



UNIVERSIDAD POLITÉCNICA DE PUEBLA

PROGRAMA ACADÉMICO DE POSGRADO

**Análisis de datos abiertos pertenecientes a
instituciones judiciales mediante vigilancia
tecnológica para su gestión en bases de datos
relacionales**

TESIS PARA OBTENER EL GRADO DE

MAESTRÍA EN INGENIERÍA EN GESTIÓN E INNOVACIÓN
TECNOLÓGICA

PRESENTA:

ING. JOSÉ JULIÁN BARRERA ROJAS

Asesora: Dra. María Auxilio Medina Nieto

Co-asesora: M. C. Rebeca Rodríguez Huesca

Juan C. Bonilla, Puebla, México. Septiembre 2024.



UNIVERSIDAD POLITÉCNICA DE PUEBLA

MAESTRÍA EN INGENIERÍA EN GESTIÓN E INNOVACIÓN TECNOLÓGICA

**Análisis de datos abiertos pertenecientes a
instituciones judiciales mediante vigilancia
tecnológica para su gestión en bases de datos
relacionales**

TESIS REALIZADA POR:

ING. JOSÉ JULIÁN BARRERA ROJAS

23 de Septiembre del 2024

Comité evaluador

(Firma)

Dra. María Auxilio Medina Nieto

M. C. Rebeca Rodríguez Huesca

M. I. Gudelia Pilar Pérez Conde

Dra. Gabriela Sánchez Esgua

Juan C. Bonilla, Puebla, México. Septiembre 2024.



UNIVERSIDAD POLITÉCNICA DE PUEBLA

MAESTRÍA EN INGENIERÍA EN GESTIÓN E INNOVACIÓN TECNOLÓGICA

Derechos de autor: Universidad Politécnica de Puebla 2024

El contenido se distribuye bajo los términos de la Licencia Abierta
Creative Commons (CC BY-NC-ND 4.0 MX)

(Firma)

.....
Ing. José Julián Barrera
Rojas

El uso estratégico de la información pública está dirigido a impulsar soluciones y mejorar la calidad de vida de las personas. La tesis implementa actividades de vigilancia tecnológica, estudia el contexto de la Política Nacional de Datos Abiertos aprobada en 2023, analiza datos abiertos de instituciones judiciales del estado de Puebla y se enfoca en un conjunto sobre denuncias y quejas en contra de servidores públicos pertenecientes a Secretaría de la Función Pública, evalúa la calidad inherente del conjunto con base en el estándar ISO/IEC 25012. Como resultado, la calidad inherente fue aceptable, 82.64%, a pesar de que los datos corresponden al periodo 2012 a 2018, se validó el cumplimiento de los principios FAIR y se identificó la implementación de 14 mejores prácticas conforme al Consorcio de la *World Wide Web*. Posteriormente se llevó a cabo una innovación tecnológica que transformó los datos sobre quejas y denuncias en un esquema relacional. La tesis describe el procesamiento de los datos, la implementación de consultas y la automatización de tareas que garantizaron integridad referencial y acceso eficiente. Los datos transformados están disponibles para usuarios de DA como docentes y estudiantes, constituyen un material didáctico de soporte a temas afines a la gestión de los datos y a las bases de datos relacionales. La difusión de los resultados se realizó en el marco del Certamen de Innovación en Transparencia 2024 y a través de la implementación de un sitio web evaluado por herramientas automáticas. A la par del desarrollo de la tesis, se obtuvieron los productos siguientes: 1) una infografía que presenta las características de la Política Nacional de Datos manera concisa, 2) la compilación del conjunto de datos transformado y 3) el sitio web denominado *difusiondelapnda*, disponible en: <http://difusiondelapnda.uppuebla.edu.mx:8321/>.

Palabras clave

Datos abiertos, Política Nacional de Datos Abiertos, ISO/IEC 25012, principios FAIR, innovación tecnológica

The strategic use of public information is aimed at promoting solutions and improving people's quality of life. This thesis implements technological surveillance activities, studies the context of the Mexican National Open Data Policy approved in 2023. The thesis analyzes open data of judicial institutions of the state of Puebla and focuses on a set on complaints and complaints against public servants belonging to the Secretariat of Public Service, evaluates the inherent quality of the set based on the ISO/IEC 25012 standard. As a result, the inherent quality was acceptable, 82.64%, despite the fact that the data correspond to the period 2012 to 2018, compliance with the FAIR principles was validated and the implementation of 14 best practices according to the World Wide Web Consortium. Subsequently, a technological innovation was carried out that transformed the data on complaints into a relational scheme. The thesis describes the processing of the data, the implementation of queries and the automation of tasks that guaranteed referential integrity and efficient access. The transformed data are available to users such as teachers and students, they constitute didactic material to support topics related to data management and relational databases. The results were disseminated within the framework of the 2024 Transparency Innovation Contest and through the implementation of a website evaluated by automatic tools. Along with the development of the thesis, the following products were obtained: 1) an infographic that presents the characteristics of the Mexican National Data Policy in a concise manner, 2) the compilation of the transformed data set and 3) the website called difusionde la PNDA, available at: <http://difusiondelapnda.uppuebla.edu.mx:8321>.

Keywords

Open data, Mexican National Open Data Policy, ISO/IEC 25012, FAIR principles, technological innovation.

1 Planteamiento del problema	9
1.1 Introducción	9
1.2 Objetivo general	12
1.3 Objetivos específicos	12
1.4 Justificación	12
1.5 Alineación con los Objetivos del Desarrollo Sostenible	16
2 Marco teórico	17
2.1 Fundamentos jurídicos del acceso a la información pública en México	17
2.1.1 Declaración Universal de Derechos Humanos	17
2.1.2 Constitución Política de los Estados Unidos Mexicanos	17
2.1.3 Ley General de Transparencia y Acceso a la Información Pública	19
2.1.4 Constitución Política del Estado Libre y Soberano de Puebla	19
2.2 Política Nacional de Datos Abiertos	20
2.2.1 Objetivo	20
2.2.2 Grupos estratégicos	20
2.2.3 Proceso de construcción	21
2.3 Marco legal para la publicación de DA en Puebla	21
2.3.1 Ley de Transparencia y Acceso a la Información Pública del Estado de Puebla	21
2.3.2 Ley del Sistema Anticorrupción del Estado de Puebla	22
2.3.3 Acuerdo General del Pleno del Consejo de la Judicatura del Poder Judicial del Estado de Puebla, por el que establece las disposiciones en Materia de Responsabilidades Administrativas	23
2.4 Instituciones judiciales de Puebla que publican DA	24
2.5 Modelo de calidad de datos ISO/IEC 25012	25
2.6 Origen y propósito de los principios FAIR	25
2.6.1 Principios rectores FAIR	26
2.6.2 Importancia de los principios FAIR y buenas prácticas	28
2.7 Innovación y vigilancia tecnológica	29
2.8 Trabajos relacionados	30

3 Metodología	36
3.1 Diseño de infografía	36
3.2 Vigilancia tecnológica e inteligencia competitiva	41
4 Resultados	43
4.1 Infografía de la PNDA	43
4.2 Evaluación de la calidad de DA-DENUNCIAS	43
4.2.1 Evaluación con el estándar ISO/IEC 25012	43
4.2.2 Cumplimiento de los principios FAIR	46
4.2.3 Evaluación de mejores prácticas para datos en la web	48
4.3 Transformación de DA-DENUNCIAS a DENUNCIAS-RELACIONAL	48
4.3.1 Implementación de consultas	52
4.3.2 Beneficios de DENUNCIAS-RELACIONAL	56
4.3.3 Automatización de tareas	57
4.3.4 Dificultades técnicas	58
4.3.5 Uso de DENUNCIAS-RELACIONAL como material didáctico	59
4.4 Difusión a través del sitio difusiondelapnda	59
4.4.1 Evaluación	60
4.4.1.1 Auto-evaluación por equipo de desarrollo	61
4.4.1.2 Resultados con herramientas abiertas de software	62
4.5 Productos de propiedad intelectual	66
5 Conclusiones	67
Referencias	69
A Instituciones judiciales del Estado de Puebla	75
B Búsqueda de DA en la PNT	81
C Procesamiento de DA-DENUNCIAS con OpenRefine	84
C.1 Exploración de datos	84
C.2 Transformaciones en DA-DENUNCIAS	86
C.2.1 Modificaciones	87
D Importación del conjunto DA-DENUNCIAS	89

E Creación de tablas y procedimientos almacenados para DENUNCIAS-RELACIONAL	93
E.1 Creación de la base de datos	93
E.2 Limpieza de datos	93
E.3 Creación de tablas padre	96
E.3.1 Tabla falta	96
E.3.2 Tabla institucion	97
E.3.3 Tabla medio	98
E.4 Creación de tablas hijas	99
E.4.1 Tabla denuncia	99
E.4.2 Tabla queja	101

1.1	Sección de la interfaz para descargar DA-DENUNCIAS.	11
1.2	Solicitudes en la PNT del 2023 para instituciones judiciales.	13
1.3	Solicitudes en la PNT de enero a julio del 2024 para instituciones judiciales.	14
2.1	Portada de la versión aprobada de la PNDA.	22
2.2	Siglas de los principios FAIR.	26
3.1	Infografía versión 1.	37
3.2	Infografía versión 2.	38
3.3	Infografía versión 3.	39
3.4	Encabezado del instrumento de recolección de datos.	40
3.5	Información solicitada a los participantes.	40
3.6	Etapas de vigilancia tecnológica.	42
4.1	Infografía para difundir la PNDA. Parte 1.	44
4.2	Infografía para difundir la PNDA. Parte 2.	45
4.3	Contradicción en el diccionario de datos para DA-DENUNCIAS	46
4.4	Metadatos del conjunto DA-DENUNCIAS.	46
4.5	Carga de datos en SQL server.	47
4.6	Modelo ER de DENUNCIAS-RELACIONAL.	49
4.7	Esquema relacional.	50
4.8	Tipos de falta con mayor número de denuncias.	53
4.9	Medios con mayor número de denuncias.	53
4.10	Instituciones con mayor número de denuncias.	54
4.11	Fecha de la denuncia más reciente.	55
4.12	Fecha de la denuncia más antigua.	55
4.13	Evidencia del uso de DENUNCIAS-RELACIONAL como material didáctico.	60
4.14	Página Index.html, inicio del sitio difusiondelapnda.	61
4.15	Página Index.html en resolución para dispositivo móvil.	63
4.16	Errores de la página Conceptos.html detectados por Stilus.	64
4.17	Resultados de la página Contacto.html por Nu HTML Checker.	65
4.18	Resultados de la página InstitucionesJ por W3C Link Checker.	65
B.1	Acceso a la sección de información pública de la PNT.	81

B.2	Íconos para los DA disponibles en la PNT.	82
B.3	Tarjeta informativa para un conjunto de DA en la PNT.	83
B.4	Ventaja emergente para seleccionar el formato de los datos.	83
C.1	Tipos de medio.	84
C.2	Tipos de falta.	85
C.3	Inconsistencia en la falta "Uso indebido de recursos públicos".	86
C.4	Número de instituciones con denuncias.	86
C.5	Número de registros sin tipo de falta: 2,563.	88
D.1	Creación de base de datos nueva.	89
D.2	Apertura del archivo sidecV2.sql.	90
D.3	Reemplazo de comillas dobles por comillas simples.	90
D.4	Eliminación de paréntesis izquierdo.	90
D.5	Archivo con formato csv en editor de texto.	91
D.6	Archivo con formato csv en hoja de cálculo.	91
D.7	Importación de csv a SQL Server.	92

2.1	Temas en la PNT para los DA de las instituciones judiciales.	33
2.2	Consejo de la Judicatura del Poder Judicial del Estado de Puebla	34
2.3	Tribunal Superior de Justicia	34
2.4	Instituciones judiciales que publican DA en la PNT.	35
2.5	Características de calidad inherente de los datos [14].	35
3.1	Descripción de participantes.	38
4.1	Porcentajes para las métricas de calidad de los datos inherente [14].	47
4.2	Mejores prácticas implementadas.	48
4.3	Mejores prácticas no implementadas.	51
4.4	Características del equipo de cómputo.	56
4.5	Datos del análisis heurístico.	61
4.6	Resultados del análisis heurístico.	62
4.7	Resultados de <i>PageSpeed Insights</i> en celular.	62
4.8	Resultados de <i>PageSpeed Insights</i> en computadora personal.	63
A.1	Archivo Judicial	76
A.2	Centro de Justicia Alternativa	76
A.3	Centro de Convivencia Familiar Supervisada	77
A.4	Escuela Estatal de Formación Judicial	78
A.5	Instituto de la Defensoría Pública	79
A.6	Tribunal de Justicia Administrativa	80

Acrónimos

DA	Datos abiertos
DA-DENUNCIAS	Conjunto de DA sobre quejas y denuncias de la Secretaría de la Función Pública
INAI	Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales
INEGI	Instituto Nacional de Estadística y Geografía
LGTAIP	Ley General de Transparencia y Acceso a la Información Pública
PNDA	Política Nacional de Datos Abiertos
PNT	Plataforma Nacional de Transparencia
SFP	Secretaría de la Función Pública
URI	Identificador de recursos uniforme, en inglés, <i>Uniform Resource Identifier</i>
W3C	Consortio <i>World Wide Web</i>

- API** Una interfaz de programación de aplicación, del inglés *Application Programming Interface*, hace referencia a un ambiente de software que soporta la comunicación e interacción entre aplicaciones o programas
- Innovación** Introducción de un nuevo o significativamente mejorado producto, proceso, método organizativo o sistema en una empresa u organización con el propósito de mejorar la eficiencia, la eficacia, la competitividad o la respuesta a las necesidades del mercado
- Stakeholder** El término en inglés *stakeholder* hace referencia a una persona u organización que tiene interés en una organización específica

1.1 Introducción

El concepto de *datos abiertos* (DA), (*open data*, en inglés), es una “filosofía y práctica que persigue que determinados tipos de datos estén disponibles de forma libre para todo el mundo, sin restricciones de derechos de autor, de patentes o de otros mecanismos de control [1]”. En la Carta Internacional de DA, en inglés, *Open Data Charter*, los DA se definen como “datos digitales que son puestos a disposición con las características técnicas y jurídicas necesarias para que puedan ser usados, reutilizados y redistribuidos libremente por cualquier persona, en cualquier momento y en cualquier lugar” que cumplen los principios siguientes [2]:

1. Abiertos por defecto
2. Oportunos y exhaustivos
3. Accesibles y utilizables
4. Comparables e interoperables
5. Para mejorar la gobernanza y la participación ciudadana
6. Para el desarrollo inclusivo y la *innovación*

De acuerdo con [8], las características de los DA son: accesibles, integrales, gratuitos, no discriminatorios, oportunos, permanentes, primarios, legibles por máquinas, en formatos abiertos y de libre uso. Los DA son insumos del movimiento de la ciencia abierta, la cual busca fomentar la colaboración y facilitar el acceso transparente al conocimiento. En América Latina, países como Argentina, Brasil, Chile, Colombia, Costa Rica, México y Perú, desde hace una década han implementando políticas públicas e iniciativas de acceso abierto y disponibilidad de datos para garantizar el acceso a investigaciones financiadas con fondos públicos [3].

Además de la investigación, el uso estratégico de la información pública está dirigido a impulsar soluciones y mejorar la calidad de vida de las personas. En México, el Instituto Nacional de Estadística y Geografía (INEGI), es el organismo público autónomo responsable de normar y coordinar el Sistema Nacional de Información Estadística y Geográfica, así como de captar y difundir información en cuanto al territorio, los recursos,

la población y economía, que permita dar a conocer las características del país y ayudar a la toma de decisiones [4].

A la fecha de elaboración de la tesis, los DA que publica el INEGI se distribuyen conforme a la Norma Técnica para el Acceso y Publicación de Datos Abiertos de la Información Estadística y Geográfica de Interés Nacional [5], éstos se organizan como sigue [6]:

1. Subsistema de Información Demográfica y Social
2. Subsistema de Información Económica
3. Subsistema de Información de Gobierno, Seguridad Pública e Impartición de Justicia

Como política pública del Sistema Nacional de Transparencia (SNT), el gobierno en turno emitió el “Plan Nacional de Socialización del Derecho de Acceso a la Información (PlanDAI) en 2023 [7], que, a través de la organización colectiva en los tres órdenes de gobierno y junto con la sociedad civil, pretende generar capacidades que incidan sobre la atención de problemas públicos y comunitarios en nuestro país [9]”. El objetivo general de este plan es que a través de la socialización del Derecho de Acceso a la Información (DAI) y en consecuencia, del aprovechamiento de la información pública por parte de la población, principalmente por grupos sociales en situación de vulnerabilidad, se contribuya a la inclusión y a la reducción de la brecha de desigualdad en el ejercicio de otros derechos fundamentales, además de buscar la adopción de un enfoque en derechos humanos y perspectiva de género.

La problemática que se aborda en la tesis es la siguiente. Actualmente, cada persona interesada (*stakeholder*) que accede a un sitio de DA necesita comprender cómo están organizados antes de analizarlos y procesarlos para recuperar información relevante. Estas tareas requieren de capacidades relacionadas con la gestión de los datos y las Tecnologías de la Información y Comunicación (TICs), propias de desarrolladores, administradores de bases de datos, científicos de datos, profesores y estudiantes, llevan un tiempo considerable y pudieran producir resultados distintos si no se realizan sistemáticamente.

La tesis implementa la guía de vigilancia e inteligencia tecnológica descrita en [10] para analizar el conjunto de DA perteneciente a la Secretaría de la Función Pública (SFP) disponible en:

<https://www.datos.gob.mx/busca/dataset/informacion-de-quejas-y-denuncias-ciudadanas>

La Figura 1.1 muestra los elementos de la interfaz para descargar este conjunto, en adelante, se emplea el acrónimo DA-DENUNCIAS para denotarlo.

The screenshot shows the 'Información de Quejas y Denuncias Ciudadanas' page on the SFP website. The page is titled 'Datos Abiertos' and features a 'Conjunto de datos' tab. A 'Reportar' button is located in the top right corner. The main content area displays the dataset name 'Quejas y Denuncias Ciudadanas' with a '.CSV' icon and a 'Descargar' button. Below this, there is a 'Más Información' button. The 'Metadatos' section contains a table with the following data:

Campo	Valor
Fuente	http://www.funcionpublica.gob.mx
Email del publicador	ldelgado@funcionpublica.gob.mx
Frecuencia	R/P6M
GUID	informacion-de-quejas-y-denuncias-ciudadanas
Idioma	es
Inicio del periodo temporal	2012-01-02T00:00:00Z

Figura 1.1: Sección de la interfaz para descargar DA-DENUNCIAS.

DA-DENUNCIAS consta de una tabla en formato CSV¹ compuesta por 7 columnas, la primera fila almacena los encabezados y 171,230 filas de datos en el periodo comprendido del 2 de Enero del 2012 al 2 de Diciembre del 2018, no se encontraron datos recientes. Con base en los datos descriptivos o metadatos, los datos provienen del Sistema Integral de Denuncias Ciudadanas (SIDECA), a la fecha de escritura de este documento, disponible en:

<https://sidec.funcionpublica.gob.mx/#/>

La tesis considera como marco legal a los elementos de la Política Nacional de Datos Abiertos (PNDA) [13], evalúa la calidad inherente de los datos utilizando el estándar

¹CSV corresponde a la expresión *Comma Separated Values*, valores separados por comas.

ISO/IEC 25012 [14] y describe cómo transformar los datos a una representación que aprovecha las características de los gestores de bases de datos relacionales. Los objetivos son los siguientes:

1.2 Objetivo general

Analizar los DA sobre denuncias ciudadanas pertenecientes a la SFP por medio de la implementación de actividades de vigilancia tecnológica para su reuso como material didáctico de temas afines a la gestión de los datos y a las bases de datos relacionales.

1.3 Objetivos específicos

1. Describir la Política Nacional de Datos Abiertos a través de una infografía para presentar sus características de manera concisa
2. Evaluar la calidad inherente de los datos de DA-DENUNCIAS utilizando el estándar ISO/IEC 25012 para valorar su reutilización
3. Mostrar el grado de cumplimiento de los principios FAIR y de mejores prácticas implementadas en DA-DENUNCIAS
4. Difundir los resultados de la evaluación de calidad y los datos transformados en un esquema relacional desde una página web para su reuso como material didáctico

1.4 Justificación

La evolución digital y el auge de la era de la información han posicionado a los DA como recursos para fortalecer la transparencia y la rendición de cuentas en las instituciones gubernamentales. En México, poner a disposición de la sociedad datos en formatos abiertos es el eslabón más débil de la agenda de apertura gubernamental, así lo demuestran los resultados de la Métrica de Gobierno Abierto [15], donde en una escala de valoración de 0 a 1, (1 implica que todas las oficinas de gobierno evaluadas publican al menos un conjunto de DA de calidad, mientras que 0 es indicativo de lo contrario), el promedio nacional para 2021 fue de 0.16, esto es, sólo seis centésimas por encima del valor obtenido en 2019 (0.10). Otro dato indica que el 76.3% de los 1,365 sujetos obligados analizados en la métrica, no publican DA [16].

A pesar de esos indicadores, los DA son un tema en la agenda nacional. Recientemente, el Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI) presentó la iniciativa *Abramos México*, que busca construir políticas en la materia como el PlanDAI y la Política Nacional de Datos Abiertos (PNDA) [13].

Por otro lado, existe un interés de la ciudadanía por acceder a la información generada en las instituciones judiciales como muestran las Figuras 1.2 y 1.3, sección DA de la Plataforma Nacional de Transparencia (PNT). En 2023 se recibieron y atendieron 699 solicitudes, mientras que del periodo entre el 2 de Enero al 7 de Agosto del 2024, se recibieron 290 solicitudes, se han atendido 246 y 44 se encuentran en proceso [17].

The screenshot shows the 'CONSULTA DE DATOS ABIERTOS' section of the PNT website. The search filters are as follows:

- Tipo de búsqueda ***: Solicitudes de información, Recursos de revisión, Obligaciones de transparencia
- Estado o Federación ***: Puebla
- Institución**: Poder Judicial del Estado de P...
- Folio**: (empty)
- Periodo ***: Rango de fechas
- Desde**: 02/01/2023
- Hasta**: 31/12/2023
- Tipo de respuesta**: ---Selecciona---
- Estatus**: ---Selecciona---

Buttons: LIMPIAR, CONSULTAR

Search Results Summary:

0	699	0
Sujetos Obligados	Solicitudes terminadas	Solicitudes en proceso

Additional options: SELECCIONE EL FORMATO DE EXPORTACIÓN E INTRODUZCA SU CORREO ELECTRÓNICO

Formato: --Selecciona--

Correo: correo@dominio.com

Logos: INAI, CAVINAI

Figura 1.2: Solicitudes en la PNT del 2023 para instituciones judiciales.

Datos abiertos - PNT

https://www.plataformadetransparencia.or...

INICIAR SESIÓN

PLATAFORMA NACIONAL DE TRANSPARENCIA

Buscar en Toda la plataforma

CONSULTA DE DATOS ABIERTOS

Tipo de búsqueda* Solicitudes de información Recursos de revisión Obligaciones de transparencia

Estado o Federación* Puebla

Institución Poder Judicial del Estado de P...

Folio

Periodo* Rango de fechas

Desde 02/01/2024 Hasta 07/08/2024

Tipo de respuesta --Selecciona--

Estatus --Selecciona--

LIMPIAR CONSULTAR

1	246	44
Sujetos Obligados	Solicitudes terminadas	Solicitudes en proceso

SELECCIONE EL FORMATO DE EXPORTACIÓN E INTRODUZCA SU CORREO ELECTRÓNICO

--Selecciona--

correo@dominio.com

CAVINAI

Figura 1.3: Solicitudes en la PNT de enero a julio del 2024 para instituciones judiciales.

La disponibilidad de información pública en datos accesibles y comprensibles de instituciones judiciales, no sólo impulsa la confianza entre los ciudadanos, sino que podría potenciar la participación activa de la sociedad en la supervisión y fiscalización de las actividades gubernamentales. Además, la apertura de DA constituye un vehículo para la eficiencia en la gestión gubernamental al facilitar la toma de decisiones basada en evidencia, optimizar procesos internos y mejorar la prestación de servicios.

La tesis describe la implementación de actividades de vigilancia tecnológica, enumera los temas de los DA de instituciones judiciales en el estado de Puebla, evalúa la calidad de DA-DENUNCIAS y muestra cómo transformarlos en un modelo relacional. Por tanto, entre los beneficiarios se encuentran en primer lugar estudiantes y docentes, en segundo

lugar, servidores públicos e investigadores que deseen analizar los datos transformados, generar reportes o gráficas. Finalmente, integrantes de organizaciones comunitarias o civiles. Se espera construir una visión integral sobre los DA y mostrar una alternativa del reuso de la información pública en contextos académicos.

1.5 Alineación con los Objetivos del Desarrollo Sostenible

La tesis está alineada con los Objetivo del Desarrollo Sostenible número 4 y 16, las metas 4.3 y 16.10, que a la letra, señalan lo siguiente [18], [19]:

ODS 4: “Garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje durante toda la vida para todos”

Meta 4.3: “De aquí a 2030, asegurar el acceso igualitario de todos los hombres y las mujeres a una formación técnica, profesional y superior de calidad, incluida la enseñanza universitaria”

ODS 16: “Paz, justicia e instituciones sólidas”

Meta 16.10: “Garantizar el acceso público a la información y proteger las libertades fundamentales, de conformidad con las leyes nacionales y los acuerdos internacionales”

El capítulo 2 está organizado como sigue. La sección 2.1 explica los fundamentos jurídicos del derecho al acceso a la información en México. La sección 2.2 expone los elementos de la PNDA. La sección 2.3 resume el marco legal para la publicación de DA en Puebla. La sección 2.4 describe las instituciones judiciales que publican DA en la PNT de manera regular. La sección 2.5 presenta el modelo de calidad de datos ISO/IEC 25012, son de interés las características de calidad de datos inherente. Posteriormente, se describen los principios FAIR en la sección 2.6. El tipo de innovación al que corresponde la tesis se explica en la sección 2.7. Finalmente, la sección 2.8 incluye los trabajos relacionados.

2.1 Fundamentos jurídicos del acceso a la información pública en México

2.1.1 Declaración Universal de Derechos Humanos

El acceso a la información en poder de cualquier sujeto obligado es un derecho humano, reconocido en la Declaración Universal de los Derechos Humanos [11], específicamente en el Artículo 19, que establece lo siguiente:

Todo individuo tiene derecho a la libertad de opinión y de expresión; este derecho incluye el de no ser molestado a causa de sus opiniones, el de investigar y recibir informaciones y opiniones, y el de difundirlas, sin limitación de fronteras, por cualquier medio de expresión.

2.1.2 Constitución Política de los Estados Unidos Mexicanos

De la declaración Universal de Derechos Humanos se desprende la adición del derecho de acceso a la información pública en el Artículo 6 de la Constitución Política de los Estados Unidos Mexicanos [12] que dice:

Toda persona tiene derecho al libre acceso a información plural y oportuna, así como a buscar, recibir y difundir información e ideas de toda índole por cualquier medio de expresión.

En síntesis, el anterior artículo¹ prevé que, para el ejercicio del derecho de acceso a la información, la Federación y las entidades federativas, en el ámbito de sus respectivas competencias, se regirán por los siguientes principios y bases:

- I.** Principio de máxima publicidad: Toda la información en posesión de cualquier autoridad, entidad, órgano y organismo de los Poderes Ejecutivo, Legislativo y Judicial, órganos autónomos, partidos políticos, fideicomisos y fondos públicos, así como de cualquier persona física, moral o sindicato que reciba y ejerza recursos públicos o realice actos de autoridad en el ámbito federal, estatal y municipal, es pública y sólo podrá ser reservada temporalmente por razones de interés público y seguridad nacional, en los términos que fijen las leyes.
- II.** La información que se refiere a la vida privada y los datos personales será protegida en los términos y con las excepciones que fijen las leyes.
- III.** Toda persona tendrá acceso gratuito a la información pública, a sus datos personales o a la rectificación de éstos.
- IV.** Se establecerán mecanismos de acceso a la información y procedimientos de revisión expeditos que se sustanciarán ante los organismos autónomos especializados e imparciales que establece esta Constitución.
- V.** Los sujetos obligados deberán preservar sus documentos en archivos administrativos actualizados y publicarán, a través de los medios electrónicos disponibles, la información completa y actualizada sobre el ejercicio de los recursos públicos y los indicadores que permitan rendir cuenta del cumplimiento de sus objetivos y de los resultados obtenidos.
- VI.** Las leyes determinarán la manera en que los sujetos obligados deberán hacer pública la información relativa a los recursos públicos que entreguen a personas físicas o morales.
- VII.** La inobservancia a las disposiciones en materia de acceso a la información pública será sancionada en los términos que dispongan las leyes.
- VIII.** La Federación contará con un organismo autónomo, especializado, imparcial, colegiado, con personalidad jurídica y patrimonio propio, con plena autonomía técnica, de gestión, capacidad para decidir sobre el ejercicio de su presupuesto y determinar

¹Consultar el Artículo 6 Apartado A, fracción VII

su organización interna, responsable de garantizar el cumplimiento del derecho de acceso a la información pública y a la protección de datos personales en posesión de los sujetos obligados en los términos que establezca la ley . . .

La ley establecerá aquella información que se considere reservada o confidencial.

2.1.3 Ley General de Transparencia y Acceso a la Información Pública

Los artículos 23, 49 y 50 de la Ley General de Transparencia y Acceso a la Información Pública (LGTAIP) [8], establecen quiénes están obligados a proporcionar acceso a la información y detallan la implementación y operación de la PNT, respectivamente:

Artículo 23. Son sujetos obligados a transparentar y permitir el acceso a su información y proteger los datos personales que obren en su poder: cualquier autoridad, entidad, órgano y organismo de los Poderes Ejecutivo, Legislativo y Judicial, órganos autónomos, partidos políticos, fideicomisos y fondos públicos, así como cualquier persona física, moral o sindicato que reciba y ejerza recursos públicos o realice actos de autoridad en los ámbitos federal, de las entidades federativas y municipal.

Artículo 49. Los organismos garantes desarrollarán, administrarán, implementarán y pondrán en funcionamiento la plataforma electrónica que permita cumplir con los procedimientos, obligaciones y disposiciones señaladas en la presente Ley para los sujetos obligados y organismos garantes, de conformidad con la normatividad que establezca el Sistema Nacional, atendiendo a las necesidades de accesibilidad de los usuarios.

Artículo 50. La Plataforma Nacional de Transparencia estará conformada por, al menos, los siguientes sistemas:

- I. Sistema de solicitudes de acceso a la información
- II. Sistema de gestión de medios de impugnación
- III. Sistema de portales de obligaciones de transparencia
- IV. Sistema de comunicación entre organismos garantes y sujetos obligados

2.1.4 Constitución Política del Estado Libre y Soberano de Puebla

La Constitución Política del Estado Libre y Soberano de Puebla [4] aborda el tema del acceso a la información pública en su Artículo 12, fracción VII, que dice:

Artículo 12. Corresponde a las leyes encargarse de:

- VII. Garantizar el acceso a la información pública en posesión de cualquier autoridad, entidad, órgano y organismo de los poderes ejecutivo, legislativo y judicial, órganos autónomos, partidos políticos, fideicomisos y fondos públicos, así como de cualquier persona física, moral o sindicato que reciba y ejerza recursos públicos o realice actos de autoridad en el ámbito estatal y municipal, así como de proteger los datos personales y la información relativa a la vida privada, en los términos y con las excepciones que establezca la Constitución Política de los Estados Unidos Mexicanos, la Ley General de Transparencia y Acceso a la Información Pública y la Ley aplicable a la materia.

2.2 Política Nacional de Datos Abiertos

La PNDA [13] establece las bases y la ruta para la apertura de datos a nivel nacional e incluye, entre otros elementos, un diagnóstico desde la óptica institucional y la social, problemáticas identificadas, alternativas de solución, objetivos, ejes transversales, poblaciones potenciales, así como responsables y sectores relevantes para la apertura de conjuntos de datos.

2.2.1 Objetivo

Guiar a los organismos garantes (en su calidad de coordinadores) y demás sujetos obligados para que avancen en el desarrollo de la agenda de DA en México.

2.2.2 Grupos estratégicos

A la fecha, se han identificado personas e instituciones que pueden contribuir en la construcción de la PNDA. En el marco de la estrategia *Abramos México* [9], se integraron los grupos estratégicos siguientes:

Grupo impulsor (GI). Coordina y supervisa las actividades de la estrategia, contribuye en la elaboración de documentos e insumos requeridos en el proceso de construcción de la Política Nacional

Grupo especializado (GE). Se conformó durante la Conferencia Nacional de Datos Abiertos 2022 [20], constó de siete mesas temáticas integradas por personas y organizaciones nacionales e internacionales de diferentes sectores que, de acuerdo con

su experiencia, especialización y capacidad de innovación en materia de DA, contribuyeron en la redacción de la política pública y aportaron elementos para la discusión y definición de los componentes y su contenido

Grupo técnico (GT). Se integra por enlaces de las instituciones y organizaciones que forman parte del GI, desarrolla trabajos de seguimiento permanente para avanzar en el cumplimiento de la estrategia

Coordinación del grupo técnico (CGT). Se encuentra a cargo de la Dirección General de Políticas de Acceso del INAI, la cual convoca a reuniones periódicas, propone actividades, lleva el registro y seguimiento de los acuerdos adoptados durante las reuniones

2.2.3 Proceso de construcción

Con el fin de alcanzar el objetivo, la construcción de la PNDA plantea de manera pública, abierta y colaborativa procesos, acciones y actores agrupados en las etapas siguientes:

1. *Preparación y apertura.* Instalación del GI elaboración del primer borrador de la política y otros insumos.
2. *Desarrollo de trabajos.* Trabajos del GT, integración de un GE, foros regionales, consulta pública y segundo borrador de la política
3. *Cierre y presentación.* Revisión y validación final hasta la publicación de la política pública en el marco del SNT

La Figura 2.1 muestra la portada de la versión aprobada de la PNDA [13].

2.3 Marco legal para la publicación de DA en Puebla

El marco legal para la publicación de DA en Puebla se integra por las leyes descritas en las secciones 2.3.1, 2.3.2 y 2.3.3.

2.3.1 Ley de Transparencia y Acceso a la Información Pública del Estado de Puebla

En la Ley de Transparencia y Acceso a la Información Pública del Estado de Puebla [26], en su Capítulo II denominado DE LA TRANSPARENCIA PROACTIVA, el Artículo 65



Figura 2.1: Portada de la versión aprobada de la PNDA.

dice que:

Adicionalmente a las obligaciones de transparencia, los sujetos obligados procurarán publicar información que sea de utilidad para la sociedad. Esta información deberá difundirse en los medios y formatos que más convengan al público al que va dirigida, siguiendo las políticas de transparencia proactiva que deberá establecer el Instituto de Transparencia y fomentando su reutilización. La información que se publique como resultado de las políticas de transparencia deberá permitir la generación de conocimiento público útil, para disminuir asimetrías de la información, mejorar los accesos a trámites y servicios, optimizar la toma de decisiones de autoridades o ciudadanos, y deberá tener un objeto claro, enfocado en las necesidades determinadas o determinables de sectores de la sociedad.

2.3.2 Ley del Sistema Anticorrupción del Estado de Puebla

La Ley del Sistema Anticorrupción del Estado de Puebla [27] en relación a la sistematización de la información y al combate a la corrupción menciona lo siguiente:

Artículo 8. El Comité Coordinador Estatal es la instancia responsable de establecer mecanismos de coordinación entre los integrantes del Sistema Estatal y tendrá

bajo su encargo el diseño, promoción y evaluación de políticas públicas de combate a la corrupción.

Artículo 9. El Comité Coordinador Estatal tendrá las facultades siguientes:

XII. Determinación de los mecanismos de suministro, intercambio, sistematización y actualización de la información que sobre estas materias generen las instituciones competentes de los órdenes de gobierno;

XIII. Establecer una Plataforma Digital Estatal que integre y conecte los diversos sistemas electrónicos que posean datos e información necesaria para que el Comité Coordinador Estatal pueda establecer políticas integrales, metodologías de medición y aprobar los indicadores necesarios para que se puedan evaluar las mismas, conforme a las disposiciones legales aplicables

XIV. Establecer una Plataforma Digital Estatal que integre y conecte los diversos sistemas electrónicos que posean datos e información necesaria para que las autoridades competentes tengan acceso a los sistemas a que se refiere el Título Cuarto de esta Ley, conforme a las disposiciones legales aplicables

Artículo 10. Son integrantes del Comité Coordinador Estatal:

V. Una persona representante del Consejo de la Judicatura del Poder Judicial del Estado de Puebla

VII. La persona que presida la Sala Especializada en materia de responsabilidades administrativas del Tribunal de Justicia Administrativa del Poder Judicial del Estado de Puebla

Artículo 49. La Plataforma Digital del Sistema Estatal estará conformada por la información que a ella incorporen las autoridades integrantes del Sistema Estatal y contará, al menos, con el sistema electrónico siguiente: *sistema de denuncias públicas de faltas administrativas y hechos de corrupción*

2.3.3 Acuerdo General del Pleno del Consejo de la Judicatura del Poder Judicial del Estado de Puebla, por el que establece las disposiciones en Materia de Responsabilidades Administrativas

El Acuerdo General del Pleno del Consejo de la Judicatura del Poder Judicial del Estado por el que establece las disposiciones en Materia de Responsabilidades Adminis-

trativas, publicado el 10 de noviembre de 2023 referente a quejas y denuncias en contra de servidores públicos, menciona lo siguiente [28]:

Artículo 8. Las quejas y denuncias podrán presentarlas cualquier persona o servidor público respectivamente que tenga conocimiento de los hechos u omisiones que pudieren constituir presuntas faltas administrativas.

Artículo 9. Las denuncias deberán contener los datos o indicios que permitan advertir la presunta responsabilidad administrativa por la comisión de faltas administrativas y podrán ser anónimas. La denuncia o queja debe ser presentada mediante escrito en la oficialía de partes del Consejo. En caso de que, se advierta de manera oficiosa con motivo de alguna visita, inspección o auditoría, indicios de conductas probablemente constitutivas de responsabilidad administrativa, de conformidad con lo previsto en el artículo 13 de este Acuerdo, se podrá iniciar la investigación.

2.4 Instituciones judiciales de Puebla que publican DA

Los DA han sido definidos como una herramienta que busca una mayor transparencia y eficiencia del servicio público, permitiendo, por un lado, reutilizar los datos públicos para generar nuevas actividades económicas y la rendición de cuentas ante los ciudadanos, por otro lado, devolver a los ciudadanos los datos que les son propios [21]. Los DA cubren temáticas diversas como el padrón de beneficiarios de un programa social, la geolocalización de la infraestructura vial ciclista, las características de las escuelas públicas o la incidencia de un delito en una zona determinada; el propósito general, de acuerdo con [16], es que cualquiera pueda aprovecharlos según sus intereses o necesidades.

Las instituciones judiciales del estado de Puebla que publican DA se rigen por la Ley Orgánica del Poder Judicial del Estado de Puebla, ésta “tiende a estructurar adecuadamente la conformación de los órganos que integran al poder que regula, delimitando sus funciones, proveyéndolo de mayor autonomía para determinar sobre su labor y administración, actualizando áreas que la tecnología ha integrado para el mejor desarrollo de la función que le es propia [25]”. Las tablas 2.2 y 2.3 describen las funciones de las dos organizaciones principales siguientes:

1. *Consejo De La Judicatura Del Poder Judicial Del Estado De Puebla*, sus organizaciones auxiliares son:

- Archivo Judicial

- Centro de Convivencia Familiar Supervisada
- Centro de Justicia Alternativa
- Escuela Estatal de Formación Judicial
- Instituto de la Defensoría Pública
- Tribunal de Justicia Administrativa

2. Tribunal Superior de Justicia (TSJ)

El Anexo A describe la misión y funciones de las organizaciones auxiliares del Consejo De La Judicatura Del Poder Judicial Del Estado De Puebla. La Tabla 2.4 enumera las organizaciones que publican DA en la PNT. La información corresponde a los temas de la Tabla 2.1. El proceso de búsqueda de los datos se incluye en el Anexo B.

2.5 Modelo de calidad de datos ISO/IEC 25012

El modelo de calidad de datos ISO/IEC 25012 establece las características de calidad que se deben tener en cuenta al evaluar las propiedades de un producto de datos determinado [14], está compuesto por 15 características organizadas en tres categorías:

1. Calidad de datos dependiente del sistema
2. Calidad inherente de los datos
3. Calidad inherente de los datos y dependiente del sistema

La Tabla 4.1 contiene la definición de las características de calidad inherente de los datos, las características restantes (accesibilidad, conformidad, confidencialidad, eficiencia, precisión, trazabilidad, comprensibilidad, disponibilidad, portabilidad y recuperabilidad) están fuera del alcance de la tesis puesto que hacen referencia al uso de los datos dentro de un sistema de información.

2.6 Origen y propósito de los principios FAIR

En el ecosistema de la ciencia, el reto de hacer posible el uso eficiente de los datos y métodos de investigación es complejo e incluye a múltiples *stakeholders*. De acuerdo con [30], el análisis computacional para descubrir patrones significativos en conjuntos de datos masivos e interconectados se está convirtiendo en una actividad de investigación rutinaria; uno de los retos es proporcionar datos legibles por máquinas como insumo

principal para el descubrimiento de conocimiento y para que los procesos científicos funcionen sin problemas y de forma sostenible.

En enero de 2014, diferentes *stakeholders* se reunieron a petición del *Netherlands eScience Center* y el *Dutch Techcentre for the Life Sciences* (DTL) en el *Lorentz Center* de *Leiden* (Países Bajos) para reflexionar y debatir sobre cómo seguir mejorando este ecosistema. De estos debates surgió la idea de que, a través de la definición y el apoyo generalizado de un conjunto mínimo de principios y prácticas rectores acordados por la comunidad, los proveedores y consumidores de datos, tanto máquinas como personas, podrían descubrir, acceder, interoperar y reutilizar más fácilmente, con la debida citación, la información generada [30].

El 15 de marzo de 2016, se publicaron formalmente los Principios FAIR para el manejo y administración de datos científicos en la revista *Scientific Data* del *Nature Publishing Group*. Las siglas corresponden a los términos en inglés *Findable*, *Accessible*, *Interoperable* y *Reusable*, (ver la Figura 2.2) [31].

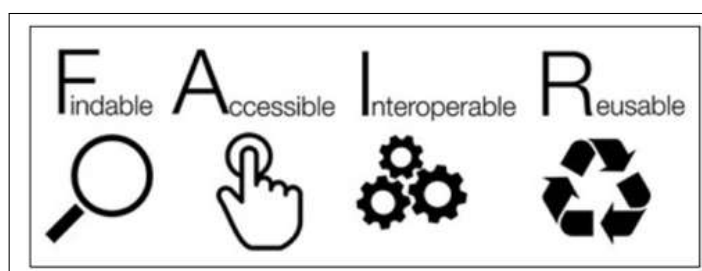


Figura 2.2: Siglas de los principios FAIR.

Los principios FAIR hacen referencia a tres aspectos [32]:

Datos: cualquier objeto digital

Metadatos: información que describe a los datos

Infraestructura: herramientas y sistemas, ofrecen “un conjunto de cualidades precisas y mensurables que una publicación de datos debería exhibir, cualidades que garanticen que los datos están localizados, son accesibles, interoperables y reutilizables [31]”

2.6.1 Principios rectores FAIR

Desde la perspectiva de la gestión de datos, los principios FAIR proporcionan una orientación gradual para avanzar hacia el objetivo final de alcanzar los conceptos especi-

ficados por sus siglas en los datos o metadatos, sin ambigüedad para las máquinas y las personas. Los principios rectores FAIR se especifican como sigue [30]:

1. Que un objeto de datos se encuentre (ser “*findable*”) implica que debe ser identificable de manera única y persistente
 - 1.1 El mismo objeto de datos debe ser localizable en cualquier momento, por lo tanto, los objetos de datos deben ser persistentes, con énfasis en sus metadatos
 - 1.2 Un objeto de datos debe contener mínimamente metadatos básicos “accionables” por máquina que le permitan distinguirse de otros objetos de datos
 - 1.3 Por lo tanto, los identificadores para cualquier concepto utilizado en los objetos de datos deben ser únicos y persistentes
2. Los datos son *accesibles* siempre y cuando puedan ser obtenidos por máquinas y personas
 - 2.1 Con la autorización apropiada
 - 2.2 A través de un protocolo bien definido
 - 2.3 Por lo tanto, las máquinas y las personas podrán juzgar la accesibilidad real de cada objeto de datos
3. Los objetos de datos pueden ser *interoperables* sólo si:
 - 3.1 Los metadatos se pueden procesar por las máquinas
 - 3.2 Los formatos de los metadatos utilizan vocabularios u ontologías comunes
 - 3.3 Los metadatos dentro del objeto de datos deben ser analizados sintáctica y semánticamente por las máquinas
4. Para que los objetos de datos sean *reusables* deben cumplir con los siguientes criterios adicionales
 - 4.1 Los objetos de datos deben cumplir con los principios *findable* e *interoperable*
 - 4.2 Los metadatos deben estar bien detallados y escritos para que puedan vincularse o integrarse automáticamente con otras fuentes de datos
 - 4.3 Los objetos de datos publicados deben referirse a sus fuentes con metadatos y procedencia bien detallados y escritos para permitir una citación adecuada

2.6.2 Importancia de los principios FAIR y buenas prácticas

[31] describe la importancia de los principios FAIR en el ámbito científico y de investigación como se indica a continuación:

Findable: Tanto las personas como las máquinas puedan encontrar los datos fácilmente, esto es, que sean localizados. Para ello, los DA se distinguen por:

- Recibir un identificador único y persistente, por ejemplo, un *Digital Object Identifier* (DOI)
- Organizar la descripción del conjunto de datos (metadatos) en campos específicos. Los metadatos se almacenan en algún repositorio o base de datos conocido o que pueda ser rastreado por un motor de búsqueda

Accessible: Una vez localizados los datos, el usuario necesita saber cómo accederlos, por lo que esta característica incluye información sobre autenticación y/o autorización de acceso. Los DA satisfacen lo siguiente:

- Los programas y protocolos informáticos de acceso son abiertos, no propietarios
- Los metadatos se conservan aunque el conjunto de datos de investigación desaparezca y facilitan el contacto con sus responsables
- Se puede implementar un registro de usuario y distintos permisos de acceso según la tipología de los datos

Interoperable: Los datos suelen estar integrados con otros datos, por lo que deben ser interoperables con otras aplicaciones. La interoperabilidad se relaciona con lo siguiente:

- Los archivos y los datos deben representarse en lenguajes y esquemas informáticos estandarizados capaces de ser leídos por máquinas
- La descripción de los conjuntos de datos se realiza mediante lenguajes informáticos controlados e identificadores únicos y persistentes
- Los conjuntos de datos se enlazan con otros que los completan y se explica en qué consiste esa relación

Reusable: El fin último de los principios FAIR es la reutilización de los datos de investigación, por eso, se requiere que los datos y los metadatos estén bien escritos y que se puedan replicar y/o combinar con otros conjuntos de datos.

Como buenas prácticas para publicar DA, en [31] se consideran las siguientes:

- La información sobre los datos es detallada y relevante
- La información sobre derechos y limitaciones de uso es explícita
- La descripción incluye el origen de los datos, sus responsables, si se han publicado previamente o no
- Los datos y los metadatos están organizados según estándares de descripción, almacenamiento e intercambio

2.7 Innovación y vigilancia tecnológica

La tesis adopta las definiciones de innovación y vigilancia tecnológica siguientes:

Innovación. “Introducción de un nuevo o significativamente mejorado producto, proceso, método organizativo o sistema en una empresa u organización con el propósito de mejorar la eficiencia, la eficacia, la competitividad o la respuesta a las necesidades del mercado [33]”

Vigilancia tecnológica. “Es un proceso organizado, selectivo y permanente, basado en la captura de información del exterior y de la propia organización sobre ciencia y tecnología en un determinado sector de interés. La información recuperada posteriormente es seleccionada, analizada, difundida y comunicada a los decisores para convertirla en los conocimientos necesarios y suficientes para tomar decisiones con menor riesgo y poder anticiparse a los cambios [34]”.

En [36], la vigilancia tecnológica se considera herramienta estratégica para la gestión de la innovación, se trata de un proceso sistemático y organizado de detección, selección y análisis de información externa e interna de una organización, referida a aspectos tecnológicos, científicos, comerciales, de mercado, legales o de innovación que permiten mejorar la toma de decisiones, reducir riesgos y anticipar los cambios o crisis. De acuerdo con [35], la vigilancia tecnológica trata de analizar el comportamiento innovador de los competidores directos e indirectos, explorar todas las fuentes de información (libros, bases de datos, patentes, etcétera), examinar los productos existentes en el mercado (tecnología incorporada), asistir a ferias y congresos para posicionarse respecto a los demás competidores y obtener así conocimiento de las competencias tecnológicas que predominarán en un futuro más o menos próximo.

La tesis implementa actividades de vigilancia tecnológica al conjunto DA-DENUNCIAS, propone una innovación tecnológica para gestionar los datos.

2.8 Trabajos relacionados

El Barómetro Global de Datos [22] resume el contexto del uso y la disponibilidad de los DA en 109 países, incluye indicadores sobre gobernanza y capacidades, identifica como usuarios de DA a los actores siguientes: medios de comunicación (36.4%), organizaciones de la sociedad civil (37.6%), los propios gobiernos (30.3%) y la academia (25.7%). En el caso de México, [23] presenta un conjunto de recomendaciones para transparentar la información del gobierno, señala que los DA brindan información sobre el funcionamiento de la seguridad y la justicia, así como la protección de los derechos humanos.

La Política de Transparencia, Gobierno Abierto y Datos Abiertos de la Administración Pública Federal de la SFP [29], integra prioridades, acciones y ejes estratégicos alineados al Plan Nacional de Desarrollo 2019-2024, su cobertura es: secretarías de estado, consejería jurídica, órganos reguladores coordinados en materia energética, organismos descentralizados, empresas de participación estatal, instituciones nacionales de crédito, organizaciones auxiliares nacionales de crédito e instituciones nacionales de seguros y de fianzas, fideicomisos. Los ejes estratégicos son:

1. Fortalecer la transparencia en la gestión pública
2. Promover un gobierno abierto
3. Impulsar el uso de DA

Esta sección expone iniciativas relacionadas con la PNDA, la evaluación de la calidad de DA y aplicaciones de vigilancia tecnológica. Entre las iniciativas se destacan las siguientes:

Infraestructura abierta. Iniciativa de transparencia y DA que promueve la adopción del Estándar de Datos de Contrataciones Abiertas (EDCA) y del estándar de datos de contrataciones abiertas para infraestructura. Busca contribuir a la generación, administración y divulgación de datos y documentos sobre los proyectos de infraestructura y las contrataciones que se llevan a cabo durante el desarrollo de éstos [24].

Contrataciones abiertas. Es una iniciativa internacional promovida por el *Open Contracting Partnership* (OCP), en español EDCA, que tiene como objetivo incrementar la oferta y la calidad de la información relacionada con los procedimientos de contratación que realizan los gobiernos. La iniciativa propicia procesos de participación entre diferentes actores incluyendo sociedad civil, periodistas y el sector privado con el fin de mejorar servicios y obras públicas [37]

DATACON. La conferencia anual DATACON 2023 es un encuentro que busca reunir a instituciones, sociedad civil, academia, sector empresarial y comunidad interesada en la agenda de DA para dialogar, colaborar y construir capacidades rumbo a la implementación de la PNDA [38]

Publicidad abierta. Iniciativa que tiene el objetivo de transparentar los montos destinados al gasto en comunicación social y publicidad oficial de las instituciones públicas [38]

Los DA tienen utilidad en el ámbito educativo, por ejemplo, en [39] se utilizan para identificar Instituciones de Educación Superior del Estado de Puebla que publican información relacionada con ciencia y tecnología (p. 61), disponible en:

<https://www.mauxmedina.com/tesisAri/Inicio.html>.

Trabajos relacionados con la aplicación del estándar ISO/IEC 25012 son: [40] propone una serie de pasos para evaluar la calidad de los datos sobre farmacias, la información está almacenada en una base de datos relacional sujeta a las políticas de acceso abierto, [41] evalúa la calidad de los datos del Módulo sobre Ciberacoso (MOCIBA) del 2021 del INEGI.

El uso y la normativa de la vigilancia tecnológica se presenta [42] y [43], respectivamente. En el primero, se resume la norma UNE 166006:2011 que regula la gestión de la vigilancia tecnológica e inteligencia competitiva en las organizaciones; su propósito es sistematizar la captura, selección, análisis y difusión de información para la toma de decisiones y seguimiento de resultados. La norma contiene los capítulos clave siguientes:

Capítulo 4. *Requisitos del sistema.* Establece la importancia de la confidencialidad, legalidad y ética en el manejo de la información entre la organización y sus *stakeholders*.

Capítulo 5. *Responsabilidades de la dirección.* Señala que la dirección requiere designar un representante para la vigilancia tecnológica e inteligencia competitiva.

Capítulo 6. *Gestión de recursos.* Detalla las competencias necesarias para el personal involucrado, incluyendo habilidades en bases de datos y minería de textos científicos

Capítulo 7. *Realización de la vigilancia.* Describe el proceso de vigilancia, desde la identificación de necesidades hasta la obtención de resultados

Capítulo 8. *Contratación de servicios.* Establece los requisitos para contratar servicios de apoyo en vigilancia tecnológica e inteligencia competitiva

Capítulo 9. *Medición, análisis y mejora.* Aborda el control de desviaciones y la reutilización de la información para mejorar procesos futuros

Por otro lado, en [43] se indica que la vigilancia tecnológica es fundamental para mantener informados a los expertos en desarrollo, gestión y mantenimiento de aplicaciones web sobre las principales vulnerabilidades, alertas de seguridad y prácticas actualizadas, con el objetivo de minimizar los incidentes de ciberseguridad. Instituciones públicas y empresas del sector privado utilizan los portales y aplicaciones web para publicar contenidos de diversos tipos, utilizar el correo institucional, realizar comercio electrónico, incluso para ofrecer un servicio a la población. Sin embargo, esto representa un problema debido a que deja abierta la puerta a los ataques informáticos que puede realizar un ciberdelincuente con la finalidad de comprometer los datos, interrumpir el servicio que se ofrece o realizar un secuestro de la información para conseguir dinero. Por lo que para la gestión de archivos físicos, electrónicos y digitales sea adecuadas, los autores de [43] proponen considerar temas relacionados con la vigilancia tecnológica y la ciberseguridad, específicamente para gestionar los datos clasificados como sensibles tales como los datos biométricos.

No.	Tema
1	Bienes y servicios
2	Catálogo y guía de archivos
3	Comité de transparencia
4	Concursos para ocupar cargos públicos
5	Condiciones generales de trabajo y sindicatos
6	Contratos de obras
7	Contratos por honorarios
8	Convenios de coordinación
9	Curricula de funcionarios
10	Declaraciones patrimoniales
11	Dictámenes financieros
12	Directorio
13	Donaciones
14	Estadísticas
15	Estructura orgánica
16	Funciones de áreas
17	Gastos de publicidad oficial
18	Gastos en comisiones oficiales
19	Indicadores de interés público
20	Indicadores de resultados
21	Información de interés público
22	Información financiera
23	Informes
24	Ingresos
25	Inventario de bienes
26	Jubilados y pensionados
27	Metas y objetivos de las áreas
28	Normatividad
29	Padrón proveedores y contratistas
30	Presupuesto del gasto público
31	Puestos y vacantes
32	Requisitos y formatos y unidad de transparencia
33	Resultados de auditorías
34	Servicios públicos
35	Servidores públicos sancionados
36	Solución de procesos en juicio
37	Sueldos
38	Tablas de aplicabilidad
39	Trámites

Tabla 2.1: *Temas en la PNT para los DA de las instituciones judiciales.*

Funciones:	El Poder Judicial tiene como objeto consolidar la justicia integral en nuestro Estado y orienta su función pública a través de los valores democráticos de legalidad, rendición de cuentas, honradez, transparencia, publicidad, accesibilidad, independencia e imparcialidad
Misión:	Impartir y administrar justicia de forma imparcial, expedita y cercana, enalteciendo la protección de los Derechos Humanos y garantizando certeza jurídica, con procesos eficaces, eficientes, transparentes, tecnológicos e informáticos que coadyuvan al combate a la corrupción fomentando una cultura de respeto, seguridad, armonía y proximidad social.
Visión:	Posicionar al Poder Judicial como referente nacional, con procesos innovadores que fortalezcan la credibilidad y confianza de la ciudadanía a través de la utilización óptima de los recursos disponibles, la mejora continua y el constante desarrollo humano y profesional de sus integrantes, garantizando el acceso a la justicia.
Página web:	https://www.pjpuebla.gob.mx/

Tabla 2.2: Consejo de la Judicatura del Poder Judicial del Estado de Puebla

Funciones:	Administrar la justicia a través de los tribunales y mecanismos alternativos para la solución de controversias con total apego a la legalidad y sujetas a los principios de imparcialidad, objetividad, excelencia, celeridad, profesionalismo e independencia para lograr un sistema de justicia sólido y confiable.
Página web:	https://pjpuebla.gob.mx/poder-judicial/tribunal-superior-de-justicia/quienes-somos

Tabla 2.3: Tribunal Superior de Justicia

No.	Institución judicial
1	Archivo Judicial
2	Centro de Convivencia Familiar Supervisada
3	Centro de Justicia Alternativa
4	Consejo de la Judicatura del Poder Judicial del Estado de Puebla
5	Instituto de Estudios Judiciales
6	Tribunal de Justicia Administrativa
7	Tribunal Superior de Justicia

Tabla 2.4: *Instituciones judiciales que publican DA en la PNT.*

Característica	Definición
Exactitud	Grado en el que los datos representan correctamente el verdadero valor del atributo deseado de un concepto o evento en un contexto de uso específico
Compleitud	Grado en el que los datos están asociados con una entidad, tienen valores para todos los atributos esperados e instancias de entidades relacionadas en un contexto de uso específico
Consistencia	Grado en el que los datos están libres de contradicción, y son coherentes con otros datos en un contexto de uso específico. Puede ser analizada en datos que se refieran tanto a una como a varias entidades comparables
Credibilidad	Grado en el que los datos tienen atributos que se consideran ciertos y creíbles en un contexto de uso específico. La credibilidad incluye el concepto de autenticidad (la veracidad de los orígenes de datos, atribuciones, compromisos)
Actualidad	Grado en el que los datos tienen atributos con la edad correcta en un contexto de uso específico

Tabla 2.5: *Características de calidad inherente de los datos [14].*

El capítulo 3 presenta el diseño de la infografía y describe las actividades de vigilancia tecnológica implementadas.

3.1 Diseño de infografía

Para determinar los elementos de una infografía que difunda la PNDA, se llevó a cabo un experimento en el que se contó con la participación de 35 estudiantes de sexto semestre de Ingeniería en Tecnologías de la Información en la Universidad Politécnica de Puebla (UPPue) seleccionados a conveniencia. Se propuso obtener la percepción de los participantes de las versiones de infografía mostradas en las Figuras 3.1, 3.2 y 3.3, respectivamente.

Los datos se recolectaron utilizando un instrumento formado por preguntas cerradas con orientación positiva y negativa, cuya respuesta se solicitó indicar mediante la selección de un valor en una escala de Likert de 5 elementos, (1 valor mínimo, 5 valor máximo). Las preguntas fueron las siguientes:

1. ¿Qué tan clara es la información?
2. ¿Qué tan difícil es la lectura?
3. ¿Cómo es la organización?
4. ¿Qué tan adecuados le parecen los colores?
5. ¿Qué tanto le desagrada el tamaño de la letra?
6. ¿Qué tan difícil es leer con los estilos de letra utilizados?
7. ¿Qué tanto se relacionan las imágenes con el contenido?
8. Tras una primer mirada, ¿le queda claro cuál es el propósito de la infografía?
9. ¿Qué información es innecesaria?

Al instrumento se integró la pregunta abierta siguiente: “¿Qué mejora haría a la infografía?”, éste se implementó como formulario de la herramienta Google Forms, las Figuras 3.4 y 3.5 muestran el encabezado y la información solicitada por participante.

Los participantes se organizaron en los 3 grupos descritos en la Tabla 3.1. Por correo electrónico se enviaron los enlaces a los 3 formularios, uno para cada versión. Los



Figura 3.1: *Infografía versión 1.*

formularios se contestaron en sesiones presenciales en los laboratorios de cómputo, la duración por sesión fue de 30 minutos, aproximadamente.

Posterior a la recolección de los datos y a su análisis, la versión 1 obtuvo una calificación de 6.5, la versión 2 de 7.4 y la versión 3 de 5.7, por tanto, la segunda se utilizó como base para elaborar la versión final que se describe en la sección 4.1.



Figura 3.2: Infografía versión 2.

Grupo	Género	Rango de edad (años)	Uso de internet por día (horas)
1	2 mujeres, 4 hombres	5 entre 20 y 25, 1 menor de 20	5 entre 4 y 8, 1 más de 8
2	5 mujeres, 9 hombres	12 entre 20 y 25, 1 entre 26 y 30, 1 mayor a 30	9 entre 4 y 8, 4 más de 8 1 menos de 4
3	4 mujeres, 11 hombres	14 entre 20 y 25, 1 entre 26 y 30	5 entre 4 y 8, 1 más de 8

Tabla 3.1: Descripción de participantes.



Figura 3.3: Infografía versión 3.

Evaluación de infografía sobre la Política Nacional de Datos Abiertos: versión 2

Este cuestionario tiene el propósito de recolectar datos sobre la percepción de una infografía que describe la Política Nacional de Datos Abiertos. Las respuestas se publicarán de forma anónima y se utilizarán como parte de una actividad de investigación del Lic. **José Julián Barrera R.**, estudiante de la Maestría en Gestión e Innovación Tecnológica (MGIT) de la Universidad Politécnica de Puebla (UPPue).

Apoyo para la evaluación: Dra. María Auxilio Medina Nieto, M. C. Rebeca Rodríguez Huesca

jose.barrera2205@uppuebla.edu.mx [Cambiar de cuenta](#)

Correo electrónico *

Registrar jose.barrera2205@uppuebla.edu.mx como el correo que se incluirá al enviar mi respuesta

Información del (a) participante
Instrucciones: favor de seleccionar la opción que lo describa de la mejor manera como participante

Figura 3.4: Encabezado del instrumento de recolección de datos.

Seleccione su categoría sexual:

Hombre

Mujer

Prefiero no decir

Otro: _____

Seleccione su rango de edad:

Menor de 20 años

Entre 20 y 25 años

Entre 26 y 30 años

Mayor a 30 años

Indique el número de horas que utiliza internet en una jornada diaria

Menos de 4 horas

Entre 4 y 8 horas

Más de 8 horas

Figura 3.5: Información solicitada a los participantes.

3.2 Vigilancia tecnológica e inteligencia competitiva

En la tesis se implementó la guía práctica de vigilancia tecnológica e inteligencia competitiva propuesta por el Observatorio Tecnológico de la Universidad de Alicante (OVTT) [10], las actividades realizadas se agruparon en las etapas siguientes:

Etapas 1: *Diagnóstico y priorización.* Se analizaron los datos de DA-DENUNCIAS

Etapas 2: *Búsqueda y captura de información.* Los DA sobre denuncias se obtuvieron de la SFP, las palabras clave fueron: *denuncias, DA, ciudadanos, ciudadanas, servidores públicos.*

Etapas 3: *Análisis de información.* DA-DENUNCIAS se procesó en primer término con la herramienta OpenRefine [50], (ver el Anexo C).

Etapas 4: *Valorización de información relevante.* Se evaluó la calidad de DA-DENUNCIAS utilizando el estándar ISO/IEC 25012 [14], así como el grado de cumplimiento en relación a los principios FAIR [31] y la implementación de buenas prácticas para datos en la web. Posteriormente, se transformaron los datos para importarse en los gestores de bases de datos MySQL¹, MariaDB² y SQL Server versión 14³.

Etapas 5: *Difusión y comunicación.* Por un lado, se elaboró una infografía para introducir la PNDA al público en general [45], ésta se distribuye gratuitamente desde el *repositorio institucional* de la UPPue. Por otro lado, como resultado del procesamiento y eliminación de redundancia en DA-DENUNCIAS, se produjo la obra derivada titulada “Esquema relacional para gestión de datos sobre quejas y denuncias de la Secretaría de la Función Pública” [46], disponible en:

<https://www.mauxmedina.com/difusiondelaPNDA/index.html>

Etapas 6: *Toma de decisiones y acciones.* [46] se utilizó como material didáctico en la asignatura “Bases de datos avanzadas” de la Maestría en Sistemas y Cómputo Inteligente de la UPPue durante una sesión virtual de clase, ver sección 4.3.5. Como acción, se determinó participar en el Certamen de Innovación en Transparencia (CIT 2024) a través del proyecto titulado “Gestión de datos sobre denuncias con herramientas abiertas de software”. Como medios de difusión se elaboraron dos videos, uno para la etapa de presentación [47] y otro para la exposición [48].

¹<https://www.mysql.com/>

²<https://mariadb.org/>

³<https://www.microsoft.com/es-mx/sql-server/sql-server-downloads>

La Figura 3.6 sintetiza las actividades principales de vigilancia tecnológica implementadas.

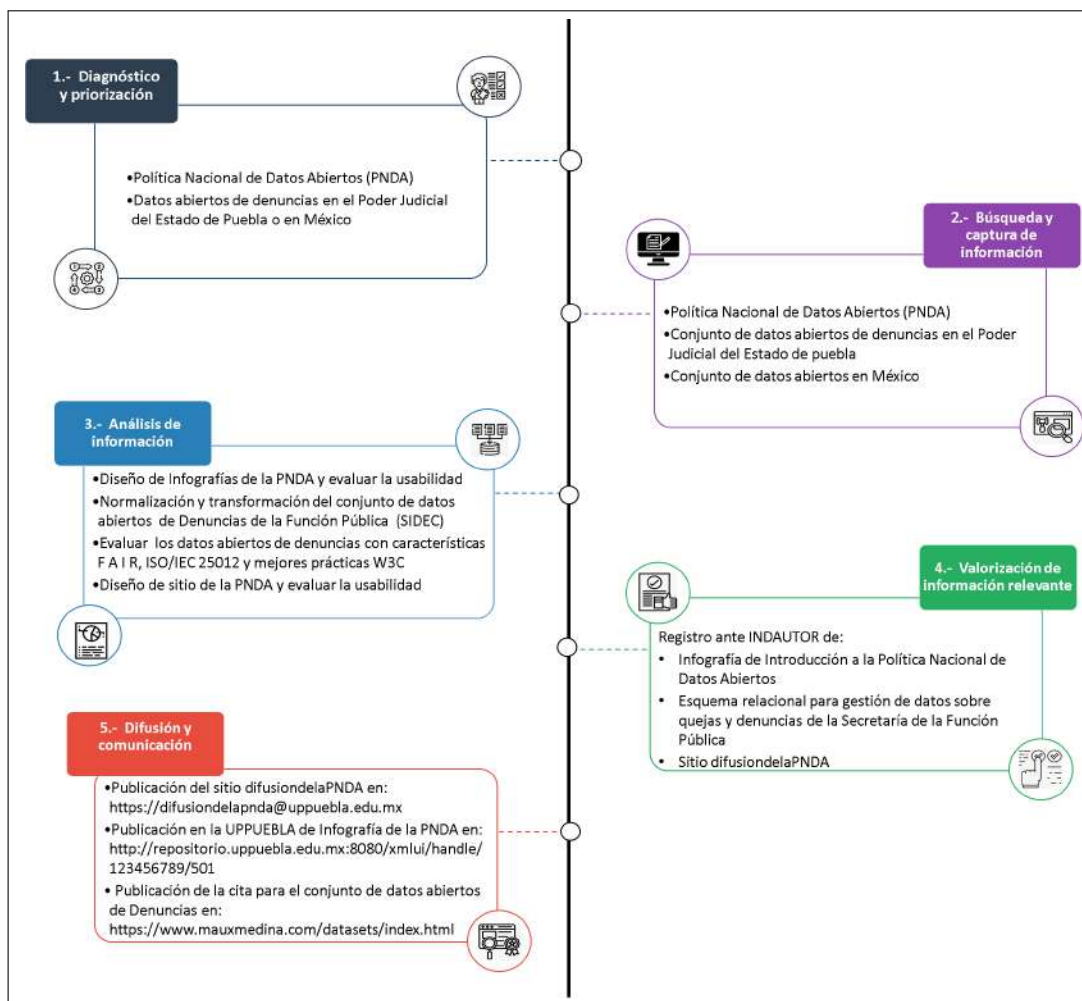


Figura 3.6: Etapas de vigilancia tecnológica.

Este capítulo 4 primero presenta la infografía para la PNDA. Después, los resultados relacionados con la evaluación de DA-DENUNCIAS. Posteriormente, se explica la transformación de DA-DENUNCIAS en ESQUEMA-RELACIONAL, se describe el sitio web denominado *difusiondelapnda* que difunde la infografía, conceptos clave y los datos transformados, se realiza la evaluación desde dos enfoques: 1) un análisis heurístico de usabilidad y 2) uso de herramientas abiertas de software. Finalmente, se enumeran los productos de propiedad intelectual generados a la par del documento de tesis.

4.1 Infografía de la PNDA

La infografía se titula “Introducción a la Política Nacional de Datos Abiertos” [45], presenta las características de la PNDA como propósito, alcance y contexto de manera concisa, contiene imágenes que permiten su reutilización, es un medio para difundir y promover la PNDA en la academia y público en general.

Las Figuras 4.1 y 4.2 muestran la parte superior e inferior, respectivamente, el archivo completo está disponible para el público en general desde el repositorio institucional de la UPPue, en la dirección electrónica siguiente:

<http://repositorio.uppuebla.edu.mx:8080/xmlui/handle/123456789/501>

4.2 Evaluación de la calidad de DA-DENUNCIAS

La evaluación de la calidad de DA-DENUNCIAS se realizó con base en el estándar ISO/IEC 25012 [14], los principios FAIR [31] y las recomendaciones para publicar datos en web [44] como se explica en las secciones 4.2.1, 4.2.2 y 4.2.3.

4.2.1 Evaluación con el estándar ISO/IEC 25012

Los porcentajes de cumplimiento para las métricas de calidad de datos inherente conforme al estándar ISO/IEC 25012 [14] de DA-DENUNCIAS son los siguientes:

Exactitud semántica. Dado que todos los datos no nulos están definidos en el diccionario de datos¹, la exactitud semántica fue del 100%.

¹datosabiertos.funcionpublica.gob.mx/datosabiertos/sidec/diccionario_sidec.xlsx



Figura 4.1: Infografía para difundir la PNDA. Parte 1.

Exactitud sintáctica. Se encontró por duplicado el nombre de la falta “Uso indebido de recursos públicos”, escrita tanto en mayúsculas y minúsculas como sólo en mayúsculas, la última aparece en 3,284 celdas (2.23%) de un total de 171,230 (ver la Figura C.3). El porcentaje de exactitud sintáctica fue 98.09%.

Completitud. El total de celdas es 1,198,610 (un millón, ciento noventa y ocho mil, seis cientos diez), producto de 171,230 filas por 7 columnas; 2,563 están vacías (0.21%) como se observa al aplicar una faceta de texto sobre el atributo TIPO DE FALTA en



Figura 4.2: Infografía para difundir la PNDA. Parte 2.

OpenRefine [50] (Figura C.5). Además, se hizo la corrección de errores ortográficos en 2,883 (ver la sección C.2). La suma es 5,446 que representan el 0.54% del total, por lo tanto, el porcentaje para la completitud fue del 99.46%.

Consistencia. Se identificó una contradicción entre la afirmación “Número de datos faltantes: ninguno” de la Figura 4.3 y las 2,563 celdas vacías para el atributo TIPO DE FALTA de la Figura C.5. Dado que el número de valores distintos para el tipo de falta (59) y la celda vacía son 60, el porcentaje para la consistencia fue: 98.33%.

Credibilidad. Los datos están disponibles en la sección *Datos Abiertos* de la plataforma Datos Abiertos de México - datos.gob.mx y se describen con los metadatos de la Figura 4.4. Debido a que la plataforma es fuente oficial de DA, la credibilidad fue del 100%.

Actualidad. Las denuncias y quejas corresponden al periodo 2 de Enero del 2012 al 2 de Diciembre del 2018. Considerando arbitrariamente como “datos actuales” si los datos corresponden a los últimos 5 años, ninguno es actual, por lo que el valor para esta característica fue del 0% (cero por ciento).

La Tabla 4.1 concentra los porcentajes para las métricas de calidad de datos inherente conforme al estándar ISO/IEC 25012 [14], el promedio para DA-DENUNCIAS fue: **82.64%**, por lo que se considera *aceptable*.

Columna	Obligatoria	Tipo	Rango	Unidad de medida
INSTITUCIÓN	SI	Alfanumérico	02-48	N/A
SIGLAS DE LA INSTITUCIÓN	SI	Alfanumérico	N/A	N/A
EXPEDIENTE	SI	Alfanumérico	2002-2018	N/A
MEDIO DE PRESENTACIÓN	SI	Catálogo	Aplicación móvil, buzón, correspondencia, medios de comunicación, medios electrónicos, personal, por procedimiento, teléfono.	N/A
TIPO DE PETICIÓN	SI	Catálogo	Denuncia / Queja	N/A
TIPO DE FALTA	SI	Catálogo	Abuso de funciones	N/A
FECHA RECEPCIÓN	SI	Fecha	02/01/2018 - 30/06/2018	N/A
5. Número de registros	171 231			
6. Número de datos faltantes	Ninguno			

Figura 4.3: Contradicción en el diccionario de datos para DA-DENUNCIAS .

Campo	Valor
Fuente	http://www.funcionpublica.gob.mx
Email del publicador	ideigado@funcionpublica.gob.mx
Frecuencia	R/P6M
GUID	informacion-de-quejas-y-denuncias-ciudadanas
Idioma	es
Inicio del periodo temporal	2012-01-02T00:00:00Z
Final del periodo temporal	2018-06-30T00:00:00Z
Modificado	2018-08-10T00:00:00Z
Nombre del publicador	SFP
Publicado	2017-10-13T16:51:02Z
Tema	Otros
Tipo del publicador	Director de Asesoría y Consulta
Diccionario de Datos	Consultar

Figura 4.4: Metadatos del conjunto DA-DENUNCIAS.

4.2.2 Cumplimiento de los principios FAIR

DA-DENUNCIAS cumple con los principios FAIR como sigue:

F: Están disponibles desde la plataforma *Datos Abiertos de México - datos.gob.mx*,

Métrica	Porcentaje
Exactitud semántica:	100.00
Exactitud sintáctica:	98.09
Completitud:	99.46
Consistencia:	98.33
Credibilidad:	100.00
Actualidad:	0.00
Promedio:	82.64

Tabla 4.1: Porcentajes para las métricas de calidad de los datos inherente [14].

- A:** Se distribuyen bajo los términos de la licencia LIBRE USO MX, están en formato csv y contienen metadatos asociados (ver la Figura 4.4).
- I:** Están publicados en formato *.csv, lo cual permite que puedan procesarse con herramientas abiertas de software como explica el Anexo D
- R:** Se usan por otras aplicaciones (ver la Figura 4.5) y para generar información, por ejemplo, sección 4.3.1 y [46]

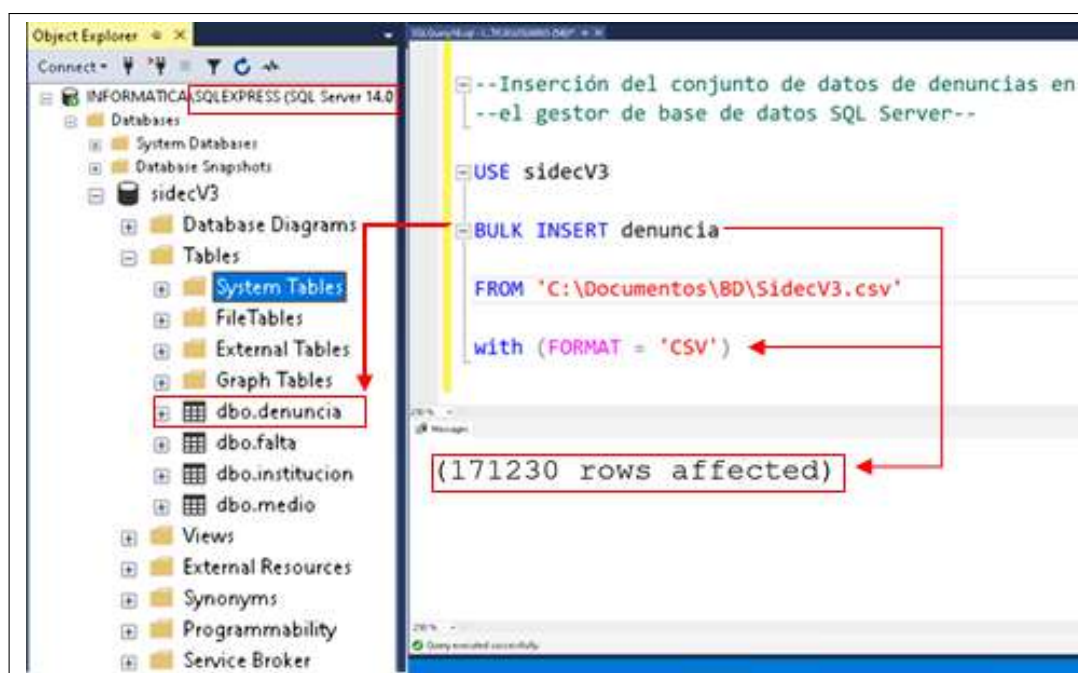


Figura 4.5: Carga de datos en SQL server.

No. de práctica	Descripción
1	Proporcionar metadatos
2	Proporcionar metadatos descriptivos
3	Proporcionar metadatos estructurales
4	Proporcionar información sobre la licencia de datos
5	Proporcionar información sobre la procedencia de los datos
9	Utilizar URIs persistentes como identificadores de conjuntos de datos
12	Utilizar formatos de datos estandarizados legibles por máquina
16	Elegir el nivel de formalización adecuado
17	Proporcionar descargas masivas
20	Proporcionar acceso en tiempo real
28	Evaluar la cobertura del conjunto de datos
29	Recopilar comentarios de los consumidores de datos
34	Seguir los términos de la licencia
35	Citar la publicación original

Tabla 4.2: *Mejores prácticas implementadas.*

4.2.3 Evaluación de mejores prácticas para datos en la web

El consorcio *World Wide Web* (W3C) recomienda un conjunto de mejores prácticas para los datos disponibles en la web [44]. Se evaluó cada una de las 35 prácticas, 14 se cumplen (40%, Tabla 4.2) y 21 no (60%, Tabla 4.3).

4.3 Transformación de DA-DENUNCIAS a DENUNCIAS-RELACIONAL

El conjunto de datos denominado “Esquema relacional para gestión de datos sobre quejas y denuncias de la Secretaría de la Función Pública”, en adelante, DENUNCIAS-RELACIONAL, es el resultado de transformar DA-DENUNCIAS en un esquema relacional que implementa el modelo entidad - relación (ER) de la Figura 4.6.

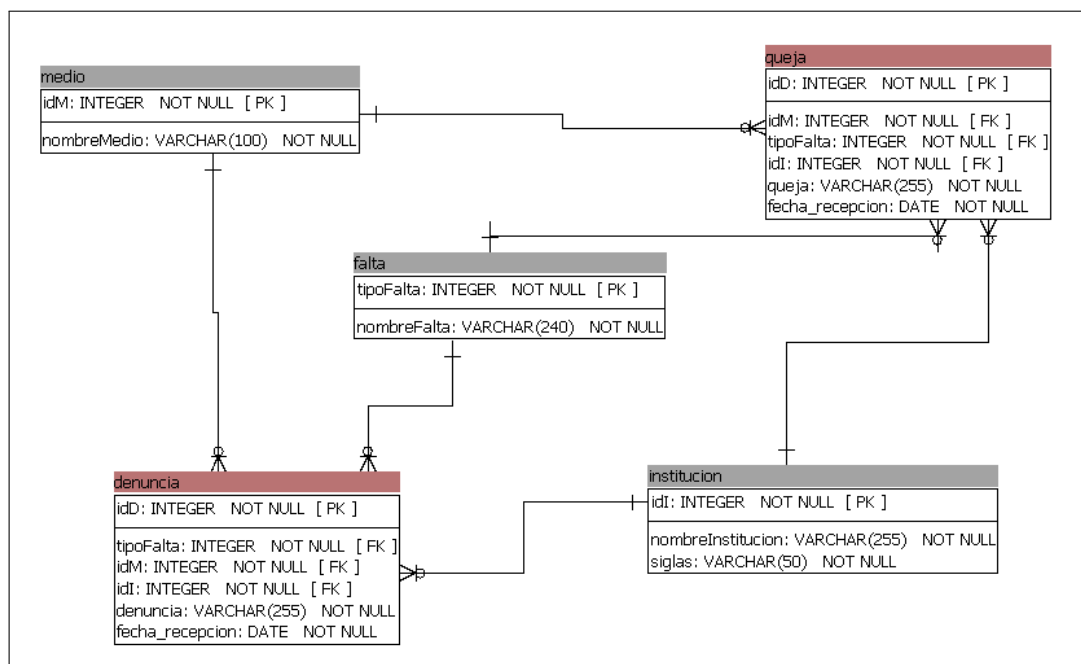


Figura 4.6: Modelo ER de DENUNCIAS-RELACIONAL.

Posterior a las transformaciones en los datos descritas en la sección C.2 del Anexo C, se importó la tabla DA-DENUNCIAS en el gestor de base de datos SQL Server versión 14.

La Figura 4.7 muestra las tablas que forman el esquema relacional. En la parte inferior de cada tabla, se incluye el número de filas, cardinalidad, abreviada como *card*. Considerar lo siguiente:

- Se realizó una fragmentación horizontal sobre la tabla denuncia, la cual originalmente contenía información tanto de quejas como de denuncias
- Los distintos valores para las entidades medio, falta e institucion corresponden a los datos del diccionario de DA-DENUNCIAS

La estructura de las tablas es como sigue:

```
CREATE TABLE medio (
    idmedio int not null primary key autoincrement,
    nombremedio varchar (100) not null
);
```

```
CREATE TABLE falta (
    idfalta int not null primary key autoincrement,
    descfalta varchar (240) not null
);
```

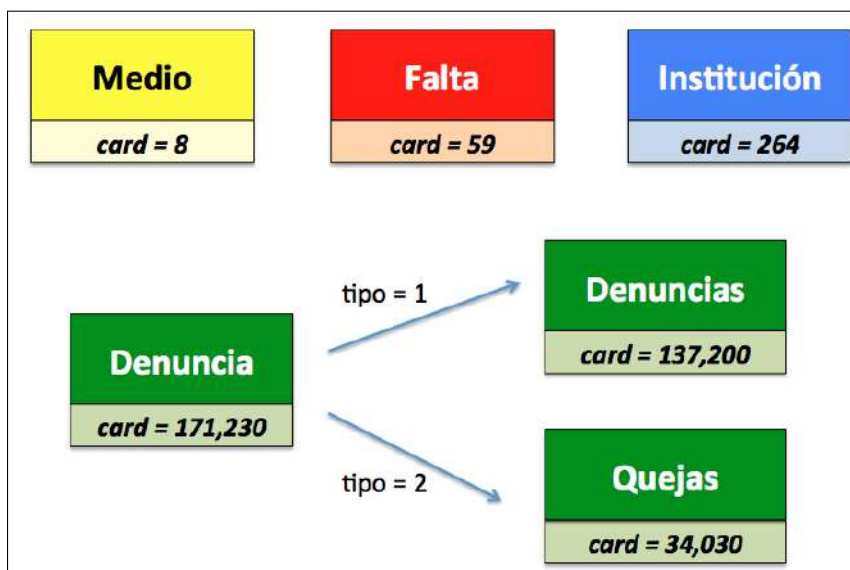


Figura 4.7: Esquema relacional.

```

CREATE TABLE institucion (
  idinstitucion int not null primary key autoincrement,
  nombreinstitucion VARCHAR(255) not null,
  siglas VARCHAR(50) not null
);

CREATE TABLE denuncia (
  idDenuncia VARCHAR(255) not null primary key,
  fechaRecepcion date not NULL,
  idmedio INT not null,
  idinstitucion INT not null,
  idfalta INT not null,
  foreign key (idmedio) references medio(idmedio),
  foreign key (idinstitucion) references institucion(idinstitucion)
  foreign key (idfalta) references falta(idfalta)
);

```

No. de práctica	Descripción
6	Proporcionar información sobre la calidad de los datos
7	Proporcionar un indicador de versión
8	Proporcionar un historial de versiones
10	Utilizar URIs persistentes como indicadores dentro del conjunto de datos
11	Asignar URIs a versiones y series de conjuntos de datos
13	Utilizar representaciones de datos neutrales a la configuración regional
14	Proporcionar datos en múltiples formatos
15	Reutilizar vocabularios, preferiblemente los estandarizados
18	Proporcionar subconjuntos para conjuntos de datos grandes
19	Utilizar la negociación de contenido para ofrecer datos disponibles en múltiples formatos
21	Proporcionar datos actualizados
22	Proporcionar una explicación para los datos que no están disponibles
23	Hacer que los datos estén disponibles a través de una API
24	Utilizar estándares web como base de las API
25	Proporcionar documentación completa para su API
26	Evitar cambios importantes en su API
27	Preservar identificadores
30	Hacer que los comentarios estén disponibles
31	Enriquecer los datos generando nuevos datos
32	Proporcionar presentaciones complementarias
33	Proporcionar comentarios al editor original

Tabla 4.3: *Mejores prácticas no implementadas.*


```

CREATE TABLE queja (
idQueja VARCHAR(255) not null primary key,
fechaRecepcion date not NULL,
idmedio INT not null,
idinstitucion INT not null,
idfalta INT not null,
foreign key (idmedio) references medio(idmedio),
foreign key (idinstitucion) references institucion(idinstitucion)
foreign key (idfalta) references falta(idfalta)
);

```

4.3.1 Implementación de consultas

Esta sección muestra la implementación de consultas en la versión 14.0 de SQL Server. En los resultados sólo se muestran los datos de las primeras 5 filas.

Consulta 1:

```

use sidecV2
SELECT top 5 count (denuncia.idfalta) as Total_de_faltas, falta.descripcionfalta
FROM denuncia,falta
where falta.idfalta=denuncia.idfalta
group by falta.descripcionfalta
order by Total_de_faltas desc

```

La Figura 4.8 muestra el resultado de la consulta 1, corresponde a los tipos de falta con mayor número de denuncias.

Consulta 2:

```

use sidecV2
SELECT top 5 count (denuncia.idmedio) as Total_de_Presentacion, medio.nombremedio as Medio_de_presen
FROM denuncia, medio
where medio.idmedio = denuncia.idmedio
group by medio.nombremedio
order by Total_de_Presentaci'on desc

```

La Figura 4.9 despliega los medios con mayor número de denuncias, resultado de la consulta 2.



Figura 4.8: Tipos de falta con mayor número de denuncias.

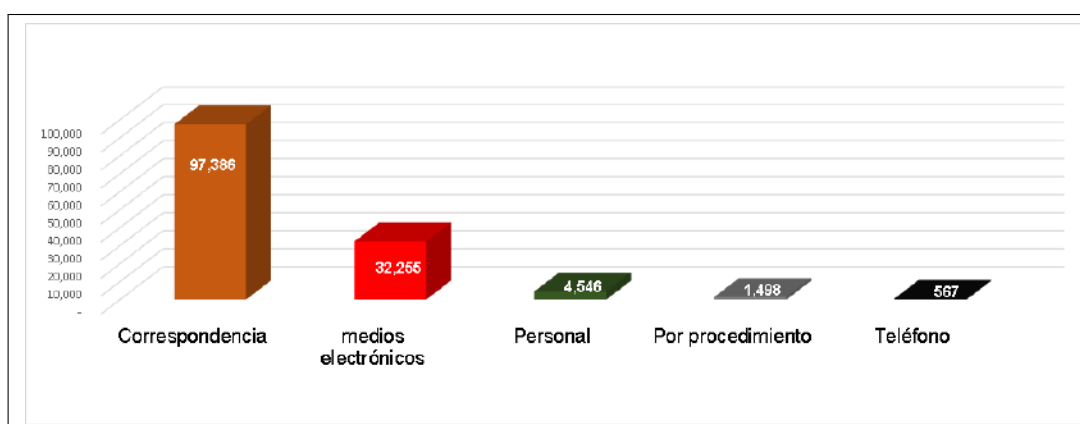


Figura 4.9: Medios con mayor número de denuncias.

Consulta 3:

```
use sidecV2
SELECT top 5 count (denuncia.idinstitucion) as Total_de_denuncias_por_institucion,
institucion.nombreinstitucion as Nombre_institucion
FROM denuncia, institucion
where institucion.idinstitucion = denuncia.idinstitucion
group by institucion.nombreinstitucion
order by Total_de_denuncias_por_institucion desc
```

La Figura 4.10 despliega los nombres de las instituciones con más denuncias, resultado de la consulta 3.

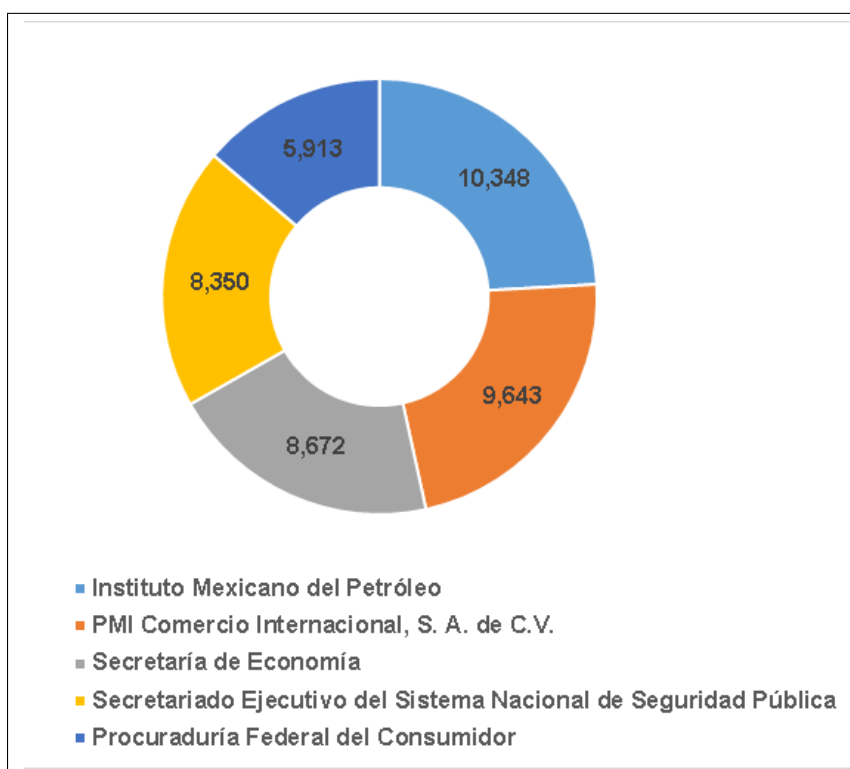


Figura 4.10: *Instituciones con mayor número de denuncias.*

Consultas 4 y 5:

```
use sidecV2
select max(fechaRecepcion) as fecha_recepcion
from denuncia
```

```
use sidecV2
select min(fechaRecepcion) as fecha_recepcion
from denuncia
```

Las Figuras 4.11 y 4.12 muestran la fecha de la denuncia más reciente y más antigua, respectivamente desde la interfaz gráfica de SQL server.

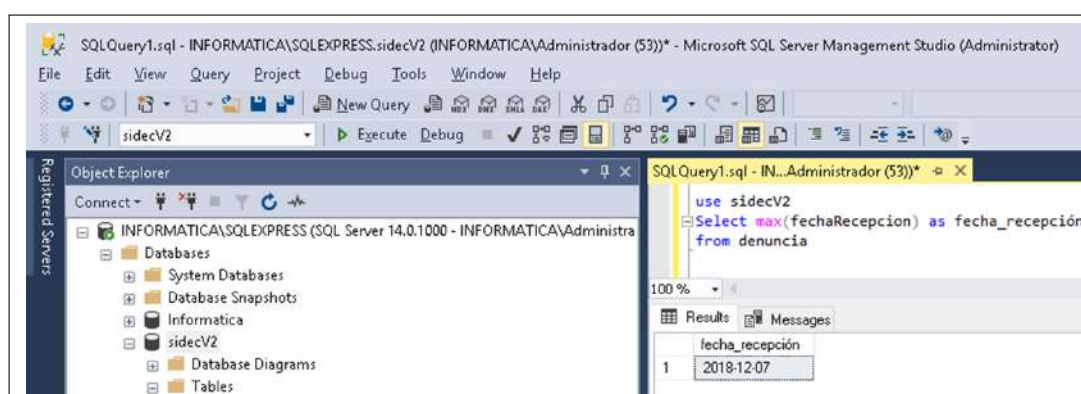


Figura 4.11: Fecha de la denuncia más reciente.

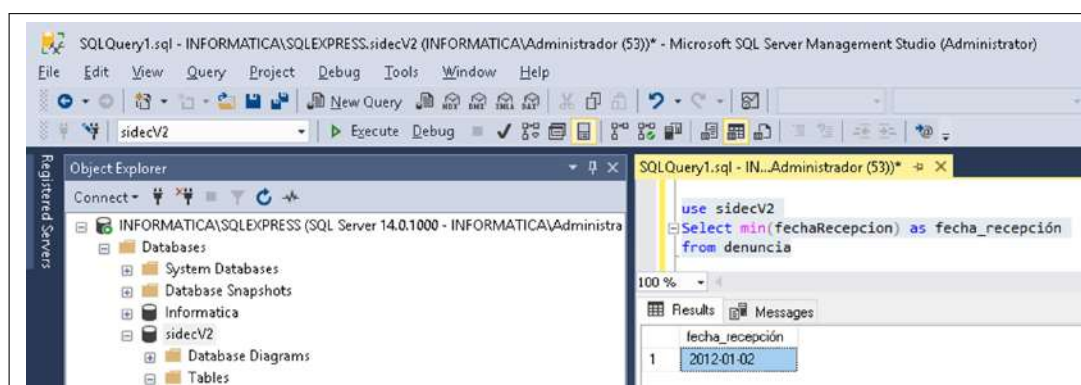


Figura 4.12: Fecha de la denuncia más antigua.

La Tabla 4.4 contiene las características del equipo de cómputo utilizado para la ejecución de las consultas, el cálculo de reducción de espacio de almacenamiento después de transformar DA-DENUNCIAS en DENUNCIAS-RELACIONAL y las pruebas con herramientas abiertas de software (sección 4.4.1.2).

No.	Componente	Descripción
1	Tipo de equipo	Computadora de escritorio
2	Sistema operativo	Microsoft Windows 10 Pro Versión 10.0.19045 Compilación 19034 Fabricante: <i>Microsoft Corporation</i> Modelo OptiPlex 7050 Tipo <i>x64-based PC</i>
3	Procesador	Intel(R) Core(TM) i7-7700 CPU @ 3.60GHz, 3600 Mhz, 4 procesadores principales, 8 procesadores lógicos
4	Versión y fecha de BIOS	Dell Inc. 1.27.0, 18/9/2023
5	Memoria física instalada (RAM)	16.0 GB
6	Disco duro de almacenamiento	(Unidades de disco estándar) Modelo ST1000DM010-2EP102 Capacidad de almacenamiento 931.51 GB

Tabla 4.4: Características del equipo de cómputo.

4.3.2 Beneficios de DENUNCIAS-RELACIONAL

Los beneficios de DENUNCIAS-RELACIONAL desde la perspectiva de las bases de datos son los siguientes:

1. Reducción en el espacio de almacenamiento considerando el formato csv: DA-DENUNCIAS ocupa 31.2 MB, en tanto, DENUNCIAS-RELACIONAL sólo 7.84 MB, es decir, el tamaño en el almacenamiento produce un ahorro del 75%
2. Modelo ER que facilita la comprensión de la información almacenada
3. Normalización de la información: se alcanza la tercera forma normal
4. Adopción de las características propias de un gestor de base de datos que apoyen

a los usuarios potenciales en el manejo de la información tales como redundancia mínima e integridad

5. Capacidad tecnológica para automatizar tareas de mantenimiento de los datos que van desde su limpieza o procesamiento, hasta la generación de reportes y estadísticas

DENUNCIAS-RELACIONAL incrementa el porcentaje de prácticas implementadas en DA-DENUNCIAS, del 40% al 48.5% al satisfacer las siguientes [44]:

Práctica 6: “Proporcionar información sobre la calidad de los datos”, los cuales cumplen la tercera forma normal y no son redundantes

Práctica 21: “Proporcionar datos en múltiples formatos” tales como csv o sql

Práctica 32: “Proporcionar presentaciones complementarias” como el modelo ER de la Figura4.6

Una desventaja es que la representación de los datos en SQL Server considerando llaves foráneas y la separación de quejas y denuncias, resultó en un incremento del 75% en el espacio de almacenamiento, de 31.2 MB a 54.7 MB.

4.3.3 Automatización de tareas

El proceso de automatización para el mantenimiento del conjunto DA-DENUNCIAS se llevó a cabo mediante los siguientes pasos:

1. Diseño de la estructura de la base de datos
2. Importación de datos en el gestor
3. Ejecución del script correspondiente
4. Limpieza de datos con OpenRefine [50] conforme a los pasos del Anexo C
5. Inserción de datos en las cinco tablas resultantes

Adicionalmente, se elaboraron *scripts con cursores* con dos propósitos principales:

Automatizar la limpieza de datos. Los scripts reducen los errores de transformación de datos (anexo C) que pudieran producirse por las personas. Notar que la cadena “se desconoce” se asigna al atributo tipo de falta en las 2,563 celdas que originalmente estaban vacías

Garantizar la integridad referencial. Se crean tablas padres e hijas con sus respectivas llaves para garantizar que exista integridad referencial en la base de datos

Manejo de errores. Los scripts contienen código para manejar errores y excepciones, por ejemplo, el error número 1048 para valores no nulos o el 1452 para cuando se intente agregar una llave foránea en la misma tabla

Entre las tareas que realizan los scripts con cursores, se encuentran las siguientes:

1. Importación de 171,230 filas que componen el conjunto DA-DENUNCIAS en formato csv
2. Procesamiento de 8,009 filas para corregir errores ortográficos y celdas en blanco en el conjunto de datos
3. Creación de las tablas padre siguientes:
 - medio: inserción de 8 filas
 - faltas: inserción de 59 filas
 - institución: inserción de 264 filas
4. Creación de las tablas hijas siguientes:
 - denuncias: con llaves foráneas e inserción de 137,200 filas
 - quejas: con llaves foráneas e inserción de 34,030 filas

Después de ejecutar los scripts en un equipo con sistema operativo Debian 11 y utilizando la versión 10.5.23 del gestor de base de datos MariaDB, la generación de la base de datos tardó aproximadamente 18 minutos, siendo la importación de datos y la creación de las tablas de denuncias y quejas las tareas que demandaron más tiempo.

4.3.4 Dificultades técnicas

Entre las dificultades técnicas se observan las siguientes:

1. El script que genera OpenRefine al exportar datos en SQL agrega comillas al nombre de los atributos, se requiere eliminarlas antes de importar los datos en algún gestor
2. Por omisión, el máximo de inserciones (instrucciones de tipo insert into) que soporta SQL Server es de 1000, por lo que los datos se cargaron utilizando el formato csv

4.3.5 Uso de DENUNCIAS-RELACIONAL como material didáctico

DENUNCIAS-RELACIONAL se usó como material didáctico en la asignatura de Bases de Datos Avanzadas (BDA) que se imparte en la Maestría en Ingeniería en Sistemas y Cómputo Inteligente (MISCI) de la Universidad Politécnica de Puebla (UPPue), periodo Mayo - Agosto del 2024. Esta asignatura fue impartida por la asesora de esta tesis.

El grupo estuvo integrado por 7 estudiantes (6 hombres, 1 mujer), todos en el rango de 25 a 35 años de edad. Se consideró un entorno virtual, esto es, se realizó una clase virtual asistida por la tecnología de Google Meets en la cual se expuso el modelo ER de DA-Denuncias y se compartió el script de DENUNCIAS-RELACIONAL. La actividad didáctica consistió en importar el script en el gestor de base de datos que cada estudiante deseara, posterior a la importación, se requirió construir fragmentos de la tabla denuncia de acuerdo al esquema de fragmentación que se haya asignado a cada estudiante. Finalmente, se solicitó construir un reporte (documento en formato *.PDF) el cual mostrara la ejecución de 5 consultas sobre el esquema fragmentado. La Figura 4.13 muestra las imágenes de los 7 esquemas de fragmentación, notar que se ocultaron los nombres de dos estudiantes. Los resultados fueron los siguientes:

- Los gestores de bases de datos utilizados fueron: MySql server (por 4 estudiantes) y MariaDB (por 3 estudiantes), los 7 estudiantes utilizan el sistema operativo Windows
- Los siete estudiantes lograron importar los datos (100%)
- Seis de los siete estudiantes (85.7%) construyeron los fragmentos y entregaron el reporte solicitado en tiempo y forma, un estudiante no entregó el reporte

4.4 Difusión a través del sitio difusiondelapnda

El sitio web denominado difusiondelapnda se implementó para difundir la infografía de la PNDA y los datos transformados en un esquema relacional, está compuesto de las páginas siguientes:

Conceptos. Contiene una definición para los conceptos de *gobernanza e innovación* [33]; incluye la cita al conjunto de datos descrito en [46]

Contacto. Despliega los correos electrónicos del autor y coautoras del sitio

Index. Es la página de inicio, muestra tres imágenes en un carrusel sobre la estrategia *Abramos México* [9], la definición de DA de la LGTAIP [8] y los principios FAIR

The screenshot shows a Classroom assignment interface. At the top, it says 'Classroom > Bases de datos avanzadas. Mayo - Agosto 2024.' Below that, there are tabs for 'Instrucciones' and 'Trabajo de los alumnos'. The main title is 'EVIDENCIA 6. Reporte de consultas' by 'MARIA AUXILIO MEDINA NIETO' (8 Jun, last modified 6 Jul), worth '100 puntos' with a due date of '14 Jul'. The instructions ask students to create a report based on a case and data assignment, describing the steps for implementing a distributed query. A note specifies that 'en el sitio 2 está el fragmento denuncias1' and that 'las faltas tienen una prioridad representada por los valores ALTA, MEDIA y BAJA'. Below the text, there are several PDF files for assignment, each with a small thumbnail showing a diagram or table. The files are: 'asignacionDeFragmentosAn...', 'asignacionDeFragmentosCh...', 'asignacionDeFragmentosGo...', 'asignacionDeFragmentosJa...', 'asignacionDeFragmentosLui...', and 'asignacionFragMiguel.pdf'.

Figura 4.13: Evidencia del uso de DENUNCIAS-RELACIONAL como material didáctico.

[31], da respuesta a las preguntas siguientes que tienen el propósito de introducir conceptos clave a los usuarios potenciales:

1. ¿Qué es *Abramos México*?
2. ¿Cuál es el propósito de la PNDA?
3. ¿A qué se le denomina DA?
4. ¿Cómo se evalúa la calidad de los DA?

Instituciones Judiciales. Permite a los usuarios acceder a los sitios de instituciones judiciales del estado de Puebla. En adelante, se denotará como InstitucionesJ

La Figura 4.14 muestra la sección superior de la página `Index.html`, la sección media despliega las preguntas y respuestas, y la inferior o pie de página, contiene enlaces a las páginas de la estrategia *Abramos México* y al sitio web del INAI.

4.4.1 Evaluación

La evaluación del sitio `difusiondelapnda` se realizó considerando dos enfoques: 1) la auto-evaluación por parte del equipo de desarrollo y 2) mediante el uso de herramientas abiertas de software. Los resultados se presentan en las secciones 4.4.1.1 y 4.4.1.2.



Figura 4.14: *Página Index.html, inicio del sitio difusiondelapnda.*

Elemento	Valor
Fecha:	14 de Agosto del 2024
Tipo de conexión:	Alámbrica
Plataforma:	Computadora personal (equipo de escritorio)
Agente de usuario:	José Julián Barrera Rojas
Tarea:	Encontrar el objetivo del sitio
Pantalla:	Página Index.html

Tabla 4.5: *Datos del análisis heurístico.*

4.4.1.1 Auto-evaluación por equipo de desarrollo

Se utilizó el instrumento *análisis heurístico para UX* del estudio Torres Burriel, éste se obtuvo al llenar un formulario de solicitud disponible en:

<https://torresburriel.com/weblog/analisis-heuristico-ux/>

La Tabla 4.5 muestra los datos del análisis heurístico. Cada heurística agrupa a varias preguntas que se contestan utilizando una escala de Likert (1 valor mínimo, 5 valor máximo). La Tabla 4.6 muestra los resultados de 9 heurísticas, no se consideraron las relacionadas con la *búsqueda* ni la *ayuda* dado que no están implementadas.

Heurística	Resultado
Generales:	5
Identidad e información	5
Lenguaje y redacción	5
Rotulado	5
Estructura y navegación	5
Organización (<i>layout</i>) de la página	5
Elementos multimedia	4
Accesibilidad	4
Control y retroalimentación	4
Promedio:	4.6

Tabla 4.6: Resultados del análisis heurístico.

Página	Rendimiento	Accesibilidad	Recomendaciones
Conceptos	90	79	82
Contacto	89	93	96
Index	95	92	96
InstitucionesJ	92	79	82

Tabla 4.7: Resultados de PageSpeed Insights en celular.

4.4.1.2 Resultados con herramientas abiertas de software

Se usaron cinco herramientas abiertas de software seleccionadas a conveniencia para producir dos versiones de *difusiondelapnda*. Los resultados expuestos en esta sección corresponden a la segunda, aunque están sujetos a interpretaciones, permiten validar de manera automática el contenido y código.

Los porcentajes obtenidos para tres métricas por la API *PageSpeed Insights*² (PSI) para dispositivos móviles y en computadoras se muestran en las Tablas 4.7 y 4.8, respectivamente.

La herramienta *Screenfly*³ se utilizó para probar la visualización de las páginas en resoluciones para dispositivo móvil, computadora personal, tableta y televisión. La Figura 4.15 muestra la página *Index.html* con la resolución de un dispositivo móvil, no hubo pérdidas de información en ninguna de las páginas en las cuatro resoluciones distintas.

Se usó la herramienta *Stilus*⁴ para detectar errores en la ortografía, gramática, estilo y tipografía de las cuatro páginas. Como ejemplo, la Figura 4.16 muestra 12 errores ortográficos para la página de *Conceptos.html*, los cuales en realidad corresponden al uso de siglas. Para esta página, no se detectó ningún error de gramática, se reporta un error

²<https://pagespeed.web.dev/>

³<https://screenfly.org>

⁴<https://www.mystilus.com/>

Página	Rendimiento	Accesibilidad	Recomendaciones
Conceptos	100	92	96
Contacto	92	93	96
Index	99	92	96
InstitucionesJ	100	85	96

Tabla 4.8: Resultados de PageSpeed Insights en computadora personal.



Figura 4.15: Página *Index.html* en resolución para dispositivo móvil.


de estilo y el uso de 2 tipos de letra distintos.

La validación del lenguaje de marcado se realizó con la herramienta Nu HTML Checker, servicio Markup validation service⁵. Como evidencia, la Figura 4.17 muestra los resultados para la página de *Contacto.html*.

La revisión de enlaces rotos se llevó a cabo con W3C Link Checker⁶. Los resultados para la página *InstitucionesJ* se observan en la Figura 4.18.

⁵<https://validator.w3.org/>

⁶<https://validator.w3.org/checklink>

 Imprimir

Errores:
ORTOGRAFÍA: 12
GRAMÁTICA: 0
ESTILO: 1
TIPOGRAFÍA: 2
 Palabras: 461
 Palabras diferentes: 274
 Caracteres con espacio: 3 288
 Caracteres sin espacio: 2 761
 Oraciones: 47
 Párrafos: 33
 Tiempo: 2.3s

© 1998-2024 MeaningCloud.

Difusión de la Política Nacional de Datos Abiertos

[🏠 Inicio](#)
[📖 Conceptos](#)

[🔍 Ejemplos de instituciones judiciales ▾](#)

[📊 Infografía de la **PNDA**](#)

ORTOGRAFÍA (A1)
 Posible sigla errónea o desconocida.

[✉ Contacto](#)

CONCEPTOS DE GOBERNANZA E INNOVACIÓN

¿Qué es gobernanza?

De acuerdo con el Programa de las Naciones Unidas para el Desarrollo, en el año 2000, se definió a la gobernanza como:

El sistema de valores, políticas e instituciones por el que una sociedad administra sus asuntos económicos, sociales y políticos

Figura 4.16: Errores de la página *Conceptos.html* detectados por *Stilus*.

Nu Html Checker

This tool is an ongoing experiment in better HTML checking, and its behavior remains subject to change

Showing results for <https://www.mauxmedina.com/difusiondelapnda/Contacto.html>

Checker Input

Show source outline image report

Check by address ▼

<https://www.mauxmedina.com/difusiondelapnda/Contacto.html>

Use the Message Filtering button below to hide/show particular messages, and to see total counts of errors and warnings.


1 message hidden by filtering

Document checking completed. No errors or warnings to show.

Used the HTML parser
Total execution time 924 milliseconds.

[About this checker](#) • [Report an issue](#) • Version: 24.8.2

Figura 4.17: Resultados de la página *Contacto.html* por Nu HTML Checker.



Interested in "developing" your developer skills? In W3C's hands-on Professional Development Program

[Dc](#)

Processing <https://www.mauxmedina.com/difusiondelapnda/Instituciones.html>

Go to [the results](#).

For reliable link checking results, check [HTML validity](#) and [CSS validity](#) first.

Back to the [link checker](#).

Status: Done. Document processed in 39.52 seconds.

```

Checking link https://es-la.facebook.com/abramos_mx
HEAD https://es-la.facebook.com/abramos_mx fetched in 0.09 seconds

Checking link https://www.mauxmedina.com/difusiondelapnda/Imagenes/PiePag/inai_footer.avif
HEAD https://www.mauxmedina.com/difusiondelapnda/Imagenes/PiePag/inai_footer.avif fetched in 1.11 seconds

Checking link https://twitter.com/UPPueblaOficial
HEAD https://twitter.com/UPPueblaOficial fetched in 0.18 seconds

Checking link https://abramosmexico.org.mx/
HEAD https://abramosmexico.org.mx/ fetched in 1.88 seconds

Processed in 39.52 seconds.
```

Results

List of broken links and other issues

There are issues with the URLs listed below. The table summarizes the issues and suggested actions by HTTP response status code

Code	Occurrences	What to do
(N/A)	6	The link was not checked due to robots exclusion rules . Check the link manually, and see also
(N/A)	1	Accessing links with this URI scheme has been disabled in link checker.

Figura 4.18: Resultados de la página *InstitucionesJ* por W3C Link Checker.

4.5 Productos de propiedad intelectual

A la par del desarrollo de la tesis, se obtuvo el registro ante el Instituto Nacional del Derecho de Autor (INDAUTOR) de los productos siguientes:

1. Infografía de la PNDA [45], contiene una descripción breve de la Política Nacional de Datos Abiertos, su objetivo, participantes y beneficios potenciales
2. El conjunto de datos denominado “Esquema relacional para gestión de datos sobre quejas y denuncias de la Secretaría de la Función Pública” [46], transforma la información de quejas y denuncias ciudadanas distribuida en una tabla en formato csv, en un esquema relacional compuesto por entidades y relaciones, el propósito es reutilizar la información pública como insumo para fortalecer el aprendizaje y la práctica de temas afines a las bases de datos impartidos en Instituciones de Educación Superior (IES)
3. El sitio web *difusiondelapnda* tiene como objetivo difundir la Política Nacional de Datos Abiertos en la comunidad académica, incluye la definición de los conceptos de *gobernanza* e *innovación* y representa un medio para difundir los productos 1 y 2

En la tesis se analizó la Política Nacional de Datos Abiertos de México aprobada en 2023, se difundió mediante la elaboración de una infografía¹ y la implementación del sitio web difusiondelapnda².

Se realizaron actividades de vigilancia tecnológica en organizaciones del Gobierno Federal y de instituciones judiciales del Estado de Puebla. Los hallazgos son los siguientes:

- La SFP distribuye DA-DENUNCIAS en la plataforma <https://datos.gob.mx>, este conjunto contiene información sobre la naturaleza de las denuncias, las entidades involucradas y las instituciones del 2 de enero del 2012 al 7 de diciembre del 2018
- En el estado de Puebla, los DA de las instituciones judiciales descritas en la sección 2.4, se acceden desde la PNT y corresponden a las temáticas de la Tabla 2.1

En la tesis se realizó la evaluación de la calidad de datos inherente de DA-DENUNCIAS con base en el estándar ISO/IEC 25012. El promedio de seis métricas resultó ser 82.64%, lo cual se consideró aceptable y por tanto, se valoró como adecuada su reutilización. Cabe señalar que la única métrica que no satisface DA-DENUNCIAS es la *actualidad*, puesto que la fecha de las denuncias tienen más de seis años. Se espera que la publicación de datos de igual o mayor calidad continúe a corto y mediano plazo.

Se mostró que los datos y metadatos de DA-DENUNCIAS cumplen con los principios FAIR y después de transformar DA-DENUNCIAS en DENUNCIAS-RELACIONAL, se obtuvo un incremento del 8.5% en la implementación de prácticas para datos en la web [44].

La tesis presentó una innovación tecnológica para la gestión de datos de DA-DENUNCIAS puesto que en DENUNCIAS-RELACIONAL se eliminó redundancia, se garantizó la integridad de los datos y se mostró que la ejecución de consultas es eficiente, además, cualquier ajuste futuro en los datos sólo requerirá modificaciones menores al modelo ER o scripts.

Se reportó el uso de DENUNCIAS-RELACIONAL como material didáctico de un curso de maestría, se espera sea reutilizado para promover el desarrollo de capacidades y habilidades entre estudiantes de licenciatura o maestría que cursan asignaturas afines a las bases de datos relacionales, por otro lado, es un ejemplo del uso innovador de los DA, representa una alternativa de gestión avanzada de la información sobre denuncias que está al alcance de los *stakeholders*

¹<http://repositorio.uppuebla.edu.mx:8080/xmlui/handle/123456789/501>

²<http://difusiondelapnda@uppuebla.edu.mx/index.html>

A la par del desarrollo de la tesis, se obtuvo el registro ante INDAUTOR de los productos titulados “Infografía de la Política Nacional de Datos Abiertos” [45] y “Esquema relacional para gestión de datos sobre quejas y denuncias de la Secretaría de la Función Pública” [46].

Como trabajo a futuro, se propone la evaluación del aprendizaje de estudiantes de ingeniería o posgrado que hagan uso de los productos desarrollados.

Referencias

- [1] Wikipedia (2024). Datos abiertos - Wikipedia, la enciclopedia libre. https://es.wikipedia.org/wiki/Datos_abiertos#Definici3n
- [2] Open Data Charter (2015). International Open Data Charter. https://opendatacharter.org/wp-content/uploads/2023/12/opendatacharter-charter_F.pdf
- [3] Babini D., Rovelli L. Tendencias recientes en las pol3ticas cient3ficas de ciencia abierta y acceso abierto en Iberoam3rica. Consejo Latinoamericano de Ciencias Sociales (CLACSO). Primera edici3n. ISBN 978-987-722-637-9.
- [4] INEGI (2024). Qui3nes somos. Instituto Nacional de Estadística y Geografía (INEGI). https://www.inegi.org.mx/inegi/quienes_somos.htm
- [5] Diario Oficial de la Federaci3n (2014). Norma T3cnica para el acceso y publicaci3n de Datos Abiertos de la Informaci3n Estadística y Geogr3fica de Inter3s Nacional. https://www.dof.gob.mx/nota_detalle.php?codigo=5374183&fecha=04/12/2014#gsc.tab=0
- [6] INEGI (2024). Programas de informaci3n. Instituto Nacional de Estadística y Geografía (INEGI). <https://www.inegi.org.mx/programas/>
- [7] Plan Nacional de Socializaci3n del Derecho de Acceso a la Informaci3n. <https://www.dof.gob.mx/2018/INAI/ACT-PUB-28-11-2018.07.pdf>
- [8] Ley General de Transparencia y Acceso a la Informaci3n P3blica. (2024). Leyes Federales vigentes. <https://www.diputados.gob.mx/LeyesBiblio/pdf/LGTAIP.pdf>
- [9] Abramos M3xico (2023). Pol3tica Nacional de Datos Abiertos. <https://abramosmexico.org.mx/comunicacion/>
- [10] Observatorio Tecnol3gico de la Universidad de Alicante (OVTT) (2020). Gu3a de vigilancia de inteligencia tecnol3gica. <https://www.ovtt.org/guias/guia-de-inteligencia-tecnologica/>.
- [11] Comisi3n Nacional de Derechos Humanos (CNDH). (2024). Declaraci3n Universal de Derechos Humanos. https://www.cndh.org.mx/sites/all/doc/Programas/Discapacidad/Declaracion_U_DH.pdf

- [12] Constitución Política de los Estados Unidos Mexicanos. (2024). Leyes Federales vigentes. <https://www.diputados.gob.mx/LeyesBiblio/pdf/CPEUM.pdf>
- [13] Condatos (2023). Política Nacinal de Datos Abiertos. Abramos México: Condatos. <https://drive.google.com/file/d/1wS0kWmxx1I68PYp81V02k4X0c22oQBXA/view>
- [14] International Organization for Standardization (ISO). (2008). Software engineering - Software product Quality Requirements and Evaluation (SQuaRE) - Data quality model. (ISO/IEC 25012:2008). <https://www.iso.org/standard/35736.html>
- [15] INAI (2021). Métrica de gobierno abierto 2021: resumen ejecutivo. Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales. Centro de Investigación y Docencia Económicas (CIDE), Ciudad de México. https://micrositios.inai.org.mx/gobiernoabierto/wp-content/uploads/2022/03/Resumen-ejecutivo_-MGA-2021.pdf
- [16] Cejudo, G., de los Cobos, P. ., Ramírez, D. L. . (2023). Midiendo la apertura gubernamental: diseño, resultados y aportaciones de la métrica de Gobierno Abierto en México. *Estado, Gobierno Y Gestión Pública*, 20 (38). <https://doi.org/10.5354/0717-8980.2022.67398>
- [17] Plataforma Nacional de Transparencia (2023). Consulta de datos abiertos. https://www.plataformadetransparencia.org.mx/web/guest/datos_abiertos
- [18] Organización de las Naciones Unidas. (2015). Objetivo 4 del Desarrollo Sostenible: garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje durante toda la vida para todos. Organización de las Naciones Unidas. <https://www.un.org/sustainabledevelopment/education/>
- [19] Organización de las Naciones Unidas. (2015). Objetivo 16 del Desarrollo Sostenible: promover sociedades justas, pacíficas e inclusivas. Organización de las Naciones Unidas. <https://www.un.org/sustainabledevelopment/peace-justice/>
- [20] Abramos México (2022). Conferencia Nacional de Datos Abiertos (DATACON). <https://abramosmexico.org.mx/wp-content/uploads/2023/03/Conferencia-Nacional-de-Datos-Abiertos-0nePager.pdf>
- [21] Vera M. M., Rocha R. D., Navarro C. A. B. (2018). Políticas públicas de datos abiertos en México. *Buen Gobierno* (25), pp. 1-26. Fundación Mexicana de Estudios Políticos y Administrativos, A.C. https://doi.org/10.35247/buengob_25_06

- [22] Barómetro Global de Datos (2022). Primera Edición Informe - Barómetro Global de Datos. ILDA. DOI: <https://doi.org/10.5281/zenodo.6488349>. <https://globaldatabarometer.org/wp-content/uploads/2022/05/GDB-Report-Spanish.pdf>
- [23] Codeando México (2021). Hacia una política nacional de datos abiertos en México | by Codeando México | Codeando México | Medium. <https://medium.com/codeandomexico/hacia-una-pol%C3%ADtica-nacional-de-datos-abiertos-en-m%C3%A9xico-520a40a9b0fc>
- [24] Infraestructura abierta (2024). Infraestructura abierta. Consejo impulsor: Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales, México evalúa: centro de análisis de políticas públicas, *Infrastructure Transparency Initiative, Open Contracting Partnership*, . <http://infraestructuraabierta.mx/>
- [25] Suprema Corte de Justicia de la Nación (SCJN) (2023). Ley Orgánica del Poder Judicial del Estado. (Ley publicada en el Número 5 del Periódico Oficial del Estado de Puebla, el viernes 6 de enero de 2023.). https://www.scjn.gob.mx/sites/default/files/justicia_constitucional_local/documento/2023-04/162309.pdf
- [26] Ley de Transparencia y Acceso a la Información Pública del Estado de Puebla. (2024). Gobierno del Estado de Puebla. <https://ojp.puebla.gob.mx/legislacion-del-estado/item/177-ley-de-transparencia-y-acceso-a-la-informacion-publica-del-estado-de-puebla>
- [27] Ley del Sistema Anticorrupción del Estado de Puebla. (2024). Gobierno del Estado de Puebla. <https://ojp.puebla.gob.mx/legislacion-del-estado/item/198-ley-del-sistema-anticorrupcion-del-estado-de-puebla>
- [28] Acuerdo General del Pleno del Consejo de la Judicatura del Poder Judicial del Estado, por el que establece las disposiciones en Materia de Responsabilidades Administrativas. (2023). Gobierno del Estado de Puebla. <https://ojp.puebla.gob.mx/legislacion-del-estado/item//6173-acuerdo-general-que-establece-las-disposiciones-en-materia-de-responsabilidades-administrativas>
- [29] Secretaría de la Función Pública. (2024). Política de Transparencia, Gobierno Abierto y Datos Abiertos de la Administración Pública Federal. https://funcionpublica.gob.mx/web/transparencia/Politica_de_Transparencia_Gobierno_Abierto_y_Datos_Abiertos_de_la_APF_2021-2024.pdf

- [30] FORCE11: the Future of Research Communications and e-Scholarship. (s.f.). Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing versión b1.0. <https://force11.org/info/guiding-principles-for-findable-accessible-interoperable-and-re-usable-dat>
- [31] Scientific Electronic Library Online (2016). Principios rectores FAIR publicados en una revista del Nature Publishing Group [en línea]. SciELO en Perspectiva. <https://blog.scielo.org/es/2016/03/16/principios-rectores-fair-publicados-en-una-revista-del-nature-publishing-group/>
- [32] Infraestructura de Datos Espaciales de la Provincia de Córdoba (IDECOR). (2023). Principios FAIR: ¿Cómo hacer tus datos accesibles e interoperables?. <https://www.idecor.gob.ar/principios-fair-como-hacer-tus-datos-accesibles-e-interoperables/>
- [33] OECD/Eurostat (2018), Oslo manual 2018: guidelines for collecting, reporting and using data on innovation, 4th edition, the measurement of scientific, technological and innovation activities, OECD Publishing, Paris/Eurostat, Luxembourg. <https://doi.org/10.1787/9789264304604-en>
- [34] Ministerio de Ciencia Tecnología e Innovación (Minciencias). (2024). Vigilancia tecnológica y/o inteligencia competitiva. Ministerio de Ciencia Tecnología e Innovación de Colombia. Glosario. <https://minciencias.gov.co/glosario/vigilancia-tecnologica-yo-inteligencia-competitiva>
- [35] Benavides V. C. A., Quintana G. C. (2006). Inteligencia competitiva, prospectiva e innovación: la norma UNE-166006 EX sobre el sistema de vigilancia tecnológica. Boletín Económico de Información Comercial Española (BICE) número 2896. <https://revistasice.com/index.php/BICE/article/view/4195/4195>
- [36] Pandolfo M. (2023). Análisis de los procesos de vigilancia tecnológica e inteligencia estratégica en la gestión de I+D+i en los organismos del Sistema Nacional de Ciencia, Tecnología e Innovación de Argentina. Universidad Nacional de Mar del Plata. Facultad de Humanidades. Departamento de Ciencia de la Información. Tesis de licenciatura en bibliotecología y documentación. <http://humadoc.mdp.edu.ar:8080/xmlui/handle/123456789/974>
- [37] Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI) (2023). Contrataciones abiertas. <https://micrositios.inai.org.mx/contratacionesabiertas/>

- [38] Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI) (2023). Política Nacional de Datos Abiertos: una ruta en colectivo. <https://datacon.mx/>
- [39] Vera A. A., Medina-Nieto M. A. (2024). Acceso a datos abiertos sobre ciencia y tecnología en Instituciones de Educación Superior. En “Objetos de aprendizaje para apoyar la docencia y el aprendizaje”. Becerra F. G., Zepeda C. C., Castillo Z. H., Carballido C. J. L., editores. Benemérita Universidad Autónoma de Puebla. ISBN: 978-607-5914-15-2
- [40] Calabrese J., Esponda S., Pasini A., Boracchia M., Pesado O. (2019). Guía para evaluar calidad de datos basada en ISO/IEC 25012. XXV Congreso Argentino de Ciencias de la Computación. <https://core.ac.uk/download/pdf/301104068.pdf>
- [41] Medina Nieto M. A., López D. E., De la Calleja M. J., Hernández V. Y. Arrieta D. D. (2024). Quality Assessment and Semantic Representation of the Module on Cyberbullying, en *IEEE Latin America Transactions*, 22 (1), pp. 6-14. DOI: 10.1109/TLA.2024.10375733
- [42] LISA News. (2022). La vigilancia tecnológica en la empresa. <https://www.lisanews.org/seguridad/que-es-la-vigilancia-tecnologica-tipos-y-ejemplos/>
- [43] . Fonseca C. J. M., Juárez Z. B. (2019). Ciberseguridad y vigilancia tecnológica: un reto para la protección de datos personales en los archivos. *Revista Académica de Investigación*, 31. <https://dialnet.unirioja.es/servlet/articulo?codigo=729555>
- [44] World Wide Web Consortium (2017). Data on the web best practices: W3C recommendation 31 January 2017. <https://www.w3.org/TR/dwbp/#intro>
- [45] Barrera R. J. J., Medina N. M. A., Rodríguez H. R., Pérez C. G. P., Sánchez E. G. (2024). Introducción a la Política Nacional de Datos Abiertos. [Dibujo (arte digital)]. Certificado del Registro Público del Derecho de Autor No. 03-2024-071712175600-01
- [46] Barrera R. J. J., Medina N. M. A., Rodríguez H. R., Pérez C. G. P., Sánchez E. G. (2024). Esquema relacional para gestión de datos sobre quejas y denuncias de la Secretaría de la Función Pública. [Compilación de datos (base de datos)]. Certificado del Registro Público del Derecho de Autor No. 03-2024-071712194700-01.
- [47] Barrera R. J. J., Medina N. M. A., Rodríguez H. R., Pérez C. G. P., Sánchez E. G. (2024). Gestión de datos sobre denuncias con herramientas abiertas de soft-

-
- ware. Certamen de Innovación en Transparencia (CIT 2024). Etapa de presentación.
https://www.youtube.com/watch?v=RxL0uNbc8_k
- [48] Barrera R. J. J., Medina N. M. A., Rodríguez H. R., Pérez C. G. P., Sánchez E. G. (2024). Gestión de datos sobre denuncias con herramientas abiertas de software. Certamen de Innovación en Transparencia. Etapa de exposiciones.
<https://youtu.be/UZy0Wk5pN-E>
- [49] Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI) (2023). Iniciativa de datos abiertos: publicidad abierta.
https://drive.google.com/file/d/1vQ4z5xQoVaK8MYkpdixFjA7YkVf1_JyM/view?pli=1
- [50] Open Refine (versión 3.8). Herramienta gratuita y de código abierto para limpiar y transformar datos. <http://openrefine.org>.

Las Tablas A.1, A.2, A.3, A.4, A.5 y A.6 describen a las organizaciones auxiliares del Consejo De La Judicatura Del Poder Judicial Del Estado De Puebla.

Funciones:	Encargado de resguardar, organizar, depurar y controlar todos los expedientes, procesos, tocas, documentos y medios magnéticos y digitales, incluyendo los que sean catalogados con valor histórico, que le remitan para su custodia los órganos jurisdiccionales y demás dependencias del Poder Judicial.
Página web:	https://pjpuebla.gob.mx/poder-judicial/consejo-de-la-judicatura/archivo-judicial

Tabla A.1: *Archivo Judicial*

Funciones:	<p>Somos el órgano encargado de ayudar a los ciudadanos a encontrar soluciones reales, pacíficas y conforme a la ley respecto de sus conflictos. Dichos acuerdos a los que tienen la posibilidad de llegar las partes con la participación de una persona facilitadora, serán plasmados en convenios que tendrán la misma fuerza y alcance que una sentencia emitida por un juez.</p> <p>De igual forma, el CJA funciona como órgano especializado en mecanismos alternativos de solución de controversias en materia penal respecto de los delitos para los que la ley prevé una salida alterna y ya han sido judicializados, esto es, que el Ministerio Público, una vez concluida su investigación, haya determinado que contaba con elementos suficientes para iniciar la etapa de juicio.</p>
Página web:	https://cja.pjpuebla.gob.mx/

Tabla A.2: *Centro de Justicia Alternativa*

<p>Funciones:</p>	<p>Entrevista psico-diagnóstica (EDX) (Artículos 3 fracción XXIII, 11, 55 párrafo segundo, y 58 primer párrafo del Reglamento Interior)</p> <p>Convivencia modalidad Entrega-Recepción (Artículos 3 fracción XXII, 6 fracción IV, 64 y 65 del Reglamento Interior)</p> <p>Convivencia modalidad Presencial (Artículos 3 fracción XII, 23 fracción IV, 24 párrafo tercero, 56, 57 y 58 del Reglamento Interior)</p> <p>Convivencia modalidad virtual (Artículos 3 fracción XIII, 42 fracción XXV, 60 a 63 del Reglamento Interior)</p> <p>Asistencias a niñas, niños y adolescentes (Artículos 3 fracción II, 44, 71, 72 y 73 del Reglamento Interior)</p> <p>Evaluación psicológica (Artículos 3 fracciones XXV y XXXI, 66 y 67 del Reglamento Interior)</p> <p>Celebración de audiencias en Cámara Gesell (Artículos 3 fracción VI, 74 y 75 del Reglamento Interior)</p> <p>Evaluación Social (Artículos 3 fracción XXIV y XXVI y 68 del Reglamento Interior)</p> <p>Verificaciones domiciliarias (Artículos 3 fracción L y 69 del Reglamento Interior).</p>
<p>Página web:</p>	<p>https://ccf.pjpuebla.gob.mx/</p>

Tabla A.3: *Centro de Convivencia Familiar Supervisada*

Funciones:	Encargado de implementar los procesos de formación, capacitación y actualización del personal de carrera judicial, así como de diseñar, organizar y aplicar los concursos de oposición para acceder a los cargos y puestos de la estructura de carrera, en términos de la Ley y de los acuerdos generales respectivos, mediante un modelo educativo que se rija por los principios de excelencia, profesionalismo, idoneidad, mérito, antigüedad, paridad de género, no discriminación e igualdad de oportunidades.
Misión:	Impartir educación y fomento de la cultura de la legalidad para la profesionalización de la Carrera Judicial, para los estudios de posgrado, educación continua e investigación, así como para auxiliar al Consejo de la Judicatura en el proceso de certificación, capacitación, actualización y selección de las personas servidoras públicas del Poder Judicial, de quienes aspiren a pertenecer a éste y de usuarios en general.
Visión:	Ser una institución educativa de excelencia académica para los estudiosos del derecho, con reconocimiento en el ámbito jurídico nacional e internacional por formar a sus egresados con habilidades necesarias para enfrentar la realidad jurídica en diferentes contextos.
Página web:	https://ccf.pjpuebla.gob.mx/

Tabla A.4: *Escuela Estatal de Formación Judicial*

Funciones:	Garantizar una asistencia jurídica de calidad en asuntos del fuero local en materia penal, familiar, civil, laboral y de justicia cívica, mediante la asesoría, gestión, representación, patrocinio o defensa de las usuarias.
Misión:	Otorgar orientación, asesoría jurídica y defensa en materia penal y de justicia para adolescentes, así como patrocinio civil, familiar, mercantil, laboral, de justicia cívica y administrativo en cualquier etapa del procedimiento legal aplicable, que permite atender a la población menos favorecida de la entidad con los principios de legalidad, independencia técnica, gratuidad, igualdad y equilibrio procesal, con la finalidad de superar las desigualdades sociales y la consolidación del estado de derecho.
Visión:	Contar con una institución con profesionales con vocación de servicio, comprometidos con la excelencia en el desempeño de sus atribuciones conduciéndose dentro de los principios de gratuidad, probidad, independencia técnica, igualdad equidad, profesionalismo y confidencialidad que genere en la sociedad la confianza en sus asuntos representados.
Página web:	https://tja.pjpuebla.gob.mx/

Tabla A.5: *Instituto de la Defensoría Pública*

Funciones:	<p>La imposición de sanciones a servidores públicos por faltas administrativas graves, así como a particulares vinculados con las mismas, de conformidad a lo que establece la Ley General de Responsabilidades Administrativas en relación al Sistema Estatal Anticorrupción.</p> <p>Conocer de los juicios que promuevan los particulares en contra de actos y resoluciones de la administración pública del Estado y de los 217 municipios, cuando consideren que dichas autoridades no han actuado conforme a derecho.</p>
Misión:	<p>Impartir justicia administrativa en las controversias que se generen por actos o resoluciones de las autoridades administrativas del estado y sus municipios, que afecten a los particulares en su esfera jurídica, así como ser el tribunal que, como parte del sistema anticorrupción del Estado, resuelva la imposición de sanciones a los servidores públicos y a particulares relacionados con ellos por faltas administrativas graves.</p>
Visión:	<p>Consolidarse como el Tribunal Anticorrupción en el Estado, respecto a las responsabilidades administrativas de servidores públicos y particulares relacionados con ellos, por faltas administrativas graves y ser el Tribunal cercano a los particulares que se consideren afectados por actos o resoluciones de las autoridades administrativas susceptibles de impugnación.</p>
Página web:	<p>https://tja.pjpuebla.gob.mx/</p>

Tabla A.6: *Tribunal de Justicia Administrativa*

Anexo B: Búsqueda de DA en la PNT

La búsqueda de DA de instituciones judiciales del estado de Puebla en la PNT se implementó como sigue:

Paso 1. Se ingresó a la PNT al colocar en el navegador la dirección siguiente:
<https://www.plataformadetransparencia.org.mx/>

Paso 2. Se elige el ícono con la leyenda Información pública que se muestra en la Figura B.1, el cual se dirige a: <https://consultapublicamx.plataformadetransparencia.org.mx/vut-web/faces/view/consultaPublica.xhtml#inicio>. En la categoría Entidad o Federación se selecciona Puebla y Poder Judicial del Estado de Puebla en la opción Institución. En la sección Obligaciones, se escoge Generales, lo cual despliega la información completa.

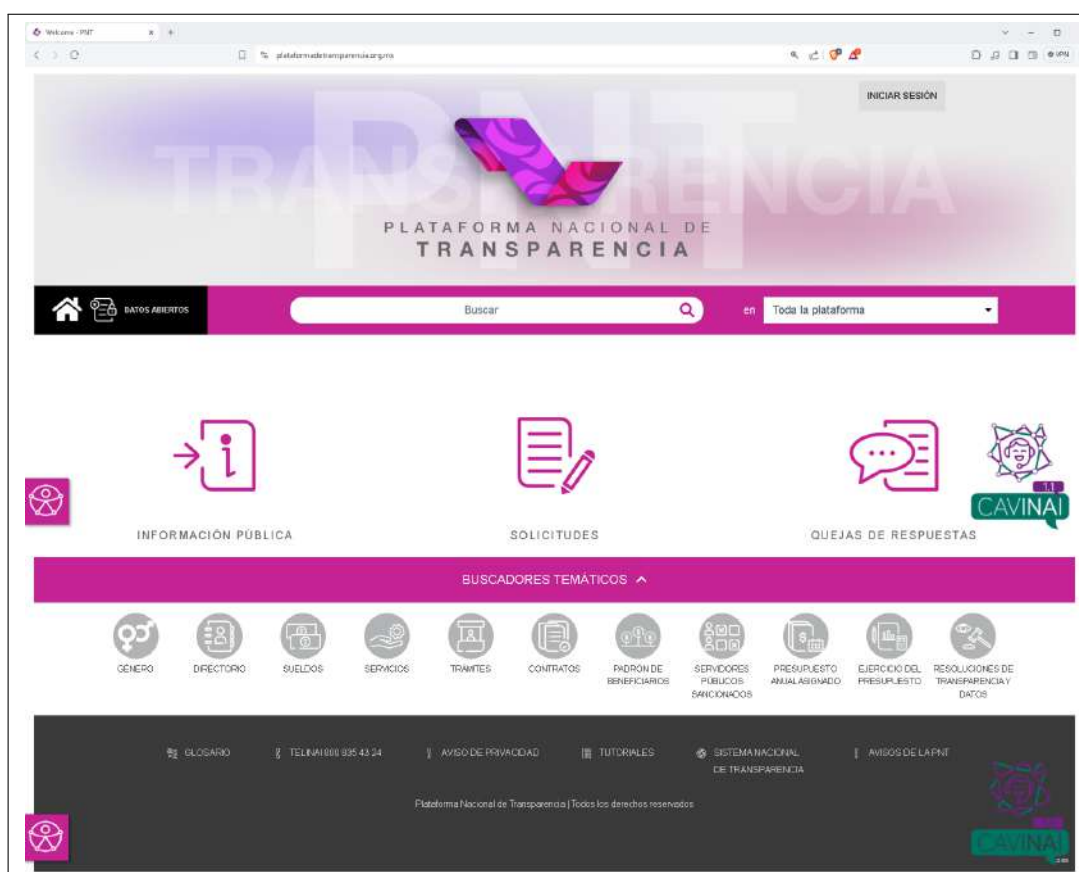


Figura B.1: Acceso a la sección de información pública de la PNT.

La plataforma contiene 38 íconos que representan los DA, 8 se muestran en la Figura B.2.

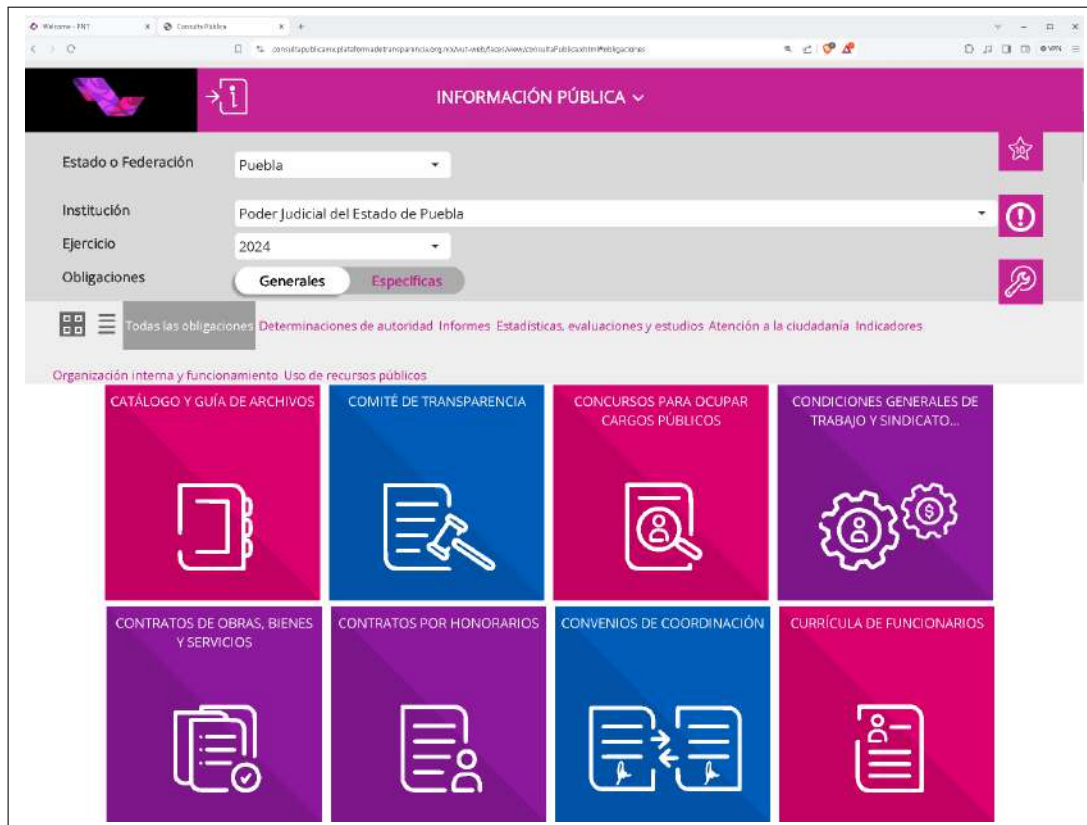


Figura B.2: Íconos para los DA disponibles en la PNT.

Paso 3. Un *click* sobre un ícono dirige al usuario a una tarjeta informativa como la que se ilustra en la Figura B.3.

Paso 4. Para obtener los datos, se presiona el botón Descarga, éste abre una ventana emergente que permite al usuario elegir el formato para los datos, .xls o .csv, (ver la Figura B.4).

Además, la plataforma cuenta con una opción para recibir los datos en el formato de archivo seleccionado por correo electrónico, para lo cual se requiere presionar el ícono con la figura de una carta. Posteriormente, se despliega una caja de texto para introducir el correo del destinatario.

Nota: la descarga de otro conjunto de DA requiere se repitan los pasos descritos.

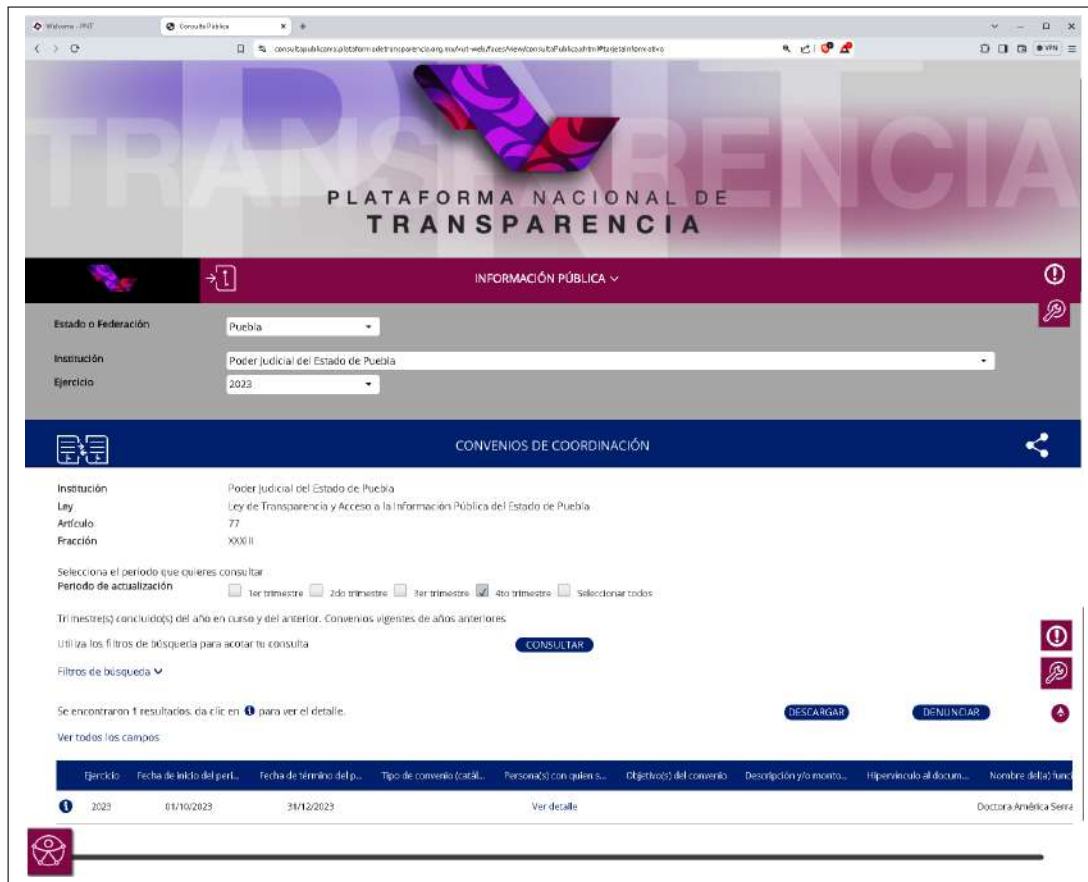


Figura B.3: Tarjeta informativa para un conjunto de DA en la PNT.

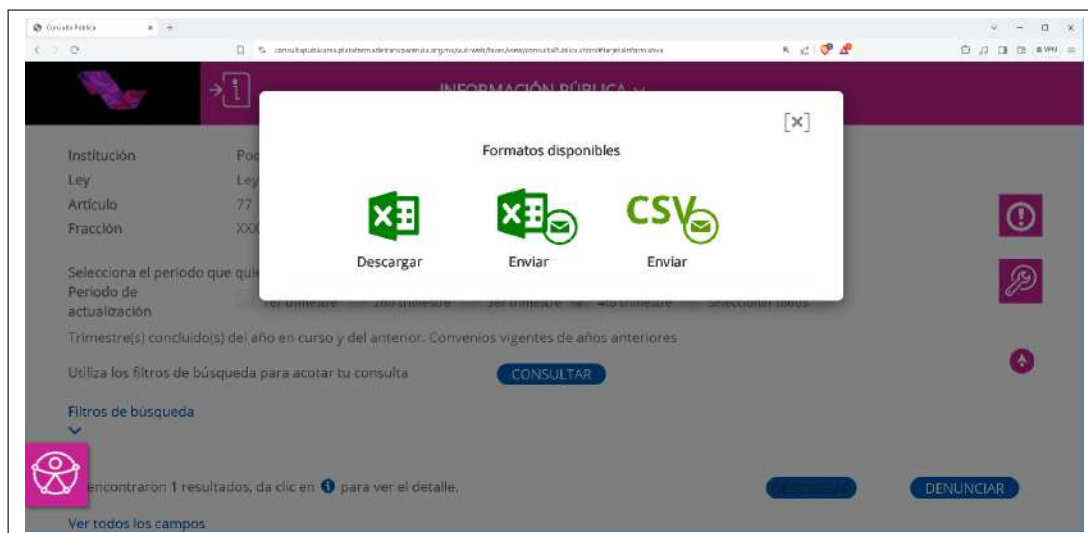


Figura B.4: Ventaja emergente para seleccionar el formato de los datos.

La herramienta OpenRefine es una aplicación de escritorio de código abierto que sirve para la limpieza y transformación de datos [50], su funcionalidad es similar a las hojas de cálculo y las base de datos puesto que opera en filas de datos que tienen celdas debajo de columnas. OpenRefine fue desarrollo de Google, tenía el nombre de Google Refine, actualmente, es un proyecto sostenido por la comunidad [?].

Las tareas sobre los datos se agrupan en *proyectos*, un proyecto se crea después de que se inicia el servidor web de manera local y se abre el navegador en la dirección <http://127.0.0.1:3333> [?].

C.1 Exploración de datos

Las Figuras C.1, C.2 y C.4 muestran aspectos relacionados con la exploración del conjunto DA-DENUNCIAS en la versión 3.8, notar que se utilizó la interfaz en el idioma por omisión, inglés.

Los números en color gris de la Figura C.1 son resultado de aplicar una faceta de texto sobre la columna MEDIO DE PRESENTACIÓN y de mostrar el resultado en orden descendente con base en el número en color gris, éste representa el número de registros o filas que contiene cada medio. No se encontró ninguna fila que no tuviera registrada alguno de los 8 medios.



Figura C.1: *Tipos de medio.*

El panel de la derecha de la Figura C.2 muestra el número total de filas: 171,231, notar que la primera fila contiene a los nombres de las columnas (encabezados). Como resultado

de ordenar los datos en la columna TIPO DE FALTA, el panel de la izquierda despliega la leyenda 59 choices, esto indica que existen 59 tipos de falta o denuncia distintos. Las faltas ordenadas por nombre se muestran en color azul, seguidas del número de filas en gris.

The screenshot shows the OpenRefine interface for a dataset named 'OpenRefine falta'. The main panel displays a facet for 'TIPO DE FALTA' with 59 choices, sorted by name. The facet is currently showing the first 10 items, which are listed in blue text with their corresponding row counts in grey text. The items are:

TIPO DE FALTA	Count
Abuso de funciones	685
Abuso o acoso sexual	2758
Actuación bajo conflicto de interés	29
Causar danos y perjuicios de manera culposa o negligente a la hacienda y patrimonio público	136
Cohecho	140
Colusión entre particulares en materia de contrataciones públicas	44
Conflicto de intereses	1796
Contratación indebida	100
Contratación indebida de ex servidores públicos	24
Desacato	76
Descuido o falta de atención en el cumplimiento de las obligaciones	38946
Desvío de recursos públicos	155
Encubrimiento	34
Enriquecimiento inexplicable	570
Enriquecimiento oculto u ocultamiento de conflicto de interés	25
Exceso en el cumplimiento de	

On the right side, a context menu is open over the 'TIPO DE FALTA' column. The menu items are:

1. Facet
2. Text filter
3. Edit cells
4. Edit column
5. Transpose
6. Sort...
7. View
8. Reconcile

A red arrow points to the 'Sort...' option in the menu.

Figura C.2: Tipos de falta.

Se encontró una inconsistencia en la falta “Uso indebido de recursos públicos”, puesto que ésta se encontraba escrita en las dos representaciones distintas que muestra la Figura C.3. Se decidió unificar el primer valor con el segundo.

La parte superior de la Figura C.4 muestra el número de instituciones que tienen denuncias, cada institución se identifica con siglas, algunas se aprecian en la parte inferior.

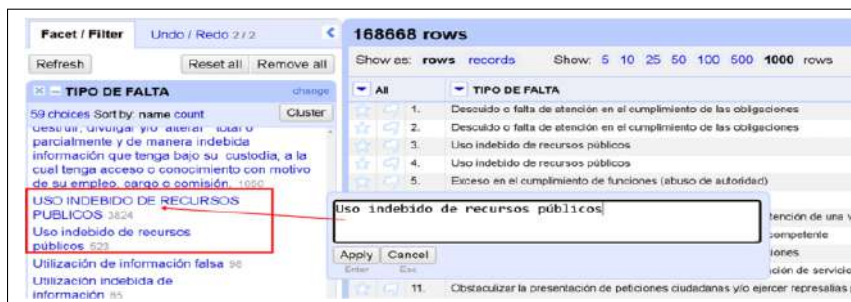


Figura C.3: Inconsistencia en la falta “Uso indebido de recursos públicos”.



Figura C.4: Número de instituciones con denuncias.

C.2 Transformaciones en DA-DENUNCIAS

Las tareas de limpieza de datos fueron: integración de acentos, uso homogéneo de mayúsculas y minúsculas, así como eliminación de espacios en blanco antes y después del texto en las celdas. Las transformaciones realizadas fueron las siguientes:

1. Se insertaron espacio en blanco en las cadenas siguientes:
 - “S.A.de C.V.” → “S. A. de C. V.”
 - “S.A.” → “S. A.”
 - “S.N.C.” → “S. A.”
 - “S.C.” → “S. C.”
2. Se colocaron acentos en la columna “siglas” en los valores de:
 - “API LAZARO CÁRDENAS”
 - “API MAZATLÁN”
3. Se agregó un punto después de Ma. en “Instituto de Investigaciones Dr. José Ma. Luis Mora”
4. Se modificó “P M I COMERCIO INTERNACIONAL SA DE CV” por “PMI Comercio Internacional, S.A. de C.V.”
5. Se modificó “SECRETARIA DE SEGURIDAD PUBLICA” por “Secretaría de Seguridad Pública”
6. Se agregó una coma en: “Instituto Nacional de Astrofísica, Óptica y Electrónica”
7. Se colocó la letra “L” en “CINETECA NACIONAL” → “CINETECA NACIONAL”
8. Se acentuó “PROMÉXICO”

C.2.1 Modificaciones

Las modificaciones realizadas sobre los datos fueron las siguientes:

- A los 2563 registros de denuncias en donde no se encontró un tipo de falta, es decir, registros con valor *blank* de la Figura C.5, se les asignó un identificador de falta igual a 100, el cual está fuera del dominio (1, 59)
- Actualmente, el “Consejo Nacional de Ciencia y Tecnología” se conoce como “Consejo Nacional de Humanidades, Ciencia y Tecnología”. También se modificó el nombre del “Instituto Federal de Acceso a la Información” (INAI)



Figura C.5: Número de registros sin tipo de falta: 2,563.

Para importar el archivo `sidecV2.sql` resultado de la normalización del conjunto DA-DENUNCIAS de la SFP en el gestor de base de datos SQL Server V14.0, se implementaron los pasos siguientes:

Paso 1. Se abrió el gestor de base de datos SQL Server V14.0 y se seleccionó la opción `New Database` para crear una base de datos nueva. Se ingresó el nombre `SidecV2`, (ver la Figura D.1).

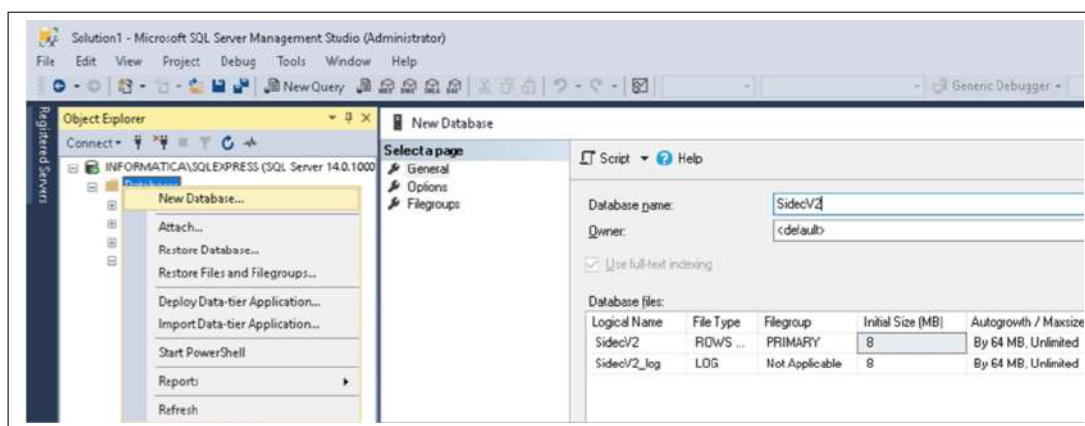


Figura D.1: Creación de base de datos nueva.

Paso 2. Una vez creada la base de datos, se seleccionó la opción `Open File` y se abrió el archivo `sidecV2.sql`, éste contiene las sentencias necesarias para crear las entidades medio, falta, institución, denuncia y queja, así como las instrucciones para insertar los datos (ver la Figura D.2).

Paso 3. Una vez abierto el archivo, se seleccionó la opción `Edit`, luego `Find and Replace` y después `Quick Replace` para reemplazar las comillas dobles por comillas simples (Figura D.3), eliminar los paréntesis (Figura D.4) y las expresiones `INSERT INTO` denuncias values y construir un archivo con formato csv como se muestra en la Figura D.5. La información del archivo de texto se almacenó en `SidecV2.csv`, (ver la Figura D.6).

Paso 4. Finalmente, se ejecutó la sentencia `BULK INSERT` para importar el archivo `SidecV2.csv` que contiene los metadatos y datos de las 171,230 denuncias (ver la Figura D.7).

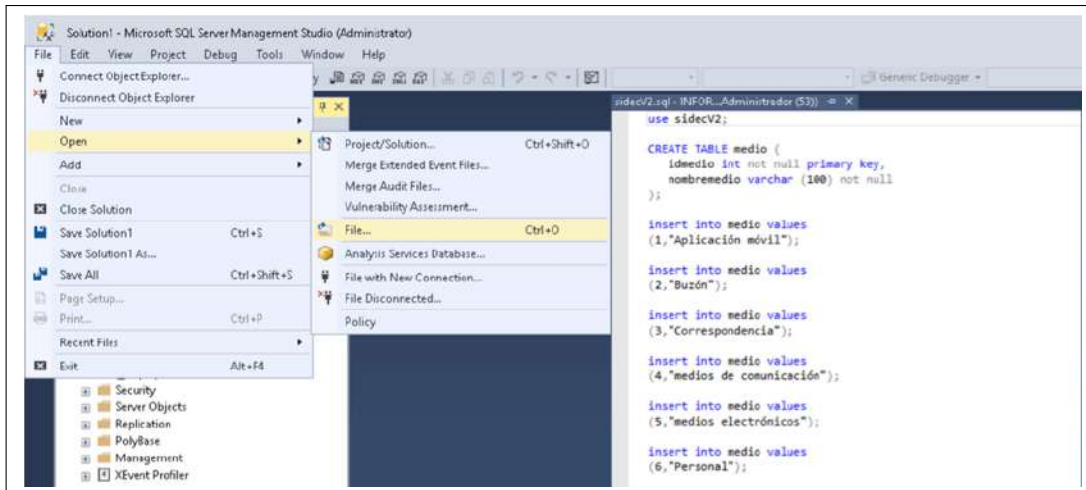


Figura D.2: Apertura del archivo sidecV2.sql.

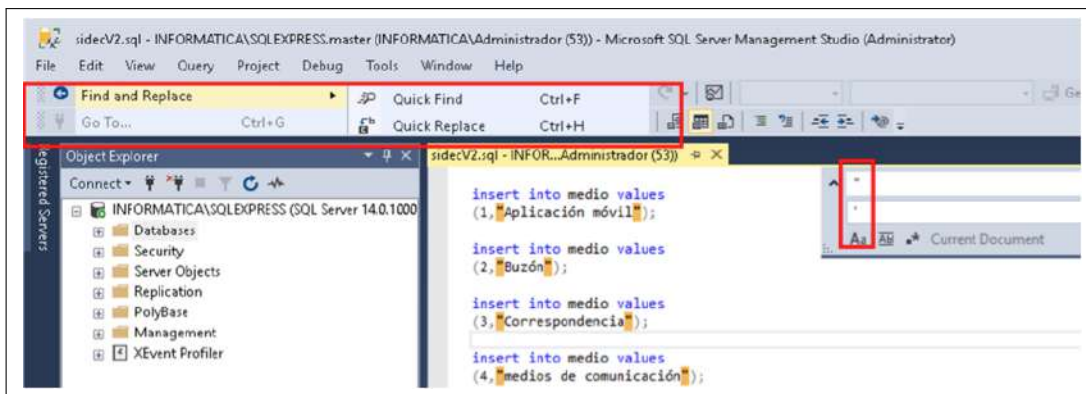


Figura D.3: Reemplazo de comillas dobles por comillas simples.

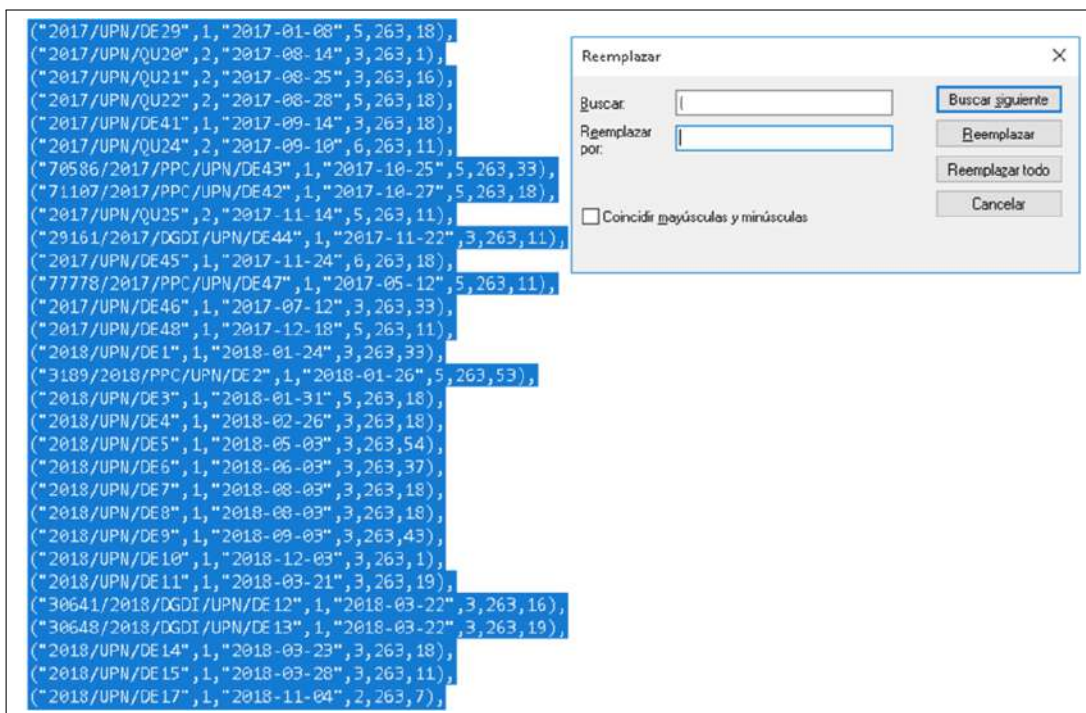


Figura D.4: Eliminación de paréntesis izquierdo.

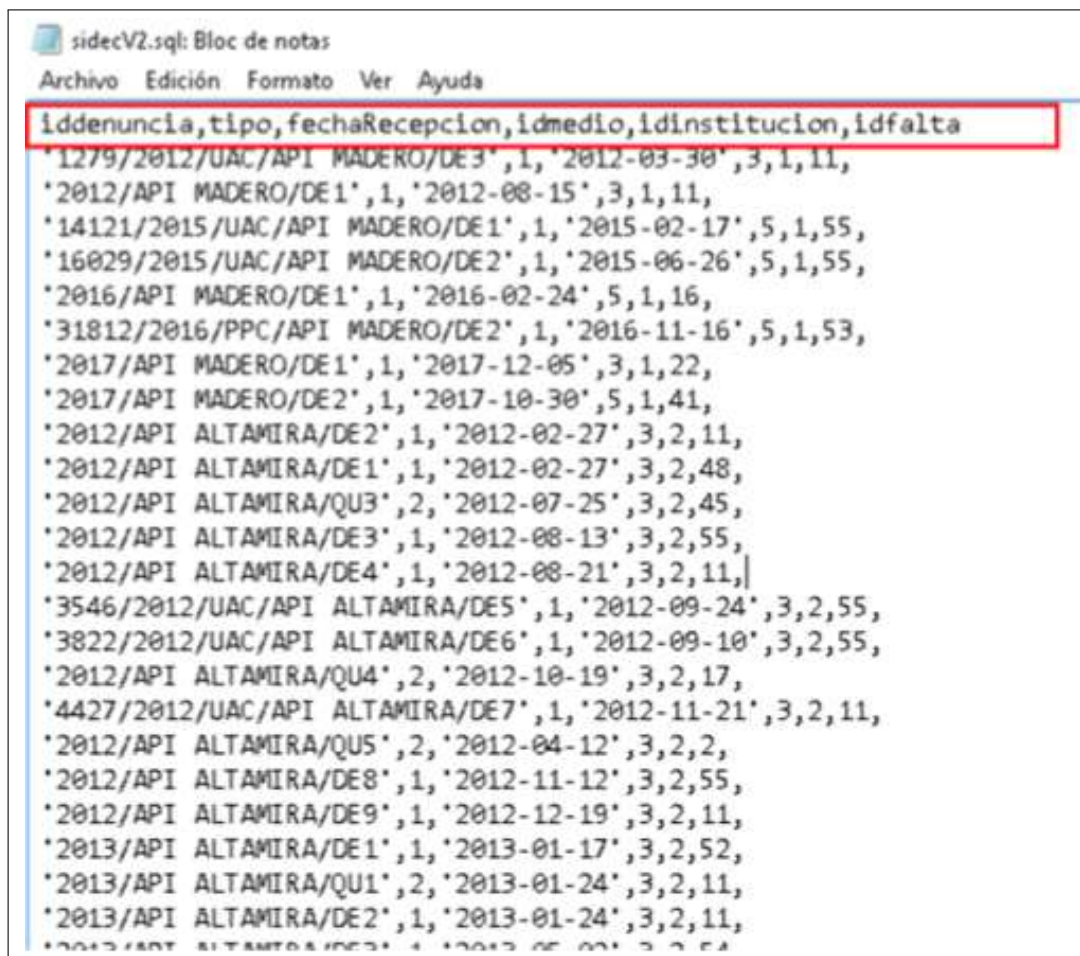


Figura D.5: Archivo con formato csv en editor de texto.

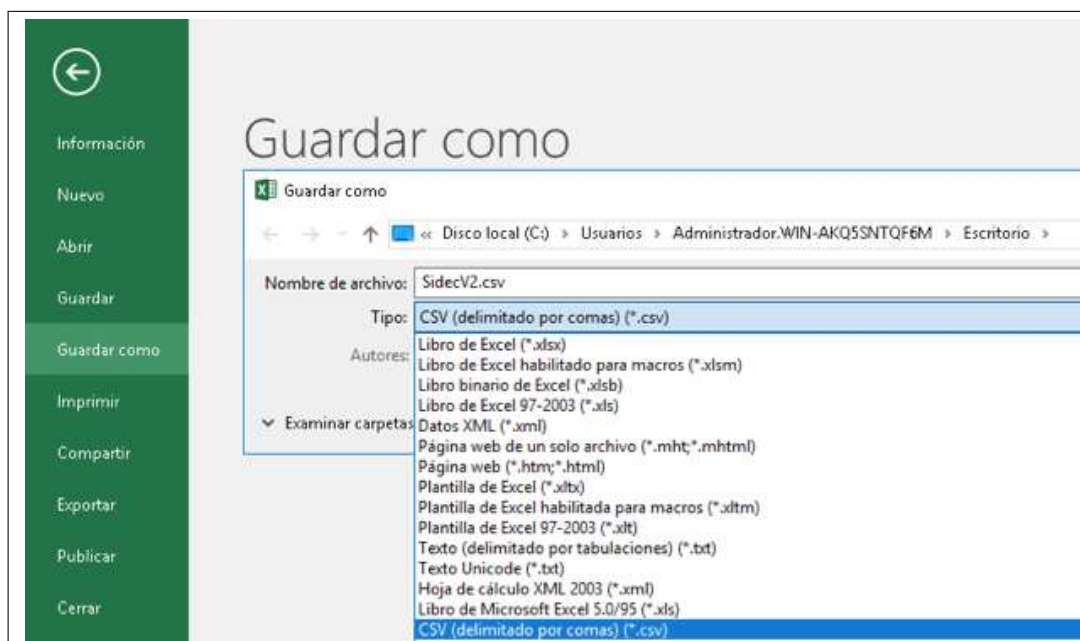


Figura D.6: Archivo con formato csv en hoja de cálculo.

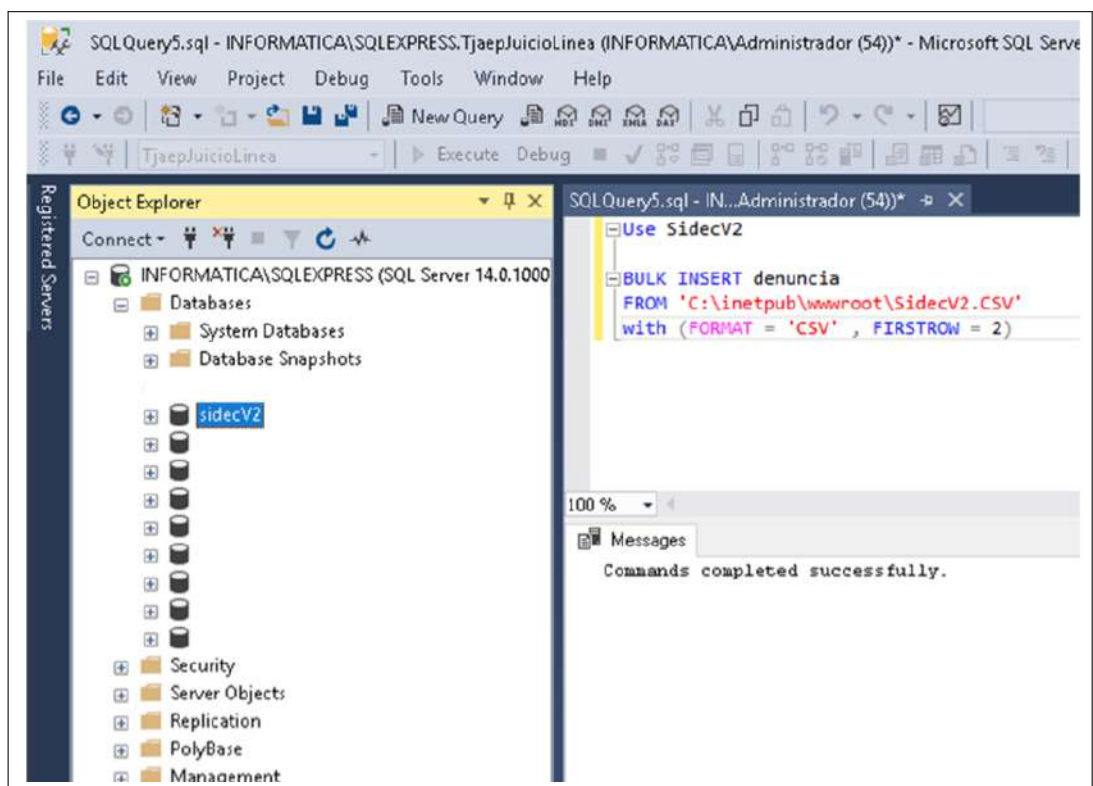


Figura D.7: Importación de csv a SQL Server.

Anexo E: Creación de tablas y procedimientos almacenados para Denuncias-Relacional

El Anexo E contiene el código para transformar DA-DENUNCIAS en DENUNCIAS-RELACIONAL, fue elaborado por la M. C. Rebeca Rodríguez Huesca, co-asesora de esta tesis y verificado por el autor.

E.1 Creación de la base de datos

Las instrucciones siguientes crean una base de datos denominada completa con una tabla del mismo nombre que almacena los datos de DA-Denuncias.

```
CREATE DATABASE IF NOT EXISTS completa;
USE completa;
CREATE TABLE completa (id INT AUTO_INCREMENT,
                        institucion VARCHAR(250) NOT NULL,
                        siglas VARCHAR(50) NOT NULL,
                        expediente VARCHAR(50) NOT NULL,
                        medio VARCHAR(30) NOT NULL,
                        tipo_peticion VARCHAR(50) NOT NULL,
                        tipo_falta VARCHAR(240) NOT NULL,
                        fecha_recepcion DATE NOT NULL,
                        PRIMARY KEY (id)
);

SOURCE /home/rebeca/sidec2.sql;
```

E.2 Limpieza de datos

El procedimiento corrige_datos automatiza las tareas de limpieza de datos realizadas inicialmente en OpenRefine, (ver el Anexo C). En el código se omiten los acentos por incompatibilidad con el editor de este documento.

```
DELIMITER $$
DROP PROCEDURE IF EXISTS corrige_datos$$
CREATE PROCEDURE corrige_datos()
BEGIN
```

```
DECLARE done INT DEFAULT FALSE;
DECLARE inst1, sig1, inst2, sig2 VARCHAR (250);
DECLARE pos INT DEFAULT 0;
DECLARE id1 INT;

DECLARE cur1 CURSOR FOR SELECT id, institucion, siglas FROM completa;
DECLARE CONTINUE HANDLER FOR NOT FOUND SET done= TRUE;

OPEN cur1;

WHILE NOT done DO
    FETCH cur1 INTO id1, inst1, sig1;
    IF NOT done THEN
        SET pos = LOCATE("S.A.de C.V.", inst1);
        IF pos <> 0 THEN
            SET inst2 = REPLACE(inst1, 'S.A.de C.V.', 'S. A. de C. V. ');
        UPDATE completa SET institucion = inst2 WHERE id = id1;
        ELSE
            SET pos = LOCATE("S.A.", inst1);
            IF pos <> 0 THEN
                SET inst2 = REPLACE(inst1, 'S.A.', 'S. A. ');
            UPDATE completa SET institucion = inst2 WHERE id = id1;
            ELSE
                SET pos = LOCATE("S.N.C.", inst1);
                IF pos <> 0 THEN
                    SET inst2 = REPLACE(inst1, 'S.N.C.', 'S. N. C. ');
                UPDATE completa SET institucion = inst2 WHERE id = id1;
                ELSE
                    SET pos = LOCATE("S.C.", inst1);
                    IF pos <> 0 THEN
                        SET inst2 = REPLACE(inst1, 'S.C.', 'S. C. ');
                    UPDATE completa SET institucion = inst2 WHERE id = id1;
                    ELSE
                        SET pos = LOCATE("Dr. Jose Ma Luis Mora", inst1);
                        IF pos <> 0 THEN
                            SET inst2 = REPLACE(inst1, 'Dr. Jose Ma Luis Mora',
                                'Dr. Jose Ma. Luis Mora ');
                        UPDATE completa SET institucion = inst2 WHERE id = id1;
                        ELSE
                            SET pos = LOCATE("P M I COMERCIO INTERNACIONAL SA DE CV", inst1);
                            IF pos <> 0 THEN
```



```

UPDATE completa SET siglas = sig2 WHERE id = id1;
ELSE
    SET pos = LOCATE("CINETECA NACIONAL", sig1);
    IF pos <> 0 THEN
        SET sig2 = REPLACE(sig1, 'CINETECA NACIONAL', 'CINETECA NACIONAL');
    UPDATE completa SET siglas = sig2 WHERE id = id1;
    ELSE
        SET pos = LOCATE("PROMEXICO", sig1);
        IF pos <> 0 THEN
            SET sig2 = REPLACE(sig1, 'PROMEXICO', 'PROMEXICO');
            UPDATE completa SET siglas = sig2 WHERE id = id1;
        END IF;
    END IF;
END IF;
END IF;
END IF;
END WHILE;
CLOSE cur1;
UPDATE completa SET tipo_falta ="Se desconoce" WHERE tipo_falta like'';
END $$
DELIMITER ;

```

E.3 Creación de tablas padre

Las tablas padre son: falta, institucion y medio, éstas se crean con los procedimientos almacenados siguientes:

E.3.1 Tabla falta

```

DELIMITER $$
DROP PROCEDURE IF EXISTS crea_faltas$$
CREATE PROCEDURE crea_faltas()
BEGIN
    DECLARE done INT DEFAULT FALSE;
    DECLARE falta1 VARCHAR (240);

    DECLARE cur1 CURSOR FOR SELECT distinct tipo_falta FROM completa;

```

```

DECLARE CONTINUE HANDLER FOR NOT FOUND SET done= TRUE;
DROP TABLE IF EXISTS falta;
CREATE TABLE falta (
    tipoFalta INT AUTO_INCREMENT NOT NULL,
    nombreFalta VARCHAR(240) NOT NULL,
    PRIMARY KEY (tipoFalta));
OPEN cur1;
WHILE NOT done DO
    FETCH cur1 INTO falta1;
    IF NOT done THEN
INSERT INTO falta (nombreFalta) VALUES (falta1);
END IF;
END WHILE;
CLOSE cur1;
END $$
DELIMITER ;

```

E.3.2 Tabla institucion

```

DELIMITER $$
DROP PROCEDURE IF EXISTS crea_instituciones$$
CREATE PROCEDURE crea_instituciones()
BEGIN
    DECLARE done INT DEFAULT FALSE;
    DECLARE inst1 VARCHAR (120);
    DECLARE sig1 VARCHAR (50);

    DECLARE cur1 CURSOR FOR SELECT distinct institucion, siglas FROM completa
    ORDER BY institucion;

    DECLARE CONTINUE HANDLER FOR NOT FOUND SET done= TRUE;
    DROP TABLE IF EXISTS institucion;
    CREATE TABLE institucion (
        idI INT AUTO_INCREMENT NOT NULL,
        nombreInstitucion VARCHAR(120) NOT NULL,
        siglas VARCHAR(50) NOT NULL,
        PRIMARY KEY (idI));

```

```

OPEN cur1;
WHILE NOT done DO
    FETCH cur1 INTO inst1, sig1;
    IF NOT done THEN
        INSERT INTO institucion (nombreInstitucion, siglas) VALUES (inst1, sig1);
    END IF;
END WHILE;
CLOSE cur1;
END $$
DELIMITER ;

```

E.3.3 Tabla medio

```

DELIMITER $$
DROP PROCEDURE IF EXISTS crea_medios$$
CREATE PROCEDURE crea_medios()
BEGIN
    DECLARE done INT DEFAULT FALSE;
    DECLARE medio1 VARCHAR (30);

    DECLARE cur1 CURSOR FOR SELECT distinct medio FROM completa
    ORDER BY medio;

    DECLARE CONTINUE HANDLER FOR NOT FOUND SET done= TRUE;

    DROP TABLE IF EXISTS medio;

    CREATE TABLE medio (
        idM INT AUTO_INCREMENT NOT NULL,
        nombreMedio VARCHAR(30) NOT NULL,
        PRIMARY KEY (idM));

    OPEN cur1;
    WHILE NOT done DO
        FETCH cur1 INTO medio1;
        IF NOT done THEN
            INSERT INTO medio (nombreMedio) VALUES (medio1);
        END IF;
    END WHILE;
    CLOSE cur1;

```

```
END $$  
DELIMITER ;
```

E.4 Creación de tablas hijas

Las tablas hijas son: denuncia y queja, éstas se crean en los procedimientos crea_denuncias y crea_quejas, los cuales implementan mecanismos de control de errores y transacciones para garantizar la integridad y consistencia de los datos.

E.4.1 Tabla denuncia

```
DELIMITER $$  
DROP PROCEDURE IF EXISTS crea_denuncias$$  
CREATE PROCEDURE crea_denuncias()  
BEGIN  
    DECLARE done INT DEFAULT FALSE;  
    DECLARE falta1 VARCHAR (240);  
    DECLARE medio1 VARCHAR (30);  
    DECLARE inst1 VARCHAR (120);  
    DECLARE id1, idF, med, instit INT;  
    DECLARE exp VARCHAR (50);  
    DECLARE fecha DATE;  
  
    DECLARE cur1 CURSOR FOR SELECT id,tipo_falta, medio, institucion, expediente, fecha_recepcion  
    FROM completa  
    WHERE tipo_peticion LIKE "Denuncia%";  
  
    DECLARE CONTINUE HANDLER FOR NOT FOUND SET done= TRUE;  
    DECLARE EXIT HANDLER FOR 1064  
    BEGIN  
        SELECT "EL ID NO SE ENCUENTRA";  
        SET done= TRUE;  
        ROLLBACK;  
    END;  
  
START TRANSACTION;  
    DROP TABLE IF EXISTS denuncia;
```



```
CREATE TABLE denuncia (  
    idD INT NOT NULL,  
    tipoFalta INT NOT NULL,  
    idM INT NOT NULL,  
    idI INT NOT NULL,  
    denuncia VARCHAR(50) NOT NULL,  
    fecha_recepcion DATE NOT NULL,  
    PRIMARY KEY (idD),  
    FOREIGN KEY (tipoFalta)  
    REFERENCES falta (tipoFalta)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION,  
    FOREIGN KEY (idM)  
    REFERENCES medio (idM)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION,  
    FOREIGN KEY (idI)  
    REFERENCES institucion (idI)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION);  
  
OPEN cur1;  
WHILE NOT done DO  
    FETCH cur1 INTO id1, falta1, medio1, inst1, exp, fecha;  
    IF NOT done THEN  
        SET idF = (SELECT tipoFalta FROM falta WHERE nombreFalta = falta1);  
SET med = (SELECT idM FROM medio WHERE nombreMedio like medio1);  
    SET instit = (SELECT idI FROM institucion WHERE nombreInstitucion = inst1);  
  
    INSERT INTO denuncia (idD, tipofalta, idM, IdI, denuncia, fecha_recepcion)  
        VALUES (id1, idF, med, instit, exp, fecha);  
  
END IF;  
END WHILE;  
CLOSE cur1;  
COMMIT;  
END $$  
DELIMITER ;
```

E.4.2 Tabla queja

```

DELIMITER $$
DROP PROCEDURE IF EXISTS crea_quejas$$
CREATE PROCEDURE crea_quejas()
BEGIN
    DECLARE done INT DEFAULT FALSE;
    DECLARE falta1 VARCHAR (240);
    DECLARE medio1 VARCHAR (30);
    DECLARE inst1 VARCHAR (120);
    DECLARE id1, idF, med, instit INT;
    DECLARE exp VARCHAR (50);
    DECLARE fecha DATE;

    DECLARE cur1 CURSOR FOR SELECT id, tipo_falta, medio, institucion, expediente, fecha_recepcion
    FROM completa
    WHERE tipo_peticion LIKE "Queja%";

    DECLARE CONTINUE HANDLER FOR NOT FOUND SET done= TRUE;
    DECLARE EXIT HANDLER FOR 1064
    BEGIN
        SELECT "EL ID NO SE ENCUENTRA";
        SET done= TRUE;
        ROLLBACK;
    END;

START TRANSACTION;
DROP TABLE IF EXISTS queja;
CREATE TABLE queja (
    idD INT NOT NULL,
    tipoFalta INT NOT NULL,
    idM INT NOT NULL,
    idI INT NOT NULL,
    queja VARCHAR(50) NOT NULL,
    fecha_recepcion DATE NOT NULL,
    PRIMARY KEY (idD),
    FOREIGN KEY (tipoFalta)
    REFERENCES falta (tipoFalta)

```

```
        ON DELETE NO ACTION
        ON UPDATE NO ACTION,
        FOREIGN KEY (idM)
        REFERENCES medio (idM)
        ON DELETE NO ACTION
        ON UPDATE NO ACTION,
        FOREIGN KEY (idI)
        REFERENCES institucion (idI)
        ON DELETE NO ACTION
        ON UPDATE NO ACTION);

OPEN cur1;
WHILE NOT done DO
    FETCH cur1 INTO id1, falta1, mediol, inst1, exp, fecha;
    IF NOT done THEN
        SET idF = (SELECT tipoFalta FROM falta WHERE nombreFalta = falta1);
        SET med = (SELECT idM FROM medio WHERE nombreMedio like mediol);
        SET instit = (SELECT idI FROM institucion WHERE nombreInstitucion = inst1);

        INSERT INTO queja (idD,tipofalta, idM, IdI, queja, fecha_recepcion)
            VALUES (id1, idF, med, instit, exp, fecha);

    END IF;
END WHILE;
CLOSE cur1;
COMMIT;
END $$
DELIMITER ;
```