionalidades de acceso a datos abiertos.

Los servicios y funcionalidades planteados que han valorado los investigadores son:

- Los datos tienen URI permanentes: 475 veces.
- Enlaces a aplicaciones relacionadas con el conjunto de datos: 475 veces.
- Herramientas para el procesamiento y análisis de datos: 478 veces.
- Visualizaciones de los datos: 479 veces.
- Canales de contacto o formularios para reutilizadores: 475 veces.
- Noticias o novedades relacionadas con los datos abiertos y sistemas de alerta de actualización del catálogo de datos: 476 veces.

- Disponibilidad de API: 469 veces.

La media obtenida para cada uno de los servicios planteados se ha calculado del mismo modo que en la pregunta 7. La evaluación de estos atributos se realiza sobre 4:

$$\text{Nota media de importancia del atributo} = \frac{\text{N}^\circ \text{ de respuestas en la matriz de valoración (No conozco hasta Muy importante)} * \text{Nota de cada valor de matriz}}{\text{Total de respuestas al atributo}}$$

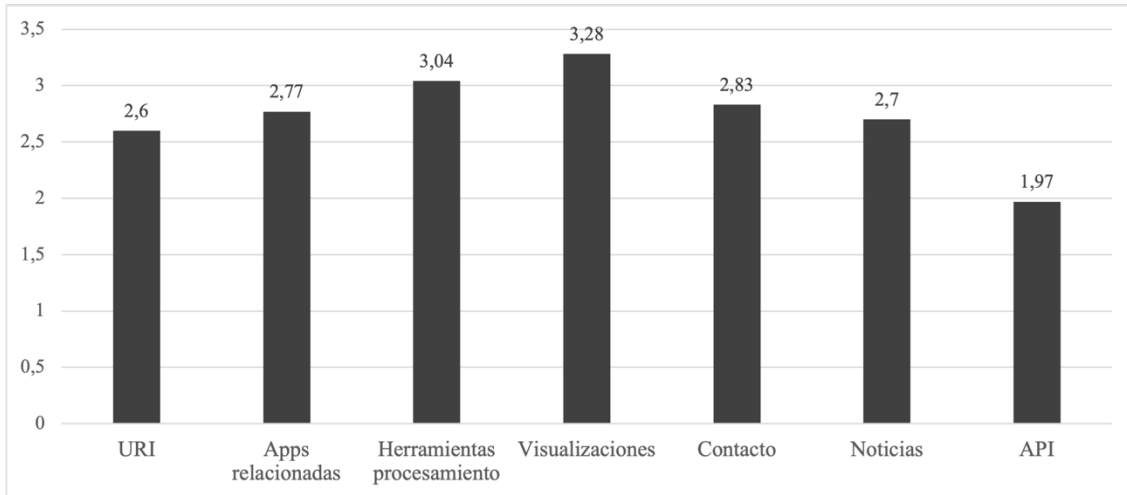


Figura 16: Valoración de servicios y funcionalidades de acceso a datos abiertos.

La funcionalidad más valorada en cuanto al acceso a los datos, con un 3,28 de nota media, son las visualizaciones de datos en los portales. Con un 3,04 de nota media, los investigadores consideran importante que haya ejemplos o enlaces a herramientas para el procesamiento y análisis de los datos.

Con casi 3 puntos, 2,83, se considera poco importante que existan canales de comunicación o contacto para aquellos reutilizadores que tengan dificultades o dudas a la hora de reutilizar la información del sector público, siguiéndole de cerca y considerándose poco importante, con 2,77 de media, que los portales dispongan de enlaces a aplicaciones relacionadas con los conjuntos de datos que sirvan como guía o ejemplo.

Tampoco consideran de vital importancia, 2,7 de media, que exista un canal de noticias o novedades relacionadas con los datos abiertos y un sistema de alerta de actualización del catálogo de datos. Se asocia a que cuando el investigador necesita un conjunto de datos, acude a la fuente de información pertinente en búsqueda concreta de lo que se necesita.

Destacan las bajas valoraciones acerca de URI's permanentes en datos y disponibilidad de API por parte de los portales de datos abiertos. Con 2,6 de nota media, que los datos tengan URI's permanentes no es una funcionalidad valorada por los investigadores y con 1,97 que



los portales de datos abiertos dispongan de un servicio API. Estas valoraciones son las más bajas ya que, en 104 (URI permanente de los datos) y 161 (disponibilidad de API web) casos respectivamente, los investigadores no conocen estas funcionalidades.

Además, en la opción de respuesta abierta, 7 investigadores aportan información sobre servicios y funciones que consideran importantes al acceder a los datos disponibles en los portales:

- Opciones múltiples de filtrado y tipos de formatos de descarga.
- Idiomas de presentación de los datos.
- Buscadores avanzados de datos.
- Inteligencia artificial para mejorar la experiencia de usuario.
- Evaluación por pares y depósitos en servidores institucionales.
- Estandarización de datos.

En la pregunta 9, el encuestado debe evaluar, con una escala de valores de 0 a 4, siendo 0 “no conozco” a 4 “siempre”, las barreras de acceso y reutilización que se encuentran a la hora trabajar con datos abiertos que se plantean en el enunciado. Se ha respondido un total de 406 veces, habiéndose omitido en 377. Se es más crítico y las administraciones encargadas de la apertura de datos deben mejorar estas barreras cuando la nota final obtenida esté más cerca del 4 de media.

Atributo	Siempre		Muchas veces		Algunas veces		Nunca		No conozco		Total	Media
	%	Total	%	Total	%	Total	%	Total	%	Total		
Tiempo	5,00	20	15,50	62	43,00	172	11,75	47	24,75	99	400	2,61
Registro	9,70	39	31,59	127	37,31	150	8,21	33	13,18	53	402	1,97
Comprensión	8,00	32	27,50	110	41,75	167	11,00	44	11,75	47	400	2,03
Calidad y estandarización	12,66	51	38,46	155	29,28	118	6,45	26	13,15	53	403	1,82
Pago	2,83	11	13,88	54	37,28	145	26,22	102	19,79	77	389	2,66
Interoperabilidad	6,89	27	31,38	123	30,61	120	8,42	33	22,70	89	392	2,31
Falta de actualización	6,00	24	33,25	133	35,75	143	11,25	45	13,75	55	400	2,07
Acceso	5,13	20	33,85	132	42,05	164	7,69	30	11,28	44	390	1,97

Tabla 31: Valoración de barreras de acceso y reutilización de datos abiertos.

Las barreras de acceso y reutilización que se plantean son las siguientes. En cada una se especifica el número de veces que se ha respondido:

- Mucho tiempo transcurrido entre la solicitud y la recepción de datos: 400 veces.
- La necesidad de registro para acceder a los datos: 402 veces.
- Acceder y comprender las condiciones de uso y de licencias y como interpretar y reutilizar los datos: 400 veces.

- Grados variables de calidad de los datos en conjuntos de datos diferentes, falta de estandarización en los formatos de los datos, información completa o de poco interés: 403 veces.
- Pago por los datos: 389 veces.
- Falta de interoperabilidad o de disponibilidad de los datos en otras AAPP / Dificultad para compararlos con otros datos: 392 veces.
- Falta de actualización de los datos: 400 veces.
- Dificultad de acceso: 390 veces.

La nota media acerca de las barreras de acceso y reutilización de los datos se calcula de la misma forma que las preguntas 7 y 8, siendo sobre 4:

$$\text{Nota media de importancia del atributo} = \frac{\text{Nº de respuestas en la matriz de valoración (No conozco hasta Muy importante)} * \text{Nota de cada valor de matriz}}{\text{Total de respuestas al atributo}}$$

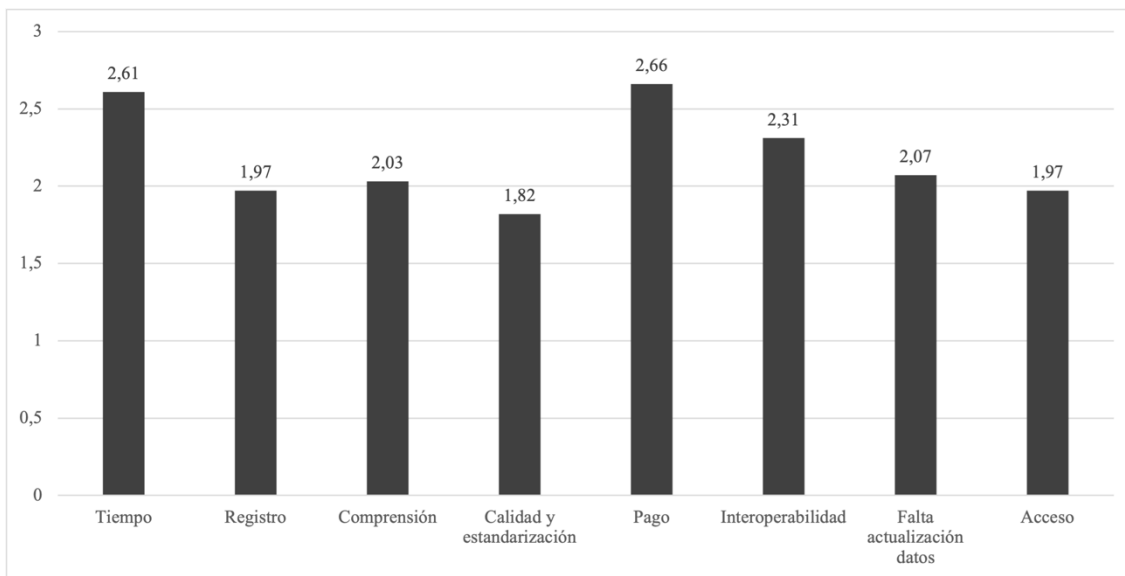


Figura 17: Valoración de barreras de acceso y reutilización de datos abiertos.

Las barreras de acceso a datos y reutilización a las que se enfrentan los investigadores y con las que más críticos han sido el pago por los datos (2,66 de nota media) y el tiempo transcurrido entre la solicitud y la recepción de los datos (2,61 de nota media). Se ha respondido a la opción “no conozco”, siendo un 0, en 77 (19,79%) y 99 (24,75%) ocasiones respectivamente.

En tercer lugar, la falta de interoperabilidad y disponibilidad de datos en otras administraciones públicas ha obtenido una nota media de 2,31, habiendo respondido a la opción “no conozco” en 89 ocasiones (22,70%).

Con 2,07 de nota media, la siguiente barrera que encuentran los investigadores es la falta de actualización de los datos. En 143 ocasiones (35,75%), la comunidad científica señala que “algunas veces” es una barrera para su uso y 133 investigadores (33,25%) señalan que es un problema al que se enfrentan “muchas veces”.

El acceso a licencias para la comprensión de las condiciones de uso y reutilización obtiene un 2,03 de nota media. En 167 ocasiones, los investigadores señalan que es una barrera “algunas veces” (41,75%), y para 110 investigadores es un problema “muchas veces” (27,50%).

Donde se detectan, estando por debajo del 2 de nota media, es en los grados variables de calidad e información incompleta con 1,82 de nota media y, con la misma nota, 1,97, la necesidad de registrarse para acceder a los datos y dificultades de acceso a la información.

La última pregunta del bloque 2, la pregunta 10, pretende conocer los tipos de datos más interesantes para investigación y trabajo de los encuestados. Se responde un total de 477 veces siendo omitida en 306 ocasiones. Cada investigador puede seleccionar más de un conjunto de datos de los propuestos. Además, disponen de una opción de respuesta abierta para introducir aquellos conjuntos de datos de interés que no estén en el enunciado.

La han contestado 477 investigadores. Cada investigador tiene la opción de elegir uno o más tipos de datos que considere necesarios para su trabajo/investigación. Para comentar los conjuntos de datos se crean 4 tablas que muestran el número de elecciones de conjuntos de datos. Con más de 100 elecciones, los conjuntos de datos más valorados por la comunidad científica española son los siguientes:

Conjunto de datos	Nº de veces elegido
Ciencia y Tecnología	220
Sector público	199
Educación	170
Medio ambiente y clima	160
Salud	151
Demografía y población	139
Economía	129
Sociedad y bienestar	111

Tabla 32: Conjuntos de datos más relevantes para los investigadores (+ 100 elecciones).

La tabla 32 muestra los conjuntos de datos de más relevancia para los investigadores encuestados. Muestra los conjuntos de datos que han sido elegidos más de 100 veces.

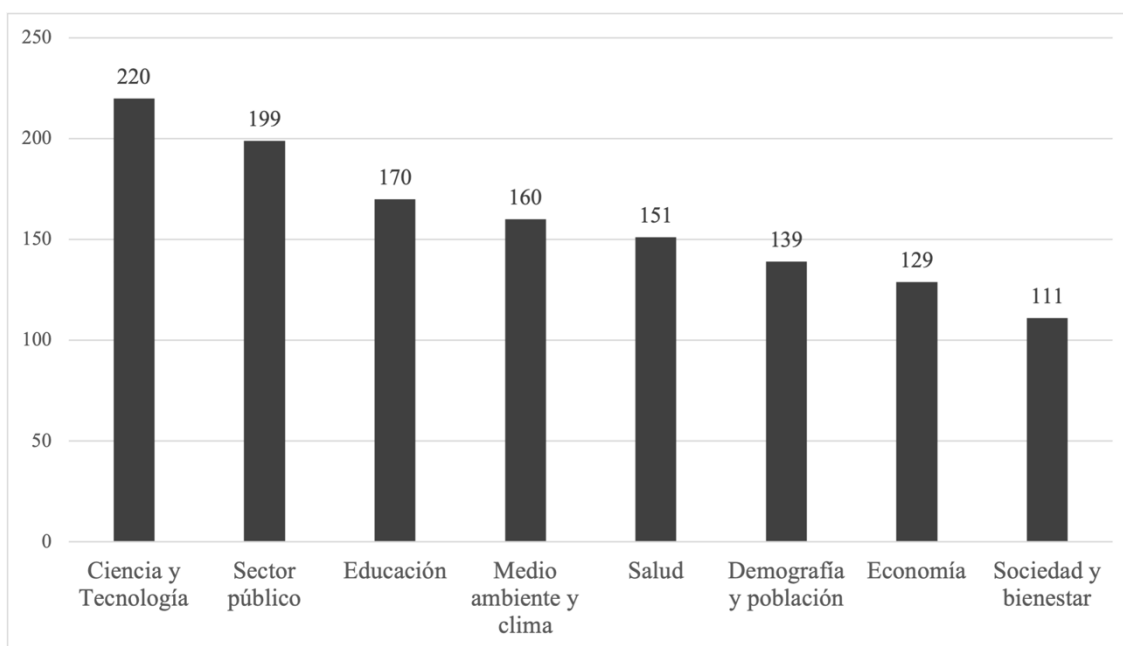


Figura 18: Conjuntos de datos más relevantes para los investigadores/as (+ 100 elecciones).

Los datos de más interés, con un total de 220 elecciones, son los relacionados en ciencia y tecnología. En segundo lugar, con 199 elecciones, datos acerca del sector público. Educación con 170, medio ambiente y clima con 160 y salud con 150 son los siguientes conjuntos de datos que ha señalado como aquellos de más relevancia. Los datos sobre demografía y población, economía y sociedad y bienestar han sido elegidos 139, 129 y 111 veces respectivamente.

Conjunto de datos	Nº de veces elegido
Cultura y ocio	83
Industria	77
Administración pública	71
Energía	69
Empleo	68
Urbanismo e infraestructuras	65
Consumo energético	58
Transporte	57
Medio rural	56
Catastro	53
Legislación y justicia	51
Vivienda	50

Tabla 33: Conjuntos de datos más relevantes para los investigadores/as (entre 50 y 90 veces seleccionados).

En la tabla 33, se muestran los conjuntos de datos elegidos entre 99 y 50 veces. Cultura y ocio, industria y administración pública son los 3 conjuntos de datos más elegidos en esta segunda tabla, con 83, 77 y 71 elecciones respectivamente.

Los siguientes conjuntos de datos que más necesitan los investigadores para llevar a cabo sus proyectos son los relacionados con energía, 69, empleo, 68 y urbanismo e infraestructuras con 65. Entre 50 y 60 elecciones se obtienen un total de 6 conjuntos de datos.

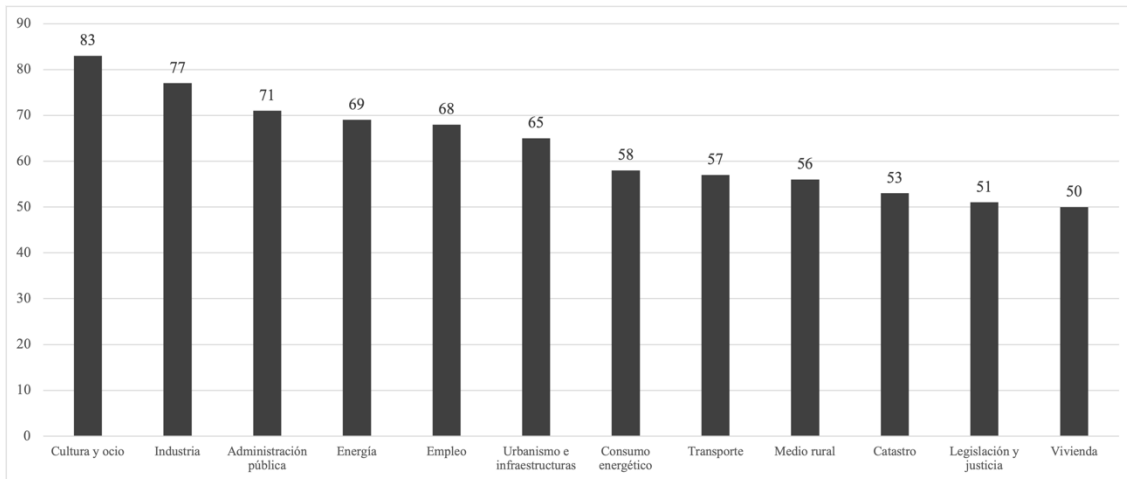


Figura 19: Conjuntos de datos más relevantes para los investigadores/as (entre 50 y 90 veces seleccionados).

Conjunto de datos	Nº de veces elegido
Turismo	48
Comercio	41
Otro (especifique)	39
Seguridad Social	32
Asociaciones	30
Importaciones/exportaciones	30
Información mercantil	27
Hacienda	26
Deporte	24
Seguridad	23
Cooperativas	17
Titularidades reales	15
Autónomos	13

Tabla 34: Conjuntos de datos más relevantes para los investigadores/as (menos de 50 elecciones).

En la tabla 34 se muestran aquellos conjuntos de datos elegidos menos de 50 veces por los investigadores. Destacan los datos de turismo y comercio, con 48 y 41 elecciones respectivamente. La opción abierta ha sido contestada un total de 39 y se muestra en la tabla siguiente. Para finalizar, con 30 o más elecciones hay 3 conjuntos de datos, con datasets sobre seguridad social, asociaciones e importaciones/exportaciones. Con menos de 30 elecciones hay un total de 7 conjuntos.

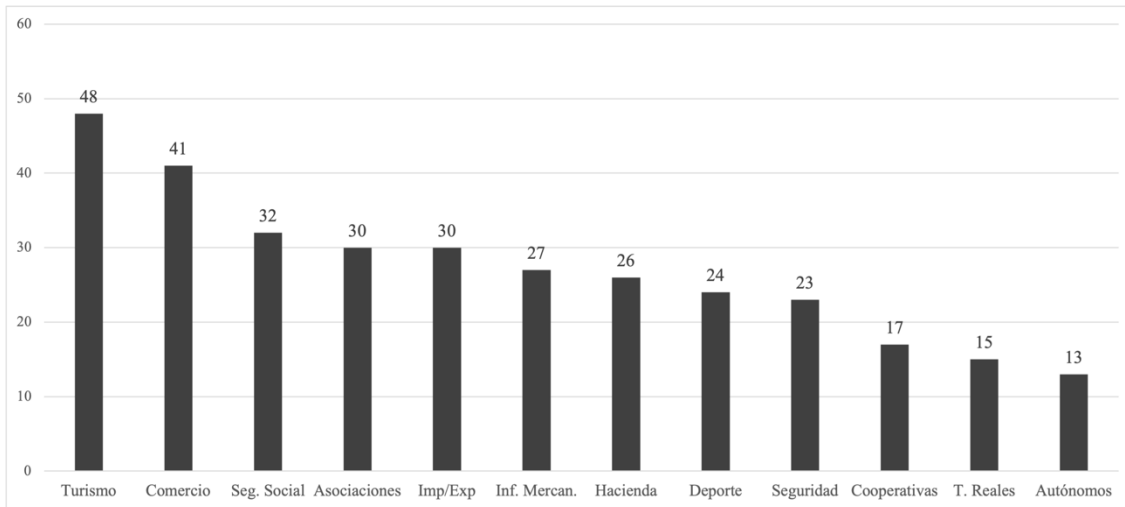


Figura 20: Conjuntos de datos más relevantes para los investigadores/as (menos de 50 elecciones).

En la categoría otros, se permitía responder de forma abierta. Se ha contestado un total de 39 veces y tras la normalización de las respuestas, las principales aportaciones de los investigadores son las siguientes:

Otros	Nº de veces elegido
Multidisciplinares	7
Arte, arquitectura e ingeniería de la construcción	5
Agricultura, alimentación y tecnología de los alimentos	4
Datos sobre historia y patrimonio histórico	3
Cartografía en sus múltiples formatos (Cartografía histórica)	3
Comunicación digital (medios de comunicación, RRSS)	3
Datos de estudios e investigaciones	2
Estadísticas internas de Universidades Públicas	1
Sísmicos	1
Salud mental	1
Igualdad, centros educativos, políticas y proyectos realizados en los centros	1
Datos experimentales de física y ciencia de materiales.	1
Residuos y otros datos medioambientales	1
Turismo	1
Participación y gobernanza	1
Materiales	1
Violencia de género, racismo y cuestiones LGTBI+	1
Patentes	1
Meteorológicos	1

Tabla 35: Conjuntos de datos obtenidos en la respuesta abierta.

Tras la normalización de las respuestas obtenidas en la opción de respuesta abierta, 7 investigadores dan la respuesta de “multidisciplinares” refiriéndose a los conjuntos de datos. Con 5 elecciones, investigadores consideran de relevancia aquellos datos sobre arte, arquitectura e ingeniería de la construcción. Con 4 elecciones, se consideran de importancia datos sobre agricultura, alimentación y tecnología de los alimentos. Con 3 elecciones hay 3

conjuntos de datos: cartografía, comunicación digital y datos sobre historia y patrimonio histórico. Dos investigadores comentan la necesidad de datos sobre estudios e investigaciones (datos de investigación) y con 1 elección existen 12 tipos de conjuntos de datos.

Bloque 3. Utilidad real y motivaciones para la reutilización de datos abiertos en la investigación.

B-3. Estandarización y limpieza de datos

El bloque 3 mide la percepción de la comunidad científica acerca de la utilidad real y motivaciones para la reutilización de datos abiertos en la investigación. Este bloque está formado por 2 preguntas, la 11 y la 12. En ambas preguntas deben evaluar una serie de afirmaciones relacionadas en la política de apertura de datos del sector público para la sociedad.

En la pregunta 11, se plantean 5 afirmaciones en las que el encuestado debe evaluar en la siguiente escala de valoración. En la opción abierta, 7 investigadores realizan aportaciones.

- Muy importante (4); Importante (3); Poco importante (2); Nada importante (1); No es útil porque no existen conjuntos de datos (0).

Para la posterior representación de la información en las visualizaciones, se codifica de la siguiente manera:

---

Fomento investigación
Apoyo
Promoción transferencia
Desarrollo
Transparencia
Otros

---

Tabla 36: Categorías de la pregunta 11 de la encuesta.

La pregunta 12 deben valorar las afirmaciones planteadas con la siguiente escala de clasificación:

- Muy importante (3); Importante (2); Poco importante (1); Nada importante (0).

Al igual que en preguntas anteriores, para la representación de la información, es necesario codificar los enunciados planteados:

---

Participación
Desarrollo
Compromiso
Innovación
Crecimiento

---

Tabla 37: Categorías de la pregunta 12 de la encuesta.



B-3. Análisis y representación

La pregunta 11 plantea 5 afirmaciones en las que los encuestados deben valorar según su experiencia, la utilidad real que tiene la apertura de datos abiertos. Esta pregunta la han respondido 458 investigadores y ha sido omitida en 325 ocasiones. Se obtiene la información contenida en la tabla siguiente, la 38:

Atributo	Nada importante		Poco importante		Importante		Muy importante		Sin utilidad <sup>83</sup>		Total	Media
	%	Total	%	Total	%	Total	%	Total	%	Total		
Fomento investigación	1,09	5	4,59	21	31,66	145	61,79	283	0,87	4	458	3,52
Apoyo	0,88	4	4,62	21	31,21	142	61,98	282	1,32	6	455	3,52
Promoción transferencia	1,32	6	7,73	35	37,97	172	51,88	235	1,10	5	453	3,38
Desarrollo	2,44	11	15,08	68	39,47	178	41,02	185	2,00	9	451	3,15
Transparencia	2,44	11	13,56	61	33,56	151	47,33	213	3,11	14	450	3,2

Tabla 38: Valoración de utilidad de los datos abiertos.

La media de valoraciones las más altas de 3,52 en dos ocasiones y la más baja un 3,15.

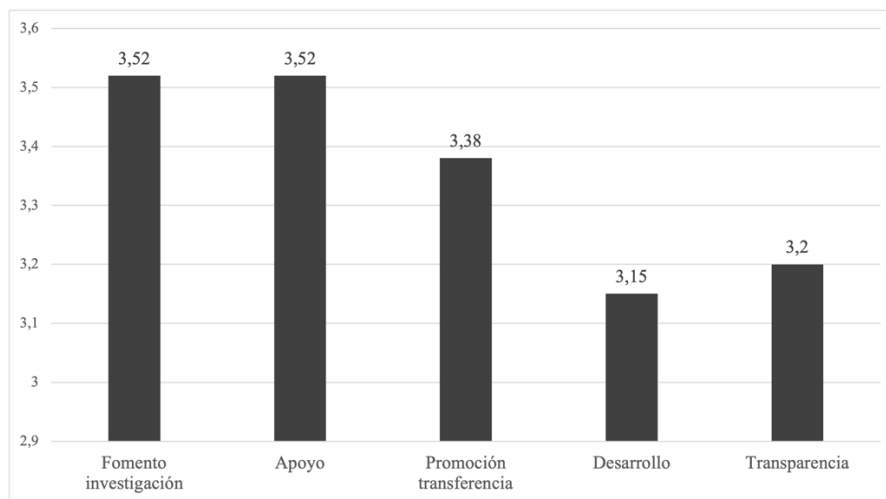


Figura 21: Calificación media en afirmaciones de utilidad de los datos abiertos.

Se han obtenido un 3,52 (sobre 4) en las afirmaciones sobre la importancia del fomento de los datos abiertos en la investigación y actividades educativas y que sirven de respaldo en las aplicaciones de la ciencia a los problemas de la sociedad. Con un 3,38, se considera de importancia el hecho de que los datos abiertos sirvan para promocionar la creación de transferencia de la tecnología.

Las dos afirmaciones menos valoradas de esta pregunta, con más de 3 de nota media, es decir, son de utilidad a la hora de abrir datos para la investigación, son con un 3,2, que permiten la

<sup>83</sup> Sin utilidad porque no existen conjuntos de datos.

rendición de cuentas y la participación informada de la sociedad civil, y con un 3,15, los datos abiertos son de utilidad para el desarrollo de nuevos productos y servicios.

Además, se han obtenido 7 respuestas en la opción abierta. Se descarta una de ellas porque está vacía. Los investigadores que han realizado aportaciones hacen hincapié en la utilidad de la apertura de datos abiertos en contextos que fomentan:

- Intercambio espontáneo de ideas y apertura de líneas de investigación.
- Replicación y comprobación de investigaciones realizadas
- Transparencia en la investigación
- Sirven para mostrar a los investigadores públicos la finalidad de su trabajo.
- Suponen el punto de partida para el diseño de nuevos productos o servicios de información estratégica y el acercamiento al paradigma del *Linked Open Data*.
- Sirven como retorno social de las inversiones públicas en ciencia y tecnología.
- Acelera el avance como sociedad.

La pregunta 12 pretende conocer el valor que tiene para los investigadores, las políticas de apertura de datos para en la sociedad. La responden en 462 investigadores y la omiten 321.

Se obtiene la información disponible en la siguiente tabla, la 39:

Atributo	Nada importante		Poco importante		Importante		Muy importante		Total	Media
	%	Total	%	Total	%	Total	%	Total		
Participación	3,28	15	13,54	62	51,97	238	31,22	143	458	3,11
Desarrollo	2,62	12	14,63	67	53,93	247	28,82	132	458	3,09
Compromiso	2,40	11	10,04	46	42,79	196	44,76	205	458	3,3
Innovación	2,17	10	7,83	36	47,39	218	42,61	196	460	3,3
Crecimiento	2,86	13	13,22	60	46,04	209	37,89	172	454	3,19

Tabla 39: Importancia de la apertura de datos en la sociedad.

Las calificaciones medias obtenidas a través de las respuestas de los investigadores están entre 3,3 (calificación media más alta, en dos atributos) y la más baja 3,09.

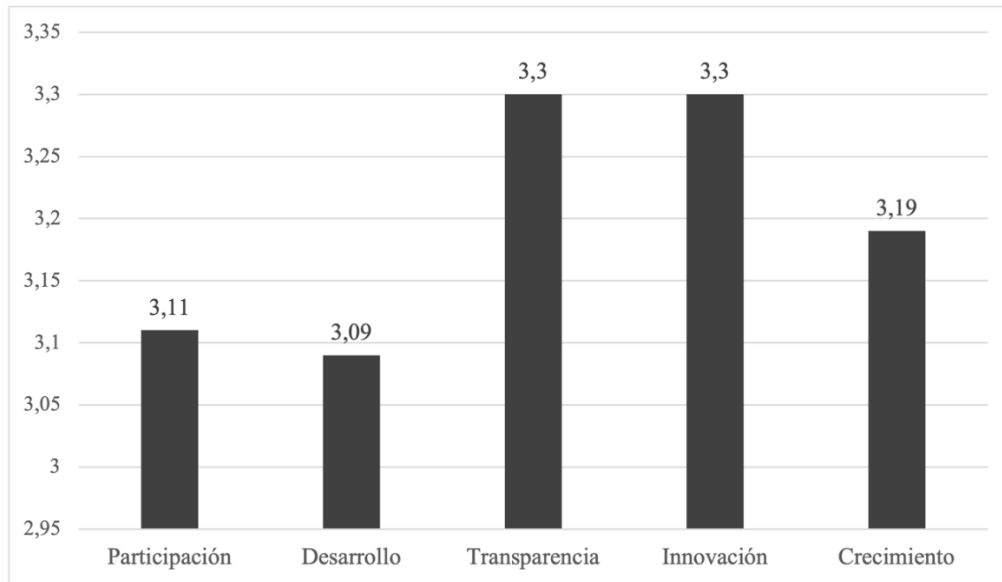


Figura 22: Importancia de la apertura de datos en la sociedad.

Los investigadores valoran las políticas de apertura de datos en la sociedad del siguiente modo. Con 3,3 de valoración media, consideran que las políticas de datos sirven para reforzar la confianza en las instituciones a través de la transparencia y aceleran la transformación digital y el uso del big data en las administraciones públicas. Con 3,19 de valoración media, estas iniciativas de apertura favorecen la digitalización y el crecimiento económico de la sociedad. Las valoraciones más bajas son, con un 3,11, el hecho de involucrar a la ciudadanía y la comunidad científica y con 3,09, la creación y desarrollo de nuevos productos y servicios.

Bloque 4. Acceso a los conjuntos de datos: fuentes de datos abiertos consultadas  
 B-4. Estandarización y limpieza de datos

El bloque 4, formado por 5 preguntas (de la 13 a la 17), estudia las fuentes de datos consultadas por los investigadores y el acceso a estos tipos de datos.

Las preguntas 13 y 14 se plantean las opciones en el enunciado de la encuesta sobre tipos de portales de datos y sobre los portales de datos abiertos autonómicos españoles.

La pregunta 15 es de respuesta abierta y se normalizan las respuestas en el apartado de los resultados de este capítulo.

En la pregunta 16 deben evaluar con la escala de valoración entre las opciones planteadas en el enunciado, siendo:

- Muy importante (4); Importante (3); Necesario (2); Poca importancia (1); Nada importante (0)

Para la posterior representación, los enunciados se codifican con una palabra clave de la siguiente manera:

---

Crecimiento-Formación
Comprensión
Desarrollo
Objetivos profesionales
Web-semántica
Acceso

---

Tabla 40: Categorías de la pregunta 16 de la encuesta.

En la pregunta 17 se plantean una serie de cuestiones en las que el investigador puede elegir tantas opciones como considere oportunas. Para su posterior representación mediante visualizaciones, se codifica de la siguiente manera:

---

Técnicas
Estandarización
Interoperabilidad
Financiación
Publicidad activa
Formación
Otros

---

Tabla 41: Categorías de la pregunta 17 de la encuesta.

B-4. Análisis y representación

La pregunta 13 busca conocer los portales de datos abiertos utilizados por los investigadores para llevar a cabo sus proyectos y trabajos. Es respondida por 365 investigadores y 418 la dejan en blanco. En el enunciado se plantean 10 tipos de portales con una opción de respuesta abierta, que es respondida en 51 ocasiones. Se obtiene la siguiente información.

Tipo de portal de datos	Nº veces elegido	%
Datos abiertos del Estado / Catálogo federado (datos.gob.es)	176	48,22
Portales de Organismos internacionales	154	42,19
Datos de portales de estadística de las administraciones	136	37,26
Portal Europeo de Datos ( <a href="https://www.europeandataportal.eu/es">https://www.europeandataportal.eu/es</a> )	129	35,34
Datos de otras CCAA u organizaciones autonómicas	111	30,41
Datos abiertos de otras instituciones públicas de la Comunidad Valenciana (ayuntamientos, diputaciones, universidades ...)	92	25,21
Datos abiertos GVA ( <a href="https://dadesobertes.gva.es/">https://dadesobertes.gva.es/</a> )	90	24,66
Datos de ayuntamientos	71	19,45
Datos de portales de transparencia de las administraciones	67	18,36
Portales de datos de iniciativa privada (empresas, entidades de la sociedad civil...)	62	16,99
Otro (especifique)	51	13,97

Tabla 42: Uso de información de portales de datos abiertos.

La tabla 42 se muestra en los dos gráficos siguientes. La figura 23, donde se plasman los tipos de portales de datos con más de 100 elecciones por parte de los investigadores y la figura 24 con menos de 100 elecciones.

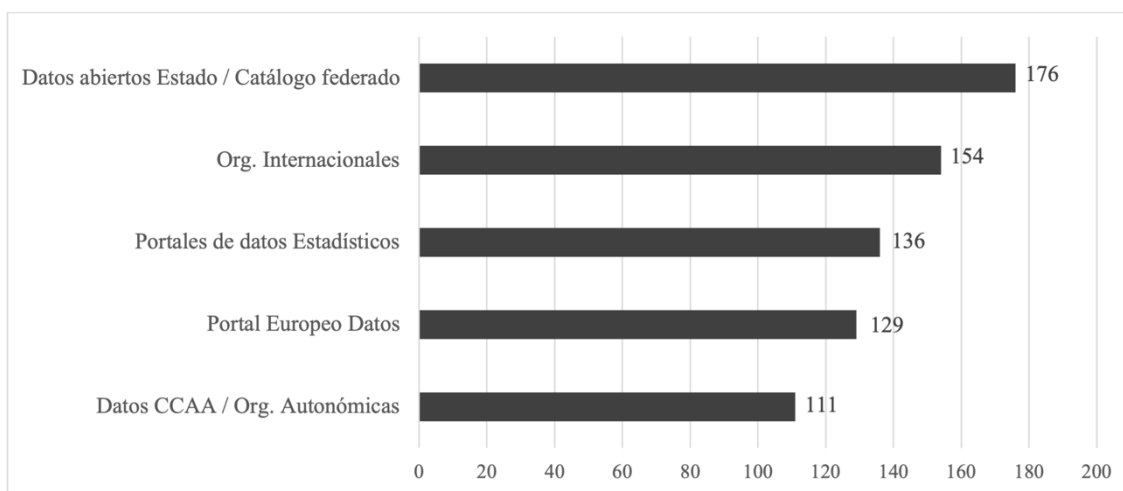


Figura 23: Tipos de portales de datos abiertos más utilizados (+ 100 elecciones).

Con 176 elecciones, el portal de datos que más utilizan los encuestados es datos.gob.es, portal de datos estatal en España que dispone de un catálogo federado de datos. Con 154 elecciones, los investigadores utilizan datos de portales de organismos internacionales. Le sigue con 136 elecciones portales con información estadísticas de las administraciones públicas, con 129, el Portal Europeo de Datos (<https://data.europa.eu/es>) y con 111, portales de datos abiertos de Comunidades Autónomas y organizaciones autonómicas.

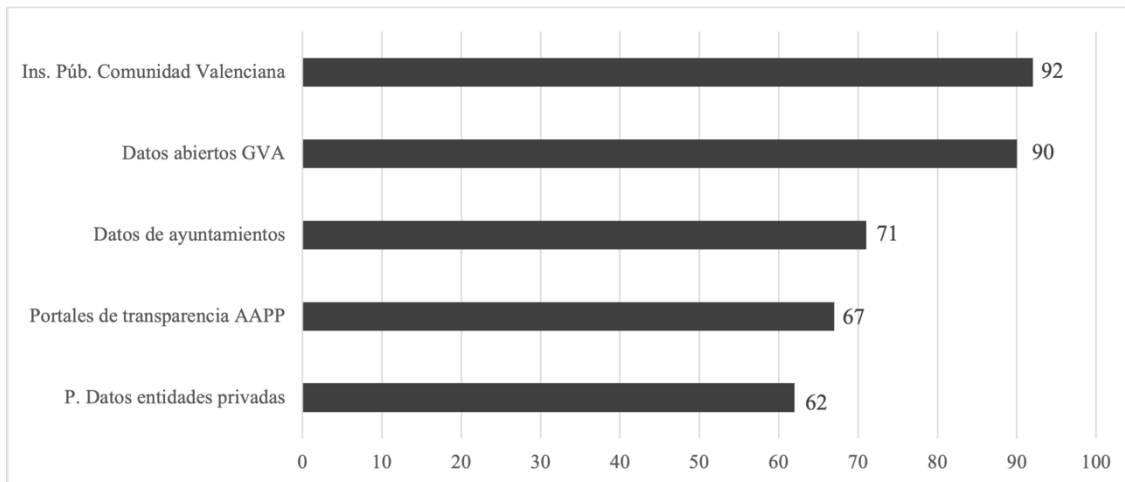


Figura 24: Tipos de portales de datos abiertos utilizados (seleccionados entre 50 y 100 ocasiones).

En la figura 24 se presentan 5 tipos de portales de datos abiertos elegidos entre 50 y 100 veces. En 92 ocasiones, los investigadores señalan que han utilizado datos de instituciones públicas de la Comunidad Valenciana. Con 90 elecciones, han utilizado datos del portal de la Generalitat Valenciana (<https://portaldadesobertes.gva.es/es>). 71 veces señalan haber utilizado datos abiertos de ayuntamientos, 67, han utilizado datos disponibles en portales de transparencia de las administraciones públicas y 62, han utilizado información disponible en portales de empresas y entidades privadas.

De las 51 respuestas obtenidas en la opción abierta, se descartan 15, cuyas respuestas no son válidas (“no”, “no he utilizado”, “ninguno”, “no procede”), quedando así con 37 respuestas.

Portal de datos abiertos	Nº veces elegido
Datos científicos, datos de revistas científicas y repositorios de investigación	18
Datos estadísticos	4
Datos internacionales	3
Datos de universidades	3
Datos europeos	2
Portales institucionales/organizaciones públicas y privadas	2
Portal nacional de turismo	1
Datos nacionales	1
REDIAM (Red Ambiental de Andalucía)	1
Referencias bibliográficas	1
Wikidata	1

Tabla 43: Otros portales de datos abiertos obtenidos en pregunta 13.

Los investigadores hacen referencia a haber utilizado datos estadísticos en 4 ocasiones, datos disponibles en portales internacionales en 3 y datos europeos en 2. Se pueden englobar dentro de las categorías creadas en el enunciado.

En 18 ocasiones los encuestados señalan que han utilizado datos científicos, de revistas científicas disponibles en repositorios de investigación.

Con 2 elecciones hacen referencia a datos europeos y portales de datos institucionales. Con 1 elección se señalan 4 tipos de conjuntos de datos.

La pregunta 14 se plantea con el objetivo de conocer el acceso y uso de información y datos abiertos disponible en portales autonómicos. Es contestada en 270 ocasiones y se omite 513.

Los portales de datos abiertos autonómicos españoles a los que más han accedido los encuestados son:

Portal de CCAA	Nº	%
Comunidad Valenciana	133	49,26
Madrid	88	32,59
Andalucía	76	28,15
Cataluña	57	21,11
País Vasco	38	14,07
Castilla y León	35	12,96
Aragón	25	9,26
Murcia	22	8,15
Navarra	22	8,15
Canarias	18	6,67
Castilla-La Mancha	18	6,67
Galicia	16	5,93
Baleares	15	5,56
Extremadura	13	4,81
Asturias	11	4,07
Cantabria	6	2,22
Ceuta	6	2,22
Melilla	6	2,22
La Rioja	5	1,85

Tabla 44: Uso de datos abiertos en portales autonómicos.

La figura 25 muestra aquellos portales de datos abiertos autonómicos a los que han accedido más de 30 investigadores.

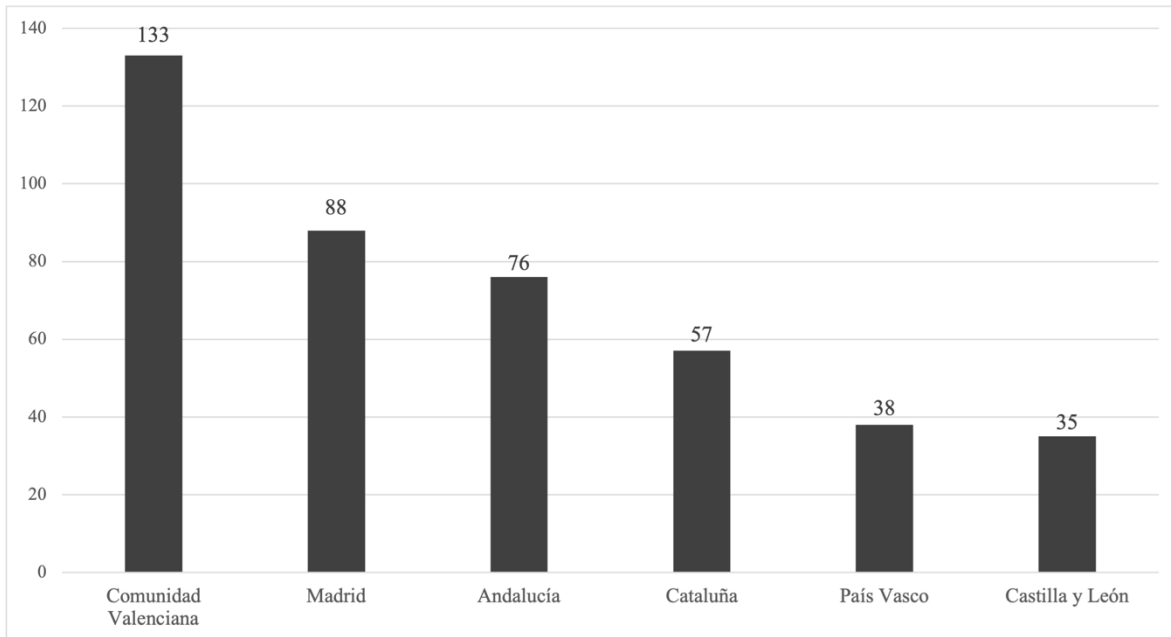


Figura 25: Portales de datos abiertos autonómicos más utilizados.

Con 133 elecciones, el portal al que más han accedido los investigadores es al de la Comunidad Valenciana. El segundo que más veces ha sido seleccionado es el de la Comunidad de Madrid con 88, seguido con 76 elecciones por el de la Junta de Andalucía.

El portal de datos abiertos de Cataluña ha sido elegido por 57 investigadores, el del País Vasco por 38 y finalmente, el de la Junta de Castilla y León un total de 35.

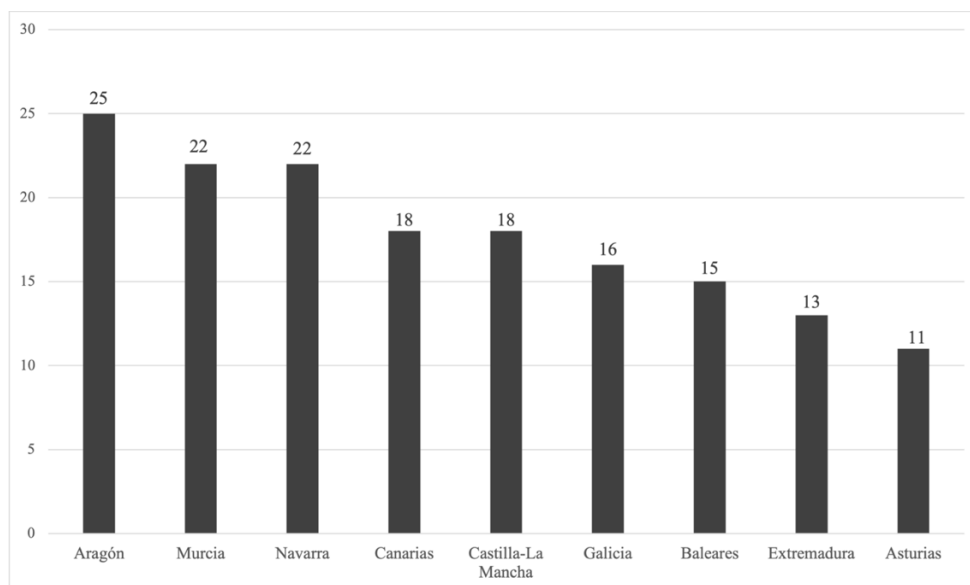


Figura 26: Uso de datos abiertos de portales autonómicos (- de 30 elecciones).



La figura 26 muestra aquellos portales que han sido seleccionados menos de 30 veces. Con 25 elecciones, el portal de datos del Gobierno de Aragón es al que más veces se ha accedido de aquellos que se muestran en esta gráfica. Lo sigue con 22 elecciones el de la Región de Murcia y Navarra. Con 18 elecciones se ha elegido a las comunidades del Gobierno de Canarias y el de Castilla-La Mancha, seguido con 16 el portal autonómico del Gobierno de Galicia, con 15 el de Baleares, con 13 el Gobierno de Extremadura y finalmente con 11, el del Gobierno de Asturias.

Con menos de 10 elecciones, los investigadores han seleccionado 6 veces el portal del Gobierno de Cantabria y de las ciudades autónomas de Ceuta y Melilla. Con 5, siendo el que menos veces ha sido elegido, el portal de datos abiertos de La Rioja.

En la pregunta 15, de respuestas abierta, se pide al investigador que especifique qué otros conjuntos de datos han utilizado de otras administraciones o fuentes. Se obtienen 133 respuestas, de las cuales se descartan 13 por no aportar información. Algunas respuestas apuntan a más de un tipo de fuentes de datos. Tras el proceso de normalización de respuestas se obtiene la siguiente información:

Fuente de datos	Total
Datos estadísticos	49
Datos nacionales	36
Datos autonómicos y locales	32
Datos universidades	6
Datos de organizaciones internacionales	6
Datos de investigación	6
Datos geoespaciales nacionales	2
Datos agrícolas	2
Datos climáticos y cartografía	1
Datos de salud	1
Datos meteorológicos	1

Tabla 45: Otras fuentes de datos utilizadas.

Tras el proceso de normalización, existe una diferencia grande entre los tipos de fuentes de datos utilizados por los investigadores. En la figura 27, se muestran los tipos de datos que más han elegido los investigadores.

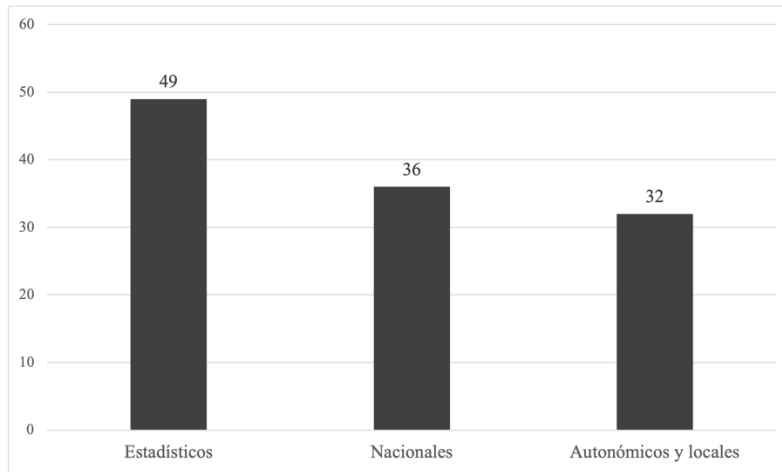


Figura 27: Fuentes datos utilizadas.

Con 49 elecciones, los investigadores hacen referencia al uso de fuentes de información con datos estadísticos. En 36 ocasiones, nombran tipos de datos disponibles en portales nacionales y con 32 elecciones, datos autonómicos y locales.

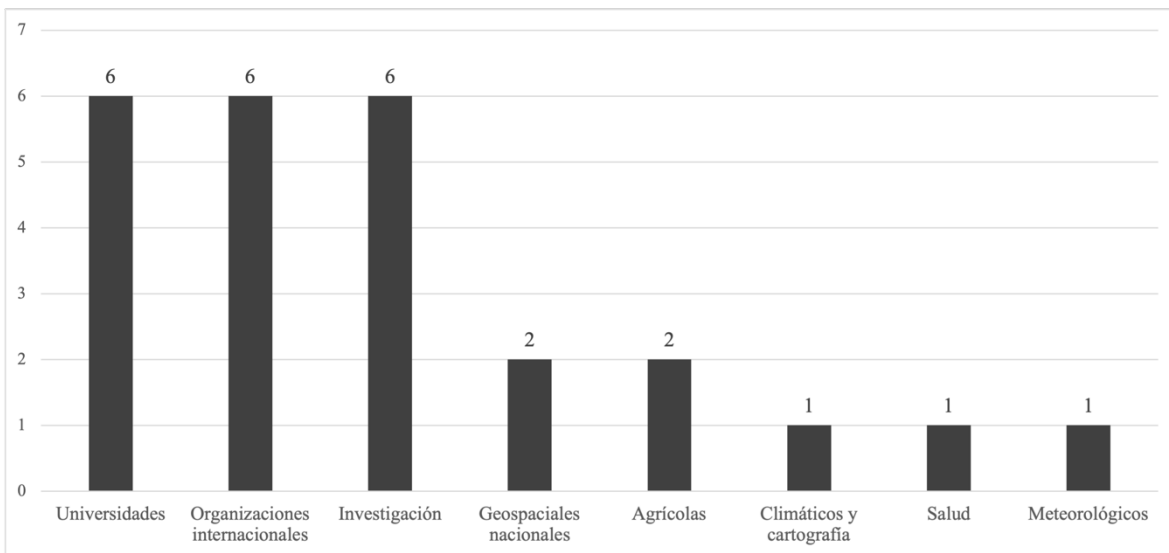


Figura 28: Otras fuentes datos utilizadas.

Con 6 elecciones, los encuestados señalan el uso de datos disponibles en portales de universidades, organizaciones internacionales y de investigación. Apuntan en 2 ocasiones el uso de datos geospaciales y agrícolas y finalmente, con 1 elección, a datos climáticos y de cartografía, de salud y meteorológicos.

La pregunta 16 mide la motivación y estímulos para trabajar y usar datos abiertos de administraciones utilizando una matriz de valoración de importancia. La responden 397 investigadores y la omiten 386.

Ítem	Nada importante (0)		Poca importancia (1)		Necesario (2)		Importante (3)		Muy importante (4)		Total	Media
	%	Total	%	Total	%	Total	%	Total	%	Total		
Crecimiento-formación	3,35	13	4,64	18	23,71	92	52,06	202	16,24	63	388	2,73
Comprensión	7,03	27	7,81	30	31,25	120	42,19	162	11,72	45	384	2,44
Desarrollo	2,54	10	5,34	21	17,30	68	52,93	208	21,88	86	393	2,86
Objetivos profesionales	4,87	19	10,77	42	28,46	111	42,05	164	13,85	54	390	2,49
Web semántica	7,92	30	10,82	41	36,41	138	36,41	138	8,44	32	379	2,27
Acceso	3,34	13	4,63	18	14,91	58	52,96	206	24,16	94	389	2,9

Tabla 46: Motivación de uso de datos abiertos.

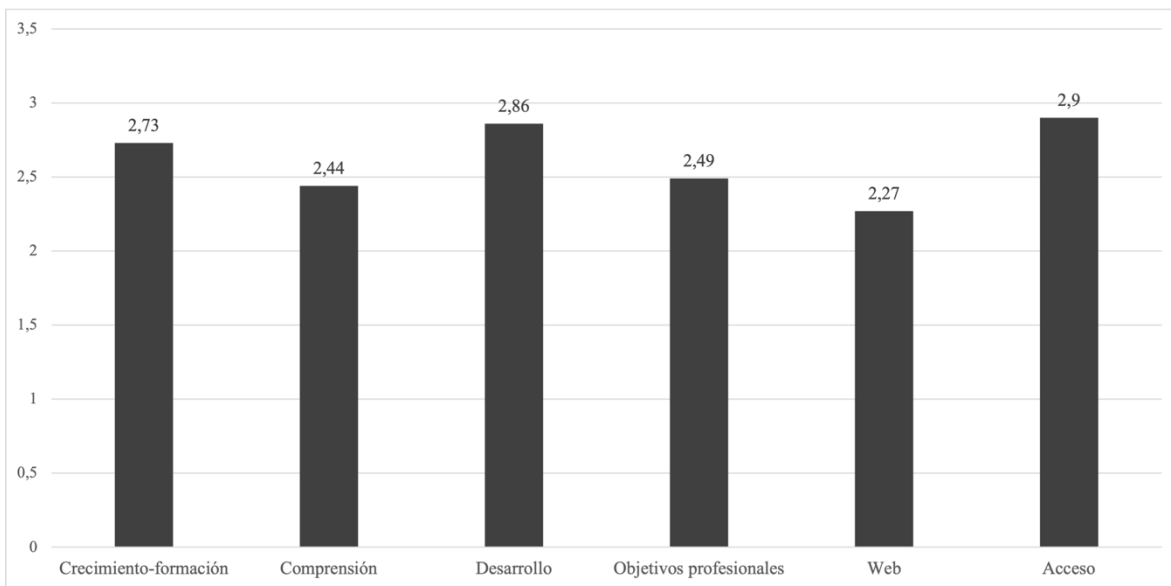


Figura 29: Motivaciones para el uso de datos abiertos.

La matriz de valoración utiliza el rango de “muy importante (4)” a “nada importante (0)”. Con 2,9 de media, el ítem mejor valorado es el que hace referencia a la mejora del acceso y comprensión de información pública por la ciudadanía. Le sigue, con 2,86, la motivación de desarrollos de servicios y productos innovadores basados en datos abiertos. Con un 2,73 de media, los investigadores ven una motivación o estímulo para utilizar información del sector público ya que les va a permitir crecer profesionalmente y formarse. Le sigue con 2,49 y relacionado con el ítem anterior, la motivación de uso de datos abiertos para el cumplimiento de objetivos profesionales. Con un 2,44 vinculan el uso de datos abiertos a una mejora en la capacidad de comprensión de funcionamiento de gobiernos y finalmente, con un 2,27, para el desarrollo de la web semántica y datos enlazados.

La última pregunta de este cuarto bloque es la 17. Mide la percepción de los investigadores ante una serie de cuestiones a las que se enfrentan los gobiernos en materia de apertura de datos abiertos. Deben seleccionar aquellos ítems que consideren de importancia ante los que

se enfrenten las administraciones. Es respondida un total de 408 veces y omitida 375. Se obtiene la información de la tabla 47 y dispone de una opción de respuesta abierta.

	Respuestas	%
Estandarización de los datos	282	69,12
Interoperabilidad administrativa	233	57,11
Falta de formación	188	46,08
Cuestiones técnicas	182	44,61
Financiación	139	34,07
Publicidad activa	110	26,96
Otro (especifique)	44	10,78

Tabla 47: Retos de apertura de datos abiertos en los gobiernos.

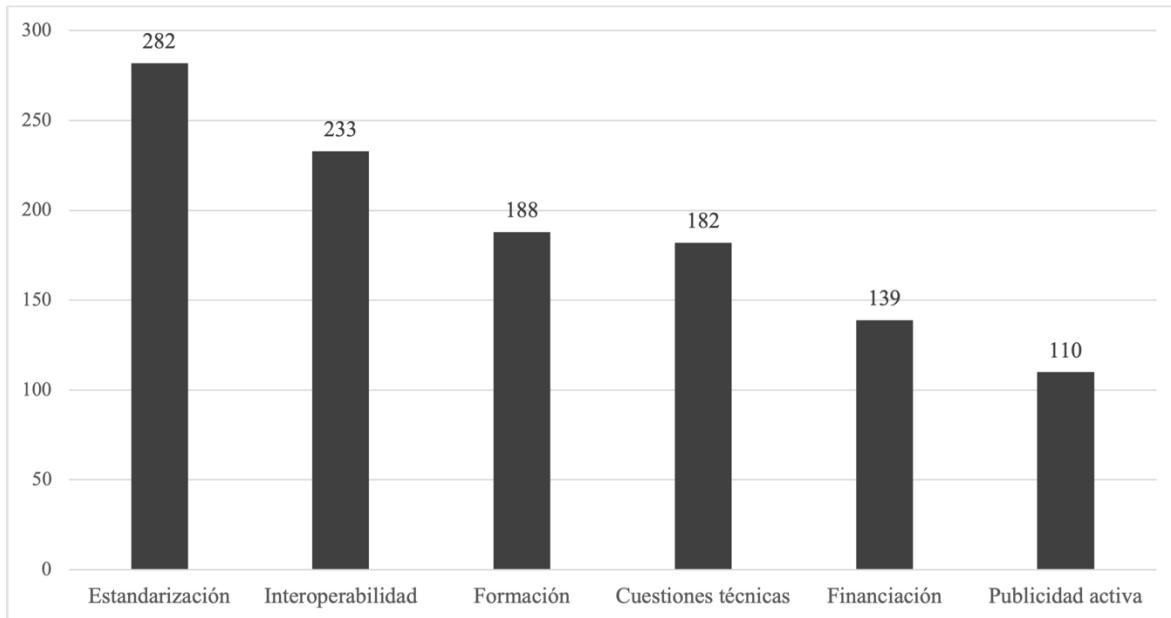


Figura 30: Retos de apertura de datos abiertos en los gobiernos.

Los investigadores consideran, con 282 elecciones, que una de las cuestiones de importancia ante las que se enfrentan los gobiernos en materia de apertura de datos es la falta de estandarización de estos. En 233 ocasiones apuntan a la falta de interoperabilidad de los mismos. La falta de formación en materia de reutilización de datos es un reto elegido en 188 ocasiones seguido de cuestiones técnicas, como la dificultad para encontrar formatos reutilizables, elegido 182 veces. Financiación, elegido en 139 ocasiones y cuestiones de publicidad activa, en 110, son los retos que menos veces han sido seleccionados por los investigadores, aun así suponiendo un % elevado de los que han contestado a esta pregunta de la encuesta.

Se obtienen además 44 respuestas en la opción abierta de las cuales se descartan 4 por no contener información. Tras el proceso de normalización de respuestas, se obtiene la siguiente información:

Reto de apertura	Nº
Falta de voluntad real de informar	11
Falta de especialización técnica	7
Falta de estrategias y diseño	6
Restricciones legales	4
Falta de acercamiento a la ciudadanía	2
Retos de transparencia y rendición de cuentas	2
Interoperabilidad y estandarización	2
Competencia desleal con otros países	1
Cuestiones éticas en la investigación	1

Tabla 48: Retos de apertura de datos abiertos en los gobiernos.

En 11 ocasiones de las 44 de respuesta abierta, los investigadores coinciden que existe una falta de voluntarismo político en materia de querer informar a la ciudadanía y por tanto es un reto al que se enfrentan las administraciones. En 7 ocasiones coinciden en que es otro reto la falta de especialización técnica y 6 veces ha sido elegida la falta de estrategias y diseño por parte de las administraciones públicas. Se señala en 4 ocasiones las restricciones legales ante las que se enfrentan tanto las propias administraciones públicas a la hora de abrir información como el reutilizador a la hora de dar uso a los datos abiertos. En 2 ocasiones señalan los retos de transparencia y rendición de cuentas, la falta de interoperabilidad y estandarización de los datos y la falta de acercamiento a la ciudadanía. Finalmente, en una ocasión señalan como retos, las cuestiones éticas de investigación y la competencia desleal con otros países que no quieren o no llevan a cabo políticas de apertura de datos abiertos.

## Bloque 5. Uso y reutilización de datos (I)

### B-5. Estandarización y limpieza de datos

El bloque 5 está formado por las preguntas 18 y 19 y es la primera parte de dos bloques que se centra en materia de reutilización de datos. La pregunta 18 plantea una serie de enunciados en los que el encuestado puede elegir tantos como haya cumplido.

La pregunta 18 se estandariza con las siguientes palabras clave:

---

Descarga
Plataforma web
Servicio web-API
App móvil
Investigación
Campañas
Integración
Mapas de situación
Visualizaciones
Estadística
Consulta
Divulgación
Seguiré usando
Sin uso
Otros

---

Tabla 49: Categorías de la pregunta 18 de la encuesta.

La pregunta 19 plantea una serie de formatos técnicos donde los encuestados eligen en la siguiente escala de valoración:

- Reutilización avanzada (3).
- Reutilización avanzada con características mejorables (2).
- Reutilización básica (1).
- Formato inadecuado para la reutilización (0).

### B-5. Análisis y representación

La pregunta 18 busca conocer cómo se han reutilizado datos abiertos. Se obtiene respuesta de un total de 387 investigadores y se omite 396 en ocasiones.

Ítem	Total	%
He descargado conjuntos de datos abiertos.	265	68,48
He utilizado datos abiertos en el desarrollo de investigación científica (artículos, informes, presentaciones)	260	67,18
He utilizado datos abiertos para análisis estadísticos	174	44,96
Planeo hacer más uso de los datos abiertos	139	35,92
He utilizado datos abiertos para consultar información pública de interés	109	28,17
He utilizado datos abierto para la creación de visualizaciones	82	21,19
He utilizado datos abiertos en el desarrollo de un sitio web	75	19,38
He utilizado datos abiertos para la creación de mapas de situación	64	16,54
He utilizado datos abiertos para facilitar la divulgación de información pública a la ciudadanía	54	13,95
He utilizado datos abiertos en el desarrollo de un servicio web o API	36	9,30
Nunca he utilizado datos abiertos	30	7,75
He integrado datos del gobierno abierto con una aplicación o sitio web existente	23	5,94
He utilizado datos abiertos en el desarrollo de una aplicación móvil	17	4,39
He utilizado datos abiertos del gobierno abierto para llevar a cabo campañas reivindicativas	10	2,58
Otro (especifique)	5	1,29

Tabla 50: Fin de la reutilización de los datos abiertos.

En la figura 31 se muestran aquellas afirmaciones planteadas en el enunciado y seleccionadas más de 100 veces.

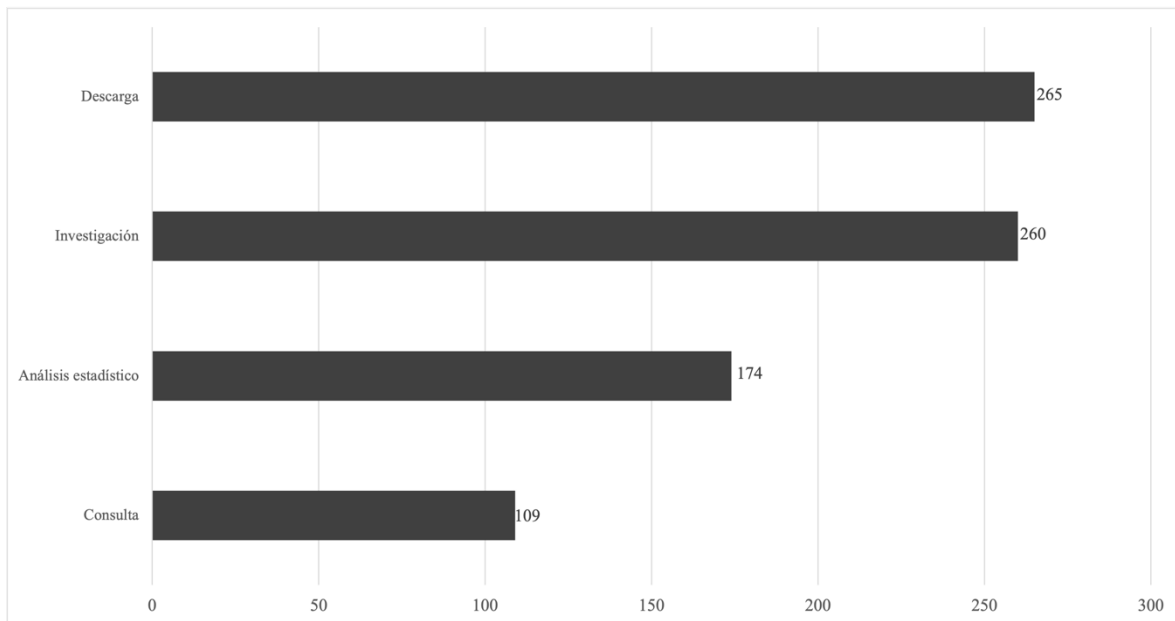


Figura 31: Fin de la reutilización de datos abiertos (+ de 100 elecciones).

Con 265 veces seleccionado de 387 investigadores que han respondido a esta pregunta, el ítem planteado en el enunciado más elegido en materia de uso de datos abiertos es la descarga de conjuntos de datos abiertos. Además, se afirma 260 veces que han utilizado datos para el desarrollo de investigación. Finalmente, en este primer gráfico, 174 señalan que han utilizado datos para análisis estadísticos y 109 han utilizado datos abiertos a modo de consulta de información pública de interés.

La figura 32 muestra las afirmaciones planteadas en el enunciado y elegidas menos de 100 veces.

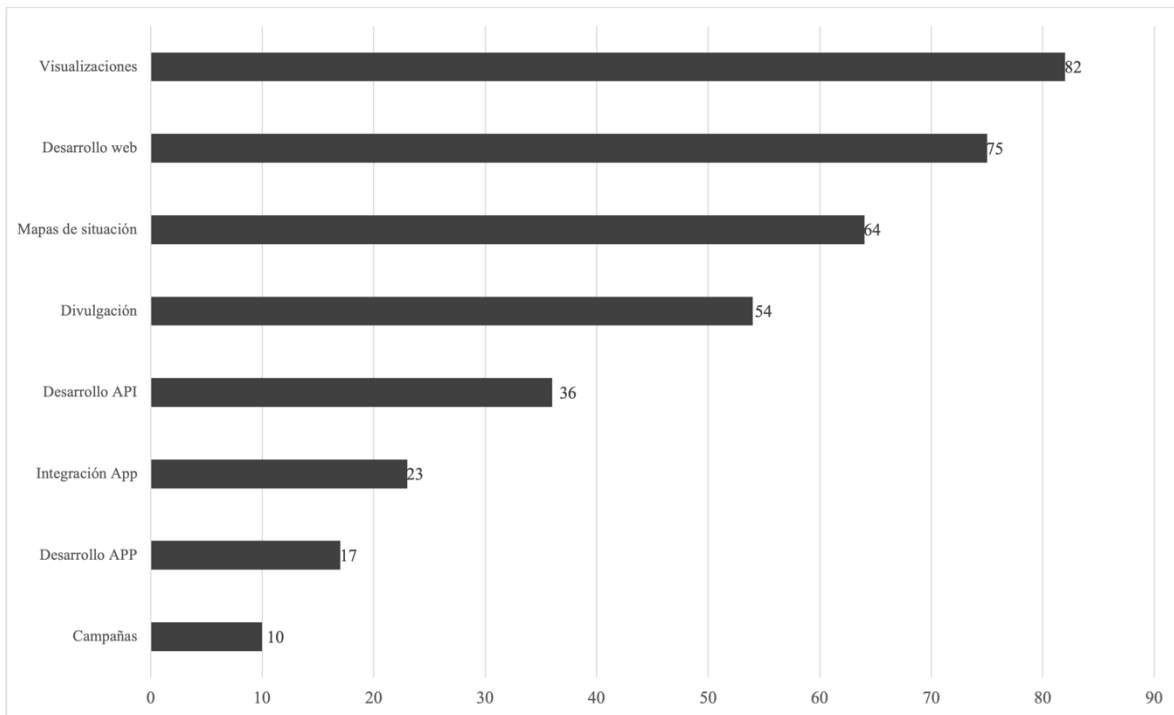


Figura 32: Fin de la reutilización de datos abiertos (- de 100 elecciones).

La forma y uso que se da a los datos abiertos, en 82 ocasiones es para la creación de visualizaciones de datos. En 75 ocasiones, los investigadores señalan que han utilizado datos para el desarrollo de un sitio web. 64 veces señalan que han creado mapas de situación basados en datos. Su uso en la divulgación de información pública en el ámbito científico es señalado en 54 ocasiones. Se ha utilizado durante el desarrollo web a través de API 36 veces, para integración de datos de gobierno de aplicaciones y sitios webs existentes en 23 ocasiones, para el desarrollo de aplicaciones 17 veces y para llevar a cabo campañas reivindicativas un total de 10 veces.

Para finalizar, 139 investigadores afirman que seguirán utilizando datos abiertos e información pública cuando sea necesario, siendo un 82% respecto al 18% (30 investigadores) que no han usado datos abiertos.



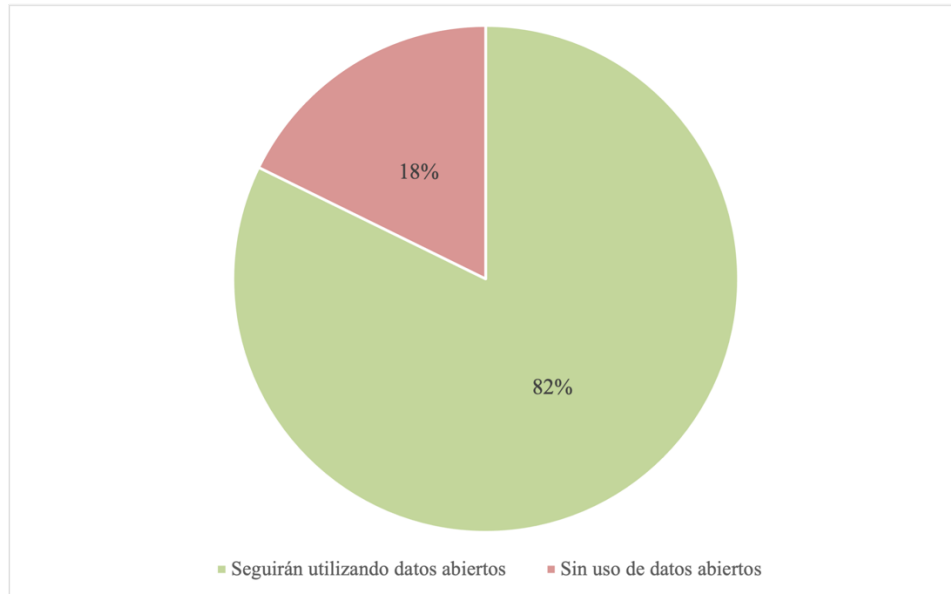


Figura 33: Uso de datos abiertos en el futuro.

En la opción abierta se obtienen 5 respuestas en las que inciden en la accesibilidad a las fuentes de información y en su uso en fines docentes y académicos.

Una parte crucial en la reutilización de información son los formatos técnicos disponibles en las fuentes de información. En la pregunta 19 los investigadores estos formatos según su idoneidad para la reutilización de información con distintos fines.

Formato técnico	Formato inadecuado (0)		Reutilización básica (1)		Reutilización avanzada mejorable (2)		Reutilización avanzada (3)		Total	Media
	%	Total	%	Total	%	Total	%	Total		
CSV	2,76	8	33,45	97	28,62	83	35,17	102	290	1,96
XLS	4,36	13	33,89	101	32,89	98	28,86	86	298	1,86
PDF	49,35	152	31,82	98	10,39	32	8,44	26	308	0,78
XML	6,83	14	31,22	64	38,05	78	23,90	49	205	1,79
JSON	6,87	9	15,27	20	35,11	46	42,75	56	131	2,14
GeoJSON	7,55	8	16,98	18	28,30	30	47,17	50	106	2,15
RDF	8,33	7	28,57	24	29,76	25	33,33	28	84	1,88
KML	8,79	8	25,27	23	36,26	33	29,67	27	91	1,87
SQL	10,16	13	22,66	29	33,59	43	33,59	43	128	1,91
RSS	16,83	17	37,62	38	28,71	29	16,83	17	101	1,46
WMS	13,25	11	28,92	24	34,94	29	22,89	19	83	1,67
WFS	11,69	9	29,87	23	32,47	25	25,97	20	77	1,73
GML	12,66	10	26,58	21	36,71	29	24,05	19	79	1,72
SHP	9,00	9	23,00	23	33,00	33	35,00	35	100	1,94

Tabla 51: Valoración de formatos técnicos.

La pregunta es contestada por 346 investigadores únicos y omitida en 437 ocasiones. Llama la atención el hecho que el formato técnico que más veces ha sido valorado lo ha sido en 308

ocasiones (formato PDF) y el que menos un total de 77 (WFS). Los formatos más valorados han sido aquellos más conocidos y utilizados por la comunidad científica.

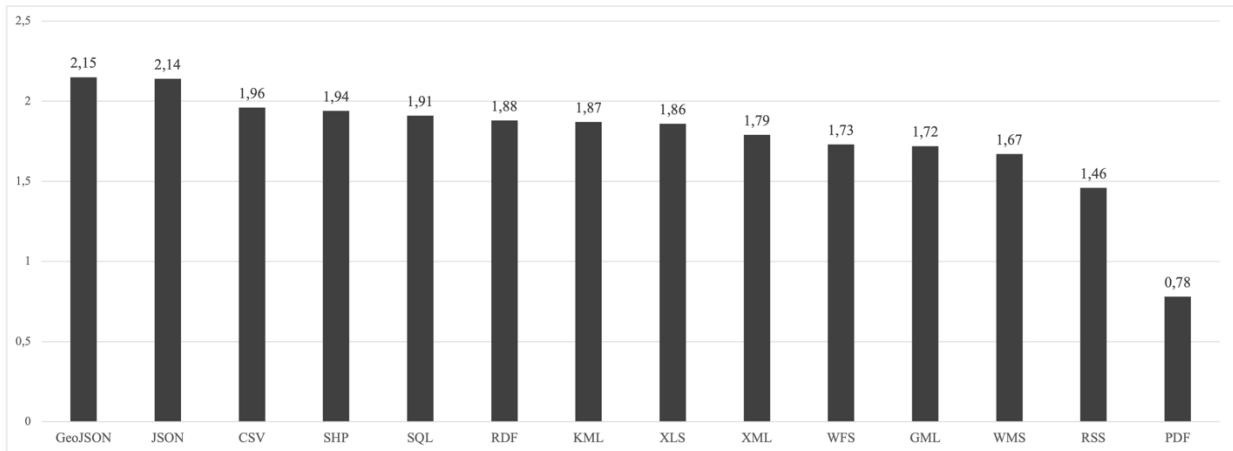


Figura 34: Valoración de formatos técnicos.

Los investigadores han considerado que los formatos con una reutilización avanzada son GeoJSON y JSON, con 2,14 y 2,14 de valoración respectivamente (sobre 3).

Entre 1,90 y 2 puntos de valoración, siendo de reutilización avanzada con características mejorables hay 3 formatos: CSV (1,96), SHP (1,94) y SQL (1,91). Entre 1,50 y 1,90 obtenemos las siguientes valoraciones en formatos: RDF (1,88), KML (1,87), XLS (1,86), XML (1,79), WFS (1,73), GML (1,72) y WMS (1,67).

Finalmente, con menos de 1,50 de valoración siendo inadecuados para la reutilización se obtiene con 1,46 el formato RSS y con 0,78 el formato PDF.

Bloque 6. Uso y reutilización de datos (II)

B-6. Estandarización y limpieza de datos

El bloque 6 está formado por 5 preguntas, de la 20 a la 24 y hace referencia al uso particular que le han dado los investigadores a los datos abiertos. La pregunta 20 es de respuesta abierta y consiste en la descripción de proyectos, investigaciones y uso que les han dado los investigadores a los datos abiertos.

La pregunta 21 tiene una serie de opciones planteadas, en cuanto al objetivo del proyecto/investigación descrita, que se deben responder en caso de que se haya respondido a la pregunta 20. Se codifica de la siguiente forma:

Estadística
Exploración
Campañas políticas
Asuntos sociales
Investigación
Desarrollo web
Desarrollo App
Herramientas internas
Visualizaciones
Mapeos
Tecnologías web
Otros

Tabla 52: Categorías de la pregunta 21 de la encuesta.

Las preguntas 22, 23 y 24 plantean una serie de opciones a elegir por el encuestado en función del uso de datos.

B-6. Análisis y representación

En la pregunta 20 el investigador describe los proyectos llevados a cabo con datos abiertos. Es de respuesta abierta y se responde 203 veces. Se obtienen 209 tipos de uso ya que algunas respuestas se pueden englobar en más de una categoría creada tras la limpieza de respuestas.

Tipo de proyecto	Nº elecciones
Publicaciones científicas y trabajos académicos	107
Proyectos investigación	50
Estudios e informes	32
Investigaciones y proyectos estadísticos	15
Docencia	12
Creación de visualizaciones	6
COVID-19	5
Varios	4
Desarrollo app y web	3
Comunicación y periodismo de datos	3
Consulta	2

Tabla 53: Uso concreto de datos abiertos.

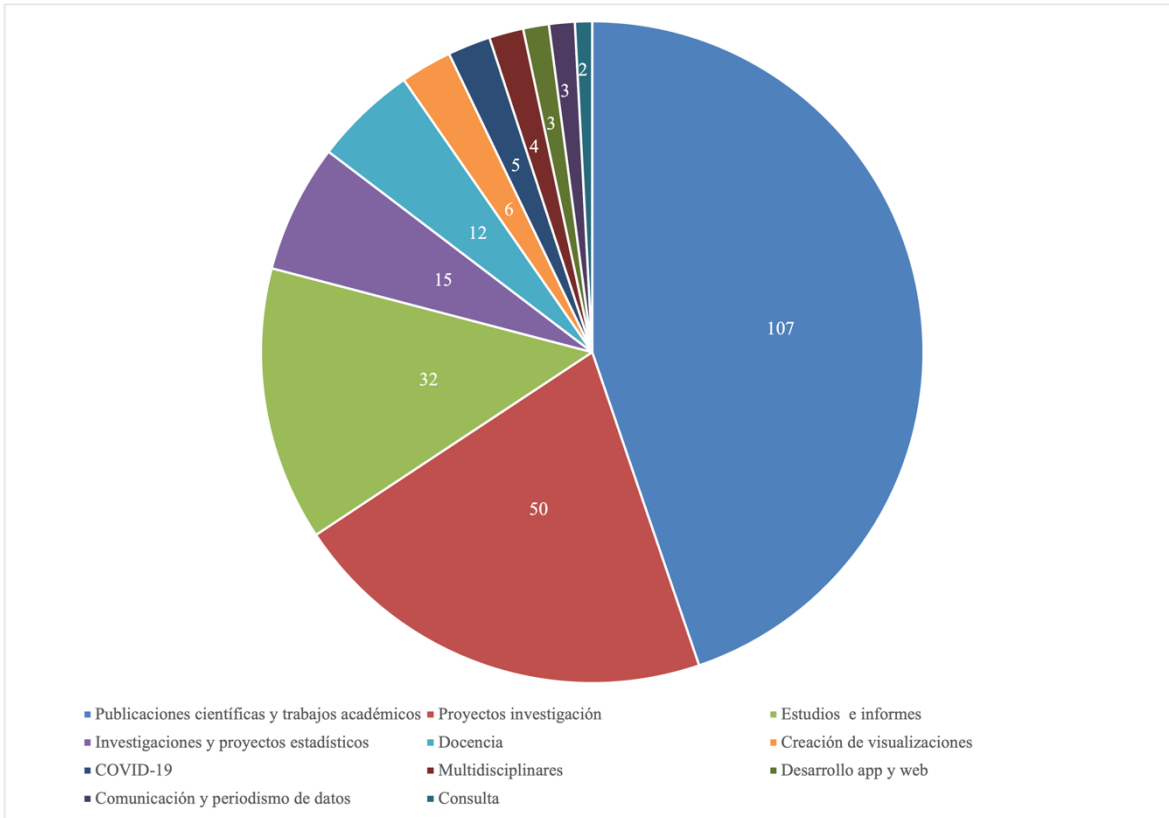


Figura 35: Uso concreto de datos abiertos.

Tras el proceso de estandarización de respuestas se obtiene, con 107 veces seleccionado por los investigadores, que el principal uso de los datos abiertos ha sido destinado a publicaciones científicas y trabajos académicos. Con poco menos de la mitad de respuestas obtenidas, 50, señalan el uso de datos abiertos para proyectos de investigación.

En 32 ocasiones señalan el uso de los datos abiertos para estudios e informes. Con 15 elecciones, se utilizan en investigaciones y proyectos estadísticos, 12 señalan su uso con fines docentes.

Con menos de 10 elecciones tenemos: con 6, uso para la creación de visualizadores de datos. Con 5, fines orientados a información acerca del COVID-19. Con 4, usos multidisciplinares. Con 3 encontramos, desarrollo de aplicaciones y sitios web y, además, comunicación y periodismo de datos y finalmente, con 2, uso a modo de consulta para satisfacción de necesidades de información.

La pregunta 21 está planteada para que la respondan aquellos que han contestado previamente a la 20. Se trata de describir el objetivo del proyecto anterior. La pregunta 20 la responden 203 personas y la 21 se responde en consecuencia a la pregunta anterior, obteniéndose un total de 233 respuestas.

Objetivo	Nº	%
Literatura científica: informes, artículos	112	48,07
Investigación estadística	110	47,21
Exploración de datos	87	37,34
Creación de visualizaciones	36	15,45
Desarrollo de mapeos	30	12,88
Otro (especifique)	27	11,59
Asuntos sociales	24	10,30
Desarrollo de un sitio web	15	6,44
Desarrollo de herramienta interna para mi organización	13	5,58
Desarrollo de app	9	3,86
Desarrollo API	6	2,58
Investigar las tecnologías web	6	2,58
Campañas políticas	2	0,86

Tabla 54: Objetivo del proyecto descrito en la pregunta 20.

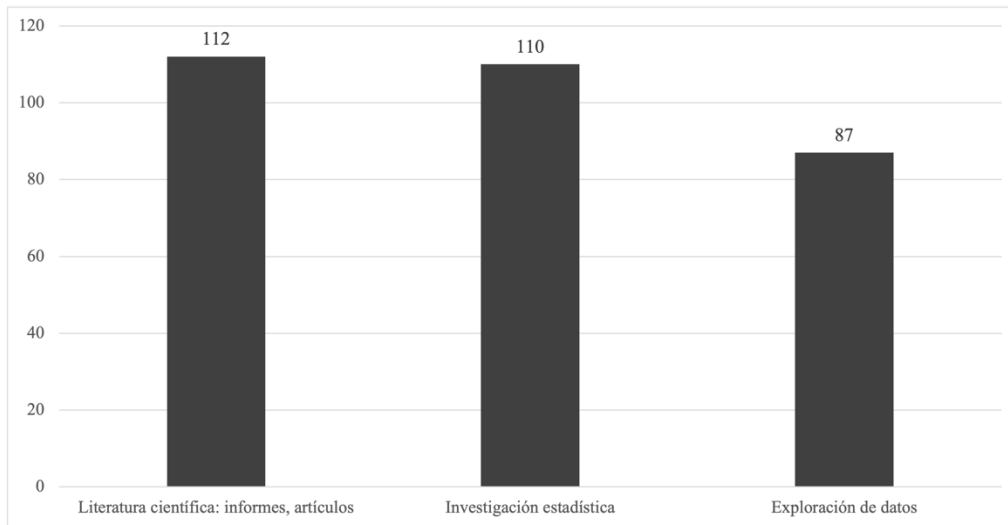


Figura 36: Objetivo del proyecto descrito en la p. 20 (+ de 50 elecciones).

La figura 36 muestra los objetivos más elegidos para el uso de datos abiertos en los proyectos descritos en la pregunta 20.

Con 112, siendo la opción más elegida, fue el uso de datos abiertos para la creación de literatura científica, seguido, con 110, para investigaciones con fines estadísticos y con 87, para exploración de datos.

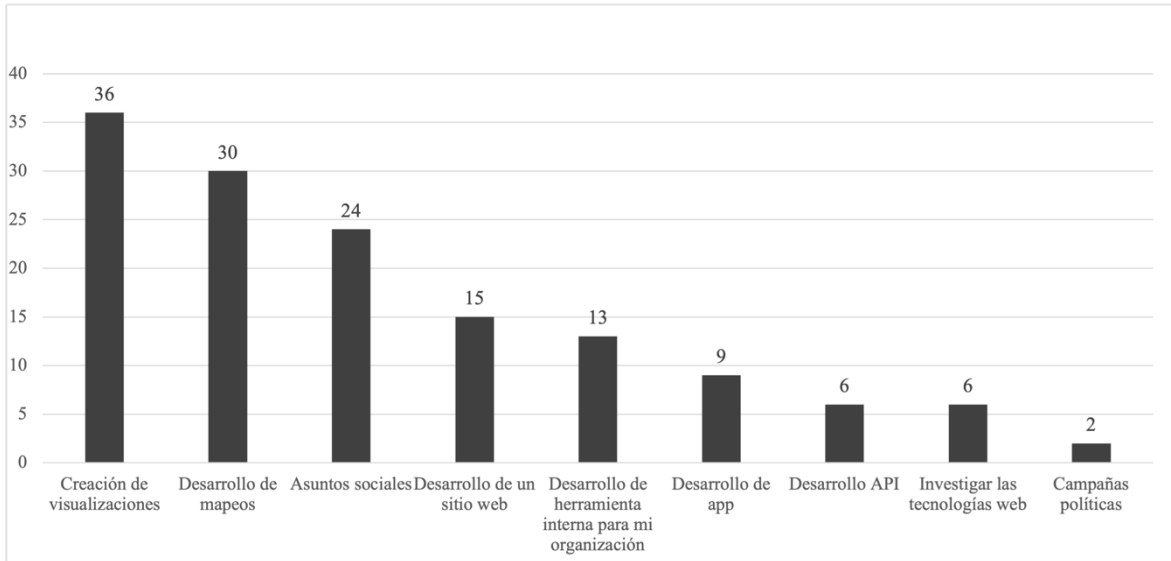


Figura 37: Objetivo del proyecto descrito en la p. 20 (- de 40 elecciones).

La figura 37 muestra los objetivos seleccionados en menos de 40 ocasiones. Con 36 elecciones, el objetivo fue la creación de visualizaciones de datos. Le sigue con 30 el desarrollo de mapeos. Con 24 se destina a asuntos sociales. En 15 ocasiones los investigadores señalan el uso para desarrollo de sitios web, con 13 para desarrollo de herramientas internas de organización, con 9 para el desarrollo de aplicaciones y con 6 para el desarrollo de API's. Finalmente, con 6 también, se destina para investigar tecnologías web y con 2 para campañas políticas.

En la opción abierta se obtienen 27 respuestas de las cuales descartamos 2 por no contener información. Se normaliza en las categorías siguientes tras el proceso de limpieza de datos:

Ítem	Nº
Investigación y ciencia	11
Informes, estudios y planes	8
Docencia	5
Desarrollo	1

Tabla 55: Objetivo del proyecto descrito en la p.20 (Respuesta abierta).

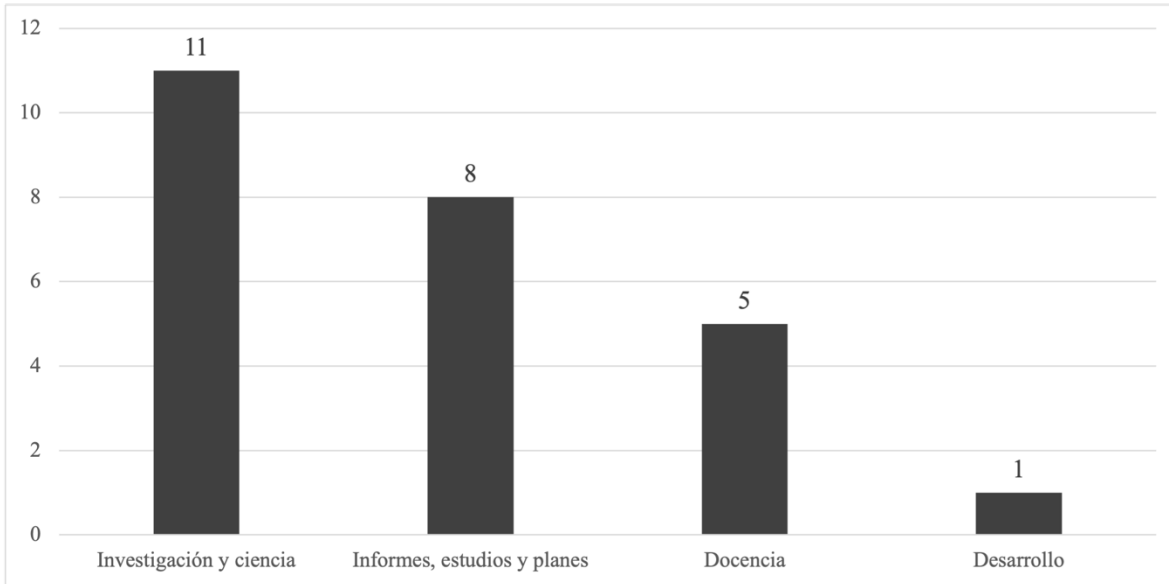


Figura 38: Objetivo del proyecto descrito en la p. 20 (respuesta abierta).

En 11 ocasiones, de forma abierta los investigadores vuelven a incidir en objetivos destinados a investigación y ciencia. Con 8 elecciones, uso destinado al desarrollo de informes, estudios y planes. Con 5, uso en fines docentes y finalmente con 1 elección, para desarrollo tecnológico.

La pregunta 22 se centra en el uso de datos abiertos en proyectos. El objetivo es conocer los ámbitos en los que se reutilizan los datos abiertos, siendo internacionales, continentales, estatales, autonómicos y locales. Es contestada por 245 investigadores y omitida en 538 ocasiones. Se obtiene la siguiente información:

Ámbito	Nº	%
Estatales	138	56,33%
Autonómicas	127	51,84%
Internacionales	118	48,16%
Locales	75	30,61%
Continental	33	13,47%

Tabla 56: Ámbito de uso de datos abiertos.

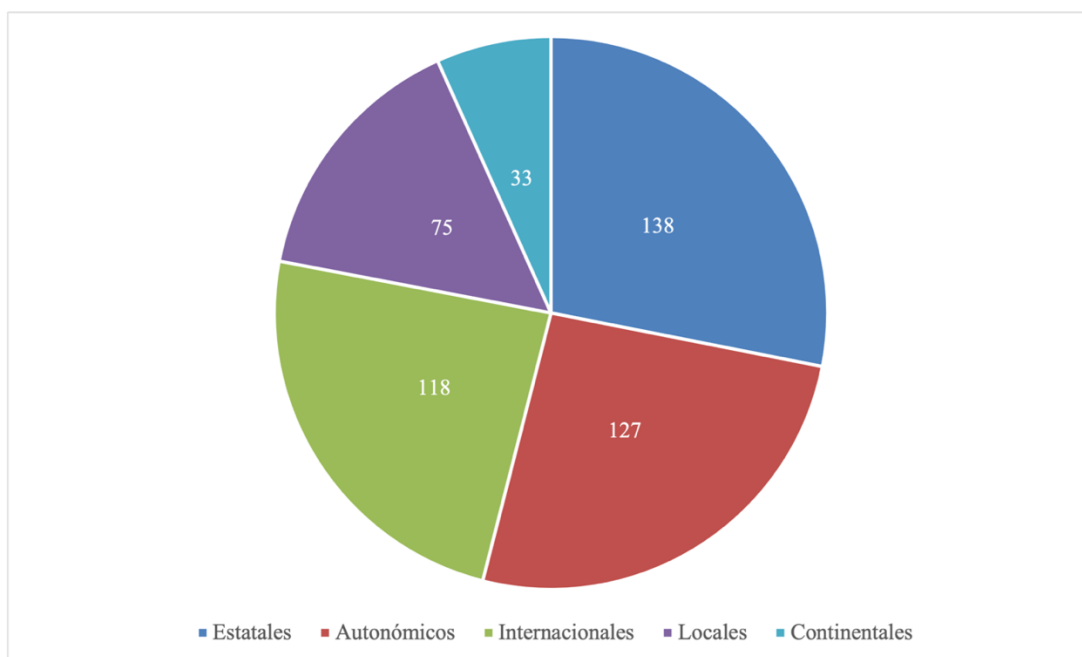


Figura 39: Ámbito de uso de datos abiertos.

El 56,33% de los investigadores que contestan a esta pregunta, siendo un total de 138 respuestas, utilizan datos abiertos en proyectos de carácter estatal. Le sigue con 127 elecciones (51,84%) proyectos centrados en ámbitos autonómicos. Con 118 elecciones (48,16%), uso de datos abiertos en proyectos internacionales. Con 75, el ámbito de uso de los datos abiertos en contexto locales y finalmente, con 33 elecciones (13,47%), se reutiliza información del sector público en proyectos continentales.

La pregunta 23 busca conocer aquellos conjuntos de datos utilizados para llevar a cabo el desarrollo de proyectos e investigaciones. Es contestada por 248 investigadores siendo omitida en 535. Se obtiene la información de la siguiente tabla:

Tipos de datos	Nº Total	%
Datos sobre el medio ambiente y el clima	78	31,45%
Población	74	29,84%
Geográficos	66	26,61%
Salud y discapacidad	54	21,77%
Datos sobre educación y habilidades	44	17,74%
Otro (especifique)	44	17,74%
Economía y comercio	41	16,53%
Mercado laboral	33	13,31%
Transporte	32	12,90%
Datos comerciales	22	8,87%
Vivienda	22	8,87%
Datos jurídicos y administrativos	21	8,47%
Ingresos y gastos del gobierno	19	7,66%
Políticos	19	7,66%
Datos sobre el crimen y la justicia	13	5,24%



Ninguno de los anteriores	9	3,63%
Análisis web	7	2,82%
Ayuda internacional y desarrollo	5	2,02%

Tabla 57: Tipos de datos utilizados en proyectos e investigaciones.

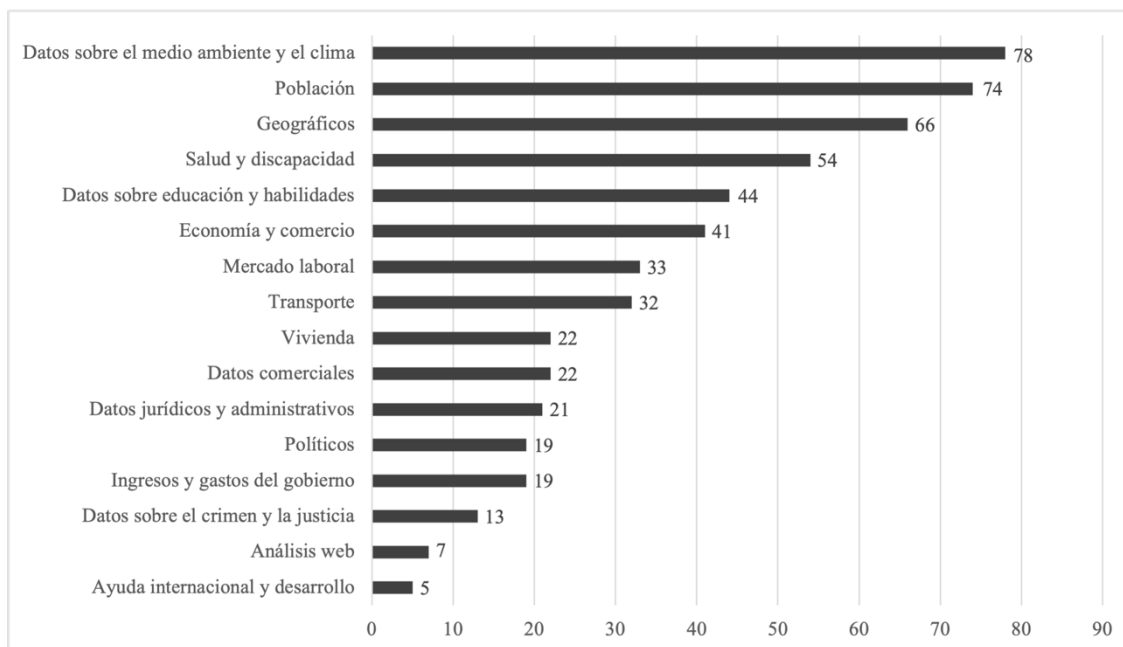


Figura 40: Tipos de datos utilizados en proyectos e investigaciones.

Con 78 elecciones, los conjuntos de datos más utilizados por los investigadores en los proyectos de investigación fueron los relacionados con medio ambiente y clima, seguido por los 74 tipos conjuntos de datos sobre población y los 66 conjuntos de datos geográficos.

Datos relacionados con salud y discapacidad son elegidos en 54 ocasiones, educación y habilidades 44 veces, economía y comercio en 41 ocasiones, mercado laboral en 33 y transporte en 32.

Con menos de 25 elecciones, los investigadores reutilizan datos sobre vivienda y datos comerciales en 22 ocasiones y datos jurídicos y administrativos en 21. Le siguen datos políticos y de ingresos y gastos del gobierno (19 elecciones), datos sobre crimen y justicia (13), datos acerca del análisis web (7) y de ayuda internacional y desarrollo (5).

En la respuesta abierta se obtienen 46 conjuntos de datos. Se presentan en la tabla 58:

Conjuntos de datos	Total
I+D, innovación, ciencia y tecnología	6
Energía	4
Finanzas y economía	4
Turismo	4
Académicos y de universidades	3
Salud (pandemia Covid19)	3
Varios	3
Cultura	2
Datos históricos, catálogos de colecciones artísticas y bienes patrimoniales	2
Demográficos	2
Deporte	2
Edificabilidad y usos	2
Accidentes de tráfico	1
Alimentación	1
Cartografía digital	1
Climáticos	1
Desempeños políticos administrativos y rendimiento de estructuras políticas	1
Información genómica	1
Materiales	1
Satélites	1
Sísmicos	1

Tabla 58: Tipos de datos utilizados en proyectos e investigaciones.

Destaca, con 6 elecciones por parte de los investigadores, datos relacionados con innovación, ciencia y tecnología. En 4 ocasiones, investigadores utilizan conjuntos de datos relacionados con energía, finanzas y economía y turismo. Con 3, datos académicos y de universidades, datos sobre salud relacionados con el COVID-19 y datos multidisciplinarios. Finalmente, con 2 elecciones, apuntan a conjuntos de datos históricos y culturales, demográficos, de deporte y sobre edificabilidad y con 1 elección, datos sobre alimentación, accidentes de tráfico, cartografía digital, materiales, satélites, información genómica y desempeños políticos.

La última pregunta de la encuesta, la 24, se centra en lenguajes, programación y tecnología utilizados para el análisis y tratamiento de datos. Es contestada 254 veces y omitida 529. Se obtiene la siguiente información:

Tecnología utilizada	Total	%
Hojas de cálculo	191	75,20
Software de estadística	114	44,88
Procesadores de texto	91	35,83
Herramientas de visualización de datos	81	31,89
Lenguajes de programación (PHP, Python,Ruby...)	65	25,59
Herramientas de mapeo de datos	55	21,65
XML	35	13,78
Otro (especifique)	28	11,02
JSON	27	10,63
Suites como Google	22	8,66
SPARQL	8	3,15
Flash	4	1,57

Tabla 59: Tecnología utilizada para el tratamiento y análisis de datos.

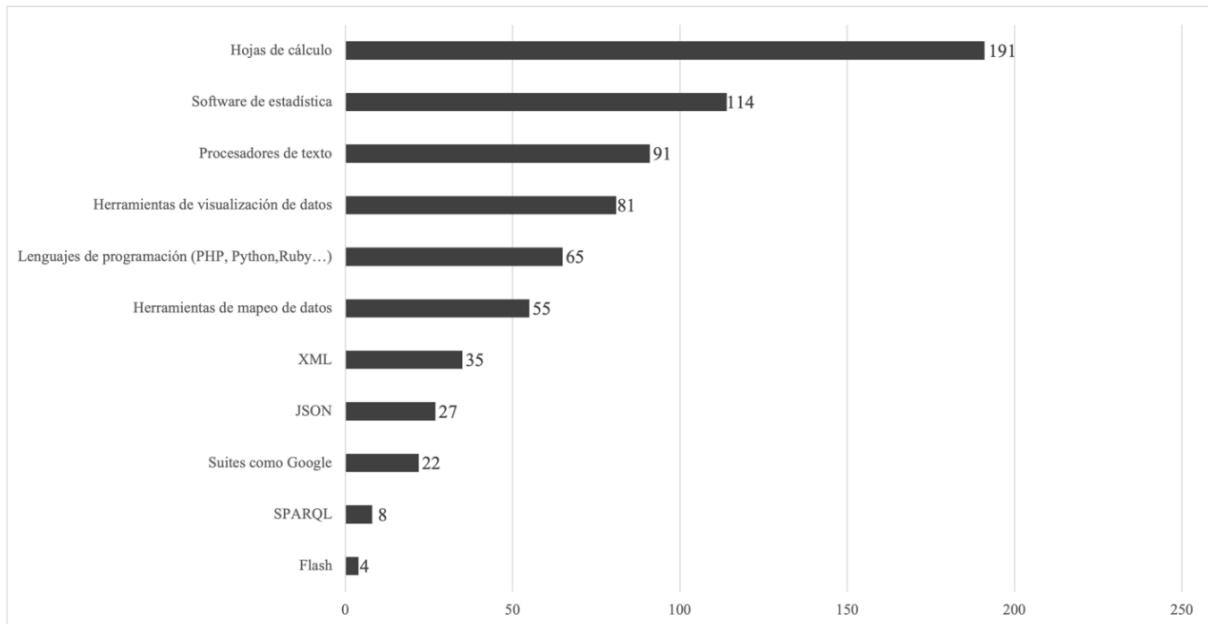


Figura 41: Tecnología utilizada para el tratamiento y análisis de datos.

Destaca habiendo sido elegido por 191 investigadores, las hojas de cálculo como tecnología para el tratamiento y análisis de datos. 114 investigadores utilizan softwares de estadística, 91 utilizan procesadores de texto, 81 herramientas de visualización de datos.

En 65 ocasiones, señalan el uso de lenguajes de programación, 55 utilizan herramientas de mapeo de datos. El uso de lenguajes como XML y JSON son elegido en 35 y 37 veces respectivamente. Finalmente, 22 investigadores señalan el uso de la suite de Google, el lenguaje SPARQL 8 veces y la tecnología Flash, 4.

Esta pregunta tiene una opción de respuesta abierta en la que se obtienen 28 aportaciones. Se descarta una por no aportar información. Se obtiene la información siguiente:

Tecnología	Total
Lenguajes programación: R, Python, C, C++, otros	9
Sistemas de información geográfica (GIS)	5
MATLAB	4
Bases de datos: Mongo DB, Apache	3
Open Refine	2
Software propio	1
Cuadros de mando y relacionales.	1
Informes ya escritos por otras entidades.	1
Análisis documental.	1

Tabla 60: Tecnología utilizada para el tratamiento y análisis de datos.

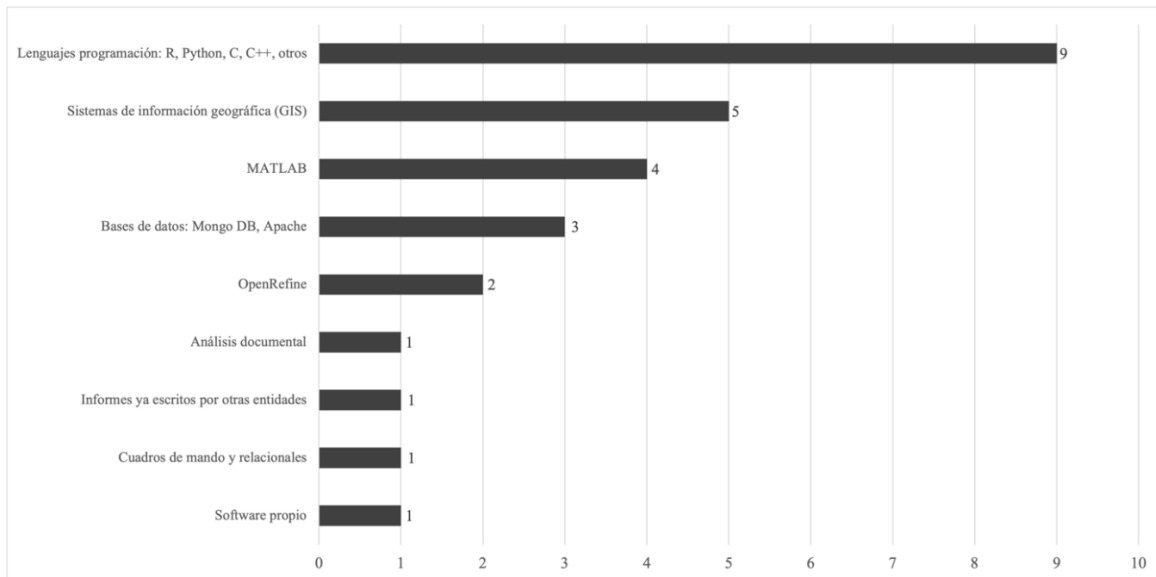


Figura 42: Tecnología utilizada para el tratamiento y análisis de datos.

Algunas respuestas obtenidas por los investigadores tienen cabida en las opciones planteadas en el enunciado. En 9 ocasiones señalan el uso de lenguajes de programación como R, Python o C. Señalan el uso de sistemas de información geográfica en 5 ocasiones. MATLAB es indicado por 4 investigadores. Bases de datos como MongoDB o Apache son señaladas en 3 ocasiones. Tratan datos con Open Refine 2 investigadores y finalmente, con una elección, hablan de cuadros de mando, informes escritos por otras entidades, softwares propios y análisis documental.

#### Desarrollo: Plataforma de visualización de datos

Con los datos obtenidos en la encuesta, además de realizar el análisis sobre la percepción, uso y reutilización de los datos abiertos en la ciencia española, se desarrolla una plataforma web para la representación de los mismos.

Está desarrollada en el lenguaje de programación R<sup>84</sup>, orientado al análisis estadístico y matemático, que permite manipular y procesar grandes volúmenes de datos, mediante la creación de visualizaciones, *dashboards*, informes automáticos, entre otras funcionalidades (Mendoza Vega, 2016). La plataforma está desarrollada en *Shiny*<sup>85</sup>, paquete de R que permite construir páginas interactivas a partir de los scripts de R.

Se puede consultar en el siguiente enlace: <https://indicame.upv.es/Encuestas/>

---

<sup>84</sup> R-Project: <https://www.r-project.org/>

<sup>85</sup> Shiny: <https://shiny.rstudio.com/>

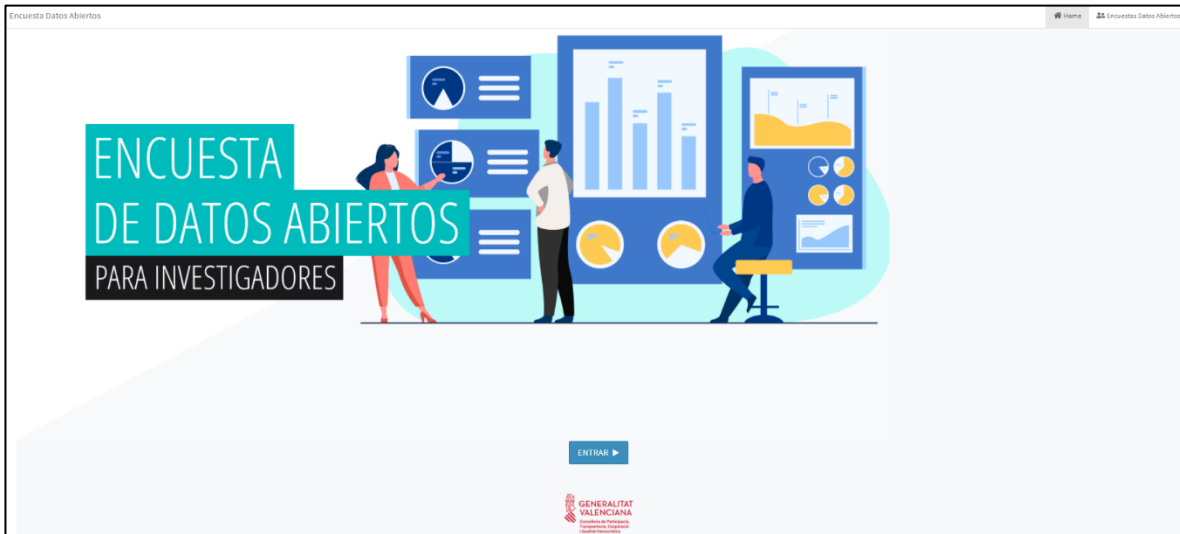


Figura 43: Página principal de la plataforma.

La información de la plataforma está estructurada igual que en la encuesta. Tiene una página principal y se puede navegar y visualizar la información a través de los 6 bloques temáticos.

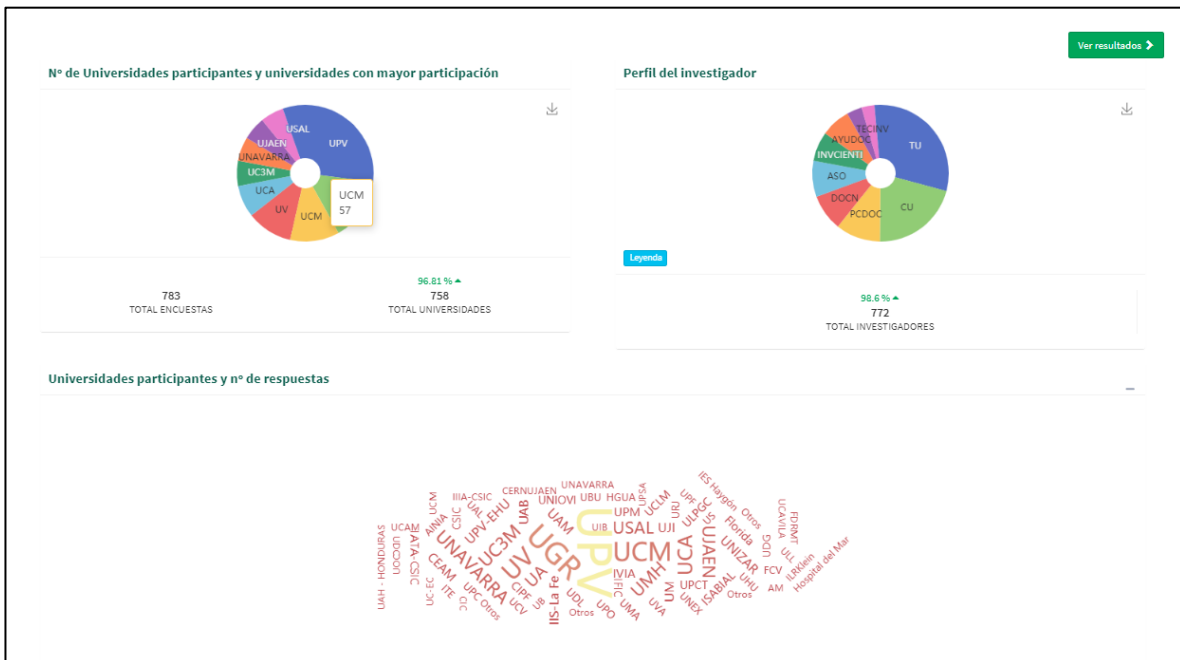


Figura 44: Visualizaciones de datos en página principal.

Encuesta de Datos Abiertos



Figura 45: Bloques de la encuesta disponibles en la plataforma.

Encuesta de Datos Abiertos

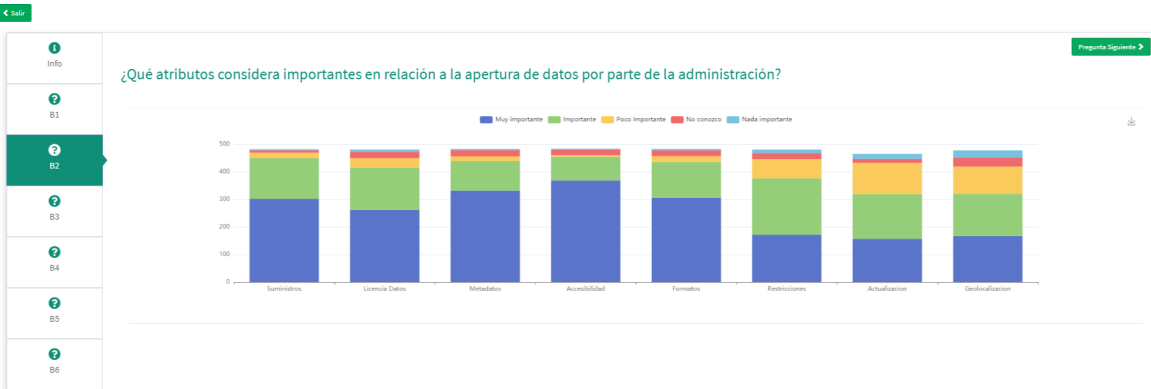


Figura 46: Ejemplos de visualizaciones de datos disponibles en la plataforma.

## Capítulo 5

# Discusión



## Discusión

### 5.1 Introducción

Como se ha dicho anteriormente, uno de los objetivos de las iniciativas de apertura de datos dentro del movimiento de gobierno abierto, a parte del fomento transparencia y rendición de cuentas de las entidades públicas con la ciudadanía, es que esa información disponible en los portales de datos abiertos sea reutilizada para generar servicios de valor añadido (Garriga-Portolà, 2011). El éxito de este tipo de políticas de apertura radica en el uso y la percepción que se tiene acerca de los datos.

Se hace necesaria una comunicación y sinergia entre las entidades que llevan a cabo la apertura de datos y la ciudadanía, potenciales reutilizadores de información del sector público. Por un lado, las administraciones deben atender a las necesidades del ciudadano tratando de conocer los conjuntos de datos que son de relevancia e interés para el desarrollo de investigaciones o productos y servicios innovadores. A su vez, deben ofrecer datos que puede que la población no perciba como necesarios pero que, a través de su consulta o navegación a través de los portales de datos abiertos, generen ideas que den paso a nuevos desarrollos.

Por tanto, no solo se trata de llevar a cabo iniciativas de apertura de datos abiertos, sino tener en cuenta qué necesita la ciudadanía, cuál es la percepción, si las fuentes de información pública son de relevancia para sus estudios y los requisitos técnicos y de contenido que deben tener los datos para que estos sean reutilizables. Si nos fijamos en la literatura científica, existen multitud de estudios que evalúan de diferentes formas la calidad de los conjuntos de datos abiertos y su nivel de reutilización en base a indicadores (Abella et al., 2017 y 2019), el sector infomediario y el valor que generan los datos (Asedie, 2021), estudios de conjuntos de datos concretos aplicados a un área determinada (Schmidt, Gemeinholzer y Treolar, 2016), o la percepción que tiene la ciudadanía en general acerca del *open data* (Davies, 2010; Estrada Zamora y Vargas-Hernández, 2018). En 2020 se realizó una encuesta dentro del marco del proyecto de Datasea, relacionada con los datos de investigación, con el objetivo de reconocer hábitos y experiencias de los investigadores españoles en ciencias de la salud en relación con la gestión e intercambio de datos brutos de investigación (Aleixandre-Benavent, Vidal-Infer, Alonso-Arroyo, Peset y Ferrer-Sapena, 2020). Estos estudios, de gran valor, utilidad y pioneros, se completan abordando una serie de perspectivas generales que permiten observar de un modo panorámico situaciones concretas de los datos abiertos. En algunas ocasiones, dentro de estos estudios, participan miembros de la comunidad científica, pero no son metodologías de investigación enfocadas directamente a la comunidad académica (salvo el último citado de Aleixandre-Benavent et al.). Aquí el primer aporte de esta tesis doctoral. Este estudio se centra específicamente en la comunidad científica y universitaria, personal altamente cualificado dentro de la sociedad. En aras de mejorar estos servicios o fuentes de información, los investigadores han respondido a la encuesta planteada

en la metodología de investigación. A través de los resultados obtenidos en el cuestionario (con opciones de respuesta abierta donde aportar críticas, inquietudes y otros comentarios), las instituciones encargadas de la apertura de datos pueden respaldarse en este estudio, teniendo en cuenta opiniones, ciertos patrones de conducta y la percepción que tienen la comunidad científica española.

El presente capítulo se estructura en dos partes. En primer lugar, se evalúa la metodología utilizada para determinar las limitaciones y alcances de la investigación. Este análisis nos va a permitir preguntarnos si los resultados obtenidos son extrapolables a una población mayor y reflexionar acerca de ella.

A continuación, se responde a las preguntas de investigación planteadas al inicio a través del cruce de datos de las distintas variables presentadas en el capítulo de resultados. Se establece así una relación entre los resultados obtenidos a través de la metodología cuantitativa con la literatura científica revisada en el marco teórico. A diferencia del capítulo anterior donde se exponen los resultados, en este, se realiza un ejercicio de pensamiento crítico para argumentar y valorar el aporte estos. El análisis se lleva a cabo desde varias perspectivas acordes con los bloques planteados en la encuesta. Estos son el tipo de investigador, conocimiento, uso y reutilización de los datos abiertos en la investigación, la percepción y utilidad real de los datos abiertos y finalmente, el acceso a fuentes de información y su uso en proyectos.

## 5.2 Limitaciones de la investigación

### 5.2.1 Limitaciones en la metodología de investigación: el estudio cuantitativo

En el área de las ciencias sociales, existen diferentes tipos de investigación y según la naturaleza de la información que se va a obtener para responder a las preguntas planteadas para resolver un problema de investigación, esta puede ejecutarse bajo dos paradigmas: el cualitativo y el cuantitativo (Sarduy Domínguez, 2007).

En esta tesis doctoral, las preguntas de investigación se han resuelto a partir del diseño de una encuesta, siendo un método cuantitativo, sujeto a limitaciones. No obstante, el uso de un método u otro tendrá ventajas y desventajas, sea cual sea la elección del método.

El propósito de los estudios cuantitativos es explicar, predecir y controlar los fenómenos a investigar a través de la obtención de datos. La primera limitación, a diferencia del método cualitativo, es que no nos va a permitir la obtención de datos narrativos extensos y, por tanto, no vamos a poder obtener un conocimiento profundo sobre el fenómeno a investigar (De Pelekais, 2000; Mills y Gay, 2019).

El diseño de los estudios cuantitativos es estructurado e inflexible y no va a permitir estudiar casos concretos de sujetos en profundidad. Además, la recolección de datos es a través de

observación no participativa. Al contrario que en los métodos cualitativos, donde los datos se obtienen a partir de una observación participativa, entrevistas que pueden ser informales y sin estructura, donde se pueden obtener notas de campo detalladas o extensas, a través de revisiones y análisis documentales, estudios de caso o los conocidos grupos focales (De Pelekais, 2000; Mills y Gay, 2019).

La investigación cualitativa centra su atención en entender a la persona como un todo, ligado a un contexto social, para comprender y dar significado al mundo de relación e interacción de las personas, además de sus motivaciones, expectativas entre otros. Además, exige el reconocimiento de múltiples realidades y trata de capturar la perspectiva del investigado. Es decir, tanto el sujeto como la fuente a investigar cobran importancia ya que los resultados tienen una fuerte dependencia de las emociones y análisis del contenido de las fuentes de información. (Mira, et al., 2004; Sarduy Domínguez, 2007).

Por tanto, en los estudios cualitativos, el investigador no solo trata de describir hechos a investigar, sino de comprender las actividades, motivaciones, valores y otros significados subjetivos mediante un análisis de la información obtenida. Es decir, se estudia los contextos situacionales y estructurales, al contrario que los estudios cuantitativos, que no van a permitir extraer al máximo toda la información ya que la investigación cualitativa establece una mejor comunicación y más profunda con los sujetos (Sarduy Domínguez, 2007).

Sin embargo, toda posición extrema es limitante para el avance. Podemos considerar ambos métodos perfectamente complementarios, siempre y cuando sea necesario y el tipo de investigación lo permita (De Pelekais, 2000). Incluso teniendo en cuenta las consideraciones expuestas, los resultados que se obtienen en esta tesis doctoral son extrapolables a otros países que quieran conocer o medir cuantitativamente la percepción de su comunidad científica ante los datos abiertos y la importancia de estos en el mundo académico y de investigación. El método cuantitativo es un método objetivo que nos va a permitir medir o cuantificar la relevancia del fenómeno en base a datos objetivos confirmativos e hipotético-deductivos.

Esta metodología nos ha permitido obtener una muestra representativa, a nivel estatal (783 respuestas a la encuesta) y, por tanto, los resultados obtenidos son extrapolables a otros territorios.

Se obtiene una primera visión panorámica en el ámbito científico-académico en nuestro país, acerca de la percepción sobre los datos abiertos procedentes de administraciones públicas y gobiernos. Nos permite conocer a indagar de qué modo son utilizados y qué requisitos deben cumplir los portales de datos abiertos para que miembros de la academia y otros investigadores reutilicen la información del sector público para el desarrollo de estudios e investigaciones innovadores.

### 5.2.2 Limitaciones de la encuesta en el desarrollo de la investigación

Tal y como se describe en el apartado de la metodología de la investigación, la encuesta está estructurada en un total de 7 bloques: uno de contexto del científico con 6 preguntas y 6 bloques de investigación con 24 preguntas. En total, 30 preguntas.

Puede considerarse una encuesta extensa y con cierta complejidad, ya que, a medida que avanzan las preguntas, estas se especializan y requieren de ciertos conocimientos técnicos y estar familiarizado en el movimiento de los datos abiertos. La encuesta trata obtener respuestas para abordar las preguntas de investigación planteadas desde las 6 perspectivas planteadas en los bloques de investigación. De ahí su extensión.

La primera pregunta<sup>86</sup> de la encuesta, acerca del conocimiento de los datos abiertos es la única cuya respuesta es obligatoria en el cuestionario. Este permite que, aquellos encuestados que no saben qué son los datos abiertos terminen la encuesta en ese momento, ya que, como se ha mencionado con anterioridad, las preguntas toman cierta dificultad que requieren de una familiarización con los datos abiertos y la reutilización de información del sector público.

El hecho que, de las 6 preguntas de contextualización, y las 24 preguntas de carácter técnico que conforman el cuestionario, solo haya 1 de respuesta obligatoria, hace que muchos investigadores no contesten a determinadas preguntas en función de su dificultad o interés. No hacer las preguntas de respuesta obligatoria es una decisión que tiene como objetivo no influir al investigador y que este se sienta en libertad absoluta de responder aquellas preguntas donde pueda aportar información que pueda ser de utilidad en la investigación.

Por tanto, y como consecuencia de esta decisión inicial de no hacer las preguntas de obligada respuesta, se detecta que las primeras preguntas tienen un alto porcentaje de respuestas (ver capítulo de resultados), disminuyendo notablemente en las últimas preguntas del cuestionario la tasa de respuesta.

Otro elemento que cabe destacar es el índice de respuestas en aquellas opciones de respuesta abierta. Hay un total de 18 opciones de respuesta abierta en la encuesta donde el investigador puede realizar comentarios y aportaciones. De las opciones de respuesta abierta, 5 son planteadas como pregunta y 13 forman parte de preguntas donde se plantean opciones dejando un espacio abierto donde el investigador puede realizar apreciaciones.

Un elemento para tener en cuenta en las metodologías de investigación cuantitativas, y que puede hacer que la tasa de respuestas no sea lo suficientemente representativa es el sesgo de

---

<sup>86</sup> Pregunta 1 de la encuesta: ¿Sabe qué son los datos abiertos? (Entendidos como filosofía que persigue que determinados tipos de datos estén disponibles de forma libre sin restricciones, en bruto, estructurados y en formatos que faciliten la reutilización).

resultados debido a la procedencia del encuestado. Es decir, que únicamente respondan al cuestionario aquellas personas que sea su campo de investigación o interés. Tal y como se muestra en el apartado de resultados, la muestra obtenida es variada y representativa, tanto por áreas de investigación, género de encuestados, rangos de edad y universidades o centros de investigación a las que están vinculados. Por tanto, esta posible limitación del estudio no se contempla gracias a la variedad y procedencia de los encuestados.

### 5.3 Análisis

#### 5.3.1 Perfil del investigador: conocimiento y reutilización de datos abiertos

En esta parte del estudio se analiza el perfil científico de quienes responden a la encuesta y el nivel de conocimiento y reutilización de datos en la carrera académica y de investigación. Nos sirve para contextualizar qué tipo de investigadores conocen y reutilizan información del sector público y con qué fines.

Tal y como se ha presentado en el capítulo de resultados, se obtienen un total de 783 respuestas, de las cuales, en materia de género, 443 han sido contestadas por hombres (un 56,58%), 327 han sido respondidas por mujeres (41,76%), 3 por intersexuales<sup>87</sup> (0,38%) y han omitido la respuesta 10 personas (1,28%). A pesar de que se trata de una metodología cuantitativa de investigación para esta tesis doctoral, teniendo en cuenta la Ley Orgánica 3/2007, del 22 de marzo, para la igualdad efectiva de mujeres y hombres, España dispone de medidas legales que garantizan la representación de las mujeres en las instituciones, comprometiéndose con la democracia paritaria, es decir, con una representación equilibrada de hombres y mujeres. Se traduce en un porcentaje de representación mínima para cualquiera de los dos sexos del 40% (Mestre, 2008).

---

<sup>87</sup> Se obtienen únicamente 3 respuestas de intersexuales, por lo que se considera una muestra residual. El análisis en este apartado se centra en los géneros femenino y masculino. Como curiosidad, los 3 intersexuales conocen y han reutilizado datos abiertos. Proceden de las áreas de investigación 2, 6 y 9 (Derecho, Historia y Matemáticas). Rango de edad de 41 a 50 (1 caso) y de 51 a 60 (2 casos).

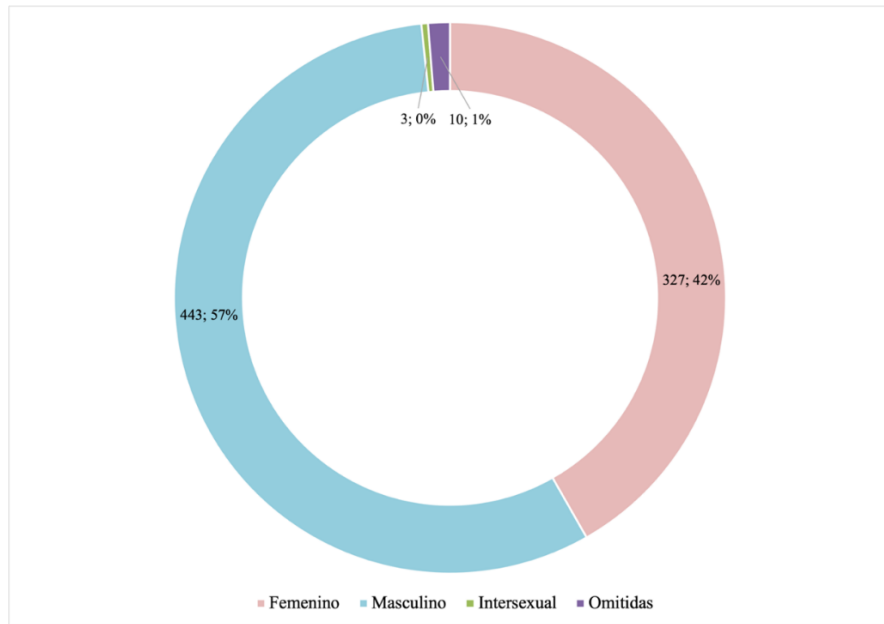


Figura 47: Distribución por género de los encuestados sobre el total de respuestas obtenidas.

De los 621 investigadores que afirman conocer qué son los datos abiertos, 362 son hombres (58,29%), 256 son mujeres (41,22%) y 3 intersexuales (0,48%). De las 57 respuestas obtenidas de quienes no conocen los datos abiertos, 32 son hombres (56,14%) y 25 son mujeres (43,86%). Por tanto, podemos afirmar que tanto en las respuestas totales obtenidas, como en la distinción entre quienes conocen o no los datos abiertos y el movimiento, cumple con los criterios de paridad representativa de género. Si nos fijamos en otros estudios, el que más cerca se encuentra de cumplir con estos criterios de paridad en las respuestas es otra investigación cuantitativa sobre datos de investigación en España, en el que el 62,4% son hombres (663) y el 37,6% de las respuestas obtenidas son mujeres (400) (Aleixandre-Benavent, Vidal-Infer, Alonso-Arroyo, Peset y Sapena, 2020). En otras investigaciones similares sobre el uso y reutilización de datos abiertos la diferencia es notoria. Por ejemplo, en el estudio desarrollado en 2010 por Tim Davies, sobre el uso de datos abiertos en Gran Bretaña, que utiliza técnicas cuantitativas para su desarrollo, la tasa de respuestas a la encuesta es, de cada 7 respuestas obtenidas, 6 son de hombres por 1 respuesta de mujer (Davies, 2010). En otros estudios como el llevado a cabo a través de una encuesta en el Foro Belmont, dirigido a una comunidad de reutilizadores de datos abiertos relacionados con el medioambiente o el estudio desarrollado en México en el que se analiza el entorno de los datos abiertos y su percepción por parte de los mexicanos a través de una encuesta Nacional de Acceso a la Información Pública y Protección de Datos Personales, no hacen referencia a distinción de género en sus respuestas (Schmidt, Gemeinholzer y Treolar, 2016; Estrada Zamora y Vargas-Hernández, 2018).

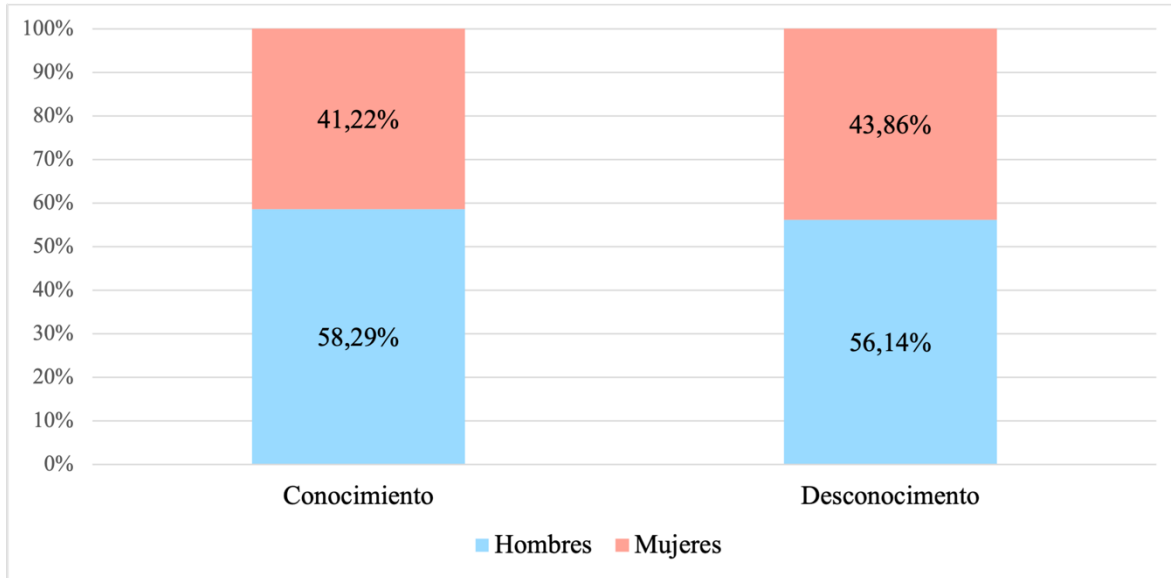


Figura 48: Porcentajes de conocimiento vs desconocimiento de datos abiertos por género.

Si nos fijamos en la plataforma de datos de EPData<sup>88</sup>, presenta una serie de visores con datos extraídos del Ministerio de Ciencia, Innovación y Universidades, en el que una de las visualizaciones representa la evolución del personal docente e investigador en las universidades, haciendo distinción por género, entre los años 2012 y 2018. Se observa, en el último año disponible, el 2018, que hay un total de 122.910 personas que forman parte del personal docente e investigador en las universidades españolas. Concretamente, 51.331 mujeres (41,76%) y 71.579 hombres (58,24%). Coincide exactamente con la tasa de respuesta obtenida en esta encuesta. Se puede concluir que, en materia de género, los resultados obtenidos son representativos si lo comparamos con la población que forman parte de la comunidad académica y personal de investigación en España.

<sup>88</sup> Plataforma de visualización de datos EPData: <https://www.epdata.es/datos/profesores-investigadores-resto-personal-docente-universidad-espana-datos-estadisticas/412>

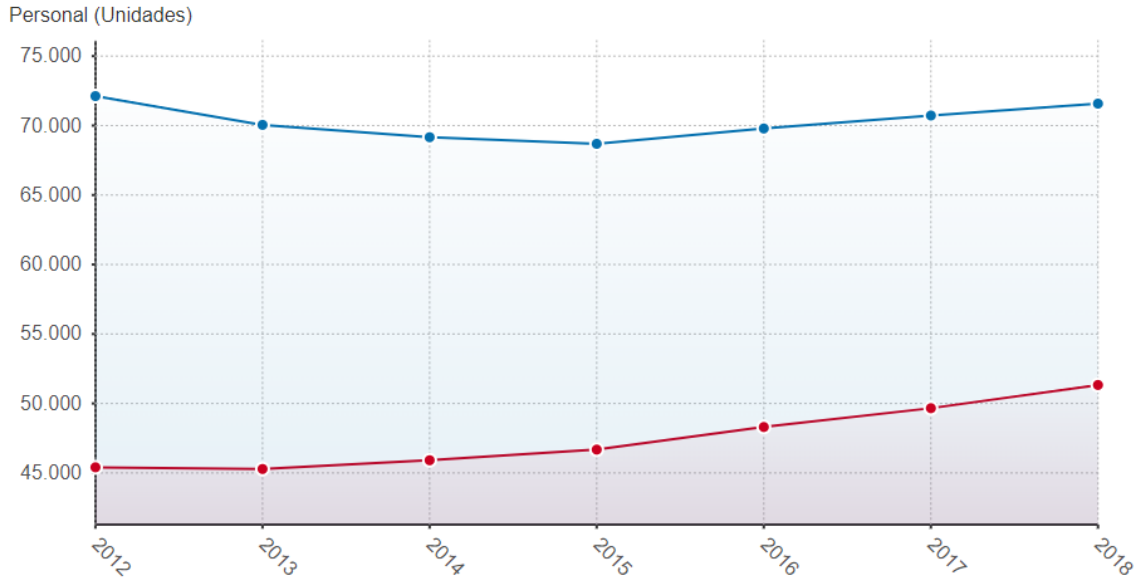


Figura 49: Población de la comunidad académica y personal de investigación en España (2012-2018)<sup>89</sup>.

Cabe recordar que esta tesis doctoral estudia el conocimiento de los datos abiertos en la comunidad científica y si estos reutilizan datos abiertos. En cuanto a la procedencia de los encuestados, existe representación de 34 de las 50 provincias y de las dos ciudades autónomas, es decir, se trata de un estudio que se centra en el uso de datos abiertos en la ciencia en territorio español. A nivel estatal, en 2019 se publicó el informe sobre reutilización de datos abiertos en España II (Abella et al., 2019), un estudio sobre el ecosistema de reutilización de datos abiertos en nuestro país. En este caso no se trata de un estudio cuantitativo si no de un análisis de calidad a través de una muestra de conjuntos de datos disponibles en portales de datos abiertos españoles. Otros estudios en los que se han llevado a cabo encuestas y otras técnicas de investigación son el ya citado informe sobre el uso y reutilización de datos abiertos en Reino Unido, que se centra en los países que conforman esta región (Davies, 2010), el estudio del Foro de Belmont, sobre el conocimiento y uso de datos sobre medioambiente, obtiene respuestas de reutilizadores de 80 países distintos o la percepción ciudadana acerca del movimiento open data en México, distribuyendo la encuesta a 14.400 viviendas en zonas urbanas con población de más de 100.000 habitantes (Schmidt, Gemeinholzer y Treolar, 2016; Estrada Zamora y Vargas-Hernández, 2018).

Se han obtenido 674 respuestas de miembros de universidades, 80 de centros e institutos de investigación y 29 omisiones. Debido al correo institucional enviado desde la Conselleria de Participación, Transparència, Cooperación y Calidad Democrática, animando a responder a la encuesta, 4 universidades valencianas aparecen entre las 11 primeras de las 47 que hay en total. Estas son la Universitat Politècnica de València (primera con más respuestas, 132),

<sup>89</sup> Fuente: Ministerio de Ciencia, Innovación y Universidades, EPData.es  
 Leyenda: Línea Roja (Mujeres); Línea Azul (Hombres).



Universitat de València (cuarta con 53 respuestas), Universidad Miguel Hernández de Elche (novena con 27 respuestas) y Universidad de Alicante (undécima con 25 respuestas). Destaca el alto volumen de respuestas de universidades como las de Universidad de Granada (segunda con 73 respuestas), Universidad Complutense de Madrid (tercera con 57 respuestas) o la Universidad de Cádiz (quinta con 37).

Se cruzan los datos de universidad a la que pertenecen los científicos con las respuestas positivas y negativas a la pregunta del conocimiento acerca de los datos abiertos. Se obtiene la información de los dos gráficos siguientes<sup>90</sup>.

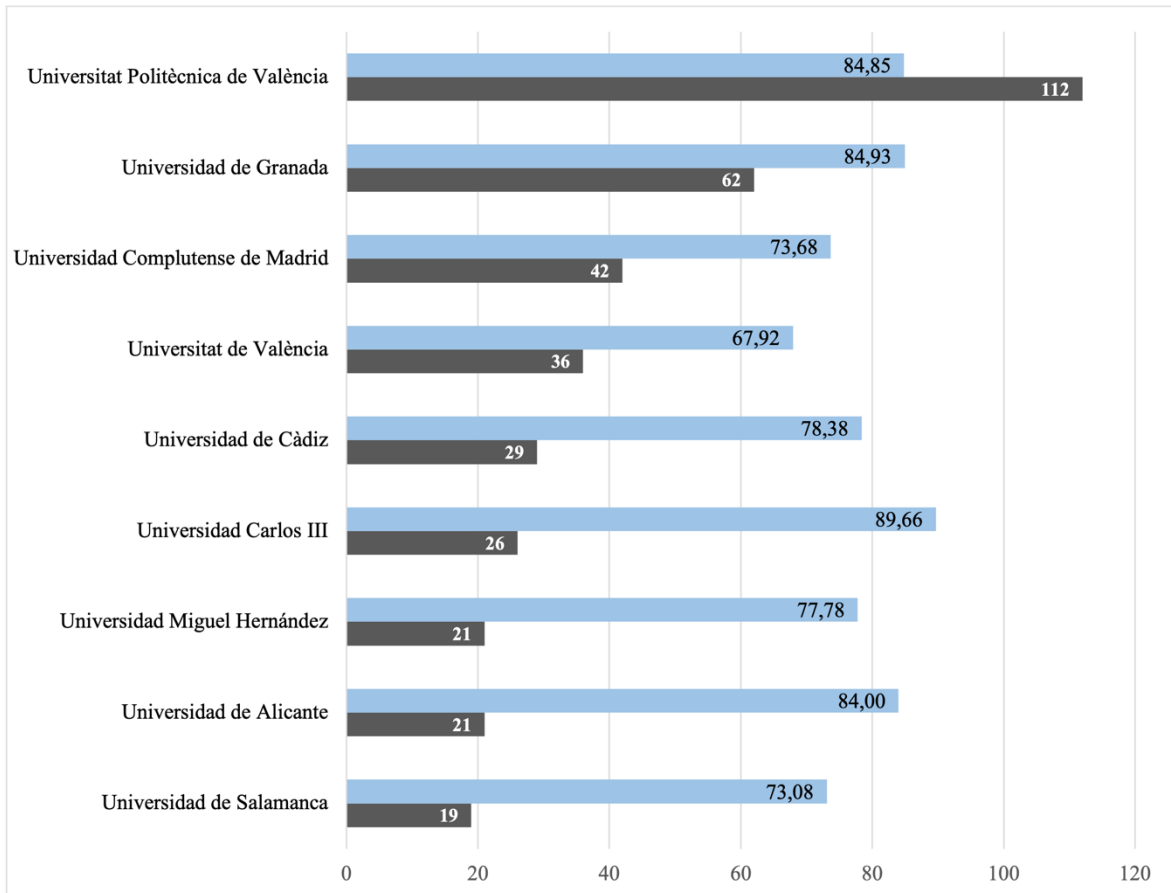


Figura 50: Conocimiento de datos abiertos por universidad.

Tal y como se describe en el párrafo anterior, 4 de las 9 universidades que más volumen de respuestas en materia de conocimiento de datos abiertos que muestra en el gráfico son valencianas. En el caso de la Universitat Politècnica de València, se obtienen un total de 112 respuestas de conocimiento de datos abiertos, lo que supone un 84,85% de las respuestas obtenidas de esta universidad. En materia de porcentaje de conocimiento por universidad

<sup>90</sup> Las barras de color negro de las figuras 50 y 51 muestran el total de respuestas por universidad, positivas (50) y negativas (51), al conocimiento de datos abiertos. Las barras azules son el porcentaje respecto al total de respuestas obtenidas de cada una de las universidades.

(que no por volumen de respuestas positivas), le siguen las universidades de Alicante, con 21 respuestas positivas siendo un 84% y Miguel Hernández de Elche, también con 21 positivas pero un 77,78% respecto a su total. Dentro de estas primeras 10 universidades con más volumen de respuestas positivas, la Universitat de València obtiene el porcentaje de conocimiento más bajo, con un total de 36 investigadores que conocen los datos abiertos y el movimiento, pero que supone un 67,92% de su total.

Destacan el volumen de respuestas obtenidas en la Universidad de Granada con 62 respuestas positivas, siendo un 84,93% de respuestas positivas de las obtenidas en esta universidad. La Universidad Complutense de Madrid con 42, supone un 73,68% y la de Cádiz con 29, es un 78,38%. Destaca el alto porcentaje que se obtiene de los investigadores residentes en la Universidad Carlos III de Madrid, que, con 26 respuestas positivas, suponen un 89,66%.

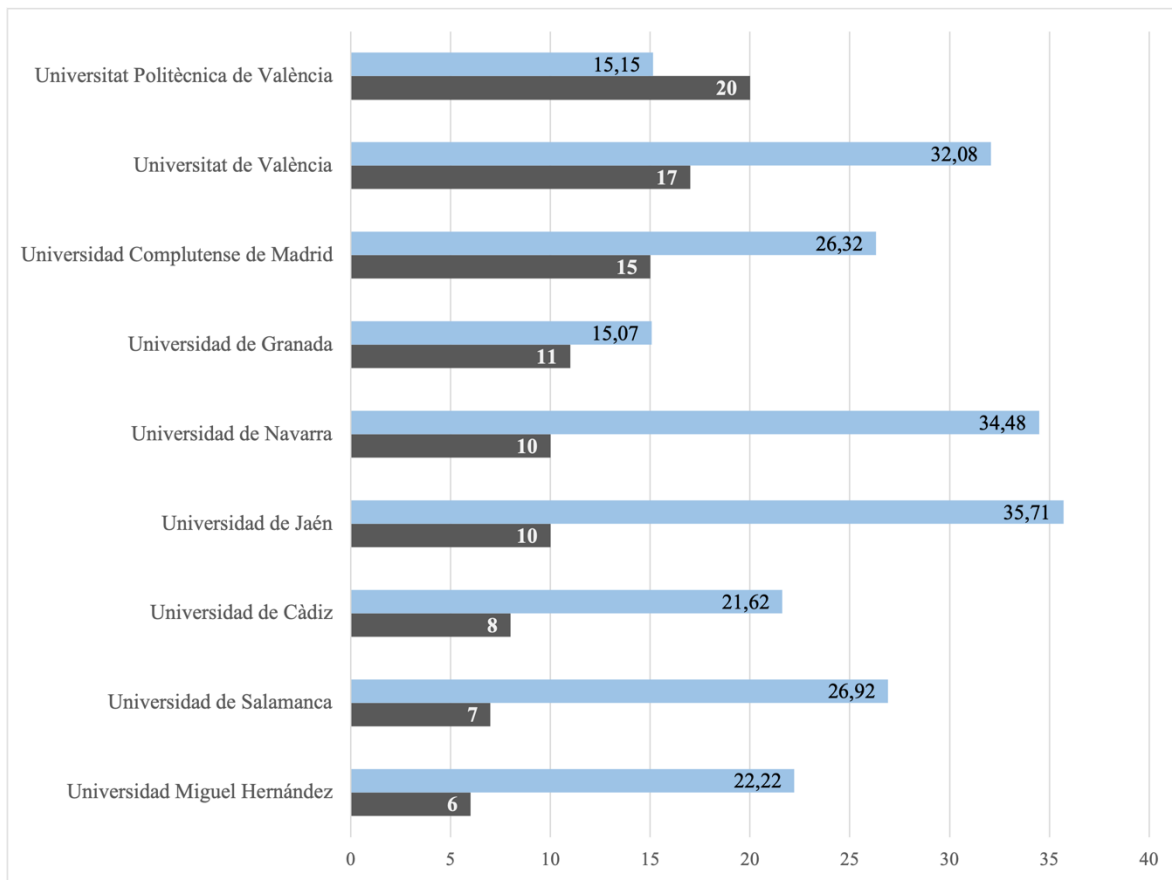


Figura 51: Desconocimiento de datos abiertos por universidad.

La procedencia por universidades o centros de investigación de quienes indican no conocer qué son los datos es diversa<sup>91</sup>. Un total de 20 científicos de la Universitat Politècnica de

<sup>91</sup> La figura 51 representa las respuestas de desconocimiento por universidad. Recoge tanto a aquellos científicos que señalan que no conocen los datos abiertos como los que omiten la respuesta. El motivo es el siguiente: se entiende que aquellos que no responden es porque no conocen los datos abiertos. De conocer el movimiento, se habría indicado positivamente.

València, (de 136 que proceden de esta universidad) afirman no saber qué son (15,15%), seguido de la Universitat de València, con 17 investigadores de 53 (siendo un 32,08%). En el caso de la Universidad Complutense de Madrid, son 15 de los 57 investigadores vinculados a esta universidad los que desconocen el movimiento (26,32%), seguida de Jaén con 10 de 28 (35,71%), Granada con 11 de 73 (15,07%), y Cádiz con 8 de 37 (21,62%). Si se observa los porcentajes de desconocimiento de datos abiertos, universidades como la Politècnica de València o la de Granada, tienen un conocimiento alto de lo que son los datos abiertos. Llama la atención la procedencia de 8 investigadores (de 80) de 8 centros de investigación (de los 21 centros que existe representación) diferentes que desconocen el movimiento y que tienen fuentes de información públicas para llevar a cabo nuevos desarrollos.

Las áreas de investigación que se han utilizado en la encuesta son las definidas por la Agencia Estatal de Investigación (AEI), siendo 19 las establecidas por la misma. Con notable diferencia, la que más veces ha resultado elegida ha sido el área 1 de Ciencias Sociales (130). Con un porcentaje de respuestas similar, se encuentran las áreas 11, Producción industria, ingeniería civil e ingenierías para la sociedad (63), el área 12, de Tecnologías de la información y las comunicaciones (63), el área 3 de Economía (61) y el área 19 de Biomedicina (59).

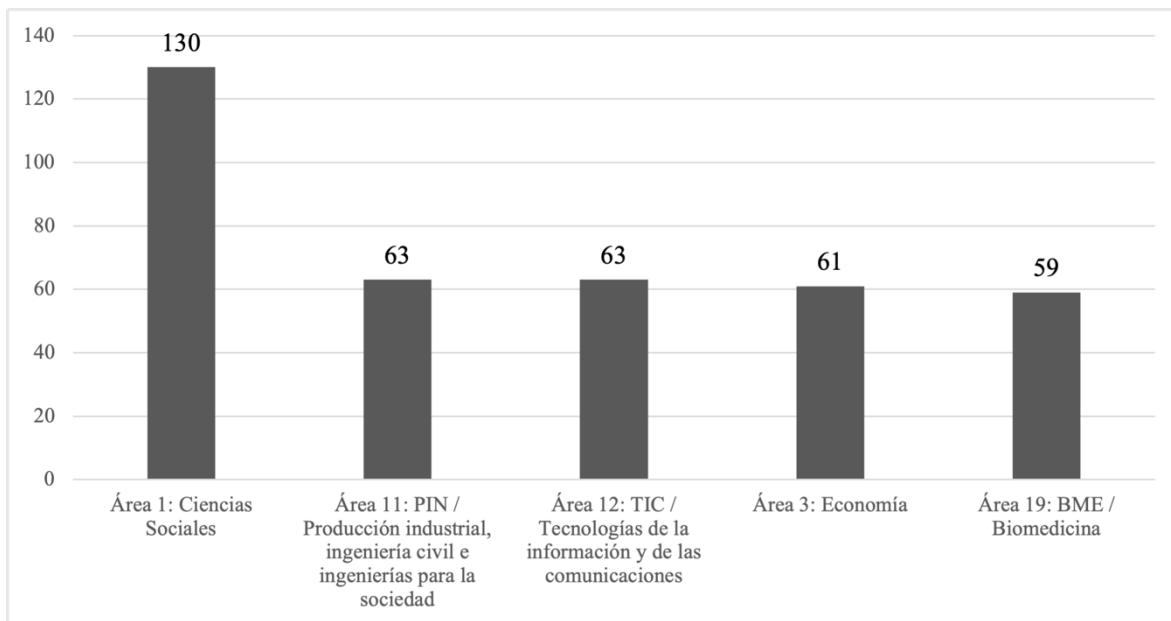


Figura 52: Áreas de investigación (+ de 50 elecciones).

Con un volumen menor, se han obtenido respuestas de las otras 14 áreas de investigación. La procedencia y el volumen de respuestas de investigadores de distintas áreas enriquece el estudio y permite obtener una visión multidisciplinar en materia de conocimiento, uso y reutilización de datos abiertos en la ciencia. Proporcionalmente, existe una diferencia entre la tasa de respuestas de investigadores procedentes de las ciencias sociales a las ciencias clásicas. La baja representación de científicos provenientes de las ciencias clásicas como la

biología, física, química, etc., es un dato para destacar, siendo que estas disciplinas en las que se utiliza la comprobación experimental de forma sistemática como parte de la aplicación del método científico, deberían hacer uso sistemático de los datos abiertos. Se plantea la cuestión de si se trata un problema de sesgo en el muestreo, o realmente responden a un interés menor por parte de este colectivo. Aun así, si lo comparamos con otras investigaciones, como el del Foro de Belmont, al tratarse de un estudio centrado en los datos relacionados con el medioambiente, se obtienen respuestas de las áreas con una estrecha relación con la naturaleza, medioambiente y salud: de ciencias de la tierra y ambientales, ciencias climáticas y atmosféricas, biológicas, físicas, ingenierías, sociales, agrícolas, veterinarias, de la salud y químicas (Schmidt, Gemeinholzer y Treloar, 2016). Se obtienen respuestas de las 19 áreas de investigación definidas previamente, con las puntualizaciones definidas.

Realizando el cruce de datos entre quienes conocen los datos abiertos y las áreas de investigación, se observa que la más elegida es el área 1 de Ciencias Sociales, con 104 (80% de conocimiento). Tiene sentido si se tienen en cuenta los conjuntos de datos disponibles en los portales de datos abiertos, potencialmente reutilizables en un área de investigación que engloba tantas disciplinas. Con 55 elecciones, la siguiente área es la 12, de Tecnologías de la Información y las Comunicaciones (87,30% de conocimiento). El área 3, de Economía es elegida en 50 ocasiones (79,37%) y en quinto lugar aparece el área 19 de Biomedicina con 42 elecciones (71,19%). Son áreas que independientemente del posterior uso y aplicación de datos en la investigación, son conscientes de este tipo de fuentes de información que pueden ser, en un momento determinado, fuentes de información potencialmente reutilizables. Dentro de las que conocemos como ciencias clásicas es el área 9, de Matemáticas, con 30 elecciones (76,92% de conocimiento), seguido del área 10 de Ciencias Físicas con 25 (78,13%). La reflexión es la misma que en el párrafo anterior: existe una baja representación de científicos procedentes de estas áreas y debemos reflexionar acerca del poco interés de este colectivo.

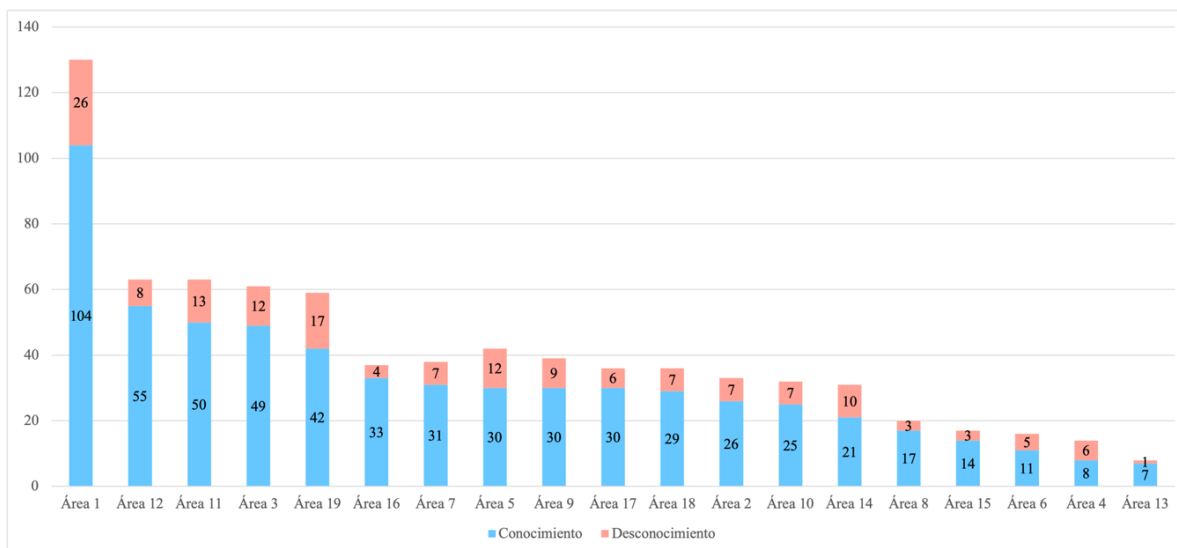


Figura 53: Conocimiento vs desconocimiento de datos abiertos por áreas.

En la misma pregunta del conocimiento acerca de los datos abiertos, se obtienen 57 respuestas negativas y 105 omitidas. Se entiende que las omitidas desconocen qué son los datos abiertos, ya que, de conocer el movimiento su respuesta hubiera sido afirmativa. El gráfico recoge las omitidas y las respuestas negativas. Al cruzar estos resultados con las áreas de investigación, el área 1, de Ciencias Sociales tiene un total de 26 científicos que desconocen los datos abiertos. El porcentaje es bajo ya que suponen únicamente un 20% del total de respuestas de esta área. Si nos guiamos por la proporción de conocimiento de datos abiertos por área, las que más conocen datos abiertos respecto al total de respuestas obtenidas del área son, el área 16, de Ciencias y Tecnologías Medioambientales (89,19% de conocimiento de datos abiertos), el área 13, de Energía y Transporte (87,50%), el área 12, de Tecnologías de la Información y las Comunicaciones (87,30%), el área 8, de Psicología (85%).

Las áreas que menos conocimiento tienen de estos recursos de información son, el área 4, de Mente, Lenguaje y Pensamiento (57,14%), el área 14, de Ciencias y Tecnologías Químicas (67,74%), seguidas de las áreas 5, de Cultura, Filosofía y Arte (71,43%) y la 19, de Biomedicina (71,19%).

Si observamos los porcentajes obtenidos en materia de desconocimiento de datos abiertos por áreas de investigación, las entidades encargadas de la apertura de información deben reflexionar sobre los conjuntos de datos que están publicando y aquellos datasets que sería interesante compartir. Este ejercicio crítico puede resultar beneficioso en aquellas disciplinas de la ciencia, como las relacionadas con las humanidades o las ciencias clásicas, ya que, tal y como apuntan los investigadores, apenas disponen de datos potencialmente reutilizables en este tipo de fuentes de información pública, y que pueden ser interesantes y de relevancia para el desarrollo de nuevos estudios.

En la universidad de desarrollo y defensa de esta tesis doctoral, la Universitat Politècnica de València, las áreas que más respuestas reciben son la 12 y la 11, de Tecnologías de la Información y las Comunicaciones y Producción Industrial, Ingeniería Civil e Ingenierías para la Sociedad, con 28 y 20 respuestas respectivamente. Destaca y concuerda con la muestra final el volumen de respuestas del área de Ciencias Sociales con 14. Teniendo en cuenta la reflexión lanzada en el párrafo anterior, se detecta un bajo volumen de respuestas en ciencias clásicas, como el área 9, de Matemáticas con 13 respuestas o el área 10 de Ciencias Físicas con 3 respuestas únicamente. En el caso de la Universitat de València, que tiene menos grados técnicos que las universidades politécnicas, se observa que el área de investigación más seleccionada es la primera, de Ciencias Sociales, con 12 elecciones de las 53 respuestas que se obtienen de esta universidad. Ciencias de la educación (área 7), Derecho (área 2) o Estudios del pasado, Historia y Arqueología (área 6) son las áreas con más presencia entre los encuestados. Observando la segunda universidad con más representación, la de Granada, el patrón de investigador es similar a los resultados totales de la encuesta, donde el área de investigación que más presencia tiene es las Ciencias Sociales, seguida de Psicología y Tecnologías de la Información y las Comunicaciones, con tasa baja de respuestas en ciencias clásicas.

Se han establecido 6 franjas de edad en la encuesta: de 22 a 30, de 31 a 40, de 41 a 50, de 51 a 60, de 61 a 65 y de más de 65. Con 249 respuestas, la franja de edad con más respuestas a la encuesta es la formada de 41 a 50 años, seguida de la franja de 51 a 60 años con 236. Si nos fijamos en la categoría laboral de quiénes participaron en la encuesta, se obtienen 209 respuestas de Profesores/as Titulares de Universidad, 150 respuestas de Catedráticos/as de Universidad y 76 de Profesores/as Contratados/as Doctores/as. Coinciden las respuestas obtenidas en la franja de edad de 22 a 30 años (68 respuestas obtenidas) con la categoría de doctorandos/as (62 respuestas). La franja de edad de 31 a 40 con 129 respuestas, coinciden con las categorías laborales de Profesores/as Asociados/as o Profesores/as Contratados/as Doctores/as con 61 y 50 respuestas respectivamente. Y a partir del rango de 41-50 se consolidan las respuestas de Profesores/as Titulares. Si observamos otros estudios, en el caso del estudio del Foro de Belmont, son 1247 quienes proporcionan información sobre su edad, teniendo una media de 43 años, la mediana representada por el grupo de edad de 41 a 45 años y siendo 20 y 90 las edades extremas (Schmidt, Gemeinholzer y Treloar, 2016). En otros estudios revisados como el análisis de los datos abiertos de Gran Bretaña (Davies, 2010), el Informe Asedie de 2021 o el estudio acerca de la percepción ciudadana de los datos abiertos en México (Estrada Zamora y Vargas-Hernández, 2018), no se hace referencia a las franjas de edad de la muestra de población estudiada por lo que se puede establecer una relación con dicho estudio.

El gráfico siguiente, muestra la edad media del profesorado en las universidades públicas españolas (Rodríguez, 2018). Si se observan las respuestas obtenidas, se pueden extrapolar

los resultados a la demografía de la población. Dentro de la muestra que ha respondido a la encuesta, existen más investigadores en el rango de edad de 61 a 65, con 55 respuestas, que del rango de 22 a 30, con 50 respuestas, hecho que reafirma las conclusiones del estudio llevado a cabo por Juan Carlos Rodríguez, de la Universidad Complutense de Madrid, en el que se observa un envejecimiento del personal docente universitario español.

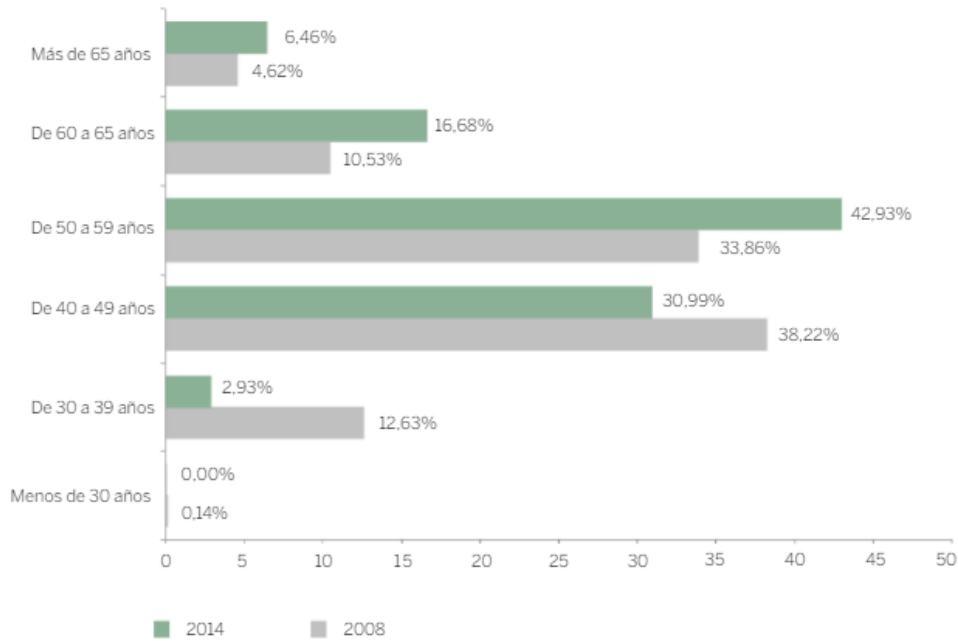


Figura 54: Edad media del profesorado en las universidades públicas españolas en 2008 vs 2014<sup>92</sup>.

En cuanto al conocimiento o desconocimiento de los datos abiertos, no existe un sesgo específico por edad. Las franjas de edad que más veces han señalado conocer el open data y su movimiento son las que están dentro de los rangos de 41 a 50, con 202 y el rango de 51 a 60 años con 185. Si comparamos la edad de la demografía española (Rodríguez, 2018) con la muestra obtenida, el volumen de respuestas es proporcional al de este estudio. Rangos de edad como el formado de 22 a 30 años recibe una tasa de respuestas baja, pero la proporción de conocimiento de datos abiertos en esta edad es alta, ya que 50 de los 68 científicos de esta edad saben qué son los datos abiertos. Lo mismo ocurre con las demás franjas de edad.

<sup>92</sup> Fuente: Juan Carlos Rodríguez, Universidad Complutense de Madrid. Demografía universitaria española (2018).

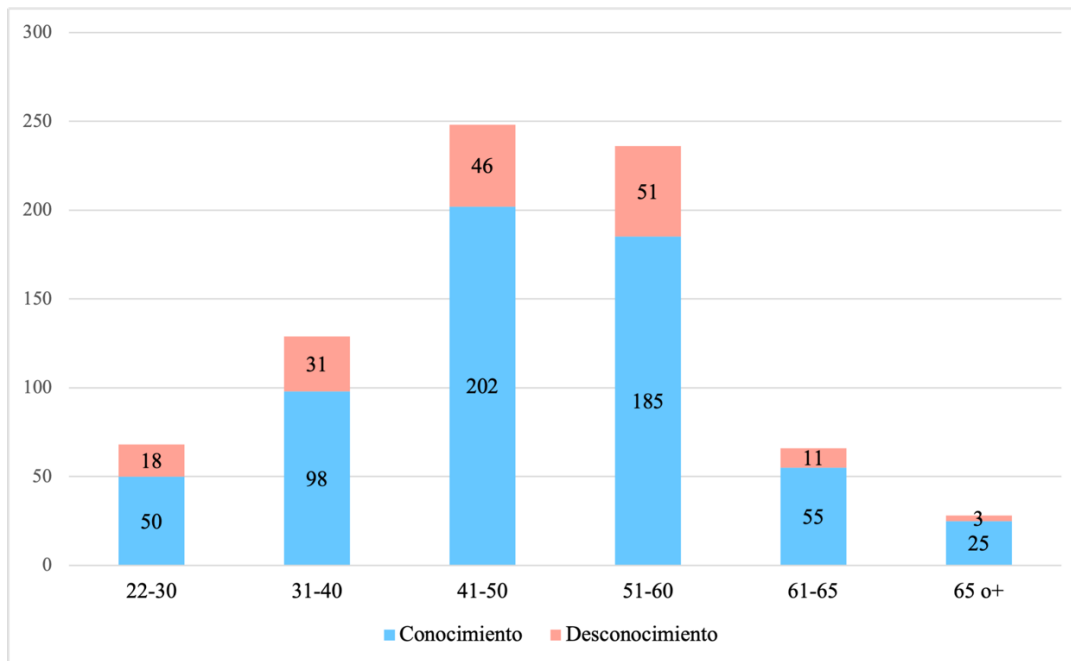


Figura 55: Conocimiento vs desconocimiento de datos abiertos por edad.

Se clasifican dentro de la categoría de desconocimiento tanto los que indican que desconocen los datos abiertos y como quienes omiten la respuesta (ya que, de conocer el movimiento hubiesen indicado sí). En el caso del desconocimiento la franja de edad que más señalan desconocer el open data es la formada por el rango de 51 a 60 años, con 51, seguido de la formada de 41 a 50, con 46. Proporcionalmente estamos hablando de porcentajes muy bajos de desconocimiento si nos fijamos en el total de respuestas obtenidas por edad. La percepción que se tiene, al ser un movimiento relativamente joven, es que sea necesario un dominio de la tecnología y se tiende a asociar que, quienes reutilicen información del sector público, no sea precisamente el rango de más de 65 años. De las 28 respuestas obtenidas dentro de esta franja, solo 3 desconocen qué son los datos abiertos. Por tanto, podemos concluir que es un movimiento que no depende del rango de edad en el que se encuentra el investigador y que esa percepción inicial es errónea.

La importancia de conocer este movimiento radica en la reutilización de los datos. De aquellos que afirman conocer los datos abiertos, 396 han reutilizado información del sector público en sus investigaciones (50,57%), 274 nunca han reutilizado datos abiertos (34,99%) y 113 omiten la respuesta de reutilización de datos. Al igual que la omisión de respuesta a la pregunta sobre conocimiento de datos abiertos se considera que el encuestado desconoce el movimiento, podemos establecer que la omisión a la pregunta acerca de la reutilización de datos, podemos considerarla negativa, siendo un 49,42% los investigadores que conocen qué son los datos abiertos, pero no los han reutilizado en sus investigaciones. Dentro de la reutilización, de los 396 que sí han reutilizado datos abiertos en algún momento, 234 son hombres (59,10%) y 162 son mujeres (40,90%). De los 274 que no han reutilizado información del sector público, 156 son hombres (56,93%) y 118 mujeres (43,07%).



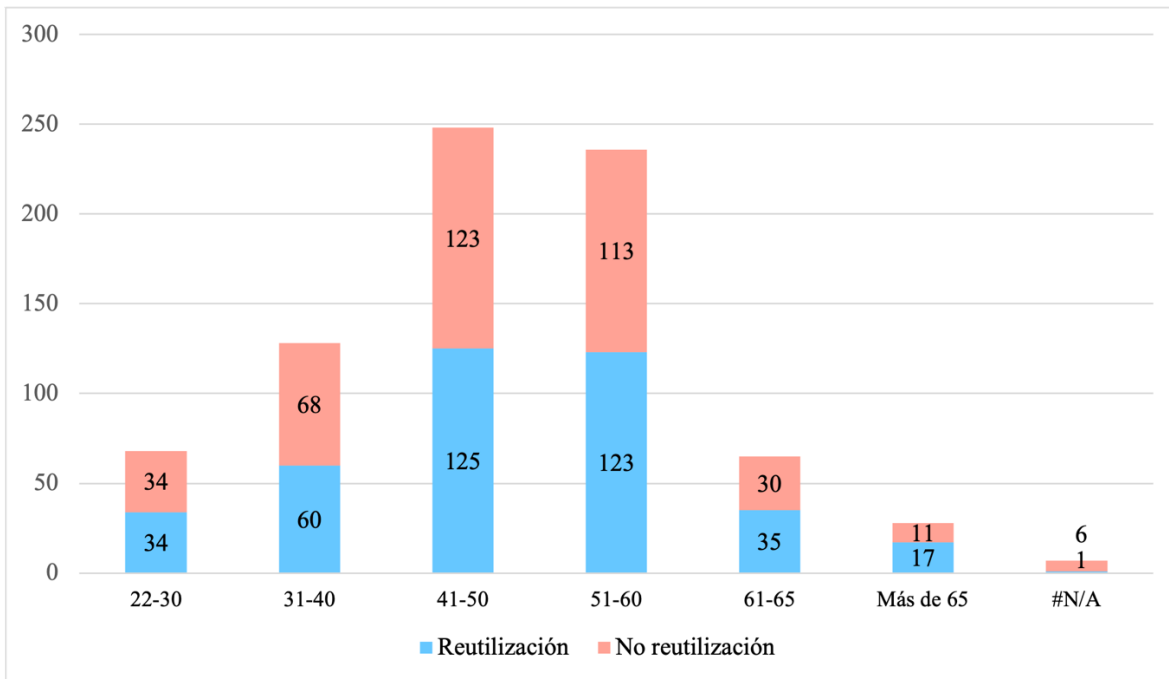


Figura 56: Reutilización vs no reutilización de datos abiertos por edad.

La pregunta acerca de la reutilización de datos abiertos se plantea a aquellos que conocen los datos abiertos. A pesar de especificarlo, responden más personas de las que afirman conocer los datos abiertos. Por ese motivo, en la figura anterior se recoge el volumen general de respuestas, es decir, por una parte, los que conocen y reutilizan información del sector público. Por otra, se engloban dentro de quienes no reutilizan datos, aquellos que desconocen el movimiento.

Es interesante también conocer aquellas áreas de investigación que más acuden a estas fuentes de información con el objetivo de extraer datos para generar conocimiento. El área 1, de Ciencias Sociales sigue siendo la más elegida como área que reutiliza datos con 75 veces seleccionada, lo que supone que un 57,69% de los investigadores, reutilizan datos. Las áreas 11 y 12, de Producción industrial, ingeniería civil e ingenierías para la sociedad y Tecnologías de la Información y las Comunicaciones, son elegidas en un total de 63 ocasiones. En el caso del área 11 reutilizan 32 de los 63 (50,79%) y el área 12, 35 de los 63 (55,56%). En el caso del área 3, de Economía, 40 de los 61 investigadores reutilizan información del sector público, siendo un 60,61% del total. Dentro de estas áreas más seleccionadas llama la atención la 19, de Biomedicina. Un total de 59 científicos contestan a la encuesta, pero solo 18 de ellos han reutilizado información, lo que supone un 30,51%. Es la quinta área de reutilización más baja tras 13, de Energía y Transporte (25%), la 4, Mente, Lenguaje y Pensamiento (28,57%), la 15, de Ciencias y Tecnologías de Materiales (29,41%) y la 8, de Psicología (30%). Proporcionalmente, el área que más datos reutiliza es la 16, de Ciencias y Tecnologías Medioambientales, con un 70,27%.

Ciencias clásicas como el área 10, de Ciencias Físicas apenas reutilizan datos, es decir, 12 de los 20 científicos de esta área señalan haber utilizado datos del sector público para sus investigaciones (37,50% de reutilización). De ahí la reflexión anterior acerca de los tipos de datos a publicar por las entidades públicas para que todas las disciplinas tengan a su disposición en este tipo de fuentes, información potencialmente reutilizable.

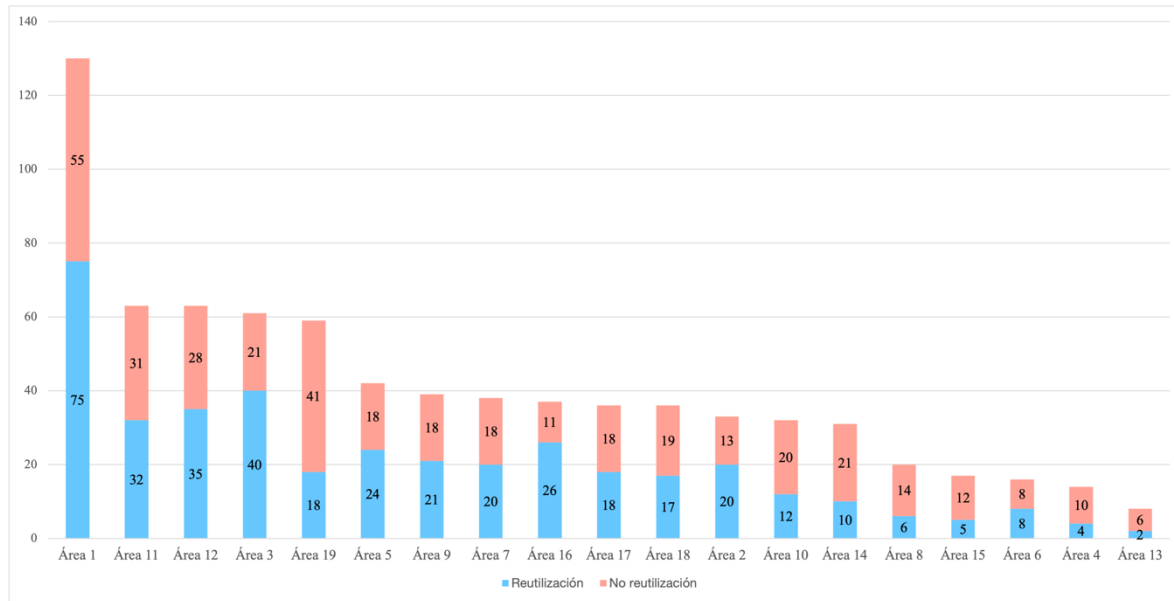


Figura 57: Reutilización de datos abiertos por área de investigación.

Si observamos la tasa de respuesta en materia de conocimiento de datos abiertos, el porcentaje obtenido es alto (79,31%). Estas iniciativas de apertura de información por parte de organizaciones gubernamentales y administraciones públicas son uno de los pilares del gobierno abierto, ya que no solo fomentan la transparencia y rendición de cuentas con la ciudadanía, sino que permiten llevar a cabo actividades que promueven la participación ciudadana. La importancia de los datos abiertos no solo radica en el conocimiento del movimiento o la localización de las fuentes. La importancia de este tipo de iniciativas radica en el uso de los mismos. Quien omite la respuesta se entiende que no conoce los datos abiertos y que no han reutilizado información del sector público. De esta manera, se puede considerar que el nivel de reutilización de los datos abiertos en la comunidad universitaria es de apenas el 50% de los encuestados. Se empieza a hablar del movimiento y cobra fuerza a finales de la segunda parte de la década de los 2000 (Garriga-Portolà, 2011), llegándose a afirmar que los datos abiertos podrían ser más importantes que el código abierto (O'Really, 2007; Muñoz, 2016). Teniendo en cuenta que los miembros la comunidad universitaria son una población con capacidad de localizar fuentes de información para el desarrollo de estudios y, por tanto, con conocimientos suficientes para tratar, interpretar y manejar grandes volúmenes de información, el dato obtenido en la encuesta en materia de reutilización se puede considerar

insuficiente, si se tiene en cuenta que, aquellos que responden al cuestionario tienen cierto interés en los datos abiertos.

Los investigadores que conocen los datos abiertos y las fuentes de información, pero no han reutilizado datos, apuntan a una falta de datos de interés en sus áreas de conocimiento para poder llevar a cabo proyectos, investigaciones o aplicarlos en fines docentes (138 respuestas). Si se observan los resultados ya analizados en materia de conocimiento y reutilización de datos por áreas de investigación, las entidades encargadas de la apertura de información deben reflexionar acerca de la publicación de determinados conjuntos de datos para que estas áreas puedan reutilizar información del sector público. Los científicos apuntan al desconocimiento acerca de cómo reutilizar la información (94). Quizás sería interesante instruir a aquellos investigadores que quieren reutilizar información o publicar ejemplos de qué se puede hacer con los datos. En materia de aspectos técnicos de los datos, hacen referencia a la falta de estandarización de estos y el trabajo que supone la normalización para el desarrollo de investigaciones (82). Finalmente, apuntan hacia la falta de actualización de los conjuntos de datos disponibles en los portales de datos abiertos, que tiene como consecuencia una desconfianza en fiabilidad de los datasets (35).

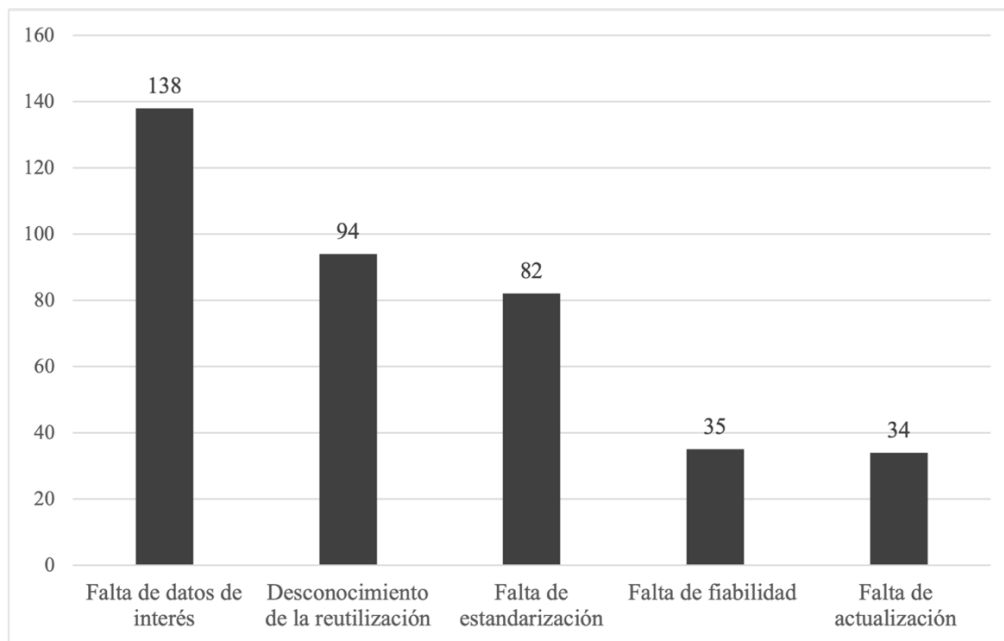


Figura 58: Motivo de la no reutilización de datos abiertos.

La falta de actualización, según el “*Manual práctico para mejorar la calidad de los datos abiertos*” publicado en Datos.Gob.es puede deberse a dos motivos: en primer lugar, algunos datasets son publicaciones puntuales que no vuelven a actualizarse ni tener un mantenimiento. El segundo motivo se debe a cambios estratégicos y falta de recursos. Afecta consecuentemente a la disponibilidad, completitud y actualidad (Datos.Gob.es, 2017a). Esta falta de actualización de la información en los portales es uno de los motivos de desconfianza

a la que apuntan los reutilizadores y consecuentemente, quedan señaladas las acciones realizadas por parte de las instituciones encargadas de la apertura y puesta a punto de los datos (Gértrudix, Gertrudis-Casado y Álvarez-García, 2016).

En cuanto a la relación de los investigadores con los datos, aquellos que conocen el movimiento de datos abiertos, están involucrados con ellos de distintos modos. En 304 ocasiones, los investigadores apuntan que son usuarios de datos para informarse. Señalan en 244 que son reutilizadores de datos y que les ha servido para el desarrollo de productos o servicios de valor añadido. Un total de 150 investigadores afirman ser proveedores de datos y únicamente 6 señalan ser administradores de datos. Es decir, 566 de los 783 investigadores que han respondido la encuesta se nutren de la información disponible en portales de datos abiertos en el desarrollo de su actividad profesional y académica. Las entidades encargadas de la apertura de datos pueden encontrar en la academia un apoyo a la hora de llevar a cabo hojas de ruta de apertura de información, ya que, tal y como se observa en los resultados, existe una representación de proveedores de datos. Pueden además asesorar en materia de búsquedas de otras fuentes de información, en herramientas para la limpieza y estandarización de datos, en explotación de datos y creación de visores que pueden servir como ejemplo e inspiración no solo a la ciudadanía y reutilizadores, si no a la propia administración.

El modo de conocer los datos abiertos viene dado de formas distintas. Artículos de investigación (321 veces elegido), canales de comunicación de los proveedores de datos y mediante talleres de formación, seminarios y congresos (159 elecciones) es el modo en el que los investigadores llegan a conocer y por tanto trabajar con datos abiertos. Canales de comunicación (88) y a través de reutilizadores (85) son otros modos en que los investigadores entran en contacto con los datos abiertos. Si observamos la forma en la que los miembros de la academia llegan a conocer este movimiento, se hace necesaria por parte de las administraciones, un modo de publicidad activa que fomente el conocimiento y el uso de los datos. Deben ser conscientes que son una fuente de información relevante no solo en la investigación, sino también en el cada vez más grande sector infomediario (Asedie, 2021).

Centrándonos en aquellos que desconocen los datos abiertos se busca el motivo, planteando una opción de respuesta abierta. De las 168 obtenidas, únicamente 78 son válidas. Los investigadores apuntan al desconocimiento de fuentes de datos, falta de información, así como del uso y potencial de reutilización. Hacen referencia además a la falta de difusión y publicidad activa desde las propias administraciones públicas encargadas de la apertura de estas fuentes de información y finalmente, apuntan a que no es necesario en su actividad docente o de investigación y de ahí el desconocimiento.

Por tanto, con las respuestas obtenidas, se observa que sería necesaria una instrucción a los miembros de las comunidades universitarias en cuanto a cultura de la reutilización de

información. Existe una falta de uso de datos abiertos y desconocimiento de estos de una parte de los docentes universitarios. Se hace imprescindible adquirir y publicitar información acerca de estos recursos que son los datos potencialmente reutilizables en el ámbito universitario (Díez y Millán, 2011). Esta formación proporcionada a los investigadores permitiría adaptar estos nuevos escenarios y reestructurar espacios encargados de promover metodologías que mejoren el rendimiento académico, de investigación e incluso, aumentar el presupuesto de las universidades para el avance tecnológico (Sánchez, 2016). Se deben llevar a cabo nuevas metodologías que impulsen la alfabetización digital que fomenten el empoderamiento participativo y de reutilización (Gértrudix-Barrio, Gálvez de la Cuesta, Said-Hung y Durán-Medina, 2016). Así se podrá llegar a considerar ciudadanos digitales correctamente formados y con capacidad de movimiento y resolución en entornos tecnológicos y con cultura de reutilización de información (Cerezo, 2010), es decir, ciudadanos digitalmente alfabetizados, familiarizados con el *Opendata*, bien informados y con capacidad de decisión (Álvarez-García, Gértrudix-Barrio y Rajas-Fernández, 2014). Por tanto, este asunto es particularmente relevante si nos referimos al colectivo universitario, quienes deberían encabezar estos avances.

Docentes universitarios e investigadores tienen la misión de crear oportunidades aprendizaje a través de las tecnologías de la información y la comunicación (UNESCO, 2008). Los datos abiertos necesitan de tecnología para usarse eficazmente. Un buen uso de los datos abiertos requiere de una capacitación tecnológica por parte de los docentes para saber a qué fuentes de datos acceder, tener la capacidad de detección de calidad de estos para la consulta o reutilización, conocer técnicas y mecanismos para la extracción de datos y finalmente, tratarlos y analizarlos para la reutilización (Rivas-Rebaque, Gértrudix-Barrio, De-Cisneros-De-Britto, 2019).

### 5.3.2 Apertura de datos: uso y reutilización de la información del sector público

Tanto los portales de datos abiertos como los datasets disponibles, deben contener una serie de atributos y requisitos mínimos para que la información sea reutilizable. Hoy en día disponemos de un volumen inmenso de datos a nuestro alcance y, sin embargo, el hecho de que sean más asequibles no es sinónimo de sencillez a la hora de acceder y reutilizarlos (Datos.Gob.es, 2017a).

Es frecuente que, al intentar localizar conjuntos de datos, estos estén disponibles en portales de información institucionales, pero sea complicado encontrarlos. Sucede en ocasiones que están divididos o distribuidos en distintas secciones de una web o que la propia estructura de la página web es la que dificulta la localización y acceso a lo que se está buscando (Datos.Gob.es, 2017a). Los encuestados señalan que tiene un peso importante accesibilidad a los conjuntos de datos, la facilidad de localizarlos y conseguirlos. Además, se considera

elemental que estos no estén sujetos a más restricciones que las de atribución de autoría de la fuente y se tiene muy en cuenta el tipo de licencia de datos para su posterior reutilización.

A pesar de la importancia de puesta a punto y accesibilidad a los datos, es necesario realizar una distinción entre cantidad y localización de datos disponibles y calidad de estos (Abella, Ortiz-de-Urbina-Criado y De-Pablos-Heredero, 2018). El éxito de estos entornos de datos abiertos viene dado por la calidad de los datasets disponibles más que por el hecho de que sean accesibles (Janssen, Charalabidis y Zuiderwijk, 2012). La puesta a punto de los datos no tiene sentido si no tiene como objetivo generar conocimiento a través de productos o servicios que aporten un valor añadido (Ferrer-Sapena, Peset y Aleixandre-Benavent, 2011). En el caso de la encuesta desarrollada en el Foro de Belmont, se considera de vital importancia, con el porcentaje más elevado de los ítems a evaluar (82% de importancia), la calidad de los conjuntos de datos.

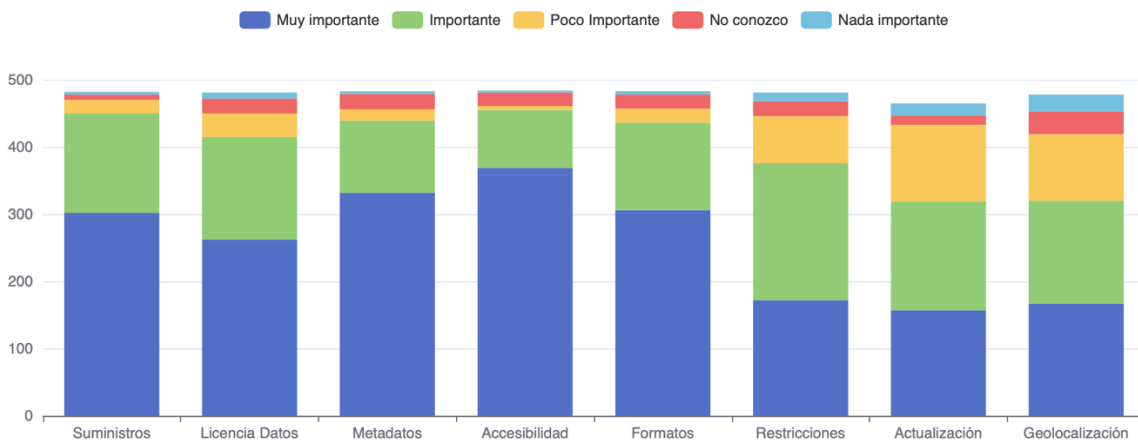


Figura 59: Atributos de importancia en la apertura de datos.

Uno de los aspectos que influye directamente en la calidad de los datasets, coincide en la importancia señalada en el cuestionario en cuanto a disponibilidad de metadatos, debiendo ser completos, estructurados y útiles para poder interpretar los datos. Aun así, existe un problema frecuente: metadatos completos, pero, al acceder a los conjuntos de datos, estos se encuentran desactualizados o incompletos. Por tanto, el tener bien definidos los metadatos no es sinónimo de datos exactos, consistentes y actualizados (Datos.Gob.es, 2017a). En el caso del estudio de datos abiertos en la investigación del medioambiente, un 78% de los encuestados consideran de vital importancia una buena definición de los metadatos, hecho que coincide en los resultados de este estudio (Schmidt, Gemeinholzer y Treloar, 2016).

Una cuestión fundamental en la reutilización de datos del sector público es la posibilidad de comparar mismos conjuntos de datos de distintas fuentes. Los encuestados señalan haber encontrado barreras en cuanto a estandarización de los datos disponibles y una falta de interoperabilidad para poder establecer cruces y comparaciones. En la interoperabilidad entra en juego desde el contenido de los datasets, cómo están estructurados y los campos que

contiene y, además, el tipo de formato técnico, que es lo que nos va a permitir cruzar y combinar con otros conjuntos de datos. Los datos que se obtienen en esta encuesta son similares a la percepción de los resultados obtenidos en el foro de Belmont sobre datos relacionados con la investigación del medioambiente. En el caso de este estudio no se plantea como una barrera si no como una funcionalidad que se espera de los datos. Un 51% de los encuestados no consideran fundamental que los conjuntos de datos sean interoperables contra un 40% que lo considera de mucha relevancia. Sí que apuntan a la importancia de la provisión de datos sin restricciones (concretamente un 67%) (Schmidt, Gemeinholzer y Treloar, 2016). La percepción de los científicos encuestados en este estudio es similar en materia de disponibilidad de los datos. En el caso concreto de esta investigación se evalúa con una escala Likert de 5 franjas y la comunidad científica define de importancia la disponibilidad sin restricciones de determinados conjuntos de datos.

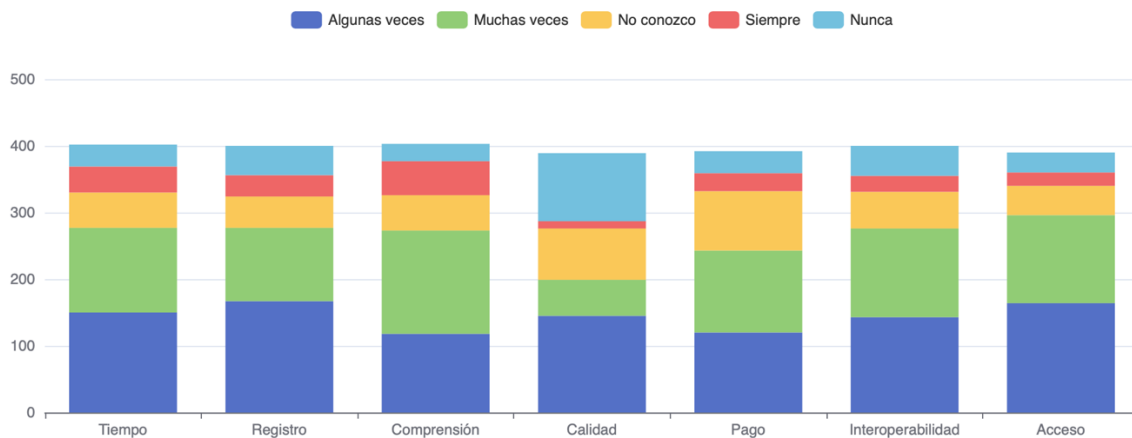


Figura 60: Barreras de acceso y reutilización de datos abiertos.

Una barrera a la que se enfrentan los investigadores y comunidad de reutilizadores es la temporal. Transcurre mucho tiempo entre la solicitud de datos a las instituciones encargadas de la apertura de estos y la recepción de estos. Como consecuencia, dificulta la reutilización y el desarrollo de nuevas investigaciones o creación de nuevos servicios. Estas barreras, según apunta el informe de “Reutilización de Información del Sector Público: Conceptos básicos, beneficios del Open Data y barreras para su aplicación” son, en muchos casos, culturales. Existe una falta de conciencia desde las entidades acerca del potencial de los datos abiertos además de un temor a la pérdida de control de estos. Cambiar este tipo de mentalidad en las administraciones requiere un compromiso político al más alto nivel y un diálogo constante entre reutilizadores (empresas, particulares, comunidad académica) y administradores de datos (administración pública) (Comisión Europea, 2011; Datos.Gob.es, 2017b).

Tal y como se ha analizado, en primer lugar, la tasa de respuestas de investigadores que desconocen qué son los datos abiertos es alta al tratarse de personal cualificado. Existe además un número considerable de investigadores que son conscientes que tienen en los

portales de datos abiertos una fuente de información de relevancia para sus estudios pero que no reutilizan información por diversas razones. Y finalmente, se aprecia a través de las respuestas de relación del investigador con los datos abiertos y cómo llega a conocerlos, que existe una falta de publicidad activa desde las instituciones encargadas de la apertura. Son los propios portales de datos abiertos quienes deben fomentar el uso de la información disponible. En ocasiones, hacen falta ejemplos prácticos, visualizaciones y herramientas dentro de los portales para que, no solo los investigadores, si no ciudadanía en general, puedan ver qué puede hacer con los datos, a través de modelos o pautas<sup>93</sup>. Así lo señalan los investigadores. Apuntan que son necesarias visualizaciones en los portales de datos que estimulen la ciudadanía y reutilizadores a llevar a cabo estudios y que se sientan capaces de realizarlas. Señalan que es de importancia tener a su disposición ejemplos de herramientas para el procesamiento de los datos y formularios de contacto con las instituciones encargadas de la apertura por si son necesarias explicaciones para la interpretación de estos. Finalmente, otra fórmula de fomento de reutilización de datos son los enlaces a aplicaciones desarrolladas con información del sector público y a canales de noticias que se encarguen de recopilar novedades de fuentes de datos y nuevas investigaciones llevadas a cabo con datos.

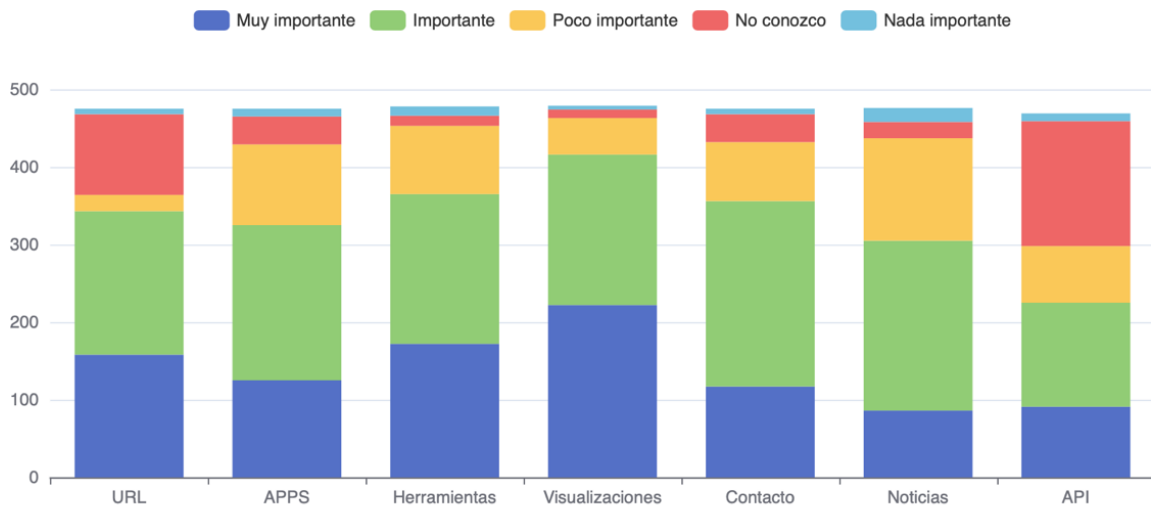


Figura 61: Importancia de servicios y funcionalidades de los datos abiertos.

El uso de los conjuntos de datos disponible es sinónimo de utilidad y necesidad para la investigación y el desarrollo. Los tipos de datos más utilizados con los de ciencia y tecnología (220 elecciones), datos sobre el sector público (199), educación (170), medioambiente y clima (160), salud (151), demografía y población (139), economía (129) y datos sobre sociedad y bienestar (111).

<sup>93</sup> Ejemplo práctico de uso de datos abiertos para el desarrollo de una investigación: Ferrer Sapena, A.; Calabuig, J.M.; Sánchez-Pérez, E.A. y Vidal-Cabo, C. (2020). Gobierno abierto y acceso a la información: un estudio de caso sobre el impacto en la economía local. *BiD: textos universitaris de Biblioteconomia i Documentació*. (45) pp. 1-25. <https://doi.org/10.1344/BiD2020.45.17>



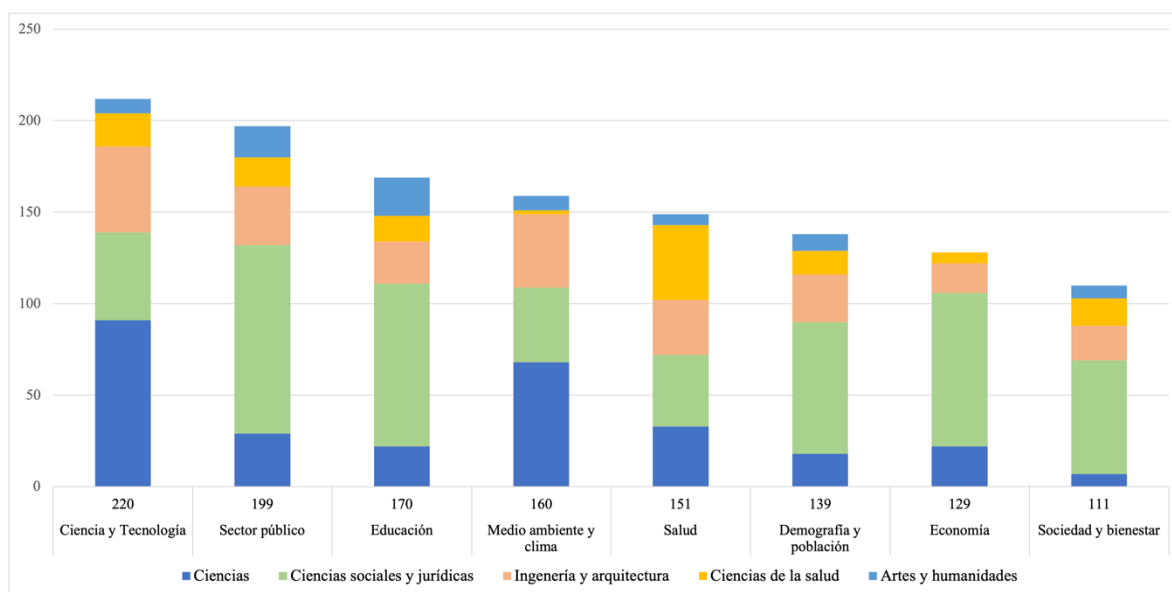


Figura 62: Conjuntos de datos (+100 elecciones) por áreas generales de investigación.

Para realizar este análisis en concreto, las 19 áreas de investigación utilizadas en este estudio se engloban en las 5 siguientes: Ciencias, Ciencias Sociales y Jurídicas, Ingeniería y Arquitectura, Ciencias de la Salud y Artes y Humanidades.

Los tipos de datos utilizados, una vez clasificados por áreas generales de investigación, son coherentes con las áreas de investigación a las que pertenecen los científicos. Por ejemplo, conjuntos de datos como el de sector público, educación, economía o demografía, son datasets utilizados por investigadores del área de ciencias sociales y jurídicas. O en el caso de datos sobre salud, coincide con investigadores del área de ciencias de la salud.

Por tanto, se debe seguir trabajando en estos conjuntos de datos para que estas áreas que se nutran de estas fuentes de información, pero es imprescindible hacer hincapié en la necesidad de apertura de datasets relacionados con las ciencias clásicas, como las matemáticas, física, química, biología, y aquellas disciplinas relacionadas con las humanidades, como historia, cultura, arte o filosofía.

### 5.3.3 Utilidad y motivaciones para la reutilización de datos abiertos en investigación

Distintos investigadores y fuentes institucionales que llevan a cabo estudios, recomendaciones y materiales formativos para la apertura y uso de datos abiertos, señalan la necesidad y utilidad que tienen los datos abiertos en nuestra sociedad. Se habla de innovación, de mejoras o creación de servicios de valor añadido, de participación ciudadana, de transparencia por parte de administraciones y control democrático, de mejora en la eficiencia y eficacia de servicios ofrecidos por los gobiernos, de impacto en políticas públicas

o de nuevos conocimientos a partir de fuentes de datos combinadas y patrones que se observan en grandes volúmenes de datos (Open Knowledge Foundation, 2017; Open Data Barometer, 2017; Barros, 2016; Ferrer-Sapena, Peset y Aleixandre-Benavent, 2011; Abella, Ortiz de Urbina Criado, De Pablos Heredero, Vidal-Cabo y Ferrer Sapena, 2019; Aleixandre-Benavent, Ferrer-Sapena y Peset, 2019, entre otros).

Los resultados obtenidos coinciden con la revisión bibliográfica donde los investigadores hacen hincapié en el impacto e importancia que tiene la apertura de datos en la producción científica y en la academia, en el fomento de la investigación y una mejora en las actividades educativas, sirviendo de respaldo a los problemas que se encuentra la sociedad.

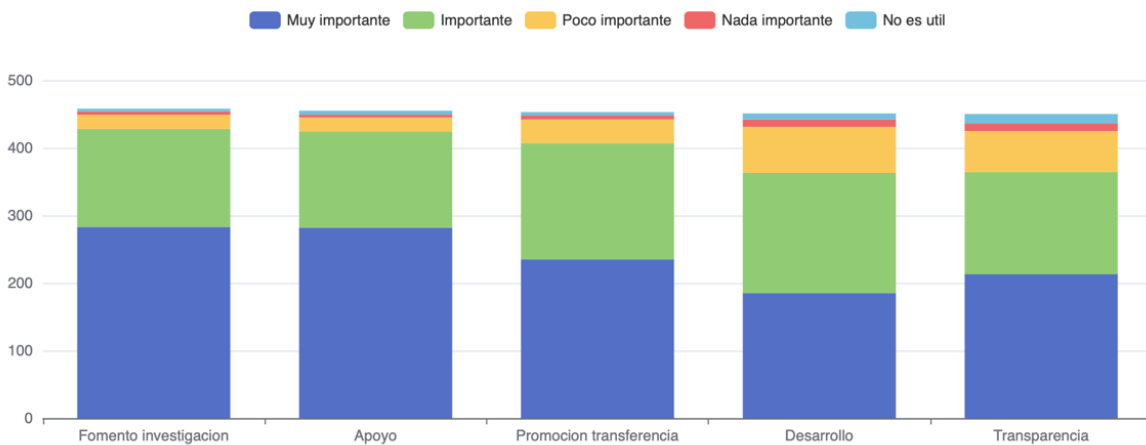


Figura 63: Utilidad de apertura de datos abiertos.

Se refuerza el discurso acerca de la transferencia de conocimiento y tecnología generado desde las universidades y centros de investigación, como vehículo transmisor al tejido social y productivo. La apertura de información contribuye a incrementar el valor generado a través de investigaciones y permite dotar con una aplicación práctica a los estos estudios generados. Quienes se benefician de estos procesos de transferencia es la sociedad, ya que permite el desarrollo de una economía basada en el conocimiento y que garantiza el bienestar de la ciudadanía (Rubiralta y Bellavista, 2003; Fundación Cotec, 2003).

Los investigadores encuentran estimulante el hecho que la apertura de información sea un camino para el intercambio espontáneo de ideas y para establecer nuevas líneas de investigación. Permite replicar y comprobar investigaciones realizadas con información disponible en fuentes públicas y mostrar a investigadores y funcionarios públicos cual es la finalidad de su trabajo además de suponer un punto importante en cuanto a transparencia en la investigación. Puede servir como punto de partida para el diseño de servicios de información estratégica y acercamiento al paradigma del *Linked Open Data*<sup>94</sup>. Finalmente,

<sup>94</sup> Datos enlazados, mediante tecnologías de web semántica, entendidos como datos estructurados para que puedan ser interconectados.

la apertura de información pública sirve como retorno social de las inversiones en ciencia y tecnología y acelera el avance de la sociedad.

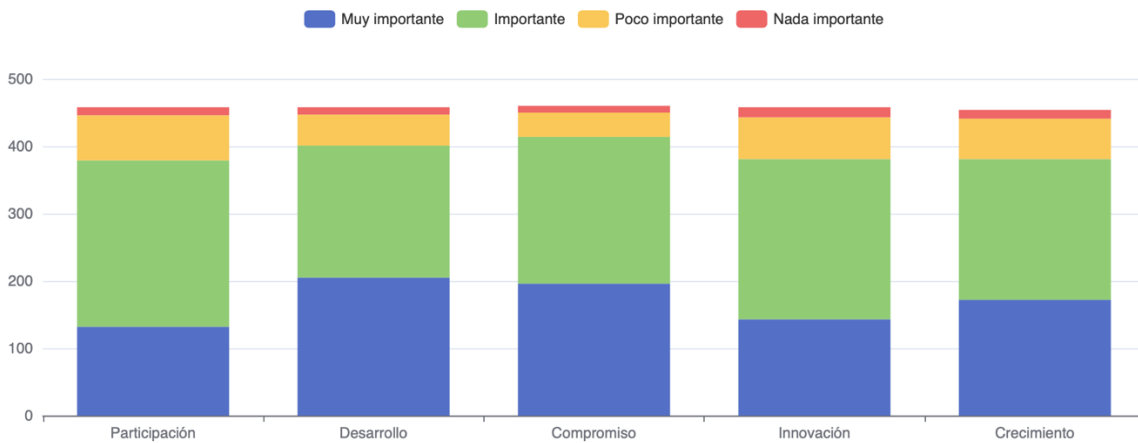


Figura 64: Importancia de la apertura de datos abiertos para la sociedad.

Consideran de vital importancia las mejoras en materia de accesibilidad a la información pública para tener una sociedad informada. El uso de datos es sinónimo de un desarrollo del que la población se beneficia y que, además, sirve para que los reutilizadores tengan un crecimiento profesional, de formación en materia de reutilización y tecnología y, por tanto, un reconocimiento laboral si se consiguen objetivos profesionales (Garriga-Portolà, 2011; Abella et al., 2019; Asedie, 2021).

#### 5.3.4 Acceso a los conjuntos de datos: fuentes de datos abiertos consultadas

En 2 años, del 2017 al 2019, el número de portales de datos abiertos incrementa un 171% en España. Paradójicamente, en este mismo período disminuyen de un 40% a un 25,2% el número totales de DMS (herramientas de gestión y visualización de datos más difundida en los portales de datos abiertos) lo que refleja una disminución a nivel global de madurez en los portales de datos abiertos (Abella et al., 2019). Observando este fenómeno, se entiende que dentro del 46,61% de encuestados que responden a la pregunta, el portal de datos abiertos al que más han accedido los investigadores sea el portal del estado (Datos.Gob.es), que contiene iniciativas de datos abiertos de distintas administraciones del estado (es decir, sirve como puerta de acceso a portales de datos abiertos de otras administraciones).

En el estudio realizado en 2019 sobre la reutilización de datos abiertos en España, las propias administraciones encargadas de la apertura se consolidaron como el segmento que más datos reutiliza por encima de académicos y entidades sin ánimo de lucro (Abella et al., 2019). Por tanto, las entidades encargadas de la apertura de información deben reflexionar ante este hecho, llevando a cabo ejercicios de puesta en valor y publicidad activa para que no sean únicamente miembros de las propias administraciones quienes reutilicen la información del sector público.

En España se llevan a cabo numerosas iniciativas de Gobierno Abierto y reutilización de información. El crecimiento de portales de datos abiertos en nuestro país en los últimos años es más que evidente: 11 portales operativos en 2013 con 5000 datasets, a 17 repositorios con más de 14000 conjuntos de datos en 2019 (Curto-Rodríguez, 2021). Es innegable que son fuentes de información a las que se puede sacar y pedir más rendimiento del obtenido. Llama la atención que, de 340 respuestas de investigadores de la Comunidad Valenciana (245 de Valencia, 78 de Alicante y 17 de Castellón), solo 133 hayan accedido al portal de datos abiertos de la Comunidad o que al de la Comunidad de Madrid solo hay accedido 88 de las 117 respuestas. Como se ha concluido en el párrafo anterior, las autonomías encargadas de llevar a cabo iniciativas de gobierno abierto deben focalizar y prestar su atención en ejercicios de puesta en valor de los datos y hacer ver a la ciudadanía y en este caso, a la comunidad científica, que tienen un respaldo y una fuente de información potencialmente reutilizable en estas fuentes.

Portal de CCAA	Nº	% <sup>95</sup>
Comunidad Valenciana	133	49,26
Madrid	88	32,59
Andalucía	76	28,15
Cataluña	57	21,11
País Vasco	38	14,07
Castilla y León	35	12,96
Aragón	25	9,26
Murcia	22	8,15
Navarra	22	8,15
Canarias	18	6,67
Castilla-La Mancha	18	6,67
Galicia	16	5,93
Baleares	15	5,56
Extremadura	13	4,81
Asturias	11	4,07
Cantabria	6	2,22
Ceuta	6	2,22
Melilla	6	2,22
La Rioja	5	1,85

Tabla 61: Uso de datos abiertos en portales autonómicos.

En las opciones de respuesta abierta, investigadores indican que no utilizan datos abiertos de administraciones públicas porque no existen conjuntos de datos sobre sus líneas de investigación en estas fuentes. Debemos reflexionar sobre, si institutos y otros centros de investigación financiados con recursos públicos, deben depositar los datos obtenidos en sus estudios e investigaciones en portales de datos abiertos, a modo de “rendición de cuentas” con la ciudadanía (ya que son estudios desarrollados con financiación pública), si deben

<sup>95</sup> % calculado sobre el total de respuestas a la pregunta 14. Los investigadores pueden seleccionar más de una opción.

depositarlos únicamente en repositorios de investigación, o si es necesario depositarlos en ambos sitios.

A pesar de que la motivación principal para el uso de datos abiertos para los miembros de la academia es la investigación y nuevos desarrollos, existen otros estímulos para llevar a cabo la reutilización. Señalan que el uso de estas fuentes es sinónimo de crecimiento y formación en el ámbito profesional, que va a permitir llegar a cumplir con objetivos laborales. Además, permite mejorar el acceso y comprensión de la información pública disponible y en menor medida, va a permitir desarrollar o conocer la web semántica a través de datos enlazados entre administraciones.

Para llevar a cabo nuevas estrategias en materia de gestión y apertura de datos, se deben trazar planes de disponibilidad, integridad, usabilidad y seguridad.

### 5.3.5 Uso y reutilización de datos abiertos

El uso de datos abiertos ha sido variado. Al tratarse de un estudio centrado en la comunidad científica, la reutilización de los datos es destinada a generar de conocimiento. Y este conocimiento es el valor añadido que aporta la reutilización de datos procedentes, en este caso, del sector público, siendo uno de los pilares del gobierno abierto.

Un alto porcentaje de quienes de los encuestados (68,48%) afirman haber descargado conjuntos de datos de los portales y su uso ha permitido desarrollar publicaciones y estudios científicos, además de para realizar análisis estadísticos. También ha servido para la creación de nuevos servicios basados en datos abiertos como herramientas y otros visores de datos, desarrollo de sitios/portales web, para la creación de mapas de situación, para integrar datos gubernamentales con aplicaciones o páginas web existentes, para llevar a cabo campañas políticas o con fines académicos y docentes. Señalan el uso de datos a modo curiosidad y consulta de información pública de interés.

A pesar del uso de datos con fines distintos, se reduce prácticamente a la mitad el porcentaje obtenido respecto al del uso, cuando se plantea la opción de utilizar datos abiertos en el futuro (disminuye de un 68,48% a 35,92%). Este fenómeno puede atender a distintas razones. En primer lugar, tal y como se ha apuntado en párrafos anteriores, a las dificultades en materia de accesibilidad, de localización de fuentes de datos y al tiempo transcurrido entre la solicitud de datos a las instituciones públicas y la recepción de estos. La competitividad y producción científica masiva hace que estos indicadores afecten directamente a la valoración de los conjuntos de datos, y, por ende, al uso y reutilización. Las entidades deben llevar a cabo iniciativas de apertura de información y, además, trabajar en la accesibilidad y poner a disposición del ciudadano, del modo más sencillo de localizar, los conjuntos de datos. Aun así, debemos reflexionar acerca del modo en el que estamos trabajando actualmente debido

a un entorno más competitivo y menos colaborativo, hecho que puede estar lastrando el progreso.

Los investigadores hacen referencia a que la no reutilización de datasets se debe a una no necesidad de datos disponibles en los portales de datos abiertos. Es otro elemento que puede hacer reflexionar a las entidades encargadas de la apertura de información acerca de qué información debe contener los portales de información pública. Por ejemplo, observando las áreas de investigación a las que pertenecen los investigadores encuestados, observamos una amplia variedad de respuestas, donde la más destacada es la de Ciencias Sociales debido a que contienen distintas áreas dentro de ella. Este tipo de investigador no ha tenido problema en poder consultar o reutilizar datos abiertos porque sí que han encontrado información en los portales. Pero debemos reflexionar acerca del motivo de que aquellos institutos y centros de investigación que reciben recursos y financiación pública no realizan ejercicios de apertura de datos para que los investigadores puedan llevar a cabo nuevas investigaciones o replicar aquellas ya desarrolladas, además de ser un ejercicio de transparencia y rendición de cuentas a la ciudadanía.

Otra dificultad que han señalado los investigadores la falta de estandarización y como consecuencia, la dificultad para trabajar y cruzar datos de la misma temática de otras administraciones. Si nos fijamos en las valoraciones de los investigadores sobre los formatos técnicos de los conjuntos de datos, coinciden con los establecidos por la métrica para la evaluación de reutilización de datos abiertos MELODA (Abella, Ortiz de Urbina-Criado y De Pablos-Heredero, 2014 y 2019). GeoJSON y JSON, seguidos de RDF y SQL son los formatos mejor valorados tanto por la métrica MELODA como por los investigadores. CSV y KML reciben buenas valoraciones ya que son formatos con una semántica que facilita la reutilización. Finalmente, el formato PDF es el peor valorado por los investigadores coincidiendo con la métrica. Llama la atención que MELODA incluya dentro de su nivel máximo de reutilización (nivel 4) el formato RSS y, sin embargo, haya recibido 1,46 puntos sobre 3, en la encuesta, por debajo de otros formatos como XLS, que ha recibido 1,86 puntos o SHP, con 1,94 puntos y que ambos están en el nivel 2 de MELODA.

Con la valoración del nivel de reutilización de datos, en función del tipo de formato, sí se detectan formatos más propensos a ser reutilizables. Aquellos formatos que son procesables automáticamente son más reutilizados. Pero formatos propietarios como XLS son muy utilizados en ciencia, debido a un conocimiento general de la herramienta. No se observa una reutilización mayor o menor por área o franja de edad como consecuencia de obtener la información en determinados formatos técnicos. Las entidades encargadas de la apertura ofrecen los datos en uno o varios formatos y el reutilizador es el encargado de procesarlos de la manera más cómoda para llevar a cabo estudios.

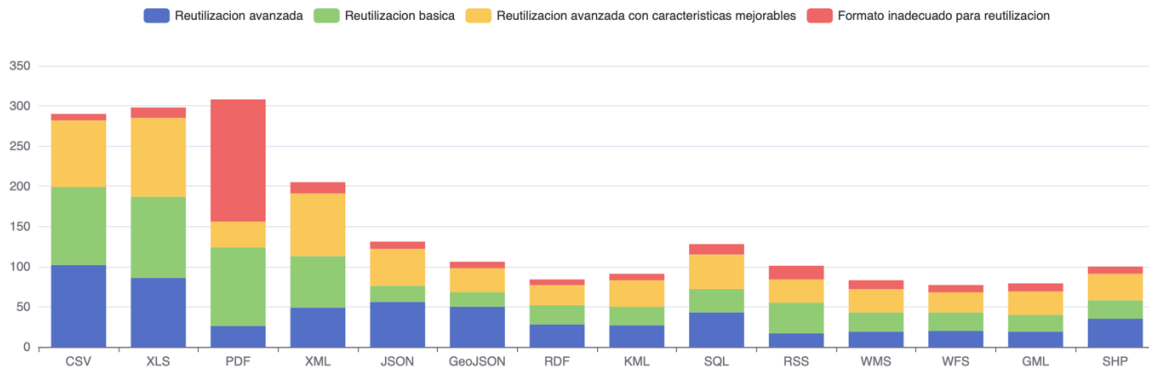


Figura 65: Nivel de reutilización por formatos técnicos.

En cuanto al producto final derivado de la reutilización de los datos abiertos, destacan las publicaciones científicas y otros trabajos académicos. Con la mitad de las respuestas, este uso de datos se destina a proyectos de investigación, a estudios e informes y docencia. Es decir, los datos abiertos tienen un impacto directo en la producción científica del país y en la consecución de méritos de investigación por parte del docente investigador universitario. El producto final se relaciona con el ámbito del proyecto. El número de elecciones del tipo de proyectos estatales, autonómicos e internacionales es muy similar. Y los conjuntos de datos utilizados es acorde a los proyectos de investigación que se fomentan y llevan a cabo desde distintos ámbitos de la administración, destacando los datos sobre medio ambiente y clima, población, geográficos y salud.

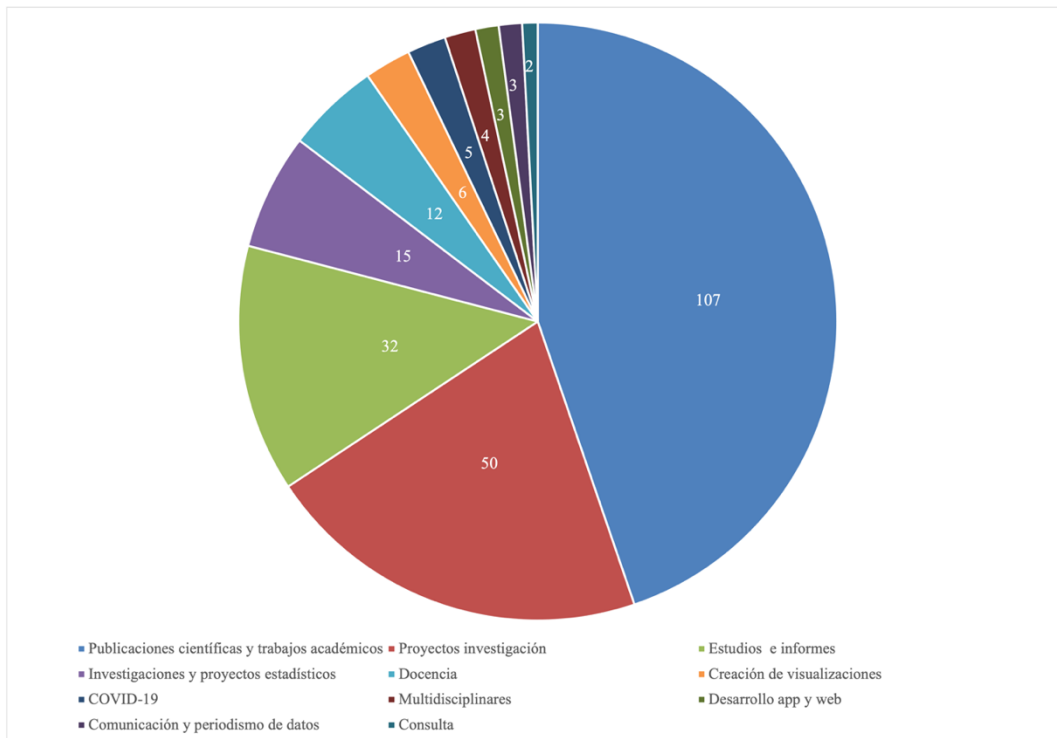


Figura 66: Uso concreto de datos abiertos.

Finalmente, en cuanto a procesamiento y tratamiento de la información, destaca la elección de hojas de cálculo, los softwares de estadística, procesadores de texto y herramientas de visualización de datos. Son los modos de trabajo más elegidos por los investigadores para llevar a cabo la minería de datos, la limpieza y el desarrollo.

Tecnología utilizada	Total	%
Hojas de cálculo	191	75,20
Software de estadística	114	44,88
Procesadores de texto	91	35,83
Herramientas de visualización de datos	81	31,89
Lenguajes de programación (PHP, Python,Ruby...)	65	25,59
Herramientas de mapeo de datos	55	21,65
XML	35	13,78
Otro (especifique)	28	11,02
JSON	27	10,63
Suites como Google	22	8,66
SPARQL	8	3,15
Flash	4	1,57

Tabla 62: Tecnología utilizada para el tratamiento y análisis de datos.



# Conclusiones, propuestas de futuro y reflexiones finales.

## Conclusiones

### 6.1 Conclusiones de la investigación

Las preguntas planteadas al inicio de la investigación surgen a raíz de una falta de literatura científica acerca de la percepción o el conocimiento que tiene la comunidad científica española sobre qué son los datos abiertos y cómo reutilizar información procedente del sector público. Cabe recalcar que este estudio se centra en el movimiento open data, como pilar del gobierno abierto, y en la disponibilidad de *datasets* que se pueden encontrar en portales de datos abiertos de entidades públicas.

Existen multitud de estudios que evalúan de diferentes formas la calidad de los conjuntos de datos abiertos y su nivel de reutilización en base a indicadores, o el sector infomediario y el valor que generan, estudios sobre conjuntos de datos concretos aplicados a un área determinada, o la percepción que tiene la ciudadanía en general acerca del *open data*. Estos estudios, de gran valor, utilidad y pioneros, se completan abordando una serie de perspectivas generales que permiten observar, de un modo panorámico, situaciones concretas de los datos abiertos. En algunas ocasiones, dentro de estas investigaciones, participan miembros de la comunidad científica, pero no son metodologías enfocadas directamente a la academia. Esta es la primera aportación de esta tesis doctoral. El estudio se centra en los investigadores y miembros de la comunidad universitaria, personal altamente cualificado dentro de la sociedad. En aras a mejorar estos servicios o fuentes de información, han sido críticos, y a través de las respuestas obtenidas en el cuestionario (aportando valor a través de opciones de respuesta abierta), las instituciones encargadas de la apertura de datos pueden respaldarse en este estudio, teniendo en cuenta opiniones, patrones de conducta y la percepción que tiene la comunidad científica española.

En los primeros informes y estudios realizados apenas se dispone de información. Son investigaciones pioneras, que arrojan una visión inicial necesaria, y sirven de punto de inflexión en los primeros pasos de trabajo de apertura de datos, además de servir como guías de actuación para las administraciones. No obstante, arrojan conclusiones muy optimistas y positivas. Sin embargo, para poder establecer conclusiones y recomendaciones tan específicas, se necesita tiempo de trabajo y observación. Se debe tener en consideración que se trata de un movimiento relativamente joven, y ser conscientes de que muchas administraciones no han llevado a cabo políticas de apertura de información y, por ende, no disponen de portales de datos abiertos. Pero no se trata únicamente de llevar a cabo políticas de apertura como fomento del gobierno abierto y reutilización. Existen casos de entidades que desarrollan iniciativas de apertura de datos, que no han tenido en cuenta la importancia del mantenimiento de las plataformas, del cumplimiento con la frecuencia de actualización definida en los *datasets*, de ciertos aspectos técnicos en materia de formatos o de cuidar el contenido y calidad de los conjuntos de datos para que sean procesables e interoperables.

En el caso de esta tesis, es la comunidad científica encuestada la que nos permite reforzar los discursos generales previos ya que, aparte de ser personal con conocimientos suficientes para la búsqueda, tratamiento y análisis de información, pueden ser en muchos momentos, potenciales reutilizadores de datos abiertos, desarrollando servicios e investigaciones basadas en datos, que aporten un valor añadido a la sociedad.

La muestra obtenida a través de la metodología planteada es representativa. Se obtienen 783 respuestas y, por tanto, los resultados y las conclusiones se pueden extrapolar al resto de la demografía docente universitaria española. Existe representación de 34 de las 50 provincias españolas y de las dos ciudades autónomas, y en concreto, de 47 universidades y 21 centros de investigación. Fijándonos en estudios previos acerca de la demografía española, se obtienen datos similares en cuanto a franjas de edad y categoría laboral a la que pertenecen los investigadores encuestados.

En materia de áreas de investigación se observa un sesgo de respuestas por parte de los investigadores que pertenecen al área 1, de Ciencias Sociales, con el volumen más alto de respuestas, situación que puede darse por la afinidad con este tipo de estudios cuantitativos, y porque reconocen su interés. Sin embargo, en el resto de las áreas no se observa un sesgo de respuesta tan claro, aunque existe. Se obtienen más respuestas de áreas que conocen y reutilizan datos que de áreas que no reutilizan o desconocen el movimiento.

Existe paridad de respuestas en materia de género ya que se obtiene un 41,76% de respuestas de mujeres, un 56,58% de respuestas de hombres, un 0,38% de intersexuales y un 1,28% sin respuesta. Por tanto, esta investigación garantiza una representación paritaria entre mujeres y hombres.

Al inicio de este estudio se plantean una serie de objetivos a abordar en el período temporal que ocupa esta tesis doctoral. Para ello, como se explica en el capítulo del método de investigación, se propone una metodología de investigación cuantitativa. Tras el análisis realizado con los datos obtenidos en el cuestionario procedentes de la muestra de la comunidad científica española, se da respuesta a los objetivos planteados y sirven para extraer las siguientes conclusiones.

Uno de los objetivos de esta tesis doctoral era conocer si los investigadores saben y entienden qué son los datos abiertos. Tras la pregunta planteada en la encuesta, se obtiene que un 79,31% afirman conocer este tipo de fuentes de información como recursos potencialmente reutilizables en ciencia. Por tanto, el porcentaje de conocimiento de datos abiertos se puede considerar elevado. Se debe prestar atención a aquellos investigadores que han indicado no conocer los datos abiertos, y a los que omiten la respuesta. Estos suponen un 20,69% del total. Apuntan hacia un desconocimiento de fuentes de información, de uso y tratamiento de los datos y del potencial que tienen, así como a una falta de difusión y publicidad activa por

parte de las administraciones. Es necesario que las entidades encargadas de llevar a cabo iniciativas de gobierno abierto a través de la apertura de datos presten atención a este tipo de cuestiones.

De aquellos investigadores que afirman conocer los datos abiertos, no interesa únicamente que estén al corriente del movimiento. La importancia radica en entenderlo, conocer las fuentes de datos y cómo se debe tratar la información para llevar a cabo la reutilización. El porcentaje de reutilización obtenido es bajo (50,57%) si se tiene en cuenta que el docente universitario es personal altamente cualificado dentro de la sociedad y conoce el método científico y las fuentes de datos e información, además de tener la capacidad de trabajar e interpretar datos para el desarrollo de nuevos estudios o servicios de valor añadido. Cuando se indaga para obtener respuestas acerca de la no reutilización de datos, los encuestados señalan varias razones. En primer lugar, apuntan a una carencia de datos de interés en sus áreas de investigación. Llama especialmente la atención del área 19, Biomedicina, con un alto volumen de conocimiento de datos abiertos pero un bajo nivel de reutilización. Otras disciplinas clásicas como el área 10 de Ciencias Físicas, la 14, de Ciencias y Tecnologías Químicas, la 8, de Psicología o la 4, de Mente, Lenguaje y Pensamiento apenas reutilizan datos procedentes del sector público apuntando también a una escasez de información. Por tanto, se debe reflexionar acerca de la falta de datos disponibles en los portales para tratar de abordar más áreas científicas.

Señalan también un desconocimiento de cómo llevar a cabo la reutilización. Se hacen necesarios ejemplos, guías prácticas de reutilización además de visores y herramientas en los portales de información pública que permitan una instrucción y sirvan de inspiración para aquellos investigadores menos familiarizados con el movimiento y la cultura del dato.

Es de vital importancia empezar a trabajar con la estandarización de determinados conjuntos de datos y que puedan compararse a otros procedentes de otras fuentes gracias a la normalización de estos. Permitirá una mayor reutilización con menor esfuerzo por parte de los reutilizadores ayudando a mejorar, además, a la interoperabilidad entre administrativa.

Apuntan también a una falta de fiabilidad de los datos, entre otros motivos, debido a errores y falta de actualización de los conjuntos de datos. No se trata únicamente de llevar a cabo políticas de apertura. Es necesario cumplir con la frecuencia de actualización estipulada en los portales ya que, la fiabilidad y la temporalidad son características vitales de la definición de los datos abiertos. Por ende, el grado de cumplimiento afectará positiva o negativamente en la reutilización.

Para que los datos sean reutilizables, existen atributos que deben cumplir los portales y los datasets y son muy valorados por los reutilizadores. Para generar valor añadido con los datos, consideran determinante una fácil localización y que no estén sujetos a restricciones, más

allá de la atribución de autoría de los datos. Consecuentemente, la variedad de tipologías de datasets y el suministro por parte de las administraciones encargadas de la apertura es otro punto que deben tener en consideración las entidades encargadas de la apertura, ya que, cuantos más conjuntos de datos de distintas áreas haya disponibles, más información será potencialmente reutilizable. Es relevante una correcta definición de los metadatos que permitan obtener más información y comparar el contenido de los datasets con otros conjuntos de datos relacionados. Finalmente, se valora positivamente la disponibilidad y variedad de formatos técnicos disponibles. Los investigadores valoran no solo los formatos procesables automáticamente por máquina, si no aquellos formatos técnicos que conocen, aunque sean “menos reutilizables” que otros de los cuales no tienen ningún conocimiento.

Uno de los problemas a los cuales se enfrenta la comunidad científica es la carga de trabajo y los plazos de entrega. Es fundamental trabajar en la accesibilidad de la información y que transcurra el menor tiempo posible entre la solicitud y la recepción de los datos. Existe un problema cultural generalizado, por parte de las administraciones, de temor a la pérdida de control de los datos y de no conciencia de su potencial. Por ello, en ocasiones, se establecen tarifas a las que se enfrentan los investigadores, que son señaladas como otra barrera a la hora de la reutilización de la información. Teniendo en cuenta la definición de dato abierto, deben estar libres sin ningún tipo de restricción o con un precio asequible.

Las administraciones deben rediseñar el contenido de determinados conjuntos de datos para que puedan ser fácilmente entendidos y de calidad. Como consecuencia, serán reutilizados más fácilmente. Las dificultades de comprensión de datos es otra barrera que han señalado los encuestados, a la que se enfrentan cuando empiezan a trabajar con los conjuntos de datos.

Los portales de datos abiertos de entidades públicas pueden llegar a ser fuentes de información de utilidad para determinadas áreas de la ciencia. Es cierto que un alto porcentaje de los investigadores encuestados son conscientes del movimiento. Pero como se ha señalado anteriormente, el objetivo no solo es estar al corriente de qué son los datos abiertos, sino ser conscientes de que se dispone de este tipo de fuentes de información potencialmente reutilizables. Se concluye que efectivamente son fuentes de información útiles en ciencia, sobre todo en áreas como las Ciencias Sociales, la Producción Industrial, la Ingeniería Civil y las Ingenierías para la Sociedad. También, en Tecnologías de la Información y las Comunicaciones, la Economía o las Ciencias y Tecnologías Medioambientales, áreas cuyos investigadores más han señalado que reutilizan datos. Se debe seguir trabajando en esta línea. Es interesante reflexionar sobre aquellas consideraciones que han aportado otros científicos procedentes de áreas que apenas reutilizan información del sector público porque no disponen de datos de interés en sus áreas de investigación.

En cuanto a los tipos de portales de datos abiertos más utilizados, los encuestados señalan el Catálogo Nacional de Datos Abiertos ([Datos.gob.es](http://datos.gob.es)) como el más consultado, ya que es un

punto de acceso único a conjuntos de datos de los portales de los distintos niveles de la administración: nacionales, autonómicos, locales, de universidades y otro tipo de fuentes financiadas con recursos públicos. Señalan también el acceso a portales de datos abiertos de organizaciones internacionales, vinculados directamente a proyectos de investigación en los que están involucrados. Uno de los usos que más se han señalado ha sido para la realización de investigaciones con fines estadísticos. Le sigue la elección de conjuntos de datos disponibles en portales autonómicos, entre ellos el de la Comunitat Valenciana. En una escala menor la comunidad científica indica un uso de datos abiertos disponibles en portales de entidades locales y entidades de iniciativa privada.

A pesar de que el nivel de reutilización se ha considerado bajo, sí que se existe producción científica derivada de la reutilización de información del sector público. Se traducen en publicaciones científicas y trabajos académicos, en proyectos de investigación, en docencia, o en el desarrollo de investigaciones comunicadas a través de artículos de investigación, para proyectos y con fines docentes, además de la creación de plataformas de visualizaciones de datos.

Para llevar a cabo la reutilización con los distintos fines especificados, se ha obtenido información acerca de los conjuntos de datos con más uso e interés para los investigadores. Los datos que señalan utilizar los docentes son aquellos relacionados con la ciencia y la tecnología, el sector público, la educación, el medio ambiente y clima, la salud, la demografía y población, o sociedad y bienestar. Tiene sentido si observamos las áreas científicas de donde proceden las respuestas. Esta información puede resultar muy valiosa para las administraciones porque, efectivamente, confirman que estas tipologías de datos son consultadas y reutilizadas para generar conocimiento a través de estudios científicos.

Paralelamente, un aspecto a mejorar es la publicación de nuevos conjuntos de datos relacionados con las áreas científicas que señalan una menor consulta y reutilización de datos de fuentes públicas. Debemos tener en cuenta que existen institutos de investigación y otros centros de trabajo, financiados con recursos públicos, que generan conocimiento en todas las disciplinas científicas. Es importante reflexionar acerca de si estos datos generados en estudios e investigaciones deben publicarse en portales de datos abiertos de las entidades que lo financian, además de en los repositorios de datos de investigación.

## 6.2 Limitaciones y propuestas de futuro

Los estudios cuantitativos, a pesar de aportar mucha información a partir de las respuestas al cuestionario (en este caso concreto se ha generado un gran volumen de datos) son, en algún sentido, investigaciones poco flexibles, y que no van a permitir estudiar casos concretos en profundidad. Tras la realización de este estudio, se obtiene una primera visión panorámica acerca de la percepción del dato abierto y de su reutilización. Quedan abiertas preguntas

concretas sobre marcos específicos y áreas particulares del conocimiento, que podrán abordarse en estudios posteriores.

Existen distintas líneas de investigación que se pueden considerar próximamente. Por un lado, se puede seguir profundizando en esta investigación a través de una metodología cualitativa. Es decir, seleccionar una muestra aleatoria de la comunidad científica con procedencias y vínculos a universidades de todo el país, para conocer de un modo más detallado, inquietudes y otras aportaciones de los científicos a través de una entrevista abierta o en grupos de discusión.

Otra investigación que puede ser interesante, con método indistinto, está relacionada con las áreas de investigación “reutilizadoras” y “no reutilizadoras”. De la mano de las áreas reutilizadoras, se puede profundizar en aquellos aspectos técnicos y de contenido que son relevantes para la reutilización. Pensando en áreas no reutilizadoras, puede ir orientada a la búsqueda de fuentes de información, y a llevar a cabo nuevas iniciativas de apertura de datos, cuya procedencia sea, por ejemplo, institutos o centros de investigación financiados con recursos públicos y que son generadores de datos potencialmente reutilizables.

Se pueden desarrollar estudios comparativos. Es decir, tomar como punto de partida los resultados de esta investigación, establecer indicadores, y comparar en los portales de datos abiertos de distintas administraciones, los contenidos, y cómo se presentan los datasets.

Este tipo de estudios no serán efectivos por sí solos, si no van acompañados de políticas y prácticas que ayuden a mejorar los servicios. Las entidades encargadas de la apertura deben reforzarse con normativas que fomenten nociones de gobernanza, reforzando la cultura del dato.

### 6.3 Reflexiones finales

Esta tesis doctoral se ha centrado en conocer la percepción, el conocimiento y el uso de los datos abiertos dentro de la comunidad científica española. Tras analizar la información obtenida a través del cuestionario, se muestran datos objetivos en materia de conocimiento y reutilización de datos, las barreras a las que se enfrentan los investigadores cuando acceden a estas fuentes de información, los atributos que se espera que cumplan los datos procedentes de entidades públicas, las tipologías de datos que resultan de interés y son reutilizados, los productos y servicios de valor añadido generados, que se traducen en producción científica basada en datos abiertos.

No ha sido su objetivo principal analizar los beneficios económicos que pueden generar un sector como el infomediario, tema sobre el que se han realizado otros estudios que se han comentado a lo largo de esta tesis. Tampoco se ha realizado ningún análisis concreto sobre

conjuntos de datos relacionados con un área de investigación determinada, aunque se tiene presente en todo momento que estos estudios son interesantes, y sobre los que queda constancia en la revisión realizada en el marco teórico.

Este estudio toma en cuenta la opinión del investigador, ciudadano cualificado dentro de la sociedad, para reivindicar la necesidad de seguir invirtiendo en la creación de portales de datos abiertos como fuentes de información potencialmente reutilizables en ciencia y generadores de conocimiento. Las entidades encargadas de la apertura pueden tener en este documento una referencia para ver el modo de mejorar las hojas de ruta que están desarrollando en la apertura y compartición de datos.

Algunas percepciones obtenidas son similares a otras que han aportado anteriormente estudios sobre datos abiertos. Sin embargo, hay cuestiones concluyentes y que afirman ser relevantes en otras investigaciones que, en este caso, no son consideradas significativas para la reutilización de datos. Por ejemplo, los investigadores no consideran decisivo para la producción científica que los datos tengan una frecuencia de actualización en tiempo real, o que estén geolocalizados; sin embargo, valoran que haya guías de reutilización, ejemplos, que sean accesibles y sin coste.

Las administraciones públicas pueden extraer de este trabajo algunas ideas sobre la forma de proceder en el desarrollo de las políticas de datos abiertos. La investigación sugiere que se trata de seguir trabajando con el objetivo no solo de abrir datos, sino de que los datos que proporcionen sean de calidad. Se entiende que es una gran inversión, ya que, si nos fijamos en el precio de los datos y de las infraestructuras necesarias en el sector privado, son poco económicos. Por tanto, no solo se trata de llevar a cabo políticas de apertura de información como pilares del gobierno abierto, que fomenten transparencia administrativa, sino de que, como se ha insistido en varias ocasiones, sean utilizados para generar conocimiento de valor añadido, y para este caso en concreto, producción científica de calidad. Es necesario realizar un trabajo desde el diseño del dato: situando el origen en las bases de datos que los almacenan, centrándose en el contenido de estos, y realizando una labor de diseño de descriptores y metadatos para que sean accesibles, interoperables, y finalmente, utilizados. Naturalmente, el esfuerzo por la compartición de buenos datos ayudaría a la reutilización.



## Bibliografia

---

## Bibliografía

- Abella, A., Ortiz-de-Urbina-Criado, M., De-Pablos-Heredero, C. (2014). Meloda, métrica para evaluar la reutilización de datos abiertos. *El Profesional de la Información*, 23(6), 582–588. DOI: <https://doi.org/10.3145/epi.2014.nov.04>
- Abella, A., Ortiz-de-Urbina-Criado, M., De-Pablos-Heredero, C. (2017). “La reutilización de datos abiertos: una oportunidad para España”. Madrid: Fundación Cotec para la Innovación. URL: <https://cotec.es/proyecto/reutilizacion-de-datos-abiertos-una-oportunidad/ee68d249-51dd-4e53-8a3c-b6e32442ffc3>
- Abella, A., Ortiz-de-Urbina-Criado, M., De-Pablos-Heredero, C. (2018). Indicadores de calidad de datos abiertos: el caso del portal de datos abiertos de Barcelona. *El Profesional de la Información*, 27(2), 375-382. DOI: <https://doi.org/10.3145/epi.2018.mar.16>
- Abella, A., Ortiz de Urbina Criado, M., De Pablos Heredero, C., Vidal-Cabo, C. y Ferrer Sapena, A. (2019). “La reutilización de datos abiertos en España II”. Madrid: Desidedatum, Universidad Rey Juan Carlos y Cátedra de Transparencia y Gestión de Datos UPV. URL: [https://www.desidedatum.com/wp-content/uploads/2019/12/La\\_reutilizacio%CC%81n\\_datos\\_abiertos\\_en\\_espan%CC%83\\_a\\_2019.pdf](https://www.desidedatum.com/wp-content/uploads/2019/12/La_reutilizacio%CC%81n_datos_abiertos_en_espan%CC%83_a_2019.pdf)
- Abella, A., Ortiz-de-Urbina-Criado, M., De-Pablos-Heredero, C. (2020). Meloda 5: A metric to assess open data reusability. *El Profesional de la Información*, 28(6). DOI: <https://doi.org/10.3145/epi.2019.nov.20>
- Acosta, S. I., y Plata Pineda, O. J. (2016). Participación, colaboración, transparencia y datos abiertos. Tras la definición de gobierno abierto. TEUKEN BIDIKAY. *Revista Latinoamericana De Investigación En Organizaciones, Ambiente Y Sociedad*, 7(9), 209-222. URL: <https://revistas.elpoli.edu.co/index.php/teu/article/view/1016>
- Alcalá, M. y Anglada, L. (2019). FAIR x FAIR. Una aproximación pragmática a la gestión de datos de investigación en modo FAIR. *Anuario ThinkEPI*, 13, e13e05. DOI: <https://doi.org/10.3145/thinkepi.2019.e13e05>
- Aleixandre-Benavent, R., Ferrer-Sapena, A. y Peset, F. (2019). Compartir recursos útiles para la investigación: datos abiertos (open data). *Elsevier Médica*. DOI: <https://doi.org/10.1016/j.edumed.2019.07.004>

- Aleixandre-Benavent, R., Vidal-Infer, A., Alonso-Arroyo, A., Peset, F. y Ferrer-Sapena, A. (2020). Research Data Sharing in Spain: Exploring determinants, practices and perceptions. *MDPI* 5(2), 29. DOI: <https://doi.org/10.3390/data5020029>
- Alguacil-Sanz, M. (2015). De la gestión inteligente al compromiso de servicio. *Archivamos*, 97. URL: <https://publicaciones.acal.es/archivamos/article/view/511>
- Álvarez-García, S., Gétrudix-Barrio, M. y Rajas-Fernández, M. (2014): La construcción colaborativa de bancos de datos abiertos como instrumento de empoderamiento ciudadano. *Revista Latina de Comunicación Social*, 69, pp. 661-683. DOI: 10.4185/RLCS-2014-1029
- Álvarez-García, S., Gertrudis-Casado, M. C. y Gétrudix-Barrio, M. (2021). Comunicación digital basada en datos abiertos y la Agenda 2030 en Latinoamérica: nuevas oportunidades. *Comunicación y Sociedad*, e7615, pp. 1-20. DOI: <https://doi.org/10.32870/cys.v2021.7615>
- Angulo López, E. (2011). Política fiscal y estrategia como factor de desarrollo de la mediana empresa comercial sinaloense. Un estudio de caso. México. Estudio cuantitativo. Universidad Autónoma de Sinaloa. Tesis Doctoral. Pp. 115-116 URL: <https://defiscal.posgrado.fca.uas.edu.mx/wp-content/uploads/2020/07/eleazar-angulo.pdf>
- Anteló Martínez, A.R. (2016). Elaboración de disposiciones de carácter general en la Administración Local, tras la entrada en vigor de la Ley 39/2015. *Revista de Estudios de la Administración Local y Autonómica*, (6) pp. 105-121. DOI: <https://doi.org/10.24965/reala.v0i5.10391>
- Arquero Avilés, R. y Marco Cuenca, G. (2014). El Portal de datos abiertos de la Unión Europea. *Revista General de Información y Documentación*. 24(1), pp. 99-118. DOI: [http://dx.doi.org/10.5209/rev\\_RGID.2014.v24.n1.45384](http://dx.doi.org/10.5209/rev_RGID.2014.v24.n1.45384)
- ASEDIE (2016). Informe del Sector Infomediario: 2016. URL: [https://repositorio.aemet.es/bitstream/20.500.11765/6217/1/OpenData\\_LHC.pdf](https://repositorio.aemet.es/bitstream/20.500.11765/6217/1/OpenData_LHC.pdf)
- ASEDIE (2020). Informe del Sector Infomediario: 2020. URL: <http://www.asedie.es/assets/informe-asedie-sobre-el-sector-infomediario-2020-vf.pdf>
- ASEDIE (2021). Informe del Sector Infomediario: 2021. URL: <https://static1.squarespace.com/static/600a99c4d2a8133c3599fc67/t/607ff089596f9b27bfc11636/1618997397041/ISI21+VF++web.pdf>

- Atenas, J., Piedra, N., Priego, E. y Havemann, L. (2015). Cómo los Datos Abierto pueden ayudar a la educación universitaria. *BID: Mejorando vidas*. URL: <https://blogs.iadb.org/conocimiento-abierto/es/como-los-datos-abiertos-pueden-ayudar-a-potenciar-la-educacion-universitaria/>
- Barros, A. (2016). Datos Abiertos: ¿qué son y cuál es su utilidad? *La Promesa del Gobierno Abierto*. URL: [https://issuu.com/clog\\_unlp/docs/datos-abiertos-que-son-y-cual-es-su](https://issuu.com/clog_unlp/docs/datos-abiertos-que-son-y-cual-es-su)
- Baz Rodríguez, J. (2019). La Ley Orgánica 3/2018 como marco embrionario de garantía de los derechos digitales laborales. Claves para un análisis sistemático. *Trabajo y derecho: nueva revista de actualidad y relaciones laborales*, (54), 49-78.
- Berners-Lee, Tim (2010). 5 estrellas Datos Abiertos. URL: <https://5stardata.info/es/>
- Bertot, J.C., Gorham, U., Jaeger, P.T.; Sarin, L.C. and Choi, H. (2014). Big data, open government and e-government: Issues, policies and recommendations. *Information polity*, 19(1), pp. 5-16. DOI: <https://doi.org/10.3233/IP-140328>
- Boletín Oficial del Estado (2017). Real Decreto 1/2017, de 13 de enero, por el que se modifica el Real Decreto 1067/2015, de 27 de noviembre, por el que se crea la Agencia Estatal de Investigación y se aprueba su Estatuto. BOE núm. 12, de 15 de enero de 2017, pp. 3524-3525. URL: [https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2017-404](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2017-404)
- Bonsón, E., Torres, L., Royo, S. y Flores, F. (2012). Local e-government 2.0: Social media and corporate transparency in municipalities. *Government Information Quarterly*, 29(2), 123-132. DOI: <https://doi.org/10.1016/j.giq.2011.10.001>
- Cabo, D. y Magallón-Rosa, R. (2013). Nuevos retos para las Administraciones Públicas. Datos, cultura cuantitativa y calidad democrática. *Telos: Cuadernos de comunicación e innovación*, (95), 77-86. URL: <https://telos.fundaciontelefonica.com/archivo/numero095/datos-cultura-cuantitativa-y-calidad-democratica/?output=pdf>
- Cadena-Iñiguez, P., Rendón-Medel, R., Aguilar-Ávila, J., Salinas-Cruz, E., Cruz-Morales, F., y Sangerman-Jarquín, D. M. (2017). Métodos cuantitativos, métodos cualitativos o su combinación en la investigación: un acercamiento en las ciencias sociales. *Revista mexicana de ciencias agrícolas*, 8(7), 1603-1617. URL: <https://www.redalyc.org/pdf/2631/263153520009.pdf>

- Cadena Vela, S.G. (2019). Marco de referencia para la publicación de datos abiertos comprensibles basado en estándares de calidad. Universidad de Alicante. Tesis Doctoral, TIN2016-78103-C2-2R.
- Calderón, C. y Lorenzo, S. (2010). *Open government: gobierno abierto*. Jaén, España: Algón Editores MMX, 2010. URL: <https://libros.metabiblioteca.org/handle/001/163>
- Campos, Concepción (2018). 10 puntos que debes conocer (YA) de la nueva LOPD y GDD. URL: <https://concepcioncampos.org/10-puntos-que-debes-conocer-ya-de-la-nueva-lopd-y-gdd/>
- Campos Domínguez, E. y Corojan, A. (2013). Estado del arte del gobierno abierto: promesas y expectativas. En Ramírez-Alujas, Á., Hofmann, A., y Bojórquez Perenieta, J.A. (Coords.): *La promesa del Gobierno Abierto* (pp. 119-136). México: ITAIP e InfoDF. <http://www.itaip.org.mx/slide/pdf/lpga.pdf>
- Carretero Sánchez, S. (2012). Democracia virtual y participación ciudadana en la legitimidad del poder político. *Revista telemática de filosofía del derecho*, 15, 105-131. URL: <https://dialnet.unirioja.es/servlet/articulo?codigo=3902419>
- Carta Internacional de Datos Abiertos. Wikipedia. URL: [https://es.wikipedia.org/wiki/Carta\\_Internacional\\_de\\_los\\_Datos\\_Abiertos](https://es.wikipedia.org/wiki/Carta_Internacional_de_los_Datos_Abiertos)
- Casas Anguita, J., Repullo Labrador, J.R. y Donado Campos, J. (2003). La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos (I). *Atención primaria*, 31(8), 527-538. URL: <https://www.elsevier.es/es-revista-atencion-primaria-27-pdf-13047738>
- Casas Anguita, J., Repullo Labrador, J.R., y Donado Campos, J. (2003). La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos (II). *Atención primaria*, 31(9), 592-600. DOI: [https://doi.org/10.1016/S0212-6567\(03\)79222-1](https://doi.org/10.1016/S0212-6567(03)79222-1)
- Cerezo, J.M. (2010). Hacia la ciudadanía digital. Innovación para el empoderamiento ciudadano a través de las TIC. *Fundación Cibervoluntarios*. URL: <https://es.scribd.com/document/64190834/Innovacion-para-el-Empoderamiento-Ciudadano-a-traves-de-las-TIC-Cibervoluntarios>
- Cerrillo, A. (2014). Los principios de los datos abiertos en la legislación española. *Revista de Internet, Derecho y Política - Revista d'Internet, dret i política*, (19), 62-77. URL: <https://dialnet.unirioja.es/servlet/articulo?codigo=5582974>

- Cerrillo-Martínez, A. (2018). Datos masivos y datos abiertos para una gobernanza inteligente. *El Profesional de la Información*, 27(5), pp. 1128-1135. DOI: <https://doi.org/10.3145/epi.2018.sep.16>
- Chinchilla Peinado, J.A. (2016). Validez y eficacia. Avances y retrocesos en la Ley 39/2015, del Procedimiento Administrativo Común de las Administraciones Públicas. Su incidencia en la Administración Local. *Cuadernos de Derecho Local: Fundación Democracia y Gobierno Local*. URL: [http://repositorio.gobiernolocal.es/xmlui/bitstream/handle/10873/1719/06\\_CHINCHILLA\\_P86\\_122\\_QDL\\_41.pdf?sequence=1](http://repositorio.gobiernolocal.es/xmlui/bitstream/handle/10873/1719/06_CHINCHILLA_P86_122_QDL_41.pdf?sequence=1)
- Clabo, N. y Ramos-Vielba, I. (2015). Reutilización de datos abiertos en la administración pública en España y uso de licencias-tipo. *Revista Española de Documentación Científica*, 38(3): e097. DOI: <http://dx.doi.org/10.3989/redc.2015.3.1206>
- Coglianesse, C. (2009). The Transparency President? The Obama Administration and Open Government. *Governance, Issue Governance*, 22(4), 529-544. Garsington Road, Oxford: Wiley Periodicals, Inc. DOI: <https://doi.org/10.1111/j.1468-0491.2009.01451.x>
- Comisión de las Comunidades Europeas (1998). *La información del sector público: un recurso clave para Europa. Libro Verde sobre la información del sector público en la sociedad de la información*. URL: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:51998DC0585>
- Comisión de las Comunidades Europeas (2009). *Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones: Reutilización de la información del Sector Público. Revisión de la Directiva 2003/98/CE*. URL: [http://dgtic.caib.es/www/caibdata/normativa/Comunicacion\\_Directiva\\_2003\\_98\\_CE.pdf](http://dgtic.caib.es/www/caibdata/normativa/Comunicacion_Directiva_2003_98_CE.pdf)
- Comisión de las Comunidades Europeas (2017). *Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Marco Europeo de Interoperabilidad – Estrategia de aplicación*. Bruselas. URL: <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52017DC0134&from=LT>
- Comisión Económica para América Latina y el Caribe-CEPAL. (2018a). *Agenda 2030 y los Objetivos de Desarrollo Sostenible: Una oportunidad para América Latina y el Caribe*. CEPAL. URL: <https://www.cepal.org/es/publicaciones/40155-la-agenda-2030-objetivosdesarrollo-sostenible-oportunidad-america-latina-caribe>

- Comisión Europea (2011). Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las regiones: Datos abiertos. Un motor para la innovación, el crecimiento y la gobernanza transparente. Bruselas. URL: <https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52011DC0882&from=EN>
- Comunicado de la Comisión sobre la Reutilización de la Información del Sector Público. Revisión de la Directiva 2003/98/CE (sin carácter legislativo). URL: <https://datosabiertos.jcyl.es/web/jcyl/RISP/es/Plantilla100Detalle/1284162108799/Normativa/1284163264241/Redaccion>
- Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones (2020). *Una Estrategia Europea de Datos. COM/2020/66 final*. URL: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52020DC0066>
- Consejo de Transparencia y Buen Gobierno (2015). *Código ético. Consejo de Transparencia y Buen Gobierno*. URL: [https://www.consejodetransparencia.es/dam/jcr:7ac8878b-1184-40e1-b226-da3fda3566c1/Codigo\\_etico.pdf](https://www.consejodetransparencia.es/dam/jcr:7ac8878b-1184-40e1-b226-da3fda3566c1/Codigo_etico.pdf)
- Cook, T. D., y Reichardt, C. S. (1986). *Métodos cualitativos y cuantitativos en investigación evaluativa*. Madrid: Morata. URL: [https://www.fceia.unr.edu.ar/geii/maestria/2014/DraSanjurjo/12de20/Cook\\_Reichardt.pdf](https://www.fceia.unr.edu.ar/geii/maestria/2014/DraSanjurjo/12de20/Cook_Reichardt.pdf)
- Coroan, A. y Campos Domínguez, E.M. (2011). *Gobierno abierto: alcance e implicaciones*. Fundación Ideas. URI: <http://uvadoc.uva.es/handle/10324/42441>
- Cotino Hueso, L. (2012). Derecho y “Gobierno Abierto”. La regulación de la transparencia y la participación y su ejercicio a través del uso de las nuevas tecnologías y las redes sociales por las administraciones públicas. Propuestas concretas, *Monografías de la Revista Aragonesa de Administración Pública*, n.º XIV, pp. 51-92. URL: <https://dialnet.unirioja.es/servlet/articulo?codigo=5527789>
- Cotino Hueso, L. (2014). El derecho fundamental de acceso a la información, la transparencia de los poderes públicos y el gobierno abierto. Retos y oportunidades. En: J. Valero Torrijos y M. Fernández Salmerón (eds.) Régimen jurídico de la transparencia del sector público, p. 68. URL: <https://dialnet.unirioja.es/servlet/articulo?codigo=4762086>

- Curto Rodríguez, R. (2017). *Datos abiertos y rendición de cuentas en las comunidades autónomas españolas: niveles de divulgación determinantes*. Universidad de Oviedo. Tesis Doctoral, 2017.
- Curto-Rodríguez, R. (2021). Análisis multidimensional de los portales de datos abiertos autonómicos españoles. *Revista Española De Documentación Científica*, 44(1), e284. DOI: <https://doi.org/10.3989/redc.2021.1.1745>
- Dassen, N., y Cruz, J. (2012). *Gobierno abierto y transparencia focalizada. Tendencias y desafíos para América Latina y el Caribe*. Washington, DC: bid. URL: [https://www.researchgate.net/publication/332494886\\_Gobierno\\_Abierto\\_y\\_Transparencia\\_Focalizada\\_Tendencias\\_y\\_Desafios\\_para\\_America\\_Latina\\_y\\_el\\_Caribe](https://www.researchgate.net/publication/332494886_Gobierno_Abierto_y_Transparencia_Focalizada_Tendencias_y_Desafios_para_America_Latina_y_el_Caribe)
- Data.Europa.eu (s.f.). ¿Qué son los datos abiertos? Data Europa Eu. URL: <https://data.europa.eu/es/trening/what-open-data>
- Datos.Gob.es (2014). Licencias de uso asociadas a las iniciativas de datos abiertos en España. *Gobierno de España: Iniciativa Aporta y Red.es*. URL: <https://datos.gob.es/es/noticia/licencias-de-uso-asociadas-las-iniciativas-de-datos-abiertos-en-espana>
- Datos.Gob.es (2014a). Proyecto Aporta. *Gobierno de España: Iniciativa Aporta y Red.es*. URL: <https://datos.gob.es/es/documentacion/proyecto-aporta>
- Datos.Gob.es (2014b). Caracterización del Sector Infomediario en España. Sector Público. *Gobierno de España: Iniciativa Aporta y Red.es*. URL: [https://datos.gob.es/sites/default/files/doc/file/ppt\\_infomediario\\_publico.pdf](https://datos.gob.es/sites/default/files/doc/file/ppt_infomediario_publico.pdf)
- Datos.Gob.es (2015). Plataformas de publicación de datos abiertos. *Gobierno de España: Iniciativa Aporta y Red.es*. URL: <https://datos.gob.es/sites/default/files/informe-herramientas-publicacion.pdf>
- Datos.Gob.es (2017a). Manual práctico para mejorar la calidad de los datos abiertos. *Gobierno de España: Iniciativa Aporta y Red.es*. URL: [https://datos.gob.es/sites/default/files/doc/file/manual\\_practico\\_para\\_mejorar\\_la\\_calidad\\_de\\_los\\_datos\\_abiertos\\_1.pdf](https://datos.gob.es/sites/default/files/doc/file/manual_practico_para_mejorar_la_calidad_de_los_datos_abiertos_1.pdf)



- Datos.Gob.es (2017b). Reutilización de Información del Sector Público: Conceptos básicos, beneficios del Open Data y barreras para su aplicación. *Gobierno de España: Iniciativa Aporta y Red.es*. URL: [https://datos.gob.es/elearning/Unidades\\_Didacticas/Unidad\\_1/contenidos/descargas/unidad\\_imprimible.pdf](https://datos.gob.es/elearning/Unidades_Didacticas/Unidad_1/contenidos/descargas/unidad_imprimible.pdf)
- Datos.Gob.es (2018). ¿Cuál es el nivel de madurez de los datos abiertos en España? *Gobierno de España: Iniciativa Aporta y Red.es*. URL: <https://datos.gob.es/es/noticia/cual-es-el-nivel-de-madurez-de-los-datos-abiertos-en-espana>
- Davies, T. (2010). Open data, democracy and public sector reform. A look at open government data use from data.gov.uk. *Data.Gov.uk*, pp. 1-47.
- Davies, T. (2010). Open Data Impacts Survey: <http://www.opendataimpacts.net/wp-content/uploads/2010/06/OpenDataImpacts-Survey-Form.pdf>
- DCAT-AP y datos abiertos: aportaciones y nuevas líneas de trabajo (2016). URL: [https://administracionelectronica.gob.es/pae\\_Home/pae\\_Actualidad/pae\\_Noticias/Anio2016/Febrero/Noticia-2016-02-16-DCAT-AP-datos-abiertos-aportaciones-y-nuevas-lineas-de-trabajo.html#.XKrsaeszZhH](https://administracionelectronica.gob.es/pae_Home/pae_Actualidad/pae_Noticias/Anio2016/Febrero/Noticia-2016-02-16-DCAT-AP-datos-abiertos-aportaciones-y-nuevas-lineas-de-trabajo.html#.XKrsaeszZhH)
- Decreto 56/2016, del Consell, de 6 de mayo, por el que se aprueba el Código de Buen Gobierno de la Generalitat. DOGV núm. 7781 de 13.05.2016. URL: <https://www.dogv.gva.es/es/eli/es-vc/d/2016/05/06/56/>
- Decreto 105/2017, de 28 de julio, del Consell, de despliegue de la Ley 2/2015, de 2 de abril, de la Generalitat, en materia de transparencia y de regulación del Consejo de Transparencia, Acceso a la Información Pública y Buen Gobierno. DOGV núm. 8118 de 01.09.2017. URL: [http://www.dogv.gva.es/va/disposicio?sig=007767/2017&L=1#\\_blank](http://www.dogv.gva.es/va/disposicio?sig=007767/2017&L=1#_blank)
- De Pelekais, Cira (2000). Métodos cuantitativos y cualitativos: diferencias y tendencias. *Telos*, Vol. 2(2): pp. 374-352. URL: <http://ojs.urbe.edu/index.php/telos/article/view/997/943>
- Desantes Fernández, Blanca y Pérez Alcázar, Ricardo (2015). La reutilización de la información del sector público (RISP). Su aplicación en el sector cultural. *Tabula*, 18, pp. 45-72. URL: <https://blog.bne.es/biblioteconomia/2016/08/12/la-reutilizacion-de-la-informacion-del-sector-publico-risp-su-aplicacion-en-el-sector-cultural/?print=pdf>

- Díez, E.J. y Millán, R.A. (2011). Open Data y RISP: generando valor social y económico. Transparencia e innovación en la eAdministración. Santiago de Compostela: *Colexio Profesional de Enxeñaría en Informática de Galicia*. URL: <https://goo.gl/9WxdC7>
- Directiva Europea 2003/98/CE del Parlamento Europeo y del Consejo, de 17 de noviembre de 2003, relativa a la reutilización de la información del sector público. URL: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:ES:PDF>
- Directiva 2007/2/CE del Parlamento Europeo y del Consejo, de 14 de marzo de 2007, por la que se establece una infraestructura de información espacial en la Comunidad Europea (Inspire). «DOUE» núm. 108, de 25 de abril de 2007, páginas 1 a 14 (14 págs.). Disponible en: <https://boe.es/buscar/doc.php?id=DOUE-L-2007-80587>
- Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, modifica la Directiva 2003/98/CE relativa a la RISP. URL: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:ES:PDF>
- Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo de 20 de junio de 2019 relativa a los datos abiertos y la reutilización del sector público. *Noticias Jurídicas*, 2019. URL: <https://noticias.juridicas.com/actualidad/noticias/14173-nueva-directiva-europea-sobre-datos-abiertos-y-reutilizacion-de-la-informacion-del-sector-publico/>
- Documento de trabajo de los servicios de la Comisión sobre la Comunicación de la Comisión sobre Reutilización de la Información del Sector Público (sin carácter legislativo). [https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52014XC0724\(01\)&from=LT](https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52014XC0724(01)&from=LT)
- Drüke, H. (2007). Can e-government make public governance more accountable? En Shah, A. (Ed.): *Performance accountability and combating corruption* (pp.59-87). The World Bank, Washington, D.C.
- Emmerich, G. E. (2004). “Transparencia, rendición de cuentas, responsabilidad gubernamental y participación ciudadana”. *Polis: Investigación y Análisis Sociopolítico y Psicosocial*, 2 (4), 67-90. URL: <https://www.redalyc.org/pdf/726/72620404.pdf>
- EP Data (S.F.). Profesores, investigadores y resto de personal docente en la universidad de España, datos y estadísticas. Portal de visualización de datos actualizado en noviembre de 2021. URL: <https://www.epdata.es/datos/profesores-investigadores-resto-personal-docente-universidad-espana-datos-estadisticas/412>

- Estrada Zamora, C. y Vargas-Hernández, J. G. (2018). Percepción ciudadana del open data y la innovación en México. *RITI Journal*, 6(11) Pp. 86-95. URL: <https://dialnet.unirioja.es/servlet/articulo?codigo=7107376>
- European Commission (2020). The Economic Impact of Open Data: Opportunities for value creation in Europe. Disponible en: <https://www.europeandataportal.eu/sites/default/files/the-economic-impact-of-open-data.pdf>
- European Data Portal (S.F.). Definición Datos Abiertos. URL: <https://www.europeandataportal.eu/en/providing-data/goldbook/open-data-nutshell>
- European Data Portal (2015-2021). Open Data Maturity in Europe. Report. URL: <https://data.europa.eu/es/impact-studies/open-data-maturity>
- Ferrer Sapena, A.; Calabuig, J.M.; Sánchez-Pérez, E.A. y Vidal-Cabo, C. (2020). Gobierno abierto y acceso a la información: un estudio de caso sobre el impacto en la economía local. *BiD: textos universitaris de Biblioteconomia i Documentació*. (45) pp. 1-25. <https://doi.org/10.1344/BiD2020.45.17>
- Ferrer-Sapena, A., Peset, F. y Aleixandre-Benavent, R. (2011). Acceso a los datos públicos y su reutilización: open data y open government". *El Profesional de la Información*, 20(3), pp. 260-269. DOI: 10.3145/epi.2011.may.03
- Ferrer Sapena, A. y Peset, F. (2021). Ciencia abierta: la importancia de los datos abiertos en la I+D. En: *II Ciclo Diálogos en el Territorio. Universitat d'Alacant*. URL: [http://eprints.rclis.org/41900/1/Presentacion\\_alicante%20fp\\_md\\_def25to.pdf](http://eprints.rclis.org/41900/1/Presentacion_alicante%20fp_md_def25to.pdf)
- Ferrer-Sapena, A., y Sánchez-Pérez, E. (2013). Open data, big data: ¿hacia dónde nos dirigimos? *Anuario ThinkEPI 2013*, 7, 150-156. URL: <http://eprints.rclis.org/21006/>
- Fundación Cotec (2003). Nuevos mecanismos de transferencia de tecnología. *Fundación Cotec para la Innovación Tecnológica*. Madrid, 2003. Pág. 49.
- Galindo Ayuda, F. (2014). La regulación de los datos abiertos. *Ibersid*, 8, pp. 13-18. URL: <https://ojs.ibersid.eu/index.php/ibersid/article/view/4169/3791>
- Gallego-Álvarez, I., Rodríguez-Domínguez, L. y García-Sánchez, I.-M. (2010). Are determining factors of municipal E-government common to a worldwide municipal view? An intra-country comparison. *Government Information Quartely*, 27(4), pp. 423-430. DOI: <https://doi.org/10.1016/j.giq.2009.12.011>

- Gamero Casado, E. (2009). Interoperabilidad y Administración electrónica: conéctense, por favor. *Revista de Administración Pública*, (179), 291-332. URL: <https://recyt.fecyt.es/index.php/RAP/article/view/45833>
- García García, J. (2014). Gobierno abierto: transparencia, participación y colaboración en las Administraciones Públicas. *Innovar Revista de Ciencias Administrativas y Sociales*, 24(54), pp. 75-88. URL: <https://www.redalyc.org/pdf/818/81832222006.pdf>
- García-Valdecasas Algüacil, M. (2016). Resumen de la Ley 39/2015, de 1 de octubre, del procedimiento administrativo común de las administraciones públicas. URL: <https://www.notariosyregistradores.com/web/normas/destacadas/resumen-de-la-ley-392015-de-1-de-octubre-del-procedimiento-administrativo-comun-de-las-administraciones-publicas/>
- Garriga-Portolà, M. (2011). ¿Datos abiertos? Sí, pero de forma sostenible. *El Profesional de la Información*. 20(3), pp. 298-303. DOI: <https://doi.org/10.3145/epi.2011.may.08>
- Gértrudix-Barrio, F., Gálvez de la Cuesta, M.C., Said-Hung, E. y Durán-Medina, J.F. (2016). Alfabetización digital, competencias mediáticas y Open Data. *RELATEC*, 15(2), pp. 113-121. DOI: <https://doi.org/10.17398/1695-288X.15.2.113>
- Gértrudix, M., Gertrudis-Casado, M.C. y Álvarez-García, S. (2016). Consumption of public institutions' open data by Spanish citizens. *El Profesional de la Información*, 25(4), pp. 535-544. DOI: <https://doi.org/10.3145/epi.2016.jul.03>
- González Bustos, M. Á. (2016). Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público [BOE n.º 164, de 10-VII-2015]. *AIS: Ars Iuris Salmanticensis*, 4 (1), 195-197.
- González Tapia, María Luisa (2018). Los derechos digitales en la Ley Orgánica 3/2018. *Diario La Ley*, (9324), 1.
- Gurstein, M. (2011). Open data: Empowering the empowered or effective data use for everyone? *First Monday*, 16(2). DOI: <https://doi.org/10.5210/fm.v16i2.3316>
- Hernández, O. (2012). *Estadística Elemental para Ciencias Sociales*. (Tercera Edición). San José, Costa Rica: Editorial Universidad de Costa Rica.

- Hernández-Pérez, T. y García-Moreno, M.A. (2013). Datos abiertos y repositorios de datos: nuevo reto para los bibliotecarios. *El Profesional de la Información*, 22(3), pp. 259-263. DOI: <http://dx.doi.org/10.3145/epi.2013.may.10>
- Hernández Sampieri, R., Fernández Collado, C. y Pilar Baptista L. (2010). Metodología de la investigación. México. Ed. Mc Graw Hill.
- Hípola, P. y Muñoz-Cañavate, A. (2015). Evolución del mercado de la información electrónica: desde los orígenes hasta el auge del sector infomediario. *Revista Española de Documentación Científica*, 38(2), e082. DOI: <http://dx.doi.org/10.3989/redc.2015.2.1174>
- ICSU-IAP-ISSC-TWAS (2016). Datos abiertos en un mundo de grandes datos: un acuerdo internacional (versión abreviada). URL: [http://www.science-international.org/sites/default/files/reports/datos-abiertos-acuerdo\\_short\\_es.pdf](http://www.science-international.org/sites/default/files/reports/datos-abiertos-acuerdo_short_es.pdf)
- International Open Data Charter (2015). Principles. URL: [https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter\\_F.pdf](https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf)
- International Science Council; IAP & TWAS (2016). Open data in a big data world. Recurso web. URL: <http://www.science-international.org/>
- ISO 19115:2003. Geographic information — Metadata. Modificada en 2014. URL: <https://www.iso.org/standard/26020.html>
- ISO 19115-1:2014. Geographic information — Metadata — Part 1: Fundamentals. URL: <https://www.iso.org/standard/53798.html>
- Janssen, M., Charalabidis, Y. y Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information systems management*, 29(4), pp. 258-268. DOI: <https://doi.org/10.1080/10580530.2012.716740>
- Kaufman, E. (2017). El papel de la Alianza para el Gobierno Abierto. Estado Abierto. *Revista sobre el Estado, la administración y las políticas públicas*, 1(2), pp. 71-116. URL: <http://181.209.22.113/index.php/EA/article/view/35/35>
- La Porte, T.M., Demchak, C.C. y De Jong, M. (2002). Democracy and bureaucracy in the age of the web. Empirical Findings and theoretical speculations. *Administration & Society*, 34(4), 411-446. DOI: <https://doi.org/10.1177/0095399702034004004>

- Labajo González, E. (2015). El método científico. El método pericial. Universidad Complutense de Madrid: Máster en Pericia Sanitaria. URL: <https://docplayer.es/21550185-El-metodo-cientifico-608104-el-metodo-pericial-master-en-pericia-sanitaria-2015-2016-dra-elena-labajo-gonzalez.html>
- Ley 56/2007, de 28 de diciembre, de Medidas de Impulso de la Sociedad de la Información. «BOE» núm. 312, de 29/12/2007. URL: <https://www.boe.es/buscar/act.php?id=BOE-A-2007-22440>
- Ley Orgánica 3/2007, del 22 de marzo, para la igualdad efectiva de mujeres y hombres. BOE nº 71, de 23 de marzo de 2007. Referencia: BOE-A-2007-6115. URL: <https://www.boe.es/buscar/pdf/2007/BOE-A-2007-6115-consolidado.pdf>
- Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. «BOE» núm. 276, de 17 de noviembre de 2007, páginas 47160 a 47165 (6 págs.). URL: <https://www.boe.es/eli/es/l/2007/11/16/37/dof/spa/pdf>
- Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. «BOE» núm. 295, de 10 de diciembre de 2013. URL: <https://www.boe.es/buscar/pdf/2013/BOE-A-2013-12887-consolidado.pdf>
- Ley 2/2015, de 2 de abril, de la Generalitat, de Transparencia, Buen Gobierno y Participación Ciudadana de la Comunitat Valenciana. DOGV núm. 7500 de 08.04.2015. URL: [http://www.dogv.gva.es/datos/2015/04/08/pdf/2015\\_3137.pdf](http://www.dogv.gva.es/datos/2015/04/08/pdf/2015_3137.pdf)
- Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. «BOE» núm. 164, de 10 de julio de 2015, páginas 57436 a 57450 (15 págs.). URL: <https://www.boe.es/eli/es/l/2015/07/09/18/dof/spa/pdf>
- Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas. «BOE» núm. 236, de 02/10/2015. URL: <https://www.boe.es/buscar/pdf/2015/BOE-A-2015-10565-consolidado.pdf>
- Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. «BOE» núm. 236, de 02/10/2015. URL: <https://www.boe.es/buscar/act.php?id=BOE-A-2015-10566>
- Ley 2/2016, de 4 de marzo, de la Generalitat, de modificación de la Ley 2/2015, de 2 de abril, de Transparencia, Buen Gobierno y Participación Ciudadana de la Comunitat Valenciana. DOGV núm. 7735 de 07.03.2016. URL: [http://www.dogv.gva.es/datos/2016/03/07/pdf/2016\\_1599.pdf](http://www.dogv.gva.es/datos/2016/03/07/pdf/2016_1599.pdf)

- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. «BOE» núm. 294, de 6 de diciembre de 2018, páginas 119788 a 119857 (70 págs.). URL: <https://www.boe.es/buscar/doc.php?id=BOE-A-2018-16673>
- Libro Verde sobre la Información del sector público en la Sociedad de la Información (1999). Consultado en Portal de Gobierno Abierto de Navarra: [https://gobiernoabierto.navarra.es/sites/default/files/opendata/libro\\_verde\\_informacion.pdf](https://gobiernoabierto.navarra.es/sites/default/files/opendata/libro_verde_informacion.pdf)
- Llinares, J. (2010). Las 10 claves para entender la colaboración en el modelo open government. En: Calderón, C. y Lorenzo, S. (editores). *Open Government: gobierno abierto*. Alcalá la Real: Algón. ISBN: 0978-84-937218-5-5.
- López, P. L. (2004). Población muestra y muestreo. *Punto cero*, 9(8). URL: [http://www.scielo.org.bo/scielo.php?pid=s1815-02762004000100012&script=sci\\_arttext](http://www.scielo.org.bo/scielo.php?pid=s1815-02762004000100012&script=sci_arttext)
- López-Roldán, P. y Fachelli, S. (2015). *Metodología de la Investigación Social Cuantitativa. La encuesta*. Bellaterra (Cerdanyola del Vallès): Dipòsit Digital de Documents, Universitat Autònoma de Barcelona. Capítulo II.3. Edición digital: <http://ddd.uab.cat/record/163567>
- Luque, C. (2019). Open Data y FAIR Data: diferencias y similitudes. OGoov. URL: <https://www.ogoov.com/es/blog/open-data-y-fair-data-diferencias-y-similitudes/>
- Marcos-Martín, C. y Soriano-Maldonado, S.L. (2011). Reutilización de la información del sector público y open data en el contexto español y europeo. Proyecto Aporta. *El Profesional de la Información*, 20(3), pp. 291-297. DOI: 10.3145/epi.2011.may.07
- Martín González, Y. (2013). Política europea de reutilización de la información del sector público. De la norma jurídica al portal de datos abiertos. *Revista Universitaria Europea*, 19, pp. 113-134. URL: <http://www.revistarue.eu/RUE/112013.pdf>
- Martínez Rodríguez, N. (2019). Protección de datos y derechos digitales. Salamanca. *Ars Iuris Salmanticensis*, (7) pp. 254-259.
- Mendoza Vega, J. B. (2016). R para principiantes. Universidad Nacional Autónoma de México: Libro en Bookdown.org. URL: <https://bookdown.org/jboscomendoza/r-principiantes4/>

- Mestre, T. (2008). Cuotas voluntarias y legales en España: La paridad a examen. *Revista Española de Investigaciones Sociológicas (REIS)*, 123, pp. 123-150. URL: [http://www.reis.cis.es/REIS/PDF/REIS\\_123\\_051215167211333.pdf](http://www.reis.cis.es/REIS/PDF/REIS_123_051215167211333.pdf)
- Mills, G. E., y Gay, L. R. (2019). Educational research: Competencies for analysis and applications. *Pearson*. One Lake Street, Upper Saddle River, New Jersey 07458. URL: <https://eric.ed.gov/?id=ED594591>
- Ministerio de Universidades (2021). Datos y cifras del sistema universitario español. Publicación 2020-2021. URL: [https://www.universidades.gob.es/stfls/universidades/Estadisticas/ficheros/Datos\\_y\\_Cifras\\_2020-21.pdf](https://www.universidades.gob.es/stfls/universidades/Estadisticas/ficheros/Datos_y_Cifras_2020-21.pdf)
- Mira, J. J., Pérez-Jover, V., Lorenzo, S., Aranaz, J., y Vitaller, J. (2004). La investigación cualitativa: una alternativa también válida. *Atención primaria*, 34(4), 161-166. DOI: [https://doi.org/10.1016/S0212-6567\(04\)78902-7](https://doi.org/10.1016/S0212-6567(04)78902-7)
- Moneo, Antonio (2014). ¿Qué es un portal de datos abiertos y para qué sirve? *Banco Interamericano de Desarrollo*. URL: <https://blogs.iadb.org/conocimiento-abierto/es/que-es-un-portal-de-datos-abiertos-y-para-que-sirve/>
- Morales García, A. M., Martínez Cardama, S., García López, F., y Caridad Sebastián, M. (2016). La transparencia en la política española: Análisis de su influencia y presencia en medios sociales. *Ibersid*, 10 (2), 91–96. DOI: <https://doi.org/10.54886/ibersid.v10i2.4350>
- Moreira, Izabela y Claussen, Mario Vinicius (2011). Políticas de transparencia en la administración pública brasileña. *Revista Reforma y Democracia* (51). URL: <https://www.redalyc.org/articulo.oa?id=357533683005>
- Müggenburg Rodríguez, M. C. y Pérez Cabrera, I. (2007). Tipos de estudio en el enfoque de investigación cuantitativa. *Enfermería Universitaria*, 4(1), 35-38. URL: <https://www.redalyc.org/pdf/3587/358741821004.pdf>
- Muñoz, Lourdes (2016). Origen y potencia del Open Data. *Iniciativa Open Data Barcelona*. URL: <https://iniciativabarcelonaopendata.cat/es/2016/05/origen-y-potencia-del-open-data/>
- Naessens, H. (2010). Ética pública y transparencia. *XIV Encuentro de Latinoamericanistas Españoles: Congreso Internacional*, Sep. 2010, Santiago de Compostela, España. pp. 2113-2130. Halshs-00531532.



- Naser, A. y Concha, A. (2012). Datos Abiertos: un nuevo desafío para los gobiernos de la región. *Instituto Latinoamericano y del Caribe de Planificación Económica y Social (ILPES)*. URL: [https://repositorio.cepal.org/bitstream/handle/11362/7331/1/S1200084\\_es.pdf](https://repositorio.cepal.org/bitstream/handle/11362/7331/1/S1200084_es.pdf)
- National Institutes of Health (2003). Final NIH statement on sharing research data. URL: <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html>
- Nemesio, Sylvana (2015). Metadatos de datos abiertos: estudio de caso en Uruguay. *BiD: textos universitarios de biblioteconomía i documentació*, 34. DOI: <https://dx.doi.org/10.1344/BiD2015.34.17>
- Nicandro Cruz-Rubio, C. (2015). ¿Qué es (y que no es) gobierno abierto? Una discusión conceptual. *Eunomía. Revista en Cultura de la Legalidad*. 8, pp. 37-53. URL: <https://e-revistas.uc3m.es/index.php/EUNOM/article/view/2475>
- Obama, B. (2009). Transparency and Open Government, memorandum for the heads of executive departments and agencies. URL: [https://web.archive.org/web/20161215181923/https://www.whitehouse.gov/the\\_press\\_office/TransparencyandOpenGovernment](https://web.archive.org/web/20161215181923/https://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment)
- Obama, B. (2009). Executive Office of the President, Office of Management and Budget, Memorandum for the Heads of Executive Departments and Agencies: Open Government Directive. Washington, DC. URL: [https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/memoranda\\_2010/m10-06.pdf](https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf)
- OCDE (2017). Recomendación OCDE sobre Gobierno Abierto (2017). URL: <https://www.oecd.org/gov/recomendacion-del-consejo-sobre-gobierno-abierto-141217.pdf>
- OCDE (2021). Recommendation of the Council concerning Access to research data from public funding. URL: <https://www.oecd.org/sti/recommendation-access-to-research-data-from-public-funding.htm>
- ONTSI (2015). *Estudio de caracterización del sector infomediario en España 2014. Parte I: Sector infomediario público*. Madrid: ONTSI. URL: [http://www.ontsi.red.es/ontsi/sites/default/files/estudio\\_infomediario\\_publico\\_2014.pdf](http://www.ontsi.red.es/ontsi/sites/default/files/estudio_infomediario_publico_2014.pdf)

- Open Data Barometer (2017). Informe global. Cuarta edición. World Wide Web Foundation. URL: <https://opendatabarometer.org/doc/4thEdition/ODB-4thEdition-GlobalReport-ES.pdf>
- Open Data Barometer (2018). El barómetro de los datos abiertos. Edición de los Líderes. De la promesa al progreso. URL: <https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report-ES.pdf>
- Open Data Charter. Definición Datos Abiertos. URL: <https://opendatacharter.net/principles-es/>
- Open Data Handbook. Definición Datos Abiertos. URL: <https://opendatahandbook.org/guide/es/what-is-open-data/>
- Open Data Handbook. ¿Cómo abrir los datos? URL: <https://opendatahandbook.org/guide/es/how-to-open-up-data/>
- Open Government Partnership. URL: <https://www.opengovpartnership.org/>
- Open Knowledge Foundation. Definición Datos Abiertos. URL: <https://okfn.org/opendata/how-to-open-data/>
- Open Knowledge Foundation (2017). El manual de Open Data: ¿Por qué datos abiertos? *Open Data Handbook*. URL: <https://opendatahandbook.org/guide/es/why-open-data/>
- Organización de los Estados Americanos (S.D.). Guía de mecanismos para la promoción de la transparencia y la integridad en las Américas. URL: [https://www.oas.org/es/sap/dgpe/guia\\_egov.asp](https://www.oas.org/es/sap/dgpe/guia_egov.asp)
- Organización para la Cooperación y el Desarrollo Económico (2014). URL: <https://www.oecd.org/gov/digital-government/2014-open-government-data-survey.pdf>
- Organización para la Cooperación y el Desarrollo Económico (2017). URL: <https://www.oecd.org/gov/2016-OECD-Survey-on-Open-Government-Data-3.0.pdf>
- Organización para la Cooperación y el Desarrollo Económicos (2017). Recomendación del Consejo sobre Gobierno Abierto. URL: <https://www.oecd.org/gov/recomendacion-del-consejo-sobre-gobierno-abierto-141217.pdf>

- ONTSI (2014). *Estudio de caracterización del sector infomediario en España 2014. Parte I: Sector infomediario público*. Madrid: ONTSI. URL: <https://ictlogy.net/bibliography/reports/projects.php?idp=2803&lang=ca>
- ONTSI (2015). *Estudio de caracterización del sector infomediario en España 2014. Parte II: Reutilización de información del sector privado*. Madrid: ONTSI. URL: <https://ictlogy.net/bibliography/reports/projects.php?idp=2804>
- Oszlak, O. (2013). Gobierno abierto: hacia un nuevo paradigma de gestión pública. *Red de Gobierno Electrónico de América Latina y el Caribe - Red GEALC*. Colección de documentos de trabajo sobre e-Gobierno (5). URL: <https://www.oas.org/es/sap/dgpe/pub/coleccion5RG.pdf>
- Osorio-Sanabria, M.A., Amaya-Fernández, F.O. y González-Zabala, M. (2019). Análisis de datos abiertos de instituciones de educación Entorno superior colombianas como apoyo a la relación Universidad-Entorno. *Entramado*. 16 (1), 272–284. DOI: <https://doi.org/10.18041/1900-3803/entramado.1.6127>
- Oviedo, E., Mazón, J. N., y Zubcoff, J. J. (n.d.). *Modelo de calidad y madurez para portales de datos abiertos*. [http://www.agendadigital.gob.es/planes-actuaciones/Bibliotecaciudadesinteligentes/1.Plan/Plan\\_Nacional\\_de\\_Ciudades\\_Inteligentes.pdf](http://www.agendadigital.gob.es/planes-actuaciones/Bibliotecaciudadesinteligentes/1.Plan/Plan_Nacional_de_Ciudades_Inteligentes.pdf)
- Parks, Wallace. (1957). “The open government principle: applying the right to know under the Constitution”. *The George Washington Law Review*, 6(1), 1-2. Washington: Washington University Law School.
- Plan de Acción Europeo sobre Administración Electrónica (2011-2015): *Aprovechamiento de las TIC para promover una administración pública inteligente, sostenible e innovadora*. URL: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0743:FIN:ES:PDF>
- Peset, F., Ferrer-Sapena, A. y Subirats-Coll, I. (2011). Open data y Linked open data: su impacto en el área de bibliotecas y documentación. *El Profesional de la Información*, 20(2), 165–174. DOI: <https://doi.org/10.3145/epi.2011.mar.06>
- Peset, F., Aleixandre-Benavent, R., Blasco-Gil, Y. y Ferrer-Sapena, A. (2017). Datos abiertos de investigación. Camino recorrido y cuestiones pendientes. *Anales de documentación*. 20(1). DOI: <https://doi.org/10.6018/analesdoc.20.1.272101>

- Pineda, B., De Alvarado, E. L. y De Canales, F. (1994). Metodología de la investigación: manual para el desarrollo de personal de salud. Segunda edición. *Organización Panamericana de la Salud*. Washington. URL: <https://iris.paho.org/handle/10665.2/3132>
- Pintos, C. y Marín, J. L. (2013). Open data: reutilización de la información pública. Netbiblio (Madrid): *Instituto Nacional de Administración Pública. Innap Innova*. URL: <https://bci.inap.es/material-open-data-reutilización-de-la-información-pública-carlos-pinto-libro-electrónico>
- Pipino, L. L., Lee, Y. W. y Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211-218. URL: <https://doi.org/10.1145/505248.506010>
- Pita Fernández, S., y Pértegas Díaz, S. (2002). Investigación cuantitativa y cualitativa. *Cad Aten Primaria*, 9, 76-78. URL: [https://www.fisterra.com/mbe/investiga/cuanti\\_cuali/cuanti\\_cuali2.pdf](https://www.fisterra.com/mbe/investiga/cuanti_cuali/cuanti_cuali2.pdf)
- Portal de Administración Electrónica (2015). Reutilización de la información del Sector Público. *Gobierno de España*. URL: [https://administracionelectronica.gob.es/pae/Home/pae\\_Estrategias/pae\\_Gobierno\\_Abierto\\_Inicio/pae\\_Reutilizacion\\_de\\_la\\_informacion\\_en\\_el\\_sector\\_publico.html#.YbHYgi8rxpQ](https://administracionelectronica.gob.es/pae/Home/pae_Estrategias/pae_Gobierno_Abierto_Inicio/pae_Reutilizacion_de_la_informacion_en_el_sector_publico.html#.YbHYgi8rxpQ)
- Portal de Datos Abiertos de la Junta de Castilla y León. Normativa. URL: <https://datosabiertos.jcyl.es/web/es/iniciativa-datos-abiertos/normativa.html>
- Portal de Datos Abiertos de la Universidad de Cantabria. Marco legal. URL: <https://web.unican.es/opendata/Paginas/Marco-Legal.aspx>
- Raffino, M. E. (2020). Métodos de investigación. Argentina. *Concepto.de*. URL: <https://concepto.de/metodos-de-investigacion/>
- Ramírez-Alujas, A.V. (2012). Gobierno Abierto y Modernización de la Gestión Pública: Tendencias actuales y el (Inevitable) Camino que Viene - Reflexiones Seminales. *Open Government and Modernization of Public Management: Current Trends and the (Inevitable) Way Forward - Seminal Reflections*. (January 3, 2012). *Revista Enfoques: Ciencia Política y Administración Pública*, 9(15), pp. 99-125. URL: <https://ssrn.com/abstract=1979133>

- Ramírez-Alujas, A.V. (2012). Gobierno abierto (Open Government). Espacios del conocimiento: Sujeto, verdad, heterotopías. A 30 años de la muerte de Michael Foucault. *Actas VIII Escuela Chile-Francia. Cátedra Michel Foucault*. Pp. 153-174. URL: <https://uchile.cl/u131564>
- Ramírez-Alujas, A.V. (2012). Gobierno abierto. Economía. *Revista en Cultura de la Legalidad*. (5), pp. 201-216. URL: <https://e-revistas.uc3m.es/index.php/EUNOM/article/view/2180/1116>
- Ramos Simón, L. F. (2003). La reutilización de la información del sector público. Aproximación al contenido de la propuesta de directive 2002. *Revista general de información y documentación*, 13 (2), 59-96. URL: <https://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=16171792>
- Rangel-Carrillo, A. M., Maestre-Góngora, G. P., y Osorio-Sanabria, M. A. (2020). Principios, lineamientos, dimensiones y atributos para la evaluación de calidad de Datos Abiertos de Gobierno. *Aibi Revista De investigación, administración E ingeniería*, 8(S1), 54-65. DOI: <https://doi.org/10.15649/2346030X.950>
- Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica. «BOE» núm. 25, de 29 de enero de 2010, páginas 8089 a 8138 (50 págs.). URL: <https://www.boe.es/buscar/doc.php?id=BOE-A-2010-1330>
- Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica. «BOE» núm. 25, de 29 de enero de 2010, páginas 8139 a 8156 (18 págs.). URL: <https://www.boe.es/boe/dias/2010/01/29/pdfs/BOE-A-2010-1331.pdf>
- Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal. «BOE» núm. 269, de 8 de noviembre de 2011, páginas 116296 a 116307 (12 págs.). URL: [https://www.boe.es/diario\\_boe/txt.php?id=BOE-A-2011-17560](https://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-17560)
- Real Decreto 919/2014, de 31 de octubre, por el que se aprueba el Estatuto del Consejo de Transparencia y Buen Gobierno. «BOE» núm. 268, de 5 de noviembre de 2014, páginas 90451 a 90462 (12 págs.). URL: <https://www.boe.es/boe/dias/2014/11/05/pdfs/BOE-A-2014-11410.pdf>

- Reglamento de Gobierno Abierto: Transparencia (2019). BOP de fecha 11 de abril de 2019. URL: [https://www.valencia.es/ayuntamiento/ayuntamiento.nsf/0/FD4E2C53553EE404C125858F002F8736/\\$FILE/13.%20Aprovacion%20inicial%20Pleno.pdf?OpenElement&lang=1](https://www.valencia.es/ayuntamiento/ayuntamiento.nsf/0/FD4E2C53553EE404C125858F002F8736/$FILE/13.%20Aprovacion%20inicial%20Pleno.pdf?OpenElement&lang=1)
- Reglamento de Transparencia y Participación Ciudadana (2012). Aprobado por acuerdo plenario de: 28-09-2012. Modificación aprobada por acuerdo plenario de 24-04-2015. Derogación del Título II por acuerdo plenario de 25-06-2020. URL: [http://gobiernoabierto.valencia.es/wp-content/uploads/2014/10/R\\_Participacion-Ciudadana\\_2016\\_cas.pdf](http://gobiernoabierto.valencia.es/wp-content/uploads/2014/10/R_Participacion-Ciudadana_2016_cas.pdf)
- Ren, G.J. y Glissmann, S. (2012). Identifying information assets for open data. The role of business architecture and information quality. En: *Procs. of the IEEE 14th Intl conf on commerce and enterprise computing*, Sept., pp. 94-100. DOI: <http://doi.ieeecomputersociety.org/10.1109/CEC.2012.23>
- Rivas-Rebaque, B.; Gétrudix-Barrio, F. y de Cisneros de Britto, J.C. (2019). La percepción del docente universitario ante el uso y valor de los datos abiertos. *Educación XXI*, 22(2), pp. 141-163, DOI: <https://doi.org/10.5944/educxx1.21317>
- Rodrigo de Castro, L. (2019). Los principios consolidados por la Carta Internacional de datos abiertos a partir de la idea de gobierno abierto. Análisis comparado de su implementación normativa en Estados Unidos y la Unión Europea. *Universidad CEU-San Pablo*: Tesis Doctoral. URL: <https://dialnet.unirioja.es/servlet/tesis?codigo=292634>
- Rodríguez, J.C. (2018). Demografía universitaria española: aproximación a su dimensión, estructura y evolución. *Fundación Europea Sociedad y Educación / Santander Universidades* Editors: Mercedes de Esteban, José Manuel Moreno. Pp. 42-52. URL: [https://www.researchgate.net/publication/323837984\\_La\\_edad\\_media\\_del\\_profesorado\\_de\\_la\\_universidad\\_publica\\_en\\_Espana\\_una\\_perspectiva\\_de\\_medio\\_plazo](https://www.researchgate.net/publication/323837984_La_edad_media_del_profesorado_de_la_universidad_publica_en_Espana_una_perspectiva_de_medio_plazo)
- Rodríguez Peñuelas, M. A. (2010). Métodos de investigación. México. *Ed. Universidad Autónoma de Sinaloa*. Métodos Cuantitativos: pp. 32.
- Romero Gómez, F. (2015). Reflexiones sobre la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas: hacia una nueva posición de los ciudadanos en el procedimiento. *Revista Aranzadi Doctrinal* (11). URL: [https://urbamalaga.com/wp-content/uploads/2016/03/Bib\\_Reflexiones-sobre-la-Ley-39-2015-de-1-de-octubre-del-Procedimiento-Administrativo-\\_BIB\\_2015\\_17292-1.pdf](https://urbamalaga.com/wp-content/uploads/2016/03/Bib_Reflexiones-sobre-la-Ley-39-2015-de-1-de-octubre-del-Procedimiento-Administrativo-_BIB_2015_17292-1.pdf)

- Rubiralta, Mario y Bellavista, Joan (2003). Nuevos mecanismos de transferencia de Tecnología. *Libro N° 9 de la Colección de Encuentros Empresariales CONEC. Fundación Cotec*, 2003. URL: <https://www.madrimasd.org/cultura-cientifica/ciencia-cultura/ensayo/nuevos-mecanismos-transferencia-tecnologia-debilidades>
- Salas Ocampos, Danelly (2020). Selección de la muestra en la investigación cuantitativa. *Investigaliacr*. URL: <https://investigaliacr.com/investigacion/seleccion-de-la-muestra-en-la-investigacion-cuantitativa/>
- Sánchez, C.M. (2016). Stephen Hepell: “En el 2030 nos reiremos de muchas cosas de los colegios de hoy”. *XLsemanal*. Madrid: Fundación Telefónica. URL: <https://www.xlsemanal.com/actualidad/20161017/stephen-hepell-educacion-futuro.html>
- Sarduy Domínguez, Y. (2007). El análisis de información y las investigaciones cuantitativa y cualitativa. *Revista Cubana de Salud Pública*, Vol. 33(3). URL: <https://www.redalyc.org/pdf/214/21433320.pdf>
- Schmidt B., Gemeinholzer B. y Treloar A (2016). Open Data in Global Environmental Research: The Belmont Forum’s Open Data Survey. *PLoS ONE*. 11(1): e0146695. DOI: <https://doi.org/10.1371/journal.pone.0146695>
- Serrano-Cinca, C., Rueda-Tomás, M. y Portillo-Tarragona, P. (2008). Factors influencing e-disclosure in local public administrations. *Environment and planning. C: Government & policy*, 27(2), 335-378. DOI: <https://doi.org/10.1068/c07116r>
- Serrano, Jesús (2019). “Cómo los datos abiertos están cambiando el mundo”. *OGooV*. URL: <https://www.ogooov.com/es/blog/datos-abiertos-cambiando-mundo/>
- Tascón, Mario (2013). “Introducción: Big Data. Pasado, presente y futuro”. *Telos: Cuadernos de comunicación e innovación*, ISSN 0213-084X, N°. 95, 2013, págs. 47-50. URL: <https://dialnet.unirioja.es/servlet/articulo?codigo=4423775>
- Tamayo, M. (2007). *Metodología de la Investigación*. México: Limusa.
- Tavarez, C. (2020). Cuantitativo vs cualitativo en una tesis de doctorado. *PhD Institute Accredited Degree Experts*. URL: <http://phdinstitute.org/cuantitativo-vs-cualitativo/>
- Thiel, S.V. y Leeuw, F.L. (2002). The Performance Paradox in the Public Sector. *Public Performance & Management Review*, 25(3), 267-281. URL: <https://www.jstor.org/stable/3381236>

- Tinati, R., Carr, L., Halford, S. y Pope, C. (2012). Exploring the impact of adopting open data in the UK government. *Digital Futures*. URL: <https://eprints.soton.ac.uk/id/eprint/344808>
- Torres Menárguez, A. (3 de abril de 2017). “En las aulas de Harvard ninguna pregunta se considera estúpida”. *El País*. URL: [https://elpais.com/economia/2017/03/27/actualidad/1490624613\\_488525.html](https://elpais.com/economia/2017/03/27/actualidad/1490624613_488525.html)
- Torres Menárguez, A. (23 de enero de 2017). “El profesor del siglo XXI tiene que enseñar lo que no sabe”. *El País*. URL: [https://elpais.com/economia/2017/01/15/actualidad/1484514194\\_176496.html](https://elpais.com/economia/2017/01/15/actualidad/1484514194_176496.html)
- Torres-Salinas, D., Robinson-García, N., y Cabezas-Clavijo, Á. (2012). Compartir los datos de investigación en ciencia: introducción al data sharing. *El Profesional de la Información*, 21(2), 173–184. DOI: <https://doi.org/10.3145/epi.2012.mar.08>
- Trelles, A., Altman, M., Magar, E. y McDonald, M.P. (2016). Datos abiertos, transparencia y redistribución en México. *Política y gobierno*, 23(2), 331-364. URL: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S1665-20372016000200331&lng=es&tlng=es](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-20372016000200331&lng=es&tlng=es)
- Trigueros, J. S. (2015). “Los antecedentes del gobierno abierto: una mirada retrospectiva en la evolución de la administración pública”. *Revista Enfoques: Ciencia Política y Administración Pública*, 13(23), 67-84. URL: <https://www.redalyc.org/pdf/960/96043202005.pdf>
- UNE 148004:2018. Información Geográfica: Datos geográficos abiertos. URL: <https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0059630>
- UNE 178301:2015. (Versión corregida en fecha 2015-07-29). Ciudades Inteligentes: Datos Abiertos (Open Data). URL: <https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0054318>
- UNESCO (2008). “Estándares de competencias TIC para docentes”. París. *Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura*. URL: <http://eduteka.icesi.edu.co/articulos/EstandaresDocentesUnesco>



- Universidad de Extremadura (2021). “Datos de investigación: gestión, datos abiertos (Open Data): Guía diseñada para entender qué son los datos de investigación y herramientas para gestionarlos”. *Servicio de Bibliotecas. Universidad de Extremadura*. URL: [https://biblioguias.unex.es/datos\\_de\\_investigacion](https://biblioguias.unex.es/datos_de_investigacion)
- Uvalle Berrones, R. (2015). Fundamentos de la transparencia en la sociedad contemporánea. *Revista Mexicana de Ciencias Políticas y Sociales*. Universidad Nacional Autónoma de México. Nueva Época, Año LXI, (226). pp. 199-220. DOI: [https://doi.org/10.1016/S0185-1918\(16\)30008-3](https://doi.org/10.1016/S0185-1918(16)30008-3)
- Valero-Torrijos, J. (2014). Acceso, reutilización y gestión avanzada de la información en el ámbito de la administración sanitaria. *Régimen jurídico de la transparencia del sector público: del Derecho de acceso a la reutilización de la información* / coord. por Julián Valero Torrijos, Manuel Fernández Salmerón, 2014, ISBN 978-84-9059-433-9, págs. 631-667.
- Velasco Rico, C. (2014). La cuestión competencial en la Ley 19/2013, de 9 de diciembre: ¿de qué margen disponen las CC. AA. en materia de transparencia, acceso a la información pública y buen gobierno? Análisis de la cuestión tras la Ley 19/2013. *Revista Jurídica de Castilla y León*, 33. URL: <https://www.jcyl.es/web/jcyl/AdministracionPublica/es/Plantilla100Detalle/1215245063566/1215245063566/1284319395584/Redaccion>
- Verhulst, S. and Young, A. (2016). Open data impact. When Demand and Supply Meet. Key Findings of the Open Data Impact Case Studies. URL: <https://www.thegovlab.org/static/files/publications/open-dataimpact-key-findings.pdf>
- Villegas Corona, Víctor Alejandro (2018). Transparencia, cultura cívica y participación ciudadana en democracias, análisis normativo y programático del caso mexicano. *Dilemata*, 27, pp. 183-202. URL: <https://dialnet.unirioja.es/descarga/articulo/6543300.pdf>
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>
- World Bank Open Data (S.F.). Definición Datos Abiertos. URL: <https://www.bancomundial.org/es/news/feature/2018/10/01/open-data-futuro>

Zuiderwijk, A. & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), 17-29. DOI: <https://doi.org/10.1016/j.giq.2013.04.003>

# Anexos

## Anexos

### Anexo I. Muestra de población: universidades y centros de investigación

---

Universitat Politècnica de València  
Universitat de València  
Universidad de Alicante  
Universitat Jaume I de Castelló  
Universitat Miguel Hernández (Elche)  
Universidad Autónoma de Madrid  
Universidad Carlos III de Madrid  
Universidad Católica de Ávila  
Universidad Católica de Valencia San Vicente Mártir  
Universidad Católica San Antonio  
Universidad Central del Ecuador  
Universidad Complutense de Madrid  
Universidad da Coruña  
Universidad de Almería  
Universidad de Burgos  
Universidad de Cádiz  
Universidad de Castilla-La Mancha  
Universidad de Extremadura  
Universidad de Granada  
Universidad de Huelva  
Universidad de Jaén  
Universidad de La Laguna  
Universidad de Las Palmas de Gran Canaria  
Universidad de Málaga  
Universidad de Murcia  
Universidad de Oviedo  
Universidad de Salamanca  
Universidad de Sevilla  
Universidad de Valladolid  
Universidad de Zaragoza  
Universidad del País Vasco  
Universidad Nacional Autónoma de Honduras  
Universidad Pablo de Olavide - Sevilla  
Universidad Politécnica de Cartagena  
Universidad Politécnica de Madrid  
Universidad Pontificia de Salamanca  
Universidad Pública de Navarra  
Universidad Rey Juan Carlos  
Universitat Autònoma de Barcelona

Universitat de Barcelona  
 Universitat de Girona  
 Universitat de les Illes Balears  
 Universitat de Lleida  
 Universitat Oberta de Catalunya  
 Universitat Politècnica de Catalunya  
 Universitat Pompeu Fabra

---

Tabla 63: Universidades encuestadas.

---

Instituto de Investigación Sanitaria La Fe  
 Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM)  
 Instituto Valenciano de Investigaciones Agrarias (IVIA)  
 Instituto de Investigación Sanitaria y Biomédica de Alicante (ISABIAL)  
 Instituto de Agroquímica y Tecnología de Alimentos (IATA - CSIC)  
 Centro de Investigación Príncipe Felipe (CIPF)  
 Instituto de Física Corpuscular (IFIC - CSIC)  
 Consejo Superior de Investigaciones Científicas (CSIC)  
 Hospital General Universitario de Alicante  
 Fundació de la Comunitat Valenciana (FCV)  
 Instituto Tecnológico de la Alimentación (AINIA)  
 Instituto Tecnológico de la Energía (ITE)  
 Agencia de Medios  
 CERN - Organización Europea para la Investigación Nuclear  
 CIC bioGUNE - Centro de Investigación Cooperativa en Biociencias  
 Fundació Docència i Recerca Mútua Terrassa  
 Hospital del Mar - Valencia  
 IES Haygón  
 Instituto de Investigación en Inteligencia Artificial - (IIIA-CSIC)  
 Instituto LRKlein  
 Instituto Universitario Menéndez Pidal (UCM)  
 Instituto de Investigación Sanitaria La Fe

---

Tabla 64: Centros de investigación encuestados.

## Anexo II. La encuesta.

### Parte I. Texto introductorio, consentimiento y contacto

#### **Encuesta.**

En primer lugar, gracias por participar en esta encuesta sobre datos abiertos y reutilización de información. Se realiza en el marco de colaboración entre la Universitat Politècnica de València y la Conselleria de Participación, Transparencia, Cooperación y Calidad Democrática. El objetivo es estudiar la percepción y conocimiento de los investigadores sobre los portales de datos abiertos, para potenciar la apertura y reutilización de los datos con acciones específicas.

Los resultados de esta encuesta permitirán investigar y establecer acciones concretas para potenciar la apertura y reutilización de los datos.

Tiene un total de 24 preguntas y se rellena en 10 minutos, en función de si quiere aportar información adicional.

Se garantiza el anonimato y confidencialidad de la encuesta.

#### **Consentimiento.**

Se le invita a participar en un estudio de investigación para conocer su opinión y experiencias como investigador respecto del uso y reutilización de los datos de los portales de transparencia y datos abiertos de las administraciones públicas en sus labores de investigador. Esta encuesta forma parte de los estudios que realiza la Cátedra de Transparencia y Gestión de Datos de la Universitat Politècnica de València. Los resultados de la encuesta se publicarán una vez finalizado el estudio y estarán accesibles en abierto para cualquier persona que quiera consultarlos y formarán parte de una publicación.

Confidencialidad y participación voluntaria.

Las respuestas (individuales) de esta encuesta serán utilizadas con fines de investigación y se almacenarán de forma segura. Se protegerá la identidad de las personas que participen en la encuesta y ninguna de las respuestas se atribuirá a nadie. La participación en el estudio es voluntaria. Si en algún momento desea finalizar su participación y/o eliminar su información, contacte con el investigador de este estudio.

#### **Contacto.**

Para cualquier pregunta, contacte con Christian Vidal Cabo (chvica@inf.upv.es), Técnico de investigación de la UPV o al correo de la Cátedra de Transparencia y Gestión de Datos (ctranspa@upv.es).

---

## Parte II. La encuesta

Bloque de información personal y laboral.

### **Datos generales.**

#### **¿En qué provincia reside?**

- Respuesta abierta (provincias España)

#### **En qué rango de edad se encuentra.**

- 22-30
- 31-40
- 41-50
- 51-60
- 61-65
- + de 65

#### **Sexo.**

- Femenino
- Masculino
- Intersexual
- Otro

#### **¿A qué universidad, centro, instituto o entidad de investigación pertenece?**

- Respuesta abierta. Universidades españolas, centros e institutos de investigación.

#### **Categoría laboral.**

- Profesor/a Asociado/a
- Profesor visitante
- Profesor Ayudante
- Profesor Ayudante Doctor
- Profesor Contratado Doctor
- Profesor Titular de Universidad
- Catedrático de Universidad
- Científico Titular
- Investigador Científico
- Profesor de Investigación
- Doctor
- Doctorando
- Técnico de investigación
- Otro (especificar)

**¿A qué área de investigación pertenece y centra su producción científica? Puede elegir más de una opción<sup>96</sup>.**

- Área 1: Ciencias Sociales
- Área 2: Derecho
- Área 3: Economía
- Área 4: MLP/ Mente, Lenguaje y Pensamiento
- Área 5: FLA/ Cultura, Filosofía, Literatura y Arte
- Área 6: PHA/ Estudios del pasado: Historia y Arqueología
- Área 7: EDU/ Ciencias de la educación
- Área 8: PSI/ Psicología
- Área 9: MTM /Matemáticas
- Área 10: FIS/ Ciencias físicas
- Área 11: PIN/ Producción industrial, ingeniería civil e ingenierías para la sociedad
- Área 12: TIC/ Tecnologías de la Información y las Comunicaciones
- Área 13: EYT/ Energía y Transporte
- Área 14: CTQ/ Ciencias y Tecnologías Químicas
- Área 15: MAT/ Ciencias y Tecnologías Materiales
- Área 16: CTM/ Ciencias y Tecnologías Medioambientales
- Área 17: CAA/ Ciencias Agraria y Agroalimentarias
- Área 18: BIO/ Biociencias y biotecnología
- Área 19: BME/ Biomedicina
- Otro (especifique)

---

<sup>96</sup> Áreas extraídas de la Agencia Estatal de Investigación del Ministerio de Ciencia e Innovación: <https://www.aei.gob.es/areas-tematicas/areas-tematicas>



Bloque 1, de contexto inicial: Conocimiento acerca de los datos abiertos.

**1. ¿Sabe qué son los datos abiertos?** (Entendidos como filosofía que persigue que determinados tipos de datos estén disponibles de forma libre sin restricciones, en bruto, estructurados y en formatos que faciliten la reutilización).

- Sí
- No

**2. ¿Ha reutilizado datos/información del sector público para el desarrollo de investigaciones, proyectos o servicios en su actividad profesional?**

- Sí
- No

**3. En caso de que la respuesta a la pregunta 1 haya sido negativa, ¿cuál es el motivo?**

Respuesta abierta. Ejemplo: Desconozco la existencia de este tipo de portales y de cómo reutilizar la información.

**4. En caso de que la respuesta a la pregunta 2 haya sido negativa, ¿cuál es el motivo?**

\*Múltiple respuesta

- 4.1 Conozco el movimiento de los datos abiertos, pero no sé cómo reutilizar los datos.
- 4.2 Falta de fiabilidad de los datos.
- 4.3 Falta de actualización de los datos.
- 4.4 Falta de información interesante y reutilizable para mis investigaciones/proyectos.
- 4.5 Falta de estandarización de los datos (mismos tipos/conjuntos de datos en diferentes portales que vienen de maneras distintas. Falta de interoperabilidad entre administraciones).
- 4.6 Otros motivos: especifique cuáles...

Codificación pregunta 4.

4.1 Desconocimiento reutilización

4.2 Fiabilidad

4.3 Actualización

4.4 Falta de dataos

4.5 Estandarización

4.6 Otros

**5. Si conoce qué son los datos abiertos, dentro de su actividad profesional y de investigación, ¿de qué manera se encuentran involucrado con los datos abiertos y de investigación? \*\*\*Múltiple respuesta.**

- 5.1 Usuario de datos (para informarme).
- 5.2 Usuario de datos para producir servicios, productos o análisis propios (Reutilizador).
- 5.3 Proveedor de datos (en portales de datos abiertos o repositorios de investigación).
- 5.4 Administrador de datos.
- 5.5 Ninguna de las anteriores.
- 5.6 Otros: Especifique cual.

Codificación pregunta 5.

- 5.1 Usuario
- 5.2 Reutilizador
- 5.3 Proveedor
- 5.4 Administrador
- 5.5 Ninguna
- 5.6 Otros

**6. En caso de que la pregunta 1 haya sido afirmativa, ¿cómo conoció el movimiento de los datos abiertos?**

- 6.1 A través de artículos de investigación.
- 6.2 Medios de comunicación convencionales.
- 6.3 En talleres de formación/seminarios/congresos.
- 6.4 Canales de comunicación y/o redes sociales de las administraciones públicas y otros proveedores de datos.
- 6.5 Redes sociales.
- 6.6 A través de reutilizadores.
- 6.7 Otros: Especifique cual.

Codificación pregunta 6. Conocimiento de datos abiertos

- 6.1 Investigación
- 6.2 Medios comunicación
- 6.3 Formación
- 6.4 Proveedores datos
- 6.5 Redes sociales
- 6.6 Reutilizadores
- 6.7 Otros

\*Si en la pregunta 1 ha respondido negativo puede finalizar su encuesta aquí. Muchas gracias por su participación. \*

Bloque 2. Importancia de la apertura de datos.

**7. ¿Qué atributos considera importantes en relación a la apertura de datos por parte de la administración?**

\*Escala de clasificación: muy importante (4), importante (3), poco importante (2), Nada importante (1), no conozco (0). En caso de que no conozca el atributo puede dejarlo en blanco.

	0	1	2	3	4
7.1 Suministro de datos sin restricciones					
7.2 Tipo de licencia de los datos para su reutilización					
7.3 Que contenga metadatos (completos, estructurados y útiles para interpretarlos)					
7.4 Que sea fácil encontrarlos y conseguirlos (accesibilidad)					
7.5 Que estén estandarizados (formato, variables...)					
7.6 Posibilidad de restricción a algunos tipos de datos (datos sensibles)					
7.7 Datos a tiempo real					
7.8 Datos georreferenciados					

Tabla 65: Pregunta 7 de la encuesta.

Codificación pregunta 7. Atributos.

- 7.1 Suministro
- 7.2 Licencias datos
- 7.3 Metadatos
- 7.4 Accesibilidad
- 7.5 Formatos
- 7.6 Restricciones
- 7.7 Actualización
- 7.8 Geolocalización

**8. Dentro de los portales de datos abiertos, ¿qué servicios o funcionalidades considera más importantes al acceder a los datos?**

\*Escala de clasificación: muy importante (4), importante (3), poco importante (2), Nada importante (1), no conozco (0). En caso de que no conozca la funcionalidad puede dejar el valor sin contestar.

	0	1	2	3	4
8.1 Los datos tienen URI permanentes					
8.2 Enlaces a aplicaciones relacionadas con el conjunto de datos					
8.3 Herramientas para el procesamiento y análisis de datos					
8.4 Visualizaciones de los datos					
8.5 Canales de contacto o formularios para reutilizadores					
8.6 Noticias o novedades relacionadas con los datos abiertos y sistemas de alerta de actualización del catálogo de datos					
8.7 Disponibilidad de API					

Tabla 66: Pregunta 8 de la encuesta.

Codificación pregunta 8. Servicios y funcionalidades.

- 8.1 URI
- 8.2 Apps relacionadas
- 8.3 Herramientas procesamiento
- 8.4 Visualizaciones
- 8.5 Contacto
- 8.6 Noticias
- 8.7 API
- 8.8 Otro (Abierta)

**9. Evalúe las barreras que encuentra para acceder y reutilizar datos abiertos.**

\*Escala de clasificación: siempre (4), muchas veces (3), algunas veces (2), nunca(1), no conozco (0).

Si no ha utilizado nunca datos abiertos de la administración pase a la siguiente pregunta.

	0	1	2	3	4
9.1 Mucho tiempo transcurrido entre la solicitud y la recepción de datos					
9.2 La necesidad de registro para acceder a los datos					
9.3 Acceder y comprender las condiciones de uso y de licencias y como interpretar y reutilizar los datos					
9.4 Grados variables de calidad de los datos en conjuntos de datos diferentes, falta de estandarización en los formatos de los datos, información completa o de poco interés					
9.5 Pago por los datos					
9.6 Falta de interoperabilidad o de disponibilidad de los datos en otras AAPP / Dificultad para compararlos con otros datos					
9.7 Falta de actualización de los datos					
9.8 Dificultad de acceso					

Tabla 67: Pregunta 9 de la encuesta.

Codificación pregunta 9. Barreras.

9.1 Tiempo

9.2 Registro

9.3 Comprensión

9.4 Calidad y estandarización

9.5 Pago

9.6 Interoperabilidad

9.7 Falta de actualización

9.8 Acceso

**10. ¿Qué tipo de datos considera de más interés para su trabajo?**

\*Puede seleccionar más de un conjunto de datos.

<input type="radio"/> Sector público	<input type="radio"/> Transporte
<input type="radio"/> Medio ambiente y clima	<input type="radio"/> Turismo
<input type="radio"/> Educación	<input type="radio"/> Hacienda
<input type="radio"/> Empleo	<input type="radio"/> Ciencia y tecnología
<input type="radio"/> Legislación y justicia	<input type="radio"/> Comercio
<input type="radio"/> Demografía y Población	<input type="radio"/> Energía
<input type="radio"/> Salud	<input type="radio"/> Seguridad
<input type="radio"/> Economía	<input type="radio"/> Asociaciones
<input type="radio"/> Urbanismo e infraestructuras	<input type="radio"/> Autónomos
<input type="radio"/> Administración pública	<input type="radio"/> Catastro
<input type="radio"/> Sociedad y bienestar	<input type="radio"/> Consumo energético
<input type="radio"/> Cultura y ocio	<input type="radio"/> Cooperativas
<input type="radio"/> Medio rural	<input type="radio"/> Importaciones/Exportaciones
<input type="radio"/> Vivienda	<input type="radio"/> Información Mercantil
<input type="radio"/> Industria	<input type="radio"/> Seguridad social
<input type="radio"/> Deporte	<input type="radio"/> Otros: Especifique...

Pregunta 10.

Histograma 1 – 10 categorías más votadas

Histograma 2 – Categoría otros

Bloque 3. Utilidad real y motivaciones para la reutilización de datos abiertos en la investigación.

**11. Según su experiencia, valore la utilidad que tiene la apertura de datos en las siguientes afirmaciones.**

\*Escala de clasificación: muy importante (4), importante (3), poco importante (2), nada importante (1), No es útil porque no existen conjuntos de datos (0).

	0	1	2	3	4
11.1 Fomenta la investigación y/o actividades educativas					
11.2 Apoya las aplicaciones de la ciencia a los problemas de la sociedad					
11.3 Promociona la creación de la capacidad de transferencia de la tecnología					
11.4 Promueve la creación de nuevos productos y servicios.					
11.5 Permite la rendición de cuentas y la participación informada de la ciudadanía y la sociedad civil.					
11.6 Otros: Detallar					

Tabla 68: Pregunta 11 de la encuesta.

Codificación pregunta 11. Utilidad de los datos abiertos.

- 11.1 Fomento investigación
- 11.2 Apoyo
- 11.3 Promoción transferencia
- 11.4 Desarrollo
- 11.5 Transparencia
- 11.6 Otros



**12. Valore la importancia de las políticas de datos abiertos en la sociedad.**

Valore el nivel de utilidad real de estas políticas de apertura de datos.

\*Escala de clasificación: muy importante (3), importante (2), poco importante (1), nada importante (0).

	0	1	2	3
12.1 Participación ciudadana: involucrar a la ciudadanía y la comunidad científica				
12.2 Creación y desarrollo de nuevos productos y servicios				
12.3 Reforzar la confianza en las instituciones a través de la transparencia.				
12.4 Acelerar la transformación digital y el uso del <i>big data</i> en las administraciones públicas				
12.5 Favorecer la digitalización y el crecimiento económico.				

Tabla 69: Pregunta 12 de la encuesta.

Codificación pregunta 12. Importancia en sociedad.

- 12.1 Participación
- 12.2 Desarrollo
- 12.3 Compromiso
- 12.4 Innovación
- 12.5 Crecimiento

Bloque 4. Acceso a los conjuntos de datos: fuentes de datos abiertos consultadas.

**13. ¿Utiliza o ha utilizado en sus proyectos/investigación, algún tipo de portal de datos abiertos de los que se enumeran a continuación?**

\*Selección múltiple

- Datos Abiertos de la Generalitat Valenciana (<https://dadesobertes.gva.es/>).
- Datos Abiertos de otras instituciones públicas de la Comunidad Valenciana.
- Datos abiertos de organismos internacionales.
- Portales/fuentes de datos europeas.
- Datos abiertos del gobierno (datos.gob.es).
- Datos de otras Comunidades Autónomas y organizaciones autonómicas.
- Datos abiertos de ayuntamientos.
- Empresas.
- Otros.

**14. ¿Ha accedido a alguno de los portales de datos abiertos autonómicos siguientes?**

\*Selección múltiple

Portal de datos abiertos:

- Junta de Andalucía
- Aragón
- Asturias
- Cantabria
- Castilla La Mancha
- Castilla y León
- Cataluña
- Ceuta
- Madrid
- Comunidad Valenciana
- Extremadura
- Galicia
- Melilla
- Navarra
- País Vasco
- Región de Murcia
- La Rioja
- Baleares
- Islas Canarias

**15. Si ha utilizado portales de datos abiertos de otras administraciones u otras fuentes de datos, por favor, especifique cuáles:**

Respuesta abierta: Ejemplo. Instituto nacional de estadística.

**16. ¿Cómo son de importantes las siguientes motivaciones para trabajar o utilizar sistemas de datos abiertos de las administraciones?**

	Nada importante (0)	Poca importancia (1)	Necesario (2)	Importante (3)	Muy importante (4)
16.1 Crecimiento y formación					
16.2 Capacidad de comprensión del funcionamiento de gobierno					
16.3 Desarrollo de servicio o producto innovador basado en datos					
16.4 Cumplimiento de objetivos laborales					
16.5 Desarrollo de la web semántica y datos enlazados					
16.6 Mejorar el acceso y la comprensión de la información pública por la ciudadanía.					

Tabla 70: Pregunta 16 de la encuesta.

Codificación pregunta 16. Motivaciones de uso.

- 16.1 Crecimiento-formación
- 16.2 Comprensión
- 16.3 Desarrollo
- 16.4 Objetivos profesionales
- 16.5 Web semántica
- 16.6 Acceso

---

**17. En materia de apertura de datos, ¿cuáles son, en su opinión, las cuestiones más importantes a las que se enfrentan los gobiernos hoy en día?**

\*Selección múltiple

Por ejemplo: falta de estandarización de datos, apertura con formatos técnicos reutilizables...

- 17.1 Cuestiones de carácter técnicas: dificultad para encontrar formatos reutilizables.
- 17.2 Falta de estandarización de los datos.
- 17.3 Falta de interoperabilidad administrativa.
- 17.4 Financiación.
- 17.5 Cuestiones de publicidad y puesta en valor de los datos.
- 17.6 Falta de formación en materia de reutilización.
- 17.7 Otras: especifique cuáles.

Codificación pregunta 17. Cuestiones de importancia.

17.1 Técnicas

17.2 Estandarización

17.3 Interoperabilidad

17.4 Financiación

17.5 Publicidad activa

17.6 Formación

17.7 Otras

---

Bloque 5. Uso y reutilización de datos (I).

**18. Existen numerosas formas de reutilizar los datos abiertos disponibles en los portales. Seleccione las opciones que haya hecho alguna vez para reutilizar los datos.**

\*Selección múltiple

- 18.1 He descargado conjuntos de datos abiertos.
- 18.2 He utilizado datos abiertos para desarrollar una plataforma web.
- 18.3 He utilizado datos abiertos en el desarrollo de un servicio web o API.
- 18.4 He utilizado datos abiertos en el desarrollo de una aplicación móvil.
- 18.5 He utilizado datos abiertos en el desarrollo de investigación científica (artículos, informes, presentaciones).
- 18.6 He utilizado datos abiertos del gobierno abierto para llevar a cabo campañas reivindicativas.
- 18.7 He integrado datos del gobierno abierto con una aplicación o sitio web existente.
- 18.8 He utilizado datos abiertos para la creación de mapas de situación.
- 18.9 He utilizado datos abierto para la creación de visualizaciones.
- 18.10 He utilizado datos abiertos para análisis estadísticos.
- 18.11 He utilizado datos abiertos para consultar información pública de interés.
- 18.12 He utilizado datos abiertos para facilitar la divulgación de información pública a la ciudadanía.
- 18.13 Planeo hacer más uso de los datos abiertos.
- 18.14 Nunca he utilizado datos abiertos.
- 18.5 Otras: Especifique cuáles.

Codificación pregunta 18. Reutilización.

- 18.1 Descarga
- 18.2 Plataforma Web
- 18.3 Servicio web - API
- 18.4 App móvil
- 18.5 Investigación
- 18.6 Campañas
- 18.7 Integración
- 18.8 Mapas de situación
- 18.9 Visualizaciones
- 18.10 Estadística
- 18.11 Consulta
- 18.12 Divulgación
- 18.13 Seguiré usando
- 18.14 Sin uso
- 18.15 Otros

**19. Los datos pueden ser liberados en formatos diferentes. Pensando en la reutilización, valore los formatos técnicos de estos.**

	Formato inadecuado para reutilización (0)	Reutilización básica (1)	Reutilización avanzada con características mejorables (2)	Reutilización avanzada (3)
CSV				
XLS				
PDF				
XML				
JSON				
GeoJSON				
RDF				
KML				
SQL				
RDF				
RSS				
WMS				
WFS				
GML				
SHP				

Tabla 71: Pregunta 19 de la encuesta.

## Bloque 6. Uso y reutilización de datos (II).

### Uso particular de datos abiertos.

En caso de que no se haya realizado ningún uso, puede saltar esta sección y habrá finalizado la encuesta. Si ha utilizado datos del gobierno abierto para múltiples proyectos, seleccione uno y descríbalos brevemente. Si quiere que este estudio tenga en cuenta otros proyectos, podrá proporcionar información adicional sobre otro estudio que haya llevado a cabo.

**20. Describa el proyecto, investigación, desarrollo de producto, etc., que haya llevado a cabo con datos abiertos:** Ejemplo: La descarga de datos abiertos de la Comunidad Autónoma relacionados con la salud me ha servido para publicar un artículo en una revista.

### 21. ¿Cuál era el objetivo del proyecto anterior?

\*Selección múltiple.

- 21.1 Investigación estadística
- 21.2 Exploración de datos
- 21.3 Campañas políticas
- 21.4 Asuntos sociales
- 21.5 Literatura científica: informes, artículos
- 21.6 Desarrollo de un sitio web
- 21.7 Desarrollo API
- 21.8 Desarrollo de app
- 21.9 Desarrollo de herramienta interna para mi organización
- 21.10 Creación de visualizaciones
- 21.11 Desarrollo de mapeos
- 21.12 Investigar las tecnologías web
- 21.13 Otros: Especifique cuáles

Codificación pregunta 21. Objetivo del proyecto.

- 21.1 Estadística
- 21.2 Exploración
- 21.3 Campañas políticas
- 21.4 Asuntos sociales
- 21.5 Investigación
- 21.6 Desarrollo web
- 21.7 Desarrollo API
- 21.8 Desarrollo App
- 21.9 Herramientas internas
- 21.10 Visualizaciones
- 21.11 Mapeos
- 21.12 Tecnologías web
- 21.13 Otros

---

**22. El uso de datos abiertos en el/los proyecto/s en lo que he participado, se centra en cuestiones:**

\*Selección múltiple.

- 22.1 Locales
- 22.2 Autonómicas
- 22.3 Estatales
- 22.4 Continentales
- 22.5 Internacionales
- 22.6 Otras: Especificar

**23. ¿Qué tipos de datos utilizó para su proyecto/investigación?**

\*Selección múltiple.

- Datos comerciales
- Datos sobre el crimen y la justicia
- Datos sobre educación y habilidades
- Datos sobre el medio ambiente y el clima
- Salud y discapacidad
- Vivienda
- Ingresos y gastos del gobierno
- Mercado laboral
- Población
- Transporte
- Análisis web
- Ayuda internacional y desarrollo
- Datos jurídicos y administrativos
- Economía y comercio
- Geográficos
- Políticos
- Ninguno de los anteriores
- Otros: especifique cuales

**24. Para el desarrollo y reutilización, ¿qué lenguajes/programas/tecnologías ha utilizado para el análisis, tratamiento de los datos?**

\*Selección múltiple

- Hojas de cálculo
- Procesadores de texto
- Lenguajes de programación (PHP, Python, Ruby...)
- JSON
- XML
- SPARQL
- Suites como Google
- Flash
- Herramientas de mapeo de datos
- Herramientas de visualización de datos



- Software de estadística
- Ninguno de los anteriores
- Otros: especifique cuales

### Anexo III. Recopiladores

Los 25 recopiladores son conjuntos de direcciones de correo electrónico que permite crear la plataforma Survey Monkey para la distribución de los envíos de correo, con 22791 direcciones de 47 universidades españolas, y la fecha del envío.

Recopilador 1: Fecha del envío: 20/04/2021

- Universidad Católica de Ávila
  - Total direcciones de correo: 100
- Universidad Politécnica de Cartagena
  - Total direcciones (directores de departamento): 47
- Universidad de Murcia
  - Total direcciones (directores de departamento): 149
- Universidad de Santiago de Compostela
  - Total direcciones: 156
- Universidad Rey Juan Carlos I
  - Total direcciones: 59
- Universidad de Extremadura
  - Total direcciones (directores de departamento): 40

Recopilador 2. Fecha del envío: 20/04/2021

- Universidad de la Rioja (UNIR)
  - Total direcciones (departamentos): 11
- Universidad País Vasco
  - Total direcciones (directores departamento): 157
- Universidad Politécnica de Madrid
  - Total direcciones (varios departamentos): 220
- Universidad de Alcalá
  - Total direcciones (Departamentos): 23
- Universidad Carlos III
  - Total direcciones (Personales, varios departamentos, biblioteconomía, periodismo, estadística, matemáticas): 155

Recopilador 3. Fecha del envío: 21/04/2021

- Universidad de Girona
  - Total direcciones (directores grupos de investigación): 105
- Universidad de Castilla la Mancha
  - Total direcciones (varios personales varios departamentos): 136
- Universidad de Salamanca
  - Total direcciones (varios personales varios departamentos): 211

Recopilador 4. Fecha del envío: 21/04/2021

- Universitat Politècnica de Catalunya (UPC)
  - Total direcciones (Varios personales varios departamentos): 262
- Universidad de Las Palmas de Gran Canaria (ULPGC)
  - Total direcciones (Varios personales varios departamentos): 459
- Universidad de Oviedo
  - Total direcciones (Varios personales varios departamentos): 413

Recopilador 5. Fecha del envío: 21/04/2021

- Universidad de Zaragoza (UZ)
  - Total direcciones (Varios personales varios departamentos): 432
- Universidad de Málaga (UM)
  - Total direcciones (Varios personales varios departamentos): 500

Recopilador 6. Fecha del envío: 21/04/2021

- Universidad de Granada
  - Total direcciones (Varios personales varios departamentos): 480
- Universidad Complutense de Madrid
  - Total direcciones (Varios personales varios departamentos): 529

Recopilador 7. Fecha del envío: 22/04/2021

- Universitat Rovira i Virgili
  - Total direcciones (Varios personales varios departamentos): 236
- Universitat Pompeu Fabra
  - Total direcciones (directores departamento): 8
- Universidad Autónoma de Madrid
  - Total direcciones (Varios personales varios departamentos): 195
- Universitat Autònoma de Barcelona
  - Total direcciones (Varios personales varios departamentos): 194

Recopilador 8. Fecha del envío: 22/04/2021

- Universidad de Burgos
  - Total direcciones (Varios personales varios departamentos): 249
- Universidad de la Laguna (Tenerife)
  - Total direcciones (Varios personales varios departamentos): 257
- Universidad Pablo de Olavide (Sevilla)
  - Total direcciones (Varios personales varios departamentos): 134
- Universidad Pablo de Málaga
  - Total direcciones (Varios personales varios departamentos): 375

Recopilador 9. Fecha del envío: 22/04/2021

- Universidad de Jaén
  - Total direcciones (Varios personales varios departamentos): 395
- Universidad de Huelva
  - Total direcciones (Varios personales varios departamentos): 51
- Universidad de Almería
  - Total direcciones (Varios personales varios departamentos): 66
- Universidad de Cádiz
  - Total direcciones (Varios personales varios departamentos): 339

Recopilador 10. Fecha del envío: 23/04/2021

- Universidad de Valladolid. Instituto de Estudios Europeos
  - Total direcciones (Varios personales varios departamentos): 63
- Universidad de Salamanca. Instituto de Neurociencias
  - Total direcciones (Varios personales varios departamentos): 102
- Universidad de Valladolid. Instituto Universitario CINQUIM
  - Total direcciones (Varios personales varios departamentos): 27

- Universidad de Valladolid. IPS
    - Total direcciones (Varios personales varios departamentos): 67
  - Universidad de Valladolid. Ins. Mat.
    - Total direcciones (Varios personales varios departamentos): 85
  - Universitat Oberta de Catalunya
    - Total direcciones (Varios personales varios departamentos): 235
  - Universitat de Lleida
    - Total direcciones (Varios personales varios departamentos): 186
- Recopilador 11. Fecha del envío: 23/04/2021
- Universitat Rovira i Virgili
    - Total direcciones (Varios personales varios departamentos): 272
  - Universidad Complutense de Madrid
    - Total direcciones (Varios personales varios departamentos): 265
  - Universidad Carlos III
    - Total direcciones (Varios personales varios departamentos): 320
- Recopilador 12. Fecha del envío: 26/04/2021
- Universidad Politécnica de Madrid
    - Total direcciones (Varios personales varios departamentos): 163
  - Universidad País Vasco
    - Total direcciones (Varios personales varios departamentos): 148
  - Universidad de Murcia
    - Total direcciones (Varios personales varios departamentos): 171
  - Universidad de Cartagena
    - Total direcciones (Varios personales varios departamentos): 217
- Recopilador 13. Fecha del envío: 26/04/2021
- Universidad de Castilla la Mancha
    - Total direcciones (Varios personales varios departamentos): 157
  - Universidad Autónoma de Madrid
    - Total direcciones (Varios personales varios departamentos): 252
  - Universitat de Lleida
    - Total direcciones (Varios personales varios departamentos): 393
- Recopilador 14. Fecha del envío: 26/04/2021
- Universidad de la Laguna
    - Total direcciones (Varios personales varios departamentos): 230
  - Universidad de Burgos
    - Total direcciones (Varios personales varios departamentos): 234
  - Universidad de Oviedo
    - Total direcciones (Varios personales varios departamentos): 219
  - Universidad de Zaragoza
    - Total direcciones (Varios personales varios departamentos): 199
  - Universidad Pablo Olavide (Sevilla)
    - Total direcciones (Varios personales varios departamentos): 172
- Recopilador 15. Fecha del envío: 27/04/2021
- Universidad de Jaen
    - Total direcciones (Varios personales varios departamentos): 265

- Universidad de Granada
    - Total direcciones (Varios personales varios departamentos): 672
- Recopilador 16. Fecha del envío: 27/04/2021
- Universidad de Las Palmas de Gran Canaria
    - Total direcciones (Varios personales varios departamentos): 297
  - Universidad de Murcia
    - Total direcciones (Varios personales varios departamentos): 340
  - Universitat Autònoma de Barcelona
    - Total direcciones (Varios personales varios departamentos): 270
- Recopilador 17. Fecha del envío: 28/04/2021
- Universidad de Almería
    - Total direcciones (Varios personales varios departamentos): 38
  - Universidad de Cádiz
    - Total direcciones (Varios personales varios departamentos): 700
  - Universitat Autònoma de BCN
    - Total direcciones (Varios personales varios departamentos): 445
- Recopilador 18. Fecha del envío: 28/04/2021
- Universidad de Salamanca
    - Total direcciones (Varios personales varios departamentos): 270
  - Universidad de
    - Total direcciones (Varios personales varios departamentos): 704
- Recopilador 20. Fecha del envío: 30/04/2021
- Universidad Autónoma de Madrid
    - Total direcciones (Varios personales varios departamentos): 195
  - Universidad de Salamanca
    - Total direcciones (Varios personales varios departamentos): 111
  - Universitat Autònoma de BCN
    - Total direcciones (Varios personales varios departamentos): 314
  - Universidad de las Palmas de GC
    - Total direcciones (Varios personales varios departamentos): 424
- Recopilador 21. Fecha del envío: 30/04/2021
- Universidad de la Laguna
    - Total direcciones (Varios personales varios departamentos): 324
  - Universidad de Granada
    - Total direcciones (Varios personales varios departamentos): 767
- Recopilador 22. Fecha del envío: 30/04/2021
- Universidad de Granada
    - Total direcciones (Varios personales varios departamentos): 656
  - Universidad
    - Total direcciones (Varios personales varios departamentos): 427
- Recopilador 23. Fecha del envío: 3/5/2021
- Universidad de Jaen
    - Total direcciones (Varios personales varios departamentos): 398
  - Universidad de las Palmas de GC
    - Total direcciones (Varios personales varios departamentos): 472

- Universidad de la Laguna
    - Total direcciones (Varios personales varios departamentos): 182
- Recopilador 24. Fecha del envío: 3/5/2021
- Universitat Autònoma de BCN
    - Total direcciones (Varios personales varios departamentos): 762
  - Universidad de Salamanca
    - Total direcciones (Varios personales varios departamentos): 103
  - Universitat Rovira i Virgili
    - Total direcciones (Varios personales varios departamentos): 310
- Recopilador 25. Fecha del envío: 4/5/2021
- Universidad Carlos III de Madrid
    - Total direcciones (Varios personales varios departamentos): 357
  - Universidad de Castilla la Mancha
    - Total direcciones (Varios personales varios departamentos): 201
  - Universidad Autónoma de Madrid
    - Total direcciones (Varios personales varios departamentos): 50
  - Universidad Complutense de Madrid
    - Total direcciones (Varios personales varios departamentos): 335
  - Universidad de Granada
    - Total direcciones (Varios personales varios departamentos): 1082