





INFOTEC CENTRO DE  
INVESTIGACIÓN E INNOVACIÓN EN  
TECNOLOGÍAS DE LA INFORMACIÓN Y

DIRECCIÓN ADJUNTA DE INNOVACIÓN Y  
CONOCIMIENTO  
GERENCIA DE CAPITAL HUMANO  
POSGRADOS

**“UNA PROPUESTA  
BASADA EN TIC  
PARA PROMOVER  
LA DISPONIBILIDAD  
DE DATOS  
ABIERTOS EN EL  
GOBIERNO DEL  
ESTADO DE  
AGUASCALIENTES”**

PROPUESTA DE INTERVENCIÓN  
Que para obtener el grado de MAESTRO EN  
GESTIÓN DE INNOVACIÓN DE LAS  
TECNOLOGÍAS DE INFORMACIÓN Y  
COMUNICACIÓN

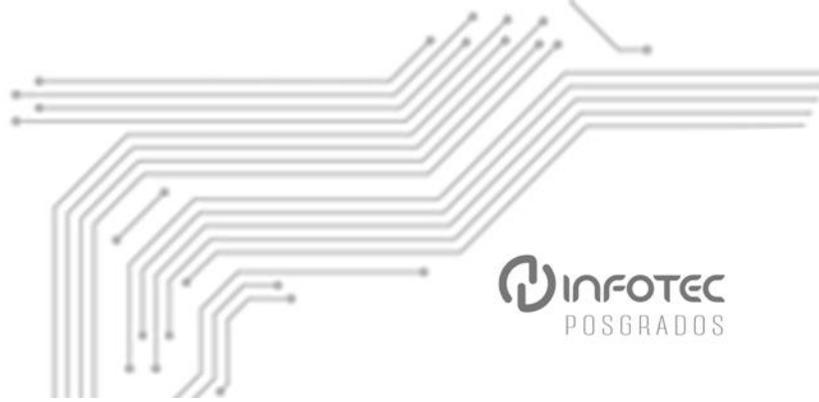
Presenta:

**Baltazar Ruiz García**

Asesor:

**Dr. Christian Iván Becerril-Velasco**

Ciudad de México, noviembre, 2020.



**AUTORIZACIÓN DE IMPRESIÓN Y NO ADEUDO EN BIBLIOTECA**  
**MAESTRÍA EN GESTIÓN DE INNOVACIÓN DE LAS TECNOLOGÍAS DE**  
**INFORMACIÓN Y COMUNICACIÓN**

Ciudad de México, 12 de julio de 2021.  
*INFOTEC-DAIC-GCH-SE-0226/2021.*

La Gerencia de Capital Humano / Gerencia de Investigación hacen constar que el trabajo de titulación intitulado

**UNA PROPUESTA BASADA EN TIC PARA PROMOVER LA**  
**DISPONIBILIDAD DE DATOS ABIERTOS EN EL GOBIERNO DEL**  
**ESTADO DE AGUASCALIENTES**

Desarrollado por el alumno **Baltazar Ruiz García** y bajo la asesoría del **Dr. Christian Iván Becerril-Velasco**; cumple con el formato de biblioteca. Por lo cual, se expide la presente autorización para impresión del proyecto terminal al que se ha hecho mención.

Asimismo se hace constar que no debe material de la biblioteca de INFOTEC.

Vo. Bo.



**Lic. Juan Ramón Abarca Damián**  
Coordinador de Biblioteca

**Anexar a la presente autorización al inicio de la versión impresa del trabajo referido que ampara la misma.**

## Agradecimientos

En primer lugar, a Dios por darme la vida, salud y oportunidad de poder perseguir un sueño.

A mi esposa Laura e hija Cindy por su paciencia y comprensión, así como por todo el amor y motivación.

A mi madre y a mis hermanos por siempre estar ahí cuando los he necesitado.

A mis amigos y compañeros de clase: Amanda Janeth González Macías, María del Carmen Hernández Martínez, Delfino Saldaña Hernández y Juan Gabriel López Sánchez, por su ánimo y acompañamiento en las clases y en las noches de desvelo.

A mis superiores laborales: Juan Francisco Larios Esparza y Alejandro Viay Danielli por su solidaridad, soporte e impulso.

Por último, pero no menos importante, a mis maestros, en especial a mi asesor Dr. Christian Iván Becerril-Velasco por su orientación y grandes enseñanzas, a la Dra. Juana Hernández Chavarría por la revisión del presente trabajo y al Dr. Héctor Edgar Buenrostro Mercado por su permanente apoyo en el transcurso de la MGITIC.

## Tabla de contenido

<b>Introducción.....</b>	<b>1</b>
<b>Capítulo 1. Marco conceptual.....</b>	<b>8</b>
1.1 Gobierno abierto .....	8
1.2 Origen de los datos abiertos .....	10
1.3 Concepto y características de los datos abiertos.....	11
1.4 Riesgos y su mitigación en la publicación de datos abiertos .....	16
1.5 Entorno internacional .....	17
1.6 Datos abiertos en América Latina y el Caribe .....	19
1.7 Datos Abiertos en México .....	21
1.8 Aspectos críticos para la publicación de datos abiertos.....	22
1.9 Plataformas de datos abiertos .....	24
1.9.1 Modelos de datos .....	25
1.9.1.1 Ciclo de vida de los datos abiertos.....	25
1.9.1.2 Cadena de valor de los datos.....	27
1.9.1.3 Procesos de datos .....	28
1.9.2 Requerimientos no funcionales para una infraestructura de datos abiertos gubernamentales.....	29
1.9.2.1 Aspectos técnicos .....	30
1.9.3 Requerimientos funcionales para una infraestructura de datos abiertos gubernamentales.....	32
1.9.4 Métricas de la infraestructura .....	33
1.9.4.1 Calidad de la Información.....	33
1.9.4.2 Calidad del sistema.....	34
1.9.4.3 Calidad del servicio .....	34
1.9.5 Comparativa de plataformas de datos abiertos .....	35
1.9.6 Plataformas de datos abiertos en México.....	38
1.9.6.1 datos.gob.mx.....	39
1.9.6.2 datos.jalisco.gob.mx.....	40
1.9.6.3 datos.puebla.gob.mx .....	41

<b>Capítulo 2. Metodología de la investigación .....</b>	<b>44</b>
2.1 Pregunta de investigación .....	44
2.2 Objetivos .....	44
2.2.1 Objetivo general .....	44
2.2.2 Objetivos específicos.....	44
2.3 Abordaje metodológico .....	45
2.4 Entorno de Investigación.....	47
<b>Capítulo 3. Caso de estudio Aguascalientes .....</b>	<b>51</b>
3.1 Recolección de datos (DoRCU) .....	51
3.1.1 Titulares de dependencias y entidades gubernamentales.....	51
3.1.2 Información recabada de los responsables de Informática .....	53
3.2 Inventario de sistemas .....	55
3.3 Propuesta de implementación.....	57
3.3.1 Requerimientos para el diseño de la plataforma .....	57
3.3.2 Características y especificaciones técnicas de los datos para ser publicados como datos abiertos. ....	62
3.3.3 Actores involucrados en la generación.....	64
3.3.4 Conjuntos de datos .....	66
<b>Conclusiones.....</b>	<b>70</b>
<b>Bibliografía.....</b>	<b>73</b>
<b>ANEXO 1 – Guía de entrevista.....</b>	<b>82</b>
<b>ANEXO 2 – Cuestionario de entrevista.....</b>	<b>83</b>
<b>ANEXO 3 – Información del área de Informática .....</b>	<b>84</b>

## Índice de figuras

Figura 1. Creación de portales de datos abiertos en Europa.....	18
Figura 2. Ciclo de vida de datos abiertos. ....	26
Figura 3. Cadena de valor de los datos abiertos. ....	28
Figura 4. Conectividad de dependencias y entidades del GEA con SAE.....	59
Figura 5. Esquema de servidores y red.....	62
Figura 6. Equipo de trabajo para generación y publicación de datos abiertos.....	65

## Índice de cuadros

Cuadro 1. Descripción de los elementos de gobierno abierto.....	8
Cuadro 2. Portales de datos abiertos en América Latina. ....	20
Cuadro 3. Métodos y herramientas en cada fase del ciclo de vida de los datos abiertos. ....	26
Cuadro 4. Categorías de requerimientos funcionales.....	32
Cuadro 5. Sitios de datos abiertos relevantes por plataforma. ....	36
Cuadro 6. Comparativa de características por plataforma.....	37
Cuadro 7. Comparativa de aspectos técnicos por plataforma.....	38

## Siglas y abreviaturas

<b>API</b>	Interfaz del programa de aplicación, por sus siglas en inglés.
<b>CKAN</b>	Red integral de archivos de conocimiento, por sus siglas en inglés.
<b>CMOV</b>	Coordinación General de Movilidad
<b>CPLAP</b>	Coordinación General de Planeación y Proyectos
<b>CSV</b>	Formato de valores separados por comas, por sus siglas en inglés.
<b>DKAN</b>	Red integral de archivos de conocimiento basado en Drupal.
<b>DoRCU</b>	Documentación de Requerimientos Centrada en el Usuario.
<b>DRUJAL</b>	Distribuidor de Drupal para los Portales <i>Web</i> en el Gobierno del Estado de Jalisco.
<b>ECLAC</b>	Comisión Económica para América Latina y el Caribe, por sus siglas en inglés.
<b>GEA</b>	Gobierno del Estado de Aguascalientes.
<b>GODAN</b>	Siglas en inglés de Datos Abiertos Globales para Agricultura y Nutrición.
<b>HTML</b>	Lenguaje de Marcado de Hiper-Texto.
<b>HTTP</b>	Protocolo de Transferencia de Hiper-Texto.
<b>IEA</b>	Instituto de Educación de Aguascalientes.
<b>IEGEMER</b>	Instituto Estatal de Gestión Empresarial y mejora Regulatoria.
<b>IIS</b>	Siglas en inglés de Servicios de Información de Internet
<b>INEGI</b>	Instituto Nacional de Estadística y Geografía.
<b>INFOTEC</b>	Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación.
<b>IoT</b>	Traducido como Internet de las cosas.
<b>ISSEA</b>	Instituto de Servicios de Salud de Aguascalientes.
<b>JSON</b>	JavaScript Object Notation. o Notación de Objeto JavaScript
<b>OCDE</b>	Organización para la Cooperación y el Desarrollo Económicos.
<b>PDF</b>	Siglas en inglés de Formato de Documento Portable

<b>RDF</b>	Traducido como Marco de Descripción de Recursos
<b>REST</b>	Representational State Transfer o Transferencia de Estado Representacional.
<b>SaaS</b>	Traducido como Software como un Servicio
<b>SAE</b>	Secretaría de Administración del Estado de Aguascalientes
<b>SECTURE</b>	Secretaría de Turismo del Estado
<b>SEDEC</b>	Secretaría de Desarrollo Social
<b>SEFI</b>	Secretaría de Finanzas.
<b>SEGUOT</b>	Secretaría de Gestión Urbanística, Ordenamiento Territorial, Registral y Catastral
<b>SI</b>	Sistemas de Información
<b>SITEUR</b>	Sistema de Tren Eléctrico Urbano
<b>SOP</b>	Secretaría de Obras Públicas.
<b>SPARQL</b>	Traducido como Lenguaje de consultas semántico para bases de datos.
<b>SSMAA</b>	Secretaría de Sustentabilidad, Medio Ambiente y Agua
<b>SWB</b>	Traducido como Constructor de Web Semántica
<b>TIC</b>	Tecnologías de la Información y Comunicación
<b>URI</b>	Identificador uniforme de recursos, por sus siglas en inglés.
<b>URL</b>	Localizador uniforme de recursos, por sus siglas en inglés.
<b>W3C</b>	Consorcio WWW.
<b>WWW</b>	Aunque no de manera literal, puede traducirse como red mundial de información en páginas HTML y accedida vía HTTP.
<b>XML</b>	Traducido como Lenguaje de Marcado Extensible.

## Glosario

### “B”

**Big data:** El término *big data* se refiere al almacenamiento de una gran cantidad de datos, usualmente sin una estructura determinada y de manera veloz y veraz.

### “C”

**Cloud computing:** Es un esquema de computación en el que las aplicaciones y los datos no se encuentran almacenados en los equipos propios y tangibles de una empresa u organización, sino que se encuentran ubicados en algún lugar de la Internet. Usualmente se provee como un servicio y no como un bien por parte de empresas especializadas.

### “D”

**Data linking:** Es una tecnología por medio del cual se interrelacionan datos en la *World Wide Web* tomando en cuenta los estándares establecidos para ello.

**Digital Security:** Establecimiento de medidas preventivas y correctivas de seguridad en torno a las Tecnologías de la Información y Comunicación. Las medidas y acciones comprenden un amplio rango de aspectos, que van desde el aspecto meramente humano del acceso físico a lugares exclusivos, hasta el uso sofisticado de algoritmos que protegen el almacenamiento y transferencia de datos. El término también puede encontrarse con el nombre de ciberseguridad.

### “E”

**Entornos virtuales:** En los inicios de la computación y del modelo cliente-servidor, lo normal era encontrar un equipo físico central realizando las funciones de servidor, mientras que un número limitado de equipos se conectaba como clientes. En los denominados entornos virtuales actuales, el rol de servidor puede ser desempeñado por una fracción de un servidor físico, pero también puede ser desempeñado por más de un equipo físico para potenciar las capacidades. De esta manera, es posible la asignación de mayores recursos a un servidor para el cumplimiento de su objetivo. Estos recursos no necesariamente se encuentran físicamente en la

organización o empresa, sino que pueden ser parte del esquema de *cloud computing*.

## “H”

**Hadoop:** Hadoop se usa por empresas como Google o Yahoo, junto con otras formas de desarrollo, para el manejo y computación de datos bajo el esquema de *big data* (Hernández-Dominguez & Hernández-Yeja, 2015).

**Hashing:** Es una técnica que utiliza funciones para la conversión de datos en otro tipo, de manera que el resultado permita una mayor seguridad en el transporte y almacenamiento de datos, así como el acceso a los mismos.

## “I”

**Industrial Internet:** Es la aplicación de *Internet of Things* al ambiente industrial y productivo, conectando sensores y maquinaria a equipos de cómputo locales o en *cloud computing* para una mayor productividad.

**Internet of Things:** Es la integración de dispositivos electrónicos de uso cotidiano a la Internet, de manera que se obtenga un mayor control de los mismos, logrando así un mayor bienestar o capacidades adicionales. Esta integración se hace principalmente por medio de la red local, pero transfiriendo los datos que generan diversos sensores e instrumentos. Un claro ejemplo de Internet de las cosas en el hogar, es la integración de un refrigerador a la Internet.

## “J”

**JavaScript:** Es un lenguaje de programación de amplio uso que permite funcionalidades adicionales en las páginas *Web*.

## “M”

**Minería o data mining:** Son técnicas que se usan en las Tecnologías de la Información para analizar y encontrar patrones, similitudes o tendencias en grandes cantidades de datos o de *big data*.

## “O”

**Open Refine:** Anteriormente conocido como *Google Refine*, es una herramienta para trabajar con datos sin estructura.

## “R”

**Rapidminer:** Herramienta de las Tecnologías de la Información que sirve para la minería de datos, con una interfaz gráfica.

## “S”

**Servidores:** En el entorno de las Tecnologías de Información y Comunicación, los servidores son equipos de grandes capacidades de procesamiento, capaces de proveer servicios a un número máximo de clientes. Entre los servicios más comunes que se pueden encontrar en los servidores son los de aplicaciones, bases de datos, impresión, transferencia de archivos y *Web*.

## “T”

**TIC:** Son las iniciales de las Tecnologías de la Información y Comunicación, y se refieren a aquellas tecnologías que facilitan el acceso a la información por medio de medios de comunicación, como lo puede ser el Internet. El alcance de estas tecnologías es bastante amplio y aplica a las diferentes situaciones o dispositivos en las que podemos encontrar la información.

*“To achieve the goals of sustainable development, critical data must be open and available for reuse by anyone, anywhere, anytime. This is the vision.”*

Sir Tim Berners-Lee, inventor del *World Wide Web*, y director fundador de *Web Foundation*, sobre la Carta Internacional de Datos Abiertos, Asociación Global para el desarrollo sostenible de datos. Septiembre 2015.

## Introducción

El presente trabajo se realiza como parte de los requisitos de obtención de grado de la Maestría en Gestión de Innovación de las Tecnologías de Información y Comunicación, impartida en la sede Aguascalientes de INFOTEC (Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación), con la modalidad de “Propuesta de intervención”, debido a que existe la intención de implementarla en el Gobierno del Estado de Aguascalientes.

Teniendo a los datos abiertos como elemento esencial del trabajo, el documento se divide en cuatro capítulos o secciones principales. En la primera parte o introducción, se plantean la justificación y el planteamiento del problema central, la ausencia de una plataforma web que contenga tanto los elementos de generación de datos abiertos, así como de publicación a través de un portal web en el Gobierno del Estado de Aguascalientes.

En el capítulo 1, se desarrollan las generalidades de los datos abiertos: su origen, características, riesgos, así como su uso en el entorno internacional, en América Latina y en México. Como parte de este capítulo, también se describen las generalidades de las plataformas de datos abiertos, abarcando aspectos relevantes como modelos de datos, requerimientos funcionales y no funcionales, además de las métricas de la infraestructura.

En el capítulo 2, correspondiente a la Metodología de la Investigación, se justifica el tipo de investigación y se enumeran los resultados esperados considerando las variables que se desprenden de la pregunta. Se plantean los objetivos, general y específicos, orientados a la generación de una plataforma web de datos abiertos. Lo anterior se realiza tomando en cuenta mi experiencia laboral de más de 15 años en el gobierno estatal y tres de ellos como responsable de los sistemas y portales web en la Secretaría de Administración, con el cargo de Coordinador de Sistemas de Información.

En el tercer y último capítulo, se desarrolla el caso de estudio para el Gobierno del Estado de Aguascalientes. Se describen el entorno de la investigación,

los resultados de las entrevistas, los requerimientos no funcionales aplicables cubriendo los aspectos técnicos básicos, así como el inventario de sistemas identificados. Como parte final del capítulo, se enlistan las acciones a realizar, las cuales integran la propuesta de implementación, considerando los objetivos previamente establecidos.

Para la generación de la bibliografía y la inserción de citas se utiliza la plataforma Mendeley, incluido su aplicativo de escritorio y el agregado para Microsoft Word<sup>1</sup>.

## **Justificación**

Los datos abiertos es uno de los temas con mayor relevancia y popularidad en la actual era de las Tecnologías de la Información y Comunicación, comparándolo con temas como *big data*, *smart cities*, *Internet of things*, etc., y como evidencia basta realizar una consulta en alguna plataforma de publicación de libros y artículos de ciencia como [springer.com](http://springer.com), [elsevier.com](http://elsevier.com) o [scholar.google.com](http://scholar.google.com).

Los conceptos de datos abiertos y gobierno abierto están muy relacionados en los aspectos de transparencia, participación ciudadana, tecnología e innovación (Amozorrutia, Augusto, & Ditadi, 2017). Sin embargo, los datos abiertos son una herramienta que permite alcanzar los objetivos del gobierno abierto, el cual supone una simetría en el diálogo que debe existir entre ciudadanos y su gobierno (Ferrer-sapena & Peset, 2011).

A nivel internacional, el término datos abiertos (*open data*) apareció por primera vez en un documento de una agencia científica americana, en donde se promovía el libre intercambio de información entre los países para el entendimiento y análisis de un fenómeno global que involucra la atmósfera, los océanos y la biósfera (Chignard, 2013). Chignard (2013), menciona que durante la administración de Obama nace el sitio [data.gov](http://data.gov) con la finalidad de que la información se divulgara rápidamente en formatos estándar y el público la pudiera encontrar y usar.

---

<sup>1</sup> Mendeley es una plataforma gratuita que permite gestionar las fuentes y recursos bibliográficos utilizados en un documento. Su sitio oficial es <https://www.mendeley.com/>

En la región de Latinoamérica y el Caribe se cuenta con la mayor cantidad de adopciones de la Carta Internacional de Datos Abiertos, tanto a nivel nacional como a nivel local; además, las plataformas de datos abiertos han crecido anualmente y se han organizado eventos internacionales como *ConDatos* (Munte-Kunigami & Serale, 2018).<sup>2</sup>

En el año 2015, el presidente de los Estados Unidos Mexicanos, Enrique Peña Nieto, firmó el decreto por el que se establece la regulación en materia de datos abiertos y establece también la creación del portal de Internet datos.gob.mx (Presidencia de la República, 2015), contando actualmente con 40,727 datos de 280 instituciones en diferentes categorías. En este portal federal de datos abiertos, al buscar archivos relacionados con Aguascalientes, se obtiene una lista de diferentes conjuntos de datos generados por distintas instituciones y dependencias federales y paraestatales, pero ningún conjunto generado por el Gobierno del Estado de Aguascalientes.

En el Estado de Aguascalientes, el Instituto Estatal de Gestión Empresarial y Mejora Regulatoria (IEGEMER), ahora parte de la Secretaría de Desarrollo Económico (SEDEC) a nivel Dirección General, es el enlace con datos.gob.mx, pero es la Secretaría de Administración del Estado (SAE) la responsable del mantenimiento y actualización del Portal Web del Gobierno del Estado de Aguascalientes (GEA) con la información que proveen las distintas dependencias y entidades del gobierno estatal. La sección de transparencia de dicho portal es un elemento muy importante a considerar en evaluaciones en materia de datos abiertos, como la que se lleva a cabo liderada por el Dr. Rodrigo Sandoval Almazán, de la Universidad Autónoma del Estado de México, en la que se tratan de identificar componentes de datos abiertos en los distintos portales gubernamentales estatales (Sandoval, 2018).

---

<sup>2</sup> El sitio <https://opendatacharter.net> indica el compromiso que actualmente 69 gobiernos y 49 organizaciones tienen para usar los datos abiertos con el fin de lograr mejores soluciones a las problemáticas actuales. Por otro lado, el sitio <https://condatos.org/> indica que ConDatos es la principal conferencia anual de América Latina y el Caribe para tratar asuntos sobre datos abiertos y gobierno abierto.

Siendo el Portal *Web* del GEA la puerta de acceso a los ciudadanos, a las empresas y a la población en general para obtener información, realizar trámites, transacciones y la participación ciudadana, es de vital importancia la publicación de datos abiertos que puedan contribuir a fomentar la innovación, a generar desarrollo económico en distintos ámbitos, a generar nuevos productos y a la reducción de las desigualdades (Muenta-Kunigami & Serale, 2018).

## **Planeamiento del problema**

Los beneficios económicos y sociales del uso los datos abiertos ya han sido ampliamente discutidos (Carrara, Chan, Fischer, & van Steenberg, 2015; Universida Nacional de Quilmes, 1996). Un ejemplo de esto es la Unión Europea, en la que para 2016 se esperaba que el tamaño de mercado directo de datos abiertos fuera de 55.3 billones de euros para el grupo EU28+, y que en 2020 sea de 75.7 billones (Carrara et al., 2015). Investigaciones señalan que los datos abiertos coadyuvan a alcanzar los beneficios potenciales para mejorar la confianza en el gobierno, asegurar la equidad de acceso a la formulación de políticas públicas e incluso son un elemento importante para lograr los objetivos de desarrollo sostenible definidos por las Naciones Unidas (OCDE, 2010; Open Data Watch, 2017). Adicionalmente, se han identificado motivadores como el mejoramiento de las relaciones públicas con el gobierno, así como mejorar los datos y procesos gubernamentales (Rothenberg, 2012).

Los datos abiertos pueden ser inter-operados entre entes y organizaciones con plataformas tecnológicas, las cuales se utilizan para reducir la brecha de comunicación entre quienes generan los datos y quienes los consumen (Charalabidis et al., 2018). También pueden transparentar las actividades del gobierno, así como facilitar el desarrollo de aplicaciones en beneficio de la sociedad (Estrada, Solís, Vieyra, Najera, & Vázquez, 2013). La interfaz de la plataforma de datos abiertos con el usuario es un portal *web*, que permite la publicación de datos abiertos, una vez que se han generado, pre-procesado, limpiado y almacenado (Charalabidis et al., 2018). Características como sostenibilidad a largo plazo, transparencia, apertura, disponibilidad y agilidad son deseables en una plataforma (Broad, Tennison, Starks, & Scott, 2015).

A nivel internacional, se tienen algunos ejemplos de los beneficios de las plataformas de datos abiertos, como la *Open Data Initiative* de Kenia (Rahemtulla et al., 2011). En África, el Caribe, Australia y Holanda, se tienen informes de mejoramiento de granjas y descubrimiento de nuevos mercados (GODAN, 2017). En América Latina y el Caribe existen algunos ejemplos de plataformas, como la presentada por Uruguay, en donde por medio de la interfaz de un portal se pueden obtener beneficios en salud con la iniciativa denominada *A tu servicio*.<sup>3</sup>

En México, como evidencia del empoderamiento de los ciudadanos por el uso de los datos abiertos, se tiene el caso del programa *Mejora tu escuela*, con el cual se toman mejores decisiones con los datos de este portal (Munte-Kunigami & Serale, 2018; Sandoval-Almazán et al., 2017).<sup>4</sup> Así mismo, a través de la coordinación de la estrategia digital nacional, se desarrolló la plataforma *Labora* para generar hasta 25 millones de pesos anuales con el uso innovador de datos (Munte-Kunigami & Serale, 2018).<sup>5</sup>

En el contexto local, Jalisco tiene un avance considerable a través de su plataforma *web datos.jalisco.gob.mx*, habilitada por una política contemplada en la estrategia Jalisco digital y orientada a impulsar el desarrollo tecnológico y la innovación (Grupo Banco Mundial, 2015). Jalisco es uno de los estados más innovadores del sector público (SITEUR, 2017) y el mejor posicionado en el *Ranking* de portales estatales de transparencia 2018, contando con 468 conjuntos de datos en diferentes categorías (Sandoval, 2018).<sup>6</sup>

En el Gobierno del Estado de Aguascalientes no se cuenta con una plataforma dedicada a los datos abiertos como en los casos mencionados. La disponibilidad de datos se limita al cumplimiento de obligaciones de transparencia correspondientes a la rendición de cuentas y a la publicación de archivos en formato

---

<sup>3</sup> *A tu Servicio* es una plataforma que permite visualizar y comparar indicadores de prestadores de servicios de salud de Uruguay, cuya página oficial es <https://atuservicio.msp.gub.uy/>

<sup>4</sup> *Mejora tu Escuela* es una plataforma cuya intención es transformar la educación en México por medio de la participación ciudadana; su página oficial es <http://www.mejoratuescuela.org/>

<sup>5</sup> *Labora* promueve la cultura de la innovación con el uso de datos. La página oficial es <http://www.labora.io/>

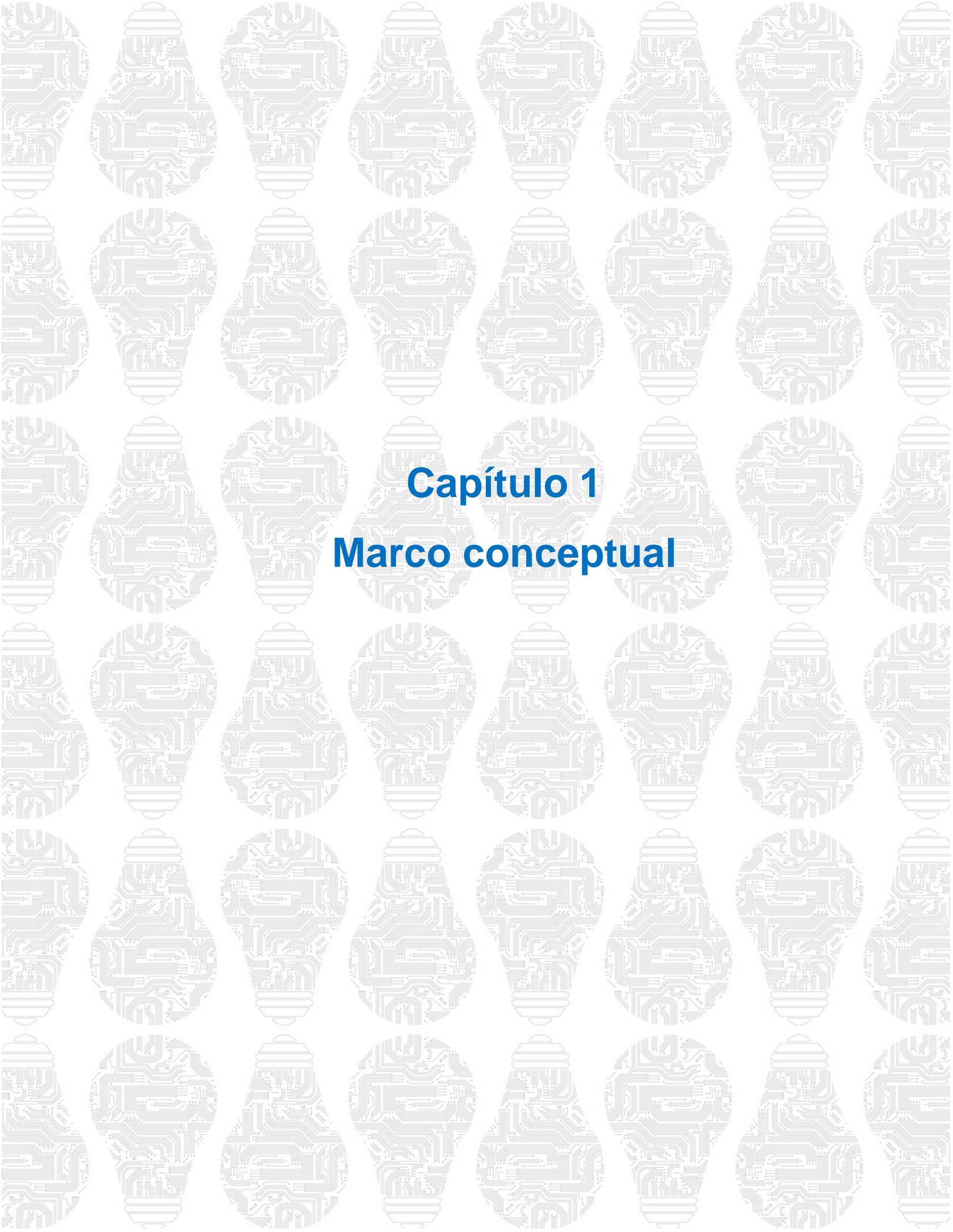
<sup>6</sup> Publicado por la revista U-GoB en <https://u-gob.com> y realizado en colaboración con investigadores del Centro de Docencia Económica, de la Universidad Autónoma del Estado de México y de la Benemérita Universidad Autónoma de Puebla.

PDF y Excel. Contrario a Jalisco, el portal de transparencia del Estado de Aguascalientes ocupó el lugar 27 de 32 en el ranking mencionado, aunque mejoró un poco al subir a la posición 20 en la evaluación correspondiente al 2019 (Sandoval, 2019).

Al no contar con una plataforma dedicada a los datos abiertos, en Aguascalientes se pierde la oportunidad de interacción entre proveedores y usuarios de datos dentro del ciclo de creación-uso de datos, respectivamente: creación-adquisición; preproceso-proceso; limpieza-uso; almacenamiento-colaboración; publicación-retroalimentación (Charalabidis et al., 2018). De igual manera, se pierde la disponibilidad de información para posibles inversiones, el mejoramiento de procesos administrativos y el mejoramiento de datos.

Pese a que en la implementación de una iniciativa de datos abiertos se ven involucrados elementos conceptuales, tecnológicos, organizacionales y legales (directivas y políticas) (Charalabidis et al., 2018). Para el caso de este trabajo, nos enfocaremos en los dos primeros con la intención de plantear una propuesta que pueda ser factible de implementar en el corto plazo, considerando el estado actual de los elementos organizacionales y legales.

Dado lo anterior, surge la interrogante sobre ¿Qué elementos conceptuales y tecnológicos es necesario considerar para el diseño de una plataforma informática en el Gobierno del Estado de Aguascalientes que permita la creación, procesamiento, almacenamiento y publicación de datos abiertos?



**Capítulo 1**  
**Marco conceptual**

## Capítulo 1. Marco conceptual

### 1.1 Gobierno abierto

Para abordar el tema de los datos abiertos, primero mencionaré que estos son un componente de gobierno abierto (Council of Europe, 2018). El gobierno abierto se basa en los principios de transparencia informativa, participación y colaboración ciudadana, con el elemento clave de la apertura de datos (Ramos-Alderete, 2016). La transparencia junto con la participación ciudadana y la rendición de cuentas como principios clave del gobierno abierto, demanda innovación por parte de las autoridades para interactuar con los ciudadanos, así como facilidades para el acceso a la información, en particular a través de la gestión de registros y datos abiertos (Council of Europe, 2018). Esto supera la idea básica que pudiera tenerse de que la transparencia se encasilla a la rendición de cuentas por parte de los sujetos obligados de gobierno. El cuadro 1 permite una mejor apreciación de los diferentes conceptos de la transparencia y la relación que se tiene entre ellos.

<b>Elemento de Gobierno Abierto</b>	<b>Componentes</b>
<b>Transparencia</b>	<ul style="list-style-type: none"><li>• <b>Acceso a la información</b> También conocido como derecho a la información. Provee derechos legales a los ciudadanos para pedir y recibir información del gobierno.</li><li>• <b>Datos abiertos</b> Grandes cantidades de datos que pueden ser libremente usados, modificados y compartidos por cualquiera.</li><li>• <b>Gestión de registros</b> El manejo de los recursos asegura que los datos sean accesibles y confiables, permitiendo que la apertura sea posible.</li></ul>
<b>Participación</b>	<ul style="list-style-type: none"><li>• <b>Espacio cívico</b> Contar con un espacio libre de restricciones para hablar, acceder información y participar en la toma de decisiones públicas.</li><li>• <b>Participación civil</b></li></ul>

	<p>El gobierno se involucra proactivamente con la sociedad civil durante el ciclo de generación de políticas, desde la identificación de prioridades hasta la implementación y evaluación de políticas.</p> <ul style="list-style-type: none"> <li>• <b>Protección de denunciantes</b></li> </ul> <p>Los denunciantes alertan a las organizaciones y a la sociedad en general sobre las irregularidades en las que surgen.</p>
<p><b>Rendición de cuentas</b></p>	<ul style="list-style-type: none"> <li>• <b>Auditoría</b></li> </ul> <p>Es fundamental para garantizar que el dinero local sea recolectado, administrado y gastado adecuadamente por el gobierno.</p> <ul style="list-style-type: none"> <li>• <b>Códigos de ética</b></li> </ul> <p>Proporciona una base importante para desafiar la negligencia profesional y la corrupción en el gobierno. Es lo que se espera de los servidores públicos.</p> <ul style="list-style-type: none"> <li>• <b>Escrutinio</b></li> </ul> <p>Ayuda a garantizar que los responsables de la toma de decisiones respondan y rindan cuentas a los residentes por sus decisiones.</p>

**Cuadro 1.** *Descripción de los elementos de gobierno abierto.*  
Fuente: Elaboración propia a partir de Council of Europe (2018).

De manera general, el cuadro 1 muestra que los elementos de la transparencia son el acceso a la información, los datos abiertos y la gestión de registros; los elementos de la participación son el espacio cívico, la participación civil y la protección de informantes; finalmente, los elementos de la rendición de cuentas son la auditoría, los códigos de ética y el escrutinio. El Congreso Europeo de Autoridades Locales y Regionales describe y ejemplifica cada uno de los elementos mencionados, de una manera muy simple, con la única intención de ubicar el contexto de los datos abiertos, el cual es nuestro tema de interés.

En la implementación de una iniciativa de datos abiertos, se ven involucrados elementos conceptuales, tecnológicos, organizacionales y legales (directivas y políticas) (Charalabidis et al., 2018). Para el caso de este trabajo, nos enfocaremos a los dos primeros: los elementos conceptuales, a fin de comprender las

generalidades de los datos abiertos, y los elementos tecnológicos, para establecer los requerimientos de una plataforma de generación y publicación de datos abiertos. En cuanto a los elementos organizacionales, las instituciones gubernamentales se encuentran con aspectos como la falta de procedimientos, así como amenazas de violaciones en la liberación de datos y riesgos de mal uso de los datos, para lo cual se requieren cambios (Charalabidis et al., 2018). Respecto a los elementos legales, se observa que a menudo se tienen políticas a un alto nivel, abarcando observancia incluso transnacional en algunos casos, para garantizar la transparencia y el acceso a la información, pero también se tienen políticas a nivel administrativo, abarcando los aspectos del entorno, contenido, implementación, evaluación y retroalimentación de los datos abiertos (Charalabidis et al., 2018).

Los datos abiertos tienen varias características que los identifican, pero un aspecto relevante de los datos son las plataformas y entornos virtuales que se encuentran entre quienes proveen los datos y quienes los consumen (Charalabidis et al., 2018). Estas plataformas pueden ser el resultado de diferentes modelos de datos abiertos que aplican a diferentes contextos, dependiendo de la manera que se requiera manejar los datos (por ejemplo: crear, organizar, publicar, usar, realimentar) (Charalabidis et al., 2018).

En los títulos siguientes se describen los aspectos generales tanto de los datos abiertos, como de las plataformas y portales para su difusión.

## **1.2 Origen de los datos abiertos**

Para conocer un poco de la historia de los datos abiertos, de su origen, y ubicarnos mejor en el entorno, se tiene la publicación realizada por el Paris Innovation Review, citada en el sitio [www.data.gov](http://www.data.gov). En la publicación Chignard (2013), menciona que, aunque el término de datos abiertos no tiene ni 20 años y que es relativamente nuevo, fue desde los años 1940 que Robert King Merton tuvo la idea que los datos de investigaciones científicas fueran compartidos de manera libre para el bien común. A principios del siglo XXI los científicos coinciden con la floreciente sociedad de la Tecnología de la Información/Internet; para ese tiempo, tanto los científicos como el público en general asumen que todo el conocimiento científico deber ser

para el bien público, entrando en debates por la propiedad intelectual (Chignard, 2013).

El término datos abiertos apareció por primera vez en un documento de una agencia científica americana, en donde promovían el libre intercambio de información entre los países para el entendimiento y análisis de un fenómeno global que involucra la atmósfera, los océanos y la biósfera (Chignard, 2013). En el 2007, en Sebastopol, California, treinta activistas de Internet se reunieron para definir el concepto de datos públicos abiertos y para que fuera adoptado por los candidatos presidenciales de los Estados Unidos de América, entre los que destacaban Tim O'Reilly y Lawrence Lessig (Chignard, 2013). Chignard continúa mencionando que parecía un sueño en ese entonces, pero en 2009 cuando el presidente Barack Obama tomó la oficina de la Casa Blanca firmó tres memorándums presidenciales, dos concernientes a gobierno abierto, de la cual los datos abiertos es uno de los pilares, indicando:

“La información mantenida por el Gobierno Federal es un activo nacional. Mi Administración tomará las medidas apropiadas, consistentes con la ley y la política, para divulgar información rápidamente en formularios que el público pueda encontrar y usar fácilmente” (Obama, 2009, p. 1).

El nacimiento del sitio web data.gov durante la administración de Obama, dio la patada inicial en la compartición de datos gubernamentales e inspiró la búsqueda de información gratuita (GODAN, 2017).

### **1.3 Concepto y características de los datos abiertos**

Los conceptos de *cloud computing*, *smart cities*, *big data*, *Internet of things*, *digital security*, *industrial Internet* y *open data*, tienen una finalidad muy específica y pueden tener un uso con diferente orientación, pero están ligados de manera muy estrecha y en ocasiones tener alguna dependencia uno del otro. En todos ellos hay una gran dependencia hacia la infraestructura de las Tecnologías de la Información

y Comunicación (TIC) –enlaces de comunicación, servidores de aplicaciones y de datos- y tienen un gran requerimiento de inversión, de personal capacitado y de una brecha digital estrecha.<sup>7</sup>

Aún con su inevitable dependencia hacia el hardware, las infraestructuras de datos abiertos pueden tener requerimientos funcionales enfocados mayormente en el software (Charalabidis et al., 2018) y pueden implementarse en los equipos disponibles en la mayoría de los gobiernos. La complejidad de su implementación parece ser más un aspecto que tiene que ver con la adopción de estándares, la definición de procesos, la capacitación del personal, la generación de inventarios de datos y con el desarrollo de plataformas de software que faciliten la generación de archivos y su publicación (México Digital, 2014).

*The World Bank* señala que “Los datos son considerados abiertos si cualquiera puede accederlos libremente, usarlos, re-usarlos y distribuirlos, para cualquier propósito, sin restricciones” (2019, p.1).

Para crear datos abiertos, primero se debe disponer de datos y es por ello que los ámbitos de los datos abiertos y de *big data* tienen áreas comunes, teniendo una cadena de valor de los datos que va desde los datos crudos hasta los servicios agregados, transitando por validación de datos, agregación de datos, análisis de datos, productos y servicios derivados de los datos (Carrara et al., 2015). Tanto los servicios agregados como los productos y servicios son los que finalmente representan un mayor impacto económico en las sociedades que tienen acceso a los datos abiertos.

Entrando en materia de las características de los datos abiertos, estos son identificados como tales cuando cumplen con las siguientes características:

- “Gratuitos. - Se obtienen sin entregar a cambio contraprestación alguna.
- No discriminatorios. - Son accesibles sin restricciones para el usuario.
- De libre uso. - Citar la fuente de origen es el único requerimiento para ser usados libremente.

---

<sup>7</sup> La brecha digital se define como “la línea divisoria entre el grupo de población que ya tiene la posibilidad de beneficiarse de las TIC y el grupo que aún es incapaz de hacerlo” (CEPAL, 2003, p. 7)

- Legibles por máquinas. - Están estructurados, total o parcialmente, para ser procesados e interpretados por equipos electrónicos de manera automática.
- Integrales. - Contienen, en la medida de lo posible, el tema que describen a detalle y con los metadatos necesarios.
- Primarios. - Proviene de la fuente de origen con el máximo nivel de degradación posible.
- Oportunos. - Son actualizados periódicamente conforme se generan.
- Permanentes. - Se conservan en el tiempo; es decir, las versiones históricas relevantes para uso público se mantienen disponibles a través de identificadores adecuados para tal efecto” (Gobierno de la República, 2015, p. 1).

Cada una de estas características tiene una explicación sencilla y bastante lógica sobre su significado y las implicaciones de ello para el término de datos abiertos. En especial el hecho de que sean primarios y que sean procesables por las computadoras, lo que facilita que cualquier persona física o moral pueda tomar esos datos para explotarlos a través de aplicaciones desarrolladas por ellos mismos o por terceros para agregarle valor a los mismos datos, pero desde luego que cada una de las características mencionadas son las que aportan algo de valor adicional.

En la literatura y en el ciberespacio podemos encontrar algunas variaciones en cuanto a sus características, pero coinciden de manera esencial, como en el documento de la Comisión Económica para América Latina y el Caribe (CEPAL), que menciona las siguientes características:

- “Están completos.
- Son primarios y no procesados.
- Son oportunos y se entregan lo antes posible.
- Son utilizables para cualquier propósito.
- Son procesables por máquina.
- El acceso es no discriminatorio, abierto y anónimo.
- Están en un formato no propietario.
- Están libres de licencias, patentes, marcas, etc.” (CEPAL, 2014, p. 19).

Los datos abiertos pueden confundirse fácilmente con el concepto de transparencia en su orientación regulatoria del gobierno, referente a la información financiera. Principalmente relacionada con la rendición de cuentas de la función pública para propiciar la toma de decisiones públicas (Congreso del Estado de Aguascalientes, 2016), la transparencia también existe como transparencia regulatoria y como derecho a la información gubernamental (Kosack & Fung, 2014). En este último tipo, la Comisión Económica para América Latina y el Caribe menciona que:

“los tres actores involucrados en la apertura y el empleo de este tipo de datos son el gobierno, que abre y publica los datos en los formatos necesarios para la reutilización; desarrolladores de aplicaciones y; usuarios, que se benefician de una mayor transparencia en la administración pública, algo que puede facilitar su participación en este proceso y ayudar a empoderarlos” (CEPAL, 2014, p. 20).

En este enfoque, los usuarios de los datos abiertos son empoderados y con ello pueden obtener un beneficio del uso, consiguiendo el valor otorgado por las aplicaciones generadas por los desarrolladores. Estos tres actores se vinculan por medio de la infraestructura que soporta la plataforma de datos abiertos. En esta perspectiva, la aportación relevante del gobierno es la apertura y el empleo de datos administrativos para generar y publicar datos abiertos; por parte de los desarrolladores, se aporta la capacidad técnica de generar aplicaciones, con el dominio de algún lenguaje de programación y del marco aplicable; y por parte del usuario, el acceso a los datos suministrados por el gobierno, contando con un acceso a Internet, a un equipo de cómputo, a las aplicaciones suministradas por los desarrolladores y el adiestramiento necesario para conjugar los elementos mencionados. En este estudio, nos centramos en el primer elemento de los tres mencionados: el gobierno, con el empleo de alguna plataforma de generación y publicación de datos.

Los datos abiertos pueden ser presentados, con algún grado de madurez, a través de portales. En el mismo documento desarrollado por la Comisión Económica para América Latina y el Caribe, se menciona que Tim Berners-Lee, a través de la World Wide Web Consortium, ha desarrollado un sistema de medida para calificar la manera en que las bases de datos han sido calificadas como abiertas:

- **Sin estrella:** no se puede acceder a los datos con una licencia abierta, incluso si están en línea.
- **Una estrella:** los datos están accesibles en la Web y son legibles por el ojo humano, pero no por un agente de software, ya que están en un formato cerrado y no se pueden reutilizar fácilmente.
- **Dos estrellas:** se puede acceder a los datos en la Web en un formato estructurado y legible por máquina. Se pueden procesar, exportar y publicar con facilidad, pero los reutilizadores aún deben confiar en software propietario como Microsoft Word o Excel.
- **Tres estrellas:** la reutilización de datos no depende del software propietario, ya que se usa el formato de valores separados por comas (CSV) en lugar de Excel.
- **Cuatro estrellas:** los datos están dentro de la Web en lugar de sobre la Web; el uso de un identificador uniforme de recursos (URI) proporciona un control fino y granular de los datos, por ejemplo, mediante marcadores y enlaces.
- **Cinco estrellas:** los datos no solo están en la Web, sino que también están vinculados a otros datos, aprovechando al máximo los efectos de la red. Esta interconexión coloca los datos en contexto y significa que pueden ser descubiertos por otras fuentes, como enlaces a Wikipedia (CEPAL, 2014, p. 20).

Con la disponibilidad de estas bases de datos en alguno de estos niveles de madurez, existen beneficios directos e indirectos como consecuencia de los datos abiertos en sus distintas etapas de la cadena de valor de los datos (RDA Europe, 2014), como el hecho de realizar transacciones de mercado en la forma de ingresos

y valor agregado bruto, generación de empleos, ahorro en costos, nuevos bienes y servicios, incremento en la eficiencia de los servicios públicos, transparencia, participación cívica, conciencia política y acceso a la información (Carrara et al., 2015).

#### **1.4 Riesgos y su mitigación en la publicación de datos abiertos**

Como señala la CEPAL (2014), los tres actores involucrados en la apertura y el empleo de datos abiertos son el gobierno, los desarrolladores de aplicaciones y usuarios. Siguiendo esta línea, enfocados en el lado de gobierno como publicador de datos, es fundamental seguir la recomendación de mantenerlo de manera simple con el mayor uso y mayor impacto posibles (Open Knowledge Network, 2017). Así como existen beneficios en la publicación de datos abiertos, también existen riesgos y dificultades al abrir los datos por aspectos como:

- Publicación de datos contra la ley.
- Infracción de secreto comercial.
- Infracción a la privacidad.
- Riesgo a la seguridad de la infraestructura.
- Publicación de datos e información incorrecta.
- Publicación de datos inexactos.
- Malinterpretación de los datos.
- Ausencia de consumidores de datos.
- Sujetos con menos voluntad de cooperar.
- Superposición de datos.
- Número incrementado de requerimientos de datos (Kucera & Chlapek, 2014).

Aunque no de manera fácil, se pueden realizar algunas acciones para mitigar estos riesgos considerando:

- Monitoreo y evaluación de la demanda de los datos.
- Correcta especificación de los conjuntos de datos.

- Evaluación del cumplimiento.
- Armonización de los datos.
- Control de calidad de los datos a ser publicados.
- Catálogo de datos internos.
- Catálogo de datos externos.
- Enlace a datos ya publicados.
- Términos de uso de los datos debidamente formulados.
- Deberes claramente explicados.
- Continuo monitoreo de los impactos de las iniciativas de datos abiertos (Kucera & Chlapek, 2014).

Todo lo anterior se centra principalmente en aspectos de procesos de manejo de los datos y en el contenido de los datos mismos.

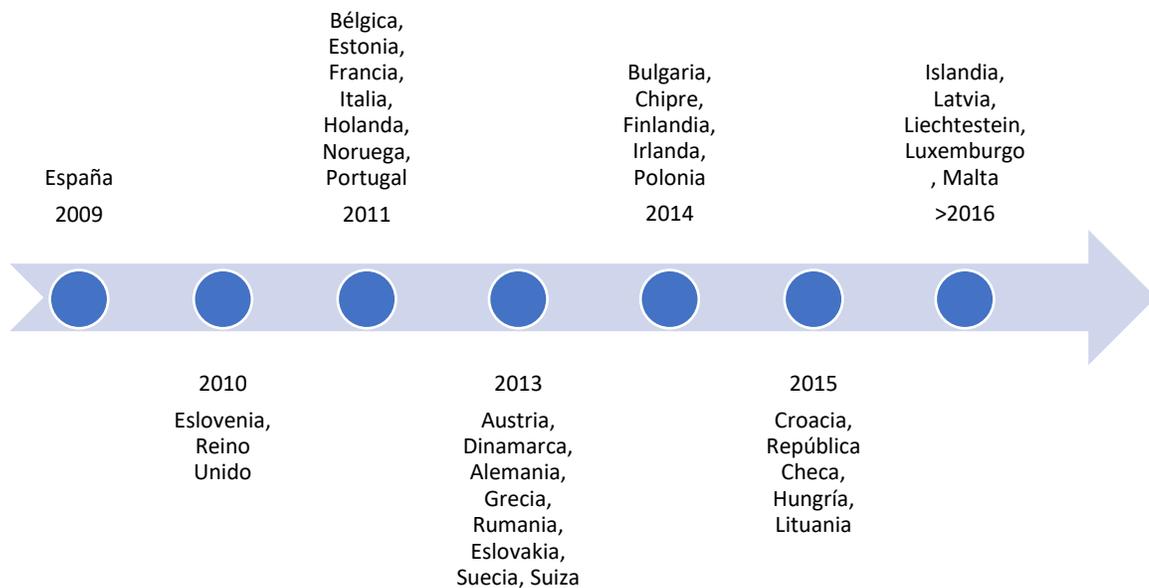
## 1.5 Entorno internacional

A nivel internacional, derivado del punto inicial en los Estados Unidos de América con su plataforma data.gov, se cuenta con múltiples plataformas y organizaciones que, entre otras funciones, se dedican a impulsar y medir el avance de los países en el uso de los datos abiertos. Entre ellas destacan Open Data 500, Open Data Initiative, Open Data Barometer, Portal Europeo de Datos y EU Open Data Portal<sup>8</sup>.

Es de resaltar que de acuerdo a Open Data Barometer, los Estados Unidos de América cuentan con una puntuación de 64 de 100 posibles en esa evaluación que contempla aspectos de preparación, implementación e impacto emergente (World Wide Web Foundation, 2019). Reino Unido con 76 puntos, se mantiene con la mayor puntuación a nivel internacional, empatado con Canadá que ha tenido una tendencia a la alza y subiendo 18 puntos (World Wide Web Foundation, 2019).

---

<sup>8</sup> Los sitios *Web* de estas plataformas las podemos encontrar en las direcciones: <http://www.opendata500.com/>; <http://opendatatoolkit.worldbank.org/en/starting.html>; [https://opendatabarometer.org/?\\_year=2016&indicator=ODB](https://opendatabarometer.org/?_year=2016&indicator=ODB); <https://www.europeandataportal.eu> y en <https://data.europa.eu/euodp/data/>



**Figura 1. Creación de portales de datos abiertos en Europa**  
 Fuente: Creación propia con base en Carrara et al. (2015).

En el 2011 el Diario Oficial de la Unión Europea emitió la decisión de la comisión 2011/833/UE que establece un portal regional para reutilizar documentos en formatos legibles por máquina (Comisión Europea, 2011).

Pese a que el documento de decisión de la comisión se emitió en 2011, desde el 2009 España generó su plataforma de datos abiertos, siendo la primera en la Unión Europea (Carrara et al., 2015). Al respecto, tenemos línea de tiempo de la Figura 1.

Algunos de los beneficios del establecimiento de plataformas de datos abiertos en la Unión Europea es que se tiene mayor acceso a la información, lo que a la vez estimula el desarrollo de servicios y modelos de negocio innovadores (Carrara et al., 2015). En esa misma región, se espera que entre 2016 y 2020 se generen 25,000 empleos relacionados directamente con los datos abiertos; se espera también una disminución de 5.5% en fatalidades en los caminos, un ahorro de 1.7 billones de euros para la administración pública, etc. (Carrara et al., 2015).

En Australia, se tienen evidencias de casos de éxito del uso de los datos abiertos, como el ejemplo de AusGOAL, una organización de datos abiertos del

sector público que ha beneficiado a los sectores agrícola y de nutrición (GODAN, 2017).

En África, de igual manera, se cuenta con el ejemplo de Open Data Kenya, que permitió al país en ser el primero de ese continente en lanzar una iniciativa de datos abiertos, con un portal en línea de 160 conjunto de datos, y que se espera que ofrezca mayor transparencia y una mejor interacción entre el gobierno y los ciudadanos (Rahemtulla et al., 2011).

## **1.6 Datos abiertos en América Latina y el Caribe**

En América Latina existe una gran diversidad entre los conjuntos de datos disponibles de cada país; además no todos los países cuentan con una plataforma de datos abiertos, y hay algunos, como El Salvador, que cuentan con un portal nacional pero que es administrado por voluntarios y no por el gobierno (Steinberg & Castro, 2017).

En 2017, se estimó que países como México, Colombia y Brasil contaban con alrededor de 20,000, 6,000 y 2,500 conjuntos de datos respectivamente, mientras que Argentina, El Salvador y Panamá listaban menos de 100 conjuntos de datos (Steinberg & Castro, 2017). Los portales de Bolivia, Colombia, México y Uruguay muestran cuáles son los conjuntos de datos más descargados y los más visitados, mientras que otros portales como los de Argentina, Brasil, Chile, Costa Rica, Ecuador, El Salvador, Panamá, Paraguay y Perú no proveen esa información debido a que, o usan plataformas de software desactualizadas o porque no se implementan bien las versiones actuales (Steinberg & Castro, 2017), lo que resalta la importancia de las plataformas de TIC requeridas para la generación y publicación de datos abiertos.

En el 2015, se estrenó en México, durante la cumbre de la Alianza para el Gobierno Abierto, la Carta Internacional de Datos Abiertos, la cual es una iniciativa multilateral y colaborativa para implementar los principios de los datos abiertos y establecer estándares y buenas prácticas en la materia a nivel mundial (Naser & Rosales, 2016).

Aunque el tema de plataformas se desarrollará un poco más adelante, en el cuadro 2 visualizamos las direcciones *Web* de los portales de datos abiertos de algunos países de América Latina, indicando el tipo de plataforma que utiliza, así como conjuntos de datos y Apps disponibles en el 2016. En ella se distinguen México, Chile y Colombia en conjuntos de datos, además Brasil, Colombia y Perú destacan en la generación de aplicaciones. En el cuadro 2 también sobresale que la plataforma más utilizada en los países listados es CKAN.

País	Portal	Plataforma	Datasets	Apps	Apps Cívicas más descargadas
Argentina	<a href="http://www.datos.gob.ar">http://www.datos.gob.ar</a>	CKAN	4	0	BA Cómo llego, BA Móvil y BA EcoBici
Brasil	<a href="http://dados.gov.br/">http://dados.gov.br/</a>	CKAN	1124	121	Basometro, Para onde foi o meu dinheiro, y Radar parlamentar
Chile	<a href="http://datos.gob.cl/">http://datos.gob.cl/</a>	CKAN	2427	29	iTransantiago, iBencinas, iFarmacias
Colombia	<a href="http://www.datos.gov.co">http://www.datos.gov.co</a>	CKAN	2108	94	Verifíquese cédula, Autocuidate y Colombia. Si Vigila
Costa Rica	<a href="http://datosabiertos.gob.go.cr">http://datosabiertos.gob.go.cr</a>	JUNAR	1998	5	Kolbi, Seguridad Pública y Touit Biodiversidad
Ecuador	<a href="http://catalogo.datosabiertos.gob.ec">http://catalogo.datosabiertos.gob.ec</a>	CKAN	63		Puntos Licencia Ecuador, Quinto en tu mano y Guía Quito
Guatemala	<a href="http://datos.transparencia.gob.gt">http://datos.transparencia.gob.gt</a>	LIBRE	100		Elecciones 2015 y Multas GT
México	<a href="http://datos.gob.mx">http://datos.gob.mx</a>	LIBRE	2758	36	Metro DF, Guía PEMEX y Mi Policía
Paraguay	<a href="http://www.datos.gov.py">http://www.datos.gov.py</a>	CKAN	114	6	Fonacide, Contrataciones PY y Buscador de Funcionarios públicos.
Perú	<a href="http://www.datosabiertos.gob.pe/">http://www.datosabiertos.gob.pe/</a>		130	57	SAT, Congreso del Perú y Ministerio de Justicia y Derechos Humanos
República Dominicana	<a href="http://datos.gob.do/">http://datos.gob.do/</a>	CKAN	256		Combustible RD, Multas RD, y Aduanas RD
Uruguay	<a href="https://catalogodatos.gub.uy/">https://catalogodatos.gub.uy/</a>	CKAN	123	31	Atuservicio.uy, PorMiBarrio.uy y QuéSabés.uy

**Cuadro 2. Portales de datos abiertos en América Latina**  
Fuente: tomado de (Naser & Rosales, 2016, p. 20)

## 1.7 Datos Abiertos en México

Actualmente México es el mejor posicionado de los países de América Latina y el Caribe con su plataforma [datos.gob.mx](http://datos.gob.mx), y con la posición número seis supera incluso a países del primer mundo como Estados Unidos de América y Alemania (World Wide Web Foundation, 2019). Esto se debe a los altos niveles de preparación, implementación e impacto emergente de la mencionada plataforma y hasta la última evaluación ha tenido una tendencia a la alza en cuanto a la puntuación final (World Wide Web Foundation, 2019).

Para la conformación de esta plataforma, en el año 2015 el presidente Enrique Peña Nieto firmó el *Decreto por el que se establece la regulación en materia de Datos Abiertos* (Presidencia de la República, 2015) .<sup>9</sup>

En el mismo decreto se establece que el modelo de datos abiertos facilitará la publicación de dichos datos por parte de las dependencias y entidades de la Administración Pública Federal, así como de las empresas productivas del Estado (Presidencia de la República, 2015).

El decreto establece también que el portal [www.datos.gob.mx](http://www.datos.gob.mx) estaría a cargo de la Unidad de Gobierno Digital (UGD) y en colaboración de la Coordinación de la Estrategia Digital Nacional (CEDN). En el portal se establece el inventario único y catálogo de datos abiertos puestos a disposición de la población por parte de la Administración Pública Federal. Además la UGD genera con base en sus atribuciones una guía de implementación de la política de datos abiertos en los gobiernos estatales (Presidencia de la República, 2015).

Como una manera de calificar la disponibilidad de datos abiertos en los Estados del país, la revista digital U-GOB realiza anualmente una evaluación que se enfoca en los portales estatales de transparencia. Esta evaluación contiene el aspecto de datos abiertos dentro de cinco componentes clave de la transparencia y en la edición del 2018 lamentablemente se observa que ha habido un retroceso en el promedio de los portales estatales en este componente al registrar 38.88 puntos,

---

<sup>9</sup> El Decreto hace referencia y se sustenta en diferentes documentos normativos que incluyen la Constitución Política de los Estados Unidos Mexicanos y la Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental, entre otros originados en la administración de Enrique Peña Nieto.

8.62 puntos menos que el año 2017 (Sandoval, 2018). Con las puntuaciones más altas en datos abiertos se ubican Jalisco, Guanajuato y Estado de México, mientras que San Luis Potosí y Yucatán se ubican en los últimos lugares (Sandoval, 2018). Lo anterior demuestra una crisis en materia de datos abiertos hacia los estados del país, pues no siguen la tendencia que el gobierno federal ha ofrecido y muestran opacidad en los datos de transacciones económicas y evidencian prácticas de corrupción (Sandoval, 2018).

Pese a todo el avance en los índices internacionales de datos abiertos, las políticas de gobierno abierto no facilitan por sí solas el acceso o la utilización de los datos abiertos por parte de los ciudadanos, para lo cual se necesita ejecutar planes de acción y políticas públicas específicas de datos abiertos; se debe centrar la gestión en el ciudadano, mejorar las reformas de relación y comunicación con la ciudadanía y facilitar la participación ciudadana (Vera Martínez, 2018). Sin embargo, este trabajo se centra en el aspecto de publicación de los datos y no contempla el aseguramiento del acceso y uso de los datos publicados por parte de la ciudadanía, debido al alcance establecido por los recursos con los que se cuenta. Aun así, dentro de la temática de plataformas de datos abiertos que se expone a continuación, en el título de modelos de datos se vislumbra brevemente la relación entre la publicación y el uso de los datos abiertos.

## **1.8 Aspectos críticos para la publicación de datos abiertos**

De acuerdo a documentos consultados, realizados a partir del estudio, impacto y beneficios de iniciativas de datos abiertos a nivel local, regional y transnacional, existen una serie de aspectos críticos que se deben tomar en cuenta para una nueva iniciativa, los cuales son:

- a) **Legislación, regulación y licencias.** Tener un marco para la publicación, tomando en cuenta la legislación sobre protección de datos y privacidad, así como el establecimiento de penalidades por el incumplimiento.
- b) **Estrategia y apoyo político.** Desarrollar una estrategia en la que participe la alta gerencia, con apoyo a los generadores de políticas de publicación de datos abiertos y a los líderes de los grupos en los que estas políticas se

implementarán; generar una compatibilidad entre las políticas y el marco organizacional.

- c) **Apoyo a la gestión y procesos de publicación dentro de las agencias gubernamentales.** Definir el proceso para la publicación de datos identificando los conjuntos de datos a publicar (con base en su valor potencial de uso y beneficio), los datos que no se publicarán, los estándares y el personal involucrado.
- d) **Formación y apoyo a los funcionarios.** Capacitar a los funcionarios en la publicación de datos abiertos, así como el establecimiento de un centro virtual para proveer ayuda y desarrollar campañas de difusión.
- e) **Evaluación de la iniciativa de datos abiertos.** Definir métricas y evaluarlas.
- f) **Sustentabilidad de la iniciativa de datos abiertos.** Crear una estrategia para mantener publicados los conjuntos de datos, asegurando la continuidad en la provisión y transparencia en las condiciones en las cuales la publicación de datos se desarrolla.
- g) **Colaboración.** Realizar reuniones internas para discutir y evaluar el proceso de publicación, así como reuniones interorganizacionales para la gestión y aprendizaje de otras iniciativas.
- h) **Plataformas de datos abiertos, herramientas y servicios.** Integrar la plataforma en el sistema de gestión de contenido para acelerar el proceso, implementando una interfaz clara con funcionalidades de búsqueda, limpieza, visualización y análisis.
- i) **Accesibilidad, interoperabilidad y estándares.** Utilizar estándares para los datos, metadatos, licencias y protocolos de intercambio y direccionamiento; integrar esquemas de metadatos, con varios tipos de metadatos y multilinguaje; implementar API's (Interfaz de Programación de Aplicación) (Parycek, Zuidoewijk, Charalabidis, & Janssen, 2015).

Cada uno de los puntos listados puede tener un mayor o menor impacto, dependiendo del nivel de la iniciativa: local, regional o transnacional. En nuestro caso, en el nivel regional, se destaca un estudio para la iniciativa particular de *New*

York, el cual establece que los puntos relevantes en este tipo de plataformas son los identificados como a) legislación, regulación y licencias; d) formación y apoyo a los funcionarios y h) plataformas de datos abiertos, herramientas y servicios (Parycek et al., 2015).

## **1.9 Plataformas de datos abiertos**

Aún antes de la creación de plataformas especializadas en datos abiertos, los gobiernos han establecido plataformas digitales y aplicaciones para presentar información de una manera amigable y con calidad (Sandoval-Almazán et al., 2017). Aunque la recomendación es contar con portales para la publicación de datos abiertos (Munte-Kunigami & Serale, 2018), en la literatura podemos encontrar la equivalencia entre una plataforma (*web*) y un portal (*web*), haciendo referencia a uno u otro de manera sinónima, como en *los datos abiertos en América Latina y el Caribe* de Munte-Kunigami & Serale (2018), y en *evaluación sobre la disponibilidad de datos abiertos en el Estado de Jalisco* de Grupo Banco Mundial (2015). También tenemos que, aunque la palabra portal puede tener varios significados, en Internet “un portal es una plataforma basada en web que recopila información de diferentes fuentes en una única interfaz de usuario y presenta a los usuarios la información más relevante para su contexto” (Liferay, 2019, párr. 1). En esta definición, se identifican dos componentes importantes como parte de los objetivos de una plataforma de datos abiertos: en primer lugar, se identifican las acciones encaminadas hacia la publicación de datos por parte de las fuentes; en segundo lugar, se identifican las acciones encaminadas hacia la explotación de datos por parte del usuario.

Adicional a los conceptos de plataforma y portal web, se tiene el enfoque de un ‘ecosistema’, en el que en lugar de limitar el uso de los datos en un solo sentido, es decir del publicador hacia el usuario, este hace un uso y mejoramiento de los datos para volverlos a poner disponibles en la plataforma con un mayor valor (Pollock, 2011). El enfoque de ecosistema, implica un estudio adicional focalizado en los usuarios, lo que no es la intención de este trabajo.

Para llevar a cabo la apertura del gobierno se requiere la construcción de plataformas para publicar datos que puedan ser accedidos por los diferentes interesados, y es por ello que diferentes gobiernos las han desarrollado para exponer sus datos públicos (Estrada et al., 2013). Tomando en cuenta los componentes de gobierno abierto mencionados por Council of Europe (2018), en la literatura y en la *World Wide Web* se pueden encontrar guías para establecer plataformas orientadas a la publicación de datos abiertos, desde el aspecto meramente informativo hasta la promoción de la colaboración entre el gobierno y las organizaciones de la sociedad civil (World Wide Web Foundation, 2016).

Con la intención de facilitar la identificación de algunos de los conceptos involucrados en la base sustantiva de las plataformas de publicación de datos, se mencionan a continuación elementos relacionados directamente con la manera en que se generan y se publican los datos, para su potencial uso por parte de la población.

### **1.9.1 Modelos de datos**

Como ya se mencionó en el apartado de *claves de gobierno abierto*, las plataformas pueden ser el resultado de diferentes modelos de datos abiertos que aplican a diferentes contextos, dependiendo de la manera que se requiera manejarlos (Charalabidis et al., 2018). Estos modelos los podemos apreciar en diferentes terminologías como lo son los conceptos de **ciclo de vida de datos abiertos**, la **cadena de valor de los datos** y los **procesos de datos**, los cuales tienen diferentes propósitos: estructurar el manejo de los datos en sí, la creación de valor por el uso de datos abiertos, y el enfoque en los procesos administrativos, respectivamente (Charalabidis et al., 2018).

#### **1.9.1.1 Ciclo de vida de los datos abiertos**

De acuerdo a Moller (2013), los datos son creados, exportados, publicados, importados, usados, transformados y reusados. Bajo el concepto de ciclo de vida de datos, cuyo enfoque se centra en las prácticas del manejo de datos desde la creación hasta su uso, se identifican dos ciclos con fases que transitan de manera

paralela entre quien publica y quien explota o hace uso de los datos: crear/adquirir; pre-procesar/procesar; curar/re-usar; almacenar/colaborar; publicar/retroalimentar (Charalabidis et al., 2018). Estas dos fases se pueden visualizar claramente en la figura 2, en la cual se ubican las acciones correspondientes al proveedor de datos en la parte interna del circuito y a las acciones de los usuarios en la parte externa de la figura.



**Figura 2.** Ciclo de vida de datos abiertos.  
Fuentes: Adaptado de Charalabidis et al. (2018, p. 20).

Para completar cada una de estas fases, se tienen una serie de herramientas y métodos (Charalabidis et al., 2018) que son dignos de un estudio detallado para cada uno de ellos, pero que en este trabajo no se profundizarán. Estas herramientas y métodos se sintetizan en el cuadro 3, indicando tanto las herramientas como los métodos que pueden ser usados para completar las acciones correspondientes a cada una de las fases conceptualizadas en la figura 2.

Fase del ciclo de datos	Herramientas	Métodos
<b>Crear/Adquirir:</b> El proceso de crear datos	Sensores; IoT; SI; Humanos; conexión con otros datos abiertos; <i>Hadoop</i> para <i>big data</i>	Creación de datos automatizada; captura manual; Vínculos con portales de datos abiertos
<b>Pre-proceso:</b> El proceso gerencial de definición de calidad de datos.	Normas detalladas de metadatos; métricas y modelos de evaluación;	Conceptualización y establecimiento de objetivos; plan de evaluación y calidad

	Identificación única (URIs y URLs)	de datos; esquema de metadatos de 3 capas para portales.
<b>Curado:</b> El proceso de cumplir la calidad de datos requerida y los requerimientos legales.	Herramienta externa de refinar datos abiertos vinculados; herramientas nativas e individuales.	Estructuración; anonimización; refinamiento de metadatos; cambio de formato de datos; limpieza de datos.
<b>Almacenamiento/obtención:</b> El proceso de toma de decisiones de almacenamiento.	Centros de datos; SPARQL repositorios para datos vinculados; bases de datos para <i>big data</i> ; enlaces con otros conjuntos de datos.	Versionador; <i>Data Linking</i> ; bases de datos orientadas a columnas para <i>big data</i> .
<b>Publicar:</b> El proceso cubriendo elementos legales.	Capacidad de subida.	Plan de publicación; Licenciamiento de acceso abierto; Derechos de propiedad intelectual.
<b>Recuperar/Adquirir:</b> El proceso de adquisición de datos a través de portales.	Portales de Datos Abiertos (Por ejemplo, Portal Europeo de Datos.	Técnicas de búsqueda políglota; API's
<b>Procesar:</b> El proceso de análisis de datos.	Herramientas externas de procesamiento de datos; <i>Open Refine</i> ; <i>Rapidminer</i> ; Excel;	Enriquecimiento de datos; Crear datos abiertos enlazados; diferente combinación de conjunto de datos; minería; Hashing
<b>Uso:</b> El proceso de presentar el resultado del análisis.	Herramientas de visualización internas y externas; paquetes estadísticos; enlace con artefactos externos.	Análisis estadístico; visualización de mapa; visualización de gráfica; analíticos visuales; diagrama de grupo.
<b>Colaborar:</b> El proceso de comunicar con otros usuarios de datos.	Espacio de colaboración y flujo de trabajo; capacidades y herramientas de Web 2.0.	Intercambio de notas / correo electrónico / ideas; crear grupos de intereses comunes.
<b>Retroalimentación:</b> El proceso de evaluar y proveer retroalimentación a los proveedores de datos	Capacidades y herramientas de Web 2.0.	Rating de calidad de datos; evaluación de la publicación.

**Cuadro 3. Métodos y herramientas en cada fase del ciclo de vida de los datos abiertos.**

Fuente: Tomada de Charalabidis et al. (2018, p. 21).

### 1.9.1.2 Cadena de valor de los datos

Bajo el concepto de cadena de valor de los datos, es importante entender los diferentes tipos de re-uso de los datos abiertos. En este modelo, los datos son validados cuando son creados o recolectados; después pueden ser analizados

cuando son liberados a través de un portal, haciendo posible la combinación de varias fuentes de información o hacer visualizaciones, al tiempo que durante todo el proceso los datos son almacenados y preservados (Carrara et al., 2015). Este modelo puede ser representado como se ilustra en la figura 3. En el modelo, tanto los sectores públicos como privados crean bienes y servicios agregados para el usuario ‘final’, en los que pueden estar involucradas transacciones de mercado, generación de empleos, ahorro en costos, nuevos bienes y servicios para un sector, así como incremento en la eficiencia de los servicios públicos, transparencia, participación cívica, conciencia política y acceso a la información para el otro (Carrara et al., 2015). La validación y constancia de estos elementos queda fuera del alcance el presente trabajo.



**Figura 3.** Cadena de valor de los datos abiertos.  
Fuente: Tomado de Carrara et al. (2015, p. 28)

### 1.9.1.3 Procesos de datos

Bajo este concepto, se tiene el enfoque en los procesos administrativos, atendiendo los retos como la inversión de las organizaciones gubernamentales en recursos, tiempo y esfuerzo en la apertura de datos para el beneficio de otros (Charalabidis et al., 2018). Los responsables de la formulación de políticas de las organizaciones pueden utilizar la información que los usuarios obtuvieron del análisis de los datos

gubernamentales, pero a menudo ambas partes no se comunican sobre las lecciones que pudieran ser aprendidas (Zuiderwijk, 2015).

La implementación de alguno de los tres modelos expuestos depende en gran medida de los objetivos que se quieran perseguir y de los recursos con los que se cuente para alcanzarlos. Un punto importante en este sentido es que adicionalmente a los riesgos relacionados con la publicación de datos expuestos por Kucera & Chlapek (2014), se puede tener la problemática técnica de una falta de infraestructura y recursos adecuados, lo cual es necesario para la apertura de datos (Charalabidis et al., 2018).

Una infraestructura de datos abiertos gubernamentales se puede definir como “un sistema compartido, (cuasi) público, en evolución, que consiste en una colección de elementos sociales interconectados (por ejemplo, operaciones de usuario) y elementos técnicos (por ejemplo, herramientas y tecnologías de análisis de datos abiertos, servicios de datos abiertos) que permiten el uso conjunto de datos abiertos gubernamentales” (Zuiderwijk, 2015, p. 45). Todos los elementos deben de funcionar de manera adecuada, ya que el mal funcionamiento de uno de ellos puede afectar a la infraestructura completa, por ejemplo, que las infraestructuras estén más orientadas a la apertura de los datos que al uso de los mismos (Charalabidis et al., 2018).

En la implementación de una plataforma de datos abiertos, se deben tomar en cuenta tanto los requerimientos no funcionales como los funcionales, los cuales se explican a continuación.

### **1.9.2 Requerimientos no funcionales para una infraestructura de datos abiertos gubernamentales**

Los llamados requerimientos no funcionales son aquellos que no se refieren directamente con las funciones específicas de un sistema sino a las propiedades del mismo, como fiabilidad, tiempos de respuesta, capacidad de almacenamiento, con los que se debe contar por defecto y que surgen de las restricciones en el presupuesto, a las políticas de la organización, a la necesidad de interoperabilidad con otros sistemas, etc. (Innova-T, 2019b).

Estos requerimientos no funcionales contemplan aspectos como los de seguridad, estándares, las restricciones de diseño e implementación, los costos, normatividad, organización, infraestructura, presupuesto, entre otros. Estos aspectos, deberían emerger como resultado del análisis de requerimientos, ya que esta etapa permite detectar la presencia de áreas no especificadas (Báez & Barba-Brunner, 2001).

Parte de estos requerimientos se contemplan en la sección de aspectos técnicos que se desarrolla enseguida.

### **1.9.2.1 Aspectos técnicos**

Los conocimientos y capacidad del personal responsable, tanto de los aspectos técnicos como del uso y explotación de las características de las plataformas, es un aspecto muy importante a tomar en cuenta en la elección de una u otra plataforma. En principio, las plataformas o portales de datos abiertos se encuentran como un elemento TIC del e-gobierno, el cual tiene al conocimiento como uno de sus elementos más importantes (Purón, 2013).

En el mismo contexto técnico, se cuenta con los catálogos de datos, que no es más que el conjunto de datos o *datasets* disponibles en una iniciativa de datos abiertos, constituidos como la pieza central y que pueden ser establecidos de una manera simple o de manera compleja en diferentes modelos (The World Bank, 2019). Sin importar su complejidad, los catálogos de datos comparten unas características en común:

- **Facilidad de acceso.** - Sin requerir autenticación, los usuarios deben de acceder a los datos de manera fácil, rápida e intuitiva.
- **Búsqueda.** - Los datos deben ser encontrados fácilmente, usando un motor de búsqueda y organizados por tema u organización.
- **Acceso a datos legibles por máquina.** - Los datos están disponibles en formatos que pueden ser procesados por las computadoras.
- **Metadatos.** - Datos adicionales como la fecha de publicación y atribución se despliegan para cada conjunto de datos.

- **Licencias de datos claras.** - Las licencias de uso se muestran para un conjunto de datos. Algunos gobiernos y organizaciones internacionales han desarrollado licencias a medida, como la Licencia de Gobierno Abierto del Reino Unido y los Términos de uso del Banco Mundial<sup>10</sup>.
- **Pre-visualización y visualización de datos.** - Algunos catálogos de datos muestran una vista previa de los datos como parte de la funcionalidad de la plataforma.
- **Cumplimiento de estándares.** - La mayoría de los catálogos de datos abiertos tienen soporte incorporado para varios estándares, como formatos de datos y metadatos. Generalmente hacen que cada conjunto de datos esté disponible como una URL única y permanente, lo que permite citar y vincular los datos directamente.
- **Interfaz de programación de aplicación (API por sus siglas en inglés).** - Permiten a los desarrolladores de software acceder al catálogo de datos abiertos, y a menudo a los datos en sí, a través del software, mediante consultas en tiempo real a las bases de datos.
- **Seguridad.** - Implementación de medidas de seguridad para evitar que los datos sean modificados sin autorización (The World Bank, 2019).

Los catálogos de datos pueden ser presentados en tres modelos básicos: en una **plataforma única**, en la cual el servidor web y el servidor de archivos se encuentran fusionados en un mismo equipo; modelo de **servidores separados**, en donde el servidor web y el servidor de archivos se encuentran en instancias separadas; modelo de **catálogos federados**, en el que se mantienen separadas las instancias de servidor web y servidor de archivos, pero además el catálogo de datos se conforma del uso de API's y datos almacenados en otras organizaciones y no directamente en los servidores principales, como en los dos primeros modelos (The World Bank, 2019). La elección del modelo adecuado forma parte de los requerimientos no funcionales a definir en una iniciativa de datos abiertos.

---

<sup>10</sup> <https://data.worldbank.org/summary-terms-of-use>

### 1.9.3 Requerimientos funcionales para una infraestructura de datos abiertos gubernamentales

Para entrar en materia en este tema, los requerimientos funcionales se refieren a aquellos requerimientos que definen una función específica que indica cómo puede usarse un sistema (Stellman & Greene, 2005). Esto implica que es conveniente definir la infraestructura de datos abiertos que los usuarios encuentren funcional y usable (Zuiderwijk, 2015).

Los requerimientos funcionales “pueden ser agrupados dentro de cinco categorías principales, las cuales son:

- Búsqueda y localización de datos.
- Análisis de los datos.
- Visualización de los datos.
- Interacción sobre los datos.
- Análisis de calidad de los datos.” (Charalabidis et al., 2018, p. 97).

Cada una de estas categorías se explica brevemente dentro del cuadro 4.

<b>Categoría</b>	<b>Descripción</b>
<b>Búsqueda y localización de datos</b>	Este tipo de requerimientos funcionales cubren aspectos de fragmentación de datos, terminología estándar, soporte para la búsqueda, incluso en otros lenguajes, filtrado y ordenamiento de los resultados de la búsqueda.
<b>Análisis de los datos</b>	Los requerimientos funcionales de análisis de datos abarcan tópicos sobre el contexto de los datos, apoyo para el uso de los datos, heterogeneidad de los datos, y soporte para el análisis de los datos.
<b>Visualización de los datos</b>	Estos requerimientos funcionales comprenden la manera en que se facilitan herramientas de visualización de los datos, incluyendo aquellas para verlos en mapas de manera geo-referenciada.
<b>Interacción sobre los datos</b>	Este tipo de requerimientos funcionales incluyen consideraciones sobre la interacción entre proveedores, creadores de políticas y usuarios, considerando conversaciones y discusiones, herramientas para la comunicación y herramientas para el rastreo de cambios en los conjuntos de datos.

<b>Análisis de calidad de los datos</b>	Requerimientos funcionales con respecto a la dependencia de la calidad de los datos abiertos, baja calidad de los datos, y a los cambios y variación en la calidad
---	--

**Cuadro 4.** *Categorías de requerimientos funcionales.*

Fuente: Elaboración propia con base en Zuiderwijk (2015).

Una vez mencionadas las generalidades de los grupos de requerimientos funcionales, hay que tener en cuenta que “la infraestructura de datos tiene el objetivo principal de mantener a la sociedad informada y, por lo tanto, contribuye directamente a una mayor accesibilidad y gobernanza con respecto a los datos” (Charalabidis et al., 2018, p. 95).

También hay que considerar que para que estos elementos puedan establecerse, previamente se debe de cumplir con un marco para evaluar la preparación institucional, el cual contiene elementos como preparación para el gobierno, asuntos legales y de política, liderazgo adaptativo, preparación de recursos, capacidad de innovación, compartición de información, colaboración y compromiso, preparación para el gestión del cambio y preparación de la gente (Charalabidis et al., 2018).

#### **1.9.4 Métricas de la infraestructura**

Las infraestructuras pueden ser evaluadas por una serie de métricas que indican el grado de preparación que tienen para la publicación de datos. Estas métricas tienen que ver con la **Calidad de la Información** (*Datasets, Metadata, Linked Data*), la **Calidad del Sistema** (Capacidades de la Plataforma, Facilidad de Uso, Desempeño, Métricas adicionales) y la **Calidad del Servicio** (Licenciamiento, Retroalimentación y colaboración) (Charalabidis et al., 2018).

##### **1.9.4.1 Calidad de la Información**

En cuanto a las métricas relacionadas con la calidad de la información, estas se enfocan en aspectos relacionados con las características propias de los datos abiertos, como la unicidad, que sean primarios, que estén en línea, que sean gratuitos, oportunos, confiables, accesibles, objetivos y con, interpretabilidad,

etc. Las métricas se agrupan en tres dimensiones principales: conjunto *de datos* – evaluación de calidad de los datos-, *metadatos* –evaluación de la calidad de los metadatos-, *datos vinculados* –evaluación de la vinculación de los datos cuando aplique- (Charalabidis et al., 2018). Las dos últimas dimensiones contribuyen en gran medida a la característica de que los datos sean leídos e interpretados por la máquina.

#### **1.9.4.2 Calidad del sistema**

Las métricas orientadas a la calidad del sistema se enfocan en aspectos sobre el funcionamiento de las plataformas y se agrupan en tres dimensiones: 1. capacidades de la plataforma –en su relación con los conjuntos de datos-, 2. facilidad de uso de la misma –cubriendo aspectos de diseño e interacción de manera subjetiva-, y 3. desempeño – abarcando métricas como eficiencia, eficacia y API's-; adicionalmente se puede agregar una cuarta dimensión en la que se evalúan aspectos de interacción con el usuario, permitiéndole enriquecer los datos para volverlos a poner a disposición con un valor agregado (Charalabidis et al., 2018). Como se ha mencionado, el aspecto de uso de los datos y por ende de interacción con el usuario, queda fuera del alcance de este estudio.

#### **1.9.4.3 Calidad del servicio**

Las métricas orientadas a la calidad del servicio se agrupan en dos dimensiones: licenciamiento –cubriendo aspectos relacionados al uso de los conjuntos de datos- y la dimensión de retroalimentación y colaboración –relacionadas con aspectos que cubren comunicación y colaboración entre proveedores y consumidores de los datos-; es uno de los aspectos más importantes con respecto al uso permitido y explotación de los datos abiertos en la generación de valor económico y social (Charalabidis et al., 2018).

Cada una de estas métricas tiene un objetivo en particular en cuanto a la operación de la plataforma de datos abiertos. El mayor cumplimiento de cada uno de los puntos garantiza el objetivo general de mantener a una sociedad informada (Charalabidis et al., 2018).

### 1.9.5 Comparativa de plataformas de datos abiertos

En la sección de datos abiertos en América Latina y el Caribe expuesta anteriormente, es de notar que la mayoría de las plataformas de datos abiertos en estos países se identifican como CKAN. Existe además una variante de CKAN basada en Drupal para la publicación de datos abiertos denominada DKAN (DKAN, 2019c). En la presente sección se realiza un breve análisis de las plataformas CKAN y DKAN, ampliamente difundidas y utilizadas por los gobiernos de América Latina (Naser & Rosales, 2016). Además de ser muy populares, estas plataformas tienen la característica de ser de libre uso, lo cual es un aspecto importante en los lugares en donde el presupuesto es limitado.

Como se observa en el trabajo de Naser y Rosales (2016), un gran número de gobiernos de América Latina se encuentra utilizando la plataforma de datos abiertos CKAN. De igual manera, el sitio oficial de DKAN (2019d) muestra un amplio listado de los sitios que utilizan su plataforma para la publicación de datos abiertos.

CKAN es un *software* de libre uso y de arquitectura abierta y que se puede obtener desde <https://ckan.org>. Por su parte, la versión denominada DKAN, se puede integrar a administradores de contenido *web* de código abierto como Drupal. CKAN es una plataforma que ha habilitado a portales de alto perfil, como los de Estados Unidos y Reino Unido, pero DKAN es un complemento para hacer los datos más abiertos y accesibles (DKAN, 2019a).

CKAN son las siglas de *Comprehensive Knowledge Active Network*, es decir, Red Integral de Archivos de Conocimiento en su traducción al español. Mientras que, como se mencionó, DKAN, *Drupal-based Knowledge Active Network*, es una variante de CKAN basada en Drupal para la publicación de datos abiertos (DKAN, 2019c). Ambas plataformas cuentan con la ventaja económica de que se basan en el modelo colaborativo de código abierto. Bajo este mismo modelo, adicionalmente se cuenta con una suite de productos que sirven como una plataforma para el desarrollo de aplicaciones y portales semánticos, elaborada por INFOTEC: *Semantic Web Builder* (SWB) (INFOTEC, 2016b). La web semántica es una extensión de la web actual que permite la interpretación de la información tanto por agentes humanos como por agentes computarizados, utilizando formatos estándar

como XML (siglas en inglés de Lenguaje de Mercado Extensible) y constituyendo la base sintáctica de RDF (siglas en inglés de Marco de Descripción de Recursos) (Lamarca, 2019).

Existen otras plataformas SaaS (*Software as a Service*) que funcionan en la 'nube', por lo que su uso podría implicar un costo de operación adicional a los costos que se tienen por el uso de la infraestructura de TIC local. Ejemplos de estas plataformas SaaS son Junar<sup>11</sup>, OpenDataSoft<sup>12</sup>, Socrata<sup>13</sup> y ArcGIS Open Data<sup>14</sup> (The World Bank, 2019). Esta última plataforma posee además las características de un Sistema de información Geográfica (GIS, por sus siglas en inglés).

Utilizando la información de Naser y Rosales (2016) y de DKAN (2019d) se elabora el cuadro 5 para mostrar los sitios gubernamentales que utilizan ambas plataformas. En cuanto a SWB, no se tiene localizada alguna plataforma gubernamental que la utilice para publicar datos abiertos.

Plataforma	Sitios gubernamentales
CKAN	Argentina, Brasil, Chile, Colombia, Costa Rica, Ecuador, Guatemala, México, Paraguay, Perú, República Dominicana, Uruguay.
DKAN	El Salvador; Buenos Aires, Argentina; Rosario, Argentina; IT PAMI, Argentina; Bogotá Secretaría de Cultura, Recreación y Deporte, Colombia; Narino, Colombia; IMSS, México; CONCYTEC, Perú; Ministerio de Educación, Perú; Ministerio de Salud, Perú; Ministerio del Trabajo; Departamento de Educación de São Paulo, Brazil; La Paz, Bolivia.

**Cuadro 5.** *Sitios de datos abiertos relevantes por plataforma.*

Fuente: Elaboración propia con base en Naser & Rosales, (2016) y DKAN (2019d).

La información referente a los sitios de CKAN se concentra en los gobiernos nacionales, mientras que la información mostrada por DKAN se centra mayormente

<sup>11</sup> <http://www.junar.com/>

<sup>12</sup> <http://www.opendatasoft.com/>

<sup>13</sup> <http://www.socrata.com/>

<sup>14</sup> <http://opendata.arcgis.com/>

en organizaciones o gobiernos locales. De cualquier manera, se aprecia una gran aceptación por ambas plataformas en los países de América Latina.

Tanto por ser las más populares, como por ser las más accesibles, se hace una comparativa de características de las plataformas CKAN, DKAN y SWB, tomado como referencia lo indicado en las páginas web de las propias plataformas y expresando el resultado en el cuadro 6.

<b>Funciones</b>	<b>CKAN</b>	<b>DKAN</b>	<b>SWB</b>
API's	x	x	
Almacenamiento de información	x	x	x
Escalable	x	x	x
Datos de diferentes portales	x	x	
Almacenamiento de archivos	x	x	x
Herramientas geoespaciales	x	x	
Gestión de contenidos	x	x	x
Buscador	x	x	x
Temas	x	x	x
Pre visualización de contenido	x	x	
Multilinguaje		x	
Blog		x	
Control de usuarios	x	x	x
Integración con Drupal		x	
Gráficas		x	
Reportes			x
Monitoreo (accesos, memoria, conexiones DB, usuarios)		x	x
Publishflow			x
Gestor de sitios	x	x	x

**Cuadro 6.** *Comparativa de características por plataforma.*

Fuente: Elaboración propia con base en DKAN (2019<sup>a</sup>), DKAN (2019b) y CKAN (2019b)

De acuerdo a las fuentes del cuadro 6, la plataforma más completa, en cuanto a las características ofrecidas, es la DKAN. Sin embargo, existen algunos otros elementos que habrá que tomar en cuenta en la elección de una plataforma de datos abiertos, como pueden ser el entorno de los sistemas operativos, el lenguaje del código y las bases de datos soportadas y que a su vez deben estar en el

conocimiento y personal con que se cuenta para administrar dicha plataforma. En este sentido, se tiene el cuadro 7 que muestra una comparativa de este tipo de aspectos técnicos de las tres plataformas:

<b>Aspecto técnico</b>	<b>CKAN</b>	<b>DKAN</b>	<b>SWB</b>
Sistema Operativo	Ubuntu 12.04	Linux	Windows, Linux
Lenguaje de programación	Java	PHP 5.3	Java
Base de datos	PostgreSQL/Redis	MySQL 5.0.15	MySQL, Oracle
Servidor Web	Apache 2 / NGINX	Apache 2.x	Tomcat 6.x

**Cuadro 7.** *Comparativa de aspectos técnicos por plataforma.*

Fuente: Elaboración propia con base en CKAN Docs (2018), DKAN Docs (2019) e INFOTEC (2016a)

El cuadro 7 nos muestra que existe una diversidad en los requisitos técnicos de las plataformas planteadas para datos abiertos. Es importante señalar que las tres plataformas pueden implementarse con software de uso libre, lo que facilita su adquisición bajo el aspecto económico.

En México, como caso de éxito a nivel regional y mundial con su alta puntuación en *Open Data Barometer*, CKAN es la herramienta utilizada para la publicación de datos abiertos. Tomando esto en consideración, la integración de los conjuntos de datos contenidos en una plataforma estatal con la plataforma federal mexicana es un aspecto importante a tomar en cuenta para la selección de plataforma de datos abiertos a implementar en el gobierno local en Aguascalientes.

### **1.9.6 Plataformas de datos abiertos en México**

Puesto que la intención de este trabajo es proponer una plataforma de datos abiertos, un aspecto importante de la documentación es la situación que guardan este tipo de plataformas en México. Existe un número limitado de plataformas funcionales en el país, con las características especificadas en este capítulo, como

las disponibles en Baja California Sur, Jalisco, Estado de México, Puebla, San Luis Potosí y Sinaloa.<sup>15</sup>

Esta sección se documenta con el fin de ser un aspecto más en la selección de la plataforma a implementar, así como de la estrategia a seguir para lograrlo. En el entorno internacional, México es un caso de éxito en cuanto a la publicación de datos abiertos, con su buena posición dentro del Open Data Barometer con su plataforma federal ubicada en [datos.gob.mx](http://datos.gob.mx) (World Wide Web Foundation, 2019). En el entorno nacional, de las plataformas disponibles en el país mencionadas en el párrafo anterior, el Estado de Jalisco cuenta con una plataforma con una buena calificación en el componente de datos abiertos del Ranking de Portales de Transparencia (Sandoval, 2018) y su plataforma se localiza como [datos.jalisco.gob.mx](http://datos.jalisco.gob.mx). Por su parte, la plataforma de Puebla es digna de analizar, ya que cuenta con un buen número de conjuntos de datos disponibles tanto en su plataforma, como en la plataforma federal. A continuación, se hará una breve reseña de cada uno de ellos.

#### **1.9.6.1 datos.gob.mx**

La iniciativa de datos del gobierno federal tiene su soporte normativo en el Decreto por el que se establece la regulación en materia de datos abiertos emitido en 2015 (Gobierno de la República, 2015). La entidad responsable de la plataforma es la Unidad de Gobierno Digital de la Secretaría de la Función Pública. Actualmente tiene 40,727 datos de 280 instituciones (GOB.MX, 2019) y se basa en la plataforma CKAN (CKAN, 2019a). Actualmente se ubica en el lugar seis del Open Data Barometer (World Wide Web Foundation, 2019). Cuenta con una Guía para la implementación de datos abiertos que contempla las etapas de la preparación de los datos abiertos y la publicación de los mismos, en la que se establecen aspectos como los metadatos mínimos requeridos, niveles de acceso a datos, estándares de formatos a usar, acceso a las fuentes de datos por medios de API y servicios *Web*

---

<sup>15</sup> Las plataformas pueden encontrarse en <http://datos.bcs.gob.mx/>, <https://datos.jalisco.gob.mx/>, <http://datos.edomex.gob.mx/>, <http://datos.puebla.gob.mx/>, <https://datos.slp.gob.mx/>, <https://transparencia.sinaloa.gob.mx/p/datos-abiertos>, [https://www.tamaulipas.gob.mx/datosabiertos/datos\\_abiertos/](https://www.tamaulipas.gob.mx/datosabiertos/datos_abiertos/),

con protocolo de intercambio y manipulación de datos denominado REST, estándares de intercambio de información y de notación de instantes, etc. (México Digital, 2014).

#### **1.9.6.2 datos.jalisco.gob.mx**

La iniciativa de datos del gobierno del Estado de Jalisco tiene su soporte normativo en el Decreto por el que se establece la regulación en materia de datos abiertos emitido en 2015 (Presidencia de la República, 2015), ya que no cuenta con una legislación local en la materia. La entidad responsable de la plataforma es la Dirección General de Tecnologías de la Información de la Secretaría de Administración del Estado de Jalisco. Actualmente tiene 477 conjuntos de datos organizados en 36 grupos (Gobierno de Jalisco, 2019b). Jalisco se ubica en el primer lugar del Ranking de Portales Estatales de Transparencia 2018 (Sandoval, 2018).

La plataforma cuenta con la misma referencia de la Guía para la implementación de datos abiertos federal y emitida por la Unidad de Gobierno Digital, que contempla las etapas de la preparación de los datos abiertos y la publicación de los mismos, en la que se establecen aspectos como los metadatos mínimos requeridos, niveles de acceso a datos, estándares de formatos a usar, acceso a las fuentes de datos por medios de API y servicios *Web* con protocolo de intercambio y manipulación de datos denominado REST, estándares de intercambio de información y de notación de instantes, etc. (México Digital, 2014). En el aspecto técnico, la plataforma cuenta con las siguientes características: (Gobierno de Jalisco, 2019a):

- Distribución: Drujal 2.0
- Administrador de contenido: Drupal 7
- Index Server: Solr Solr 4 – Apache Tomcat
- Apache
- PHP
- Servidor de base de datos: MySQL

La generación y publicación de catálogos de datos se realiza con base en la guía publicada en la página <https://datos.gob.mx/guia/>. Esta guía indica, sobre todo en sus apartados 'Publica' y 'Manuales y Plantillas', la estructura técnica en la que se deben de preparar los datos, contemplando conjuntos, metadatos y formatos de presentación.

Es importante hacer notar que tanto la plataforma [datos.gob.mx](https://datos.gob.mx) como la plataforma [datos.jalisco.gob.mx](https://datos.jalisco.gob.mx) cuentan con personal dedicado por completo a la gestión de datos abiertos.

### **1.9.6.3 datos.puebla.gob.mx**

Al igual que la iniciativa de Jalisco, la iniciativa de datos del gobierno del Estado de Puebla tiene su soporte normativo principal en el decreto por el que se establece la regulación en materia de datos abiertos emitido en 2015, ya que no cuenta con una legislación local en materia de datos abiertos, además, cuenta con la misma referencia de la guía para la implementación de datos abiertos federal y emitida por la Unidad de Gobierno Digital (Gobierno de Puebla, 2019).

Los datos abiertos encontrados en la plataforma [datos.gob.mx](https://datos.gob.mx) muestra un gran número de conjuntos de datos, más de 500, originarios del Estado de Puebla (Datos.gob.mx, 2020), es por ello que se ha considerado como una buena experiencia en la implementación de los datos abiertos. En la actualidad la plataforma de datos abiertos del Estado de Puebla se puede encontrar en <http://datos.puebla.gob.mx/> y tiene disponibles 376 conjunto de datos (Gobierno de Puebla, 2020a).

La entidad responsable de la plataforma es la Coordinación Estatal de Transparencia y Gobierno Abierto. Puebla se ubica en el lugar décimo séptimo del ranking de portales estatales de transparencia 2018 (Sandoval, 2019).

La plataforma cuenta con una sección dedicada a los desarrolladores, en la que se pueden encontrar referencias a herramientas informáticas para la limpieza de datos, extracción, mapeo y publicación de datos, de manera similar a como se presenta por parte del Gobierno Federal. Las herramientas indicadas para tales

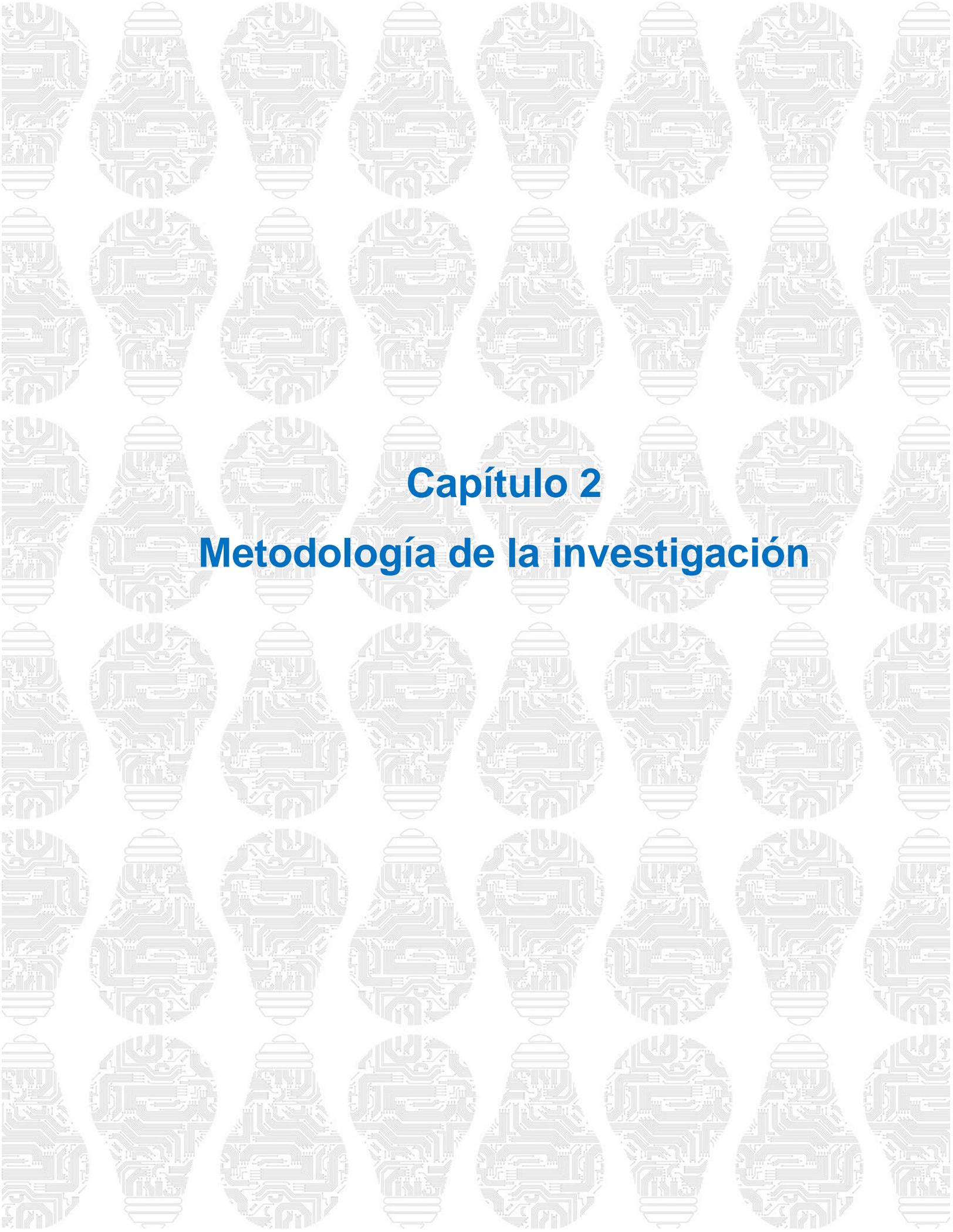
tareas son: OpenRefine (limpieza), import.io (extracción), ArcGIS y Carto (Georeferencia), DKAN (publicación) (Gobierno de Puebla, 2020b).

Para la carga masiva de datos se recomienda realizar la lectura de los apartados adecuados en las páginas *web* de DKAN y se especifican los accesos a los datos por medio de API's y servicios Web con protocolo de intercambio y manipulación de datos REST, con formato JSON (Gobierno de Puebla, 2020b).

Para el aspecto técnico de la publicación de datos abiertos, se intuye que las características de la infraestructura serían similares a los requerimientos de DKAN, aunque esto no pudo ser validado con el personal responsable de la Coordinación Estatal de Transparencia y Gobierno Abierto. Los requerimientos son (DKAN Docs, 2019, p. 1):

- MySQL: versión mínima 5.0.15+ con PDO
- Antes de la instalación, crear base de datos MySQL
- PHP: versión mínima 5.3.x
- Apache: versión mínima 2.x
- Git

Para concluir este capítulo, es conveniente destacar que la elección de una de las plataformas analizadas en este capítulo: CKAN, DKAN y SWB, dependió en gran medida de los resultados alcanzados durante el abordaje metodológico a desarrollar en el siguiente capítulo. Tomando esto en consideración, la suma tripartita que integran las entrevistas realizadas, los casos expuestos de plataformas de datos abiertos en México y lo asimilado en el marco conceptual, perfila que la plataforma a implementar en el GEA es la compuesta por la solución de datos abiertos DKAN, ejecutada en un servidor *web* Apache 2.x con servidor de base de datos MySQL 5.0.15, sistema operativo Linux con distribución CentOS 7 y un entorno de lenguaje de programación PHP 5.3. Estas soluciones deberán ejecutarse en un ambiente de servidores separados, considerando la infraestructura actual y presente en las oficinas del GEA.



**Capítulo 2**  
**Metodología de la investigación**

## Capítulo 2. Metodología de la investigación

Una vez que se tiene la exposición de las generalidades de los datos abiertos y de las plataformas para publicarlos en el capítulo anterior, el presente capítulo está dedicado a la búsqueda de una propuesta de plataforma de datos abiertos específica para el caso del Gobierno del Estado de Aguascalientes.

Se plantearán el objetivo general, así como el abordaje metodológico, con el cual se transitará por la investigación hasta llegar a una propuesta con los hallazgos encontrados.

### 2.1 Pregunta de investigación

La pregunta general de este trabajo es **¿Qué elementos conceptuales y tecnológicos es necesario considerar para el diseño de una plataforma informática que permita la creación, procesamiento, almacenamiento y publicación de datos abiertos en el Gobierno del Estado de Aguascalientes (GEA)?**

### 2.2 Objetivos

Con base en la pregunta expuesta, se tienen los siguientes objetivos:

#### 2.2.1 Objetivo general

Diseñar una plataforma informática que permita la creación, procesamiento, almacenamiento y publicación de datos abiertos en el Gobierno del Estado de Aguascalientes (GEA).

#### 2.2.2 Objetivos específicos

- Identificar los requerimientos conceptuales y tecnológicos para el diseño de esta plataforma.
- Determinar los actores involucrados en las diferentes fases del proceso de generación de datos abiertos.

- Detallar las características y especificaciones técnicas de los datos para ser publicados como datos abiertos.
- Definir los conjuntos de datos a publicar por parte de las dependencias y entidades del GEA.

## **2.3 Abordaje metodológico**

Para alcanzar los objetivos del presente trabajo, se realizó una investigación con un enfoque cualitativo, ya que no se establece una medición numérica (Hernández-Sampieri, Fernández-Collado, & Baptista-Lucio, 2010). La técnica utilizada será la triangulación, la cual consiste en el uso de diferentes fuentes, métodos o ambientes, aunque no necesariamente a tres partes (Okuda & Gómez-Restrepo, 2005). Las fuentes para este trabajo son tres: i. la literatura existente y documentos oficiales; ii. entrevistas a responsables de casos de éxito y; iii. la metodología Documentación de Requerimientos Centrada en el Usuario (DoRCU) de Ingeniería de Requerimientos, por tratarse de un desarrollo de una plataforma informática.

Los usuarios objetivo en este caso son, en primer lugar, el personal de alto mando de las diferentes dependencias y entidades dentro de una muestra dirigida, y en segundo lugar el personal técnico e informático de las mismas dependencias y entidades del GEA.

La DoRCU consiste en un proceso iterativo en el que se obtienen los requerimientos, contemplando aspectos de definición, técnicas y herramientas (Báez & Barba-Brunner, 2001), y cuenta con las siguientes etapas (Báez & Barba-Brunner, 2001):

- Elicitación de requerimientos. - En esta etapa se obtiene el conjunto de todas las partes involucradas.
- Análisis de requerimientos. - En esta etapa se detecta la presencia de áreas no especificadas.
- Especificación de requerimientos. - Esta etapa es un proceso de descripción del requerimiento a un nivel de detalle.
- Validación de requerimientos. - En esta etapa se verificará que los requerimientos sean consistentes y completos.

Durante la etapa de elicitación se podrá identificar a los informantes y a los involucrados en general, así como a la documentación pertinente tanto interna como externa, la cual se pueden componer de procedimientos, formatos, políticas, regulaciones, instrucciones de trabajo, etc. (Innova-T, 2019a). Los informantes son los proveedores de datos y quienes contextualizan a los mismos (Hernández-Sampieri et al., 2010).<sup>16</sup> Al realizar la elicitación, se ejecuta un proceso mediante el cual se obtienen los requisitos funcionales del sistema y de los usuarios, plasmando las reglas del negocio, las especificaciones de calidad, así como las restricciones que imperan en el entorno (Innova-T, 2019a).

En la etapa de Análisis se podrán identificar los requerimientos no funcionales como los aspectos de seguridad, estándares; las restricciones de diseño e implementación, los costos, normatividad, organización, infraestructura, presupuesto, entre otros. Con información suficiente de la etapa de elicitación, se puede hacer un análisis de la información, lo que, en la práctica, implica frecuentemente nuevos ciclos de elicitación, buscando completar la información que pudo haber quedado pendiente de las primeras sesiones (Fatto Consultoría y Sistemas, 2016).

Durante la etapa de especificación, principalmente, se esperan las definiciones y características técnicas de los datos, tomando en cuenta ciertas métricas y estándares adoptados de manera internacional.

Al finalizar la etapa de validación, se podrán tener identificados los conjuntos de datos abiertos recomendados para su publicación por parte de las dependencias y entidades del GEA.

Iniciando desde la etapa de elicitación, la recolección de datos no estructurados se hará mediante entrevistas a proveedores de datos con una muestra dirigida por el factor de importancia de los datos y tomando como referencia algunos casos similares en la región. Desde el punto de vista de DoRCU, los proveedores de datos son los usuarios de la plataforma desde el momento en que harán uso de ella para agregar y actualizar conjuntos de datos. Los consumidores

---

<sup>16</sup> La elicitación no es una palabra aceptada por la Real Academia Española y deriva del inglés *to elicit*, que tiene las acepciones de adquirir, conseguir, obtener, asegurar, derivar.

de datos abiertos no serán contemplados en esta propuesta, al menos de manera inicial, ya que se observarán casos de éxito y guías de otras instituciones.

Siguiendo la metodología de Documentación de Requerimientos Centrada en el Usuario (DoRCU) de Ingeniería de Requerimientos, se ha previsto la recolección de datos no estructurados tanto con entrevistas a titulares de dependencias y entidades del GEA, como con información procedente del personal de las áreas de Informática de las mismas, con una muestra dirigida por conveniencia, tomando en cuenta el factor de importancia de los datos y las referencias en algunos casos similares.

Además, se localizaron casos de éxito de plataformas en cada una de las entidades federativas por medio de las páginas *web* oficiales en cada una de ellas, para consultar la disponibilidad de datos estatales publicados en la plataforma federal (Gobierno de México, 2019). Derivado de esta búsqueda, se identificó que los gobiernos estatales de Jalisco, Puebla y Ciudad de México contaban con un mayor número de conjuntos de datos publicados, así como una mejor presentación en sus plataformas (Gobierno de Jalisco, 2019b; Gobierno de la Ciudad de México, 2020; Gobierno de Puebla, 2020a). Lo anterior se aprecia, en parte, considerando los subtotales de conjunto de datos por temas (Gobierno de Puebla, 2020b), así como la valoración de otros estudios que destacan la apertura de datos en el estado de Jalisco (Munte-Kunigami & Serale, 2018) y los esfuerzos pioneros desde 2015 del Estado de Puebla (Naser & Rosales, 2016). Derivado de ello, se realizaron entrevistas por conveniencia con los responsables de dichas plataformas para complementar la revisión documental.

La guía de la entrevista realizada se fundamenta en los aspectos desarrollados en el marco conceptual, así como en los objetivos general y específicos puntualizados en el presente capítulo. De esta manera, la guía se puede visualizar en el anexo 1 de este documento.

## **2.4 Entorno de Investigación**

Una vez expuesto el marco conceptual y tomando como base los objetivos y el abordaje metodológico planteados, se procede a realizar una propuesta de diseño

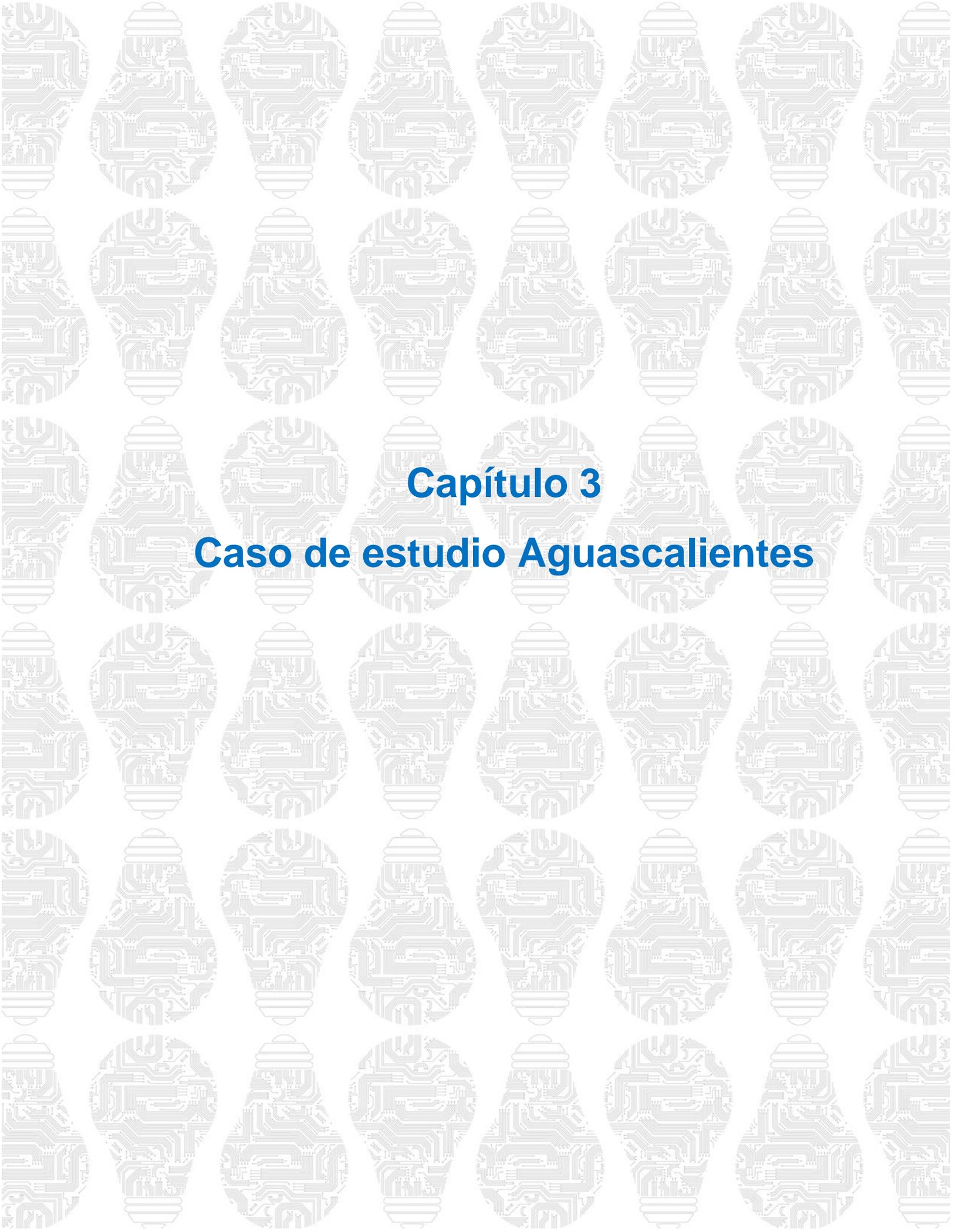
de una plataforma informática que permita la creación, procesamiento, almacenamiento y publicación de datos abiertos en el GEA. A través del estudio de las plataformas de datos abiertos de los estados de Jalisco y de Puebla, así como con la plataforma federal datos.gob.mx, se identificaron los conjuntos de datos prevalentes por temática o categoría, para vincularlos con las dependencias y entidades del gobierno local que pudieran generar datos similares. De esta forma, las temáticas existentes en la publicación de datos abiertos, pueden ser relacionadas con dependencias y entidades del GEA, como son:

- CMOV. Coordinación de Movilidad del Estado.
- CPLAP. Coordinación General de Planeación y Proyectos.
- IEA. Instituto de Educación de Aguascalientes.
- ISSEA. Instituto de Servicios de Salud del Estado de Aguascalientes.
- SAE. Secretaría de Administración del Estado.
- SECTURE. Secretaría de Turismo del Estado.
- SEDEC. Secretaría de Economía.
- SEFI. Secretaría de Finanzas.
- SEGUOT. Secretaría de Gestión Urbanística, Ordenamiento Territorial, Catastral y Registral.
- SOP. Secretaría de Obras Públicas.
- SSMAA. Secretaría de Sustentabilidad, Medio Ambiente y Agua.

Como resultado de lo anterior, se decide establecer una muestra consistente en los titulares de las dependencias y entidades mencionadas. Una vez hecho esto, se diseña la entrevista con preguntas que tienen que ver con los datos generales del personal, como son nombre, cargo, dependencia a la que pertenece. Los cuestionamientos base, encaminados a conseguir los objetivos establecidos, tienen que ver con indagar el nivel de conocimiento del funcionario sobre la temática de datos abiertos, la relación que estos tienen con la transparencia, la disponibilidad de datos abiertos en el Estado de Aguascalientes o en otros estados, la utilidad que tienen, la identificación de las partes involucradas en su generación, el presupuesto

o recursos que se tienen o podrían asignar para su publicación. Otras interrogantes se enfocan en la identificación de riesgos, inconvenientes o impedimentos legales que pudieran existir. Los aspectos contemplados en la entrevista se pueden identificar más claramente en el Anexo 1 – Guía de entrevista. Las entrevistas se realizan mediante la grabación de la conversación entre entrevistador y entrevistados, con la autorización de estos últimos para realizar dicha acción, para posteriormente transcribir el audio para su análisis. Las preguntas comprendidas en las entrevistas se pueden consultar en el Anexo 2 – Cuestionario de entrevista.

Finalmente, la identificación de requerimientos no funcionales orientados a los aspectos de seguridad, estándares, las restricciones de diseño e implementación, los costos, normatividad, organización, infraestructura y presupuesto se hace en colaboración con los encargados de los asuntos administrativos y de los encargados de los asuntos de los aspectos de tecnologías de la información en las dependencias y entidades mencionadas. Este análisis se hace con la ayuda del documento que se indica como Anexo 3 – Información del área de Informática.



**Capítulo 3**  
**Caso de estudio Aguascalientes**

## Capítulo 3. Caso de estudio Aguascalientes

### 3.1 Recolección de datos (DoRCU)

Como se mencionó en el capítulo de metodología, la recolección de datos se realizó tanto con los datos proporcionados por los titulares de dependencias y entidades gubernamentales, como con la información proporcionada por los responsables de las áreas de Informática.

#### 3.1.1 Titulares de dependencias y entidades gubernamentales

Aunque de hecho se realizaron las entrevistas en la totalidad de las dependencias y entidades del GEA especificadas, no todos los titulares a los que iba dirigida pudieron atender por cuestiones de agenda, por lo que en su lugar designaron a otros responsables de áreas clave, contando con la autorización verbal para citarlos en la mayoría de los casos. En el cuadro 8 se muestra a los entrevistados:

Dependencia o entidad	Cargo	Nombre
CMOV	Coordinador de Movilidad del Estado	Sin autorización para publicar datos.
CPLAP	Director General de Planeación Estratégica Territorial	Yami Euclides Reyes Delgado
IEA	Subdirector de Tecnologías de la Información	Felipe Leonardo Valero Delgado
ISSEA	Director de Planeación	Ulises Santiago Morales Espino
SAE	Secretario de Administración del Estado	Juan Francisco Larios Esparza
SECTUR	Secretario de Turismo	Humberto Javier Montero de Alba
SEDEC	Director General de Gestión Empresarial y Mejora Regulatoria	Juan Francisco Javier Hernández Pliego
SEFI	Sub-secretaria de Egresos	Virginia Adriana Calzada Ortiz
SEGUOT	Secretario de Gestión Urbanística, Ordenamiento Territorial, Registral y Catastral	Armando Roque Cruz
SOP	Director del Instituto de Infraestructura Física y Educativa	Alfredo Gallegos Padilla
SSMAA		Sin autorización para publicar datos.

**Cuadro 8.** Entrevistas realizadas.

Fuente: Elaboración propia.

Derivado de lo recabado durante las entrevistas, se pueden destacar los siguientes puntos:

- Existe una tendencia hacia el desconocimiento sobre el tema de los datos abiertos y una confusión para ubicarlos en el componente denominado Acceso a la información del elemento Transparencia de Gobierno Abierto.
- Una vez aclarada la confusión y de haber explicado su naturaleza y su ubicación dentro del Gobierno Abierto y la utilidad potencial que tienen, se encontró que no existe una publicación de datos abiertos en las dependencias y entidades del GEA, aunque se pudieron identificar algunas fuentes de información que podrían compartirse en el formato adecuado. Las fuentes probables de datos se pueden observar en la sección 3.2.6 Inventario de sistemas.
- El posible uso de los datos durante las entrevistas se inclina hacia un cumplimiento de obligaciones ubicadas en el elemento Rendición de cuentas de Gobierno abierto, más que hacia un uso para fomentar el desarrollo económico y social.
- Se observa que, en caso de implementarse una plataforma, en las dependencias y entidades del GEA no existen recursos económicos para la asignación de personal, bienes y servicios para tal fin. Algunos entrevistados sugirieron que, de implementarse, la parte tecnológica debería de estar ubicada en la Secretaría de Administración, la cual actualmente administra la Red Gubernamental de Tecnologías de la Información y Comunicación.
- Según los entrevistados, la generación de datos abiertos deberá realizarse con el mismo personal con que se cuenta, agregando nuevas funciones a su descripción de puestos, debido a la falta de presupuesto mencionado en el punto anterior.
- En el caso de los titulares entrevistados, la mayoría argumentó que tiene facultad para decidir sobre cuáles datos abiertos pudieran generarse y publicarse en la dependencia o entidad gubernamental a la que pertenecen.

- En cuanto a los riesgos, se expresó poca preocupación al respecto, pero mientras algunos manifestaron inquietud hacia aspectos externos como el mal uso o interpretación de los datos por parte de contrincantes políticos, algunos lo hicieron en cuanto a la dificultad de generarlos por la falta de personal y por las múltiples actividades que ya desempeñan los posibles asignados a tal tarea. Hubo también mención sobre el uso ventajoso que pudieran tener los datos disponibles, sobre todo en el aspecto de asignación de proyectos a proveedores, contratistas y desarrolladores inmobiliarios.
- En relación a los aspectos legales, la inclinación fue hacia consultar al área jurídica, aunque algunos manifestaron que en la perspectiva en la que se presentaba el proyecto, no veían inconveniente.

### **3.1.2 Información recabada de los responsables de Informática**

Los aspectos de seguridad, estándares, restricciones de diseño e implementación, costos, normatividad, organización, infraestructura y presupuesto, aunque de alguna manera fueron cuestionados durante las entrevistas, se plantearon y consultaron al personal de las áreas de Informática de las dependencias y entidades estatales, a través de las reuniones que mensualmente se tienen como parte de las actividades del aún no formalizado Comité de Informática del Gobierno del Estado de Aguascalientes. Como parte de esta consulta, se observó lo siguiente:

- El aspecto de seguridad informática es un tema relevante en cada una de las entidades y dependencias del GEA, pero no se proporcionó información detallada ni se autorizó a publicar información al respecto.
- Se utilizan los Manuales de Lineamientos y Procedimientos de TI vigentes en el Gobierno del Estado de Aguascalientes, como fundamento para los estándares utilizados.
- Al igual que el aspecto de seguridad, no se autorizó el detalle sobre la información del diseño de la red, más allá de saber si se tiene conectividad hacia la red gubernamental o no.

- Los organigramas del personal informático en las dependencias y entidades del GEA, con frecuencia se limitan a dos personas para proporcionar soporte técnico en equipos de cómputo y sistemas de información. Las áreas operativas se componen en su mayoría de capturistas de datos y registros, supervisadas o coordinadas por jefes de departamento, pero no existen áreas especializadas de explotación de la información.
- El personal de Informática en cuestión, no tiene la facultad para decidir sobre la implementación, y la forma en la que debería realizarse, de datos abiertos.
- Existe un apoyo por parte de ellos siempre y cuando se cuente con el soporte de sus superiores, quienes en su mayoría son los coordinadores o directores del área administrativa.
- Desconocen si se asignaría un presupuesto específico para el proyecto, pero lo perciben difícil. De igual manera, desconocen la normatividad aplicable en la materia de los datos abiertos.
- Finalmente, el personal tiene el conocimiento para establecer portales web genéricos y para el manejo de datos, pero requieren de una capacitación especializada para preparar los datos a fin de transformarlos en datos abiertos.

Se observa, adicionalmente, que la infraestructura de TIC es muy limitada en la mayoría de las dependencias y entidades del GEA. Para sistemas robustos y de gran relevancia, se hace uso de la infraestructura establecida en la SAE. En esta última dependencia, se cuenta con el denominado modelo de servidores separados, con el servidor web y el servidor de bases de datos ubicados en instancias separadas.

Los servidores TIC en la SAE funcionan con los sistemas operativos Microsoft Windows y Linux CentOS; utilizan lenguajes de programación basados en Microsoft Visual Studio y JavaScript; almacenan los datos en Microsoft SQLServer y en MySQL; por último, ofrecen los servicios de internet a través de la solución de IIS.

### 3.2 Inventario de sistemas

La determinación de los conjuntos de datos a publicar se basa en los sistemas informáticos y en los registros administrativos que se usan en las dependencias y entidades del GEA. Para la obtención de esta información, como parte de la búsqueda de los conjuntos de datos recomendados a publicar por parte de las dependencias y entidades del GEA en una plataforma de datos abiertos, se solicitó a los encargados de Informática que conforman el Comité de Informática del Gobierno del Estado de Aguascalientes el inventario de sistemas y registros que se usan en cada uno de sus centros de trabajo.

El hallazgo sobre la indagación de sistemas es que un número considerable de registros se basa en hojas de cálculo implementadas en Microsoft Office, lo que imposibilita la generación automática de datos abiertos, como puede hacerse si los datos se encuentran almacenados en sistemas de bases de datos como lo pueden ser MS-SQL Server, Oracle y MySql, entre otros.

Los sistemas o registros que destacan en las dependencias son los siguientes:

- CPLAP. Coordinación General de Planeación y Proyectos.
  - Proyectos gubernamentales.
  - Registro de Estadísticas.
    - Sociales.
    - Gobierno y Seguridad Pública.
    - Administración y Finanzas.
    - Economía.
    - Infraestructura, Medio Ambiente y Territorio.
- SAE. Secretaría de Administración del Estado.
  - Proveedores.
  - Adquisiciones.
  - Personal
- SEDEC. Secretaría de Economía.
  - Apoyos económicos y financiamiento.
  - Empresas e industria.
  - MIPyMES.
  - Empleo.
  - Actividad económica.

- Inversión Extranjera Directa.
- SEFI. Secretaría de Finanzas.
  - Padrón vehicular.
  - Presupuesto.
  - Contabilidad.
  - Ingresos.
  - Egresos.
- SOP. Secretaría de Obras Públicas.
  - Proyectos de obra pública.
  - Inversión en obra pública.
  - Padrón de Contratistas.
- SEGUOT. Secretaría de Gestión Urbanística, Ordenamiento Territorial, Catastral y Registral.
  - Padrón de desarrollos urbanos.
  - Padrón de desarrolladores.
  - Cartografía catastral.
- IEA. Instituto de Educación de Aguascalientes.
  - Catálogo de centros educativos en diferentes niveles, públicos y privados.
  - Registro de alumnos en los diferentes niveles educativos, público y privado.
  - Resultado de evaluaciones.
  - Registro de bibliotecas públicas.
- ISSEA. Instituto de Servicios de Salud del Estado de Aguascalientes.
  - Centros de salud y hospitales.
  - Adquisiciones de medicamentos e insumos.
  - Expediente clínico electrónico.
  - Registro de morbilidad y mortalidad.
- SECTURE. Secretaría de Turismo del Estado.
  - Centros de entretenimiento y atracciones.
  - Recorridos turísticos.
  - Inversión en la Feria Nacional de San Marcos.
  - Hoteles.
- SSMAA. Secretaría de Sustentabilidad, Medio Ambiente y Agua.
  - Registro de calidad del aire.

- Ubicación de áreas naturales protegidas, parques y jardines.
- Registro de incendios forestales.
- Centros de Verificación.
- SSP. Secretaría de Seguridad Pública.
  - Licencias de conductor emitidas.
  - Incidencia delictiva.
  - Registro de llamadas de emergencia.
- CMOV. Coordinación de movilidad.
  - Rutas de transporte público.
  - Ciclopistas.
  - Catálogo de paraderos.
  - Padrón de vehículos de transporte público.

### 3.3 Propuesta de implementación

Una vez que se han cubierto ampliamente los aspectos conceptuales a lo largo del trabajo, especialmente en el primer capítulo con la revisión de la literatura especializada en la materia, los aspectos tecnológicos se consolidan tanto con la consulta de plataformas de datos abiertos con sus elementos de modelos de datos y requerimientos no funcionales, así como con el acercamiento con los titulares e informáticos de las dependencias y entidades gubernamentales tomando en cuenta la metodología DoRCU. Con base a estos dos aspectos, se tiene la siguiente propuesta para el Gobierno del Estado de Aguascalientes, agrupada en puntos de interés planteados en la sección de objetivos.

#### 3.3.1 Requerimientos para el diseño de la plataforma

1. **Establecer un marco legal.** Tal y como se sugiere por Parycek *et al.* (2015) lo primordial es contar con un marco regulatorio para la publicación de datos abiertos, tomando en cuenta la legislación sobre protección de datos y privacidad, así como el establecimiento de penalidades por el incumplimiento. En las plataformas de éxito examinadas, estas se adhieren al *decreto por el que se establece la regulación en materia de datos abiertos* emitido en 2015, así como a la Guía de Implementación de la Política de

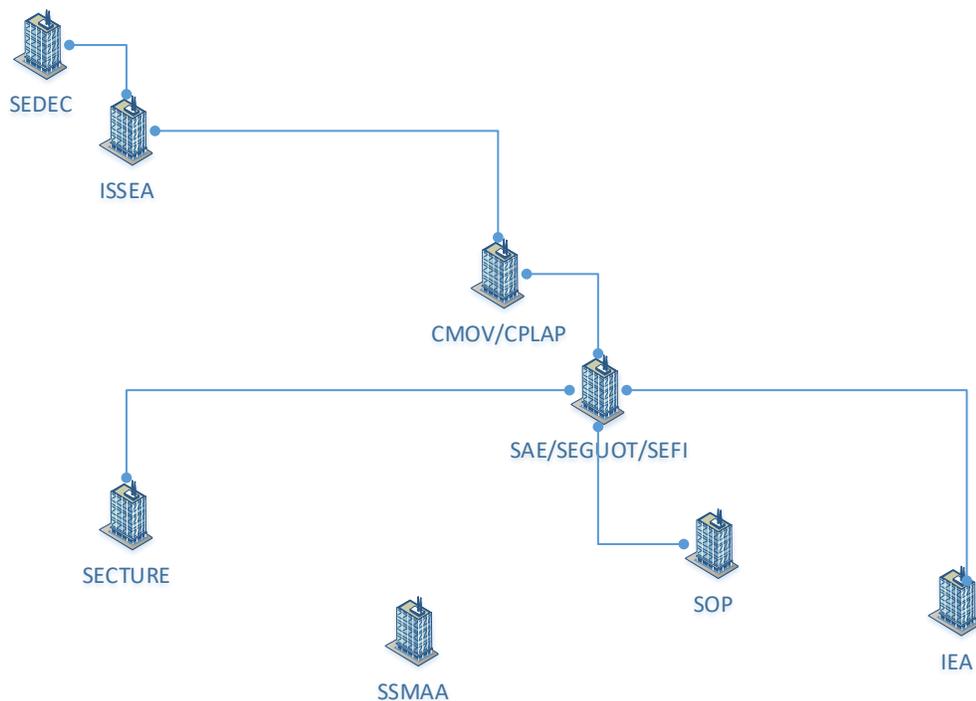
Datos Abiertos y la Ley General de Transparencia y Acceso a la Información Pública, principalmente. Algunos de los entrevistados, como en el caso de Obras Públicas y Desarrollo Económico, se denota una preocupación sobre la privacidad y la discrecionalidad de los datos, así como por el hecho de que algunos datos deben ser considerados restringidos y que no se pueden habilitar tan fácilmente, por cuestiones de seguridad y por procesos de licitación.

- 2. Hacer uso del recurso humano existente en las dependencias y entidades gubernamentales involucradas.** Pese a que en los casos de éxito de plataformas de bases de datos expuestas se tiene personal dedicado por completo a las mismas, se propone hacer uso del personal existente tanto en la parte operativa como en los asuntos de TIC. Lo anterior con base en lo recabado de las entrevistas, en especial de la realizada en la Secretaría de Administración, ya que se argumenta que se tienen que aprovechar las estructuras existentes con la mentalidad de hacer más con menos. Para ello será necesario agregar nuevas funciones a la descripción de puesto de cada uno de los involucrados, asignando comisiones específicas para el proyecto de datos abiertos.

Los organigramas de cada dependencia y entidad del GEA pueden ser consultados en [www.aguascalientes.gob.mx](http://www.aguascalientes.gob.mx).

- 3. Hacer uso de la infraestructura existente de TIC.** Al igual que en el punto anterior, la sugerencia es hacer uso de la infraestructura de TIC en las dependencias y entidades gubernamentales para la implementación de la plataforma de datos abiertos y para la explotación de sistemas y bases de datos. Lo anterior tomando en cuenta a que se deduce una falta de presupuesto adicional para dedicarlo a un nuevo proyecto. Esta es una recomendación hecha por los entrevistados, en especial en SAE, en donde se tiene la mentalidad de hacer más con menos.

Aunque el objetivo es publicar datos de la totalidad de dependencias y entidades del GEA, de manera inicial las listadas en el cuadro 8 se encuentran integradas a la Red Gubernamental, con excepción de la SSMAA, lo que facilita el uso de sitio TIC de la SAE como punto central en la concentración y publicación de los datos disponibles. Una representación de esta red puede ilustrarse en la figura 4, en la cual se ilustra la conectividad con que se cuenta hacia la SAE, por medio del Protocolo de Internet (IP por sus siglas en inglés).



**Figura 4.** *Conectividad de dependencias y entidades del GEA con SAE.*  
Fuente: Elaboración propia.

- 4. Elaborar una campaña de difusión.** Con base en la observación sobre el desconocimiento del tema de los datos abiertos en la mayoría del personal entrevistado, se recomienda realizar campañas de difusión en cada una de las dependencias y entidades del GEA. La formación y apoyo a los funcionarios en materia de datos abiertos, indicados por Parycek et al. (2015), forma parte primordial de los aspectos críticos de la publicación de datos abiertos. Haciendo uso de las recomendaciones anteriores, se deberá

hacer uso del personal existente, asignando nuevas funciones y formando un comité para dicha campaña, con especial apoyo de la Coordinación de Comunicación Social del Estado de Aguascalientes. Para coadyuvar en esta tarea, se pueden utilizar plataformas especializadas en la educación, como puede ser Moodle, así como la misma plataforma de datos abiertos, para facilitar la capacitación en línea, como se observa en las plataformas de éxito examinadas.<sup>17</sup>

5. **Adoptar la plataforma DKAN.** El resultado mostrado en el cuadro 6 referente a la comparativa de características por plataforma (CKAN, DKAN y SWB) indica que la plataforma *open source* DKAN tiene un mayor número de características en comparación con las otras dos. DKAN se encuentra disponible en <https://getdkan.org/> y es una plataforma más robusta en relación a CKAN, siendo estas dos las más usadas en América Latina. Además, se encuentra presente en dos plataformas de datos abiertos estatales de las más completas en el país, tal como se expuso en la sección de plataformas de datos abiertos en México, permitiendo la integración con Drupal como sistema de gestión de contenido. Para una apropiación de la plataforma, se deberá de realizar el diseño gráfico, maquetado o prototipo, de acuerdo a las especificaciones de imagen institucional establecidos en la administración en funciones. En este sentido, se deberán proveer por lo menos los siguientes recursos de TI:

- a. Servidor Web Apache 2.x, el cual puede ser obtenido de <http://www.apache.org/>
- b. Servidor de base de datos MySQL 5.0.15, disponible en <https://www.mysql.com/>
- c. Considerar el sistema operativo Linux con distribución CentOS 7, que se puede conseguir en <https://www.centos.org/download/>

---

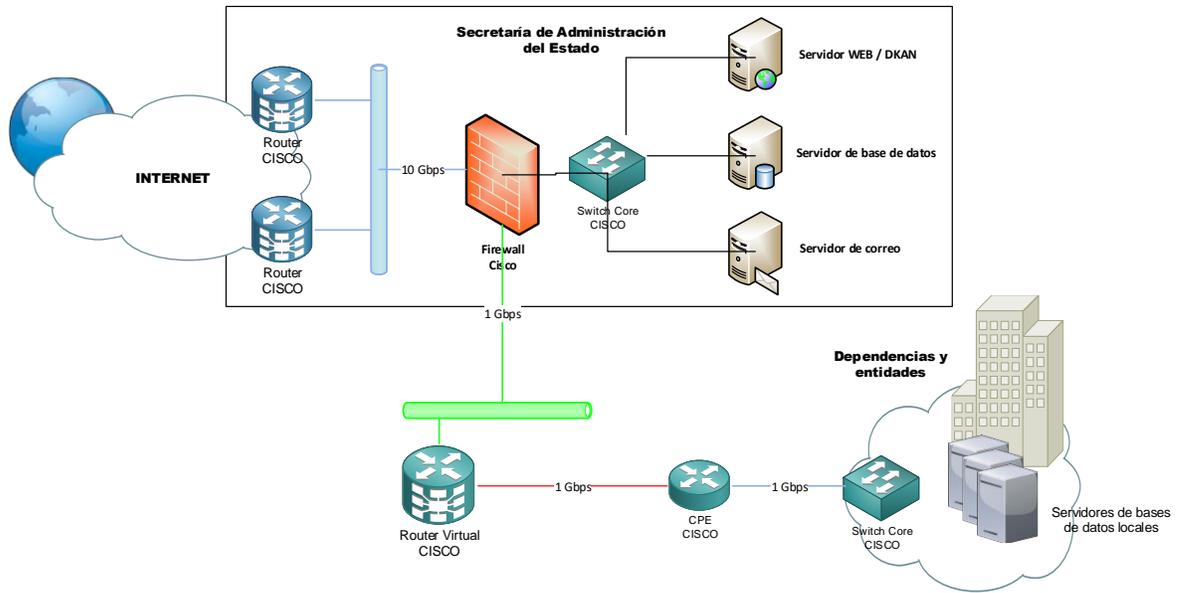
<sup>17</sup> Moodle es una plataforma educativa que se basa en un sistema de gestión de contenido. Su página oficial es <https://moodle.org>

d. Entorno de lenguaje de programación PHP 5.3, descargable en [https://www.php.net/releases/5\\_3\\_0.php](https://www.php.net/releases/5_3_0.php)

Estos recursos de TIC deberán de establecerse en SAE, como punto central de la conectividad de las dependencias y entidades gubernamentales estatales, y como parte de la infraestructura existente mencionada anteriormente.

6. **Establecer el modelo de servidores separados.** Como parte de los requerimientos no funcionales mencionado en la sección de aspectos técnicos, el modelo de servidores separados, en donde el servidor web y el servidor de archivos se encuentran en instancias separadas, es el ideal en el caso del GEA en virtud de que es el modelo actual en SAE. Adicionalmente, el hacer uso de la infraestructura existente es una recomendación de los entrevistados y un gran número de fuentes de datos se encuentran centralizados en SAE. Una vez que se consiga una madurez del proyecto, se podrá hacer una migración al modelo de catálogos federados, con datos almacenados en otras organizaciones y no directamente en los servidores principales.

En este punto se deben asignar los recursos, del inventario existente actual, para disponer de un servidor web y con uno servidor de archivos en instancias separadas, considerando las especificaciones del punto anterior. A manera ilustrativa de este aspecto, se presenta la figura 5, la cual representa la distribución de los servicios.



**Figura 5.** Esquema de servidores y red.  
Fuente: Elaboración propia

La figura 5 muestra, contenido dentro del recuadro, la infraestructura de la Secretaría de Administración del Estado, con su conexión a internet y el equipo de seguridad y comunicación informática antes de los servidores separados de bases de datos y aplicación, e incluso otro tipo de servidores como el de correo a manera de ejemplo. La conexión mostrada en color verde indica la conectividad que se tiene con las otras dependencias y entidades gubernamentales que hacen uso de los servicios de SAE y que en su momento y oportunidad podría usarse para el consumo de datos locales por medio de API's, siguiendo el modelo de catálogo de datos federados.

### 3.3.2 Características y especificaciones técnicas de los datos para ser publicados como datos abiertos.

Para la definición de estos aspectos, se deben tomar en cuenta las siguientes recomendaciones, tomando en cuenta los aspectos encontrados en la literatura especializada, así como en las plataformas de datos abiertos en México:

7. **Adoptar el modelo de ciclo de vida de los datos abiertos.** Aunque las herramientas y métodos expuestos en el cuadro 3 son muy variados, se

deberán de adoptar los recursos disponibles y existentes en el GEA, tal como se recomienda por los entrevistados.

8. **Adecuar los datos.** Los conjuntos de datos disponibles en las plataformas de datos abiertos en México analizadas, toman en cuenta la adecuación de los datos a las especificaciones indicadas en la plataforma <https://datos.gob.mx/guia/>, considerando los pasos detallados disponibles en las secciones de Planea, Publica, Perfecciona, Promueve (México Digital, 2014). Este punto incluye la documentación de los metadatos y adjuntar el diccionario de datos en cada conjunto de datos a publicar.

La guía publicada en la plataforma de datos.gob.mx respalda el punto referente a la accesibilidad, interoperabilidad y estándares, referido por Parycek et al. (2015), el cual indica que se deben utilizar estándares para los datos, metadatos, licencias y protocolos de intercambio y direccionamiento.

9. **Diseñar la plataforma.** En este punto se consideran las características en común que comparten los catálogos de datos, tal como se desglosa en la sección de aspectos técnicos: facilidad de acceso, búsqueda, acceso a datos legibles por máquina, metadatos, licencias de datos claras, pre-visualización y visualización de datos, cumplimiento de estándares, interfaz de programación de aplicación y seguridad, expuestos. En este sentido, la plataforma DKAN es un punto de referencia excelente, puesto que cuenta con los puntos mencionados. El contar con una plataforma, es uno de los aspectos críticos para la publicación de datos abiertos (Parycek et al., 2015).

Los tres puntos anteriores forman parte también de los requerimientos funcionales expuestos en el punto 1.9.3, relativos a la búsqueda y localización de datos, análisis de los datos, visualización de los datos, interacción sobre los datos y análisis de calidad de los datos. De esta manera, el modelo de ciclo de vida de los datos abiertos, la guía propuesta por el Gobierno de México y la plataforma *open source* DKAN representan, en conjunto, la solución a los aspectos funcionales de la plataforma de datos abiertos en propuesta.

### 3.3.3 Actores involucrados en la generación

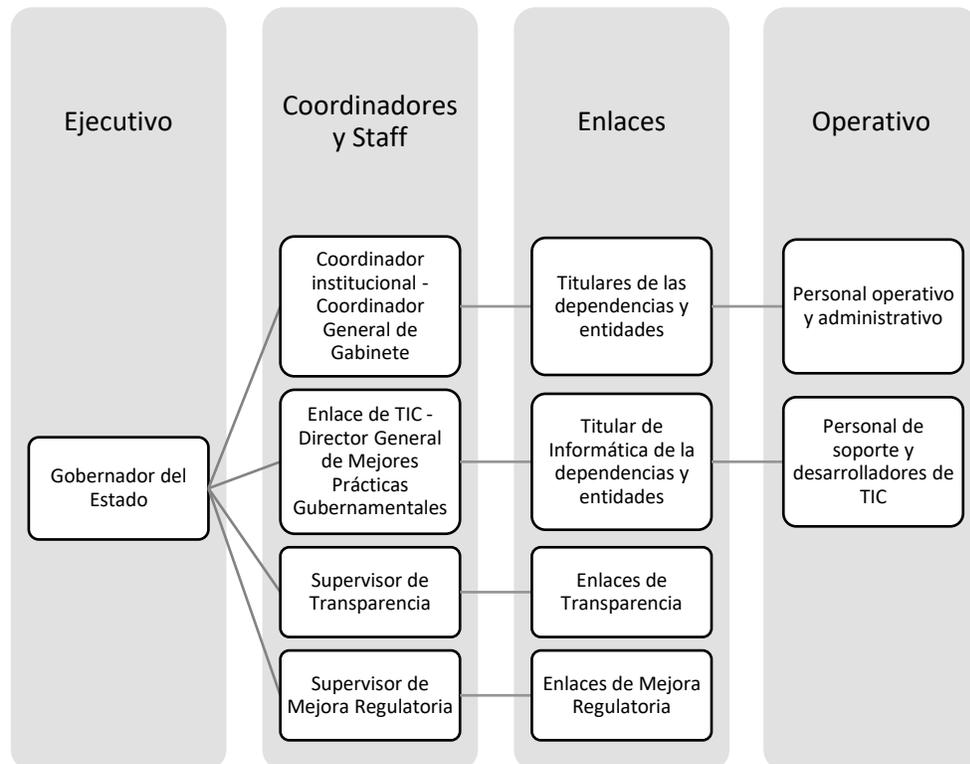
10. **Establecer un equipo de trabajo responsable.** La estrategia y apoyo político para desarrollar una estrategia en la que participen la alta gerencia y los líderes de los grupos, forma parte de los aspectos críticos para la publicación de datos abiertos, como lo establecen Parycek *et al* (2015). Establecer un equipo de trabajo responsable para coordinar la generación y publicación de datos abiertos, que incluya al titular del Gobierno del Estado y a funcionarios de alto mando, así como al personal técnico y operativo de las dependencias y entidades del GEA, es esencial para un proyecto de datos abiertos.

Comenzando con la participación del Gobernador del Estado, con la intención de contar con el respaldo e impulso que su figura permite, es necesario contar con dos coordinadores, uno orientado a la parte operativa y administrativa, denominado Coordinador institucional; otro orientado a los aspectos de las TIC, denominado Coordinador de TIC. Para el primer caso, el Coordinador General de Gabinete será el elemento esencial de impulso y enlace con los titulares, los cuales a su vez darán instrucciones al personal operativo y administrativo en el acceso y explotación de los datos. Para el segundo caso, el Director General de Mejores Prácticas Gubernamentales (DGMPG) será el encargado de coordinar a los titulares de Informática, quienes a su vez contarán con el apoyo del personal de soporte y de desarrollo, en caso de que exista.

Adicionalmente, es importante que participen como Supervisor de Transparencia y Supervisor de Mejora Regulatoria, el titular de la Unidad de Transparencia de la CE y el titular de la Dirección General de Mejora Regulatoria, respectivamente. Estos a su vez contarán con el apoyo de los enlaces de transparencia y mejora regulatoria. Es importante notar que esta relación laboral ya se encuentra en operación para fines distintos a los datos

abiertos y que no es necesario contratar personal adicional para el equipo de trabajo propuesto.

Lo anterior puede ser conceptualizado en la figura 6, deberán designarse cada una de las personas que se verán involucradas en el proceso de generación y publicación de datos abiertos.



**Figura 6.** *Equipo de trabajo para generación y publicación de datos abiertos*  
Fuente: Elaboración propia.

**11. Identificar al personal involucrado en la apertura de datos**, tomando en cuenta su relación con los sistemas identificados y priorizados en colaboración con los titulares de las dependencias y entidades del GEA. Con base en la descripción de puesto actual del personal, delimitar las acciones en relación a esta apertura y asignar responsabilidades. Se tendrá consideración del inventario de sistemas listado previamente en este documento.

Como parte del punto de formación y apoyo a los funcionarios de los aspectos críticos para la publicación de datos abiertos, es obligado realizar capacitación específica con respecto a su función, tomando en cuenta el ciclo de vida de los datos abiertos que se propone en el punto 7.

La formación de un equipo de trabajo, forma parte de los pasos contemplados en la guía para el cumplimiento del Decreto de Datos Abiertos, publicada en la página <https://datos.gob.mx/guia/>.

#### **12. Designación de los equipos de Análisis y de Desarrollo de sistemas.**

Tomando en cuenta la recomendación de hacer uso de los recursos existentes, los equipos de análisis y de desarrollo de sistemas deberán emanar principalmente de la Dirección General de Mejores Prácticas Gubernamentales de la SAE, como ente central de las TIC en el GEA. Adicionalmente se puede contar con el apoyo del personal de Informática presente en las dependencias y entidades gubernamentales involucradas en el inventario de sistemas mencionado anteriormente.

El equipo de análisis de sistemas se encargará de hacer el diseño, maquetado o prototipo de la plataforma, como parte de la personalización de la plataforma DKAN. El equipo de desarrollo de sistemas o programadores se encargará de hacer la adaptación de la plataforma de código abierto seleccionada, así como su implementación en los servidores Web de la SAE.

Adicionalmente, se debe de contar con el personal en funciones del área de soporte a las tecnologías de información, perteneciente a la DGMPG, quienes se encargan de la administración de la infraestructura y servicios TIC en SAE.

#### **3.3.4 Conjuntos de datos**

**13. Revisar los inventarios de sistemas.** Es necesario que se tomen en cuenta los sistemas de información existentes e identificados en colaboración con los titulares de las dependencias para priorizar los datos que serán abiertos.

Para ello es importante asimilar la experiencia de las plataformas de datos abiertos existentes en el país. En este punto se detectó un interés y disposición por parte de los entrevistados, quienes en su mayoría argumentaron que tiene facultad para decidir sobre cuáles datos abiertos pudieran generarse y publicarse en la dependencia o entidad gubernamental a la que pertenecen.

**14. Crear listados de los conjuntos de datos.** A partir del punto anterior, se tienen que crear los listados de los conjuntos de datos que se generarán por dependencia, tomando en cuenta los ya identificados en la sección *Inventario de sistemas* de este documento, como resultado de la adopción de experiencias de otras plataformas de datos abiertos en el país, así como de los listados de sistemas proporcionados por los responsables de Informática de las dependencias y entidades del GEA.

**15. Actualizar los datos.** Como sugieren Parycek *et al.* (2015) hay que establecer una estrategia para mantener publicados los conjuntos de datos, asegurando la continuidad en la provisión y transparencia en las condiciones en las cuales la publicación de datos se desarrolla, definiendo la frecuencia especificada en cada conjunto de datos.

**16. Asegurar la calidad de los datos.** Es obligado realizar revisiones aleatorias y periódicas, definidas por comisiones específicas para determinar la calidad de los datos publicados. Este punto se contempla en la sección de riesgos y mitigación en la publicación de datos abiertos, expuesta en este documento.

La calidad es una característica importante considerada en el modelo del ciclo de vida los datos abiertos propuesto en el punto 7. Además el análisis de calidad de los datos forma parte de los requerimientos funcionales (Charalabidis et al., 2018).

Una vez desglosada la propuesta de implementación, para concluir este capítulo recordemos que el objetivo principal de nuestra investigación era diseñar

una plataforma informática que permita la creación, procesamiento, almacenamiento y publicación de datos abiertos en el Gobierno del Estado de Aguascalientes (GEA). Basados en lo anterior, principalmente en lo expuesto en los capítulos 1 y 2, el diseño de la plataforma contempla los siguientes elementos:

- Características generales de los datos abiertos.
- Grado de madurez de los datos abiertos.
- Riesgos en la publicación de datos abiertos que pueden ser mitigados con diferentes acciones.
- Estándares establecidos a nivel internacional.
- Normatividad.
- Estrategia política, al igual que apoyo y formación de los funcionarios y servidores públicos.
- Plataformas *open source disponibles*.
- Guías de implementación.
- Infraestructura, hardware y software, de TIC.
- Definición de métricas.
- Modelos de datos.
- Catálogos de datos a publicar, así como sus características y el modelo en que serán presentados.

The background of the entire page is a repeating pattern of lightbulbs. Each lightbulb is filled with a complex, grey circuit board pattern, representing technology and innovation. The lightbulbs are arranged in a grid, with some appearing slightly more prominent than others.

## Conclusiones

## Conclusiones

Existe un gran número de fuentes de datos disponible en las organizaciones gubernamentales que pueden ser utilizadas en beneficio de las propias administraciones, pero también por parte de los ciudadanos con diferentes intenciones, que van desde la transparencia y mejora de los gobiernos hasta la estimulación del desarrollo económico y social. De estos propósitos, el que puede ser medido con relativa facilidad es el de la transparencia, y es donde el Gobierno del Estado de Aguascalientes debe trabajar para mejorar.

La apertura de los datos gubernamentales es un tema relativamente nuevo y aun así existe un gran número de investigaciones y trabajos científicos en torno a los datos abiertos. Sin embargo, pese a todo el potencial que tienen, en Aguascalientes todavía no se tiene suficiente conocimiento de ellos tanto para propiciar la transparencia, por parte de los servidores públicos, como para lograr los objetivos de crecimiento económico, por parte de los ciudadanos. El tema de datos abiertos se relaciona y traslada usualmente hacia el derecho a la información en lugar de ubicarlos como datos para ser usados hacia los objetivos mencionados. Se necesitan cambios a nivel organizativo, administrativo, operativo y técnico que resulten en la formulación de políticas y en el cambio de paradigmas tradicionales.

Los datos abiertos tienen características muy bien definidas y difundidas en la literatura, como se ha mencionado, pero, quizá lo más relevante es que son gratuitos, de libre uso, primarios, legibles por máquinas y oportunos. Tienen además, diferentes grados de madurez en la medida en que estos son presentados a través de plataformas o portales.

Pese a que en la región de Latinoamérica y el Caribe se cuenta con la mayor cantidad de adopciones de la Carta Internacional de Datos Abiertos, son países de otras regiones las que cuentan con una mejor preparación para la publicación de datos abiertos. Cabe destacar que, gracias a las acciones en la materia por parte del expresidente Enrique Peña Nieto, México cuenta con una plataforma de datos abiertos muy completa. Sin embargo, mantener dicha plataforma en condiciones operativas favorables requiere de una constante actualización de los datos y del

compromiso de todos los involucrados; también es recomendable el monitoreo constante de las evaluaciones para no tener un estancamiento en las acciones requeridas.

En el Estado de Aguascalientes, se vislumbran tres aspectos críticos para la implementación de una nueva iniciativa de datos abiertos: 1) legislación, regulación y licencias; 2) formación y apoyo a los funcionarios y 3) plataformas de datos abiertos, herramientas y servicios. Aunado a esto, existe una gran necesidad de difundir los fundamentos y objetivos de los datos abiertos a los servidores públicos.

Para el aspecto legislativo, regulación y licencias, es recomendable que en el corto plazo el GEA tenga su soporte normativo en el decreto por el que se establece la regulación en materia de datos abiertos emitido en 2015 a nivel federal; pero, a largo plazo, lo conveniente será que se establezca la normatividad en materia de datos abiertos a nivel local, tomando los documentos mencionados como punto de partida.

Relativo al punto de Formación y apoyo a los funcionarios, es primordial que se establezca un programa con acciones de difusión y capacitación en materia de datos abiertos. Aunque dirigida de manera universal hacia los funcionarios públicos del GEA, se enfatizarán dichas acciones hacia el equipo de trabajo o comité responsable de coordinar la generación y publicación de datos abiertos, conformado por personal de alto mando, técnico y operativo en las dependencias y entidades involucradas.

En cuanto al aspecto técnico de plataformas de datos abiertos, herramientas y servicios, se tiene que el ciclo de vida de los datos abiertos es el más adecuado para el GEA para crear y publicar datos abiertos. Debido a la estandarización de este tipo de datos, en la literatura y en la *World Wide Web* se pueden encontrar múltiples guías para establecer plataformas orientadas a la publicación de datos abiertos, siendo la de Datos Abiertos de México (<https://datos.gob.mx/guia>) una de las más indicadas para el caso del GEA por el potencial intercambio de datos entre ambas plataformas y por la estructuración de los datos en sí.

Existen varias métricas de la infraestructura de una plataforma de datos abiertos, abarcando aspectos de calidad de la información, capacidad del sistema y

calidad del servicio, los cuales contemplan la conformación de los datos, su presentación y la retroalimentación por los usuarios, respectivamente.

La disponibilidad de plataformas abiertas, como DKAN, facilita en gran medida la implementación de una iniciativa de datos abiertos, tomando como fuentes de datos los múltiples sistemas existentes en las dependencias y entidades gubernamentales. El hecho de contar con una red gubernamental de telecomunicaciones, centralizada por la Secretaría de Administración del Estado, posibilita la adecuada actualización de conjuntos de datos, una vez que se establezca una plataforma de datos abiertos. El esquema basado en servidores distintos para el almacenamiento y la publicación de los datos abiertos, además de la disponibilidad en SAE del *software* necesario para estos dos roles, incrementa la factibilidad de la implementación de una solución de datos abiertos. De igual forma se aprecia la presencia de personal vinculado a las TIC con la capacidad de realizar dicha implementación con algo de capacitación especializada.

La experiencia de otros gobiernos locales, demostrada a través de plataformas propias, es una guía adicional para alcanzar de manera más rápida los objetivos que se establezcan.

Para el caso del Estado de Aguascalientes, se identifica que no existen recursos económicos, tecnológicos ni humanos adicionales para implementar una iniciativa de datos abiertos, por lo que se deben de hacer los ajustes necesarios en la organización y las funciones del personal involucrado, así como en la infraestructura de TIC existente.

Es imperativo estar conscientes de los riesgos que existen en la publicación de datos abiertos a través de una plataforma, sobre todo en aspectos que contravengan la ley, invadan la privacidad y pongan en riesgo la infraestructura de TIC que la sustenten. Para minimizar estos riesgos, hay que considerar las acciones que se pueden realizar y que se han planteado en este documento de manera superficial.

Aunque en general es bueno aprender de las experiencias de otros gobiernos, se debe de tener en cuenta que la mejor plataforma es la que se ajusta en mejor medida al contexto local.

## Bibliografía

- Amozorrutia, A. B., Augusto, C., & Ditadi, S. (2017). Gobierno Abierto , Datos Abiertos ¿ Existe una relación con la gestión documental ? Análisis en cuatro países de Latinoamérica. *Legajos*, (13), 111-143. Recuperado 6 de febrero de 2009 de <http://www.archivos.gob.mx/Legajos/pdf/Legajos13/08GobiernoAbierto.pdf>
- Báez, M. G., & Barba-Brunner, S. I. (2001). Metodología DoRCU para la Ingeniería de Requerimientos, 210-222.
- Broad, E., Tennison, J., Starks, G., & Scott, A. (2015). Open Data Institute Prepared for the 3rd International Open Data Conference Ottawa , May 2015, (May). Recuperado 15 de enero de 2019, de <https://theodi.org/article/who-owns-our-data-infrastructure/>
- Carrara, W., Chan, W. S., Fischer, S., & van Steenberg, E. (2015). *Creating Value through Open Data*. (Digital Agenda for Europe, Ed.). European Commission. Recuperado 15 de marzo de 2019 de <https://doi.org/10.2759/328101>
- CEPAL (Comisión Económica para América Latina y el Caribe). (2003). *Los caminos hacia una sociedad de la información en América Latina y el Caribe*. Naciones Unidas , 2195, 75. Santiago de Chile: Naciones Unidas. Recuperado 16 de enero de 2019, de [https://repositorio.cepal.org/bitstream/handle/11362/2354/2/S034237\\_es.pdf](https://repositorio.cepal.org/bitstream/handle/11362/2354/2/S034237_es.pdf)
- CEPAL (Comisión Económica para América Latina y el Caribe). (2014). Big data and open data as sustainability tools: A working paper prepared by the Economic Commission for Latin America and the Caribbean. *Naciones Unidas Comisión Económica Para América Latina Y El Caribe*.
- Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Höchtl, J., & Ferro, E. (2018). *The World of Open Data*. (Springer, Ed.), *Springer* (1.<sup>a</sup> ed.). Springer International Publishing. Recuperado 6 de junio de 2018, de <http://www.springer.com/gp/book/9783319908496>
- Chignard, S. (2013). A brief history of Open Data. Recuperado 5 de abril de 2018,

de <http://parisinnovationreview.com/articles-en/a-brief-history-of-open-data>  
CKAN (Comprehensive Knowledge Archive Network). (2019a). CKAN, the world's leading Open Source data portal platform. Recuperado 31 de julio de 2019, de <https://ckan.org/>

CKAN (Comprehensive Knowledge Archive Network). (2019b). CKAN Features. Recuperado 28 de julio de 2019, de <https://ckan.org/features/>

CKAN Docs (Comprehensive Knowledge Archive Network). (2018). Installing CKAN from package. Recuperado 28 de julio de 2019, de <https://docs.ckan.org/en/2.8/maintaining/installing/install-from-package.html>

Comisión Europea. (2011). Decisión de la Comisión de 12 de diciembre de 2011 relativa a la reutilización de los documentos de la Comisión (2011/833/UE). *Diario Oficial de la Unión Europea*, (13), 39-42.

Congreso del Estado de Aguascalientes LXII Legislatura. (2016). Ley de Transparencia y acceso a la información pública del Estado de Aguascalientes. México. Recuperado 17 de febrero de 2019 de [http://www.ags.gob.mx/transparencia/docs/art55/fracc\\_1/estatal/LEY DE TRANSPARENCIA Y ACCESO A LA INFORMACIÓN PÚBLICA DEL ESTADO DE AGUASCALIENTES Y SUS MUNICIPIOS.pdf](http://www.ags.gob.mx/transparencia/docs/art55/fracc_1/estatal/LEY_DE_TRANSPARENCIA_Y_ACCESO_A_LA_INFORMACION_PUBLICA_DEL_ESTADO_DE_AGUASCALIENTES_Y_SUS_MUNICIPIOS.pdf)

Council of Europe. (2018). *Transparency and open government* (Congress of local and regional authorities Vol. 35), (14), 1-25. Recuperado 25 de julio de 2019, de <https://rm.coe.int/transparency-and-open-government-governance-committee-rapporteur-andre/16808d341c>.

Datos.gob.mx. (2020). Datos Abiertos - Búsqueda Puebla. Recuperado 25 de junio de 2020, de <https://datos.gob.mx/busca/dataset?organization=estado-de-puebla>

DKAN (Drupal-based Knowledge Archive Network). (2019a). Comparing DKAN and CKAN. Recuperado 28 de julio de 2019, de <https://dkan.readthedocs.io/en/latest/introduction/dkan-ckan.html>

DKAN (Drupal-based Knowledge Archive Network). (2019b). DKAN Open Data Platform. Recuperado 28 de septiembre de 2019, de <https://getdkan.org/>

DKAN (Drupal-based Knowledge Archive Network). (2019c). DKAN Overview.

Recuperado 28 de julio de 2019, de <https://docs.getdkan.com/en/latest/introduction/>

DKAN (Drupal-based Knowledge Archive Network). (2019d). DKAN sites. Recuperado 28 de julio de 2019, de <http://docs.getdkan.com/en/latest/introduction/dkan-sites.html>

DKAN Docs (Drupal-based Knowledge Archive Network). (2019). DKAN Installation Basics. Recuperado 28 de julio de 2019, de <https://dkan.readthedocs.io/en/latest/installation/basic.html>

Estrada, H., Solís, J., Vieyra, S., Najera, K., & Vázquez, B. (2013). Plataforma de Datos Abiertos: una herramienta para promover la transparencia de la información gubernamental. Recuperado 18 de abril de 2019, de [https://infotec.repositorioinstitucional.mx/jspui/bitstream/1027/135/1/Articulo\\_SW\\_LOD\\_2013.pdf](https://infotec.repositorioinstitucional.mx/jspui/bitstream/1027/135/1/Articulo_SW_LOD_2013.pdf)

Fatto Consultoría y Sistemas. (2016). Elicitación de Requisitos. Recuperado 10 de octubre de 2019, de <https://www.youtube.com/watch?v=VOmLc1q4Md8>

Ferrer-sapena, A., & Peset, F. (2011). Acceso a los datos públicos y su reutilización: open data y open government. *El profesional de la información*, 20(3), 260-269. Recuperado 11 de mayo de 2019, de <https://doi.org/10.3145/epi.2011.may.03>

GOB.MX (Sitio oficial de Gobierno de México). (2019). Descubre datos abiertos de tu gobierno. Recuperado 31 de julio de 2019, de <https://datos.gob.mx/>

Gobierno de Jalisco. (2019a). Acerca de. Recuperado 31 de julio de 2019, de <https://datos.jalisco.gob.mx/acerca-de>

Gobierno de Jalisco. (2019b). Datos Abiertos Jalisco. Recuperado 31 de julio de 2019, de <https://datos.jalisco.gob.mx/>

Gobierno de la Ciudad de México. (2020). Portal de datos de la Ciudad de México. Recuperado 6 de julio de 2020, de <https://datos.cdmx.gob.mx/pages/home/>

Gobierno de la República. (2015). ¿Qué son los Datos Abiertos? Recuperado 3 de abril de 2018, de <https://datos.gob.mx/blog/que-son-los-datos-abiertos?category=casos-de-uso&tag=educacion>

Gobierno de México. (2019). Descubre datos abiertos de tu Gobierno. Recuperado 8 de agosto de 2019, de <https://datos.gob.mx/>

Gobierno de Puebla. (2019). Gobierno Abierto. Recuperado 10 de julio de 2019, de <http://cetga.puebla.gob.mx/contenido/gobierno-abierto>

Gobierno de Puebla. (2020a). Datos Abiertos - Dataset. Recuperado 10 de julio de 2020, de <http://datos.puebla.gob.mx/search/type/dataset>

Gobierno de Puebla. (2020b). Desarrolladores. Recuperado 10 de julio de 2020, de <http://datos.puebla.gob.mx/desarrolladores>

GODAN (Global Open Data for Agriculture and Nutrition). (2017). *Success Stories*. Oxfordshire. Recuperado 20 de abril de 2019, de [https://www.godan.info/sites/default/files/documents/Success\\_Stories\\_2.pdf](https://www.godan.info/sites/default/files/documents/Success_Stories_2.pdf)

Grupo Banco Mundial. (2015). *Evaluación sobre la disponibilidad de Datos Abiertos en el Estado de Jalisco*. Guadalajara, Jal. Recuperado 12 de junio de 2019, de [https://iieg.gob.mx/contenido/proyectosiieginegi/Jalisco\\_ODRA\\_Junio\\_2015\\_RAS\\_MX\\_WB.pdf](https://iieg.gob.mx/contenido/proyectosiieginegi/Jalisco_ODRA_Junio_2015_RAS_MX_WB.pdf)

Hernández-Dominguez, A., & Hernández-Yeja, A. (2015). Acerca de la aplicación de MapReduce + Hadoop en el tratamiento de Big Data. Recuperado 13 de junio de 2020, de [http://scielo.sld.cu/scielo.php?pid=S2227-18992015000300004&script=sci\\_arttext&tlng=en](http://scielo.sld.cu/scielo.php?pid=S2227-18992015000300004&script=sci_arttext&tlng=en)

Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, M. del P. (2010). *Metodología de la Investigación* (Quinta edi). México: McGraw Hill.

INFOTEC (Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación). (2016a). Requerimientos de instalación. Recuperado 28 de septiembre de 2019, de [http://www.semanticwebbuilder.org.mx/es\\_mx/swb/Requerimientos\\_de\\_Instalacion](http://www.semanticwebbuilder.org.mx/es_mx/swb/Requerimientos_de_Instalacion)

INFOTEC (Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación). (2016b). SemanticWebBuilder. Recuperado 28 de julio de 2019, de [http://www.semanticwebbuilder.org.mx/es\\_mx/swb/SWB\\_Platform](http://www.semanticwebbuilder.org.mx/es_mx/swb/SWB_Platform)

Innova-T (Sitio web dedicado a innovación tecnológica). (2019a). ¿En que consiste la elicitación de requerimientos? Recuperado 26 de octubre de 2019, de <http://innova-t.co/lessons/en-que-consiste-la-elicitation-de-requerimientos/>

Innova-T (Sitio web dedicado a innovación tecnológica). (2019b). ¿Qué es un

- requerimiento no funcional? Recuperado 31 de julio de 2019, de <http://innovat.co/topic/que-es-un-requerimiento-no-funcional/>
- Kosack, S., & Fung, A. (2014). Does Transparency Improve Governance? *Annual Review of Political Science*, 17(1), 65-87. Recuperado 25 de agosto de 2019, de <https://doi.org/10.1146/annurev-polisci-032210-144356>
- Kucera, J., & Chlapek, D. (2014). Benefits and Risks of Open Government Data. *Journal of Systems Integration*, 30-41. Recuperado 22 de agosto de 2019, de <https://doi.org/10.20470/jsi.v5i1.185>
- Lamarca, M. J. (2019). Hacia la Web Semántica. Recuperado 28 de julio de 2019, de [http://www.hipertexto.info/documentos/web\\_semantica.htm](http://www.hipertexto.info/documentos/web_semantica.htm)
- Liferay. (2019). ¿Qué es un portal web? Recuperado 29 de marzo de 2019, de <https://www.liferay.com/es/resources/l/web-portal>
- México Digital. (2014). Guía de implementación de datos abiertos. Ciudad de México: datos.gob.mx. Recuperado 11 de septiembre de 2019, de [https://www.gob.mx/cms/uploads/attachment/file/31770/Guia-de-implementacion-de-datos-abiertos-v1\\_01.pdf](https://www.gob.mx/cms/uploads/attachment/file/31770/Guia-de-implementacion-de-datos-abiertos-v1_01.pdf)
- Moller, K. (2013). Lifecycle models of data-centric systems and domains. *Semantic web*, 4(1), 67-88. Recuperado 5 de febrero de 2019, de <https://content.iospress.com/articles/semantic-web/sw060>
- Muente-Kunigami, A., & Serale, F. (2018). Los datos abiertos en América Latina y el Caribe. Recuperado 13 de enero de 2019, de <https://doi.org/10.1080/13506280500162490>
- Naser, A., & Rosales, D. (2016). Panorama regional de los datos abiertos Avances y desafíos en América Latina y el Caribe. *Gestión pública. Instituto Latinoamericano y del Caribe de Planificación Económica y Social (ILPES)*, 86, 125. Recuperado 13 de enero de 2019, de <https://doi.org/10.1088/0953-8984/22/44/446001>
- Obama, B. (2009). MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES. SUBJECT: Transparency and Open Government (p. 1). The White House. Recuperado 11 de noviembre de 2018, de <https://obamawhitehouse.archives.gov/the-press-office/transparency-and->

open-government

- OCDE (Organización para la Cooperación y el Desarrollo Económicos). (2010). OCDE Building an open and innovative government for better policies and service delivery, Paris, 8-9 Junio, 10.
- Okuda, M., & Gómez-Restrepo, C. (2005). Métodos en investigación cualitativa: triangulación. *Revista Colombiana de Psiquiatría*, XXXIV(1), 119-124. Recuperado 22 de febrero de 2020, de <http://www.scielo.org.co/pdf/rcp/v34n1/v34n1a08.pdf>
- Open Data Watch. (2017). *OPEN DATA TO SUPPORT THE SUSTAINABLE DEVELOPMENT*. Recuperado 23 de agosto de 2019, de <https://opendatawatch.com/wp-content/uploads/2017/03/UNSC2017-OpenDataSupportsSDGs.pdf>
- Open Knowledge Network. (2017). Advancing the State of Open Data Through Dialogue. Recuperado 3 de abril de 2019, de <https://index.okfn.org>
- Parycek, P., Zuiderwijk, A., Charalabidis, Y., & Janssen, M. (2015). Critical Factors for Open Data Publication and Use : A Comparison of City-level , Regional , and Transnational Cases, (January 2016). Recuperado 10 de marzo de 2019, de <https://doi.org/10.29379/jedem.v7i2.397>
- Pollock, R. (2011). Building the (open) data ecosystem. Recuperado 3 de abril de 2019, de <https://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem/>
- Presidencia de la República (2015). Decreto por el que se establece la regulación en materia de Datos Abiertos, Pub. L. No. DOF: 20/02/2015, 2. México: Diario Oficial de la Federación. Recuperado 10 de diciembre de 2018, de [http://www.dof.gob.mx/nota\\_detalle.php?codigo=5382838&fecha=20/02/2015](http://www.dof.gob.mx/nota_detalle.php?codigo=5382838&fecha=20/02/2015)
- Purón, G. (2013). Las dimensiones del conocimiento en la adopción del e-gobierno: Un análisis factorial confirmatorio. *Gestión y Política Pública*, (449), 211-257.
- Rahemtulla, H., Kaplan, J., Gigler, B.-S., Cluster, S., Kiess, J., & Brigham, C. (2011). *Open Data Kenya*. Washington DC. Recuperado 25 de marzo de 2019, de [https://www.researchgate.net/publication/314236877\\_Open\\_Data\\_Kenya](https://www.researchgate.net/publication/314236877_Open_Data_Kenya)
- Ramos-Alderete, J. L. (2016). Gobierno Abierto : una visión de sus antecedentes , surgimiento y conceptualización, (24), 1-31.

- RDA (Research Data Alliance) Europe. (2014). The Data Harvest: How sharing research data can yield knowledge , jobs and growth, (December). Recuperado 6 de junio de 2020, de <https://rd-alliance.org/sites/default/files/attachment/The Data Harvest Final.pdf>
- Rothenberg, J. (2012). *Case study International Benchmark - Open data and use of standards*. Recuperado 8 de febrero de 2019, de [http://workspace.unpan.org/sites/internet/Documents/Internationale\\_benchmark\\_v1\\_03\\_final.pdf](http://workspace.unpan.org/sites/internet/Documents/Internationale_benchmark_v1_03_final.pdf)
- Sandoval-Almazán, R., Luna-Reyes, L. F., Luna-Reyes, D. E., Gil-García, J. R., Cid-Puron, G., & Picazo-Vela, S. (2017). *Building Digital Government Strategies: Principles and Practices*. Recuperado 4 de mayo de 2020 de, <https://doi.org/10.1007/978-3-319-60348-3>
- Sandoval, R. (2018). Ranking de portales estatales de Transparencia 2018. Recuperado 14 de enero de 2019, de <https://u-gob.com/ranking-de-portales-estatales-de-transparencia-2018/>
- Sandoval, R. (2019). Ranking de Transparencia 2019: en la encrucijada. Recuperado 10 de octubre de 2019, de <https://u-gob.com/ranking-de-transparencia-2019-en-la-encrucijada/>
- SITEUR (Sistema de Tren Eléctrico Urbano). (2017). Jalisco se confirma como capital de la innovación; recibe 10 de los 50 premios a las más innovadoras del sector público 2017. Recuperado 11 de febrero de 2019, de <http://www.siteur.gob.mx/noticias/item/jalisco-se-confirma-como-capital-de-la-innovación-recibe-10-de-los-50-premios-a-las-más-innovadoras-del-sector-público-2017.html>
- Steinberg, M., & Castro, D. (2017). The State of Open Data Portals in Latin America. Recuperado 13 de enero de 2019, de <https://www.datainnovation.org/2017/07/the-state-of-open-data-portals-in-latin-america/>
- Stellman, A., & Greene, J. (2005). *Applied software project management*. (Sebastopol, Ed.). O'Reilly Media.
- The World Bank. (2019). Technology Options. Recuperado 31 de julio de 2019, de

<http://opendatatoolkit.worldbank.org/en/technology.html>

Universidad Nacional de Quilmes. (1996). La difusión de la tecnología. *Redes*, III(8), 119-161. Recuperado 13 de noviembre de 2019, de <https://www.redalyc.org/pdf/907/90711321004.pdf>

Vera Martínez, M. (2018). Los datos abiertos y el plan de acción de gobierno abierto en México 2013-2015. *Nósis. Revista de Ciencias Sociales y Humanidades*, 27(54), 1-18. Recuperado 18 de abril de 2019, de <https://doi.org/10.20983/noesis.2018.2.1>

World Wide Web Foundation. (2016). Open Data How to Guides. Recuperado 8 de julio de 2019, de <https://webfoundation.org/research/open-data-how-to-guides/>

World Wide Web Foundation. (2019). Global Report. Recuperado 13 de enero de 2019, de <https://opendatabarometer.org/4thedition/report/>

Zuiderwijk, A. M. G. (2015). *Open Data Infrastructures The design of an infrastructure to enhance*. ('s-Hertogenbosch, Ed.). The Netherlands: Uitgeverij BOXPress. Recuperado 17 de marzo de 2019, de <https://repository.tudelft.nl/islandora/object/uuid%3A9b9e60bc-1edd-449a-84c6-7485d9bde012>

# ANEXOS

## ANEXO 1 – Guía de entrevista

1. Datos generales del entrevistado
2. Conocimiento general sobre los datos abiertos
3. Relación de los datos abiertos con la transparencia y gobierno abierto.
4. Conocimiento sobre fuentes de datos abiertos disponibles de manera local y general.
5. Utilidad de los datos abiertos.
6. Toma de decisión y personal involucrado en la generación y publicación de los datos abiertos en el entorno interno.
7. Datos disponibles en la dependencia o entidad gubernamental.
8. Requerimientos de presupuesto, personal e infraestructura para la generación y publicación de datos abiertos.
9. Identificación de riesgos.
10. Impedimentos legales.

## ANEXO 2 – Cuestionario de entrevista

1. ¿Cuál es su nombre?
2. ¿Cuál es su cargo?
3. ¿A cuál dependencia pertenece?
4. ¿Tiene conocimiento del tema de datos abiertos? ¿Qué sabe de ellos?
5. ¿Cuál considera que es la relación entre transparencia, datos abiertos y gobierno abierto? ¿Son lo mismo? ¿Cuál es la diferencia?
6. ¿Conoce datos abiertos que estén disponibles? ¿Cuáles? ¿En dónde se pueden encontrar?
7. ¿De qué manera pudieran ser de utilidad esos datos? ¿Para quién?
8. ¿Tiene facultad para decidir sobre cuáles datos abiertos pudieran generarse y publicarse en la dependencia o entidad?
9. ¿Cuál es el personal involucrado para ello?
10. ¿Cuáles datos abiertos pudieran generarse y publicarse en la dependencia o entidad en la que labora? – ¿Pueden tomarse los registros administrativos como base?
11. ¿Qué requiere para publicar estos datos? ¿Personal, infraestructura?
12. ¿Cuáles son los riesgos o inconvenientes que identifica?
13. ¿Existe algún impedimento legal para la apertura de los datos?

## ANEXO 3 – Información del área de Informática

Dependencia o entidad: \_\_\_\_\_

Nombre del responsable informático: \_\_\_\_\_

Componente	Situación	Responsable del tema
Seguridad informática		
Estándares utilizados		
Diseño de la red		
Organización de personal		
Procesos vigentes		
Infraestructura TIC disponible		
Costos en TIC		
Presupuesto disponible		
Normatividad vigente		
Capacitación técnica WEB		