

Informe final de hallazgos

Centro de Investigación y Docencia Económicas

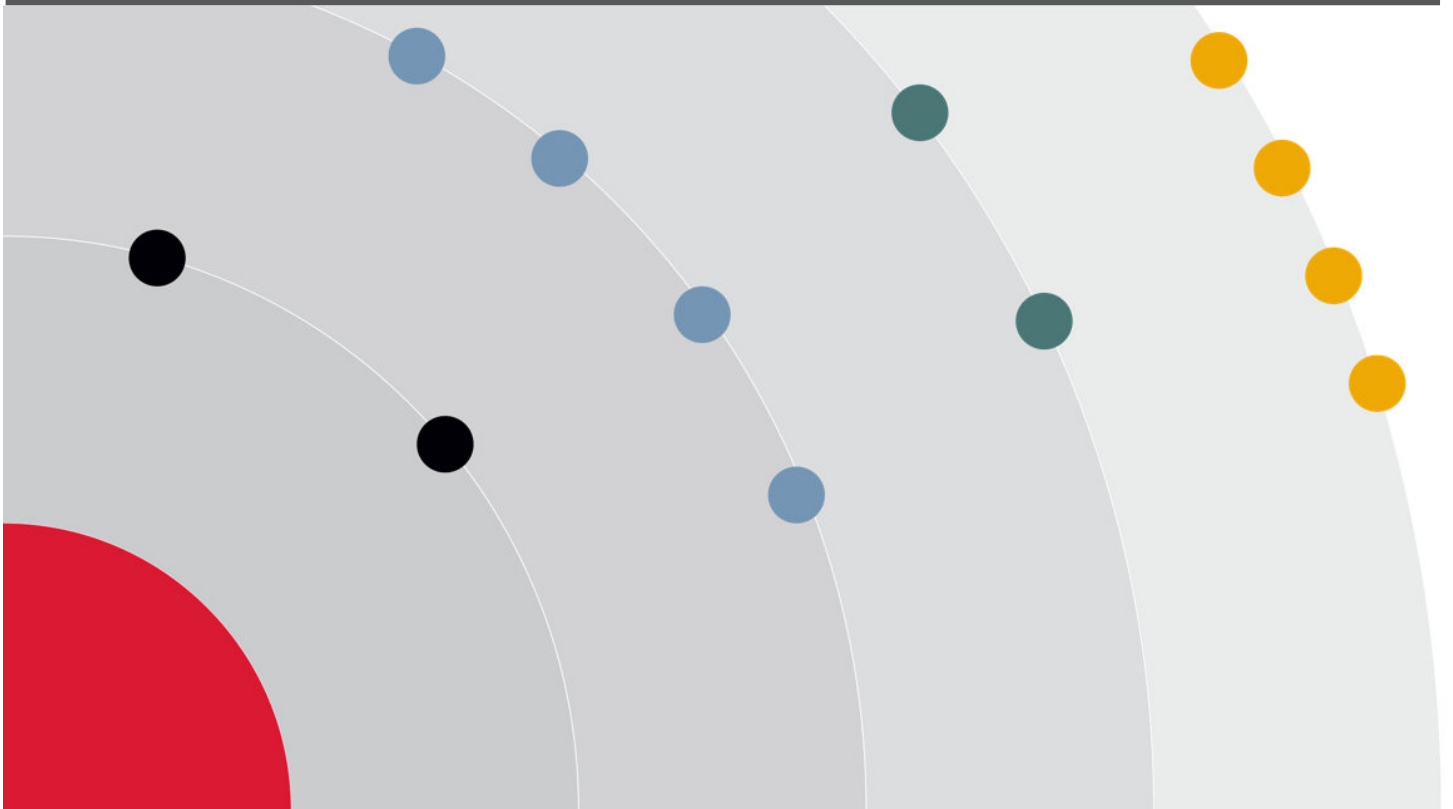


Inventario de sistemas electrónicos de información y bases de datos existentes a cargo de los entes públicos del orden federal que servirán como base para la conformación de la Plataforma Digital Nacional



PREPARADO PARA
LA SECRETARÍA EJECUTIVA DEL SISTEMA NACIONAL ANTICORRUPCIÓN

**“NO ES POSIBLE ADMINISTRAR LO INNOMBRABLE.
Y, COMO COROLARIO DE ELLO, NO SE PUEDE
GESTIONAR LO QUE NO SE DEFINE EXPLÍCITAMENTE.”
GWEN THOMAS**



“La Gobernanza de datos es un sistema de decisión de derechos y responsabilidades para procesos vinculados a la generación de información ejecutados conforme a modelos preestablecidos que describen quién puede tomar qué acciones con base en qué información, y cuando, bajo qué circunstancias y utilizando qué métodos y procesos.”

DATA MANAGEMENT INSTITUTE

CONTENIDOS

Introducción	1
Métodos de levantamiento de información	4
Perspectiva metodológica del levantamiento de información	2
Alcances del estudio	8
Panorama de sistemas relevantes	13
Hallazgos clave por área temática	60
Diagnóstico del panorama de la gestión de datos	76
Recomendaciones	100
Buenas prácticas	104
Aplicación de buenas prácticas	109
Anexos	117

INTRODUCCIÓN

UNA MIRADA FRESCA AL ENTORNO

¿POR QUÉ ANALIZAR EL ENTORNO EXISTENTE?

La publicación de la Ley General del Sistema Nacional Anticorrupción (LGSNA o Ley) en 2016 creó una nueva serie de responsabilidades y atribuciones sobre actores nuevos y existentes del entorno institucional mexicano. En el Título cuarto de esta Ley se establece la necesidad de construir una Plataforma Digital Nacional que servirá, entre otras cosas, para administrar la información requerida para la operación del Sistema Nacional Anticorrupción (SNA) con base en los siguientes tipos de sistemas:

1. Sistema de evolución patrimonial, de declaración de intereses y constancia de presentación de declaración fiscal;
2. Sistema de los servidores públicos que intervengan en procedimientos de contrataciones públicas;
3. Sistema nacional de servidores públicos y particulares sancionados;
4. Sistema de información y comunicación del Sistema Nacional y del Sistema Nacional de Fiscalización;
5. Sistema de denuncias públicas de faltas administrativas y de hechos de corrupción, y
6. Sistema de información pública de contrataciones.

Esta definición temática integra y define un nuevo ecosistema de información que se desarrollará bajo el auspicio de las instituciones sujetas a esta normatividad. Un ecosistema de información hace referencia al conjunto de soluciones operativas, tecnológicas y de gestión en conjunto con la infraestructura

tecnológica y las metodologías implementadas para el manejo de todos los aspectos del ciclo de vida de la información en un área de interés definida y común.

Sin embargo, este ecosistema no surge en un vacío. Actualmente, múltiples entes federales desarrollan actividades de levantamiento, tratamiento y explotación de información dentro de las áreas temáticas clave definidas por la LGSNA. Estas actividades fueron resultado de políticas independientes de regulación institucional que se implementaron en respuesta a coyunturas localizadas. Por tal motivo, no tuvieron oportunidad de operar bajo un marco común ni desarrollarse de manera homogénea. Todo ello, aunado a diferencias de objetivos institucionales y de asignaciones presupuestarias ha resultado en una estructura poco uniforme y en sistemas de distinto nivel de madurez.

¿QUÉ EVALUAMOS DEL ENTORNO?

Considerando esta situación no homogénea, la implementación de cualquier medida conducente a la integración de la Plataforma Digital Nacional del Sistema Nacional Anticorrupción requiere, primero, el desarrollo de un diagnóstico en 4 líneas de análisis:

1. Detección de los sistemas electrónicos y bases de datos con datos de las áreas temáticas clave
2. Diagnóstico de la completitud de los esfuerzos por recuperar información en las áreas temáticas establecidas por la Ley
3. Diagnóstico del nivel de preparación que muestra cada sistema para formar parte de acciones de interconexión, interoperabilidad y transferencia

4. Diagnóstico de la confiabilidad de los datos en los sistemas

La primera línea de análisis representa la descripción del universo de sistemas considerados. Por otro lado, las segunda, tercera y cuarta líneas de análisis buscan, de manera conjunta, diagnosticar que situación guarda el panorama de la gestión de datos en las áreas temáticas definidas por la Ley. Con la atención de estas líneas de investigación, el presente estudio busca hacer un diagnóstico comprensivo de la situación que guarda el ecosistema de información delimitado por el artículo 49 de la LGSNA.

¿CÓMO SE REALIZAN ESTOS DIAGNÓSTICOS?

Detección de sistemas y bases de datos relevantes

La detección de sistemas y bases de datos se desarrolló mediante la aplicación de cuestionarios a las instituciones participantes. Este proceso resultó en un amplio listado de sistemas y bases de datos cuyo contenido es relevante para el ecosistema de información de la LGSNA.

Diagnóstico de completitud temática

La completitud temática se determinó mediante una clasificación de los sistemas por perfil de información que reportaron capturar dentro de las áreas temáticas de interés del estudio. Con base en esta información, se calculó la proporción de las áreas temáticas que cubría cada institución con sus actividades de levantamiento, procesamiento y explotación de información.

Diagnóstico de complementariedad informática

Uno de los objetivos fundamentales del estudio es evaluar en qué medida los sistemas y bases de datos en poder de las instituciones participantes cuentan con las condiciones para formar parte de los esfuerzos de transferencia, interconexión e interoperabilidad que se cree que potencialmente se podrían requerir para el cumplimiento de la LGSNA.

Por tal motivo, como parte del estudio se definió un estimador que permitiera evaluar la preparación con que cuentan los sistemas detectados en materia normativa y documental para integrarse en estos esfuerzos. Esta medida de preparación se conceptualizó como índice de complementariedad informática y su cálculo se basa en los reactivos reportados por las instituciones como parte del proceso de detección y levantamiento de información.

Diagnóstico de confiabilidad de datos

La confiabilidad de la información es una consideración central para determinar el valor de un conjunto de datos. Por lo tanto, se consideró central al estudio la creación de una herramienta de diagnóstico que permitiera valorar la medida en que la información depositada en los sistemas y bases de datos detectados está protegida por procesos y políticas que aseguren su integridad, validez y veracidad. Esta métrica se conceptualizó bajo el nombre de índice de confiabilidad de datos y se nutrió de la información reportada por las instituciones en los cuestionarios del proceso de levantamiento de información.

¿CÓMO SE ORGANIZA EL DOCUMENTO?

El presente documento está organizado en 5 secciones.

La primera sección describe los métodos de levantamiento de información utilizados para la detección de sistemas y bases de datos con el apoyo de las instituciones participantes en el estudio.

La segunda sección define los alcances del estudio con relación a las organizaciones participantes y temáticas abordadas. Asimismo, se especifican los casos en los que no se pudo realizar el levantamiento de información o en los que éste se realizó de manera parcial.

La tercera sección presenta el listado de sistemas detectados, así como consideraciones con respecto a su análisis. Asimismo, se presenta un diagnóstico descriptivo de las características generales de los sistemas electrónicos y bases de datos detectados por las instituciones participantes.

La cuarta sección presenta el diagnóstico del panorama de gestión de datos. En este apartado, se exponen los hallazgos derivados del diagnóstico y se profundiza en las definiciones metodológicas y en los resultados del cálculo de los índices de completitud temática, complementariedad informática y confiabilidad de datos.

Finalmente, en la quinta sección se presenta una serie de recomendaciones clave basadas en las mejores prácticas a nivel internacional que buscan asegurar una apropiada gestión integral de los datos.

SECCIÓN I. MÉTODOS DE LEVANTAMIENTO DE INFORMACIÓN



MÉTODOS DE LEVANTAMIENTO DE INFORMACIÓN

El análisis del entorno existente requiere de una gran cooperación con las personas involucradas en la operación y gestión de cada repositorio de información. Por ello, se seleccionaron metodologías que permitieran al equipo de análisis explotar el conocimiento de los expertos de cada institución.

Con el fin de cubrir el espectro de necesidades de información de este estudio se utilizaron 3 metodologías de levantamiento: la aplicación de un cuestionario en línea, visitas de levantamiento *in situ* y entrevistas a profundidad con actores clave.

LEVANTAMIENTO DE CUESTIONARIOS ENTRE ACTORES CLAVE

La construcción de un perfil sobre una estructura de información requiere conocimiento actualizado, preciso y veraz. Por este motivo, se desarrollaron cuestionarios de diagnóstico de los repositorios de datos. Los reactivos de estos cuestionarios se integraron alrededor de las áreas temáticas clave de información definidas en el alcance del proyecto, a saber:

1. Objetivo
2. Características del sistema y de la plataforma en la que está montada
3. Número de registros
4. Condiciones de la información
5. Tasa de crecimiento de los registros de la base de datos
6. Fechas o periodos de actualización de la información
7. Layout de las bases de datos
8. Uso de la información contenida

9. Origen de la información contenida
10. Nivel de Integración con la política de datos abiertos.
11. Clasificación de la información en términos de la LFTAIP
12. Antigüedad del sistema
13. Datos de los responsables de la información.



En algunas de las dimensiones se tomó como referencia la Herramienta para la Evaluación de la Calidad de los Registros Administrativos (HECRA) desarrollada por Federico Segui Stagno. Sin embargo, estas se ajustaron para los fines específicos del estudio.¹

¹ Sergui Stagno, Federico. (2012). Guía de la herramienta para la evaluación de la calidad de Registros Administrativos a ser usados con fines estadísticos. Asistencia Técnica del Banco Mundial prestada al Gobierno de Yucatán. Yucatán, México.

La batería de preguntas de los cuestionarios se validó con un pequeño grupo de expertos antes de iniciar el levantamiento. Estos dieron su retroalimentación y, en algunos casos, se ajustaron los reactivos a fin de atender sus inquietudes. La aplicación de los cuestionarios se llevó a cabo mediante una plataforma en línea en la que se generaron direcciones electrónicas donde las instituciones obtuvieron acceso y realizaron las actividades de carga de información en atención a los 130 reactivos de ambos cuestionarios.



LEVANTAMIENTO EN SITIO

Generalmente existe una brecha entre la percepción interna de las actividades de una unidad y aquella que resulta de una revisión presencial por un actor externo. Con eso en mente, se realizó un levantamiento en sitio para que los participantes y los investigadores pudieran enfatizar en aspectos de la operación diaria que pudieran escaparse a un cuestionario formal.²

ENTREVISTAS ABIERTAS-SEMIESTRUCTURADAS

Como complemento a los encuentros en sitio, las entrevistas abiertas-semiestructuradas permiten a los investigadores homogeneizar lenguaje, contexto y conocimiento para profundizar en áreas de interés de la investigación. Por tal razón se llevaron a cabo 14 entrevistas a profundidad con servidores públicos para reconocer la diversidad de las experiencias asociadas a cada uno de los casos evaluados.



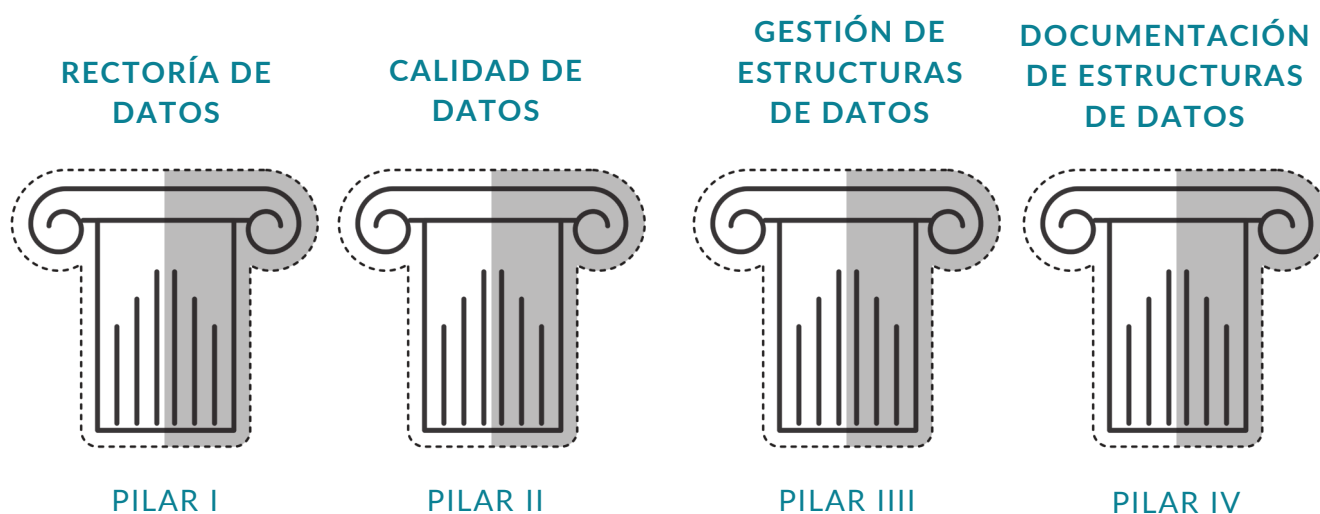
² El protocolo de entrevista utilizado se puede consultar en el anexo IV al final del presente documento.

PERSPECTIVA METODOLÓGICA DEL LEVANTAMIENTO DE INFORMACIÓN

La construcción de un inventario de los sistemas existentes dentro del ecosistema de información creado por la LGSNA requiere el respaldo de un marco conceptual robusto que permita priorizar y categorizar los campos de información que se rescatan sobre los sistemas existentes. Por tal motivo, el levantamiento de datos realizado como parte del estudio echó mano de una estructura conceptual que permitiera, además de la detección de sistemas, desarrollar análisis anclados en el conocimiento de gestión de datos.

PILARES DE UNA BUENA GESTIÓN DE DATOS

La estructura de información seleccionada para el levantamiento de información se basa en las dimensiones de buenas prácticas en materia de gestión de datos. Conocidas como 4 pilares de gestión de datos, estas dimensiones evalúan campos clave de información sobre el manejo y operación de estructuras de datos.



Cada uno de los pilares permite una revisión de algún aspecto clave de la gestión de datos. Asimismo, abre la puerta para un análisis más robusto que contraste lo observado con las mejores prácticas en el entorno internacional. Con el fin de establecer una base para este análisis, el levantamiento de información, adicionalmente a la detección de información básica de los sistemas, se vinculó con estas dimensiones.

SECCIÓN II. ALCANCES DEL ESTUDIO



ALCANCES DEL ESTUDIO

El estudio tiene como área de enfoque la presentación de un panorama claro de la política pública alrededor de 6 áreas temáticas clave de información definidas en los supuestos del artículo 49 de la Ley General del Sistema Nacional Anticorrupción.

1. Información de evolución patrimonial de servidores públicos
2. Información de servidores públicos que intervienen en procedimientos de contrataciones
3. Información de servidores públicos sancionados
4. Información de particulares sancionados
5. Información de denuncias públicas de faltas administrativas y hechos de corrupción
6. Información pública de contrataciones

El alcance del estudio excluye la información relativa al sistema de información y comunicación del Sistema Nacional Anticorrupción y del Sistema Nacional de Fiscalización indicados en el cuarto numeral del artículo 49 de la Ley toda vez que estos no están plenamente desarrollados y, por lo tanto, no están en condiciones para incluirse en el alcance del análisis.



SUJETOS CONSIDERADOS

Los sujetos del estudio fueron 18 instituciones del ámbito federal. Estas instituciones comprenden una amplia gama de perfiles incluyendo tribunales y organismos autónomos, así como instituciones del poder ejecutivo, legislativo y del poder judicial. Dado su papel y atribuciones dentro de la Administración Pública Federal (APF), con la revisión de los sistemas de la Secretaría de la Función Pública se dio cobertura a las dependencias de la APF.

Situación legal	Institución
Organismos públicos autónomos	Banco de México
	Comisión Federal de Competencia Económica
	Comisión Nacional de los Derechos Humanos
	Instituto Federal de Telecomunicaciones
	Instituto Nacional de Estadística y Geografía (INEGI)
	Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Públicos
	Instituto Nacional Electoral
	Instituto Nacional para la Evaluación de la Educación
Poder Ejecutivo (concentradora)	Secretaría de la Función Pública
Poder judicial	Consejo de la Judicatura Federal
	Suprema Corte de Justicia de la Nación
	Tribunal Electoral del Poder Judicial de la Federación
Poder legislativo	Auditoría Superior de la Federación
	Cámara de Diputados
	Senado de la República
Tribunal autónomo	Tribunal Federal de Justicia Administrativa
	Tribunal Superior Agrario

SUJETOS CONSIDERADOS FUERA DEL ESTUDIO

De las instituciones consideradas inicialmente no fue posible contar con la participación de 4 instituciones:

1. Banco de México³
2. Consejo de la Judicatura federal⁴
3. Tribunal Superior Agrario⁵
4. Tribunal Federal de Conciliación y Arbitraje⁶

Esta falta de participación se debió a causas operativas y de gestión internas a las instituciones y no por motivos atribuibles al estudio. No obstante, se considera que la muestra analizada contiene una riqueza y variedad analítica que permitió extraer lecciones de validez transversal al universo de sistemas en el ecosistema de información de las áreas temáticas asociadas a la ley.

³ El Banco de México hizo una entrega parcial de información. Por tal motivo, se tomó en cuenta la existencia de sus sistemas para fines de conteo en materia de cobertura temática según lo reportado. No obstante, no fue considerado en el resto del análisis toda vez que no se tuvo la información necesaria al momento de evaluar los datos para este informe.

⁴ El Consejo de la Judicatura Federal no hizo entrega de la información solicitada. Se tuvo comunicación con los enlaces designados sin alcanzar en ningún momento el objetivo del levantamiento de información. Por tanto, se dejó fuera de los procesos de levantamiento y análisis de información.

⁵ El Tribunal Superior Agrario no pudo ser contactado por el Secretariado Ejecutivo del SNA con motivo de las afectaciones a sus instalaciones por el sismo del 19 de septiembre de 2017. Con en este hecho, se determinó de manera conjunta entre la SESNA y el equipo de investigación del CIDE, que no se avanzaría con el levantamiento de información en esta institución.

⁶ El Tribunal Federal de Conciliación y Arbitraje reportó que no contaba con sistemas propios para administrar su información por su dependencia operativa y de gestión del poder ejecutivo en estas áreas temáticas. Toda vez que la información del ejecutivo se levanta, procesa y explota por la SFP se consideró apropiado eliminarlos como sujetos del estudio y evitar duplicidades de información en el levantamiento o una exposición de información sesgada con motivo del diseño institucional y las atribuciones de la institución.

LÍNEA BASE SUJETOS DE ESTUDIO

En el estudio sólo se consideraron repositorios de información cuyos datos han sido tratados sistemáticamente y que por consiguiente cuentan con condiciones básicas para operarse, analizarse y explotarse como bases de datos formales.

A fin de alcanzar este resultado, el alcance del estudio se limitó con base en 4 criterios para la inclusión de información de los entes federales:

1. La información está arreglada conforme a una lógica consistente y definida en una matriz de datos.
2. La información de cada registro está organizada y distribuida en múltiples categorías o dimensiones.
3. La información de cada registro puede explotarse y extraerse de manera independiente entre sí sin perder valor.
4. El archivo electrónico que contiene la información fue creado en formato electrónico y no es una digitalización (escaneo o fotografía) de documentación original o productos de información que se turnaron por terceras partes.

LÍMITES DEL ALCANCE DEL ESTUDIO

Como parte del desarrollo de esta investigación no se realizó ningún tipo de obtención ni análisis de las bases de datos físicas de los entes sujetos del estudio. Por lo tanto, el levantamiento de información se realizó estrictamente con base en la información provista por los funcionarios responsables de ello en el marco del proyecto. En ningún momento se solicitó a las instituciones participantes en el estudio la entrega de registros provenientes de sus bases de datos.

Asimismo, es importante resaltar que el CIDE no realizó verificaciones o evaluaciones de la veracidad de la información provista. Por lo tanto, la información se recibió de buena fe y se tomó como válida para todos los casos.

SECCIÓN III. PANORAMA DE SISTEMAS RELEVANTES



PANORAMA DE SISTEMAS RELEVANTES

La evaluación exploratoria del ecosistema de información permitió la construcción de un diagnóstico descriptivo en el que, por vez primera, se pudo hacer un inventario comprensivo de los sistemas electrónicos y bases de datos en materia de lo establecido por la LGSNA.

El ecosistema de información cuenta con un total de 70 sistemas electrónicos y bases de datos que atienden los requisitos temáticos y operativos fijados al inicio del levantamiento. Enseguida, se muestra el listado completo de los sistemas a los que se hace referencia.

LISTADO DE SISTEMAS DETECTADOS DURANTE LA FASE INICIAL DE LEVANTAMIENTO.



1

Nombre del sistema

Registro Único de Proveedores



BANCO DE MÉXICO

1

Sistema de Declaración Patrimonial Electrónica (DEPTEL)

2

Portal de Contrataciones

3

Registro de denuncias

4

Registro de Procedimientos Administrativos Insaturados y Sanciones

5

Registro de Servidores Públicos facultados para suscribir actos



Nombre del sistema

- 1 DeclaraUEC⁷
- 2 Portal de adquisiciones del Congreso
- 3 Sistema de Registro y Control Patrimonial



Nombre del sistema

- 1 DeclaraNet COFECE
- 2 Registro de Inconformidades
- 3 Registro de Procedimientos Administrativos de Responsabilidades
- 4 Registro de quejas y/o denuncias
- 5 Registro de Sanción a Proveedores



Nombre del sistema

- 1 Delsa System
- 2 Meta 4
- 3 Registro y control de las relaciones patrimoniales
- 4 Sancionados con inhabilitación firme
- 5 Sistema de Control de Almacén y Activo Fijo
- 6 Sistema de gestión de adquisiciones
- 7 Sistema integral de administración de recursos financieros

⁷ El sistema DeclaraUEC, operado por la Cámara de Diputados se le imputó a la ASF en el análisis posterior, toda vez que toda la información contenida en la plataforma es de personal de la ASF.



INSTITUTO FEDERAL DE
TELECOMUNICACIONES

Nombre del sistema

- 1 DeclaraNet Plus (IFT)
- 2 Registro de Servidores Públicos Sancionados (IFT)



Nombre del sistema

- 1 DeclarINEGI
- 2 Registro de Servidores Públicos Sancionados (RESPS)
- 3 SIDEQ (Módulo de quejas y denuncias)
- 4 Véndale al INEGI



Nombre del sistema

- 1 DeclaraNet
- 2 Plataforma Nacional de Transparencia



Nombre del sistema

- 1 Base de datos de proveedores sancionados
- 2 CompraINE
- 3 DeclaraINE
- 4 Sistema de Información de Quejas, Denuncias y Responsabilidades
- 5 Sistema Integral de Gestión Administrativa (SIGA) – Módulo de compras (PO) y Módulo de Tesorería (XTR)
- 6 Sistema Integral de Gestión Administrativa (SIGA) – Módulo de Recursos Humanos



Nombre del sistema

- 1 Buzón electrónico de quejas y denuncias
- 2 DeclaralNEE
- 3 Registro de licitantes y proveedores sancionados
- 4 Sistema de Nómina DELSA
- 5 Sistema Integral Financiero



Nombre del sistema

- 1 Padrón de Proveedores y Contratistas
- 2 Procedimientos Contratación- Padrón Proveedores y Contratistas (SIPOT)
- 3 Procedimientos de Contratación
- 4 Registro de impedimentos para contratar
- 5 Registro de quejas o denuncias
- 6 Registro de Servidores Públicos Sancionados (SCJN)
- 7 Sistema de Declaración Patrimonial



Nombre del sistema

- 1 Sistema de Registro de Proveedores Sancionados
- 2 Sistema de registro de servidoras/es públicas/os con sanciones

Nombre del sistema	
1	Archivo de datos relevantes de contratos reportados en CompraNet
2	CompraNet – im Módulo de Información e Inteligencia de Mercado para las Contrataciones Públicas
3	CompraNet (como plataforma transaccional)
4	CompraNet –hc (Módulo para la integración y consulta del historial en materia de contrataciones y su cumplimiento)
5	DeclaraNet Plus (SFP)
6	Directorio de proveedores y Contratistas Sancionados
7	Directorio de Unidades Compradoras
8	OMEXT (Omisiones y extemporáneos en la presentación de la declaración de situación patrimonial)
9	Órganos Internos de Control (OIC Supervisores CNET)
10	Programas Anuales de Adquisiciones, Arrendamientos, Servicios y Obras Públicas (PAAASOP)
11	Registro de Servidores Públicos Sancionados
12	Registro Único de Proveedores y Contratistas (RUPC)
13	Responsables de Unidades Compradoras
14	Sistema de desarrollo organizacional de órganos de vigilancia y control
15	Sistema de Procedimiento Administrativo de Sanción a Proveedores y Contratistas
16	Sistema de Procedimientos Administrativos de Responsabilidades
17	Sistema integral de inconformidades
18	Sistema Integral de Quejas y Denuncias Ciudadanas (SIDECA)
19	Sistema para la Evaluación de Resultados de los Órganos de Vigilancia y Control



Nombre del sistema

- 1 Control de Expedientes Procesales (CONEXP)



Nombre del sistema

- 1 Servidores Públicos Sancionados

El total de sistemas detectados sirvió como insumo para el cálculo de la cobertura temática en el ecosistema completo de información. No obstante, de los repositorios de información detectados y considerados al cierre del análisis solamente 52 se utilizaron para el análisis completo desarrollado acerca de las características de los sistemas electrónicos y bases de datos reportados. El motivo de lo anterior se expone a continuación.

CRITERIOS PARA LA DEPURACIÓN DEL LISTADO DE SISTEMAS ELECTRÓNICOS Y BASES DE DATOS DETECTADOS

Todos los sistemas detectados fueron sujetos al proceso de análisis. Sin embargo, luego de la revisión de relevancia temática se eliminaron del análisis 8 sistemas electrónicos y bases de datos que, conforme a la información provista por las instituciones que los sometieron, no contienen información dentro de las áreas temáticas establecidas por la LGSNA.

	Nombre del sistema	Institución responsable
1	Delsa System	Comisión Nacional de Derechos Humanos
2	Meta 4	Comisión Nacional de Derechos Humanos
3	Sistema Integral de Gestión Administrativa (SIGA) – Módulo de Recursos Humanos	Instituto Nacional Electoral
4	Sistema de Nómina DELSA	Instituto Nacional para la Evaluación Educativa
5	Órganos Internos de Control (OIC Supervisores CNET)	Secretaría de la Función Pública
6	Sistema para la Evaluación de Resultados de los Órganos de Vigilancia y Control	Secretaría de la Función Pública
7	Sistema de Desarrollo Organizacional de Órganos de Vigilancia y Control	Secretaría de la Función Pública
8	Control de Expedientes Procesales (CONEXP)	Tribunal Federal de Conciliación y Arbitraje

Adicionalmente, durante el proceso de levantamiento en sitio se observó con apoyo del personal especializado, que cuatro de las bases de datos reportadas por la Secretaría de la Función Pública son reportes directos de la base de datos del sistema CompraNet. Toda vez que el análisis y consideración de esta información duplicaría la información de CompraNet y que sobre-representaría los sistemas de aquella institución, se decidió eliminarlos del análisis.

	Nombre del sistema	Institución responsable
1	Archivo de datos relevantes de contratos reportados en CompraNet	Secretaría de la Función Pública
2	Registro Único de Proveedores y Contratistas (RUPC)	Secretaría de la Función Pública
3	Directorio de Unidades Compradoras	Secretaría de la Función Pública
4	Responsables de Unidades Compradoras	Secretaría de la Función Pública

Adicionalmente, los sistemas reportados por el Banco de México no se pudieron integrar por completo al análisis toda vez que dicha institución no capturó a tiempo la información detallada de su funcionamiento para su procesamiento por parte del CIDE. Sin embargo, se dejaron para fines de conteo de cobertura temática e institucional.

	Nombre del sistema	Institución responsable
1	Registro Único de Proveedores	Auditoría Superior de la Federación
2	Sistema de Declaración Patrimonial Electrónica (DEPTEL)	Banco de México
3	Portal de Contrataciones	Banco de México
4	Registro de denuncias	Banco de México
5	Registro de Procedimientos Administrativos Insaturados y Sanciones	Banco de México
6	Registro de Servidores Públicos facultados para suscribir actos	Banco de México

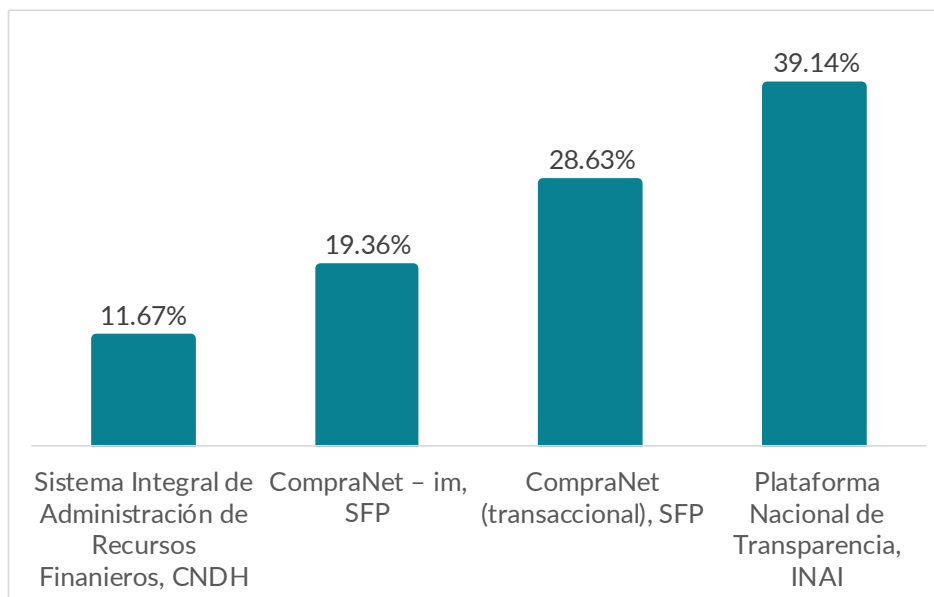
Esto resulta en un total de 52 sistemas considerados para fines de análisis dentro del presente estudio que contienen información de temáticas relevantes para el diagnóstico del ecosistema de información.

DISTRIBUCIÓN DE INFORMACIÓN POR VOLUMEN DE DATOS

CUATRO SISTEMAS CONCENTRAN EL 99.3% DEL VOLUMEN DE DATOS TOTALES RECOLECTADOS.

¿CÓMO SE DISTRIBUYE LA CARGA DE INFORMACIÓN ENTRE LOS SISTEMAS EVALUADOS?

La información de los sistemas evaluados muestra una alta concentración del volumen de datos en 4 sistemas principales. Estos sistemas ocupan, en conjunto, un 99.39 por ciento de los datos totales reportados. El porcentaje relativo al total de los registros reportados que ocupa cada uno de estos sistemas se muestra en la gráfica a continuación.^{8,9}

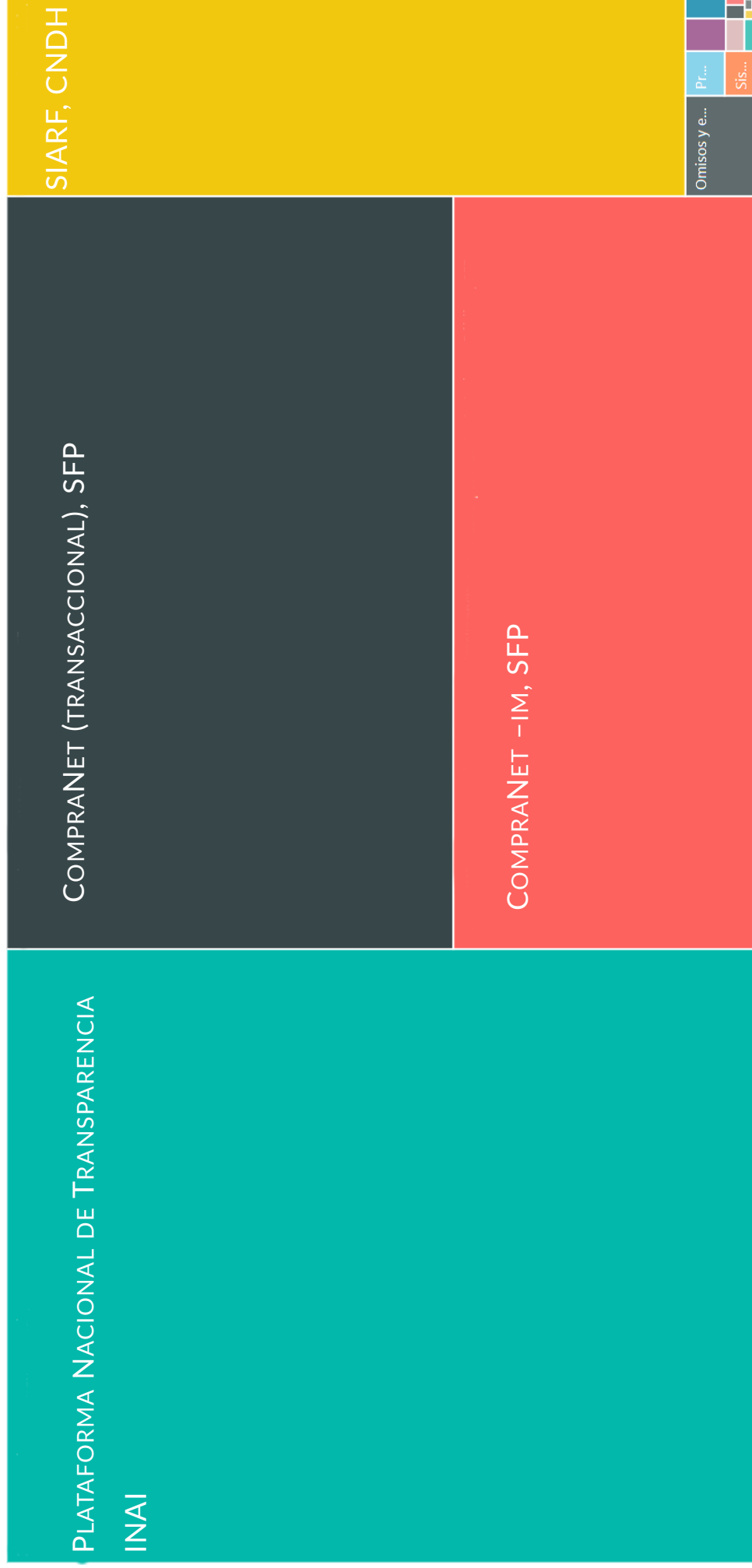


Como comparativo, el quinto lugar en mayor cantidad de datos contiene el 0.6 por ciento del total de los registros reportados. Más aún, el resto de los sistemas aporta el 1.21 por ciento del total de los datos reportados.¹⁰ Este desbalance muestra las diferencias en recursos, requerimientos y amplitud de la cobertura requerida para cada sistema electrónico o base de datos y se puede observar en el diagrama de árbol a continuación. No obstante, resalta que estas diferencias no se empatan perfectamente con las diferencias en buenas prácticas implementadas.

⁸ La CNDH reportó que este sistema contiene información de servidores públicos sancionados. No obstante, el volumen de datos específicos a las temáticas dentro del alcance del estudio no se pudo determinar puntualmente. Por lo tanto, el volumen de datos reportado podría encontrarse sesgado al contener registros de datos de otras temáticas no relevantes de toda la institución, en particular todos los registros administrativos y financieros de la institución, que podrían escapar del alcance del estudio.

⁹ La Plataforma Nacional de Transparencia se incluye en el análisis, ya que contiene información materia de este estudio, sin embargo, se debe considerar que el volumen de información que representa puede quedar sobreestimado, ya que en muchos casos, la Plataforma no es la fuente original de información sistematizada, sino un sistema al que las instituciones le reportan datos de muchas temáticas acerca de su operación y no sólo de los temas objeto de este estudio. Adicionalmente, se debe tomar en cuenta que en algunos casos la información contenida en la Plataforma puede estar en sistemas que pertenecen a otras instituciones, ya consideradas en el estudio como fuente original.

DIAGRAMA DE ÁRBOL. SISTEMAS DETECTADOS POR VOLUMEN DE INFORMACIÓN.



¹⁰ En total, 10 sistemas no reportaron datos del número de registros en sus bases de datos o reportaron contar con cero registros. Estos son los siguientes: Servidores Públicos Sancionados del TFJA; Sistema de gestión de Adquisiciones de la CNDH; Sistema Integral de Inconformidades de la SFP; Procedimientos de Contratación (Padrón de proveedores y Contratistas SIPOT) de la SCJN; Portal de Adquisiciones del Congreso de la Cámara de Diputados; DeclaraNet Plus de la SFP; Registro de Sanción a proveedores de la COFECE; Véndale al INEGI del INEGI; DeclaraNet Plus del IFT y el Buzón electrónico de quejas y denuncias del INEE

ANÁLISIS DEL ECOSISTEMA DE INFORMACIÓN

Los sistemas que formaran parte del ecosistema de información de la LGSNA muestran algunas características comunes. Estas características ya dibujan un panorama preliminar de la situación que enfrentaría cualquier esfuerzo en materia de transferencia, interconexión o interoperabilidad que quisiera aprovechar estos repositorios de datos.

Con el fin de presentar este panorama y sus tendencias, esta sección busca mostrar los principales hallazgos para cada uno de los pilares de gestión de datos. Inicialmente, se presentan los resultados generales del ecosistema de información y, posteriormente, se desagrega este diagnóstico por área temática.

PILAR I. RECTORÍA DE DATOS

La rectoría o gestión de datos es el conjunto de actividades conectadas a la planeación, desarrollo, implementación y administración de sistemas o repositorios de información para la adquisición, almacenamiento, resguardo, recuperación, explotación, diseminación, almacenamiento y eliminación de datos.¹¹ Este pilar hace referencia a todo lo concerniente a la asignación clara de responsabilidades en cada una de las dimensiones de la gestión y resguardo de información dentro de una organización.

TECNOLOGÍAS DE MANEJO DE DATOS

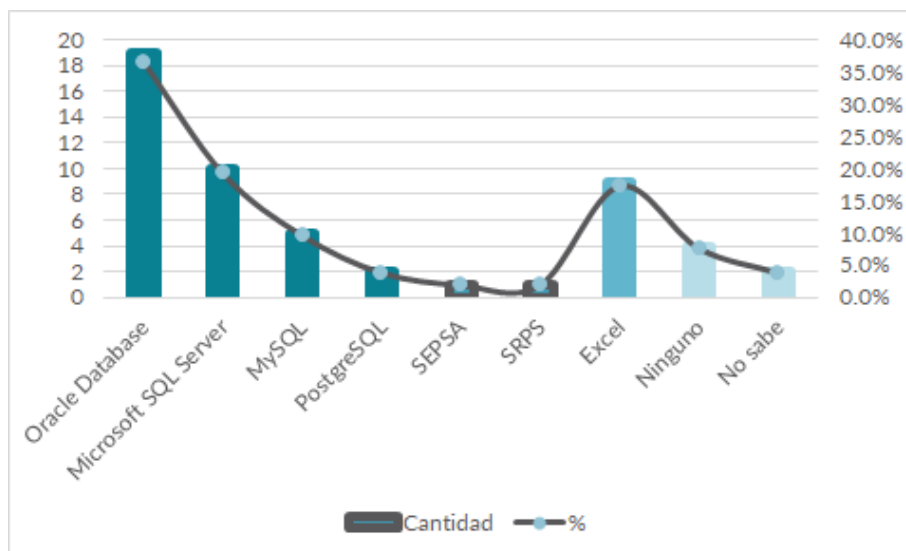
EXISTE UNA AMPLIA PENETRACIÓN DE TECNOLOGÍAS DE INFORMACIÓN PARA EL MANEJO DE DATOS EN PODER DE LAS INSTITUCIONES UNA VEZ QUE ADQUIERES LA INFORMACIÓN.

¿Cómo se explota la información capturada?

La mayoría de los sistemas detectados utilizan manejadores de bases de datos para sus repositorios de información. Esto indica un avance en el progreso del tratamiento de datos en las áreas temáticas que han sido atendidas. En este respecto se puede señalar que el 71.2 por ciento (37) de los repositorios de información reportados como parte de sistemas electrónicos o bases de datos utilizan tecnología de manejadores de bases de datos que requiere para su operación algún nivel de tratamiento formal de datos.¹²

¹¹ La gestión o rectoría de datos puede referirse a fuentes análogas o digitales de datos. No obstante, para fines del presente reporte se hace referencia estrictamente a los repositorios digitales de información y las políticas asociadas toda vez que existen algunas diferencias en las agendas de trabajo y metodologías y se consideró que los repositorios digitales cuentan con ventajas y mayor potencial de explotación de información.

¹² En este cálculo no se consideran los sistemas que no tienen relevancia temática para el estudio (8) y aquellas bases de datos reportadas que son meramente reportes de otros sistemas (5). Lo anterior con motivo de que el monto estaría afectado artificialmente y alteraría la perspectiva de la penetración de tratamiento de datos en el entorno.



Como se puede observar en la gráfica, el manejador de bases de datos utilizado más comúnmente es Oracle Database que se utiliza en 36.5 por ciento de los repositorios de información, seguido por Microsoft SQL Server (19.2%) y MySQL (9.6%). Por otro lado, 15.3 por ciento (9 sistemas o bases de datos) de las bases de datos se manejan en hojas de cálculo.¹³

DESEMPEÑO POR ÁREA TEMÁTICA

Evolución patrimonial

En materia de evolución patrimonial, la mayor parte de los sistemas utilizan un manejador de bases de datos. De los 13 sistemas en el área temática, el 92.3 por ciento de los repositorios está sujeto a un manejador de bases de datos. Solamente un sistema no indicó contar con una herramienta de este tipo. Sin embargo, toda vez que es una versión del sistema DeclaraNet se puede asumir con seguridad que este es el caso.

Motivos para reflexionar: Una de las instituciones participantes lleva los registros de sus quejas y denuncias en archivos de procesador de textos que se capturan en un archivo y se manda a imprimir para integrarse al expediente físico del caso. Una vez que se imprime el archivo, se cierra destruye el archivo electrónico. La información sólo existe como parte del expediente físico. Como complemento, se lleva un listado de los casos en activo que mantiene el encargado del área y que se va limpiando una vez que se consideran listos para almacenamiento.

¿Qué pasaría si se afectan los archivos físicos?

¿Cómo se justifica no a provechar las capacidades de computación del mundo moderno ?

Servidores públicos sancionados

En materia de servidores públicos sancionados, la mayoría de los sistemas utilizan un manejador de base de datos. No obstante, es una de las áreas temáticas con menos extensión en el uso de estas herramientas. En esta área, el 66.7 por ciento de los sistemas cuenta con un manejador de bases de datos.

¹³ El sistema Véndale al INEGI reportó utilizar dos manejadores distintos de bases de datos por lo que el total agregado reportado en la gráfica supera el límite superior de sistemas y bases de datos indicado en el análisis.

Contrataciones

En materia de contrataciones, la mayoría de los sistemas utilizan un manejador de base de datos. De los 23 sistemas en el área temática, el 65.2 por ciento se encuentra sujeto a un a manejador de base de datos.

Sanciones a proveedores

En materia de sanción a proveedores, la mayoría de los sistemas utilizan a un manejador de base de datos. De los 23 sistemas en el área, el 63.6 por ciento está sujeto a un manejador de base de datos.

Quejas

En materia de quejas y denuncias, la mayoría de los sistemas no utilizan un manejador de base de datos. De los 13 sistemas en el área temática, sólo el 46.2 por ciento de los sistemas gestionan su información con el apoyo de manejadores de bases de datos. Esta es el área temática con la penetración más baja de estas herramientas.

VIGENCIA DE REPOSITORIOS DE INFORMACIÓN

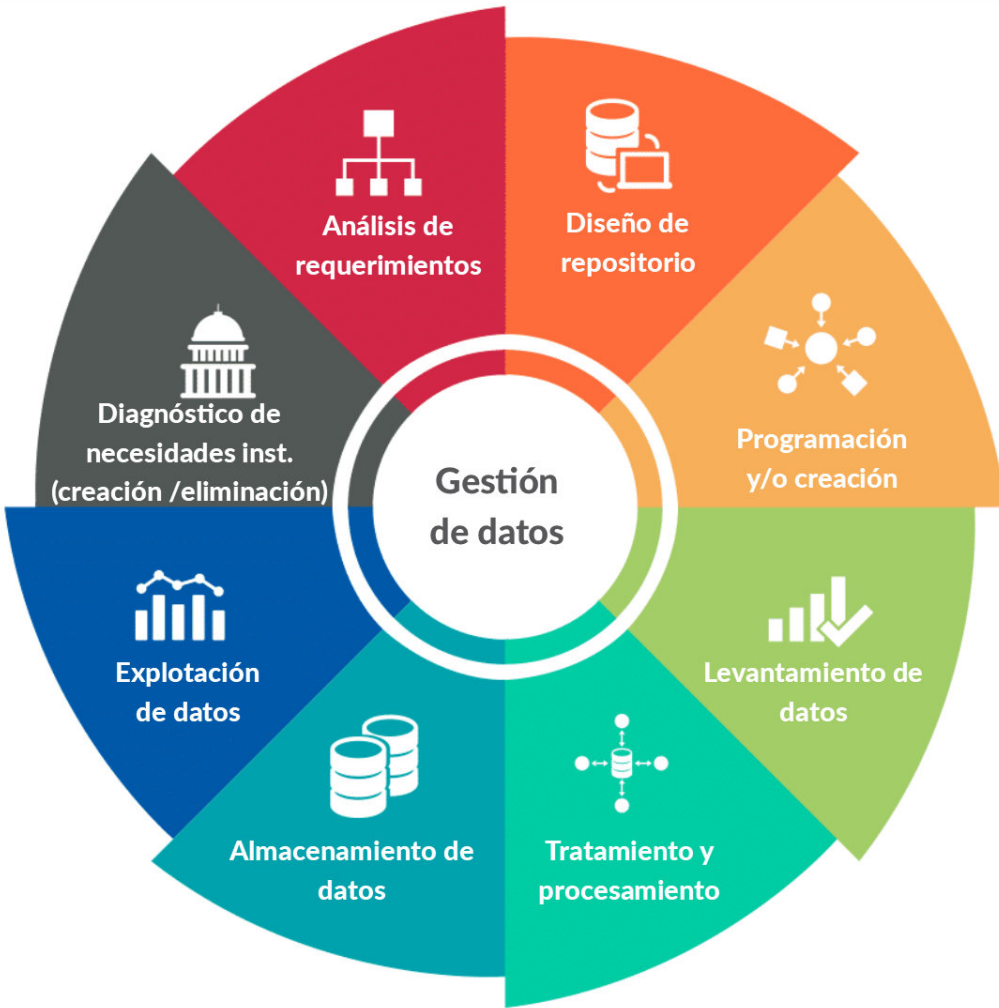
SE PREVÉ QUE CASI TODOS LOS SISTEMAS Y BASES DE DATOS MANTENDRÁN SU VIGENCIA Y VALIDEZ.

Pregunta: ¿Se continuará con la actualización de los sistemas detectados?

Los repositorios de datos siguen un ciclo de vida conforme cumplen los objetivos institucionales que motivaron su creación. Este ciclo de vida pasa por algunas fases clave en el ciclo de vida de la gestión de datos.

Este ciclo de vida esta marcado por un ciclo de gestión de datos que es dirigido desde el diagnóstico de necesidades institucionales y que comprende el ejercicio de las facultades ejecutivas dentro de la institución para definir si las necesidades de datos requieren la creación de una nueva estructura de datos, el ajuste de una existente o la eliminación de una preexistente.

Alrededor de esta función ejecutiva, el ciclo de vida de datos progresa desde la fase del análisis específico de requerimientos en materia de datos hasta la explotación de estos. Entre estos dos puntos se llevan a cabo múltiples actividades de gestión en materia del diseño de los repositorios, creación del repositorio conforme a tecnologías seleccionadas, levantamiento de datos, tratamiento y procesamiento de datos, almacenamiento de datos y, finalmente, la explotación de los mismos para los fines conforme a los que fue creado. Todo lo anterior hace necesario determinar si los repositorios de datos reportados se encuentran aún en la fase productiva de su ciclo de vida.



Con el fin de realizar este diagnóstico, se buscó confirmación por parte de las instituciones que administran los sistemas de que estos continuarían siendo actualizados. Del total de sistemas detectados, sólo una institución indicó que el sistema reportado dejaría de ser actualizado en el futuro. En su oportunidad, la CNDH indicó que su "Sistema de gestión de adquisiciones" no seguiría siendo actualizado. Esto muestra que en el futuro inmediato los sistemas detectados se mantendrán relativamente estables.

Es importante señalar que durante las entrevistas se observó que algunas de las instituciones parti-

cipantes prevén el desarrollo de sistemas que podrían reemplazar a los sistemas existentes en el mediano plazo. Estos casos no son representativos de una tendencia general, sin embargo, cualquier esfuerzo de integración, interoperación o interconexión entre sistemas requerirá considerar estos posibles cambios dentro del inventario de repositorios de datos.

PROCESOS DE ADQUISICIÓN DE INFORMACIÓN

LA INFORMACIÓN CAPTURADA EN LOS REPOSITORIOS DE INFORMACIÓN ES PREDOMINANTEMENTE UN REGISTRO EXTEMPORÁNEO DE EVENTOS.

¿En qué momento se registran los eventos de datos reflejados en las bases de datos?

Toda vez que los sistemas son repositorios de información, es importante contar con un diagnóstico sobre las fuentes de las que se alimentan. Ello permite entender cómo se registra la información y de la coordinación que se requiere para su integración. Estas dimensiones definen la frontera entre sistemas de reporte y transaccionales. Los sistemas de reporte están diseñados para el manejo y gestión de información que se obtiene de diversas fuentes y se vacía en la base de datos.

En contraste, los sistemas transaccionales están diseñados para operar algún proceso sustantivo de las actividades de la organización y, por lo tanto, obtienen su información de manera constante conforme se desarrollan las actividades cotidianas de la organización.

De entre los sistemas detectados, el 59.6 por ciento (31) son registros extemporáneos de la información registrada. Por otro lado, el 40.4 por ciento (21) restante de los sistemas recuperan la información de manera contemporánea como parte de procesos de trabajo. Pero, ¿se implementan las mejores prácticas aplicables según su perfil?

Las mejores prácticas no se implementan frecuentemente dentro del ecosistema de información de la LGS-NA. Del total de los sistemas y bases de datos evaluados que operan bajo el perfil de reporte, sólo 9.6 por ciento (3) reportan que realizan validaciones de veracidad mediante auditorías de la información, esto es 5.7 por ciento del total del universo de sistemas detectados. Por otro lado, el 61.2 por ciento (19) de los sistemas de reporte están sujetos a otro tipo de validaciones de veracidad de los datos (36.5 por ciento del universo).

¿Por qué suelen ser más confiables los sistemas transaccionales?

Como resultado de la cercanía a los procesos de trabajo, los sistemas transaccionales cuentan, generalmente, con procesos integrados de validación de información que son requeridos por el imperativo operativo de que la información que contiene sea válida, veraz y precisa. En contraste, los sistemas de reporte no siempre cuentan con evaluaciones de este tipo, que requieren la dedicación de tiempo adicional del personal que administra los sistemas. Con motivo de lo anterior, en materia de rectoría de datos, los sistemas de reporte suelen ser menos confiables que los sistemas transaccionales.

En este respecto es importante señalar que se observó que en estos casos las validaciones de veracidad generalmente se limitaban a modalidades de validación que permitan asegurar la integridad de la información (para más detalle véase la subsección de control de veracidad del Pilar II). Aunado a esta situación, se debe tener en mente que el 25.8 por ciento (8) de los sistemas de reporte no cuentan con ningún tipo de validación. Lo anterior, implica que el 87 por ciento de los sistemas de reporte no cuentan con medidas avanzadas de rectoría de información.

Por su parte, los sistemas transaccionales no se desempeñan mejor en términos de validación de veracidad. Sólo 5 de los 21 sistemas transaccionales (23.8 por ciento) implementan medidas formales de auditoría de datos.

Lo descrito, subraya la necesidad de una política de rectoría de datos que no solo implemente medidas de administración de tecnologías de información. Por el contrario, es necesario que como parte del desarrollo y maduración del ecosistema de información se prevean mecanismos de gestión de datos que se acomoden a las necesidades de cada perfil de sistema.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

En materia de evolución patrimonial existe una distribución relativamente homogénea de sistemas operando bajo ambos perfiles de operación. Si bien se reportó para la mayoría de los sistemas (8) de evolución patrimonial que se operan los repositorios de información bajo el modelo transaccional de captura de información (61.5 por ciento) la diferencia respecto de los sistemas de reporte es de 3 sistemas.

Servidores públicos sancionados

El área temática de servidores públicos sancionados está marcada por una presencia predominante de sistemas de reporte. De los 12 sistemas detectados en esta área temática, los equipos que administran estos repositorios de información indicaron que 9 de ellos operan bajo el sistema de reporte.

Contrataciones

En materia de contrataciones, los sistemas detectados predominantemente funcionan bajo el modelo de reporte. Se reportó que de los 23 sistemas detectados en esta área, 15 capturan su información de manera extemporánea. Esto es interesante toda vez que las contrataciones tienen más oportunidad de implementar flujos de trabajo automatizados y de tal forma generar información transaccional confiable.

Sanciones a proveedores

En materia de sanción a proveedores, hay una distribución relativamente equitativa entre los sistemas que operan bajo la modalidad transaccional y de reporte. Conforme a lo señalado por las instituciones participantes, de los 11 sistemas detectados en el área temática, 6 de 11 son transaccionales y 5 de reporte.

Quejas

En materia de quejas y denuncias, los sistemas se manejan casi por completo bajo el esquema de reporte. Entre los 13 sistemas en el área temática, las instituciones reportaron que 11 de ellos se manejan mediante reporte de información. Esto se conecta con las dinámicas de operación en esta área temática que descansan en reportes de los procesos de seguimiento a las quejas y denuncias sobre servidores públicos.

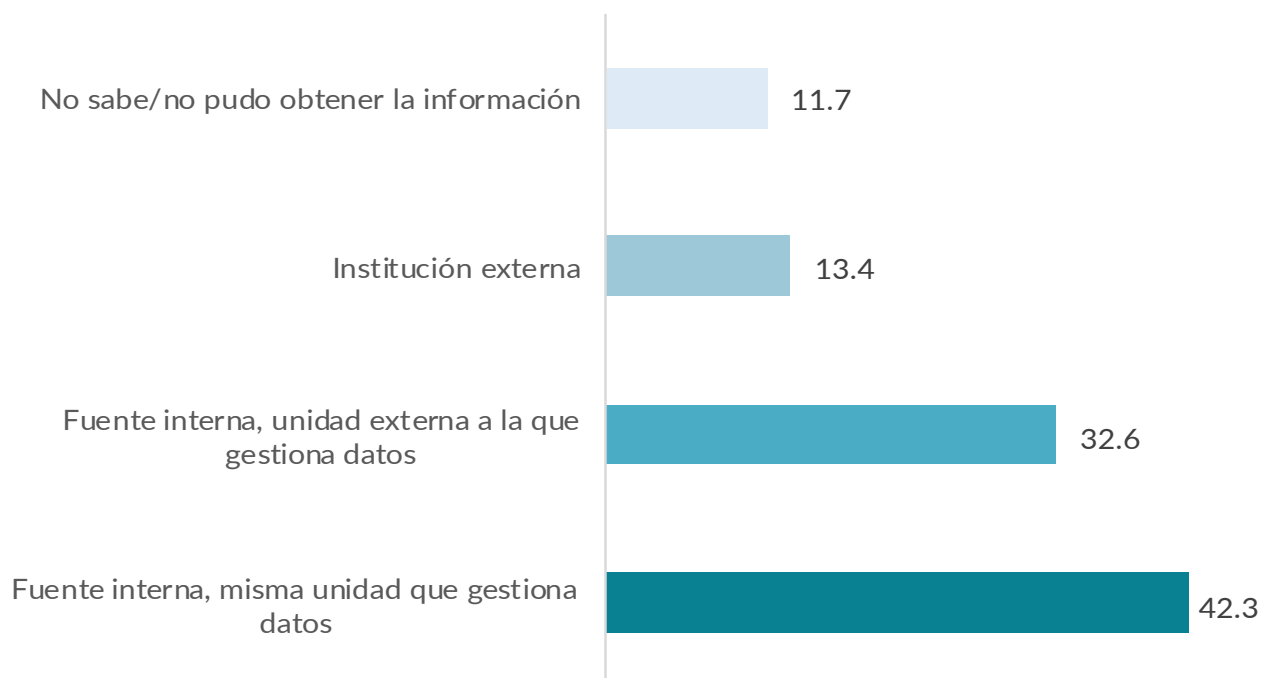
RELACIÓN CON FUENTES DE DATOS

EL GRUESO DE LA INFORMACIÓN PROCESADA SE GESTIONA Y CAPTURA DENTRO DE LAS MISMAS INSTITUCIONES

Pregunta: ¿Quién provee la información a los gestores de las bases de datos?

Los sistemas y bases de datos detectados dentro del ecosistema se dividen de manera homogénea entre aquellos que consolidan el levantamiento y la gestión de datos y aquellos que separan estas dos funciones. Por un parte, 42.3 por ciento (22) de los repositorios detectados tienen como fuente de información a la misma unidad que gestiona los procesos de levantamiento de información. Por otro lado, el 46 por ciento (24) de estos tienen como fuente de información a una tercera parte ya sea dentro o fuera de la misma institución.

De entre los sistemas y bases de datos que se nutren de terceras partes, el 32.6 (17) por ciento de los repositorios de información tienen como fuente de información una unidad dentro de la misma institución separada operativamente de aquella que gestiona, trata y analiza los datos. Mientras tanto, el 13.4 por ciento (7) de ellos obtienen su información de una institución externa.



RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

En materia de evolución patrimonial, se reportó que el grueso de los sistemas tiene como origen de su información a una unidad interna a la institución, pero externa al equipo que gestiona el sistema. De los 13 sistemas en el área temática 9 (69.2 por ciento) obtienen su información de esta manera.

Servidores públicos sancionados

En materia de servidores públicos sancionados las instituciones indican que el grueso de los sistemas tiene como punto de origen de su información a la misma unidad que gestiona los sistemas. Según la información provista, de los 12 sistemas, 8 son su propia fuente de datos, 3 tienen a otra unidad en la misma institución que provee esta información y sólo 1 (registro de servidores públicos sancionados del IFT) obtiene su información de una unidad externa a la institución.

Contrataciones

En materia de contrataciones, el origen de la información se distribuye equitativamente entre los tipos de fuente de datos. Por un lado, se reportó que los sistemas cuya información proviene del exterior de la institución representa el 26 por ciento (6) del total al igual que aquellos cuyo origen es interno, pero en una unidad externa a la que gestiona los datos. Por otro lado, los administradores de los sistemas indicaron que los sistemas que tienen una operación de extracción y gestión de datos consolidada en una misma unidad representan el 30.4 por ciento (7) del entorno. Sin embargo, la diferencia es irrelevante. Finalmente, 4 sistemas no pudieron obtener la información o no saben.

Sanciones a proveedores

En materia de sanción a proveedores, el origen de la información es predominantemente interno (72.7 por ciento). Sin embargo, dentro de esta categoría los sistemas se distribuyen equitativamente entre aquellos cuyas fuentes están en la misma unidad que gestiona el sistema y aquellos que consolidan la gestión y son la fuente de información (36.3 por ciento cada uno). En contraste, sólo un sistema tiene fuente externa de datos para su operación.

Quejas

En materia de quejas y denuncias, lo más común es que la información provenga de la misma unidad que administra el sistema. En esta área temática se reportó que 9 de los 13 sistemas tienen este arreglo, mientras que sólo un sistema obtiene sus datos de otra unidad dentro de la misma institución. En esta área, sólo se detectó que un sistema obtiene información de fuentes externas (Registro de quejas o denuncias de la SCJN).

ACCESO A LA INFORMACIÓN

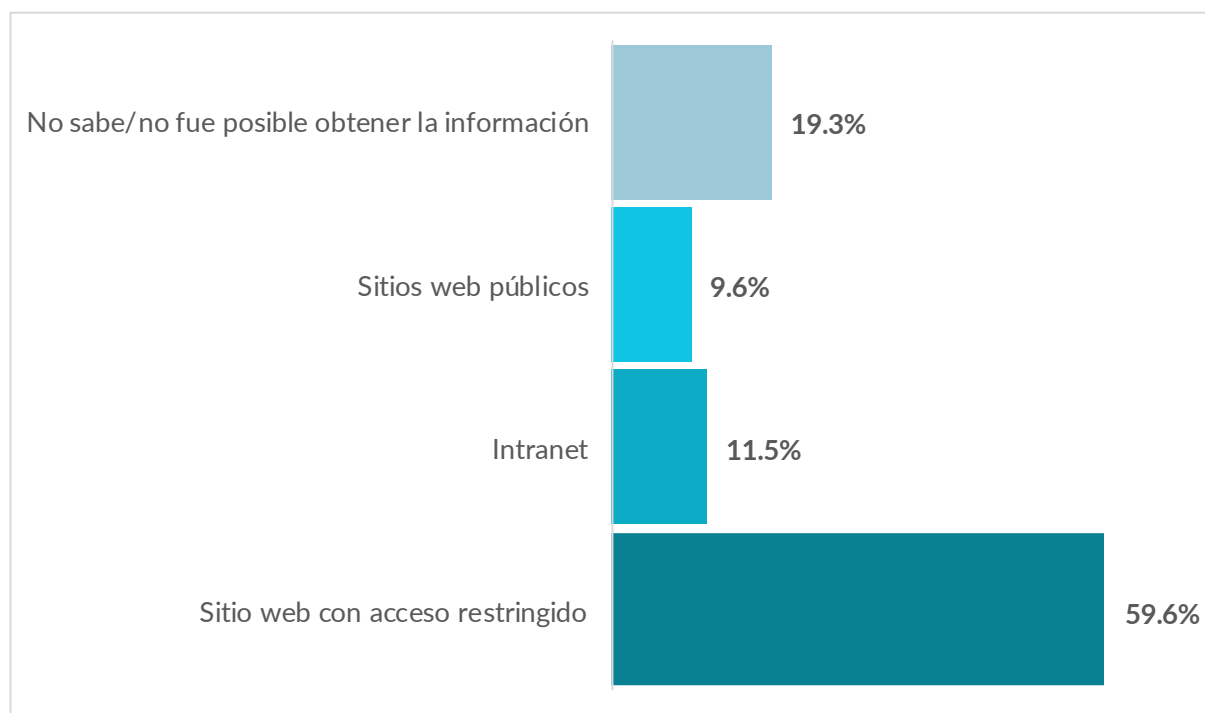
LOS SITIOS WEB CON ACCESO RESTRINGIDO REPRESENTAN LA MODALIDAD MÁS COMÚN PARA CONTROLAR EL ACCESO Y GESTIÓN DE LA INFORMACIÓN

¿Qué tipo de plataforma se utiliza más comúnmente?

El grueso de los repositorios de información detectados utiliza tecnología basada en sitios web con acceso restringido para almacenar, operar y gestionar la información levantada. Estos representan el 59.6 por ciento de los repositorios evaluados. En contraste, el segundo modelo más utilizado es la intranet que fue implementado en 11.5 por ciento (6 casos) de aquellos sistemas electrónicos o bases de datos considerados.

En tercer lugar, se encuentran los sitios web públicos con una penetración de 9.6 por ciento (5 casos) de sistemas o bases de datos.¹⁴

MODALIDAD DE ACCESO A REPOSITORIOS DE DATOS.



RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

En materia de evolución patrimonial, se reportó que los sistemas existentes funcionan predominantemente bajo la forma de sitios web con acceso restringido. Los datos indican que de los 13 sistemas del área temá-

¹⁴ Se reportó por parte de las instituciones que 10 sistemas o bases de datos están en formato de sitio web público. No obstante, 4 casos son hojas de cálculo y uno de los casos no tiene información de áreas temáticas relevantes para el estudio.

tica, 84.6 por ciento (11) funcionan en este formato. Sólo dos repositorios de información manejan su acceso en redes de intranet institucional.

Servidores públicos sancionados

En materia de servidores públicos sancionados, se reportó que los repositorios de información operan bajo el formato de sitios web con acceso restringido. De los 12 sistemas, 83.3 por ciento (10) funcionan de esta manera. Sólo 2 sistemas estructuran su operación en el modelo de intranet.

Contrataciones

En materia de contrataciones, el formato de operación más común de los sistemas es el uso de sitios web con acceso restringido. Los administradores de los sistemas indican que de los 23 repositorios de datos del área temática, 17 funcionan en este formato. Mientras tanto, tres funcionan bajo el modelo de intranet, dos no saben (CompraNet) y uno es un sitio web público (base de datos de sancionados INE).

Sanciones a proveedores

En materia de sanción a proveedores, se señaló que la operación se hace mediante sitios web con acceso restringido. Los datos muestran que de los 11 sistemas del área temática, 9 operan en este formato y dos mediante intranet.

Quejas

En materia de quejas y denuncias, se indicó que los sitios operan predominantemente bajo el formato de sitio web con acceso restringido, 11 de los 13 siguen este formato mientras que 2 son intranet.

ALREDEDOR DE LA MITAD DE LA INFORMACIÓN DE LOS SISTEMAS ELECTRÓNICOS Y BASES DE DATOS ESTÁ DISPONIBLE FUERA DE LA UNIDAD O DEPARTAMENTO ENCARGADA

¿Qué nivel de acceso tiene la información?

En su mayoría, la información de los repositorios de información es accesible a las unidades que la gestionan y a terceros interesados. El 44.2 por ciento (23) de los sistemas contienen al menos alguna parte de su información que es accesible a terceros. El 11.5 por ciento (6) de los sistemas reportados tiene disponible toda la información a terceros. En contraste, el 40.3 por ciento (21) de la información no es accesible a nadie fuera de las unidades o departamentos que las gestionan.

La distribución de los sistemas más importantes en términos de volumen tiene un perfil variado de disponibilidad de su información conforme a lo establecido en las normas de transparencia. Estas se alinean entre tres perfiles de accesibilidad que se muestran en el diagrama en la siguiente página.

PLATAFORMAS CON MAYOR VOLUMEN DE DATOS SEGÚN SU NIVEL DE ACCESIBILIDAD A LA INFORMACIÓN.

Sin disponibilidad

- Sistema Integral de Administración de recursos Financieros, CNDH

Disponibilidad parcial

- Plataforma Nacional de Transparencia, INAI

Disponibilidad total

- CompraNet -im, SFP
- CompraNet (plataforma transaccional), SFP

Es interesante observar que de los 21 sistemas que no tienen consumidores de información fuera de la unidad que gestiona la información únicamente 5 sistemas (23.8 por ciento) tienen el 100 por ciento de la información sujeta a clasificación o reserva de datos conforme a lo establecido en la Ley General de Transparencia y Acceso a la Información Pública (LGTAIP).

Dentro del grupo de sistemas y bases de datos detectadas, el 33.3 por ciento tiene más del 50 por ciento de su información clasificada o reservada. No sorprende en este respecto, que únicamente un sistema tenga consumidores de información fuera de la unidad que lo administra, a saber, el Sistema de procedimientos administrativos de responsabilidades de la SFP.

A la luz de esta información vale la pena considerar quienes son los consumidores de información que consultan estos datos. Los datos muestran un comportamiento bastante uniforme. De entre los 21 sistemas que mantienen inaccesible su información, el consumo interno de información se realiza en 57.1 por ciento de los casos por personal de unidades administrativas de las instituciones participantes y en 42.8 por ciento de los casos por áreas de contraloría. Estos dos principales tipos de consumidor interno para información con acceso restringido.

Por otro lado, 28 sistemas tienen toda o parte de la información accesible a terceros. Es interesante observar que entre estos el 50 por ciento (14) no pudo puntualizar qué porcentaje de los campos en su base de datos están reservados bajo la LGTAI. El 50 por ciento restante se dividió equitativamente entre sistemas con menos del 50 por ciento de sus campos reservados y aquellos con 50 por ciento o más de sus campos reservados.

En materia de rectoría de datos es importante implementar medidas que dejen espacios de oportunidad para este tipo de vacíos de información en áreas claves de control de la gestión de la información. Esto se debe corregir mediante políticas de alto nivel que subsanen estos problemas.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

En materia de evolución patrimonial, la información reportada sugiere que los sistemas manejan se distribuyen casi de manera homogénea. Según las respuestas de los participantes, de los 13 sistemas detectados, el 53.8 por ciento (7) ofrecen acceso parcial a la información capturada mientras que los 6 restantes no permiten acceso a nadie fuera de la unidad o departamento que gestiona la información. En este respecto es destacable que nadie ofrece acceso total a la información.

Servidores públicos sancionados

La accesibilidad de los datos en el área temática de servidores públicos sancionados se distribuye casi por la mitad entre el acceso a terceras partes y aquel ofrecido a terceras partes. Según los datos, de los 12 sistemas detectados, el 58.3 por ciento (7) ofrecen acceso a terceras partes mientras que 5 mantienen la información únicamente accesible a los miembros de la unidad o departamento que gestiona el sistema.

Contrataciones

En materia de contrataciones, se reportó que el 65.2 por ciento (15) de los sistemas tienen información accesible a terceros, ya sea a nivel parcial o total. La información indica que de los 23 sistemas detectados en esta área temática con acceso disponible fuera de la unidad gestora, el 43.4 por ciento ofrece un acceso parcial a la información capturada mientras que el 21.7 ofrece acceso total. En el 30.4 restante de los casos se reportó que no se ofrece acceso a terceros.

Sanciones a proveedores

En materia de sanción a proveedores, los sistemas detectados se encuentran divididos casi por la mitad entre sistemas que ofrecen acceso parcial a terceros y aquellos que no ofrecen ningún tipo de acceso. De los 11 sistemas detectados, el 45.4 por ciento (5) de los sistemas ofrece acceso parcial a terceras partes mientras que 36.3 por ciento (4) no tiene acceso a terceras partes fuera de la unidad gestora del sistema. Sólo 1 sistema ofrece acceso total, a saber, el Directorio de proveedores y contratistas sancionados de la SFP.

CONSUMIDORES DE INFORMACIÓN

¿se conocen los consumidores de información?

EL GRUESO DE LOS SISTEMAS TIENEN DEFINIDOS A SUS CONSUMIDORES DE INFORMACIÓN

Las dinámicas de generación y procesamiento de datos en un sistema con una buena rectoría de datos deben estar profundamente asociadas a los consumidores de información. Por tal motivo, es importante observar en qué medida las instituciones tienen claridad sobre sus consumidores de información.

En este respecto, el grueso de las unidades encargadas de sistemas de información fue capaz de definir la ubicación de sus consumidores de información al interior y exterior de su entorno de operación. En el 82.6 por ciento (43) de los casos evaluados las instituciones fueron capaces de definir si contaban con consumidores de información al interior o al exterior de las instituciones. Por tal motivo, se observa con extrañeza que para el 17.3 por ciento (9) de los repositorios detectados no había claridad sobre la presencia de consumidores de su información dentro o fuera de sus equipos de trabajo.

Repositorio de información	Institución que lo administra	Consumidores internos	Consumidores externos
Véndale al INEGI	Instituto Nacional de Estadística, Geografía e Informática	Identificado	Respuesta desconocida
Plataforma Nacional de Transparencia	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	Identificado	Respuesta desconocida
Procedimientos de Contratación	Suprema Corte de Justicia de la Nación	Identificado	Respuesta desconocida
Sistema de Gestión de Adquisiciones	Comisión Nacional de Derechos Humanos	Identificado	Respuesta desconocida
CompralNE	Instituto Nacional Electoral	Identificado	No identificado
DeclaraNet	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	Identificado	Respuesta desconocida
Procedimientos Contratación- Padrón Proveedores y Contratistas (SIPOT)	Suprema Corte de Justicia de la Nación	No identificado	No identificado
DeclaraNet Plus	Secretaría de la Función Pública	Respuesta desconocida	Identificado
DeclaraNet Plus	Instituto Federal de Telecomunicaciones	Respuesta desconocida	Respuesta desconocida

De los 9 casos en que un sistema no pudo identificar claramente si contaba con consumidores de información al interior o el exterior de la institución administradora 8 de ellos no pudieron determinar si contaban con consumidores al exterior de la organización. En contraste, sólo en 3 casos no se pudo determinar si había consumidores al interior de la organización.

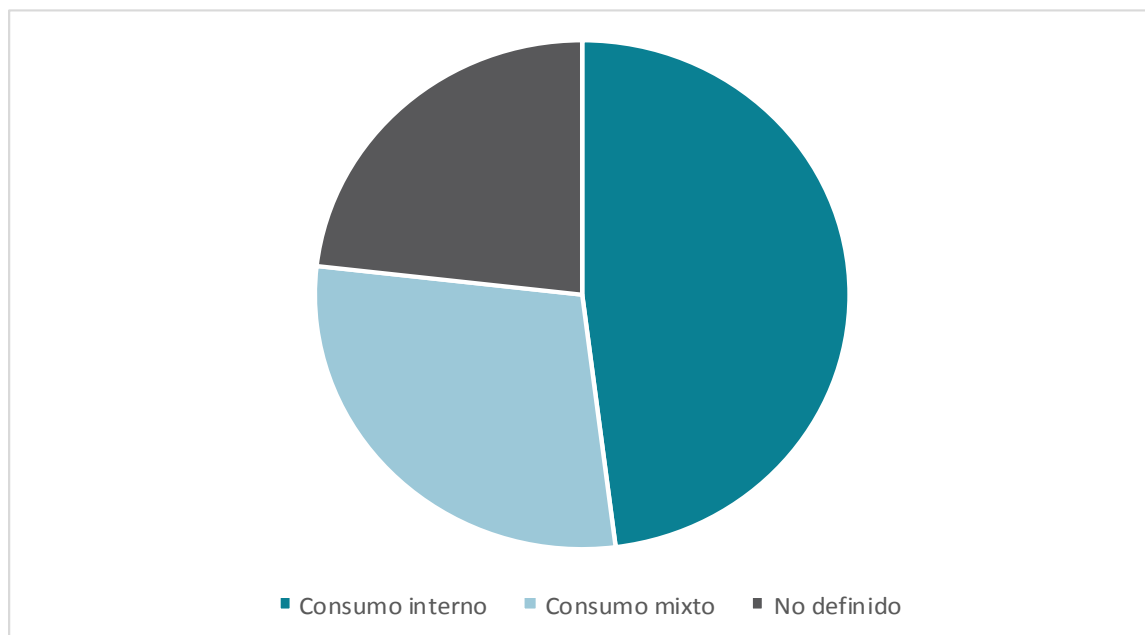
Adicionalmente a los casos anteriormente señalados, se registraron 3 casos en los que se indicó que no se tenían consumidores de información diagnosticados ni dentro ni fuera de la institución que administra el sistema.

Repositorio de información	Institución que lo administra
Buzón electrónico de quejas y denuncias	Instituto Nacional para la Evaluación de la Educación
DeclaralNEE	Instituto Nacional para la Evaluación de la Educación
Registro de licitantes y proveedores sancionados	Instituto Nacional para la Evaluación de la Educación

Dentro de las mejores prácticas en materia de rectoría de datos está considerada la clara definición de las audiencias de la información como parte de los procesos de definición de la plataforma. Por tal motivo, la falta de claridad en audiencias en 23 por ciento (12) de los sistemas evaluados, aún de manera informal o no documentada, es un llamado de atención para la implementación de mejores prácticas en esta materia.

En un escenario óptimo, el equipo de la institución estaría en posibilidades de consultar documentación del sistema en la que se definirían y delimitarían las audiencias internas y externas, así como los límites de su acceso a la información. Por lo tanto, se debe de buscar implementar políticas comprensivas en materia de rectoría de datos que aseguren que esta situación no solo no se observe, sino que se asegure que la definición se haga conforme a una metodología estructurada y homogénea.

PERFILES DE CONSUMO DE INFORMACIÓN EN EL ECOSISTEMA DE INFORMACIÓN DE LA LGSNA.



En contraste con estos casos, la dinámica observada entre aquellas unidades que pudieron precisar si contaban con consumidores de información al interior y exterior de sus instituciones se ajustó a dos perfiles.

1. Consumo interno. Solamente cuenta con consumidores dentro de la misma institución.
2. Consumo mixto. Sujeto a un consumo de información tanto por equipos internos a la organización como a audiencias externas a la misma.

De entre estos perfiles, el consumo interno tiene una presencia predominante. El 48 por ciento de los repositorios (25) de información se ajustaron a este perfil. En contraste, el 28.8 por ciento de los sistemas y bases de datos (15) tienen un tipo de consumo mixto. Esta diferencia señala que el grueso de los sistemas desarrollados se diseñan para satisfacer únicamente las necesidades de información de las instituciones que levantan y procesan la información.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

De la información se puede extraer que los sistemas de evolución patrimonial tienen consumidores en el interior de las unidades que gestionan los sistemas. Los gestores de sistemas indicaron que de entre los 13 sistemas dentro del área temática 69.2 por ciento (9) tienen consumidores al interior de las unidades que los gestionan. Por otro lado, el 30.7 por ciento (4) no tienen usuarios definidos. Asimismo, ninguno de los sistemas tiene usuarios existentes fuera de las unidades o departamentos que las gestionan.

Servidores públicos sancionados

En materia de servidores públicos sancionados, los repositorios de información atienden a usuarios internos. Según la información provista, de entre los 12 sistemas del área temática 91.6 por ciento (11) tienen usuarios al interior de la institución que los gestiona. De entre estos sistemas 4 también tienen usuarios al exterior de la institución que administra la información.

Contrataciones

En materia de contrataciones, el grueso de las instituciones tiene consumidores de información dentro de la institución que las gestiona. En revisión de los datos de cuestionarios se observó que de entre los 23 sistemas del área temática 82.6 por ciento (19) de ellos tienen usuarios al interior de la institución administradora. Sin embargo, de entre estos, sólo el 22.5 por ciento (7) también tiene consumidores al exterior de la institución. No se observaron instituciones que sólo atendieran a usuario externos.

Sanciones a proveedores

En materia de sanción a proveedores, se observó que la mayoría de los sistemas tienen consumidores de información al interior de la institución, pero sólo una porción menor de ellos tiene consumidores al exterior. Del total de los sistemas en el área temática sólo 1 sistema no tenía usuarios al interior de la institución. De entre estos sistemas, sólo el 36.3 por ciento también atiende a consumidores externos.

Quejas

En materia de quejas y denuncias, se reportó que la mayoría de los sistemas y bases de datos tienen consumidores al interior de las instituciones que los gestionan. Según los datos provistos, sólo 1 de los 13 sistemas en el área temática no tiene consumidores al interior de la institución que gestiona el sistema. De entre estos, 3 de los 13 sistemas tienen consumidores de información al exterior de la institución que gestiona los sistemas y bases de datos.

LA MAYORÍA DE LOS SISTEMAS TIENEN INFORMACIÓN RESERVADA. SIN EMBARGO, LOS SISTEMAS CON MAYOR VOLUMEN DE DATOS SON PÚBLICOS

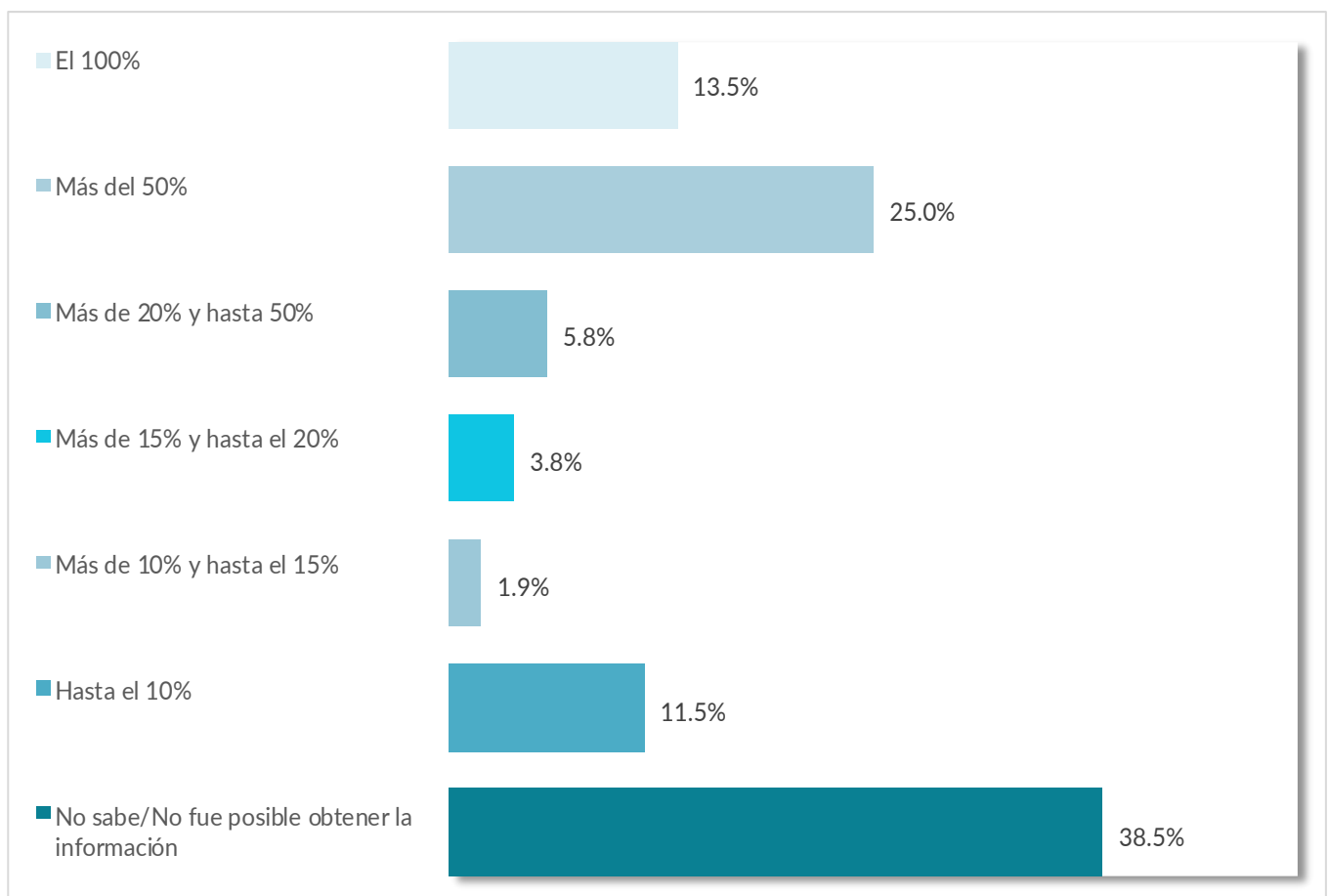
¿En qué medida los campos de información están reservados?

Una parte significativa de los sistemas y bases de datos detectados tienen 50 por ciento o más de sus campos de información reservados conforme a lo establecido en la normatividad de transparencia. El 25 por ciento (13) de los repositorios de información tienen más del 50 por ciento de sus campos reservados mientras que el 13.4 por ciento (7) tienen el 100 por ciento de sus campos de información reservados.

El resto de los sistemas que indicaron esta información (12) tienen entre el 10 hasta el 20 por ciento de su información reservada mediante los mecanismos establecidos por la normatividad aplicable.

No obstante, el impacto de estas limitaciones es relativamente menor, toda vez que ninguno de los 4 sistemas más grandes en términos de volumen de datos se encuentra en la categoría de reserva total o parcial de datos. La información de los módulos registrados del sistema CompraNet (CompraNet-im y CompraNet transaccional) cuentan con una restricción de hasta el 10 por ciento de los datos y aunque no se especifica, la información de la Plataforma Nacional de Transparencia está al menos parcialmente disponible en línea. Esto indica que, en términos de volumen, el grueso de la información está disponible al público.

NIVEL DE RESERVA/CLASIFICACIÓN DE INFORMACIÓN EN EL ECOSISTEMA DE INFORMACIÓN DE LA LGSNA.



RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

El grueso de la información contenida en los repositorios de información en materia de evolución patrimonial se encuentra bajo clasificación o reserva. Según las respuestas a cuestionarios, de los 13 sistemas en el área temática, el 61.5 por ciento tiene más del 50 por ciento de la información reservada. Por otro lado, el 15.3 por ciento (2) de los sistemas en el área temática tienen entre 20 y 50 por ciento de los campos reservados. No hay sistemas plenamente transparentes toda vez que el tema es sensible. El resto de los sistemas no pudieron determinar puntualmente el porcentaje de los campos que se encuentran reservados.

Servidores públicos sancionados

En materia de servidores públicos sancionados, no se pudo determinar el monto de campos de información reservada o clasificada para la mayoría de los sistemas. Según lo reportado, de los 12 sistemas del área temática, el 50 por ciento (6) de los repositorios de información no tienen un diagnóstico preciso del porcentaje de campos reservados. Por otro lado, 25 por ciento de los sistemas tienen el 100 por ciento de la información clasificada o reservada conforme a lo establecido en la ley. El resto de los sistemas existentes tienen entre 10 y más del 50 por ciento de los campos reservados.

Contrataciones

En materia de contrataciones, no se pudo determinar el porcentaje de campos reservados para la mayoría de los repositorios de datos. Se observó que de los 23 sistemas dentro del área temática, 47.8 por ciento (11) de ellos no pudieron puntualizar la proporción de los campos que están reservados en las bases de datos. Para el 21.7 por ciento (5) de los repositorios restantes se tiene hasta el 10 por ciento de los campos reservados y el 13 por ciento (3) tienen el 100 por ciento de los campos de información clasificada o reservada. El resto de los sistemas tienen entre 10 y más del 50 por ciento de la información clasificada.

Sanciones a proveedores

En materia de sanciones a proveedores, no se pudo determinar el porcentaje de campos de información reservada para el grueso de los sistemas. Según la información en los cuestionarios, de los 11 sistemas en el área temática, el 63.6 por ciento (7) no pudo definir precisamente la proporción de campos reservados. El porcentaje restante de los sistemas tienen entre 10 y el 100 por ciento de información reservada.

Quejas

En materia de quejas y denuncias, el grueso de los repositorios tiene el 100 por ciento de la información clasificada o reservada. Se reportó que de los 13 sistemas del área temática, el 46.1 por ciento (6) tienen el 100 por ciento de información clasificada o reservada. Por otro lado, el 30.7 por ciento (4) tienen más del 50% de la información reservada. El resto de los sistemas tienen entre 10 y 50 por ciento de información reservada. Es notable que ninguno de los sistemas es completamente público.

PILAR II. CALIDAD DE DATOS

La fuerza que sostiene las actividades vinculadas al levantamiento, procesamiento, explotación y eliminación de información es la calidad de la información contenida en todos los repositorios de datos. La precisión, completitud y consistencia de la información a través de todas las fuentes de datos son las características de iniciativas exitosas. El pilar de calidad de datos atiende todas las actividades, procesos y protocolos establecidos para realizar validaciones de veracidad y validez de la información, así como algunos protocolos de acción para asegurar la integridad de los datos.

HOY LAS PRÁCTICAS DE PROTECCIÓN DE LA INTEGRIDAD DE LA INFORMACIÓN ABARCAN UN ESPECTRO QUE CUBRE PRACTICAS RIESGOSAS HASTA BUENAS PRÁCTICAS

¿Qué prácticas se observan en el ecosistema de información actualmente?

Como se ha mencionado, existe un progreso desigual a través de las instituciones evaluadas en materia de los protocolos implementados para proteger la confiabilidad de la información almacenada. Pero ¿qué prácticas existen actualmente?

Detectamos un amplio espectro de procesos y procedimientos que asegura en mayor o menor medida la integridad de los datos preservados. De entre esta gama se pudo construir un listado de seis modelos generales con diferentes niveles de protección a la integridad de la información. Estos se enlistan a continuación de menor a mayor nivel de protección.

1. Registros físicos preservados con medidas de seguridad físicas.
2. Manejo de bases de datos en hojas de cálculo con edición directa a campos sin guardado de versiones históricas del documento.
3. Sistemas electrónicos con edición directa a bases de datos sin bitácoras de actividad ni guardado de versiones históricas de bases de datos.
4. Bases de datos guardadas en hojas de cálculo con edición directa a campos con registros físicos que replican registros electrónicos.
5. Sistemas electrónicos con edición de información con bitácoras de actividad con medidas informales de protección a la información, pero sin protocolos establecidos que regulen la actividad alrededor de los repositorios de datos
6. Sistemas electrónicos con edición de información con bitácoras de actividad con medidas formales de protección a la información y con protocolos formales de acceso, edición y consulta de información sobre los repositorios de datos.

Como resultado de la observación de estas prácticas, es importante establecer espacios y herramientas que permitan que las instituciones se coordinen para la implementación de una ruta crítica que ordene la transición del primer modelo hacia el sexto. Esta ruta crítica se puede apoyar en normatividad y lineamientos que incentiven la transición hacia mejores esquemas de protección de datos. Es importante señalar que la inacción en esta materia implica el mantener vivas susceptibilidades importantes de seguridad de la información.

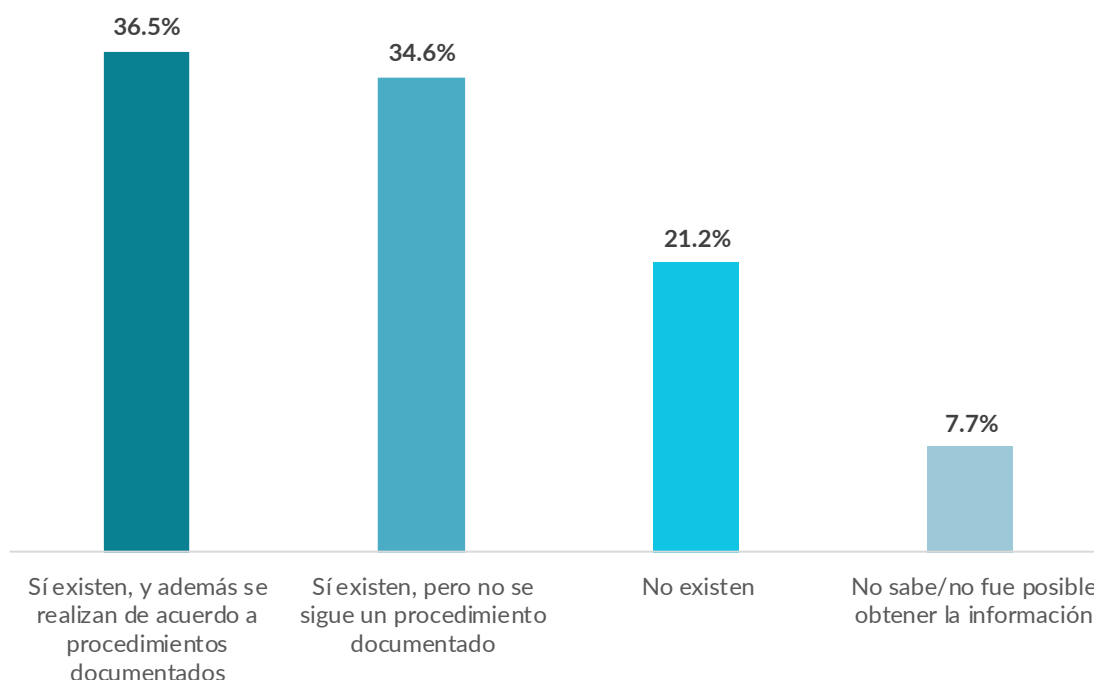
CONTROL DE VALIDEZ

SE TIENE UN BUEN AVANCE EN POLÍTICAS DE CONTROL DE VALIDEZ, SIN EMBARGO, SE DEBE REFORZAR LA FORMALIZACIÓN DE ESTOS PROCESOS

¿Se asegura que la información se ajuste a los parámetros de captura?

Una de las medidas básicas del aseguramiento de la calidad de la información es la implementación de procedimientos que permitan validar que la información capturada en los campos de las bases de datos se conforme a los parámetros previstos para la información. Esta medida asegura que la información tiene un nivel mínimo de validez y sienta las bases para evaluaciones más sustantivas.¹⁵

PENETRACIÓN DE CONTROLES DE VALIDEZ EN EL ECOSISTEMA DE INFORMACIÓN DE LA LGSNA.



¹⁵ En esta dimensión, 8 por ciento de los sistemas no pudieron determinar o no pudieron obtener información sobre medidas de control de validez.

Entre los sistemas detectados, 72 por ciento cuentan con procesos de validación de la información capturada. De entre estos sistemas, el 37 por ciento ha plasmado estas medidas de validación en procedimientos documentados. Sin embargo, aún el 35 por ciento de los procesos de validación no han sido formalizados. Esto representa un riesgo significativo toda vez que estas políticas son susceptibles de abandonarse con mayor facilidad en la eventualidad de rotación significativa de personal o cambios en figuras clave que respalden estas medidas.

Esta situación es más grave en dos áreas temáticas. No deja de sorprender que más del 50 por ciento de los sistemas en materia de servidores públicos sancionados y 62 por ciento en materia de quejas y denuncias no han formalizado sus procedimientos de control de validez de la información. El problema más importante asociado a esto es que aún si se quisiera, no se puede asegurar que estas medidas de validación se conformen a estándares básicos de calidad.

No obstante, existen áreas más problemáticas. En este sentido, salta a la vista que 21 por ciento de los sistemas no cuentan con políticas de control de validez de la información conforme a parámetros previstos. Esta es una situación preocupante toda vez que, al no implementar estas medidas básicas, estas instituciones reducen significativamente la capacidad de diagnosticar objetivamente la calidad de su información y, por lo tanto, de aplicar medidas correctivas en casos que requieran ajustes.

Esta situación no sucede con la misma frecuencia en las distintas áreas temáticas. De hecho, el problema es más grave en dos ámbitos. A pesar de contar con sistemas bastante desarrollados, 31 por ciento de los sistemas detectados en el área de información de evolución patrimonial no tienen medidas de validación.

De manera similar, el 27 por ciento de los sistemas de información sobre proveedores sancionados no implementan medidas de este tipo. Esto, a pesar de que un error en la base de datos puede tener repercusiones legales significativas para una institución.

Como corolario de lo anterior, se debe señalar que es evidente que se requiere la definición de una ruta hacia la implementación y formalización de políticas para el control de validez de información. Esto debe formar parte de una política apropiada de calidad de la información.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

Según las respuestas a cuestionarios, se cuenta con procesos implementados y documentados para controlar la validez de la información para la mayoría de los repositorios detectados. Sin embargo, existe una gran porción de ellos que no desarrollan ninguna actividad en este sentido. La información provista revela que de los 13 sistemas en el área temática, el 61 por ciento de los sistemas sujeta su información a procesos documentados para controlar de la información capturada. Sin embargo, de estos sólo el 38 por ciento están documentados mientras que el 23 por ciento restante se desarrollan conforme a procedimientos no

documentados. No obstante, 31 por ciento de los sistemas no tienen ningún tipo de medida de diagnóstico de validez de la información capturada.

Servidores públicos sancionados

En materia de servidores públicos sancionados, se han implementado procesos de control de validez de la información capturada para la mayoría de los repositorios detectados. Los encargados de estos indicaron que de los 12 sistemas en el área temática, el 50 por ciento tiene procesos de control de validez de la información, sin embargo, esto no están documentados. Sólo el 33 por ciento de los sistemas en esta área temática tienen procedimientos documentados de control de validez de los datos conforme a los parámetros de captura de datos. El 17 por ciento restante no tiene ningún proceso implementado.

Contrataciones

En materia de contrataciones, el grueso de los sistemas reportados tiene procesos de control de validez de la captura de información que han sido documentados. Conforme reflejan los datos, de los 23 sistemas dentro del área temática, el 73 por ciento de los sistemas tiene procesos de control de validez de captura. De entre estos, 43 por ciento están documentados mientras que 30 por ciento no lo están. En esta área temática, sólo el 17 por ciento no tiene implementado ningún tipo de control de este tipo.

Sanciones a proveedores

En materia de sanción a proveedores, se han implementado procesos de control de validez de la captura de información para la mayoría de los repositorios de información. De entre los 23 sistemas dentro del área temática, el 72 por ciento de los sistemas tiene procesos de control de validez de captura de la información conforme a los parámetros previstos. De entre estos, la mitad (36 por ciento) tiene estos procesos documentados y la otra mitad no los ha documentado. El restante 27 por ciento no tiene ningún tipo de medidas de validación de este tipo.

Quejas

En materia de quejas y denuncias, de han desarrollado medidas no documentadas de control de validez de captura de información para la mayor parte de los repositorios detectados. Según la información provista, de los 13 sistemas del área temática, el 62 por ciento tiene procesos no documentados en este respecto. En esta área, tan sólo el 15 por ciento de los sistemas con medidas de validación de este tipo ha documentado estos procedimientos. Finalmente, sólo el 15 por ciento de estos sistemas tiene validaciones de validez de los datos capturados.

LA INFORMACIÓN ESTÁ SUJETA A VALIDACIONES DE VERACIDAD. SIN EMBARGO, SE REQUIERE TRABAJAR EN LA SISTEMATIZACIÓN DE ESTAS MEDIDAS PARA ROBUSTECERLAS

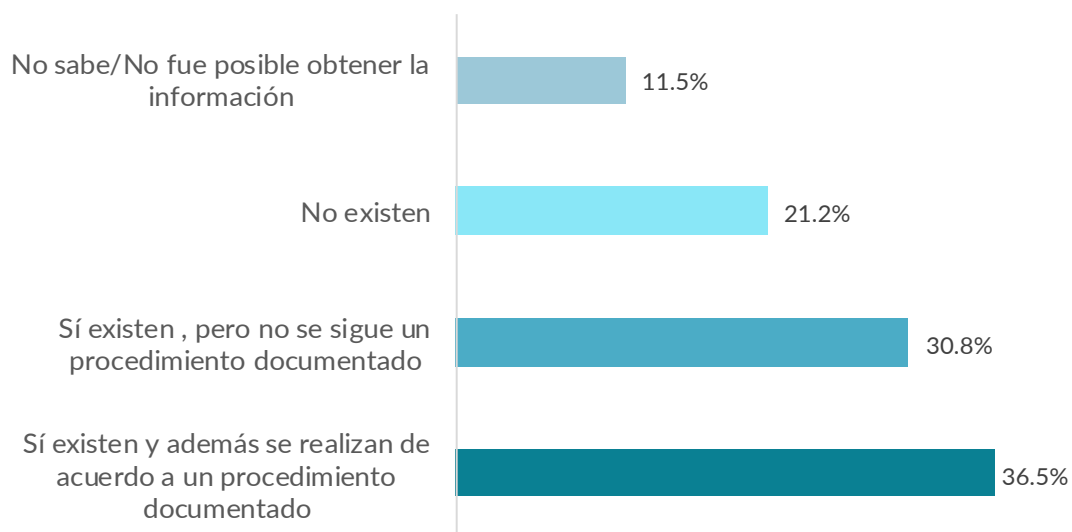
VALIDACIÓN DE VERACIDAD

¿En qué medida se asegura la veracidad de la información registrada?

El fundamento del valor de toda base de datos es la veracidad de la información. Por tanto, uno de los elementos centrales en materia de calidad de datos es que las unidades cuenten con políticas y procesos que aseguren que los datos reflejen de manera precisa la realidad. Como resultado de ello, el proceso de mejoramiento de la calidad de datos pasa por la implementación y documentación de prácticas de validación de veracidad de los datos.

Entre los sistemas detectados, se observó que el 68 por ciento cuenta con esquemas de validación de veracidad. No obstante, de entre estos sólo el 37 por ciento se desarrolla conforme a procedimientos documentados. El 31 por ciento restante opera bajo esquemas no documentados. Este diagnóstico muestra la relevancia de políticas en materia de calidad de datos. En la medida en que no se tenga una política de alto nivel en este respecto no se podrá avanzar hacia una aplicación comprensiva de medidas básicas de validación de datos.

En este respecto, los hallazgos del estudio no son menores. De entre los 19 sistemas que cuentan con medidas formales de validación de veracidad de la información, sólo poco más de la mitad (11 o 57.8 por ciento) aplica estas medidas con frecuencia. Por su parte, de los 16 sistemas sobre los cuales no se ha documentado procedimientos, sólo 7 hace validaciones con frecuencia.



Nada de lo anterior sugiere que haya un avance robusto en la materia. Sin embargo, el problema no sólo es de forma, sino que los datos sugieren que también se debe de hacer una evaluación de fondo de las prácticas implementadas. Si bien una mirada inicial a los datos apunta a que la mayoría de los sistemas tiene prácticas de validación de veracidad el cruce de variables parece indicar otra cosa.

Como parte de un modelo de gestión de datos es esencial que se formalicen en documentación los procesos de validación. Sin embargo, también es importante evaluar su calidad. En este respecto, los datos apuntan a una situación que requiere atención. Esto se debe a que, de los sistemas con procedimientos documentados de validación de veracidad, más de la mitad (68.4 por ciento) realiza estas verificaciones mediante procesos distintos a la auditoría de los datos. Más aún, entre los sistemas que no han documentado sus procedimientos de validación este porcentaje sube al 81.2 por ciento.

“A veces, sí revisamos los casos donde es obvio que se les durmió la mano en el cero. Porque no es posible que tengan más dinero que el presupuesto de la institución.”

El motivo de que esto llame la atención es que durante el proceso de levantamiento en sitio se observó que estos modelos alternativos no satisfacen los requerimientos de una gestión apropiada de datos ya que operan conforme a dos modalidades:

1. Evaluaciones casuísticas de veracidad de datos detonadas por procesos internos no vinculados con la gestión de la información. i.e. Auditoría de información de evolución patrimonial de servidores públicos como resultado de procesos de investigación.
2. Revisión de veracidad de valores extremos en campos de información. i.e. Validación con usuarios de registros que rebasan valores esperados en materia de ingresos, deudas o valor de bienes (demasiados ceros).

Estos comportamientos no representan, propiamente dicho, un esquema de validación de veracidad. Si bien logran validar información mediante la confirmación de información no representan esquemas estructurados y sistemáticos que permitan asegurar la integridad de la información. Por lo tanto, con el fin de asegurar la integridad de los datos y su valor documental se debe de buscar la implementación de esquemas robustos de validación de datos que representen medidas sustantivas que aseguren su precisión.

Finalmente, en este respecto se debe señalar que 21 por ciento de los sistemas no cuenta con ningún tipo de política de verificación de veracidad, cifra que se complementa con el 12 por ciento de los sistemas para los cuales no se pudo definir precisamente si contaba con un esquema de este tipo (no sabe o no fue posible obtener la información).

La situación descrita anteriormente es alarmante toda vez que cualquier explotación de esta información está sujeta a cuestionamientos sobre la veracidad de la información y en estos casos las instituciones no cuentan con ninguna herramienta para estimar la probabilidad de que el conjunto de datos tenga respaldo documental de algún tipo.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

Los sistemas de evolución patrimonial están sujetos a alguna forma de validación de veracidad. De entre los 13 sistemas en el área temática, se realizan validaciones de veracidad conforme a procedimientos documentados para el 46 por ciento. Mientras tanto, se reportó que el 31 por ciento que también realiza validaciones de veracidad no ha documentado sus procesos. En contraste, sólo el 8 por ciento no realiza validaciones de veracidad y 16 por ciento no pudo determinar si emprende estas actividades (no sabe o no pudo obtener la información).

Servidores públicos sancionados

En materia de servidores públicos sancionados, la mayoría de los sistemas se sujeta a validaciones de veracidad. De los 12 sistemas en el área temática, los equipos encargados de su administración indicaron que el 84 por ciento recibe alguna forma de validación de veracidad. De entre estos sistemas validados, la mitad tiene procedimientos documentados y la mitad no han sido documentados. Por otro lado, el 17 por ciento de ellos no tienen ningún procedimiento de validación.

Contrataciones

En materia de contrataciones, se cuenta con procesos de validación documentados para la mayoría de los repositorios detectados. De entre los 23 sistemas registrados en esta área, el 35 por ciento está sujeto a algún tipo de validación documentada. Por otro lado, el 30 por ciento está sujeto a procedimientos de validación no documentados. Muy cerca de este número, el 22 por ciento de los sistemas no tiene ningún tipo de validación de veracidad. Adicionalmente, el 12 por ciento no pudo determinar con precisión si han implementado procedimientos de este tipo o si están documentados (no saben o no respondieron).

Sanciones a proveedores

En materia de sanción a proveedores, se cuenta con procedimientos de validación para la mayor parte de los sistemas. El 54 por ciento de los encargados de sistemas afirmaron contar con algún proceso de validación de veracidad. Sin embargo, la mitad de ellos los ha documentado y la otra mitad no. Cada uno de estos subgrupos ocupa el 27 por ciento del total de la población. En contraste, el 45 por ciento de la población no cuenta con ningún tipo de medida de validación de la veracidad de la información.

Quejas

En materia de quejas y denuncias, la mayoría de los sistemas está sujeto a medidas de validación no documentadas. Se reportó que de los 13 sistemas en el área temática, el 54 por ciento aplica medidas no documentadas de validación de la veracidad. Mientras tanto, el 15 por ciento ha tomado acciones para documentar las medidas aplicadas para la validación de la veracidad. En contraste, 23 por ciento no ha implementado ninguna medida para validar la veracidad de la información en sus sistemas.

ROBUSTEZ Y DIAGNÓSTICOS DE CALIDAD DE LA INFORMACIÓN

34.6% DE LAS INSTITUCIONES NO FUERON CAPACES DE DETERMINAR LA MAGNITUD DE LOS VACÍOS DE INFORMACIÓN EN SUS REPOSITORIOS DE INFORMACIÓN, LAS QUE SÍ, REPORTARON MAYORITARIAMENTE QUE MENOS DEL 5% DE SUS REGISTROS DE INFORMACIÓN NO TIENEN DATOS

¿Existen vacíos de información en los repositorios de información?

El ecosistema de información muestra algunos comportamientos de riesgo muy extendidos. Es notable que el 34.6 por ciento (18) de los sistemas o bases de datos detectados no contaron con un diagnóstico o no fueron capaces de diagnosticar el volumen de registros sin información dentro de sus repositorios de información. Esto tiene implicaciones importantes toda vez que una apropiada gestión de la información comienza por contar con un claro control de los datos recabados.

Sin embargo, hay avances importantes en este respecto en otros sistemas. Dentro del ecosistema de información de la LGSNA el 55.7 por ciento (29) de estos repositorios diagnosticaron que el número de registros sin datos se encuentra entre el 0 y 5 por ciento. Los 4 sistemas más grandes en términos de volumen de datos reportaron ubicarse en este grupo. En contraste, sólo 5 por ciento (5) de los sistemas electrónicos o bases de datos considerados indicaron tener más del 5% de los registros sin datos.

Es interesante observar que todos los sistemas que diagnosticaron que tenían entre 0 y 5 por ciento de los registros sin vacíos de información también indicaron tener entre 0 y 20 por ciento de campos con vacíos de datos. Asimismo, los sistemas que pudieron determinar que sus registros estaban por encima del 5 por ciento (7.6 por ciento) indicaron también que el 20 por ciento de los registros se encontraban en el nivel más bajo de observaciones sin registros posible.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

La mayor parte de los sistemas de evolución patrimonial muestran una carencia de diagnóstico sobre los vacíos de datos en sus registros y campos. De los 13 sistemas en el área temática, solo en el 38.4 por ciento se reportó que variables y registros sin datos se encuentran en niveles bajos o nulos. En contraste, el 61.5 por ciento restante de los equipos que administran los sistemas no pudo determinar de manera precisa en qué situación se encontraban sus plataformas en materia de campos o registros.

Servidores públicos sancionados

En materia de servidores públicos sancionados, se reportó que los sistemas tienen un buen desempeño en materia de diagnóstico de medidas de vacíos de información.

La información permite observar que de los 12 sistemas en el área temática, 91.6 por ciento (11) tienen entre 0 y 20% de sus campos sin datos. Asimismo, de estos sistemas el 83.3 por ciento tiene entre 0 y 5 por ciento de los registros con vacíos de información.

Contrataciones

En materia de contrataciones, se observa progreso en el diagnóstico de vacíos de información, sin embargo, aún hay una presencia significativa de sistemas sin capacidad de diagnóstico en este respecto. Se reportó que, de entre los 23 sistemas en el área temática sólo para el 60.8 por ciento se reportó la medida en que sus bases de datos cuentan con registros vacíos. Del total de sistemas en el área, el 43.4 por ciento tiene entre 0 y 20 por ciento de sus campos vacíos y entre 0 y 5 de sus registros vacíos. Por otro lado, el 17.3 por ciento indicó que sus bases de datos superan el nivel más bajo de registros y campos vacíos, señalando que tienen más del 20 por ciento de sus campos vacíos y más del 5 por ciento de sus registros vacíos. Llama la atención que 3 de los 4 sistemas que se encuentran en esta situación corresponden a la SCJN.

Sanciones a proveedores

En materia de sanción a proveedores, el área temática está dividida casi por la mitad entre los sistemas para los que se puede diagnosticar su situación de campos y registros vacíos y aquellos que no. Se indicó que de los 11 sistemas en el área temática, el 54.5 por ciento pudo diagnosticar su situación en materia de campos y registros vacíos. De entre estos la mayoría (5 de 6) encontró que sólo entre el 0 y 20 por ciento de sus campos y el 0 y 5 por ciento de sus registros tiene vacíos de información. El resto de los sistemas no fue capaz de diagnosticar su situación ya que no les fue posible obtener la información o simplemente no respondieron.

Quejas

En materia de quejas y denuncias, para la mayoría de los sistemas se pudo diagnosticar su situación en materia de campos y registros vacíos. Los encargados indicaron que de los 13 sistemas en el área temática, el 76.9 por ciento tuvo un diagnóstico. De entre estos la mayoría (9 de 10) indicó que sólo entre el 0 y 20 por ciento de sus campos y el 0 y 5 por ciento de sus registros tiene vacíos de información. Sólo 3 sistemas no pudieron diagnosticar su situación.

UNA BUENA PARTE DE LOS SISTEMAS REPORTADOS POR INSTITUCIONES NO MANEJAN UN REGISTRO PRECISO Y CONSISTENTE DE LOS CAMPOS SUJETOS A EDICIÓN

¿Qué porcentaje de la información ha sido editada?

La información levantada indica que el grueso de los sistemas y bases de datos no cuentan con un diagnóstico preciso del porcentaje de registros de información sujetos a edición de datos.

Para el 51.9 por ciento (27) de los repositorios analizados los administradores no estuvieron en posibilidad de determinar el porcentaje de edición o desconocen la cantidad de estas intervenciones.

Esto es un problema por dos razones. Primero, porque debería tenerse claridad sobre las modificaciones de variables debido a datos faltantes o errores de inconsistencia de la información que podrían tener un impacto en la calidad de los datos analizados. Segundo, porque no contar con bitácoras de edición de datos confiables representa una mala práctica en el desarrollo de sistemas que genera vulnerabilidades en seguridad y calidad de datos.

Para casi la mitad restante de sistemas electrónicos y bases de datos, esto es el 44.2 por ciento (23), los administradores pudieron determinar que su nivel de edición de datos es de hasta el 10 por ciento.

Es importante señalar que esta ausencia de capacidad para determinar un elemento central de la política de administración de datos es un problema general para todos los sistemas identificados y no sólo para aquellos que no usan manejadores de bases de datos. La información reportada genera un foco rojo en términos de confiabilidad de la información dentro de los sistemas o bases de datos y refuerza la necesidad de establecer políticas generales de desarrollo de repositorios de datos y su proceso de maduración.

Llama la atención que de entre los sistemas que no tienen una estimación del volumen de edición de datos se reporta que en el 27 por ciento se realizan ediciones a las bases de datos. En el 52.9 de estos sistemas la edición se realiza directamente en la base de datos. Lo anterior indica que estas ediciones se hacen mediante herramientas que permitan controlar cambios a la información de manera sistemática. Los sistemas en esta categoría abarcan todo tipo de perfiles de información como contrataciones (3), evolución patrimonial (2), servidores públicos sancionados (3) y quejas y denuncias (1). Toda vez que estos sistemas, por su perfil de información, manejan datos de alta sensibilidad, es importante que se implementen mejores prácticas en la materia para asegurar una revisión sistemática de información en materia de edición de datos.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

En materia de evolución patrimonial, la mayoría de los administradores de los sistemas no pudieron diagnosticar sus niveles de edición de información. Respecto de los 13 sistemas del área temática, no se pudo indicar con precisión el nivel de edición de información del 61.5 por ciento de ellos. El 38.4 por ciento restante de los administradores de sistemas estimó que hasta el 10 por ciento de sus datos habían sido editados.

Servidores públicos sancionados

En el área de servidores públicos sancionados, la mayoría de los administradores de los sistemas no fue capaz de determinar en qué medida se habían hecho ediciones de datos en sus registros en la base de datos. De los 12 sistemas en el área temática, en el 58.3 por ciento de los casos las áreas respondientes no pudieron determinar la magnitud de la edición de datos en sus bases de datos ya sea porque no pudieron obtener la información (4) o porque se desconoce el porcentaje (3). Entre los sistemas que pudieron determinar la magnitud de estas ediciones el 100 por ciento indicó que estas habían sido de hasta el 10 por ciento.

Contrataciones

En materia de contrataciones, el universo de sistemas se divide casi a la mitad entre aquellos sobre los que se pudo diagnosticar su nivel de edición de datos y aquellos que no. Por un lado, en el 52.1 por ciento de los casos, los administradores de los sistemas pudieron diagnosticar el nivel de edición. De entre estos sistemas, el 91.6 por ciento tiene hasta el 10 por ciento de edición de datos, mientras que sólo en un caso esta cifra se encuentra entre el 15 y el 20 por ciento. Los administradores de los sistemas no pudieron diagnosticar su situación en este respecto para los casos restantes.

Sanciones a proveedores

En materia de sanción a proveedores, el área temática está dividida casi por la mitad entre los sistemas que pueden diagnosticar su situación de campos y registros vacíos y aquellos que no.

De entre los 11 sistemas en el área temática, en el 52.1 por ciento de casos reportados los administradores de los sistemas pudieron diagnosticar su volumen de edición de datos. En este subconjunto, el 91.6 por ciento indicó que se editó hasta el 10 por ciento de sus datos. Solamente un equipo administrando un sistema señaló que esta cifra se encontraba entre el 15 y el 2 por ciento de edición. En el resto de los casos el equipo de administración no pudo obtener la información.

Quejas

La mayor parte de los equipos que administran los sistemas de quejas pudieron diagnosticar el volumen de edición de datos y lo ubicaron, en su mayoría en el 10 por ciento. De entre los 13 sistemas detectados en el área, en el 61.5 por ciento (8) de los casos los equipos que lo administran pudieron estimar su volumen de edición de datos. De entre estos casos, el 75 por ciento estimó que esta cifra se encontraba entre el 0 y el 10 por ciento. En el resto de los casos se indicó que esta cifra estaba entre el 15 y el 20 por ciento. Finalmente, el 38.4 por ciento de los equipos que administran los sistemas no pudo diagnosticar el volumen de edición de datos.

PILAR III. CONSISTENCIA CONCEPTUAL DE ESTRUCTURA DE DATOS

La gestión de estructuras de datos o *master data management* es una combinación de acciones, políticas y tecnologías orientadas a la consolidación, limpieza y potencialización de los repositorios de información estratégicos a fin de que estos se sincronicen con los procesos de negocio y robustezcan las herramientas analíticas de cada institución. Estas acciones se centran directamente en el control de las estructuras de datos y las políticas y acciones requeridas para que este esfuerzo se desarrolle de manera óptima y ordenada. Este control cuenta con insumos básicos entre los cuales destacan los metadatos de la información almacenada. Los beneficios de estas acciones son orientados a la mejora de la eficiencia operacional, la calidad de los reportes de información y la toma de decisiones basada en datos. Asimismo, permiten contar con los insumos básicos de la gobernanza de datos requerida para que un ecosistema de información opere de manera ordenada y coordinada.

LEVANTAMIENTO DE METADATOS

SE REQUIERE EMPUJAR LA PROFUNDIZACIÓN Y MEJORA DEL LEVANTAMIENTO DE METADATOS YA QUE LA MAYORÍA DE LOS SISTEMAS NO DOCUMENTAN METADATOS Y CUANDO LO HACEN NO SE AJUSTAN A ESTÁNDARES INTERNACIONALES

¿Que tan generalizado está el levantamiento de metadatos?

Uno de los insumos fundamentales en el desarrollo ordenado de un ecosistema de información es la integración de documentación que permita la homogeneización conceptual y operativa de los campos y conceptos de la información capturada en áreas temáticas comunes. La documentación requerida en este respecto es conocida como documentación de metadatos. Esta información en su definición más simple es una serie de datos sobre los conjuntos de datos. Los metadatos describen una vista resumen de la información y características de los conjuntos de datos.

La principal distinción que se puede hacer entre la documentación en general de un sistema y los metadatos está basada en que los metadatos tienen una estructura que permite que sean procesados por un ordenador. Esta última condición les permite ser transmitidos más fácilmente y ser interoperables. Por tal motivo, un buen conjunto de metadatos permite a los que lo consultan a comprender el contenido de un conjunto de datos, así como sus alcances y limitaciones.¹⁶

¹⁶ En buena medida esta definición descansa en las definiciones de la CEPAL en esta materia. Para mayor información véase la bibliografía de gestión de datos de investigación disponible en el URL <https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/metadatos>.

La existencia de esta documentación en los órganos públicos gubernamentales es un buen indicador de una apropiada gestión de estructuras de datos. Por tal motivo, la falta de esta documentación en el ecosistema de datos de la LGSNA es una llamada de atención que invita a la implementación de políticas de alto nivel en materia de gestión de estructuras de datos. Esto se debe de hacer sobre todo en campos clave mediante los cuales se identifica la información y en los formatos en que se almacena información fundamental de cada área temática.

Al momento del levantamiento de información, el 61.5 por ciento de los administradores de sistemas detectados indicaron que no documentaban los metadatos de sus repositorios de información. Esta falta de documentación se encuentra en promedio en un 26.8 por ciento a través de las áreas temáticas definidas por la LGSNA. Las tres áreas con mayor cantidad de sistemas sin documentación de metadatos son las de evolución patrimonial, servidores públicos sancionados e información de quejas y denuncias.

En este respecto, es importante señalar que se observó un problema en la consistencia de la información provista por los encargados de proveer la información. De entre los 13 casos evaluados en que se afirmó contar con documentación de metadatos, en 6 de ellos también se indicó que no se contaba con documentación básica de metadatos como diccionarios de datos o diagramas entidad-relación. En ninguno de estos casos se indicó que se contara con algún otro tipo de documentación de metadatos. Por tanto, se debe de considerar que el número podría no ser representativo.

Como respuesta a esta situación se evaluó en qué casos las instituciones afirmaron contar con uno o más documentos de metadatos. Este análisis muestra un resultado aún más preocupante. Del total de sistemas detectados, sólo para el 51.9 por ciento de ellos se indicó que existe uno o más documentos de metadatos.

Sin embargo, del total de los sistemas, sólo el 26.9 por ciento de los respondientes pudo indicar que contaba con más de un documento de metadatos y señalar cuál es este documento. Mientras que, por otro lado, el 25 por ciento restante de los equipos a cargo de estos sistemas pudo indicar únicamente un tipo de documento específico de metadatos.

Por tanto, se debe considerar que la verdadera cifra de ausencia de documentación se puede encontrar en un espectro que abarca de 48.1 al 61.5 por ciento. Esto indica que existe una necesidad de llevar a cabo una revisión puntual de la documentación de metadatos a través de las instituciones participantes.

Sin embargo, esta no es la única inconsistencia observada en la materia. Entre los 32 casos en que los equipos encargados de un repositorio de información señalaron no contar con documentación de sus metadatos, el 31.2 por ciento de ellos afirmó contar con mecanismos de control de cambios en los procedimientos o definiciones de sus estructuras de metadatos. En la práctica internacional, este tipo de inconsistencia de información se observa comúnmente en dos escenarios:

1. Las unidades o departamentos que manejan la información cuentan con políticas de metadatos, pero estas no se aplican de manera uniforme a todos los sistemas.
2. La política de manejo de metadatos se aplica conforme a silos funcionales pero estas prácticas exitosas de gestión de datos no se aplican uniformemente dentro de la institución.

Dentro de los modelos de maduración de gestión de datos estos escenarios se observan por lo común en la etapa inicial de maduración en el manejo de datos en los cuales se observa que el segundo escenario es el más común. La experiencia internacional indica que la verdadera causa de esta situación es la ausencia de políticas de alto nivel en materia de gestión de estructuras de datos que deben establecer parámetros para la generación de documentación y su manejo a través del tiempo.

Como se indicó, este tipo de comportamientos anómalos o inconsistentes están generalmente asociados a un ecosistema de información en el que sus integrantes individuales no cuentan con estándares robustos, claros y abiertos en materia de gestión de datos. La información obtenida durante el estudio sugiere que en entorno de repositorios de información de la LGSNA no es una excepción. En efecto, del total de los casos reportados, sólo 3 equipos encargados de los repositorios de información indicaron que se documentan los metadatos de sus sistemas y bases de datos conforme a estándares internacionales en la materia. Sin embargo, de entre ellos uno pudo indicar solamente contar con diccionario de datos, otro no específico el tipo de documentación disponible y el tercero indicó sólo contar con manuales del sistema que no implican la formalización de metadatos.

Lo anterior, implica que, aunque para el 51.9 por ciento de los repositorios de información se cuente con algún tipo de documentación de metadatos, estos documentos pueden tener niveles muy variados de metodologías, calidad y orientación. La relevancia de esto es que estas variaciones requerirán esfuerzos de homogeneización que también tendrán que aplicarse a los estándares de metadatos además del trabajo requerido para la alineación conceptual y operativa de las estructuras de datos.

Todo lo observado, implica que no se tienen los insumos necesarios para la implementación de políticas que permitan el diagnóstico, atención e implementación de medidas correctivas respecto de la consistencia de las estructuras conceptuales de la información. Esto debe ser remediado con el fin de asegurar que el ecosistema de información pueda madurar y desarrollarse de manera armónica.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

En materia de evolución patrimonial, se cuenta con alguna documentación de metadatos para la mayoría de los sistemas detectados. De los 13 sistemas detectados con relevancia para el área temática, en el 84.6 por ciento de los casos, los administradores de los sistemas indicaron que no se cuenta con documentación de

metadatos. Sin embargo, en el 69.2 por ciento de los casos se indicó que se contaba con al menos un documento de metadatos.

Servidores públicos sancionados

En materia de servidores públicos sancionados, se observa una falta de documentación de metadatos. De entre los 12 repositorios de información, se reportó que el 83.3 por ciento (10) no cuenta con documentación de metadatos. No obstante, el 41.7 por ciento indicó que se cuenta con algún tipo de documentación de metadatos.

Contrataciones

En materia de contrataciones, se observa un progreso mínimo en la integración de documentación de metadatos. De entre los 23 repositorios de información sólo para el 47.8 por ciento se reportó que se cuenta con documentación de metadatos. En esta área temática las cifras coinciden entre las cifras reportadas de metadatos y la existencia de documentación puntual en la materia.

Sanciones a proveedores

En lo concerniente a la sanción a proveedores, hay un avance marginal en la documentación de metadatos. De entre los 11 sistemas relevantes para el área temática, se reportó que sólo para el 36.4 por ciento de ellos se integra algún tipo de documento de metadatos. Esta información es consistente con la documentación indicada por los responsables en otros reactivos.

Quejas y denuncias

En materia de quejas y denuncias, se observa un avance limitado en materia de integración de documentación de metadatos. Asimismo, se observa una diferencia importante entre lo reportado en materia de integración de metadatos y la documentación específica que se indicó que está disponible respecto de los repositorios de información detectados. De entre los 13 sistemas relevantes para el área temática sólo para el 15.4 por ciento se reportó que se cuenta con documentación de metadatos. Sin embargo, se señaló que el 38.5 por ciento de los sistemas en el área cuentan con algún tipo de documentación de metadatos (i.e. diccionario de datos o diagrama entidad relación). Por tanto, es importante realizar revisiones de esta documentación para poder determinar el estatus exacto en el que se encuentra el área temática.

PILAR IV. DOCUMENTACIÓN DE ESTRUCTURAS DE DATOS

Cualquier esfuerzo de integración, desarrollo, ampliación y diagnóstico de los niveles de madurez, estabilidad y desarrollo de un repositorio de información requiere documentación robusta. Esta documentación permite una clara determinación de riesgos, definición de huecos de información, potencial de crecimiento, diagnóstico de requerimiento de ajustes, entre otras acciones previsibles requeridas para el perfeccionamiento de las estructuras de datos. Su ausencia representa un riesgo en múltiples escenarios como rotación inesperada de personal clave.

LA MAYORÍA DE LOS SISTEMAS NO INTEGRA NINGÚN TIPO DE DOCUMENTACIÓN DE LAS ESTRUCTURAS DE DATOS

¿Qué tan bien documentadas están las estructuras de datos?

La documentación de estructuras de información es una herramienta estratégica en el diagnóstico de la maduración de las estructuras de información en múltiples dimensiones. ¿Pero qué es lo que te dice esta documentación? ¿para qué se necesita? Entre otras cosas, el análisis de estos documentos permite determinar las siguientes preguntas clave:

Compatibilidad. En qué medida las variables contenidas en una base de datos tienen los mismos campos de información (conceptos de información) y en qué medida estos se definen de la misma manera y se recuperan en campo de manera similar. i.e. ¿El CURP se codifica de la misma manera? ¿ingreso total incluye bonos y prestaciones?

Potencial de crecimiento. En qué medida la información disponible está estructurada de manera que se puedan hacer interrelaciones entre bases de datos para potencializar la capacidad de consulta de información a través de instituciones públicas gubernamentales. i.e. ¿Podemos hacer cruces de información para asegurar que un servidor público no está sancionado en ninguna institución pública gubernamental? ¿Podemos interconectar todas las bases existentes? ¿Se puede integrar toda la información en una única plataforma?

Capacidad explicativa. Cuál es la capacidad con la que contamos para hacer inferencias como resultado de las estructuras de datos. i.e. ¿Podemos hacer validaciones de la información de evolución patrimonial según lo que se ha reportado por un servidor público a través de instituciones del ejecutivo federal y órganos autónomos?

Por tanto, la disponibilidad de esta documentación tiene un impacto significativo sobre la accesibilidad y capacidad de analizar la información dentro de los repositorios de datos.

Existe una amplia variedad de formatos de documentación que se puede generar alrededor de un sistema electrónico o base de datos. Sin embargo, es importante que esta documentación cubra, al menos, los tres tipos de metadatos pues que cada uno de ellos cumple una función específica y relevante que permite localizar y acceder a un conjunto de datos, así como comprender el contexto en el que estos han sido creados y gestionados, facilitando su correcta interpretación. Estos tipos son los siguientes:

1. Metadatos descriptivos. Información sobre el contenido del recurso de información, que es de utilidad para su recuperación y comprensión. i.e. Nombre, RFC, cargo, etcétera.
2. Metadatos administrativos. Información necesaria para manejar un recurso o que se relaciona con su creación. Dentro de esta categoría se encuentran los metadatos técnicos (necesarios para decodificar y procesar los recursos de información), los metadatos de preservación (necesarios para la gestión y migración en el largo plazo), y los metadatos de derechos (referidos a la propiedad intelectual de la información).
3. Metadatos estructurales. Describen las relaciones que existen entre las partes de un conjunto de datos. Por ejemplo, un esquema que representa las relaciones entre tablas de una base de datos o una tabla de contenidos (Diagrama Entidad Relación).

Con esto en mente, y buscando recuperar la mayor cantidad de información posible sobre la documentación disponible para los sistemas y bases de datos, se solicitó a las instituciones reportar si contaban con parte o todos de tres tipos de documentación relevantes para los metadatos y para la estructura operativa de los sistemas detectados. Esta documentación permite hacer un proxy del nivel de información disponible respecto de los repositorios de información, así como de la formalización de los sistemas electrónicos que las respaldan. La documentación sobre la que se indagó es la siguiente:

1. Diccionario de datos
2. Casos de uso
3. Modelo entidad relación

Adicionalmente, se dejó un espacio para que, en caso de contar con documentación de metadatos alternativa a la señalada, las instituciones pudieran reportar la documentación disponible. Esta documentación alternativa de metadatos se presentó en 23.1 por ciento de los casos (14). Durante el análisis de la información se observó que el grueso de las entradas en este respecto (8) se refería a manuales de usuario, manuales de instalación, historias de usuario u otros documentos que si bien pueden llegar a representar documentación valiosa para fines operativos del sistema no añaden información relevante para los fines del presente análisis. Por tal motivo, sólo se consideró a 6 de las entradas de documentación alternativa que correspondían a información sobre la arquitectura de las bases de datos, diagramas de clases y otra documentación que puede jugar un papel central en la determinación de políticas de datos.

Con este espectro de información en mente, los resultados del estudio indican que no existen en el ecosistema de información políticas consistentes a nivel institucional que atiendan el levantamiento y documentación de metadatos o de integración de documentación técnica de los sistemas desarrollados para las instituciones públicas participantes.

Del total de sistemas detectados solamente el 11.5 por ciento (6) cuenta con toda la documentación antes señalada.¹⁷ Las características de estas plataformas reflejan una buena parte de los hallazgos hasta ahora observados en materia de gestión de datos. Estos sistemas pertenecen a solamente tres instituciones. No obstante, estas instituciones cuentan con más sistemas detectados en el ecosistema¹⁸ de los cuales ningún otro tiene toda la documentación de metadatos disponible.¹⁹ Más aún, la documentación de metadatos, según lo reportan estas instituciones para estos sistemas, no se hace conforme a estándares internacionales. Lo anterior, refleja la centralidad e impacto que tiene la definición de políticas de manejo de datos para la implementación homogénea de buenas prácticas.

Naturalmente, algunos sistemas levantan parcialmente la documentación a la que se hace referencia. Estos sistemas representan el 42.3 por ciento del total de sistemas detectados. Dentro de este subgrupo, los sistemas detectados cuentan con uno o dos documentos de metadatos conforme a los perfiles sobre los que se indagó. En este respecto, del total de sistemas detectados, el 23.1 por ciento cuenta con un tipo de documentación de metadatos mientras que 19.2 por ciento cuenta con dos de los tres tipos de documentación de metadatos.

Estos hallazgos en materia de documentación de las estructuras de datos reflejan la necesidad de implementar una política de alto nivel en materia de manejo de los repositorios de información. Asimismo, se debe subrayar la importancia de que esta política sea sujeta a estándares homogéneos que permitan el desarrollo armonioso del ecosistema de información.

RESULTADO POR ÁREA TEMÁTICA

Evolución patrimonial

En materia de evolución patrimonial, el grueso de los sistemas cuenta con un avance en materia de documentación de las estructuras de datos. De los 13 sistemas del área temática, el 69.2 por ciento de los sistemas detectados cuentan con 1 o 2 documentos relevantes para la gestión de las estructuras de datos. No obstante, aún con un criterio más laxo, sólo el 30.8 por ciento de estos sistemas cuenta con toda la documentación de metadatos sobre la que se solicitó información. Más aún, se debe subrayar que esta área

¹⁷ Para fines comparativos no se consideró la documentación alternativa sino como un elemento adicional de comparación toda vez que no se pudo observar un comportamiento homogéneo de documentación en este respecto.

¹⁸ La SFP detectó 12 sistemas, la CNDH detectó 4 sistemas y la SCJN detectó 7 sistemas relevantes para el estudio.

¹⁹ Para las bases de datos que se mantienen en Excel el criterio de totalidad de la documentación se redujo a la presencia/ausencia de diccionarios de datos. Esto afectó únicamente a un sistema (Procedimientos de contratación de la SCJN) al cual se le asignó una calificación de cumplir con la totalidad de la documentación.

temática no cuenta con ningún sistema que tenga toda la documentación indicada. Finalmente, el 30.8 por ciento de los sistemas en esta área no cuentan con ningún tipo de documentación.

Servidores públicos sancionados

En el área de servidores públicos sancionados, la mayoría de los sistemas no cuentan con ningún tipo de documentación de sus estructuras de datos. De los 12 sistemas en el área temática, el 58.3 por ciento no tiene ningún tipo de documentación. En contraste, el 16.7 por ciento de los sistemas ha generado toda la documentación sobre la que se indagó en el levantamiento mientras que el 25 por ciento ha generado parte de la documentación y al menos uno de los documentos atiende metadatos del sistema.

Contrataciones

En materia de contrataciones, el grueso de los sistemas no cuenta con documentación sobre sus estructuras de datos. De los 23 sistemas en el área temática, el 52.2 por ciento no cuenta con ningún tipo de documentación de su estructura de datos. Sin embargo, se observa algo de avance en esta materia. El 39.1 por ciento ha generado una parte de la documentación consultada. De entre este subgrupo, el 30.4 por ciento (7) generó toda la documentación de metadatos sobre la que se indagó. Por otro lado, el 8.7 por ciento generó toda la documentación consultada.

Sanciones a proveedores

En materia de sanción a proveedores, se tiene uno de los peores niveles de documentación de estructuras de datos. De los 11 sistemas en el área temática, el 63.6 por ciento de las unidades encargadas de los mismos reportó que no cuenta con ninguna documentación. En el 27.3 por ciento de los casos se reportó que el repositorio de información cuenta con algún tipo de documentación. De entre estos, el 27.3 por ciento generó toda la documentación de metadatos sobre la que se indagó. Finalmente, sólo el 9.1 por ciento de los sistemas cuenta con toda la documentación consultada.

Quejas

En materia de quejas y denuncias, se tiene el segundo peor desempeño en materia de documentación de las estructuras de datos. De los 13 sistemas en el área, el 61.5 por ciento de los sistemas no cuenta con ningún tipo de documentación. Del 38.5 por ciento que cuenta con documentación, 15.4 por ciento sólo tiene un documento de las estructuras de datos y estos documentos se han concentrado por completo en los metadatos del sistema. En contraste, el 23.1 por ciento de los sistemas generó el total de la documentación consultada.

HALLAZGOS CLAVE POR ÁREA TEMÁTICA

Evolución patrimonial



INFORMACIÓN DE EVOLUCIÓN PATRIMONIAL DE SERVIDORES PÚBLICOS

Los sistemas de información patrimonial sirven para cumplir lo establecido en materia de responsabilidades administrativas de los servidores públicos plasmado en la Ley General de Responsabilidades Administrativas (LGRA).

Los esquemas actuales de operación se desarrollan en tres fases que operan según calendarios de trabajo predeterminados.

1. Declaración inicial
2. Declaración de evolución patrimonial
3. Declaración de cierre

Cada una de estas fases busca capturar puntos clave de la evolución patrimonial de los servidores públicos que, en conjunto, permiten evaluar y diagnosticar si un servidor público presenta anomalías en su desarrollo patrimonial.

ACTORES DEL ENTORNO

En cumplimiento de la normatividad existente en la materia, actualmente se realizan tareas de captura, procesamiento y explotación de información relativa a la situación patrimonial de los servidores públicos en 12 instituciones públicas.

No.	Institución
1	Auditoría Superior de la Federación
2	Cámara de Diputados
3	Comisión Federal de Competencia Económica
4	Comisión Nacional de Derechos Humanos
5	Instituto Federal de Telecomunicaciones
6	Instituto Nacional de Estadística, Geografía e Informática
7	Instituto Nacional Electoral
8	Secretaría de la Función Pública
9	Suprema Corte de Justicia de la Nación
10	Banco de México*
11	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales
12	Instituto Nacional para la Evaluación de la Educación

*Se indica la información del Banco de México, aunque su sistema no fue considerado para fines de análisis

SISTEMAS DETECTADOS

La información patrimonial de los servidores públicos se gestiona dentro de 15 sistemas electrónicos y bases de datos. Estos sistemas presentaron una estructura operativa y de información con muchas características comunes. El motivo detrás de esto es que el levantamiento de información normalmente se desarrolla con el fin de poder evaluar la evolución patrimonial de las personas contratadas por las instituciones públicas para desarrollar funciones públicas.

No obstante, la información no sólo se recopila con este motivo. La información de esta área temática tiene múltiples puntos de interés. Este tipo de datos también se encuentran en sistemas de contrataciones, transparencia y de administración financiera.

No.	Sistemas de evolución patrimonial	Institución responsable
1	DeclaraINE	Instituto Nacional Electoral
2	DeclaraINEE	Instituto Nacional para la Evaluación de la Educación
3	DeclaraNet Plus	Instituto Federal de Telecomunicaciones
4	DeclaraNet Plus	Secretaría de la Función Pública
5	DeclaraNet, COFECE	Comisión Federal de Competencia Económica
6	DeclaraNet, INAI	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales
7	DeclaraUEC	Cámara de Diputados / ASF
8	DeclarINEGI	Instituto Nacional de Estadística, Geografía e Informática
9	Declaración Patrimonial Electrónico (DEPTEL)	Banco de México*
10	Plataforma Nacional de Transparencia	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales
11	Portal de Contrataciones	Banco de México
12	Registro y Control de las Relaciones Patrimoniales	Comisión Nacional de Derechos Humanos
13	Sistema de Declaración Patrimonial	Suprema Corte de Justicia de la Nación
14	Sistema de Registro y Control Patrimonial	Cámara de Diputados
15	Sistema Integral Financiero	Instituto Nacional para la Evaluación de la Educación

*Se indica la información del Banco de México, aunque su sistema no fue considerado para fines de análisis.

SISTEMA DOMINANTE

¿Qué sistema domina esta área temática? DeclaraNet

El listado de sistemas con información de esta área temática muestra cierta diversidad, aunque siendo herederos de un código fuente común, los sistemas de este entorno son, en su mayoría, sistemas réplica del Sistema DeclaraNet. Destacan 7 diferentes versiones de este sistema que fueron implementadas mediante un acuerdo interinstitucional que permitió la distribución de su código fuente.

Esta experiencia común entre las instituciones ha permitido una homogeneización importante han generado características positivas para este ecosistema de información, a saber:

1. Estructuras de datos compartidas
2. Alineación a cambios en estructuras de datos
3. Reducción de costos en el ajuste institucional a nuevos requerimientos de información
4. Creación de canales de comunicación informales entre responsables de información

INFORMACIÓN DE SERVIDORES PÚBLICOS SANCIONADOS

Los repositorios de información con datos sobre servidores públicos sancionados cumplen el propósito de registrar a funcionarios públicos que hayan sido sujetos a un proceso de responsabilidades administrativas y cuenten con fallos en firme en su contra.

Estos repositorios, en general, operan bajo dos perfiles:

1. Sistemas de gestión y/o registro del proceso de investigación y seguimiento a resultados de procesos de responsabilidades administrativas
2. Listados de servidores públicos sancionados (publicados en sitios web o resguardados para consumo institucional)

Cada uno de estos perfiles cumple una función específica acorde a las necesidades operativas y de gestión de la institución administradora del repositorio de datos.

ACTORES DEL ENTORNO

Actualmente, se realizan tareas de captura, procesamiento y explotación de información relativa a la información de servidores públicos sancionados en 11 instituciones públicas.

No.	Institución
1	Banco de México*
2	Comisión Federal de Competencia Económica
3	Comisión Nacional de Derechos Humanos
4	Instituto Federal de Telecomunicaciones
5	Instituto Nacional de Estadística, Geografía e Informática
6	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales
7	Instituto Nacional Electoral
8	Secretaría de la Función Pública
9	Suprema Corte de Justicia de la Nación
10	Tribunal Electoral del Poder Judicial de la Federación
11	Tribunal Federal de Justicia Administrativa

*Se indica la información del Banco de México, aunque su sistema no fue considerado para fines de análisis.

SISTEMAS DETECTADOS

La información sobre servidores públicos sancionados se maneja en 13 sistemas electrónicos y bases de datos. Esta área temática presenta una amplia variación en términos del tipo de información almacenada, diferencia explicada en buena medida por las divergencias entre los objetivos de estos repositorios.

Los repositorios de información están especializados en su área temática. No obstante, 30.7 por ciento de los sistemas en el área temática también contienen información sobre quejas y denuncias. Estos sistemas generalmente son registros de los procesos de investigación sobre responsabilidades administrativas que, en algunos casos, resultan en sanciones a funcionarios públicos.

No.	Nombre del sistema	Institución
1	Omisos y extemporáneos en la presentación de la declaración de situación patrimonial (OMEXT)	Secretaría de la Función Pública
2	Plataforma Nacional de Transparencia	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales
3	Registro de Procedimientos Administrativos de Responsabilidades	Comisión Federal de Competencia Económica
4	Registro de Procedimientos Administrativos Insaturados y Sanciones	Banco de México*
5	Registro de Servidores Públicos Sancionados (RESPS)	Instituto Nacional de Estadística, Geografía e Informática
6	Registro de Servidores Públicos Sancionados	Instituto Federal de Telecomunicaciones
7	Registro de Servidores Públicos Sancionados	Suprema Corte de Justicia de la Nación
8	Registro de Servidores Públicos Sancionados	Secretaría de la Función Pública
9	Sancionados con inhabilitación firme	Comisión Nacional de Derechos Humanos
10	Servidores Públicos Sancionados	Tribunal Federal de Justicia Administrativa
11	Sistema de Información de Quejas, Denuncias y Responsabilidades	Instituto Nacional Electoral
12	Sistema de registro de servidoras/es públicas/os con sanciones	Tribunal Electoral del Poder Judicial de la Federación
13	Sistema Integral de Administración de Recursos Financieros	Comisión Nacional de Derechos Humanos

*Se indica la información del Banco de México, aunque su sistema no fue considerado para fines de análisis.

SISTEMA DOMINANTE

¿Qué sistema domina esta área temática? Registro de servidores públicos sancionados, SFP

La información de servidores públicos sancionados es celosamente guardada y operada con motivo de su vinculación con los procesos de investigación y sanción a servidores públicos. Esta misma conexión impacta en el tipo de información almacenada y el perfil que se les da a estos registros. Comúnmente, estos registros son utilizados como listados que contienen información descriptiva de los expedientes a los que están asociados. Por tal motivo, las estructuras se alinean con piezas clave de información de los procesos legales.

Se debe subrayar que, según se comentó en varias de las entrevistas con actores clave vinculados a estos sistemas, se busca que estos sistemas se alineen a los campos de información contenidos en los registros de la SFP. Esto es resultado, en parte, de que el listado de Función Pública representa un referente para todas las instituciones al ser el listado público más consultado en la materia.

El sustento legal común de estos repositorios de datos ha permitido una relativa homogeneidad en el tipo de información recuperada. Como resultado, se observan algunas dinámicas positivas para este ecosistema de información.

1. Campos clave de información compartidos
2. Cultura profesional compartida
3. Normatividad común que rige procesos y dinámicas operativas

Sin embargo, esta área temática presenta niveles de integración muy dispares entre los flujos de trabajo a los que está sujeta la información y los repositorios de información. Esto es, con frecuencia se observó que existe una separación funcional entre las fuentes de información y sus gestores. Esto a veces resulta en una falta de capacidad para mantener la información actualizada, situación que podría tener un impacto importante de frente a procesos de interoperación o interconexión de la información.

Por tal motivo, antes de la integración se deben refinar los procesos de actualización de información y uniformar medidas para asegurar su vigencia. Toda vez que esta información tiene el potencial de generar repercusiones legales es importante que estos riesgos sean mitigados de manera significativa antes de una integración sustantiva de datos.

“Nosotros buscamos ajustarnos a los campos de Función Pública, así podemos comparar su información con la nuestra.”

INFORMACIÓN DE CONTRATACIONES Y COMPRAS PÚBLICAS

Los repositorios de datos de compras y contrataciones públicas tienen como objetivo registrar la información de las operaciones de compra venta de los sujetos de la ley de adquisiciones, arrendamientos y servicios del sector público, así como la ley general de transparencia y acceso a la información pública.

Esta información puede llegar a abarcar registros sobre todas las fases de las operaciones de compra, venta y renta de bienes y servicios por parte de las instituciones públicas. Por tal motivo, es común que se utilicen sistemas complejos que operan mediante módulos de un flujo de trabajo.

Los esquemas actuales de operación se desarrollan en múltiples fases que se ajustan a los calendarios de trabajo de cada proceso de compra/venta. Ejemplos de estas fases son: estudio de mercado; lanzamiento de convocatoria de licitación; atención a dudas; recepción de propuestas; evaluación de propuestas, y asignación de licitación

Cada una de estas fases requiere conjuntos de información con funciones puntuales en el contexto de los procesos definidos para estas compras.

ACTORES DEL ENTORNO

Actualmente, se realizan tareas de captura, procesamiento y explotación de información relativa a las compras y contrataciones públicas en 12 instituciones públicas.

No.	Institución
1	Auditoría Superior de la Federación*
2	Banco de México*
3	Cámara de Diputados
4	Comisión Federal de Competencia Económica
5	Comisión Nacional de Derechos Humanos
6	Instituto Nacional de Estadística, Geografía e Informática
7	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales
8	Instituto Nacional Electoral
9	Instituto Nacional para la Evaluación de la Educación
10	Secretaría de la Función Pública
11	Suprema Corte de Justicia de la Nación
12	Tribunal Electoral del Poder Judicial de la Federación

*Se indica la información del Banco de México, aunque su sistema no fue considerado para fines de análisis

SISTEMAS DETECTADOS

La información relativa a las contrataciones se gestiona dentro de 25 sistemas electrónicos y bases de datos. Los repositorios de información con datos relativos al área temática cumplen una amplia variedad de objetivos que incluyen contrataciones, transparencia y de administración financiera.

No.	Nombre del sistema	Institución
1	Registro Único de Proveedores	Auditoría Superior de la Federación
2	Portal de Contrataciones	Banco de México
3	Portal de Adquisiciones del Congreso	Cámara de Diputados
4	Registro de Inconformidades	Comisión Federal de Competencia Económica
5	Registro de quejas y/o denuncias	Comisión Federal de Competencia Económica
6	Registro de Sanción a Proveedores	Comisión Federal de Competencia Económica
7	Sistema de Control de Almacén y Activo Fijo, CNDH	Comisión Nacional de Derechos Humanos
8	Sistema de Gestión de Adquisiciones, CNDH	Comisión Nacional de Derechos Humanos
9	Véndale al INEGI	Instituto Nacional de Estadística, Geografía e Informática
10	Plataforma Nacional de Transparencia	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales
11	Base de datos de proveedores sancionados	Instituto Nacional Electoral
12	CompralNE	Instituto Nacional Electoral
13	Sistema Integral de Gestión Administrativa (SIGA) – Módulo de compras (PO) y Módulo de Tesorería (XTR)	Instituto Nacional Electoral
14	Registro de licitantes y proveedores sancionados	Instituto Nacional para la Evaluación de la Educación
15	Sistema Integral Financiero	Instituto Nacional para la Evaluación de la Educación
16	CompraNet – im Módulo de Información e Inteligencia de Mercado para las Contrataciones Públicas	Secretaría de la Función Pública
17	CompraNet como plataforma transaccional	Secretaría de la Función Pública
18	CompraNet –hc Módulo para la integración y consulta del historial en materia de contrataciones y su cumplimiento	Secretaría de la Función Pública
19	Programas Anuales de Adquisiciones, Arrendamientos, Servicios y Obras Públicas (PAAASOP)	Secretaría de la Función Pública

No.	Nombre del sistema	Institución
20	Sistema de Procedimiento Administrativo de Sanción a Proveedores y Contratistas	Secretaría de la Función Pública
21	Padrón de Proveedores y Contratistas	Suprema Corte de Justicia de la Nación
22	Procedimientos Contratación- Padrón Proveedores y Contratistas (SIPOT)	Suprema Corte de Justicia de la Nación
23	Procedimientos de Contratación	Suprema Corte de Justicia de la Nación
24	Registro de quejas o denuncias	Suprema Corte de Justicia de la Nación
25	Sistema de Registro de Proveedores Sancionados	Tribunal Electoral del Poder Judicial de la Federación

*Se indica su información, aunque el sistema no fue considerado para fines de análisis con motivo de una captura parcial de los datos.

SISTEMA DOMINANTE

¿Qué sistema domina esta área temática? CompraNet

La información sobre contrataciones y adquisiciones de bienes y servicios muestra un nivel de maduración significativo en el grueso de las instituciones consideradas en el estudio. No obstante, existen casos extremos en esta área temática en los que aún se opera mediante registros físicos (i.e. Senado de la República). Esto podría representar un reto hacia el proceso de homogeneización y procesamiento de información.

El hecho de que un sistema funcione como el estándar de uso para la mayoría de las instituciones presenta una buena oportunidad para la homogeneización del tratamiento y procesamiento de información.

La integración en este tema presenta un avance significativo toda vez que el Sistema CompraNet es utilizado de manera común por las instituciones consideradas para el estudio. No obstante, entre las instituciones que no forman parte del esquema de CompraNet aún persisten retos significativos en materia de captura, procesamiento y explotación de la información.

Por tal motivo, es posible que instituciones que no participan en el esquema de CompraNet presenten resistencias con motivo de los costos de integrarse al esquema de CompraNet y la inversión realizada en el desarrollo de herramientas alternativas para el manejo de información en esta materia.

INFORMACIÓN DE PROVEEDORES DE BIENES Y SERVICIOS SANCIONADOS

Los repositorios de información que atienden el área temática de proveedores de bienes y servicios sancionados representan registros de los particulares con sanciones en firme como resultado de un proceso administrativo. Estos sistemas se dividen en dos subtipos:

1. Sistemas de control de expedientes de averiguaciones sobre posibles fallas administrativas por parte de particulares
2. Listados de proveedores con resoluciones sancionatorias en firme

Estos tipos de sistemas generalmente representan dos partes de un todo en el cual el primer tipo nutre las entradas registradas en el segundo.

ACTORES DEL ENTORNO

Actualmente, se realizan tareas de captura, procesamiento y explotación de información relativa a la situación patrimonial de los servidores públicos en 8 instituciones públicas.

No.	Institución
1	Comisión Federal de Competencia Económica
2	Comisión Nacional de Derechos Humanos
3	Instituto Nacional Electoral
4	Instituto Nacional para la Evaluación de la Educación
5	Secretaría de la Función Pública
6	Suprema Corte de Justicia de la Nación
7	Tribunal Electoral del Poder Judicial de la Federación
8	Banco de México

*Se indica el Banco de México como una institución que realiza tareas de captura, procesamiento y explotación relativa a la situación patrimonial ya que se reportó la existencia de un sistema, sin embargo, como se mencionó con anterioridad, su sistema no fue considerado para fines de análisis debido a que no se contó con la información completa al cierre de este informe.

SISTEMAS DETECTADOS

La información en esta área temática se preserva dentro de 11 sistemas electrónicos y bases de datos. Estos sistemas presentan modelos comunes. Sin embargo, se observan diferencias significativas en materia operativa y de gestión de la información. En el área temática es común que estos repositorios de información se integren con los sistemas de contrataciones. Este acoplamiento de información se observó en 81.8 por ciento de los repositorios en el área temática.

No.	Sistema	Institución
1	Base de datos de proveedores sancionados	Instituto Nacional Electoral
2	Directorio de proveedores y Contratistas Sancionados	Secretaría de la Función Pública
3	Programas Anuales de Adquisiciones, Arrendamientos, Servicios y Obras Públicas (PAAASOP)	Secretaría de la Función Pública
4	Registro de impedimentos para contratar	Suprema Corte de Justicia de la Nación
5	Registro de licitantes y proveedores sancionados	Instituto Nacional para la Evaluación de la Educación
6	Registro de Sanción a Proveedores	Comisión Federal de Competencia Económica
7	Sistema de Control de Almacén y Activo Fijo	Comisión Nacional de Derechos Humanos
8	Sistema de Procedimiento Administrativo de Sanción a Proveedores y Contratistas	Secretaría de la Función Pública
9	Sistema de Registro de Proveedores Sancionados	Tribunal Electoral del Poder Judicial de la Federación
10	Sistema Integral de Gestión Administrativa (SIGA) – Módulo de compras (PO) y Módulo de Tesorería (XTR)	Instituto Nacional Electoral
11	Sistema Integral Financiero	Instituto Nacional para la Evaluación de la Educación

SISTEMA DOMINANTE

¿Qué sistema domina esta área temática? No hay sistema dominante

El área temática de proveedores de bienes y servicios sancionados no cuenta con un claro liderazgo como el que se observa en otras áreas. Dentro de este subsistema se observó una amplia gama de perspectivas sobre la operación y gestión de la información relevante al área.

Vale la pena observar que, en entrevista con actores clave de las instituciones participantes, se observó con frecuencia que el listado público de la SFP se utiliza como un insumo durante los procesos de contratación toda vez que este listado concentra información de las dependencias, entidades de la Administración Pública Federal y de los Gobiernos de los Estados. No obstante, esto se hace sólo como instrumento de consulta y no representa un esquema que guíe la operación de la misma manera que CompraNet lo hace en su área temática.

INFORMACIÓN DE QUEJAS Y DENUNCIAS

Los sistemas y bases de datos que contienen información sobre quejas y denuncias públicas de faltas administrativas y de hechos de corrupción están orientados a la documentación de los reportes que serán sujetos a consideración de los órganos internos de control para averiguación sobre posibles faltas administrativas o de algún otro tipo.

Estos conjuntos de datos generalmente funcionan como buzones para la recepción de denuncias del público que podrían tener una implicación en materia de faltas administrativas o hechos de corrupción. No obstante, como lo señalan los responsables de las instituciones participantes, comúnmente se utiliza como una ventanilla que además de estas funciones sirve como ruteador de preguntas e inconformidades.

ACTORES DEL ENTORNO

Actualmente, se realizan tareas de captura, procesamiento y explotación de información relativa a la información de quejas y denuncias en 8 instituciones.

No.	Institución
1	Banco de México*
2	Comisión Federal de Competencia Económica
3	Comisión Nacional de Derechos Humanos
4	Instituto Nacional de Estadística, Geografía e Informática
5	Instituto Nacional para la Evaluación de la Educación
6	Secretaría de la Función Pública
7	Suprema Corte de Justicia de la Nación
8	Comisión Federal de Competencia Económica

*Se indica la información del Banco de México, aunque su sistema no fue considerado para fines de análisis

SISTEMAS DETECTADOS

La información sobre quejas y denuncias esta contenida en 14 sistemas electrónicos y bases de datos. Es interesante observar que estos repositorios de información generalmente están dedicados a más de un objetivo operativo.

No.	Sistema	Institución
1	Registro de denuncias	Banco de México*
2	Registro de Procedimientos Administrativos de Responsabilidades	Comisión Federal de Competencia Económica
3	Registro de Inconformidades	Comisión Federal de Competencia Económica
4	Registro de quejas y/o denuncias	Comisión Federal de Competencia Económica
5	Sancionados con inhabilitación firme	Comisión Nacional de Derechos Humanos
6	SIDEQ Módulo Quejas y Denuncias	Instituto Nacional de Estadística, Geografía e Informática
7	Buzón electrónico de quejas y denuncias	Instituto Nacional para la Evaluación de la Educación
8	Sistema de Procedimientos Administrativos de Responsabilidades	Secretaría de la Función Pública
9	Sistema Integral de Inconformidades	Secretaría de la Función Pública
10	Sistema Integral de Quejas y Denuncias Ciudadanas (SIDEQ)	Secretaría de la Función Pública
11	Registro de Servidores Públicos Sancionados	Secretaría de la Función Pública
12	Registro de Servidores Públicos Sancionados	Suprema Corte de Justicia de la Nación
13	Registro de quejas o denuncias	Suprema Corte de Justicia de la Nación
14	Registro de Sanción a Proveedores	Comisión Federal de Competencia Económica

*Se indica la información del Banco de México, aunque su sistema no fue considerado para fines de análisis

SISTEMA DOMINANTE

¿Qué sistema domina esta área temática? No hay un sistema dominante.

La información de quejas y denuncias contenida en las plataformas de información comúnmente se asocia a otras piezas de información. El 57.1 por ciento de los sistemas en el área temática están dedicados a compilar información de más de un área temática. En este respecto, se observan 2 tendencias de integración de información.

1. Vinculación con información de servidores públicos sancionados, y
2. Vinculación con adquisiciones de bienes y servicios.

Durante el proceso de levantamiento se observó que la variabilidad de la información capturada como parte de los procesos de denuncia o comunicación de quejas resultan en bases de datos que requieren una gran cantidad de depuración. Lo anterior, con motivo de que estos sistemas básicamente permiten la captura de campos abiertos de información en los que el público puede capturar sus inquietudes. Este contexto provee una diferencia operativa que los distingue de otras áreas temáticas.

Es importante que se hagan esfuerzos para que los actores clave en el área temática comiencen a compartir y aprovechar las experiencias en materia de sistematización y tratamiento de datos sobre quejas y denuncias con el fin de maximizar el valor de la experiencia en la materia.

DIAGNÓSTICO DEL PANORAMA DE LA GESTIÓN DE DATOS



DIAGNÓSTICO DEL PANORAMA DE LA GESTIÓN DE DATOS

Con el objetivo de sintetizar la complejidad de la realidad operativa y de gestión de los repositorios de información detectados como parte del proyecto se construyeron índices de desempeño en materia de gestión de datos para las áreas temáticas clave del ecosistema de información de la LGSNA. Estos índices permiten una mirada homogénea a la realidad estudiada y permiten evaluar una multiplicidad de dimensiones de manera comparativa.

¿Qué medimos?

Los índices desarrollados atienden 3 ejes temáticos claves relacionados íntimamente con la gestión de datos. Estos son:

1. Completitud temática
2. Complementariedad en la gestión de datos
3. Confiabilidad de la gestión de datos

Estos índices facilitan el diagnóstico del nivel de maduración y cobertura de cada sistema con relación a las múltiples áreas temáticas definidas por la Ley. Originalmente, se propuso el diseño de un índice que evaluara directamente en las estructuras de datos el desarrollo, amplitud, homogeneidad y convergencia de los campos de datos en cada uno de los registros de información.

No obstante, como resultado de un acceso limitado a esta información²⁰ se procedió a construir una medida basada en proxys de información que permitiera desarrollar un diagnóstico de en qué medida las condiciones que rodean a la información se prestan para el desarrollo de actividades de interconexión, interoperabilidad o transferencia de sistemas.

¿Qué se busca con estas estimaciones?²¹

En conjunto, el análisis de estas áreas temáticas busca determinar a qué distancia se encuentran los sistemas evaluados de una línea base que permita que cualquier esfuerzo de integración, adopción o interconexión se desarrolle de una manera fácil y ordenada.

²⁰ Durante el proceso de levantamiento se solicitaron los diccionarios de datos y otra documentación relevante a las instituciones participantes. No obstante, sólo se recibió esta información para 3 sistemas, a saber, Plataforma Nacional de Transparencia, DeclaraNet (versión implementada por el INAI) y DeclaraUEC.

²¹ Durante el cálculo de los índices señalados en este apartado las respuestas correspondientes a No sabe, No contesto y No fue posible obtener la información obtuvieron una calificación de 0. Esto se hizo con motivo de que se interpretó el registro de respuestas como un indicador proxy de la capacidad institucional para obtener información de los sistemas.

ÍNDICE DE COMPLETITUD TEMÁTICA

El índice de completitud busca proveer una medida de la cantidad de áreas temáticas en que cada institución participante detectó por lo menos un sistema electrónico o base de datos abocado a almacenar información de dicha área. Como resultado de lo anterior, la detección de al menos un sistema dentro de un área temática otorga un valor de una para esa institución. La cobertura se evaluó como porcentaje del total de las 6 áreas temáticas analizadas en el estudio.

Esta medida de cobertura temática permite ubicar los espacios temáticos que no han sido atendidos por las instituciones participantes o en los que no se cuenta con acciones de levantamiento de información que cumplan las condiciones básicas requeridas para un tratamiento y explotación de información acordes a la realidad informática de la actualidad.²¹

SE LEVANTA LO ESTRICTAMENTE NECESARIO Y SE DESARROLLAN LAS ÁREAS CON MAYOR ESCRUTINIO PÚBLICO

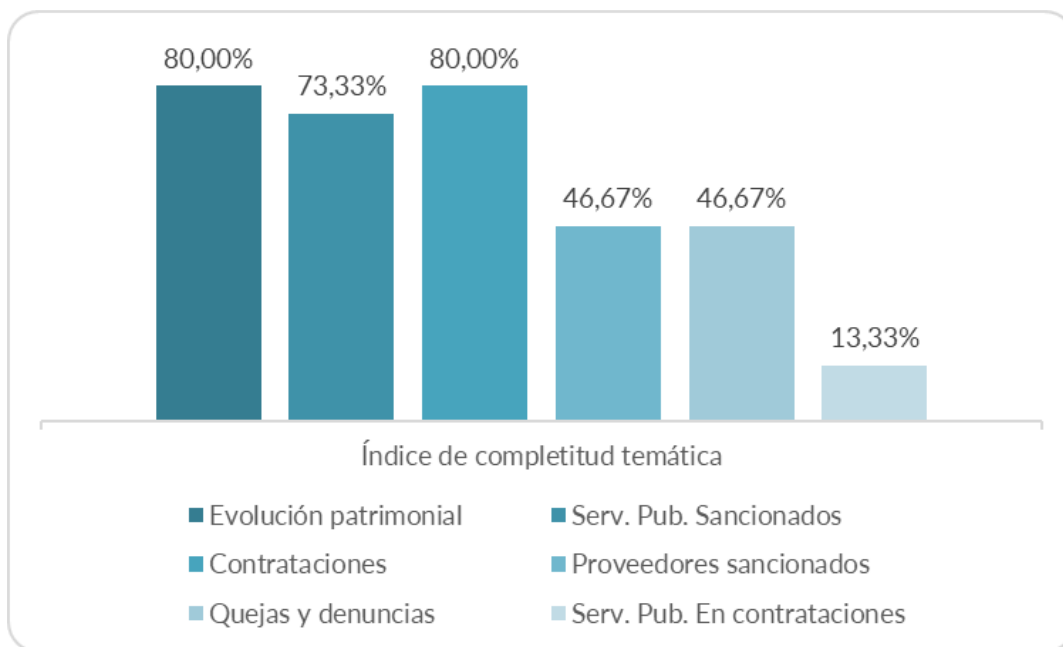
LA COBERTURA DE INFORMACIÓN MUESTRA UNA CONCENTRACIÓN EN ÁREAS DE INFORMACIÓN MÁS CONTROVERTIDAS EN LA ARENA PÚBLICA MIENTRAS QUE LAS AGENDAS DE LEVANTAMIENTO, TRATAMIENTO, EXPLOTACIÓN Y ELIMINACIÓN DE INFORMACIÓN EN ÁREAS MÁS TÉCNICAS Y DE CONFLICTOS INTERNOS NO SE DESARROLLAN O NO SE CUBREN

¿Qué información se recupera?

El ecosistema de información muestra una concentración no homogénea de sistemas electrónicos y bases de datos a través de las áreas temáticas fijadas por la LGSNA. Los datos recabados muestran una división entre la masa de sistemas considerados en el estudio. Por un lado, la información en materia de evolución patrimonial de servidores públicos, servidores públicos sancionados e información de contrataciones es recolectada, tratada y explotada en más del 50 por ciento de las instituciones en formatos legibles por máquinas.

Por otro lado, la información relativa a proveedores sancionados, denuncias públicas de faltas administrativas y hechos de corrupción, así como la información de servidores públicos que intervienen en procedimientos de contrataciones públicas sólo se recolecta, procesa y explota en menos del 50 por ciento de las instituciones estudiadas en formatos legibles por máquinas.

²² El listado de las valoraciones utilizadas para cada área temática en este índice se puede consultar en el Anexo V al final del presente documento.



Durante el periodo de entrevistas se observó que es común que las instituciones recuperen información en todas o la mayoría de estas áreas temáticas. Por este motivo, es clave no confundir las cifras presentadas como huecos absolutos de información. Por el contrario, el levantamiento en sitio reflejó que comúnmente se cuenta con registros locales en formatos físicos o en algún formato electrónico no legible por máquinas (i.e. procesador de textos, PDF, digitalizaciones de formatos físicos, imágenes).

Sin embargo, toda vez que la información detectada no se traduce posteriormente en formatos electrónicos apropiados, esta información no cuenta con un tratamiento lo suficientemente maduro como para ser parte activa de cualquier esfuerzo de transferencia, integración o interconexión de información con un actor externo.

PERSISTEN VACÍOS DE INFORMACIÓN CONCENTRADOS EN INSTITUCIONES QUE NO CUENTAN CON UN ESQUEMA DESARROLLADO DE LEVANTAMIENTO, TRATAMIENTO, EXPLOTACIÓN Y ELIMINACIÓN DE REPOSITORIOS DE DATOS.

¿Las instituciones recogen toda la información requerida?

Las instituciones evaluadas tienen grandes diferencias en la cantidad y temáticas de la información que recuperan y gestionan como parte de su operación cotidiana. Estas diferencias no parecen estar determinadas por factores estructurales del arreglo institucional del país ni por sus atribuciones normativas (constitucionales y legales o de cualquier otro tipo).

Las instituciones participantes tuvieron, en promedio, un 56.67 por ciento de la completitud temática respecto de lo que marca la LGSNA. No obstante, el que una tercera parte de las instituciones analizadas cuente con menos del 50 por ciento de cobertura temática sugiere que aún queda bastante trabajo pendiente en materia de construcción de capacidades organizacionales, técnicas y tecnológicas.

En esta métrica, sólo la Secretaría de la Función Pública contó con el 100 por ciento de información sistematizada en las temáticas señaladas por la ley antes mencionada. En buena medida, esto se debe a la posición que ocupa aquella Secretaría en el poder ejecutivo federal como punto neural de las políticas de levantamiento de información en este ámbito. En contraste, otras instituciones únicamente recuperan información de manera interna. Es posible que esta diferencia haya tenido un impacto en la maduración de los repositorios de información.

Institución	Índice completitud institucional
Senado de la República	0.00%
Tribunal Federal de Justicia Administrativa	16.67%
Auditoría Superior de la Federación*	33.33%
Cámara de Diputados	33.33%
Instituto Federal de Telecomunicaciones	33.33%
Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	50.00%
Tribunal Electoral del Poder Judicial de la Federación	50.00%
Instituto Nacional de Estadística, Geografía e Informática	66.67%
Instituto Nacional Electoral	66.67%
Instituto Nacional para la Evaluación de la Educación	66.67%
Banco de México**	83.33%
Comisión Federal de Competencia Económica	83.33%
Comisión Nacional de Derechos Humanos	83.33%
Suprema Corte de Justicia de la Nación	83.33%
Secretaría de la Función Pública	100.00%

* Para fines de claridad, es importante señalar que uno de los repositorios de información asociado a la Auditoría Superior de la Federación (ASF) se opera por parte de la Cámara de Diputados. No obstante, se decidió hacer la asociación toda vez que toda la información en ese sistema es sobre personal de la ASF y que esto se hace con motivo del arreglo institucional mexicano y no por limitantes o deficiencias en la institución. Este modelo de operación descentralizado en una tercera parte fue una excepción única a la operación convencional y auto contenida que llevaban las instituciones evaluadas en el levantamiento. Como tal, sólo parece representar un modelo de organización alternativa desde el aspecto normativo y de gestión que podría ser interesante para aquellas instituciones que no cuentan con capacidades propias para desarrollar esfuerzos en alguna de las áreas temáticas definidas por la Ley.

** A pesar de que la información enviada no fue confirmada en fases reiteradas de levantamiento de datos, se consideró la información enviada por el Banco de México estrictamente para fines de diagnóstico de la cobertura temática de los sistemas existentes.

Es importante señalar que, como se verá más adelante en las secciones de análisis del entorno por perfil temático, el hecho de que una institución cuente con sistemas en un área temática no implica que actualmente cuenta con todas las condiciones requeridas para desarrollar transferencias, interconexiones o proceder a un esquema de interoperabilidad con plataformas de su misma área temática. La presencia de sistemas sólo implica que actualmente la institución cuenta con acciones implementadas para levantar, procesar y, en ocasiones, explotar la información alrededor de un área temática.

No obstante, el índice de completitud temática no refleja la concentración de sistemas por área temática.

¿Con qué frecuencia se duplican esfuerzos en una misma área temática?

La revisión a profundidad de la concentración de sistemas y bases de datos por área temática muestra que en algunas de ellas existen traslapes en materia de levantamiento y procesamiento de información. En algunos casos, estos traslapes implican una interconexión de sistemas. Ejemplo de ello es el uso de una base de datos común que aprovechan los múltiples módulos del sistema CompraNet. Sin embargo, hay casos como el de la Comisión Federal de Competencia Económica en que el conjunto de información sobre materia de quejas y denuncias se divide en múltiples bases de datos.

Como parte de la agenda del proceso de maduración del ecosistema de información creado por la LGSNA se debe de buscar mitigar este fraccionamiento operativo de la información. Actualmente, existen circunstancias que dificultan este avance homogéneo. Es el caso en materia de disponibilidad presupuestaria y capacidad técnica para el desarrollo de sistemas electrónicos para la gestión de información a través de su ciclo de vida.

No obstante, la definición del ecosistema de información de la LGSNA provee una oportunidad para que todos los actores participantes se beneficien de esfuerzos existentes y futuros. Esto se debe a que la nueva Ley define claramente un conjunto de instituciones que se beneficiarían significativamente de la cooperación en esta materia. Lo anterior, podría apoyar esfuerzos para la creación o robustecimiento de la infraestructura tecnológica disponible de manera coordinada. De la misma manera, se puede fortalecer el capital humano técnico en el entorno de manera conjunta a fin de maximizar su impacto y asegurar que estos repositorios de información sean atendidos por personal con la experiencia y capacidad necesarios.

El mapa de calor muestra los niveles de concentración de sistemas con injerencia en las múltiples áreas temáticas definidas por la Ley. En este se puede observar que sólo 8 instituciones cuentan con sistemas con traslapes temáticos. En materia temática, destaca que el área de contrataciones tiene la mayor cantidad de instituciones con esfuerzos duplicados. En esta área 6 instituciones tienen repositorios de información que coinciden en el manejo de información de áreas temáticas

Institución	Evolución pa- trimonial	Serv. Pub. San- cionados	Contrataciones	Proveedores sancionados	Quejas y denuncias	Serv. Pub. En contrataciones
Auditoría Superior de la Federación	1	0	1	0	0	0
Banco de México	2	1	1	0	1	1
Cámara de Diputados	1	0	1	0	0	0
Comisión Federal de Competencia Económica	1	1	3	1	3	0
Comisión Nacional de Derechos Humanos	1	2	2	1	1	0
Instituto Federal de Telecomunicaciones	1	1	0	0	0	0
Instituto Nacional de Estadística, Geografía e Informática	1	1	1	0	1	0
Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	2	1	1	0	0	0
Instituto Nacional Electoral	1	1	3	2	0	0
Instituto Nacional para la Evaluación de la Educación	2	0	2	2	1	0
Secretaría de la Función Pública	1	2	3	3	4	1
Senado de la República	0	0	0	0	0	0
Suprema Corte de Justicia de la Nación	1	1	3	1	2	0
Tribunal Electoral del Poder Judicial de la Federación	0	1	1	1	0	0
Tribunal Federal de Justicia Administrativa	0	1	0	0	0	0

ÍNDICE DE COMPLEMENTARIEDAD EN LA GESTIÓN DE DATOS

La falta de políticas que normen la gestión de datos ha facilitado una gran variación en las características de los esfuerzos de sistematización de la información en cada una de las áreas temáticas que marca la LGSNA. Como resultado de ello, actualmente hay avances muy poco uniformes en materia de gestión de datos, elemento fundamental para evaluar la complementariedad de la información.

Por tal motivo, ya que no se puede realizar un diagnóstico directo de la complementariedad de los datos, se realizó un análisis basado únicamente en la detección de la existencia las políticas de gestión de datos que tienen un impacto en las condiciones para la implementación de acciones de transferencia, interconexión o interoperabilidad. Por tal motivo, se habla de complementariedad en la gestión de datos.

Esta medida de complementariedad de gestión de datos, marca en una escala de 0 a 100 la medida en que cada sistema cumple criterios básicos que, conforme a las mejores prácticas a nivel internacional, están relacionados con una apropiada gestión de datos. Asimismo, estas buenas prácticas son insumos valiosos para la implementación de medidas de transferencia, interconexión o interoperabilidad con otros sistemas electrónicos y bases de datos.

Motivos para reflexionar: Imagine el siguiente escenario. El encargado del sistema NotNet se enferma gravemente de manera inesperada y tiene que abandonar su puesto de trabajo. Su reemplazo recibe el sistema sin documentación de metadatos. Para operar, el nuevo funcionario se apoya del personal del área y logra comprender en alguna medida la estructura del sistema con unas pocas cosas que no quedan claras y que, con motivo de sus responsabilidades, no tiene tiempo de aclarar a profundidad. Luego, a resultas de cambios en personal de alto nivel, hay una rotación sustantiva de personal. El sistema se delega a un nuevo equipo que recibe la plataforma, aún sin documentación. Como parte de nuevas políticas el sistema se transfiere a una nueva área para lo cual se pide la generación de documentación. Ahora, el nuevo equipo enfrenta un reto imposible, explicar los errores y defectos “bugs” no documentados del equipo inicial, las anomalías que no pudo explicar el segundo funcionario y las dudas adicionales del nuevo equipo que recibió la plataforma. El hueco de conocimiento se sigue profundizando, hartos de los problemas que generan todos estos “inexplicables” alguien del nuevo equipo le dice a su jefe: “¿y si rehacemos el sistema desde cero?”. Lo que una buena documentación pudo haber evitado. Ser el último equipo es cuestión de suerte, pero no ser el primero es cuestión de buenas políticas de gestión de datos.

El índice de complementariedad informática considera 2 dimensiones relevantes.

1. Marco Normativo
2. Documentación del sistema

El análisis de estas dimensiones se apoyó en 9 variables que representan de manera aproximada información valiosa para el cálculo de cada una de las dimensiones señaladas.

El índice de complementariedad de datos es una estimación sintética y de la implementación de criterios básicos de gestión de datos vinculados a buenas prácticas en el entorno internacional.

NINGUNO DE LOS SISTEMAS SE ENCUENTRA EN EL LUGAR IDEAL Y POCO MENOS DEL 20 POR CIENTO DE LOS SISTEMAS DETECTADOS CONTÓ AL MENOS UNA CALIFICACIÓN DE 75 PUNTOS.

¿Qué tan bien se gestionan los datos en el ecosistema de información?

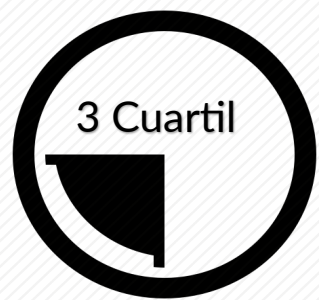
Los procesos de gestión de datos en la mayoría de las instituciones son poco homogéneos, poco robustos y no están adecuadamente regulados.

RESULTADOS DEL ÍNDICE DE COMPLEMENTARIEDAD

Los resultados de los repositorios de información dentro del estudio cubrieron un espectro de valores entre 12.5 y 75 puntos. En adelante, se presentan los resultados de cada repositorio de información según estos se posicionaron en los cuartiles de calificaciones posibles asignadas por el índice de complementariedad. conforme a la división del desempeño por sistema conforme a los cuartiles de valores posibles. Se inicia en el tercer cuartil toda vez que ningún sistema alcanzó los valores del cuartil más alto de la escala.

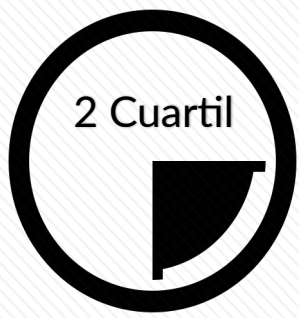


No.	Sistema	Institución	Índice
1	Sistema de Control de Almacén y Activo Fijo	Comisión Nacional de Derechos Humanos	75
2	Sistema Integral de Administración de Recursos Financieros	Comisión Nacional de Derechos Humanos	75
3	CompraNet como plataforma transaccional	Secretaría de la Función Pública	75
4	Sistema de Procedimientos Administrativos de Responsabilidades	Secretaría de la Función Pública	75
5	Sistema Integral Financiero	Instituto Nacional para la Evaluación de la Educación	75
6	Plataforma Nacional de Transparencia	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	75
7	Véndale al INEGI	Instituto Nacional de Estadística, Geografía e Informática	75
8	DeclaraNet	Comisión Federal de Competencia Económica	75
9	Sistema Integral de Quejas y Denuncias Ciudadanas (SIDECC)	Secretaría de la Función Pública	75
10	Registro de Servidores Públicos Sancionados	Secretaría de la Función Pública	75

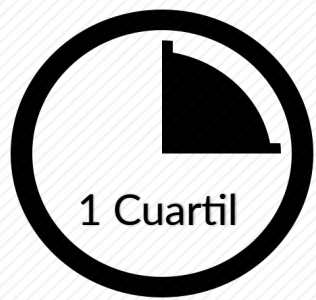


No.	Sistema	Institución	Índice
11	CompraNet -hc Módulo para la integración y consulta del historial en materia de contrataciones y su cumplimiento	Secretaría de la Función Pública	68.75
12	Programas Anuales de Adquisiciones, Arrendamientos, Servicios y Obras Públicas (PAAASOP)	Secretaría de la Función Pública	68.75
13	CompraNet - im Módulo de Información e Inteligencia de Mercado para las Contrataciones Públicas	Secretaría de la Función Pública	68.75
14	CompralNE	Instituto Nacional Electoral	62.5
15	DeclaraUEC	Cámara de Diputados / ASF	62.5
16	Sistema de Registro y Control Patrimonial	Cámara de Diputados	62.5
17	DeclaralNEE	Instituto Nacional para la Evaluación de la Educación	62.5
18	Sistema de Información de Quejas, Denuncias y Responsabilidades	Instituto Nacional Electoral	62.5
19	Sistema Integral de Gestión Administrativa (SIGA) - Módulo de compras (PO) y Módulo de Tesorería (XTR)	Instituto Nacional Electoral	56.25
20	Sistema de Declaración Patrimonial	Suprema Corte de Justicia de la Nación	56.25
21	Sistema de Procedimiento Administrativo de Sanción a Proveedores y Contratistas	Secretaría de la Función Pública	56.25

Como se puede observar el alto desempeño en el índice no está constreñido a una institución o a una misma institución o información de un área temática. Sin embargo, si se puede observar una presencia significativa de la Secretaría de la Función Pública a través de los sistemas dentro del segundo cuartil más alto posible de evaluaciones.



No.	Sistema	Institución	Índice
22	DeclaralNE	Instituto Nacional Electoral	50
23	Directorio de proveedores y Contratistas Sancionados, SFP	Secretaría de la Función Pública	50
24	DeclaraNet, INAI	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	50
25	Registro de Servidores Públicos Sancionados (RESPS), INEGI	Instituto Nacional de Estadística, Geografía e Informática	50
26	DeclarINEGI	Instituto Nacional de Estadística, Geografía e Informática	50
27	Registro de licitantes y proveedores sancionados, INEE	Instituto Nacional para la Evaluación de la Educación	50
28	Buzón electrónico de quejas y denuncias, INEE	Instituto Nacional para la Evaluación de la Educación	50
29	Procedimientos de Contratación, SCJN	Suprema Corte de Justicia de la Nación	43.75
30	Padrón de Proveedores y Contratistas, SCJN	Suprema Corte de Justicia de la Nación	43.75
31	DeclaraNet Plus, SFP	Secretaría de la Función Pública	37.5
32	Sistema de registro de servidoras/es públicas/os con sanciones, TEPJF	Tribunal Electoral del Poder Judicial de la Federación	37.5
33	Sistema de Registro de Proveedores Sancionados, TEPJF	Tribunal Electoral del Poder Judicial de la Federación	37.5
34	Portal de Adquisiciones del Congreso, Diputados	Cámara de Diputados	37.5
35	SIDEQ Módulo Quejas y Denuncias, INEGI	Instituto Nacional de Estadística, Geografía e Informática	37.5
36	Registro de Procedimientos Administrativos de Responsabilidades, COFECE	Comisión Federal de Competencia Económica	31.25
37	Registro de quejas y/o denuncias, COFECE	Comisión Federal de Competencia Económica	31.25



No.	Sistema	Institución	Índice
38	Servidores Públicos Sancionados	Tribunal Federal de Justicia Administrativa	25
39	Registro de impedimentos para contratar	Suprema Corte de Justicia de la Nación	25
40	Registro de Inconformidades	Comisión Federal de Competencia Económica	25
41	Omisos y extemporáneos en la presentación de la declaración de situación patrimonial (OMEXT)	Secretaría de la Función Pública	25
42	Registro de Servidores Públicos Sancionados	Suprema Corte de Justicia de la Nación	25
43	Registro de Sanción a Proveedores	Comisión Federal de Competencia Económica	25
44	Registro y Control de las Relaciones Patrimoniales	Comisión Nacional de Derechos Humanos	25
45	Sistema de Gestión de Adquisiciones	Comisión Nacional de Derechos Humanos	25
46	DeclaraNet Plus	Instituto Federal de Telecomunicaciones	25
47	Procedimientos Contratación- Padrón Proveedores y Contratistas (SIPOT)	Suprema Corte de Justicia de la Nación	25
48	Sistema Integral de Inconformidades	Secretaría de la Función Pública	25
49	Registro de quejas o denuncias	Suprema Corte de Justicia de la Nación	12.5
50	Registro de Servidores Públicos Sancionados	Instituto Federal de Telecomunicaciones	12.5
51	Sancionados con inhabilitación firme	Comisión Nacional de Derechos Humanos	12.5
52	Base de datos de proveedores sancionados	Instituto Nacional Electoral	12.5

En el universo de sistemas detectados, el promedio de desempeño en el índice de complementariedad en la gestión de datos fue de 47.1 puntos. Esto implica que, en promedio, la mayor parte de los sistemas no están sujetos a la mitad de las mejores prácticas en materia de gestión de datos.

El desempeño de los sistemas mostró una gran variación a través del ecosistema de información. El índice de complementariedad de gestión de datos tiene una desviación estándar de 0.204. Lo anterior, implica que los sistemas obtuvieron, en promedio, 20.4 puntos en el índice alrededor de la media del universo de sistemas analizado. Esta variación resultó en un rango que abarca de los 12.5 puntos que obtuvieron los cuatro sistemas con peor desempeño hasta los 75 puntos que lograron los 10 sistemas con mejores resultados. En efecto, estos datos indican una inconsistencia en la implementación de buenas prácticas.

LA INCONSISTENCIA EN LA IMPLEMENTACIÓN DE BUENAS PRÁCTICAS DE GESTIÓN DE DATOS PREVALECE AÚN AL INTERIOR DE INSTITUCIONES.

¿Es un problema de instituciones o de políticas?

A menudo se asume que una institución con resultados exitosos es una institución con buenas prácticas y políticas apropiadas. Sin embargo, como se ha sugerido previamente, en materia de gestión de datos esto es equivocado. Con frecuencia se presentan diferencias importantes en materia de recursos presupuestarios, capital humano, experiencia, y otras dimensiones relevantes entre las unidades funcionales de una organización. Por tal motivo, en ausencia de políticas claras y robustas de gestión de datos, los resultados de distintos proyectos de levantamiento, tratamiento y procesamiento de información llegan a tener variaciones importantes en materia de la implementación de buenas prácticas en la gestión de datos.

Los resultados del índice de complementariedad sugieren que aún persisten diferencias importantes al interior de algunas instituciones en el ecosistema de información de la LGSNA. Esto se refleja en la desviación estándar en los resultados del índice de complementariedad estimado a nivel de cada institución.

Institución	Desv. Std. Índice de complementariedad
Auditoría Superior de la Federación	N/A
Cámara de Diputados	12.5
Comisión Federal de Competencia Económica	19
Comisión Nacional de Derechos Humanos	26.9
Instituto Federal de Telecomunicaciones	6.3
Instituto Nacional de Estadística, Geografía e Informática	13.6
Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	12.5
Instituto Nacional Electoral	18.7
Instituto Nacional para la Evaluación de la Educación	10.4
Secretaría de la Función Pública	18.6
Suprema Corte de Justicia de la Nación	14.1
Tribunal Electoral del Poder Judicial de la Federación	0
Tribunal Federal de Justicia Administrativa	N/A

Los datos muestran que al interior de las instituciones existe una desviación estándar promedio de 13.8 puntos en el resultado del índice. Esto apoya la posibilidad de que aún se requiere un mayor esfuerzo para el desarrollo de comportamientos homogéneos al interior de las instituciones evaluadas en relación a su gestión de datos.

La poca uniformidad en materia de gestión de datos al interior de las instituciones ha resultado en un rango en la desviación estándar entre los sistemas a nivel institucional que abarca de 6.3 puntos hasta 26.9 puntos. Llama la atención en este respecto que las tres instituciones con la desviación estándar más alta en las muestras a nivel institucional tengan más de un sistema que alcanzó la calificación más alta observada en el índice de complementariedad en la gestión de datos (75 puntos).

Esta presencia inconsistente a nivel institucional refuerza lo señalado en la sección de buenas prácticas de la necesidad de establecer políticas de alto nivel y marcos de responsabilidades y rendición de cuentas a nivel institucional. Por su parte, la inconsistencia entre sistemas subraya la necesidad de que estos sean implementados por líderes de alto nivel dentro de la institución a fin de que puedan afectar de manera transversal a las distintas áreas y unidades de la organización sin consideración de su área de especialización.

Durante el levantamiento en sitio se observó que hay ciertas dinámicas y fenómenos recurrentes que amplifican los fenómenos observados en este respecto.

- **Licenciamiento de sistemas.** Se detectaron múltiples instancias en las que una vez que las instituciones firmaron acuerdos de licenciamiento no se solicitaron, en conjunto con el código fuente, los documentos de cada plataforma. Asimismo, en los casos en que el sistema fuente no contaba con esta documentación las instituciones receptoras no procedieron a generar esta documentación de manera unilateral.
- **Acuerdos con proveedores de servicios que limitan la obtención de documentación relevante de los sistemas.** Durante el levantamiento en sitio se comentó que en algunos casos en que los sistemas se proveen como servicios por parte de un proveedor privado los contratos contienen estipulaciones que limitan la documentación con que pueden contar como cliente del servicio. Se sugiere la regulación de este tipo de prácticas a fin de que siempre se cuente con información de metadatos sobre los repositorios de información. Esto contribuirá a la reducción, prevención y control de los riesgos asociados a las estructuras de información de las unidades del sector público. Ejemplo de ello es el riesgo que asumen las instituciones públicas y el efecto que tienen estos acuerdos sobre su capacidad de respuesta en escenarios de incumplimiento de contrato.

- Límites de recursos en los equipos que administran los sistemas. En algunos casos, los equipos que administran las plataformas cuentan con recursos materiales y humanos muy limitados. Desde la perspectiva de algunos actores, esto, en conjunto con las prioridades institucionales, reduce el valor de actividades de documentación y formalización documental de los repositorios de información frente al resto de sus actividades sustantivas.

“Si tenemos que hacerlo pues se hace, pero ahorita tenemos mucha chamba y lo podemos hacer después”

- Falta de priorización de la documentación. Durante el proceso de levantamiento se observó que no existe una percepción de prioridad asociada a la generación de documentación y formalización de protocolos vinculados a los repositorios de información.

“Si me lo piden lo hago, pero nadie me lo ha pedido porque no se ha necesitado.”

Todo lo anterior indica la necesidad de la implementación proactiva de esquemas de mejora de las prácticas en materia de gestión de datos. No obstante, es importante señalar que, a raíz de la definición de un ecosistema de información mediante la publicación de la LGSNA, estas políticas deben aplicarse de manera uniforme a todos los actores. Adicionalmente, estas políticas deben trazar rutas de maduración para los sistemas en diferentes niveles de desarrollo en materia de gestión de datos. De otra manera, las políticas podrían definir un estándar al cual cada actor buscaría llegar de manera poco estructurada y con resultados poco uniformes.

ÍNDICE DE CONFIABILIDAD DE LA GESTIÓN DE DATOS

La información, a diferencia del arte, sólo es valiosa si refleja fielmente la realidad de aquello a lo que hace referencia. Por este motivo, la evaluación de la veracidad y validez de los datos contenidos en las bases de datos es una actividad central del manejo apropiado de cualquier conjunto de datos, más aún si estos son utilizados para la toma de decisiones públicas o el desarrollo de procesos jurídicos.

En el cumplimiento de sus funciones, es común que un funcionario público tome la información contenida en las bases de datos oficiales como una representación fiel de la realidad. Es poco común que se cuestionen los datos en estos repositorios de información en ausencia de un detonador que obligue a hacerlo. Sin embargo, ¿qué tan confiables son?

La ausencia de una política de datos que integre una perspectiva de protocolos de aseguramientos de la información ha resultado en un desarrollo incipiente de procesos formales de aseguramiento de la calidad de datos en las bases de datos del ecosistema de la LGSNA. Existen algunos avances en este respecto, sin embargo, estos han demostrado ser poco uniformes y muy limitados.

Un diagnóstico de la confiabilidad de los datos en los sistemas electrónicos y bases de datos requiere un contraste con documentación de respaldo para cada uno de los repositorios de datos detectados. No obstante, este análisis está más allá del alcance del presente estudio y representa una prioridad secundaria a la detección de medidas y protocolos que aseguran que la veracidad y validez de los datos sea resultado de procesos sistemáticos y no de la buena fe de las fuentes de datos. Por tal motivo, este análisis se enfoca en la confiabilidad de la gestión de datos y no en la confiabilidad de los datos en sí mismos.

La evaluación de confiabilidad de la gestión de datos está dirigida a determinar la robustez de las acciones implementadas en materia normativa, operativa y de gestión de información para asegurar que la información contenida en los sistemas electrónicos y bases de datos sea válida y veraz en el corto, mediano y largo plazos. En una escala del 0 al 100, el índice de confiabilidad de datos provee un estimador de la medida en que los administradores de los sistemas evaluados han implementado protocolos para determinar y proteger la integridad de los datos a través de su ciclo de vida. Esto implica la implementación de medidas y acciones en 4 dimensiones.

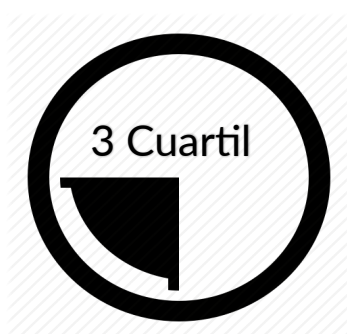
1. Planeación;
2. Normatividad regulatoria;
3. Protocolos de seguridad, y
4. Medidas de evaluación de la veracidad y la validez de la información.

El diagnóstico del avance en estas dimensiones se hizo con el respaldo de 20 reactivos que permiten una mirada dentro de estos repositorios de datos.

NINGUNO DE LOS SISTEMAS SE ENCUENTRA EN EL LUGAR IDEAL Y POCO MENOS DEL 20 POR CIENTO DE LOS SISTEMAS DETECTADOS CONTÓ AL MENOS UNA CALIFICACIÓN DE 75 EN LA EVALUACIÓN DE LAS CONDICIONES BASE CONSIDERADAS PARA EL ÍNDICE.

¿Qué penetración tienen protocolos de protección y validación de la información en los sistemas detectados?

Dentro del ecosistema de información definido por la LGSNA se puede observar que aún hay mucho por hacer con relación a la implementación de políticas de protección de la integridad y validación de datos. Los resultados observados, así como la evaluación en sitio de los protocolos de protección de datos sugieren que no se cuenta con marcos normativos robustos que aseguren la validez y veracidad de la información ni medidas integradas que protejan la integridad de los datos a través del tiempo.



No.	Sistema	Institución	Índice
1	CompraINE	Instituto Nacional Electoral	75
2	DeclararINE	Instituto Nacional Electoral	70.5
3	Sistema Integral de Gestión Administrativa (SIGA) – Módulo de compras (PO) y Módulo de Tesorería (XTR), INE	Instituto Nacional Electoral	65
4	Sistema de Control de Almacén y Activo Fijo, CNDH	Comisión Nacional de Derechos Humanos	62
5	Sistema Integral de Administración de Recursos Financieros, CNDH	Comisión Nacional de Derechos Humanos	62
6	Procedimientos de Contratación, SCJN	Suprema Corte de Justicia de la Nación	59
7	CompraNet como plataforma transaccional, SFP	Secretaría de la Función Pública	58.5
8	Sistema de Procedimientos Administrativos de Responsabilidades, SFP	Secretaría de la Función Pública	55

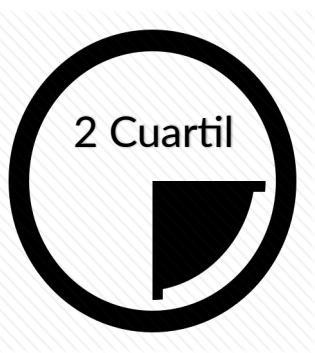
Los resultados del índice de confiabilidad muestran que ningún repositorio de información cuenta con todos los criterios para asegurar la confiabilidad de la información. Más aún, ningún sistema electrónico o base de datos cuenta con los requisitos para rankear en el cuartil más alto de puntuación del índice de confiabilidad en la gestión de datos.

De manera consistente con los hallazgos hasta ahora, las instituciones que cuentan con repositorios de información en el tercer cuartil de la escala del índice no han logrado homogeneizar estos resultados en materia de confiabilidad en la gestión de datos a través de todos los sistemas electrónicos y bases de datos.



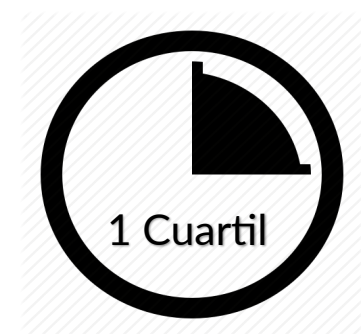
2 Cuartil

No.	Sistema	Institución	Índice
9	CompraNet –hc Módulo para la integración y consulta del historial en materia de contrataciones y su cumplimiento, SFP	Secretaría de la Función Pública	50
10	Programas Anuales de Adquisiciones, Arrendamientos, Servicios y Obras Públicas (PAAASOP), SFP	Secretaría de la Función Pública	49
11	Sistema Integral Financiero, INEE	Instituto Nacional para la Evaluación de la Educación	48
12	Plataforma Nacional de Transparencia, INAI	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	47.5
13	Padrón de Proveedores y Contratistas, SCJN	Suprema Corte de Justicia de la Nación	45.5
14	Servidores Públicos Sancionados, TFJA	Tribunal Federal de Justicia Administrativa	45
15	Véndale al INEGI	Instituto Nacional de Estadística, Geografía e Informática	44
16	DeclaraNet Plus, SFP	Secretaría de la Función Pública	42.5
17	DeclaraUEC, Diputados	Cámara de Diputados / ASF	37.5
18	Sistema de Registro y Control Patrimonial, ASF	Cámara de Diputados	37.5
19	DeclaraNet, COFECE	Comisión Federal de Competencia Económica	35
20	Sistema de Declaración Patrimonial, SCJN	Suprema Corte de Justicia de la Nación	34
21	Sistema de registro de servidoras/es públicas/os con sanciones, TEPJF	Tribunal Electoral del Poder Judicial de la Federación	34
22	Sistema Integral de Quejas y Denuncias Ciudadanas (SIDECA), SFP	Secretaría de la Función Pública	33
23	Registro de Procedimientos Administrativos de Responsabilidades, COFECE	Comisión Federal de Competencia Económica	33
24	Registro de quejas y/o denuncias, COFECE	Comisión Federal de Competencia Económica	33
25	Directorio de proveedores y Contratistas Sancionados, SFP	Secretaría de la Función Pública	32.5

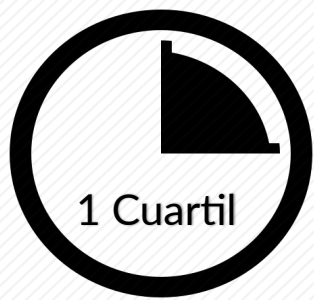


No.	Sistema	Institución	Índice
26	Registro de Servidores Públicos Sancionados	Secretaría de la Función Pública	30
27	DeclaraNet	Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	30
28	Registro de impedimentos para contratar	Suprema Corte de Justicia de la Nación	30
29	Registro de Inconformidades	Comisión Federal de Competencia Económica	30
30	Registro de Servidores Públicos Sancionados (RESPS)	Instituto Nacional de Estadística, Geografía e Informática	27.5
31	Omisos y extemporáneos en la presentación de la declaración de situación patrimonial (OMEXT)	Secretaría de la Función Pública	27.5

Finalmente, 21 repositorios de información se encuentran en el último cuartil de la escala del índice de confiabilidad en la gestión de datos. Es importante que estos sistemas electrónicos y bases de datos tomen medidas en el corto plazo para ajustar estas deficiencias y asegurar la integridad de su información.



No.	Sistema	Institución	Índice
32	DeclarINEGI	Instituto Nacional de Estadística, Geografía e Informática	25
33	Registro de Servidores Públicos Sancionados, SCJN	Suprema Corte de Justicia de la Nación	25
34	Registro de Sanción a Proveedores, COFECE	Comisión Federal de Competencia Económica	23
35	Registro de quejas o denuncias, SCJN	Suprema Corte de Justicia de la Nación	23
36	Registro de Servidores Públicos Sancionados, IFT	Instituto Federal de Telecomunicaciones	23
37	Sistema de Procedimiento Administrativo de Sanción a Proveedores y Contratistas, SFP	Secretaría de la Función Pública	22.5
38	Sistema de Registro de Proveedores Sancionados, TEPJF	Tribunal Electoral del Poder Judicial de la Federación	22.5
39	Registro de licitantes y proveedores sancionados, INEE	Instituto Nacional para la Evaluación de la Educación	20



No.	Sistema	Institución	Índice
40	Portal de Adquisiciones del Congreso, Diputados	Cámara de Diputados	20
41	Registro y Control de las Relaciones Patrimoniales, CNDH	Comisión Nacional de Derechos Humanos	20
42	DeclaralINEE	Instituto Nacional para la Evaluación de la Educación	19.5
43	SIDEQ Módulo Quejas y Denuncias, INEGI	Instituto Nacional de Estadística, Geografía e Informática	17.5
44	Buzón electrónico de quejas y denuncias, INEE	Instituto Nacional para la Evaluación de la Educación	15
45	Sistema de Gestión de Adquisiciones, CNDH	Comisión Nacional de Derechos Humanos	12.5
46	Sancionados con inhabilitación firme, CNDH	Comisión Nacional de Derechos Humanos	12.5
47	Sistema de Información de Quejas, Denuncias y Responsabilidades, INE	Instituto Nacional Electoral	10
48	Declaranet Plus, IFT	Instituto Federal de Telecomunicaciones	10
49	Procedimientos Contratación- Padrón Proveedores y Contratistas (SIPOT), SCJN	Suprema Corte de Justicia de la Nación	10
50	CompraNet - im Módulo de Información e Inteligencia de Mercado para las Contrataciones Públicas, SFP	Secretaría de la Función Pública	9
51	Base de datos de proveedores sancionados, INE	Instituto Nacional Electoral	7.5
52	Sistema Integral de Inconformidades, SFP	Secretaría de la Función Pública	5

Dentro del conjunto de los sistemas detectados como parte del levantamiento, el promedio de desempeño en el índice de confiabilidad de la gestión de datos fue de 32.9 puntos. De este se extrae que, en promedio, hay un avance marginal en materia de la implementación de mejores prácticas para una apropiada protección de la integridad y validez de los datos.

No obstante, durante el levantamiento fue muy claro para el equipo de investigación que existen variaciones importantes en materia de prácticas de protección de la integridad y validación de datos a través del ecosistema de información. Las variaciones en el índice de confiabilidad confirman esta observación. En el universo de sistemas detectados, el índice de confiabilidad tuvo una desviación estándar de 17.3. Lo anterior implica que los sistemas obtuvieron, en promedio, 17.3 puntos en el índice por debajo o por encima de la media en un rango de resultados que abarcó desde el puntaje más bajo de 5 hasta el resultado más alto de 75 puntos. Esta información respalda lo antes señalado en materia de inconsistencias en la implementación de políticas apropiadas de gestión de datos.

LAS MEDIDAS DE ASEGURAMIENTO DE LA CONFIABILIDAD DE LA INFORMACIÓN SE VEN AFECTADAS POR LAS DIFERENCIAS ENTRE SILOS FUNCIONALES DENTRO DE LAS INSTITUCIONES EVALUADAS.

¿Es un problema de instituciones o de políticas?

Durante el levantamiento de información con las instituciones participantes se observó que algunas instituciones cuentan con políticas complejas y desarrolladas de manejo y protección de la infraestructura tecnológica en la que se resguardan los datos. Entre otras cosas, estas políticas, alineadas con las líneas de acción de la Estrategia Digital Nacional, buscan construir una política de TICs para la administración pública. No obstante, como lo reflejan los esfuerzos a nivel internacional, una política de tecnologías de la información no es una política de datos.

La falta de políticas en materia de gestión de datos ha permitido resultados muy poco uniformes en materia de confiabilidad de la gestión de datos a través de los equipos de trabajo encargados de gestionar la información. Esto se refleja en la desviación estándar para cada una de las instituciones evaluadas. Si las políticas de aseguramiento y protección de integridad y validación de los datos almacenados en los repositorios de información analizados siguieran algún patrón común al interior de las organizaciones evaluadas, entonces, la desviación de cada uno de los sistemas respecto de su evaluación promedio en el índice tendería a ser cero. No obstante, la desviación estándar dentro de las instituciones es positiva para todas las organizaciones participantes. Más aún, el rango de esta desviación llega a ser de hasta 30 puntos en el caso del Instituto Nacional Electoral.

Institución	Desv. Std. Índice de confiabilidad
Auditoría Superior de la Federación	N/A
Cámara de Diputados	8.8
Comisión Federal de Competencia Económica	4.2
Comisión Nacional de Derechos Humanos	23.2
Instituto Federal de Telecomunicaciones	6.5
Instituto Nacional de Estadística, Geografía e Informática	9.7
Instituto Nacional de Transparencia, Acceso a la información y Protección de Datos Personales	8.8

Institución	Desv. Std. Índice de confiabilidad
Instituto Nacional Electoral	30.3
Instituto Nacional para la Evaluación de la Educación	13.1
Secretaría de la Función Pública	16.5
Suprema Corte de Justicia de la Nación	14.8
Tribunal Electoral del Poder Judicial de la Federación	5.7
Tribunal Federal de Justicia Administrativa	N/A

Estas diferencias al interior de las instituciones son resultado de una ausencia de políticas de alto nivel que determinen procesos y regulen los comportamientos en materia del aseguramiento de la confiabilidad de la información a través de su ciclo de vida. Lo anterior, abre una ventana de oportunidad para la ejecución *ad hoc* de protocolos de aseguramiento de la confiabilidad de los datos. Resultado cuya magnitud se puede observar en la tabla de desviación estándar a nivel institucional del índice de confiabilidad de la gestión de datos.

Uno de los resultados que llaman la atención es que instituciones con algunos de los sistemas mejor evaluados en el índice de confiabilidad tengan una desviación estándar tan alta a nivel institucional. En el entorno internacional, es común observar este tipo de fenómenos a causa de la definición de silos funcionales dentro de las organizaciones. Esto es, los grupos operativos de una institución se dividen conforme a áreas temáticas de especialización y no se cuenta con los mecanismos para asegurar que compartan lecciones de buenas prácticas.

Para evitar que este aislamiento funcional dentro de una organización impacte la gestión de datos, es importante que se integren lineamientos de operación nutridos con la experiencia de los equipos dentro de la organización. Esto permite rescatar las valiosas lecciones adquiridas y estructurar prácticas y dinámicas comunes que aseguren uniformidad en los resultados base. Es importante señalar que en este caso el desarrollo estará también interconectado a un ecosistema de información donde coexisten múltiples instituciones. Por tal motivo, es importante que se busque que haya un piso base de políticas para todas las instituciones sujetas a la LGSNA a fin de que se desarrolle el ecosistema de manera homogénea y con una misma dirección.

INTERRELACIÓN ENTRE ÍNDICES DE GESTIÓN DE DATOS

EXISTE UNA RELACIÓN POSITIVA ENTRE LOS ÍNDICES DE GESTIÓN DE DATOS.

¿Qué relación guardan los desempeños entre los índices de gestión de datos?

Hasta ahora se han presentado los resultados de los índices de complementariedad y confiabilidad como indicadores de la necesidad de mejorar la gestión de datos en el ecosistema de información. Por tanto, uno de los supuestos no explícitos de este trabajo es que las buenas prácticas en materia de complementariedad de gestión de datos mantienen una correlación positiva con las mejores prácticas en materia de confiabilidad en la gestión de datos.

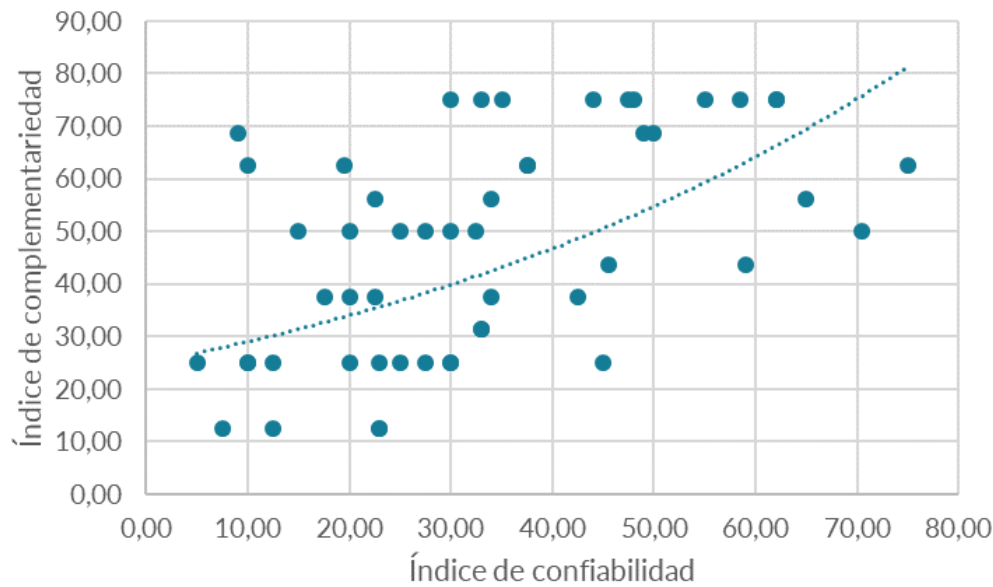
La base de este supuesto es la idea de que un equipo consciente de la importancia de implementar buenas prácticas en materia de documentación de las estructuras de datos buscará con mayor consistencia crear un esquema robusto de medidas para proteger y asegurar la confiabilidad de la información almacenada.

Sin embargo, este supuesto no debe aceptarse sin controversia. Actualmente, una buena parte de las instituciones operan sin documentación y consumen datos de buena fe asumiendo que los datos son veraces y válidos. Todo ello a pesar de que no se cuenta con mecanismos ni procesos que aseguren estas cualidades. No obstante, durante el levantamiento en sitio, fue común escuchar que no se han tenido problemas derivados de la información contenida en las bases de datos. “Nunca se ha tenido necesidad”, “no han tocado problemas”, “nunca nos ha pasado” o “ha pasado una vez nada más”.

Con motivo de lo anterior, es importante evaluar con seriedad la idea de que la situación actual es una en la que si bien los sistemas existentes no cuentan con mucha documentación de las estructuras de datos los datos son seguros, veraces y válidos. Una manera de realizar esta consideración es evaluar la ocurrencia conjunta o correlación entre indicadores que reflejen las gradaciones de esta ocurrencia. Es decir, conforme a estos indicadores ¿qué tan a menudo, un sistema con poca documentación tiene buenas prácticas para proteger la integridad de la información y viceversa?

La información disponible para el ecosistema de información de la LGSNA permite realizar esta evaluación con el apoyo de los índices de complementariedad y confiabilidad en la gestión de datos. Para tal fin, se calculó el coeficiente de correlación de Pearson que evalúa en una escala de -1 a +1 la medida en que dos variables están correlacionadas dentro de un conjunto de datos. En esta escala dos datos que no tienen ninguna relación obtendrían una valoración de 0 mientras que dos datos perfectamente correlacionados (positiva o negativamente) tendrían una calificación de -1 o +1 según la dirección de esta relación (positiva si se incrementan juntos o negativa si uno se incrementa cuando el otro se reduce).

Los resultados computados resultan en un coeficiente de Pearson de 0.531. Esto implica que existe una relación positiva entre estas variables. Sin embargo, aún hay otros factores ambientales que afectan la implementación de mejores prácticas en una u otra de las dimensiones evaluadas. Este resultado sugiere que es más común que un sistema que cuenta con mejores prácticas de documentación de las estructuras de datos cuente con un esquema más robusto de prácticas que aseguren la confiabilidad de los datos de un repositorio de información.



La dispersión de los datos de los índices de complementariedad de la gestión de datos y de confiabilidad en la gestión de datos representa gráficamente esta relación. Como se puede ver en el cruce de información de los índices, los sistemas y bases de datos con un resultado pobre en el índice de complementariedad tienen por lo general un resultado pobre en el índice de confiabilidad en la gestión de datos (cuadrante inferior izquierdo del gráfico).

Pero esta no es toda la historia. Es cierto que la distribución de datos sugiere que es poco frecuente observar repositorios de información con estructuras de datos ampliamente documentadas que cuenten con políticas de gestión de datos que aseguren la confiabilidad de la información. Sin embargo, el gran número de repositorios de datos que tienen buenos resultados en el índice de complementariedad y que tienen un desempeño pobre en el índice de confiabilidad sugiere que la documentación de las estructuras de datos es sólo un primer paso.

Esta información subraya la importancia de que cualquier política de gestión de datos que se quiera aplicar debe considerar todos los aspectos clave de la gestión de datos, a saber, diseño, planeación, levantamiento, tratamiento, procesamiento, almacenamiento y eliminación de repositorios de datos.

SECCIÓN IV. RECOMENDACIONES



RECOMENDACIONES

El principal hallazgo del estudio apunta a un avance poco homogéneo en la implementación de políticas de datos consistentes y comprensivas a nivel institucional para todas las instituciones evaluadas y a través de las áreas temáticas definidas por la LGSNA.

Sin embargo, lo verdaderamente valioso de un hallazgo de esta naturaleza no descansa en la observación de las deficiencias sino en la definición de respuestas apropiadas y relevantes; respuestas que deberán robustecer el ecosistema de información, madurándolo y empujándolo hacia una dirección de orden y empoderamiento de los actores clave que operan dentro del mismo.

El solaz de un mundo interconectado y globalizado es la existencia y disponibilidad de conocimiento resiliente y robusto. Esto es, la capacidad de consultar prácticas y procesos que han sido probados en una gran variedad de contextos y cuyos resultados se pueden evaluar. El presente análisis busca contribuir a la construcción de conocimiento en la materia que le atañe.

En el entorno internacional, la conversación sobre la gestión de datos ha avanzado significativamente. Conforme la información se transforma en uno de los activos de mayor valor que puede tener una institución, la gestión segura, confiable y organizada de los datos se ha tornado en una conversación sobre la administración del patrimonio de una organización. Por tal motivo, es clave que el ecosistema de información de la LGSNA se nutra de las mejores prácticas en lo que se ha señalado como la gestión integral de datos.

¿Qué se debe construir?

La experiencia internacional muestra claramente la ventaja que provee la gestión de datos desde una perspectiva integral y con base en principios uniformes y generalizados. Estos principios conforman una serie de buenas prácticas hacia la que nuestra propia política en materia de datos debería dirigirse, definiéndose una hoja de ruta que permita la construcción y vinculación de fuentes de información en esta materia.

Políticas con miras al largo plazo

Como parte de esta nueva hoja de ruta se debe establecer cómo redefinir las actuales políticas y protocolos de levantamiento, procesamiento y explotación de datos para construir un marco normativo sólido que permita potenciar las capacidades de detección, investigación y sanción de comportamientos ilícitos.

Al mismo tiempo, es preciso reconocer que un enfoque centrado exclusivamente en el control y la fiscalización es insuficiente. La nueva arquitectura de datos y de sistemas de información debe también generar una cultura de cumplimiento, de deontología del servicio público que desemboque en instituciones y servidores públicos con mayor integridad. Un mecanismo autorregulado, digamos, que genere los incentivos para la honestidad y la responsabilidad en el desempeño de los cargos públicos.

Redefiniendo el lugar de los datos en el gobierno

Tomando su valor con una perspectiva más amplia, el lugar de los datos en la administración pública debe ser, también, uno de empoderamiento y simplificación de las atribuciones del servidor público en lo que concierne a sus responsabilidades. Esta nueva perspectiva adoptada en el sector público debe reconocer que la transparencia y la rendición de cuentas han sido motores fundamentales en la conversación sobre los múltiples aspectos del ciclo de vida de los datos. No obstante, también debe permitir transitar a un diálogo en el que los datos sean vistos como una herramienta poderosa para mejorar la eficiencia, robustecer la calidad y promover la innovación de los servidores públicos en el desempeño de sus labores y en sus capacidades para brindar valor adicional a la política pública de manera transversal y de largo plazo.

Alcances de una nueva perspectiva de la gestión de datos

Toda política de datos bien concebida con un respaldo analítico robusto debe asegurar que existan políticas y procedimientos apropiadas para todos los puntos del ciclo de vida de la información.

Esta perspectiva debe comenzar a mostrar que las políticas de tratamiento de datos no aplican únicamente a sistemas electrónicos complejos, sino que también son relevantes para los registros internos de información a fin de proteger a las instituciones de eventos comunes como:

1. Rotación no anticipada de personal clave;
2. Enfermedad o incapacidad de personal clave;
3. Pérdidas imprevisibles de información;
4. Brechas de seguridad;
5. Mal uso de información por servidores públicos;
6. Auditorías y procesos de control interno, y
7. Crisis de comunicación pública.

Asimismo, una gestión integral de datos debe considerar todas las fases del ciclo de vida de los datos dentro de una institución. Por ejemplo, una institución con una política de gestión integral de datos debería realizar las siguientes actividades al crear un nuevo repositorio de datos:

1. Un equipo de tomadores de decisión se reúne y determina la integración de una base de datos acerca de una temática clave definida por normatividad. Integran la documentación base del repositorio de información (i.e. diccionario de datos) y definen métodos seguros para realizar su levantamiento.
2. Basado en las definiciones aplicables, las áreas designadas como responsables de llevar a cabo el levantamiento definen las actividades que aseguran consistencia de datos y, siguiendo un esquema operativo pre establecido, establecen las mejores condiciones posibles para asegurar su integridad y pertinencia.
3. La unidad encargada de gestionar los datos sigue los protocolos establecidos en los manuales para crear un proceso de recepción, tratamiento y vaciado de datos en cooperación con los departamentos de tecnologías de la información y siguiendo las mejores prácticas reconocidas por la documentación institucional.
4. Los tomadores de decisión definen los objetivos de la información y con apoyo de las unidades y personal relevante se establecen indicadores clave, métricas relevantes y mecanismos para mantener actualizado y relevante el repositorio de información. Se fijan medidas y responsables para la publicación y distribución de reportes con una definición clara de responsables y personal con acceso a datos y productos de información.
5. Se define el alcance de mediano y largo plazo de la información. Una vez que se determine que el periodo de utilidad de la información ha llegado a su fin se ejecutan los protocolos establecidos para cerrar el proceso de levantamiento según la normativa y la información existente se almacena en un repositorio designado y se establece un calendario para su potencial eliminación y medidas para su resguardo durante su tiempo de resguardo.

Resiliencia informática

El ciclo de vida de información que se describió anteriormente delinea genéricamente el proceso completo por el que debe pasar un repositorio de información cuando se tiene una perspectiva de largo plazo y políticas de manejo de información apropiadas. Estas políticas mitigan los riesgos, asignan claramente responsabilidades, tienen una clara definición a través de sus distintas etapas y cuentan con participación de equipos multidisciplinarios y multinivel en una organización. Estas actividades integran perspectivas de gobernanza de datos (*data governance*), administración de riesgos y marcos de gestión y responsabilidades informáticas (*accountability framework*), también comprendidas como gestión integral de datos, perspectiva que ve la gestión de datos como un continuo de procesos que cubren todo el ciclo de vida de la información. Esta perspectiva permite construir resiliencia informática para resistir eventos imprevisibles vinculados a los repositorios de información bajo la protección de una institución.



BUENAS PRÁCTICAS

La construcción de esta resiliencia informática y la integración de un esquema de gestión integral de datos está basada en la implementación de buenas prácticas que establecen una base sólida para la operación de las unidades encargadas de la implementación de procesos y políticas de gestión de datos. Por tanto, se recomienda que se implementen a nivel general estas medidas dentro del ecosistema de información. Su implementación ayudará a remediar la mayoría de los problemas y dinámicas perniciosas observadas.

PILAR I RECTORÍA DE DATOS

Obtener respaldo de actores estratégicos de alto nivel

La implementación de políticas de rectoría de datos requiere una gran cantidad de cambios en materia cultural, operativa y de gestión. Por tanto, se requiere apoyo y liderazgo de actores clave de alto nivel para una implementación efectiva en todos los aspectos necesarios de la organización.

Evitar la recolección duplicada de datos para asegurar rectoría transversal de información

Según la experiencia internacional el potencial más grande de desperdicio en la gestión de información se encuentra en la colección duplicada de información. En países como EE. UU. y el Reino Unido esto ha llevado a la creación de normatividad relevante para evitar este tipo de eventos.²³

Establecer controles que cubran de manera integral el ciclo de vida de los datos

La política de rectoría de datos debe ser definida desde una perspectiva integral del ciclo de vida de los datos. Esto es, no solamente afectar una parte de su ciclo de vida, sino que debe considerar desde su creación hasta su eliminación final. Es común cometer el error de crear conjuntos de datos que una vez que no son considerados útiles son dejados de lado y desechados sin una política definida y transparente. La política de rectoría de datos debe atender todos los momentos del ciclo de vida de los datos incluyendo mínimamente, el diseño, desarrollo, levantamiento, tratamiento y procesamiento, almacenamiento y eliminación de información contenida en repositorios de información independientemente de su formato de almacenamiento y nivel de tratamiento. De esta manera, como se señaló anteriormente, se alcanza una situación de resiliencia en materia de gestión de datos.

²³ Se hace referencia a la Orden Ejecutiva 12906 del ejecutivo norteamericano que requiere a las agencias gubernamentales la implementación de procesos internos para asegurar que no existan agencias dentro del gobierno federal que recolecten esta información. En Reino Unido se implementaron mecanismos tecnológicos como el Gigateway Data Locator para ubicar conjuntos de información de las agencias gubernamentales.

Establecer una política de datos sólida y consistente

Definir líneas de acción comunes a todas las unidades involucradas con datos en las cuales se establecen una serie de principios generales de alto nivel que establecen un marco operativo dentro del cual la gestión de datos se pueda desarrollar de manera ordenada. Este documento debe ser aprobado en altos niveles de política pública y debe establecer una figura de alto nivel de liderazgo que enarbole los contenidos de la política frente a actores clave. Es importante señalar que esta política de datos no puede estar contenida en una política de uso y explotación de tecnologías de información. Estas deben ser independientes toda vez que los datos utilizados por una tecnología u otra son independientes de los sistemas que la explotan para fines de análisis.

Definición clara de responsables de información con un marco de responsabilidades y rendición de cuentas (accountability framework)

Uno de los aspectos clave de la gestión integral de datos es la clara identificación de la titularidad de actores específicos sobre múltiples dimensiones del ciclo de vida de la información. Normalmente, la organización que requiere el levantamiento de datos o su compilación de otras fuentes retiene los derechos de gestión operativa y conceptual sobre los datos.

Para desarrollar esta asignación se requiere el desarrollo de un marco de responsabilidad y rendición de cuentas (*accountability framework*) que establezca con claridad la titularidad de su información. Este marco establece titularidad sobre los datos; obligaciones legales y operativas; políticas institucionales de seguridad de datos, liberación y distribución, así como los acuerdos y formalidades que se requerirán para modificar los niveles de acceso y protección a los datos.

PILAR II. CALIDAD DE DATOS

Control de validez

El aseguramiento de la calidad de los datos debe considerar políticas y medidas que aseguren que el contenido de las bases de datos se ajuste a los estándares de validez en la captura de información y que todos los campos requeridos sean llenados por parte de los sujetos obligados. En casos del guardado de información que sea transversal a través del ecosistema de instituciones con sujetos obligados es importante implementar medidas de *data scrubbing* para eliminar instancias duplicadas de datos y otras incidencias.

Validación de veracidad

Adicional al control de la validez de los datos se deben de implementar políticas que permitan la revisión de la veracidad de la información capturada en los repositorios de información de cada institución. Estas políticas deben incluir la auditoría de la información existente en bases de datos de manera aleatoria o integral según las capacidades operativas de cada unidad y el perfil de riesgo y criticidad de la información. Esto permite a las instituciones contar con protocolos que validen la información almacenada en las bases de

datos para maximizar la utilidad de la información y robustecer la confiabilidad de las mismas frente a cualquier observador interno o externo.

Control de cambios a datos

El control de cambios debe formar parte de una política robusta de calidad de datos. Esto es, se deben implementar políticas y acciones programadas para asegurar que se cuente con registros históricos de la información que permitan proteger la integridad histórica de la información. Estas medidas pueden comprender registros históricos de actividades en las bases de datos mediante bitácoras del sistema o guardado resguardado con contraseña de versiones de las bases de datos.

PILAR III. CONSISTENCIA CONCEPTUAL DE ESTRUCTURA DE DATOS

Definición de una política clara y de alto nivel de metadatos

Todos los repositorios de información deben contar con una política puntual de metadatos. Esto es una serie de principios de alto nivel que enmarquen el esfuerzo de integración de metadatos en la organización. Estos deben establecer la importancia, transversalidad y obligatoriedad de la recolección de metadatos como un recurso que debe mantenerse relevante y con actualización de manera oportuna.

Definición de estándares de metadatos

Los metadatos recolectados deben seguir estándares comunes a las organizaciones que se involucren de manera que se logre experiencia y valor compartido. En el caso de ecosistemas de organizaciones como es el caso de la información relevante para el SNA se debe de hacer una definición de estándares que incluya a todos los actores relevantes de cada entorno temático de información.

Preparar un levantamiento común de metadatos

Los líderes del esfuerzo de la política de metadatos deben asegurar que se lleve a cabo un esfuerzo coordinado para establecer el conjunto inicial de metadatos a fin de poder identificar y atender las instancias de duplicación de información, diferencias que requieran atención en conceptualización de campos en las bases de datos y comenzar los trabajos hacia la construcción de conceptos comunes en los campos de datos de la información disponible en el entorno relevante.

Establecer un grupo que asegure la vigencia y relevancia de la información de metadatos

Adicionalmente a la realización de un levantamiento de metadatos es importante establecer un grupo encargado de mantener vigente y relevante la información en los archivos de metadatos con políticas formales de actualización y gestión aplicables a estos archivos de datos. Estas responsabilidades deben formar parte del marco de responsabilidades y rendición de cuentas de cada organización.

PILAR IV. DOCUMENTACIÓN DE ESTRUCTURAS DE DATOS

Definición de una política clara y de alto nivel de documentación de sistemas y formalización de protocolos de acción

Es importante que toda institución cuente con una plataforma formal de políticas sobre documentación requerida y formalización de protocolos de acción establecidos y distribuidos en memorándums o lineamientos. Estos deben enarbolarse por actores clave de alto nivel de manera que se transmita su relevancia y obligatoriedad a todos niveles a fin de establecer un control firme sobre el comportamiento alrededor de datos de la organización.

Definición de estándares aplicables a la documentación generada sobre repositorios de datos y sistemas electrónicos

El desarrollo documental de una organización requiere la definición de estándares comunes para la generación de documentación de manera que se integren conjuntos de información común y con metodologías compartidas. Estas políticas permiten crear un conjunto de conocimiento consistente, comparable y homogéneo a través del tiempo.

APLICACIÓN DE BUENAS PRÁCTICAS



APLICACIÓN DE BUENAS PRÁCTICAS AL ENTORNO DE LA LGSNA

¿Cómo aplicar las buenas prácticas internacionales en el entorno de la LGSNA?

La mejora en las prácticas de gestión de datos requiere de la implementación de medidas que atiendan las necesidades específicas detectadas en el ecosistema de datos del LGSNA. En adelante se presentan recomendaciones que se pueden implementar en el corto y mediano plazos en respuesta a problemas específicos del ecosistema de información.

AUSENCIA DE FOROS PARA LA DIFUSIÓN DE EXPERIENCIAS EN ÁREAS TEMÁTICAS CLAVE

Problema: Se observó que aun en los casos en que se cuenta con sistemas técnicamente desarrollados que se han difundido a través de un amplio grupo de instituciones, no se han establecido grupos cooperativos para conocer y aprovechar las amplias experiencias observadas en el campo en pos del enriquecimiento de estos sistemas. Los conocimientos derivados de este tipo de ejercicio podrían ser aprovechados en todas las áreas temáticas definidas por la LGSNA.

Respuesta: Creación de grupos de trabajo conjuntos a través de todas las instituciones con atribuciones para el levantamiento de información en áreas temáticas. Estos grupos deberán recuperar las experiencias, necesidades y ventanas de oportunidad detectadas por los miembros del grupo a fin de que las instituciones encargadas de cada plataforma puedan explotar la información para poder robustecer cada sistema electrónico.

Problema: Las instituciones encargadas de la gestión de información tienen prácticas internas de manejo de bases de datos y sistemas de alto riesgo. La edición directa en bases de datos, ausencia de guardado de versiones de documentos y el acceso irrestricto y no supervisado a bases de datos son comportamientos que deben eliminarse por completo. No obstante, las medidas que atienden estos problemas deben ser homogéneas a través de las instituciones que manejan esta información e implementarse de manera consistente para todos los repositorios de información aplicables.

Respuesta: Una perspectiva de largo plazo debe desarrollar manuales y procedimientos estándar. Estos proveerán respuestas consistentes, robustas y funcionales para todos los niveles de maduración de sistemas electrónicos y bases de datos.

III FALTA DE ASIGNACIÓN DE RESPONSABILIDADES SOBRE LAS MÚLTIPLES DIMENSIONES DE LA GESTIÓN DE DATOS

Problema: Las instituciones no tienen una clara asignación de las responsabilidades asociadas a las múltiples dimensiones de la gestión de datos. Esto resulta en que, actualmente, no haya una definición clara, formal y sustentada de quién es dueño de qué responsabilidades respecto de los procesos, infraestructura, políticas de desarrollo, definiciones conceptuales, documentación y otros aspectos clave para una buena gestión integral de datos.

Respuesta: Las instituciones deben definir una matriz de responsabilidades (*accountability framework*) para cada repositorio de información conforme a los estándares internacionales en la materia. Esta matriz asigna de manera puntual y formal las responsabilidades de gestión de datos, gestión de metadatos, gestión de los aspectos operativos del levantamiento de información, gestión de aspectos normativos del levantamiento de información, gestión del desarrollo y maduración de la información y otras áreas de la gestión de datos entre los actores relevantes y pertinentes. De esta manera, se mitiga el riesgo de una posible desatención sobre áreas estratégicas del manejo de información.

IV DEFINICIÓN DE HOJAS DE RUTA PARA LA GESTIÓN DEL CICLO DE VIDA DE BASES DE DATOS Y SISTEMAS ELECTRÓNICOS

Problema: Los servidores públicos no cuentan con guías para comenzar a trabajar en la creación, desarrollo o eliminación de repositorios de datos. Por lo tanto, cada servidor público que adquiere una responsabilidad en este respecto tiene que descubrir aspectos básicos de procesos cuyos hitos ya se encuentran claramente definidos a nivel internacional.

Respuesta: Las instituciones dentro del ecosistema pueden crear, de manera independiente o conjunta, documentación que indique la ruta crítica para la gestión de los momentos del ciclo de vida de los repositorios de información. I.e. Un documento con la ruta crítica para la creación de un sistema explica que se debe elegir el objetivo de la base de datos, que la operacionalización de las variables seguirá un formato institucional cuyos componentes codifican información sobre las variables, que si se contrata a un proveedor externo para el desarrollo de software se debe solicitar la provisión de documentación sobre los diccionarios de datos, casos de uso o modelos de la arquitectura de las estructuras de datos, entre otras cosas. Esta planeación deberá atender múltiples dimensiones clave como protocolos de protección de la integridad de los datos y límites claros y formales a su acceso (límites al acceso, guardado versionado de los datos, registros de acceso a la información, etc.) acordes al nivel de maduración de cada sistema o base de datos.

V DEFINICIÓN DE HOJAS DE RUTA PARA LA TRANSICIÓN ENTRE NIVELES DE MADURACIÓN DE SISTEMAS ELECTRÓNICOS Y BASES DE DATOS

Problema: No existe claridad sobre el camino que se debe de transitar para avanzar de una base de datos en una hoja de cálculo y que no cuenta con diccionario de datos hacia un repositorio de datos sujeto a una gestión integral de datos apropiada y relevante a la magnitud y características de cada conjunto de información. Esto crea la sensación de que una buena gestión es un parámetro inalcanzable y no un continuo de políticas con hitos puntuales y viables para cualquier unidad de trabajo.

Respuesta: En el entorno internacional la práctica más común para responder a este problema es la creación de modelos de maduración para organizaciones gubernamentales. Esto implica la creación de una ruta crítica de la maduración de sistemas que es relevante a cualquier tipo de repositorio de datos y que implica medidas al alcance de cualquier unidad operativa del gobierno. Estos modelos no se centran en materia de la infraestructura tecnológica, sino que establecen hitos en el manejo de la información y señalan los escalones de maduración y las características que se requerirán para subir al siguiente escalón. Esto permite avanzar en el desarrollo de los sistemas y dar previsibilidad a los usuarios y administradores del sistema de sus recursos, necesidades y agenda pendiente.

VI DISEÑO DE ESTRUCTURAS DE GOBERNANZA DE DATOS (DATA GOVERNANCE) PARA LA ADMINISTRACIÓN DE LOS REPOSITARIOS DE INFORMACIÓN

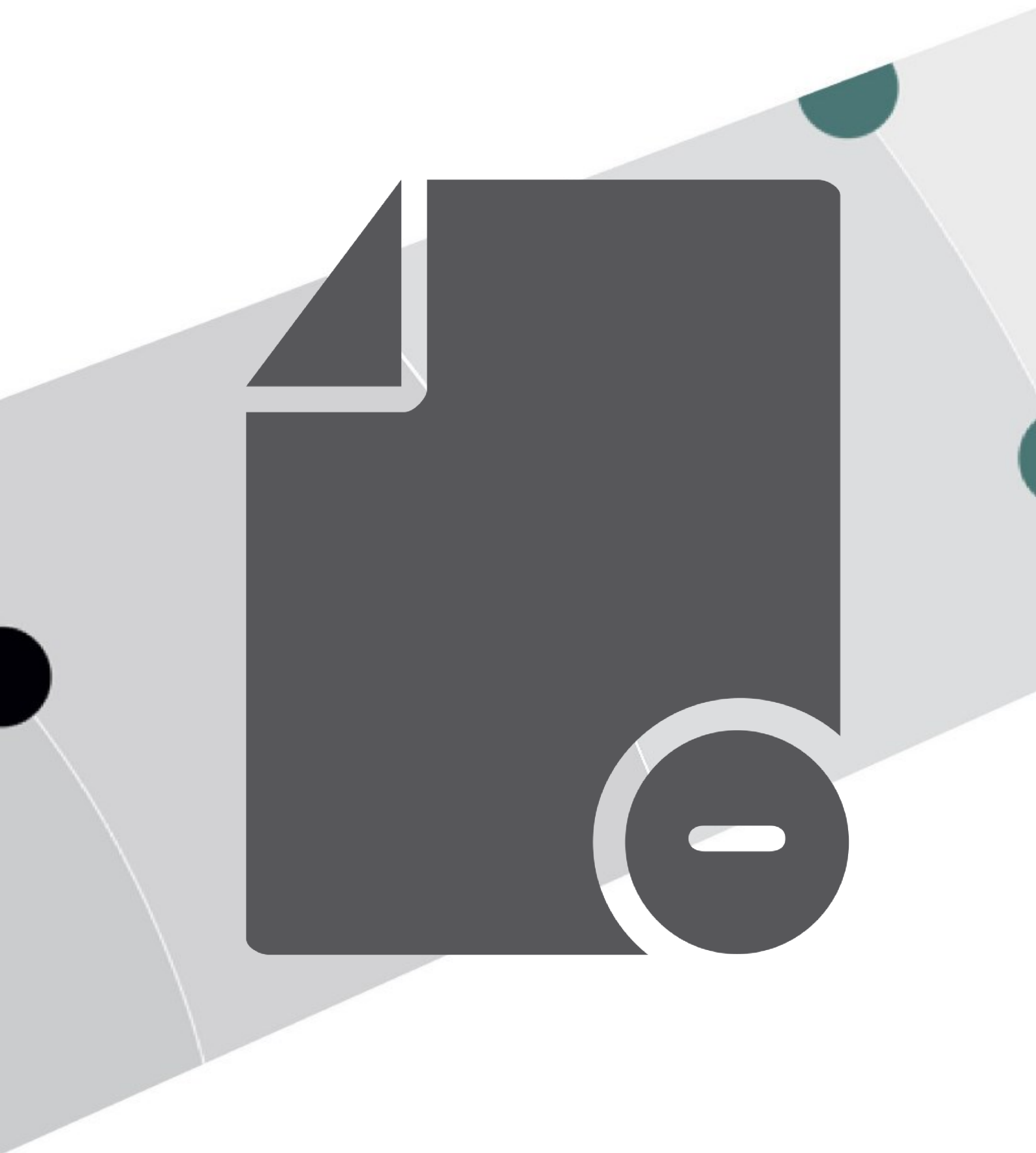
Problema: La gestión de datos requiere dirección estratégica, pero con frecuencia no se cuenta con un órgano con la tarea de planear en el mediano y largo plazos la dirección que tomarán los repositorios de datos.

Respuesta: Se necesita crear una estructura formal de toma de decisiones que tenga como misión planear los objetivos de la gestión de datos dentro de la organización. Esto implica el establecimiento de juntas institucionales o interinstitucionales que cuenten con las facultades, atribuciones y responsabilidades necesarias. Por ejemplo, definir que para el año 2019 se buscará que todos los repositorios de información cuenten con documentación de la arquitectura de datos.

VII PRÁCTICAS DE LEVANTAMIENTO DE INFORMACIÓN EN FORMATOS NO LEGIBLES POR MÁQUINAS.

Problema: Aún se observan instituciones que capturan la información de áreas temáticas relevantes para la LGSNA en formatos no legibles por máquinas (análogos y digitales). Esto no sólo implica un problema de capacidad de uso de la información para la institución, sino que impacta en el costo del proceso de transición de estos datos a un esquema legible por máquinas. Cualquier medida a implementar a futuro que desee corregir esta situación se enfrentará a un elevado costo para la integración de datos históricos, así como cambios organizacionales que son difíciles de revertir en el corto plazo.

Respuesta: Se requiere establecer políticas en el corto plazo que reduzcan la cantidad de información capturada en estos formatos a fin de mitigar la tasa de crecimiento del problema. Por ejemplo, la definición de formatos de captura de información clave para áreas temáticas con alta prevalencia de este fenómeno. Adicionalmente, se deben de definir calendarios de trabajo para la definición de rutas críticas hacia la eliminación de estas prácticas entre los integrantes del ecosistema de información.



ANEXO I. PROGRAMA DE TRABAJO

La integración del inventario de sistemas electrónicos de información y bases de datos existentes se realizó mediante la implementación de una metodología dual de detección y diagnóstico que consta de 4 etapas:

1. Desarrollo de herramientas metodológicas;
2. Aplicación de cuestionarios de diagnóstico;
3. Estudio y análisis de sistemas reportados, y
4. Análisis de información extraída.

En la primera etapa se desarrollaron herramientas analíticas para el levantamiento ordenado y metodológicamente robusto de información de los sistemas susceptibles de ser integrados en la Plataforma. Durante esta misma fase se realizaron acercamientos a las instituciones sujetas a la LGSNA para la presentación de los alcances, objetivos y limitaciones del inventario a levantarse. El objetivo de estos acercamientos fue obtener retroalimentación, alinear expectativas y promover la comprensión de las necesidades de información que serían extendidas a las instituciones del orden federal.

Durante el proceso de comunicación y de acercamiento inicial con las instituciones participantes de los poderes federales se tomaron acciones en dos direcciones. Primero, realizar una exposición presencial de los objetivos, alcances y del calendario de trabajo del estudio. Segundo, crear y distribuir material informativo y de apoyo para facilitar el proceso de detección de sistemas electrónicos y bases de datos susceptibles de ser integrados en la Plataforma.

Estas líneas de acción se ejecutaron en beneficio de todas las instituciones participantes y en estrecha relación con los enlaces designados por los asistentes al evento de presentación del proyecto. El equipo manifestó a los participantes que las necesidades de información se encuentran en áreas temáticas clave definidas por la LGSNA.

La segunda etapa, corresponde a la aplicación de las herramientas de diagnóstico diseñadas de manera previa. Esta fase del estudio contó con la participación de 17 enlaces de cada una de las 18 instituciones consideradas para la participación de los poderes federales. El único caso de excepción de una institución considerada para el estudio y con la cual no se pudo contar con información fue el Tribunal Superior Agrario que durante el periodo de levantamiento no desarrolló actividades con motivo de las afectaciones del sismo del 19 de septiembre del 2017. Esta situación fue notificada en su momento a la SESNA que, en conjunto con el equipo de investigación del CIDE, tomó la determinación de proceder sin el levantamiento de la información correspondiente a aquella institución.

La tercera etapa está definida por la implementación de una serie de levantamientos de información en sitio a través de entrevistas. Estas acciones permitieron al equipo de investigación poder construir un panorama contextual de la información provista en los cuestionarios e indagar a profundidad en las realidades operativas y técnicas de estos sistemas.

La cuarta etapa representa únicamente el periodo de sistematización, tratamiento, limpieza y análisis de la información recabada con el fin de poder extraer hallazgos clave y definir las necesidades estratégicas de información que, se pueden satisfacer, en beneficio de los procesos de decisión del SESNA.

ANEXO II. CUESTIONARIO INICIAL DE DIAGNÓSTICO

No.	Pregunta
1	Nombre de la organización a la que pertenece.
2	Nombre del responsable del llenado del cuestionario.
3	Unidad/Departamento donde desempeña su cargo.
4	Cargo que desempeña actualmente.
5	Dirección institucional de correo electrónico.
6	Teléfono de contacto.
7	Medio preferido de contacto.
8	Nombre del sistema o base de datos a registrar.
9	Nombre de la organización que administra el sistema o base de datos
10	Unidad o departamento encargado de la gestión de información.
11	Unidad o departamento que funge como fuente de datos.
12	Indique, en términos generales, quién es el dueño de información del sistema o base de datos. Esto es, el área y cargo que puede tomar decisiones sobre la estructura, acceso, uso y aprovechamiento del sistema o base de datos y la información vinculada a la misma.
13	Indique el nombre del manejador de base de datos que se utiliza.
14	En caso de no contar con un manejador de base de datos indicar el formato de almacenamiento de la información del sistema o base de datos.
15	Indique si el sistema o base de datos tiene información sobre servidores públicos y/o empleados directos o indirectos del gobierno federal.
16	Indique si el sistema o base de datos contiene información sobre la situación fiscal de los servidores públicos y/o empleados directos o indirectos del gobierno federal.
17	Indique si el sistema o base de datos contiene información sobre la situación patrimonial de los servidores públicos sancionados con relación al incumplimiento de sus obligaciones en el desempeño de sus empleos, cargos o comisiones.
18	Indique si el sistema o base de datos tiene información sobre servidores públicos sancionados con relación al incumplimiento de sus obligaciones en el desempeño de sus empleos, cargos o comisiones.
19	Indique si el sistema o base de datos contiene información sobre proveedores de bienes o servicios que hayan participado con la intención de obtener beneficios o que hayan sido beneficiados como parte de una adquisición, arrendamiento, enajenación de todo tipo de bienes, prestación de servicios de cualquier naturaleza o la contratación de obra adjudicada o licitada mediante convocatoria pública por una parte de una institución pública.
20	Indique si el sistema o base de datos contiene información sobre la situación patrimonial de estos proveedores de bienes o servicios.

No.	Pregunta
21	Indique si el sistema o base de datos contiene información sobre la situación fiscal de estos proveedores de bienes o servicios.
22	Indique si el sistema/base de datos contiene información sobre sanciones aplicadas a estos proveedores de bienes o servicios.
23	Indique si el sistema/base de datos contiene información relativa a la evaluación, asignación o documentación de cualquier tipo de adquisiciones, arrendamientos, enajenaciones de cualquier tipo de bienes, prestación de servicios de cualquier naturaleza y la contratación de obra que realicen, sean adjudicadas o licitadas mediante convocatoria pública, instituciones públicas.
24	Indique si el sistema/base de datos contiene información sobre hechos reportados mediante quejas y/o denuncias públicas de faltas administrativas o hechos de corrupción.
25	Por favor, indique la periodicidad de la actualización del sistema o base de datos.
26	Por favor, indique la fecha de la última actualización del sistema o base de datos.
27	Por favor indique si se prevé continuar con la actualización del sistema o base de datos en el futuro.
28	Por favor indique si el sistema o base de datos cuenta con la siguiente documentación
29	Indique si existen leyes o reglamentaciones que regulen la existencia del sistema o base de datos.
30	En caso de que aplique indique el marco legal aplicable.

ANEXO III. CUESTIONARIO DETALLADO DE DIAGNÓSTICO DE SISTEMAS INFORMÁTICOS.

No.	Pregunta
1	Nombre del responsable del llenado del cuestionario.
2	Unidad/departamento en el que desempeña su cargo.
3	Cargo que desempeña actualmente.
4	Dirección institucional de correo electrónico.
5	Teléfono de contacto.
6	Medio preferido de contacto.
7	Nombre del sistema o base de datos
8	Indique el año en que se creó o comenzó a registrar oficialmente el sistema o base de datos.
9	Indique cuáles son los 3 objetivos sustantivos principales del sistema o base de datos.
10	Enumere los 3 principales usos que se le dan a la información contenida en el sistema o base de datos.
11	Indique el modo de generación de la información en su sistema o base de datos.
12	¿Cuál es el método de acceso que utiliza para consultar o editar información de su sistema o base de datos?
13	¿Se tienen definidas formalmente las principales necesidades de información que actualmente tienen los dueños de la información respecto del sistema o base de datos?
14	En caso afirmativo, por favor indique el grado de satisfacción de los consumidores principales de la información con el actual arreglo de datos de su sistema o base de datos.
15	Señale la fuente de su estimación.
16	¿Existe normatividad de cualquier nivel (leyes, reglamentos, etc.) que regule la existencia, operación, validación o auditoría de la información del sistema o base de datos?
17	en caso afirmativo, por favor indique si según la normatividad aplicable de la fuente de la información está obligada a entregar los datos requeridos para el sistema o base de datos.
18	Señale la normatividad a la que se hace referencia.
19	¿Se realizan evaluaciones de la veracidad de los registros capturados en el sistema o base de datos? (i.e. auditorías)
20	¿Existe alguna ley, reglamento, norma, disposición administrativa, resolución interna o cualquier tipo de regulación que reglamente la auditoría de la información capturada en el sistema o base de datos?
21	Señale la normatividad a la que se hace referencia.
22	Indique si al menos parte de la información contenida en su sistema o base de datos es accesible a terceras partes (personas fuera de la unidad/departamento dueño de la información ya sean ciudadanos o personal de cualquier institución).
23	En caso afirmativo, indique qué tipo de tercera parte tiene acceso a la información (Puede seleccionar más de una opción).

No.	Pregunta
24	En caso de que la información sólo pueda consumirse por personal de la institución, ¿existen acuerdos de confidencialidad firmados entre terceras partes y el dueño de la información? (si es la misma unidad se puede considerar un compromiso de confidencialidad marco de la institución con sus funcionarios).
25	Indique si la información es reservada según lo establecido en la LFTAIP y demás normativa aplicable.
26	En caso de que haya información clasificada, por favor indique qué porcentaje de los campos del sistema o base de datos se encuentran bajo esta clasificación (campos del diccionario de datos).
27	En caso de que haya información clasificada, por favor indique qué porcentaje de los registros (renglones) del sistema o base de datos se encuentran bajo esta clasificación. En caso de que su sistema realice actualizaciones diarias indique el porcentaje al corte del mes previo.
28	En caso de que menos del 100% de la información tenga el estatus de clasificada, por favor indique si la información no clasificada se hace pública de manera proactiva como parte de los esfuerzos de transparencia de su institución.
29	si se hace pública la información de manera proactiva, por favor indique en que medios se pone a disposición del público.
30	Indique el número de campos/atributos con los que se cuenta su sistema o base de datos.
31	Indique el número total de registros que forman parte actualmente en su sistema o base de datos.
32	Indicar la fecha de consulta del número de registros (DD/MM/AAAA).
33	Indique el número de registros del sistema o base de datos al cierre de las últimas cinco actualizaciones de información. En caso de que se tenga una actualización diaria, o de sistemas que se actualizan en tiempo real, indique el número de registros capturados en cada uno de los últimos cinco meses (capture sus respuestas desde el más antiguo al más reciente).
34	Indique el porcentaje respecto del total que representa el número de registros capturados durante las últimas tres actualizaciones. En caso de que se tenga una actualización diaria, o de sistemas que se actualizan en tiempo real, indique el número de registros capturados en cada uno de los últimos cinco meses (capture desde el más antiguo al más reciente).
35	si es posible, indique el número de registros con los que contaba su sistema o base de datos al cierre de cada uno de los últimos tres años (capture desde el registro más antiguo al más reciente).
36	¿El porcentaje de registros sin datos en la base de datos principal consolidada supera el 5% (Considere cada celda vacía es decir la intersección entre registros y campos vacíos).
37	¿El porcentaje de registros con variables (campos) sin datos en el archivo principal consolidado del sistema/base de datos supera el 20% en alguna de las variables?
38	¿Cuál es el porcentaje de edición de datos (modificaciones de variables debido a datos faltantes o errores de inconsistencia de la información)?
39	Por favor indique el mayor porcentaje de valores atípicos (<i>outliers</i>) que haya encontrado (si corresponde).

No. Pregunta

- 40 ¿Cuál es el porcentaje de errores de codificación? (errores detectados de captura de datos en el sistema o base de datos)
-
- 41 ¿El sistema/base de datos cuenta con diccionario de datos?
-
- 42 En caso de que cuente con diccionario de datos, indique si se han agregado campos al sistema/base de datos luego de su creación original.
-
- 43 En caso afirmativo, indique si el diccionario de datos está actualizado al momento de responder el cuestionario.
-
- 44 Si aplica, indique la fecha de la última modificación del diccionario de datos.
-
- 45 En caso de que su sistema o base de datos cuente con un modelo entidad-relación, indique si ha sufrido cambios luego de su diseño original (creación de nuevas tablas o campos o cualquier otro ajuste que afecte la información sustantiva contenida).
-
- 46 Indique la fecha de la última modificación del modelo entidad-relación.
-
- 47 ¿Se documentan los metadatos del sistema/base de datos (información sobre los datos, población objetivo, métodos de captura, etc.)?
-
- 48 ¿Existe un procedimiento documentado de control de cambios en los procedimientos y/o definiciones?
-
- 49 ¿Existe algún procedimiento documentado de captura de los datos del sistema/base de datos?
-
- 50 ¿Existe algún procedimiento documentado de control de consistencia de los datos del sistema/base de datos?
-
- 51 ¿Se han documentado los mecanismos para garantizar la integridad referencial de la base de datos del sistema/base de datos?
-
- 52 Indique el principal mecanismo utilizado.
-
- 53 en caso de que la unidad que gestione el sistema /base de datos no sea la fuente de la información, ¿los términos de entrega de datos están documentados y definidos formalmente?
-
- 54 En caso de que la unidad que gestione el sistema/base de datos no sea la fuente de información, ¿existen procedimientos documentados para asegurar la protección física de los datos (mecanismos de control de acceso a servidores, centros de procesamiento de información, etc.)?
-
- 55 ¿Existe algún mecanismo establecido con el propósito de mejorar la manera en que el sistema/base de datos cubre las necesidades de sus consumidores de información? (i.e. consultas o encuestas).
-
- 56 ¿Existen controles establecidos para verificar que la información capturada en el sistema/base de datos sea veraz y precisa?
-
- 57 ¿Existen controles para verificar que la información capturada en el sistema/base de datos siga los parámetros establecidos para su captura? (control de rangos, códigos válidos, etc.)
-
- 58 ¿Existen controles establecidos para verificar que los valores extremos registrados en el sistema/base de datos no constituyan errores de captura o de algún otro tipo?
-
- 59 En caso de que se detecte un registro con datos inválidos o en blanco, ¿el dueño de la información hace alguna modificación en el sistema/base de datos?
-
- 60 En caso afirmativo, indique el tipo de modificación que se hace a datos inválidos o en blanco.

No. Pregunta

- 61 ¿Las modificaciones de los datos del sistema son realizadas de acuerdo a un procedimiento establecido?
-
- 62 Indique si la información en su sistema o base de datos se audita con regularidad.
-
- 63 En caso de que se realicen auditorías sobre la información contenida en su sistema/base de datos, señale si estas auditorías han resultado en una revisión de los datos del sistema o base de datos.
-
- 64 conforme a la última auditoría, indique el porcentaje de ajustes que sufrió la información de su sistema o base de datos como resultado de la auditoría. El porcentaje de ajustes es el número total de registros modificados, sobre el número total de registros en el sistema o base de datos.
-
- 65 En caso de que se requieran ajustes a la información en los registros, indique con que frecuencia se hace la consolidación final de información.
-
- 66 La información contenida en el sistema o base de datos, ¿tiene su origen en instituciones externas o es provista por la misma organización?
-
- 67 En caso de que la información tenga origen en la misma institución, ¿Cuántas unidades o departamentos participan?
-
- 68 Indique el nombre de las unidades que participan en la integración de la información
-
- 69 En caso de que la información sea provista por instituciones externas, ¿Cuántas instituciones reportan la información a partir de la cual se actualiza el sistema o base de datos?
-
- 70 En caso de que la información tenga origen fuera de su institución ¿Cuántas unidades participan?
-
- 71 En caso de que la información no se genere por la misma unidad/departamento que gestiona el sistema indique si hay un medio definido de entrega de los datos que serán integrados al sistema/base de datos.
-
- 72 Indique el medio de entrega acordado para la información capturada en su sistema/base de datos.
-
- 73 En caso de que haya información enviada para su posterior carga, indique el tiempo promedio de captura o actualización de la misma.
-
- 74 En caso de que la información no sea capturada como parte de un proceso de trabajo, indique si en la normatividad aplicable se establecen calendarios de trabajo para la carga de información.
-
- 75 Señale la normatividad a la que se hace referencia.
-
- 76 En caso de que su sistema/base de datos contenga información reportada y no surja de una actividad vinculada a su operación, ¿existe normatividad que regule la validación de la información reportada en el sistema/base de datos? (métodos de comprobación documental o de cualquier otra naturaleza).
-
- 77 Señale la normatividad a la que se hace referencia.
-
- 78 ¿La información contenida en el sistema/base de datos es sometida a un filtrado o tratamiento previos a su ingreso en su recipiente final?
-
- 79 ¿El sistema o base de datos opera de manera independiente o consume información de otros sistemas/bases de datos?

No. Pregunta

- 80 En caso de que se consuman datos de otros sistemas/bases de datos indique sus nombres.
- 81 Indique el nombre de la unidad que funge como fuente de la información.
- 82 ¿La unidad que funge como fuente de la información es la misma de aquella que la consume?
- 83 Indique la relación jerárquica entre el consumidor de la información y la fuente de la información.
- 84 ¿Se recuperan estadísticas de uso y explotación de su sistema/base de datos?
- 85 Indique el número de visitas únicas promedio a su sistema durante los últimos 6 meses.
- 86 Indique el número de usuarios registrados con los que contaba su sistema o base de datos al cierre de cada uno de los últimos tres años. Comience por el año más antiguo y termine en el año más cercano.
- 87 Indique el número total de visitas a su sistema desde su creación.
- 88 Indique el número total anual de visitas a su sistema para cada uno de los últimos tres años comenzando por el año más antiguo y hacia el más reciente.
- 89 Indique el promedio mensual de visitas a su sistema durante los últimos tres meses comenzando por el mes más antiguo y hacia el más reciente.
- 90 Indique el promedio anual de visitas a su sistema durante los últimos tres años comenzando por el más antiguo y hacia el más reciente.
- 91 Indique si la carga de información se realiza bajo un calendario establecido o solicita acciones a usuarios en un periodo específico del año.
- 92 Indique si el sistema o base de datos tiene un consumidor de información dentro de la institución.
- 93 Señale quién o quiénes son los consumidores de la información dentro de la institución (cargos).
- 94 Señale el número estimado de consumidores de la información dentro de la institución
- 95 Indique la frecuencia con que la información es consultada por los consumidores dentro de la institución.
- 96 Indique si hay consumidores de la información contenida en el sistema o base de datos en otras instituciones públicas.
- 97 Nombre de la institución pública que consume la información.
- 98 Señale quién o quiénes son los consumidores de la información en otras instituciones públicas (cargos).
- 99 Señale el número estimado de consumidores de la información en otras instituciones.
- 100 Indique la frecuencia con que la información es consultada por los consumidores de otras instituciones.

ANEXO IV. PROTOCOLO DE ENTREVISTA (LEVANTAMIENTO EN SITIO).

OBJETIVOS

Obtener un diagnóstico profundo de las condiciones operativas, estructurales y de alcances de cada sistema seleccionado para atenderse como parte del inventario de sistemas electrónicos de información y bases de datos existentes a cargo de los entes públicos del orden federal que servirán como base para la conformación de la Plataforma Digital Nacional.

Las entrevistas buscan diagnosticar los retos y perspectivas del entorno en los que se desenvuelven cada uno de los sistemas, así como de los comportamientos de aprovechamiento y explotación de la información por parte de los servidores públicos dueños de la información.

PROTOCOLO INICIAL

1. Presentaciones profesionales.
2. Objetivo de la investigación.
3. Breve explicación de la metodología. ¿Qué son las entrevistas semi-estructuradas y duración aproximada de la entrevista?
4. Agradecimiento por la participación.

PARTICIPANTES

Centro de Investigación y Docencia Económicas

Primer Entrevistador:

Segundo Entrevistador:

FECHA

día de mes de 2017

SEDE

Dirección

HORA DE INICIO

00:00 horas

HORA DE TÉRMINO

00:00 horas

BATERÍA DE PREGUNTAS

I. PERFIL DEL SISTEMA

Como parte del diagnóstico de las características del sistema electrónico o base de datos se debe recabar información sobre sus usos y objetivos en el marco del contexto operativo diario más allá de sus usos y objetivos formales.

1. Más allá de atender requerimientos legales y operativos, ¿cree que el sistema cumple algún otro objetivo durante su uso dentro de esta institución?
 - A. Objetivo: definir objetivos sustantivos percibidos por operativos que puedan estar desalineados con objetivos formales y que por ello no hayan sido reportados. Indicar objetivos y usos reportados
Objetivos reportados
Usos reportados
2. ¿Cree que el sistema o base de datos ha llegado a su punto de maduración más alto?
 - A. En caso negativo. ¿A dónde cree que podría crecer el sistema para robustecer su funcionamiento y cumplir mejor sus objetivos actuales?
 - B. En caso positivo. ¿Cree que hay algún otro objetivo al que este sistema/base de datos debería aspirar a cumplir?
3. Del 1 al 10, ¿qué tan exitosa es la información contenida en el sistema o base de datos para cumplir sus objetivos?
 - A. En caso de que sea el sistema actual menos de diez. ¿Qué le impide al sistema ser un diez?
4. ¿Qué operaciones realiza cotidianamente con esta información?
5. ¿Cuáles son las funcionalidades o aplicativos del sistema que más usa?
6. ¿Hay alguna parte del sistema o base de datos que sepa que está ofrecido pero que nunca ha usado

II. MARCO LEGAL

Buscar definir ventanas de oportunidad en el marco normativo que puedan presentar obstáculos o facilitadores hacia el esfuerzo de integración dentro de la Plataforma Digital Nacional.

1. ¿Existe algún nivel normativo que debería o podría modificarse para mejorar la operación o características del sistema o base de datos?
2. ¿Qué dificultades normativas podrían presentar el sistema/base de datos con miras a su integración en el marco de la Plataforma Digital Nacional.

III. VOLUMEN DE DATOS

Buscar que el entrevistado nos apoye en el diagnóstico de posibles tendencias en los datos del sistema que puedan ser relevantes para el desarrollo del sistema o base de datos en el mediano y corto plazos.

1. ¿Considera que en el corto o mediano plazos podría haber un cambio importante en el comportamiento de los datos registrados en el sistema/base de datos? En caso afirmativo, desarrolle.

- ¿Qué eventos en el entorno del sistema/base de datos podrían implicar variaciones importantes en la cantidad de información contenida en el mismo? (Considerar política pública o situación socioeconómica del país) Buscar diagnosticar riesgos en existencia o presiones de información sobre el sistema/base de datos. En caso de que sea una base de datos no tratada formalmente (Excel o similar) indagar si estos cambios podrían crear presión para formalizar la información.

IV. CALIDAD DE LA INFORMACIÓN

La información puede tener muchos problemas. Por ejemplo, errores de captura, reportes de información incorrecta (intencional o no intencional), documentación insuficiente de respaldo, malentendidos en los procesos de captura. Todos ellos representan diferentes riesgos para los tomadores de decisiones y sus responsabilidades. Indagar con el fin de diagnosticar si existen riesgos relevantes en el arreglo operativo y de gestión del sistema electrónico o base de datos.

- Del 1 al 10 que tan buenos considera que son los filtros que aseguran la calidad de la información de este sistema/base de datos.
 - En caso de que sea el sistema actual menos de diez. ¿Qué le impide al sistema ser un diez?
- ¿Han tenido algún problema por errores en la información o te preocupa que pueda presentarse uno? Desarrollar.
- ¿Qué harías tú hoy para evitar ese problema? Analizar proceso, sistema, estructura de datos.

V. CONSUMO DE INFORMACIÓN

Como parte del proceso de explotación de información por parte de los usuarios que la utilizan como información de consulta es importante definir la naturaleza de los objetivos de estas consultas.

Contrastar pregunta 9 “objetivos sustantivos principales del sistema o base de datos” con preguntas de página 9 “Consumidores de información”. Si es consumida por actores internos enfocar preguntas 92-95, en caso de que sean externos utilizar preguntas 96-100.

- Validar consistencia entre objetivos y usuarios y perfil de usuarios y frecuencias de uso reportados.

Consumidor	Pregunta	Respuesta	¿Información es
Interno/ Ex-	Prueba 1	Prueba 1	Si/No porque

VI. CORROBORAR INFORMACIÓN DE CUESTIONARIO

Matriz de preguntas a corroborar

No.	Pregunta	Respuesta a corroborar	Respuesta confirmada

VII. INDAGAR SOBRE INFORMACIÓN FALTANTE EN CUESTIONARIOS.

Como parte del proceso de explotación de información por parte de los usuarios que la utilizan como información de consulta es importante definir la naturaleza de los objetivos de estas consultas.

Matriz de información faltante

No.	Pregunta	Nueva respuesta
1	Conforme a la normatividad aplicable, ¿cuáles son los límites para hacer pública la información contenida en su sistema/base de datos?	Respuesta a registrar
2	¿Cuáles son las solicitudes de documentación más comunes asociadas a su sistema/base de datos que enfrentan de manera cotidiana el per-	
3	En una escala del 1 al 10, ¿qué tan precisa considera que es la información contenida en el sistema/base de datos?	
4	¿Con base en qué elementos asigna esta califica-	
5	En caso de que haya respondido a la pregunta 84 afirmativamente, si no reportaron estadísticas de uso, ¿con qué información cuenta?	

VIII. COMENTARIOS FINALES

IX. DESPEDIDA

ANEXO V. COMPOSICIÓN DEL ÍNDICE DE COMPLETITUD

El índice de completitud evalúa la detección de sistemas a través de las 6 áreas temáticas clave del estudio establecidas con base en la LGSNA.

No.	Área temática	Valor
1	Información patrimonial de servidores públicos	N/6
2	Información de servidores públicos que intervienen en procedimientos de contrataciones	N/6
3	Información de servidores públicos sancionados	N/6
4	Información de particulares sancionados	N/6
5	Información de denuncias públicas de faltas administrativas y hechos de corrupción	N/6
6	Información pública de contrataciones	N/6

ANEXO VI. COMPOSICIÓN DEL ÍNDICE DE COMPLEMENTARIEDAD

El eje temático de complementariedad se analizó con un índice que considera 2 dimensiones relevantes.

1. Marco Normativo
2. Documentación del sistema

El análisis de estas dimensiones se apoyó en 9 variables que representan de manera aproximada información valiosa para el cálculo de cada una de las dimensiones señaladas.

No.	Pregunta cuestionario	Parámetros	Valor
1	Indique si existen leyes o reglamentaciones que regulen la existencia del sistema o base de datos.	Marco regulatorio aplicable = 1 Sin marco regulatorio = 0	n/9
2	Indique el nombre del manejador de base de datos que se utiliza.	Base de datos con tratamiento = 1	n/9
3	En caso de no contar con un manejador de base de datos indicar el formato de almacenamiento de la información del sistema o base de datos.	Base de datos sin tratamiento = 0	
4	Por favor, indique si se prevé continuar con la actualización del sistema o base de datos en el futuro.	Se prevé continuar = 1 No se prevé continuar = 0	n/9
5	¿El sistema o base de datos cuenta con diccionario de datos?	Sí = 1 No = 0	n/9
6	En caso afirmativo, indique si el diccionario de datos esta actualizado al momento de responder este cuestionario.	Sí = 1 No = 0	n/9
7	¿Se documentan los metadatos del sistema/base de datos (información sobre los datos, población objetivo, métodos de captura, etc.)?	Si, sigue estándar internacional = 1 Sí, no sigue estándar = 0.5 No = 0	n/9
8	Por favor indique si el sistema o base de datos cuenta con la siguiente documentación.	Diagrama entidad relación=1 No diagrama entidad-relación=0	n/9
9	Por favor indique si el sistema o base de datos cuenta con la siguiente documentación.	Casos de uso=1 No casos de uso=0	n/9

ANEXO VII. COMPOSICIÓN DEL ÍNDICE DE CONFIABILIDAD.

El índice de confiabilidad se evaluó mediante el análisis de 20 reactivos que permiten una mirada a las políticas implementadas alrededor de los repositorios de datos.

No.	Pregunta cuestionario	Parámetros	Valor
1	Indique el modo de generación de la información en su sistema o base de datos	Reporteo = 0 Transaccional = 1	n/20
2	¿Existe normatividad de cualquier nivel (leyes, reglamentos, etc.) que regule la existencia, operación, validación o auditoría de la información del sistema o base de datos?	Sí = 1 No = 0	n/20
3	En caso afirmativo, por favor indique si según la normatividad aplicable, la fuente de información está obligada a entregar los datos requeridos para el sistema o base de datos	Existe un marco legal que obliga entrega de datos = 1 Existe marco legal pero no obliga entrega de datos = .5 No existe marco legal = 0	n/20
4	¿Se realizan evaluaciones de la veracidad de los registros capturados en el sistema o base de datos? (i.e. auditorías)	Sí, se realizan auditorías = 1 Sí, se realizan validaciones de veracidad, pero no auditorías = 1 No se realizan validaciones = 0	n/20
5	¿Cuál es el porcentaje de edición de datos? (modificaciones de variables debido a datos faltantes o errores de inconsistencia de la información)	Más de 50% = 1/5 21-50% = 2/5 16-20% = 3/5 11-15% = 4/5 Menos de 10 = 5/5 No sabe = 0	n/20
6	¿Cuál es el porcentaje de errores de codificación? (errores detectados de captura de datos en el sistema o base de datos)	Más de 50% = 1/5 21-50% = 2/5 11-20% = 3/5 6-10% = 4/5 Menos de 5 = 5/5 No sabe = 0	n/20
7	¿Existen controles establecidos para verificar que la información capturada en el sistema/base de datos sea veraz y precisa?	Sí existen y están documentados = 1 Si existen no están documentados = .5 No existen = 0	n/20

No.	Pregunta cuestionario	Parámetros	Valor
8	¿Existen controles para verificar que la información capturada en el sistema/base de datos siga los parámetros establecidos para su captura? (control de rangos, códigos válidos, etc.)	Sí existen y están documentados = 1 Si existen no están documentados =.5 No existen= 0	n/20
9	¿Existen controles establecidos para verificar que los valores extremos registrados en el sistema/ base de datos o constituyan errores de captura o de algún otro tipo?	Sí existen y están documentados = 1 Si existen no están documentados =.5 No existen= 0	n/20
10	¿En caso de que se detecte un registro con datos inválidos o en blanco, el dueño de la información hace alguna modificación en el sistema/ base de datos?	Sí = 1 No= 0	n/20
11	¿Las modificaciones de los datos del sistema son realizadas de acuerdo con un procedimiento establecido?	Sí existen y están documentados = 1 Si existen, no están documentados =.5 No existen= 0	n/20
12	Indique si la información en su sistema o base de datos se audita con regularidad.	Sí = 1 No= 0	n/20
13	Conforme a la última auditoría, indique el porcentaje de ajustes que sufrió la información de su sistema o base de datos como resultado de la auditoría. El porcentaje de ajustes es el número total de registros modificados sobre el número de registros en el sistema o base de datos.	Más de 50%= 1/5 21-50%= 2/5 16-20%= 3/5 11-15% = 4/5 Menos de 10 = 5/5 No sabe = 0	n/20
14	¿La información contenida en el sistema/base de datos es sometida a un filtrado o tratamiento previo a su ingreso en su recipiente final?	Sí = 1 No= 0	n/20
15	Existe un procedimiento documentado de control de cambios en los procedimientos y/o definiciones?	Sí, se informan cambios = 1 Sí, no se informan cambios = 0.5 No = 0	n/20
16	¿Existe algún procedimiento documentado de captura de los datos del sistema/ base de datos?	Existen procedimientos, están documentados y se tienen procesos que lo aseguran = 1 Existen procedimientos, están documentados pero los procesos no están asegurados = 0.6 Existen procedimientos, no están documentados = 0.3 No existen = 0	n/20

No.	Pregunta cuestionario	Parámetros	Valor
17	¿Existe algún procedimiento documentado de control de consistencia de los datos del sistema/base de datos?	Existen procedimientos, están documentados y se tienen procesos que lo aseguran = 1 Existen procedimientos, están documentados pero los procesos no están asegurados = 0.6 Existen procedimientos, no están documentados = 0.3 No existen = 0	n/20
18	¿Se han documentado los mecanismos para garantizar la integridad referencial de la base de datos del sistema/base de datos?	Sí, totalmente = 1 Sí, parcialmente = 0.5 No = 0	n/20
19	En caso de que la unidad que gestione el sistema/base de datos no sea la fuente de la información, ¿los términos de entrega de datos están documentados y definidos formalmente?	Existen términos, están documentados y se tienen procesos que lo aseguran = 1 Existen términos, están documentados pero los procesos no están asegurados = 0.6 Existen términos, no están documentados = 0.3 No existen = 0	n/20
20	En caso de que la unidad que gestione el sistema/base de datos no sea la fuente de información ¿Existen procedimientos documentados para asegurar la protección física de los datos (mecanismos de control de acceso a servidores, centros de procesamiento de información, etc.)?	Existen procedimientos, están documentados y se tienen procesos que lo aseguran = 1 Existen procedimientos, están documentados pero los procesos no están asegurados = 0.6 Existen procedimientos, no están documentados = 0.3 No existen = 0	n/20