



Global
Partnership
for Sustainable
Development Data



COLLABORATIVE
ON SDG DATA
INTEROPERABILITY

CONECTANDO COMUNIDADES DE DATOS

INTRODUCCIÓN AL MARCO DE EVALUACIÓN DE MADUREZ DE DATOS INTEGRADOS

ÍNDICE

Tabla de Contenido.....	2
Reconocimientos.....	3
Introducción.....	4
Antecedentes y contexto de la evaluación de la madurez de datos integrados.....	4
Cómo utilizar el marco de evaluación de la madurez de los datos integrados.....	5

CAPÍTULO 1

Interoperabilidad organizacional.....	7
1.1 Objetivos estratégicos.....	8
1.2 Liderazgo y gestión.....	9
1.3 Supervisión y rendición de cuentas.....	10
1.4 Conformidad legal.....	11
1.5 Ética de datos.....	12
1.6 Adquisiciones de proveedores externos.....	13
1.7 Vínculos con ecosistemas de datos.....	14

CAPÍTULO 2

Interoperabilidad humana.....	15
2.1 Custodia de datos.....	16
2.2 Preservación de la privacidad y la confidencialidad.....	17
2.3 Conocimientos y habilidades del personal.....	19
2.4 Comunicación interna y externa.....	20
2.5 Adaptabilidad.....	21

CAPÍTULO 3

Interoperabilidad de datos.....	23
3.1 Capacidades para modelación de datos y metadatos.....	24
3.2 Capacidad de organización y clasificación de datos.....	25
3.3 Difusión y apertura de datos.....	26
3.4 Análisis de datos y automatización.....	27
3.5 Protección de datos.....	28

CAPÍTULO 4

Interoperabilidad tecnológica.....	30
4.1 Infraestructura de datos.....	31
4.2 Seguridad cibernética y respuesta ante incidentes.....	32

Apéndice A: El marco de evaluación de madurez de datos integrados.....	34
---	-----------

Apéndice B: Glosario.....	43
----------------------------------	-----------

Bibliografía.....	44
--------------------------	-----------

RECONOCIMIENTOS

El marco de evaluación de madurez de datos integrados y su guía introductoria fueron escritos por Tom Orrell, director gerente de *DataReady Limited*, en nombre de la Alianza Global de Datos para el Desarrollo Sostenible, y Luis Gerardo González Morales, estadístico de la División de Estadística de las Naciones Unidas. La Guía contó con el apoyo de todos los miembros del Consorcio colaborativo sobre interoperabilidad de datos de los ODS, a quienes los autores expresan su gratitud. Los autores desean agradecer en particular a los miembros del Grupo asesor del consorcio colaborativo que apoyaron la producción del Marco de evaluación de madurez, incluyendo a Eric Swanson de *Open Data Watch*; Josh Powell y Paige Kirby de *Development Gateway*; Karen Carter del *Fondo de las Naciones Unidas para la Infancia*; Malarvizhi Veerappan del *Banco Mundial*; Mahsa Hedayati de la *Oficina de Tecnología de Información y Comunicaciones de las Naciones Unidas*; y Neil Jackson de la *Oficina de Relaciones Exteriores, Commonwealth y Desarrollo del Reino Unido*.

Los autores también quisieran extender un agradecimiento especial a Bill Anderson de *Development Initiatives*, cuya visión ayudó a establecer el tema interoperabilidad en el mapa de políticas de la revolución de los datos; Shaida Badiee y todo el equipo de *Open Data Watch*, quienes han apoyado continuamente este trabajo; Manuel Cuéllar y Enrique Jesus Ordaz López del *Instituto Nacional de Estadística y Geografía de México*, por su continuo apoyo y retroalimentación; Fredy Rodriguez en *CEPEI* (Centro de Pensamiento Estratégico Internacional); Rachael Beavan, Edafe Onerhime y sus colegas de la *Oficina de Relaciones Exteriores, Commonwealth y Desarrollo*; el *Grupo de Trabajo de Datos Abiertos de la Comisión de Estadística de las Naciones Unidas*; y Francesca Perucci de la *División de Estadística de Naciones Unidas* y Claire Melamed de la *Alianza Global de Datos para el Desarrollo Sostenible* y sus equipos, por su apoyo al Consorcio colaborativo sobre interoperabilidad de datos de los ODS durante los últimos cuatro años.

INTRODUCCIÓN

Antecedentes y contexto de la evaluación de la madurez de datos integrados

La interoperabilidad es la capacidad de unir flujos de datos de manera que tanto las máquinas como los humanos comprendan y contextualicen su contenido. Es gracias a la interoperabilidad que uno puede usar las herramientas de Microsoft Office en un iPad de Apple, enviar y recibir correos electrónicos de una variedad de proveedores de correo electrónico, o colaborar de forma remota con colegas en la misma hoja de cálculo utilizando una multitud de dispositivos. En el marco de la revolución de los datos para el desarrollo sostenible, la interoperabilidad permite superponer datos de observación de la tierra, datos administrativos o datos derivados de dispositivos móviles con datos estadísticos, lo que ayuda a lograr y monitorear el progreso hacia los Objetivos de Desarrollo Sostenible (ODS). En última instancia, la interoperabilidad permite a los tomadores de decisiones integrar y contextualizar una variedad de datos de manera que se maximice su utilidad.

Desde 2017, la Alianza Global para Datos de Desarrollo Sostenible (la Alianza Global) y la División de Estadística de Naciones Unidas han gestionado y mantenido conjuntamente un Consorcio colaborativo sobre interoperabilidad de datos de los ODS (el Colaborativo). En 2018, el consorcio colaborativo publicó dos documentos: la nota titulada “Uso de datos para Unir esfuerzos de desarrollo” (la Nota) y la guía “Interoperabilidad de datos: una guía para profesionales para unir datos en el sector del desarrollo” (la Guía). En conjunto, estos dos documentos proporcionan una descripción general de lo que la interoperabilidad significa para los estadísticos, los profesionales del desarrollo que administran datos y los especialistas en TI que trabajan en el desarrollo sostenible. Entre ellos, la Nota y la Guía definen y exploran cinco pilares de interoperabilidad identificados por el Colaborativo:

- Gestión de datos, gobernanza e interoperabilidad
- Modelos canónicos de datos y metadatos
- Clasificaciones y vocabularios
- Interfaces estandarizadas
- Implementación de enfoques de datos vinculados

Este construye sobre los fundamentos establecidos por la Guía y la Nota, e introduce el Marco de evaluación de la madurez de los datos integrados (el Marco de evaluación de madurez) desarrollado por el Colaborativo en 2019-2020 (el Anexo A contiene el Marco de evaluación de madurez completo). La Guía y la Nota fueron avalados por la Comisión de Estadística de Naciones Unidas en su 50a sesión en 2019, y ahora están comisariados por el Grupo de Trabajo de la Comisión sobre Datos Abiertos.

Cómo utilizar el marco de evaluación de la madurez de los datos integrados

El marco de evaluación de madurez presentado en este documento está diseñado para ser utilizado por estadísticos oficiales y profesionales que regularmente producen, cotejan o emplean diversos conjuntos de datos en el sector del desarrollo sostenible. Éste se basa en los conceptos definidos y establecidos en la Nota y en la Guía, y está diseñado para ser utilizado con fines de planificación estratégica y monitoreo.

El marco de evaluación de madurez es el resultado de un análisis de varios **modelos de madurez**, de uso común, incluidos, entre otros:

- El "Modelo de gobierno de la información" de IBM (IBM 2007)
- El "Modelo de madurez para datos y analítica" de Gartner (Gartner 2018)
- El "Modelo responsable de madurez de datos para organizaciones humanitarias y de desarrollo" de CARE USA (CARE USA 2019)
- El "Modelo de madurez de datos abiertos" del Open Data Institute (ODI 2015)

El Marco de evaluación de madurez se basa en elementos de todos estos modelos e intenta crear una herramienta adaptable y flexible que pueda ser utilizada por las oficinas nacionales de estadística (ONEs) y otras entidades que controlan o procesan datos en el sector de desarrollo.

El marco de evaluación de madurez (ver Anexo A) tiene tres componentes: estratos de interoperabilidad, dimensiones y niveles de madurez. Comienza con los cuatro estratos de interoperabilidad presentados en la Nota y en la Guía, que son, siguiendo la conceptualización original de John Palfrey y Urs Gasser en *Interop: The Promise and Perils of Highly Interconnected Systems* (Palfrey y Gasser 2012):



Estos cuatro estratos tienen un total de 19 dimensiones, o funciones de gestión de datos que corresponden a cada uno de ellos. La Evaluación de la madurez luego identifica las características que se encuentran en cada dimensión del nivel de madurez de una organización, empezando por el nivel de *indefinición* y progresando en madurez a los niveles de *iniciación*, *aprendizaje*, *desarrollo*

El Marco de evaluación de madurez está diseñado para usarse de manera integral o parcial. Algunas de las circunstancias en las que esta herramienta podría usarse incluyen las siguientes:

- Cuando se produce un nuevo marco o estrategia de gestión de datos en una ONE, una organización internacional o un departamento gubernamental que produce, recopila o utiliza diversos conjuntos de datos.
- Durante una revisión estratégica o actualización de un marco de gestión de datos o de un proceso operativo del negocio, lo que puede abarcar desde la formulación de una estrategia nacional de desarrollo estadístico hasta la adopción del modelo genérico de procesos de negocio estadísticos.
- Como parte de una evaluación más amplia de la madurez de la gobernanza de datos por parte de una organización internacional productora y gestidora de datos, con el fin de fortalecer su capacidad de gestión de datos.

y consolidación. Estos niveles de madurez no son mutuamente excluyentes; es totalmente plausible que un usuario esté consolidando el progreso en una dimensión, pero solo iniciando en otras. Maturity Assessment can be found in Annex A and a Glossary is located in Annex B.

Los usuarios que deseen obtener más información sobre el valor de la interoperabilidad y su relevancia para los procesos de desarrollo sostenible deben consultar la Nota y la Guía antes de utilizar este documento.

El resto de este documento se divide en cuatro capítulos, cada uno de los cuales explora las dimensiones de la Evaluación de madurez en una de las cuatro capas de interoperabilidad.

Cada dimensión está explicada y vinculada a secciones relevantes de este documento y partes de la Guía. Asimismo, se proporcionan otros recursos y ejemplos según sea necesario. El marco de evaluación de madurez completo se puede encontrar en el Anexo A; un Glosario se encuentra en el Anexo B.



CAPÍTULO 1

INTEROPERABILIDAD ORGANIZACIONAL

La interoperabilidad organizacional se puede interpretar de dos formas. Por un lado, puede referirse a cómo se distribuyen en una organización las funciones de **gobernanza** y **gestión de datos** que afectan a la interoperabilidad.¹ Alternativamente, puede verse como las formas en que organizaciones enteras, o departamentos individuales dentro de ellas, se involucran con el ecosistema de datos más amplio o el sistema estadístico nacional para decidir el grado de interoperabilidad que les gustaría lograr entre sus activos de datos colectivos. La Evaluación de Madurez identifica siete dimensiones de interoperabilidad organizacional. Estas dimensiones y sus características asociadas para cada nivel de madurez se describen en las Tablas 1.1 a 1.7.

¹ Para los propósitos de este documento, hemos seguido las definiciones de gobernanza y gestión de datos en “Data Management Body of Knowledge” (DAMA International 2017). Las definiciones para los dos términos se encuentran en el Anexo B, y una descripción general del concepto de datos, la gobernanza se establece en la Guía (González y Orrell 2018), p. 19.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
La interoperabilidad no se reconoce como un objetivo estratégico	Se reconoce la importancia de la capacidad de unir datos, pero ésta no se identifica explícitamente como un objetivo estratégico	La interoperabilidad se identifica como un objetivo estratégico dentro de las unidades técnicas de una organización, pero no fuera de ellas	La necesidad de unir datos entre sistemas se reconoce como un objetivo estratégico dentro de una estrategia de datos organizacionales. Se reconoce el valor de contar con estándares y un gobierno de datos sólido.	En las estrategias organizacionales se reconoce el valor que la integración de datos puede conferir a la toma de decisiones. La interoperabilidad forma parte de la estrategia de compromiso externo de una organización con otros productores y usuarios de datos and users

1.1 Objetivos estratégicos

La gestión de datos puede definirse en términos generales como el desarrollo, la implementación y el seguimiento de estrategias y planes que permiten aprovechar el valor de los datos de forma segura. Por lo tanto, el desarrollar una estrategia de administración de datos que establezca opciones y decisiones para trazar un curso de acción es una función central para toda organización gestora de datos (DAMA International 2017, págs. 17 y 31).

Para las organizaciones que administran, recopilan o procesan diversos conjuntos de datos, la interoperabilidad debe ser un principio clave en su estrategia de gestión de datos. Las decisiones que determinan qué conjuntos de datos deben ser interoperables y en qué grado², el nivel deseado de interoperabilidad en relación con el resto del ecosistema de datos, y otras consideraciones

similares, deben ser capturadas en la estrategia de gestión de datos.

En los últimos años, el valor de la interoperabilidad como objetivo estratégico ha sido cada vez más reconocido en el ámbito de las estadísticas oficiales. Por ejemplo, desde 2018, la ONE de México, el Instituto Nacional de Estadística y Geografía (INEGI), ha estado trabajando en una estrategia de arquitectura y gobernanza de la información. La interoperabilidad se reconoce como uno de los tres pilares de la estrategia, junto con la calidad, seguridad y confidencialidad de los datos.

Al establecer objetivos estratégicos relacionados con la interoperabilidad, es importante asignar recursos presupuestarios y humanos en los planes de implementación que los acompañan.

² Ver pág. 11 de la Guía (González y Orrell 2018) para una discusión sobre la interoperabilidad óptima.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
No existe un liderazgo definido sobre cuestiones de interoperabilidad	El liderazgo ad hoc en temas de interoperabilidad surge orgánicamente dentro de una organización, pero no está coordinado	El liderazgo en torno a la interoperabilidad emerge a través de varias unidades técnicas pero permanece fragmentado Los silos persisten	Existe una jerarquía coordinada de liderazgo sobre cuestiones de interoperabilidad Se han establecido funciones claras relacionadas con la interoperabilidad en toda la organización.	Existe un comité o consejo de gobernanza de datos con un mandato explícito para liderar en asuntos relacionados con la interoperabilidad. Los líderes y gerentes de la organización entienden el valor de los datos unidos, y la integración de datos se identifica claramente como una función en las descripciones de trabajo relevantes

1.2 Liderazgo y gestión

El Manifiesto de datos del líder (Data Leaders 2017), elaborado por una coalición de líderes de opinión en el campo de la gestión de datos, reconoce que

“Los datos ofrecen un enorme potencial sin explotar para crear una ventaja competitiva, nueva riqueza y puestos de trabajo; mejorar la asistencia sanitaria; mantenernos a todos más seguros; y de otra manera mejorar la condición humana [...] todavía no encontramos ejemplos de cambio fundamental duradero sin un liderazgo comprometido y la participación de todos en todos los niveles de la organización”.
(Data Leaders 2017)

Por lo tanto, el liderazgo es crucial para la gestión de datos y, por extensión, es fundamental tanto para las ONE en su papel de gestoras de datos (sección 2.1) como para garantizar que la interoperabilidad sea tratada como un objetivo estratégico (sección 1.1). Para que las organizaciones tomen conciencia del potencial de sus activos de datos, la alta dirección debe participar en procesos estratégicos que se relacionan con la interoperabilidad de datos y valorar el aporte de la interoperabilidad a sus organizaciones. En organizaciones más grandes, el liderazgo sobre cuestiones de interoperabilidad se estructura con frecuencia en varios niveles de gestión y se manifiesta en forma de comités o consejos de gobernanza de datos, que también son fundamentales para la supervisión y la rendición de cuentas (DAMA International 2017) (sección 1.3).

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
No hay supervisión ni rendimiento de cuentas sobre los problemas de interoperabilidad	La organización es consciente de la necesidad tener cadenas de responsabilidad para supervisar cómo se unen los datos, pero aún no se toman medidas para establecerlas.	Las estructuras de supervisión de la interoperabilidad emergen a través de diferentes unidades técnicas pero no están coordinadas o alineadas La responsabilidad sobre cómo se realizan los esfuerzos de interoperabilidad es fragmentada y poco clara	Las funciones de supervisión y rendición de cuentas sobre interoperabilidad están integradas en la estrategia de una organización y se reflejan en la estructura de liderazgo. Las cadenas fragmentadas de supervisión y rendimiento de cuentas respecto a la interoperabilidad comienzan a unirse y surgen procedimientos operativos estándar comunes.	Existe un comité o consejo de gobernanza de datos con un mandato explícito para liderar en asuntos relacionados con la interoperabilidad. Los líderes y gerentes de la organización entienden el valor de los datos unidos, y la integración de datos se identifica claramente como una función en las descripciones de trabajo relevantes

1.3 Supervisión y rendición de cuentas

Las estructuras de supervisión y rendición de cuentas son principios centrales de la gobernanza de datos y de las funciones de gestión relacionadas con el establecimiento de objetivos estratégicos (sección 1.1), liderazgo y gestión (sección 1.2) y administración de datos (sección 2.1). Asegurarse de que la supervisión y la rendición de cuentas efectivas estén integrados en los procesos organizativos ayuda a garantizar que los procesos para hacer que los datos sean interoperables y, por lo tanto, integrables, compatibles y accesibles para otros, se gestionen y mantengan correctamente. También es necesario supervisar los conjuntos de datos que pueden contener información que

revele la identidad de una persona u otros atributos confidenciales sobre personas o grupos vulnerables (secciones 2.2 y 3.5).

La supervisión y la rendición de cuentas pueden ser internas o externas. La supervisión y la responsabilidad internas se pueden estructurar en torno a consejos y comités de gobernanza de datos³, y se pueden integrar en las descripciones de funciones de los miembros del personal. La responsabilidad externa incluye garantizar que las acciones de una organización cumplan con las regulaciones legales aplicables, las pautas éticas obligatorias y las expectativas de los socios.

³ Consúltense las páginas 18 y 19 de la Guía (González y Orrell 2018) para ver una discusión sobre los consejos y comités de gobernanza de datos.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
No hay conciencia sobre las obligaciones legales aplicables relacionadas con la integración de datos interoperables	Existe un entendimiento general de que las acciones que facilita la interoperabilidad (transmisión, intercambio y (re) utilización de datos) podrían estar reguladas, pero no está claro cómo	La organización toma medidas activas para comprender mejor los requisitos de conformidad legal en relación con la retención, la transmisión, el intercambio y la utilización de datos, y los pone a disposición de los usuarios de datos.	El cumplimiento de las leyes aplicables sobre intercambio, transmisión y utilización de datos está integrado en las funciones de supervisión y rendimiento de cuentas, y se refleja en la estrategia de datos de la organización, la cual está publicada en línea	Las actividades de transmisión, intercambio y utilización de datos de la organización cumplen plenamente con las leyes aplicables y, a veces, superan los estándares legales

1.4 Conformidad legal

Pueden surgir desafíos legales y regulatorios para la interoperabilidad cuando se trata del intercambio e integración de activos de datos entre organizaciones y a través de las fronteras nacionales. Las leyes establecen límites aplicables entre lo que es una conducta aceptable y lo que no lo es. En algunos casos, éstas rigen cómo se pueden compartir los datos (por ejemplo, las leyes que regulan y establecen estándares para el reporte, la seguridad y la protección de datos) y, en otros, rigen qué datos pueden, o más a menudo no pueden, compartirse e integrarse (por ejemplo, las leyes de protección de datos y privacidad).⁴ Por lo tanto, las cuestiones legales y reglamentarias están estrechamente relacionadas tanto con los derechos de las personas sobre los datos (sección 2.2) como con los deberes de las organizaciones de proteger los datos (sección 3.5).

La transmisión internacional de datos puede ser especialmente compleja, como destaca el caso *Data Protection Commissioner v Facebook Ireland and Maximillian Schrems* (ECJ 2020).. En este caso, el tribunal determinó que el mecanismo principal que

permitía el intercambio de datos personales entre la Unión Europea y los Estados Unidos de América era deficiente porque no garantizaba la privacidad y protección de los datos de los ciudadanos de la Unión Europea. Esto tiene un claro efecto dominó sobre la capacidad de las entidades en estas dos jurisdicciones para integrar conjuntos de datos que contienen datos personales. Este caso puede requerir que las entidades limiten la capacidad de ciertos tipos de datos bajo su control para ser interoperables con otros sistemas.

Los problemas legales y de cumplimiento relacionados con la interoperabilidad de datos son complejos y multifacéticos. Frecuentemente se requiere la participación de expertos legales, para establecer, caso por caso, los posibles efectos de leyes particulares en jurisdicciones específicas sobre el deseo de una organización de permitir la interoperabilidad de conjuntos de datos que contienen información personal o sensible. Si un proyecto implica la transmisión transfronteriza de datos personales o confidenciales, es probable que se brinde asesoramiento legal. necesarios en la fase de diseño para garantizar que las actividades cumplan con las leyes aplicables.

⁴ See p. 19 and Annex B of the Guide (Gonzalez and Orrell 2018) for further information on legal and regulatory frameworks.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
No hay conciencia sobre los riesgos éticos que la interoperabilidad de datos podría representar	Existe un entendimiento general de que la unión de datos a veces puede presentar riesgos éticos, pero no está claro cómo	La organización toma medidas activas para comprender las consecuencias éticas que puede ocasionar la integración de datos y para comprender como ésta puede ocasionar daños no intencionales Se llevan a cabo evaluaciones de impacto ético de manera rudimentaria y ad hoc	Se comprenden y categorizan adecuadamente en la organización los diferentes tipos cuestiones éticas que puede generar la integración de datos Se toman las medidas adecuadas para minimizar el daños causados por violaciones a los estándares éticos Se llevan a cabo evaluaciones de impacto ético de manera rutinaria	Se comprenden bien los riesgos que representa la integración de datos y se realizan revisiones éticas a lo largo del ciclo de vida de los datos para monitorear problemas y corregir el curso, según sea necesario. Las evaluaciones de impacto ético se publican en línea y de forma transparente La organización solo integra datos una vez que ha emprendido y documentado una revisión de los riesgos potenciales y ha tomado las medidas adecuadas para mitigarlos.

1.5 Ética de datos

La interoperabilidad puede dar lugar a dilemas éticos. Estos difieren de las cuestiones legales en que es probable que se relacionen con asuntos de equidad o lo que se considera correcto o incorrecto en un contexto particular, en contraposición con lo legal o ilegal. Una cuestión de interoperabilidad ética podría ser si las empresas internacionales de procesamiento de datos que obtienen acceso a microdatos demográficos nacionales como parte de un programa de desarrollo deberían tener derecho a reutilizar y beneficiarse de esos datos en el futuro, integrándolos con otros conjuntos de datos. Dependiendo de las leyes nacionales de una jurisdicción en particular, esto podría ser legal, pero aún plantea cuestiones éticas sobre la distribución equitativa del valor extraído de los activos nacionales de datos.

En algunos países, se están adoptando marcos de ética de datos para establecer principios operativos y acciones que los servidores públicos pueden tomar para resolver dilemas éticos.

El "Marco de ética de datos" del Reino Unido (Gobierno del Reino Unido, 2020) establece una guía concisa y clara sobre la ética de datos y sirve como un ejemplo útil para otros. Grupos de la sociedad civil como el Open Data Institute también han producido herramientas fáciles de usar, en este caso el "Data Ethics Canvas" (ODI 2019a), para ayudar a las organizaciones a pensar en las implicaciones éticas de su trabajo.

A veces, las cuestiones legales y éticas pueden fusionarse y la ruta para salir de un dilema ético puede ser legal. Teniendo en cuenta el ejemplo anterior, los regímenes de concesión de licencias sólidos que limitan la forma en que se pueden reutilizar y reintegrar los datos podrían ofrecer una solución a ese desafío ético específico (sección 1.7).

Es una buena práctica evaluar periódicamente los problemas éticos relacionados con conjuntos de datos específicos (incluidos los problemas relacionados con la interoperabilidad), documentarlos y considerar los pasos para mitigarlos o resolverlos.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
La organización no es consciente del impacto de la adquisición de soluciones técnicas de interoperabilidad	<p>Existe un nivel emergente de entendimiento sobre la necesidad de integrar datos en todas las actividades relacionadas con adquisiciones de proveedores externos, pero no existe un enfoque coherente</p> <p>Existe una gran dependencia hacia contratistas externos para solventar carencias, pero no hay coordinación entre diferentes unidades de la organización, lo que a menudo resulta en la adquisición de soluciones para el manejo de datos que son incompatibles.</p>	<p>La coordinación entre diferentes unidades organizacionales en torno a la adquisición de soluciones técnicas de proveedores externos comienza a materializarse, pero aún no está formalizada</p> <p>Ocasionalmente, diferentes unidades coordinan informalmente la contratación de proveedores externos para garantizar que los nuevos sistemas de datos sean compatibles con la infraestructura de datos existente.</p>	<p>La adquisición de sistemas de datos compatibles e interoperables en una organización se formaliza y coordina</p> <p>Existe una política común de adquisiciones de proveedores externos en toda la organización, la cual requiere que el personal considere los problemas de interoperabilidad al adquirir nuevos sistemas.</p> <p>La dependencia de proveedores externos es estratégica y coordinada</p>	<p>La organización integra la adquisición de nuevos sistemas y servicios de datos en su estrategia de datos, e incluye planes a futuro</p> <p>Las diferentes unidades planifican y analizan estratégicamente sus necesidades comunes de adquisición de soluciones, y se aseguran de que cualquier sistema o servicio nuevo de datos sea compatible con la infraestructura existente y satisfaga las probables necesidades futuras.</p>

1.6 Adquisiciones de proveedores externos

La adquisición de servicios de TI, sistemas de datos u otras soluciones relacionadas con el procesamiento, el almacenamiento o el intercambio de datos plantean problemas importantes de interoperabilidad para las organizaciones.

Numerosas ONEs y ministerios, departamentos y agencias gubernamentales de todo el mundo han informado de manera anecdótica que la **dependencia hacia proveedores exclusivos** es un serio obstáculo para la interoperabilidad. La dependencia hacia proveedores exclusivos se produce cuando las organizaciones adquieren por separado diferentes soluciones patentadas de datos y de informática como parte de diferentes programas de trabajo o proyectos patrocinados por donantes, sin tener en cuenta la arquitectura de datos o la infraestructura digital existentes. Esto puede crear sistemas de información aislados que

produzcan conjuntos de datos que no interoperen o cuyas salidas de datos no se puedan integrar entre sí.

Para evitar la dependencia hacia proveedores exclusivos, las organizaciones deben utilizar un enfoque coordinado y un conjunto de pautas comunes entre sus departamentos, de manera que consideren la interoperabilidad al adquirir nuevos sistemas digitales o de datos. Los departamentos del gobierno del Reino Unido, por ejemplo, han adoptado un conjunto de principios de estándares abiertos que establecen: “Los marcos gubernamentales para la adquisición de TI deben: especificar el uso de estándares abiertos para la interoperabilidad de software, formatos de datos y documentos, a menos que haya una razón clara por la que esto no sea posible” (Gobierno del Reino Unido, 2018). Siempre que sea posible, se deben utilizar estándares de código abierto y paquetes de software para minimizar los riesgos asociados con la dependencia hacia proveedores exclusivos.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
<p>No hay conciencia sobre la forma en que se utilizan los datos en el ecosistema de datos y el papel que en ello juega la interoperabilidad de datos</p>	<p>Existe un entendimiento emergente de que la integración de datos entre las diferentes entidades que conforman un ecosistema de datos pueden generar tanto oportunidades como desafíos</p> <p>La organización interactúa con otros sectores del ecosistema de datos de manera informal y ad hoc</p>	<p>La organización comienza a adjuntar términos de licencia pro forma a los datos que publica, transmite o comparte, los cuales incluyen disposiciones sobre su integración con otros datos, pero no supervisa ni interactúa con los usuarios para garantizar su aplicación.</p> <p>La organización comienza a documentar los datos que recibe de otras organizaciones.</p> <p>La organización comienza a documentar y coordinar su involucramiento con otras partes del ecosistema de datos.</p>	<p>La organización clasifica efectivamente sus datos y los licencia para su adecuada utilización</p> <p>La organización documenta todos los datos que se comparten con ella y tiene entendimiento general de lo que se puede y lo que no se puede hacer con ellos</p> <p>La organización se involucra con otras partes del ecosistema de datos de manera coordinada, en conformidad con su estrategia de datos</p>	<p>La organización tiene un conjunto de licencias bien establecido y adaptado a sus propias necesidades, el cual define parámetros claros para la utilización de datos, incluida su integración, según el tipo de datos en cuestión</p> <p>La organización documenta todos los datos que se comparten con ella y cuenta con una guía y procedimientos claros que rigen si esos datos se pueden unir y cómo se pueden unir con otros conjuntos de datos bajo su control.</p> <p>La organización define como prioridad estratégica el involucramiento con otras partes del ecosistema de datos, y tiene un enfoque bien coordinado y con procesos claros para integrar sus datos con datos externos.</p>

1.7 Vínculos con ecosistemas de datos

Las ONEs modernas, y de hecho, muchos otros departamentos gubernamentales, organizaciones internacionales (incluidas las agencias custodias de los ODS), actores del sector privado y organizaciones no gubernamentales, pueden verse como parte de un ecosistema de datos común, una red de organizaciones interconectadas, sistemas de datos, infraestructura digital y aplicaciones.⁵

La forma en que las organizaciones eligen interactuar con otras entidades dentro de sus

propios ecosistemas de datos es una decisión estratégica con importantes consideraciones de interoperabilidad. La Sección 3.3 del marco de evaluación de madurez considera varios aspectos técnicos en juego. Los acuerdos que las organizaciones establecen con otras partes del ecosistema de datos son especialmente importantes para garantizar que cualquier tipo de interoperabilidad está pensado y documentado estratégicamente. Los acuerdos de licencia de datos (documentos legales que especifican lo que un usuario puede hacer y no hacer con los datos) son un tipo de herramienta para hacer esto.⁶

⁵ Tanto *Development Gateway* como *Open Data Watch* han producido una guía útil sobre la estructura y el valor de los ecosistemas de datos. Ver, por ejemplo, *Understanding National Data Ecosystems* (Development Gateway 2019) y *The State of Development Data Funding* (ODW 2016, p. 6).

⁶ Antes de la Comisión de Estadística de Naciones Unidas en 2020, el Grupo de Trabajo sobre Datos Abiertos elaboró un reporte titulado "Orientación sobre la implementación de datos abiertos en las oficinas nacionales de estadística" (UNSD 2020b), que incluye orientación sobre el papel de las licencias de datos abiertos en la Parte II.



CAPÍTULO 2

INTEROPERABILIDAD HUMANA

La interoperabilidad humana puede referirse a dos cosas. Por un lado, se refiere a la necesidad de garantizar que, a medida que los datos se hacen interoperables entre diferentes sistemas, aquéllos sigan siendo legibles y utilizables por usuarios humanos. Por otro lado, se refiere a la capacidad de los individuos, grupos, equipos y departamentos de individuos dentro de las organizaciones para poder comunicarse y trabajar juntos de manera que fomenten la interoperabilidad entre sus flujos de trabajo y activos de datos. La evaluación de madurez identifica cinco dimensiones de la interoperabilidad humana. Estas dimensiones y sus características asociadas para cada nivel de madurez se describen en los Cuadros 2.1 a 2.5.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
No se ha asignado ningún miembro del personal al rol de custodio de datos	La integración de datos forma parte de la función de ciertos miembros del personal, pero no se refleja en sus descripciones de trabajo y se mantiene como una actividad ad hoc	El personal de diferentes unidades tiene explícitamente funciones relacionadas con la interoperabilidad de datos, pero hay poca (o ninguna) coordinación central	Existe un plan coordinado de reclutamiento de personal que refleja las diversas dimensiones y roles relacionados con la custodia de los datos, incluida la interoperabilidad, en toda la organización Las diferentes unidades se coordinan y comunican entre sí qué datos están uniendo, cómo y por qué	Un consejo o comité de gobierno de datos de la organización supervisa un plan estratégico para la gestión de datos que incluye un plan claro para la custodia de datos, incluidas las funciones de interoperabilidad El personal de toda la organización es consciente de cómo se usan, unen y comparten los datos con otras entidades.

2.1 Custodia de datos

En el campo de la gestión de datos, la **custodia de datos** es la “práctica de administrar activos de datos en nombre de otros y en el mejor interés de la organización” (DAMA International 2017, p. 76). La custodia de datos eficaz y responsable es clave para la interoperabilidad. A menudo los custodios de datos son los encargados de implementar los procedimientos y prácticas que permiten una mayor interoperabilidad.

Los custodios de datos suelen ser miembros individuales del personal, designados estratégicamente en la organización para supervisar las funciones clave de gestión de datos en las que tienen una experiencia sustancial. Por ejemplo, el custodio de datos en jefe puede presidir consejos o comités de gobierno de datos (secciones 1.2 y 1.3), mientras que los custodios de datos técnicos pueden ser profesionales de TI que supervisan la integración de datos y demás en departamentos específicos. No es raro que las organizaciones operen jerarquías de custodia de datos con un consejo de gobierno de datos de alto nivel que delega cuestiones específicas a los jefes de departamento, quienes a su vez designan custodios de datos para implementar las decisiones del

consejo de gobierno de datos.

Dentro de la revolución de los datos para el desarrollo sostenible, el concepto de custodia de datos está siendo abordado en términos más amplios. Por ejemplo, The GovLab, con sede en la Universidad de Nueva York, ha realizado una investigación exhaustiva sobre el papel de la custodia de datos en las colaboraciones de datos y ahora mantiene un Red de Custodios de Datos (The GovLab 2020). Además, en el campo específico de las estadísticas oficiales, la custodia de datos fue el tema central del Foro de alto nivel sobre estadísticas oficiales de 2020 que precedió a la sesión anual de la Comisión de Estadística de Naciones Unidas. En la comunidad de estadísticas oficiales, el concepto de custodia de datos también se está ampliando, pasando de su aplicación a las funciones de gestión de datos individuales a la exploración del papel de las ONE como custodios nacionales de datos que coordinan todo el sistema estadístico nacional y de las entidades del ecosistema general de datos a nivel nacional (sección 1.7) (UNSD 2020a). La definición y el alcance de la custodia de datos evolucionan continuamente, a medida que las diferentes áreas involucradas adoptan enfoques de acuerdo con sus necesidades particulares.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
Hay poca o ninguna comprensión de los riesgos para la privacidad y la necesidad de preservar la confidencialidad en los conjuntos de datos interoperables	Existe un entendimiento emergente de los riesgos que representa para los individuos o grupos vulnerables el integrar distintos conjuntos de datos.	Existe conocimiento de las leyes (internacionales) de privacidad y confidencialidad aplicables, así como de principios los normativos, buenas prácticas y lineamientos relevantes, pero no se consideran ni se los siguen de manera rutinaria cuando los activos de datos de una organización se integran con otros datos o se utilizan de alguna otra manera.	Se considera de forma rutinaria las leyes, principios, buenas prácticas y lineamientos aplicables La organización realiza evaluaciones de impacto en la privacidad antes y durante los proyectos relacionados con datos, y esas evaluaciones incluyen consideraciones sobre los riesgos asociados con la interoperabilidad de los datos.	La preservación de la privacidad individual y la confidencialidad de los datos forman parte de las evaluaciones legales y éticas de la organización, y están integrados en todo el ciclo de vida de los datos. La organización se adhiere a los más altos estándares de privacidad y preservación de la confidencialidad, independientemente de los requerimientos legales locales. La organización ha integrado la preservación de la privacidad y la confidencialidad en su estrategia de datos y proporciona explícitamente orientación sobre los riesgos de los datos interoperables (incluyendo, por ejemplo, el "efecto mosaico") La organización tiene una visión de largo plazo, y conoce los riesgos potenciales de privacidad inherentes a la interoperabilidad de datos que plantean tecnologías emergentes tales como la Internet de las cosas, la verificación biométrica de identidad o los procesos automatizados generales

2.2 Preservación de la privacidad y la confidencialidad

Establecer la interoperabilidad entre sistemas requiere un cuidadoso equilibrio, para determinar, por una parte, qué datos son susceptibles de abrirse, compartirse o integrarse y, por otra, qué datos no lo son. Existen numerosas razones por las que una organización podría no querer que sus datos sean interoperables con otros conjuntos de

datos, pero una razón clave es el deseo de mantener la confidencialidad y proteger el derecho a la privacidad que las partes interesadas tienen sobre sus datos personales y sensibles.⁷

La necesidad de preservar la confidencialidad y respetar el derecho de las personas a la privacidad está bien establecida. El sexto principio fundamental de las estadísticas oficiales estipula que “los datos individuales recopilados por las agencias estadísticas para la compilación estadística, ya se refieran a personas físicas o jurídicas, serán estrictamente confidenciales y se

⁷ Véase, por ejemplo, el libro de Helen Nissenbaum (disponible en línea), *Privacy in Context: Technology, Policy and the Integrity of Social Life*.(2010) para más información

utilizarán exclusivamente con fines estadísticos” (ECOSOC 2013). La Declaración Universal de Derechos Humanos de 1948 y el Pacto Internacional de Derechos Civiles y Políticos de 1966 establecen el derecho a la privacidad de todas las personas, incluida la información relacionada con ellas.

En el caso de las organizaciones estadísticas, los registros individuales recopilados a partir de encuestas, censos, fuentes administrativas, proveedores de telecomunicaciones, etc., son muy sensibles. El contar con protocolos sólidos de privacidad y seguridad de datos es particularmente importante cuando se vinculan registros de diferentes fuentes a través de identificadores comunes de individuos, hogares, empresas o geografías en el proceso de producción de estadísticas oficiales. Esto también es fundamental para evitar cualquier daño a la reputación y mantener la confianza de todas las partes interesadas.

En el marco de la revolución de los datos para el desarrollo sostenible, y considerando el papel de las ONEs como actores clave en los ecosistemas de datos nacionales (sección 1.7) y como custodios de datos (sección 2.1), existen investigaciones en marcha para entender los límites entre la capacidad de las ONEs para abrir los datos para uso público y la necesidad de preservar la confidencialidad. En 2019, la Red de Investigación Temática sobre Datos y Estadísticas de la UN Sustainable Development Solutions Network publicó el documento “Maximizing Access to Public Data: Striking the Balance Between “Open by Default” and Targeted Data Sharing” (TReNDS 2019), el cual explora estos temas con más detalle. Además, la publicación de UN Global Pulse titulada “Risks, Harms and Benefits Assessment” (UN Global Pulse 2020) proporciona un recurso útil para las entidades que buscan mapear los riesgos asociados con sus conjuntos de datos.

Otro tema por considerar es el concepto de **datos de grupo sensibles**. Los datos de grupos sensibles son observaciones que si bien pueden no representar un riesgo para las personas si llegan a ser divulgados, sí constituyen un riesgo potencial

para grupos vulnerables. Un ejemplo podría ser la identificación potencial, en conjuntos de datos de observación de la tierra, de hogares de pueblos indígenas que habitan en los bosques. Aunque los miembros individuales de un grupo pueden no ser identificables, los datos pueden identificar a qué grupo pertenece un miembro. Si estos datos cayeran en manos de madereros criminales, por ejemplo, el grupo podría quedar expuesto a daños. Los Principios de Naciones Unidas sobre Protección y Privacidad de Datos Personales (Naciones Unidas, 2020) reconocen que la preservación de la confidencialidad puede aplicarse a datos no personales en ciertas circunstancias, particularmente “en un contexto sensible que puede poner a ciertas personas o grupos de personas en riesgo de sufrir daños” (Naciones Unidas 2020).

Una dimensión final de la preservación de la privacidad y la confidencialidad que se relaciona con la interoperabilidad es el efecto mosaico. El Centro de Datos Humanitarios, administrado por la Oficina de Naciones Unidas para la Coordinación de Asuntos Humanitarios (OCHA), ha definido que el efecto mosaico ocurre “cuando múltiples conjuntos de datos se vinculan para revelar nueva información significativa. Si bien esa información podría usarse para obtener información valiosa, malos actores podrían usarla para causar daño” (OCHA 2020). Los riesgos asociados con el efecto mosaico pueden incluir el riesgo de volver a identificar a personas o grupos vulnerables en conjuntos de datos anonimizados muy desagregados cuando se integran entre sí. Si bien cada observación individual puede ser anónima, cuando varias observaciones se vinculan, la información resultante puede identificar a las personas dentro de los conjuntos de datos. Tener en cuenta los riesgos asociados con el efecto mosaico es una parte importante de la planificación estratégica (sección 1.1). En el “Modelo de madurez de datos responsable para organizaciones humanitarias y de desarrollo” de Care USA (CARE USA 2019) se encuentra información y orientación más detalladas sobre este fenómeno.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
El personal no tiene los conocimientos o habilidades necesarias para integrar datos	Algunos miembros del personal tienen el conocimiento y las habilidades necesarios para integrar datos, pero esto no se refleja en las descripciones de sus roles y es tangencial a sus funciones principales.	El conocimiento y las habilidades relevantes para la interoperabilidad comienzan a reconocerse como parte de las descripciones de puestos de trabajo dentro de algunas unidades, pero el enfoque está fragmentado	Existe un enfoque coordinado para el fortalecimiento del conocimiento y las habilidades en la organización, que reconoce y aborda explícitamente las necesidades de interoperabilidad	Numerosas partes de la organización reconocen el valor de la interoperabilidad de datos, incluidas las unidades no técnicas. Varios cursos de capacitación relacionados con cuestiones de gobierno de datos, incluida la interoperabilidad, están disponibles para todo el personal

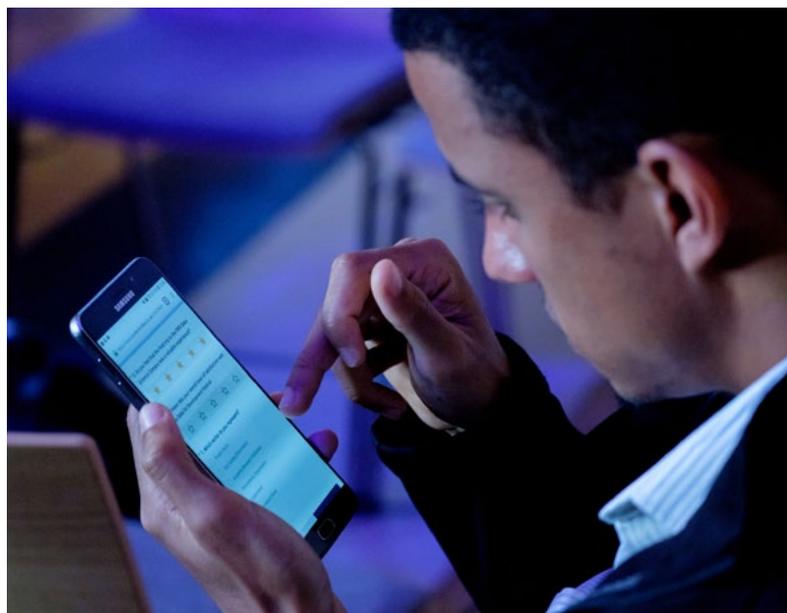
2.3 Conocimientos y habilidades del personal

Las habilidades y los conocimientos del personal son fundamentales para la interoperabilidad humana. Para las organizaciones que administran diversos conjuntos de datos y desean mejorar su interoperabilidad, es esencial que a medida que desarrollan estrategias de gestión de datos y establecen objetivos estratégicos (sección 1.1), consideran aspectos de supervisión y rendición de cuentas (sección 1.3) y consideran las funciones de custodia de datos (sección 2.1), también consideren qué habilidades y conocimientos necesitará el personal para desempeñar eficazmente sus funciones.

La capacidad de las organizaciones para integrar de manera eficiente múltiples registros de datos para generar productos de conocimiento valiosos depende cada vez más de las habilidades técnicas de los miembros del equipo.

Por ejemplo, el personal de muchas organizaciones necesita desarrollar habilidades fundamentales que les permitan diseñar, construir y mantener flujos de integración, procesamiento y difusión de datos.

En el campo de las estadísticas oficiales, la Alianza de Estadísticas para el Desarrollo en el Siglo XXI (PARIS21) publicó en 2020 *“Guidelines for Developing Statistical Capacity: A Roadmap for Capacity Development 4.0”* (PARIS21, 2020), en que se detalla las formas como las ONEs pueden apoyar la formación continua de su personal y la creación de conocimiento



Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
No hay comunicación coordinada interna o externamente que refleje el valor de la integración de datos	Los ejemplos de buenas prácticas y el valor generado como resultado de la interoperabilidad comienzan a surgir dentro de una organización, pero no se comunican interna o externamente.	El personal y las unidades comienzan a compartir ejemplos de buenas prácticas entre ellos, pero esto no está coordinado El valor de la interoperabilidad comienza a ser entendido por personal no técnico, pero aún no se comunica externamente	Los mecanismos para facilitar la comunicación interna y el intercambio de mejores prácticas en torno a la interoperabilidad forman parte de la estrategia de datos de una organización La comunicación entre unidades ayuda a traducir las mejores prácticas y ejemplos de generación de valor para audiencias externas.	La organización tiene una variedad de canales coordinados de comunicación interna entre las distintas unidades y su personal, lo que permite compartir buenas prácticas y ejemplos de generación de valor. La organización promueve el valor de la integración en todo el ecosistemas de datos y comunica activamente sus experiencias y ejemplos con otros actores.

2.4 Comunicación interna y externa

Otro principio de la interoperabilidad humana es garantizar que exista una comunicación interna y externa eficaz en torno a los problemas clave de gestión de datos que afectan la capacidad de los usuarios para utilizar e integrar los datos.

Internamente en una organización, es necesario establecer canales de comunicación formales entre las personas, los equipos y los departamentos al establecer objetivos estratégicos en la de gestión de datos (sección 1.1), funciones de liderazgo y gestión (sección 1.2), supervisión y rendición de cuentas (sección 1.3), y funciones de custodia de datos (sección 2.1).

Externamente, es esencial establecer circuitos de retroalimentación de los usuarios que permitan elevar y resolver la comunicación de la calidad de los datos, las brechas, la usabilidad y otras dimensiones de la interoperabilidad humana. De manera similar, para las audiencias no técnicas, es

importante que las organizaciones que trabajan en estadísticas oficiales o la revolución de datos en las estadísticas oficiales puedan comunicarse de manera efectiva con formuladores de políticas, quienes son los usuarios finales de sus datos. Se puede encontrar más orientación sobre el diseño centrado en el usuario en el contexto de las estadísticas oficiales en la "Guía sobre la implementación de datos abiertos en las oficinas nacionales de estadística" (UNSD 2020b), elaborada por el Grupo de trabajo de datos abiertos de la Comisión de Estadística de Naciones Unidas y presentado como documento para la 51ª sesión de la Comisión, celebrada en marzo de 2020.

La importancia de comunicar el significado y el valor de los datos de manera eficaz a audiencias no técnicas se detalla en las "Directrices para el desarrollo de la capacidad estadística: una hoja de ruta para el desarrollo de capacidades 4.0" de PARIS21" (PARIS21, 2020). Además, la Alianza Global ha publicado numerosas historias de datos, blogs y artículos de opinión en su sitio web, que transmiten la importancia de los problemas de

datos a audiencias no técnicas. Por ejemplo, “5 Useful Things I Learned About Data Storytelling: Lessons from #Visualize2030” (Jacobson, 2019) y “Joining-up Data for Universal Healthcare in Kenya: The View from the Ministry of Health” (Orrell 2020)

son dos recursos publicados en la página web de la Alianza Globalo que, respectivamente, brindan orientación sobre cómo contar historias con datos y dan un ejemplo de narración de datos centrada en la interoperabilidad.

Cuadro 2.5: Estrato de interoperabilidad humana: Dimensión de adaptabilidad

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
Los procesos relacionados con las funciones del personal y la supervisión de la interoperabilidad de los datos son rígidos y difíciles de cambiar.	Existe un entendimiento emergente dentro de la organización del valor de la adaptabilidad en las funciones y la supervisión de la gestión de datos en general, pero no hay un enfoque específico	Las diferentes unidades de la organización comienzan a reconocer formalmente la necesidad de que las funciones del personal y la supervisión de los sistemas de datos sean adaptables para garantizar la continua generación de valor a partir de sus activos de datos.	Una organización reconoce y refleja en su estrategia de datos el valor de capacitar al personal para que sea flexible en el uso de los datos, incluida la forma en que los combinan con otros datos. El personal tiene autoridad para adaptar sus procesos de trabajo y la supervisión de los activos de datos de la organización de manera que mejore su valor, incluso integrándolos.	La organización se ha convertido en líder en gestión adaptativa, el personal se siente empoderado y confía en su capacidad para adaptar su supervisión de los sistemas de datos según sea necesario, incluida la forma en que combinan los datos, para maximizar el valor generado

2.5 Adaptabilidad

La capacidad de una organización para garantizar que sus procesos internos de gestión de datos sean adaptables es una dimensión importante de la interoperabilidad humana. A medida que florecen las innovaciones digitales y de datos, las organizaciones y su personal deben asegurarse de que sus estrategias de gestión de datos mantengan el ritmo de estos desarrollos. La adaptabilidad requiere comunicación y empoderamiento de los miembros del equipo, así como flexibilidad para corregir el rumbo en cualquier momento a fin de satisfacer las necesidades cambiantes de las partes interesadas.

Pensar cómo las funciones de toma de decisiones se distribuyen entre las organizaciones es una parte clave de la planificación estratégica (sección 1.1), los vínculos con ecosistemas de datos más amplios (sección 1.7), la custodia de datos (sección 2.1), los conocimientos y habilidades del personal (sección 2.3) y comunicación organizacional interna (sección 2.4). El Capítulo 17 del “*Data Management Body of Knowledge 2nd Edition*”(DAMA International 2017) proporciona más orientación sobre la gestión de datos y la gestión de cambios organizativos.

Un buen ejemplo del valor que puede agregar ser adaptable se relaciona con la necesidad de los individuos y organizaciones de mantenerse

al tanto de la aparición de nuevos componentes de infraestructura de datos. Por ejemplo, las organizaciones que buscan mejorar la interoperabilidad, la accesibilidad y el uso de sus activos de datos deben ser conscientes de las tendencias relacionadas con la accesibilidad y aplicación de las infraestructuras digitales, y poder adaptarse a ellas o adaptarlas a sus necesidades. En términos generales, esto tiene dos dimensiones.

Primero, las innovaciones digitales y de datos ocurren muy rápidamente, con el surgimiento altamente frecuente de nuevas técnicas, herramientas y usos de los datos que surgen. Esto significa que las organizaciones que buscan mejorar la accesibilidad, la interoperabilidad y el uso de sus activos de datos deben asegurarse de realizar un seguimiento regular nuevas tecnologías e incorporarlas siempre que sea posible. Esto requiere planificación a nivel estratégico (sección 1.1), mantener vínculos proactivos con el ecosistema de datos más amplio (sección 1.7) y garantizar que los custodios de datos tengan la responsabilidad de monitorear tales innovaciones (sección 2.1)

En segundo lugar, significa que las organizaciones también deben ser conscientes de los posibles costos de oportunidad de perderse las nuevas tecnologías digitales. Si bien puede parecer contradictorio, la evidencia sugiere que a medida que se implementan más innovaciones digitales en todo el mundo, esto exacerba las desigualdades alrededor de las tecnologías digitales existentes. El “Informe sobre la economía digital 2019” de la Conferencia de las Naciones Unidas sobre Comercio y Desarrollo (UNCTAD 2019), por ejemplo, encuentra que a medida que las brechas digitales empeoran, también lo hacen las desigualdades, ya que aquellos con acceso a las últimas tecnologías van cada vez más por delante de las que no lo tienen. Comprender la dinámica de estos temas es una parte importante del proceso de planificación estratégica. Esto permite a las organizaciones aprovechar al máximo lo que tienen y planificar formas en las que se puede mejorar gradualmente el acceso a la infraestructura digital avanzada, incluidas las tecnologías que permiten o mejoran la interoperabilidad.



CAPÍTULO 3

INTEROPERABILIDAD DE DATOS

La interoperabilidad de datos requiere garantizar que los sistemas de datos y los conjuntos de datos estén diseñados para permitir la interoperabilidad. Esto significa garantizar que los datos se puedan recopilar, almacenar y procesar en formatos legibles de manera mecanizada y que se consideren las necesidades de interoperabilidad cuando se modelan y clasifican los datos. El marco de evaluación de madurez identifica cinco dimensiones de interoperabilidad de datos. Estas dimensiones y sus características asociadas para cada nivel de madurez se describen en los Cuadros 3.1 a 3.5.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
Hay poca o ninguna capacidad para modelar datos o metadatos	Existe un entendimiento emergente del valor que la modelación de datos y metadatos puede conferir a los activos de datos, pero la modelación de datos no es una prioridad para las unidades técnicas	Diversas unidades en una organización reconocen el valor de la modelación de datos y metadatos, incluida su importancia para la interoperabilidad de los datos, toman medidas para alinear sus técnicas de modelación, y comienzan a coordinar sus esfuerzos.	Las unidades técnicas coordinan la forma en que abordan la modelación de datos y metadatos, y alinean los esfuerzos para modelar los datos de manera consistente en función de sus necesidades organizacionales Las necesidades internas aún tienen prioridad sobre los grupos externos, pero los datos se modelan de manera consistente	La organización utiliza rutinariamente modelos canónicos de datos y metadatos que siguen patrones estandarizados, haciéndolos reutilizables y propicios para el intercambio de datos. La selección y aplicación de modelos canónicos se realiza a través de una planificación cuidadosa, incluso a través del compromiso con los usuarios de datos y otras entidades dentro del ecosistema de datos.

3.1 Capacidades para modelación de datos y metadatos

La modelación de datos y metadatos es la base de la interoperabilidad de datos. Es parte del proceso de diseño en el que la estructura interna y las interrelaciones entre diferentes conjuntos de datos se definen, optimizan y describen para capturar todos los datos relevantes. métricas y dimensiones comerciales (Chang 2018). El proceso de modelación de datos y metadatos incluye una serie de pasos, que van desde el descubrimiento, análisis y determinación del alcance de los requerimientos de un conjunto de datos, hasta su comunicación en forma de modelo (DAMA International 2017).

La modelación de datos debe ser adecuada para los fines de su uso. Se debe ser flexible, ya que los patrones utilizados en aplicaciones específicas o con fines analíticos a menudo implican decisiones que deben revisarse con frecuencia, tales como

aceptar cierto grado de **redundancia** a cambio de mayor facilidad de uso. Sin embargo, un enfoque casuístico para la modelación de datos pone en riesgo la capacidad de interoperar entre modelos y establecer flujos a través del ecosistema de datos.

Se necesitan patrones comunes de modelado de datos y metadatos para evitar silos organizacionales, o incluso departamentales. El Capítulo 2 de la Guía (González y Orrell 2018) profundiza en la relación entre la interoperabilidad y la modelación de datos y destaca el valor de los modelos de datos canónicos y de los estándares de la industria para permitir la interoperabilidad. De manera similar, el Capítulo 5 del “*Data Management Body of Knowledge*” (DAMA International 2017) proporciona orientación sobre cómo las organizaciones pueden mejorar sus procesos de modelado de datos.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
La organización no es consciente de la importancia de la clasificación de datos para la interoperabilidad y no tiene una idea clara de sus activos de datos.	Las unidades de la organización comienzan a inventariar sus datos. Las unidades son conscientes de la necesidad de clasificar sus datos de forma estandarizada, pero solo lo hacen caso por caso.	Existen intentos informales entre algunas unidades para usar clasificaciones comunes, pero estos no están coordinados en todas las unidades relevantes. Se utilizan clasificaciones comunes en toda la organización, pero no de manera consistente.	El uso de clasificaciones de datos se aborda en forma coordinada en toda la organización. Diferentes unidades trabajan juntas para identificar las clasificaciones más apropiadas para sus datos y garantizar que los datos bajo su control se clasifiquen adecuadamente.	La organización no solo utiliza clasificaciones de datos de manera rutinaria y adecuada, sino que también produce sus propias clasificaciones para llenar vacíos y garantizar la coherencia. La organización se involucra activamente con otras entidades en el ecosistema de datos para mejorar los sistemas de clasificación comúnmente utilizados y establecer nuevos según sea necesario La organización comunica efectivamente el valor de una clasificación de datos consistente para la interoperabilidad.

3.2 Capacidad de organización y clasificación de datos

La coordinación interna y entre organizaciones, y a través de todo el sistema estadístico nacional o del ecosistema de datos (sección 1.7) con el fin de decidir sobre el uso de clasificaciones e identificadores comunes, es clave para la interoperabilidad de un conjunto de datos. Los vocabularios estándar, las clasificaciones y los identificadores únicos forman parte de la infraestructura de datos básica de un país u organización. Ayudan a mejorar la coherencia y evitar la ambigüedad en la descripción de un conjunto de datos, al tiempo que permiten a los usuarios localizar y vincular más fácilmente elementos de datos relacionados.

Los sistemas de identificación y clasificación de datos “determinan la manera en que los datos se recopilan, procesan, analizan y comparten con los usuarios. [...] El uso de clasificaciones y vocabularios comunes permite compartir los datos de forma eficiente y que los usuarios encuentren más fácilmente información relacionada en

numerosas plataformas de datos” (Orrell 2018). Las clasificaciones estándar permiten la integración de múltiples conjuntos de datos, así como su análisis e interpretación consistentes. Proporcionan la base taxonómica para administrar y describir datos estadísticos y son componentes fundamentales de los marcos clave para la compilación de estadísticas oficiales (por ejemplo, el Sistema de Cuentas Nacionales).

Tanto los sistemas de clasificación como los de identificación deben ser interoperables y requieren una gobernanza, mantenimiento y supervisión adecuados. Esto es particularmente importante en el caso de los identificadores únicos, a fin de garantizar que sean emitidos a todas las entidades relevantes por una autoridad central reconocida, y no se emitan más de una vez, y que sigan siendo válidos durante toda la vida útil de las entidades a las que se refieren.

El Capítulo 3 de la Guía (González y Orrell 2018) explora con más detalle el valor de las clasificaciones y los vocabularios estándar para la interoperabilidad, y proporciona ejemplos de su uso entre las comunidades de las estadísticas oficiales y de la revolución de datos para el desarrollo sostenible.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
La organización no es consciente de la importancia de la clasificación de datos para la interoperabilidad y no tiene una idea clara de sus activos de datos.	Las unidades de la organización comienzan a inventariar sus datos. Las unidades son conscientes de la necesidad de clasificar sus datos de forma estandarizada, pero solo lo hacen caso por caso.	Existen intentos informales entre algunas unidades para usar clasificaciones comunes, pero estos no están coordinados en todas las unidades relevantes. Se utilizan clasificaciones comunes en toda la organización, pero no de manera consistente.	El uso de clasificaciones de datos se aborda en forma coordinada en toda la organización. Diferentes unidades trabajan juntas para identificar las clasificaciones más apropiadas para sus datos y garantizar que los datos bajo su control se clasifiquen adecuadamente.	La organización no solo utiliza clasificaciones de datos de manera rutinaria y adecuada, sino que también produce sus propias clasificaciones para llenar vacíos y garantizar la coherencia. La organización se involucra activamente con otras entidades en el ecosistema de datos para mejorar los sistemas de clasificación comúnmente utilizados y establecer nuevos según sea necesario La organización comunica efectivamente el valor de una clasificación de datos consistente para la interoperabilidad.

3.3 Difusión y apertura de datos

Si bien los capítulos 1 y 2 de este documento abordan las dimensiones del acceso a los datos, el intercambio y la apertura, también existen consideraciones técnicas en el estrato de interoperabilidad de los datos.

En la capa de interoperabilidad de datos, las consideraciones sobre cómo planificar estratégicamente la publicación de datos abiertos (sección 1.1) también se refieren a cómo se modelan (sección 3.1) y cómo se clasifican (sección 3.2) los datos. Siempre que sea posible, se debe utilizar estándares de datos y software de código abierto para la publicación de metadatos y datos abiertos, ya que esto permitirá a los usuarios integrar más fácilmente nuevos conjuntos de datos en sus sistemas.

Existe un interés creciente en las herramientas y tecnologías que permiten la publicación de datos de tal manera que fuentes de información relacionadas puedan ser identificadas e integradas

de manera automática a través de la World Wide Web. Este enfoque de "datos vinculados" facilita la interoperabilidad y la integración de datos al conectar conjuntos de datos estructurados y no estructurados con elementos de metadatos especiales diseñados para ser referenciados a través de la web. Esto, a su vez, permite la recuperación automatizada de información relevante de una gama de fuentes potencialmente enorme. La herramienta de búsqueda de conjuntos de datos de Google (Google 2020) opera un enfoque de datos vinculados utilizando estándares establecidos por el consorcio Schema.org.

Los enfoques de datos vinculados son particularmente importantes para los datos de desarrollo, porque la naturaleza "indivisible" de los Objetivos de Desarrollo Sostenible hace que sea más urgente que nunca unir los recursos de información y los activos de datos que pertenecen y son administrados por diferentes sectores y comunidades. El Capítulo 5 de la Guía (González y Orrell 2018) cubre los datos abiertos vinculados con más detalle.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
Existe poco o ningún conocimiento sobre cómo habilitar la interoperabilidad entre conjuntos de datos para realizar análisis de datos, o cómo vincular datos para entrenar algoritmos (aprendizaje automático)	Existe un entendimiento desigual entre las unidades organizacionales del papel que juega la interoperabilidad en la realización de análisis automatizado de datos. Se comprende de manera limitada cómo se debe utilizar los datos interoperables para entrenar algoritmos.	Comienza a surgir un enfoque coordinado entre las unidades organizacionales, y algunas unidades comienzan a producir conjuntos de datos depurados, de calidad garantizada y consistentes, que están disponibles para ser integrados y procesados automáticamente. Hay un esfuerzo coordinado para comprender cómo se pueden combinar los conjuntos de datos para entrenar algoritmos.	Las funciones de análisis de datos y aprendizaje automático se reflejan dentro de la estrategia de datos de una organización. Los beneficios y riesgos relativos de ejecutar análisis automatizado de datos interoperables, o usarlos para entrenar algoritmos, generalmente se comprenden bien, pero aún no existe un enfoque coherente en toda la organización.	La estrategia de datos de la organización incluye planes prospectivos sobre cómo las herramientas de análisis de datos se pueden aplicar de manera responsable a múltiples conjuntos de datos interoperables en el futuro. Existe una comprensión matizada y bien establecida de los beneficios y riesgos relativos relacionados con la ejecución de análisis automatizado de datos interoperables, o su uso para entrenar algoritmos; se aplican evaluaciones adecuadas de riesgo y costo-beneficio, según sea necesario. La organización se involucra de manera proactiva con otras entidades en el ecosistema de datos para compartir sus aprendizajes, y utiliza herramientas de análisis de código fuente abierto siempre que sea posible para permitir un escrutinio transparente.

3.4 Análisis de datos y automatización

A medida que las revoluciones digitales y de datos se han disparado en todo el mundo, también lo ha hecho la capacidad de una gama más amplia de entidades e individuos para recopilar, organizar, estructurar y analizar tipos de datos más diversos. Y al integrar más datos y tipos de datos, la interoperabilidad se vuelve más importante para garantizar que éstos se modelen (sección 3.1) de manera que permitan el procesamiento y análisis, incluso por medios automatizados.

En la comunidad de estadísticas oficiales, en los últimos años ha aumentado la cooperación

internacional en torno al análisis de datos estadísticos, incluso de manera automatizada. El Grupo de trabajo de mundial sobre big data para estadísticas oficiales (UNSD 2020c), por ejemplo, explora el papel que pueden desempeñar los conocimientos extraídos de la big data para contribuir a alcanzar los ODS. La interoperabilidad entre datos derivados de múltiples fuentes a menudo es necesaria para ejecutar técnicas de procesamiento automatizado sobre conjuntos de datos.

A medida que las ONE y otras entidades en la revolución de los datos para el desarrollo sostenible experimentan e innovan con nuevas fuentes de datos, incluso mediante el uso de técnicas de procesamiento automatizado para analizarlas, es importante que los riesgos asociados con

estas prácticas se comprendan y gestionen adecuadamente.

Las organizaciones deben ser especialmente conscientes de los riesgos asociados con el efecto mosaico, que se analiza con más detalle en la sección 2.2. El "Modelo de madurez de datos

responsable para organizaciones humanitarias y de desarrollo" de CARE USA (CARE USA 2019) también contiene más información sobre los riesgos asociados con el procesamiento automatizado, incluido el efecto mosaico.

Cuadro 3.5: Estrato de interoperabilidad de datos: Dimensión de protección de datos

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
Existe poca o ninguna comprensión o conocimiento de los vínculos entre la interoperabilidad de datos y las técnicas de protección de datos, que incluyen anonimización, seudonimización y cifrado.	Existe cierto conocimiento y comprensión de la necesidad de proteger los datos que se combinarán con otros datos, incluso mediante el uso de técnicas adecuadas de seudonimización, anonimización y cifrado, según sea necesario, pero este conocimiento no es uniforme, y las técnicas de protección de datos no se aplican de manera consistente.	Diversas unidades aplican rutinariamente técnicas apropiadas de protección de datos a sus conjuntos de datos antes de integrarlos, pero hay poca o ninguna coherencia en la forma en que se aplican esas técnicas Existe cierta comprensión, pero limitada, de los riesgos de reidentificación inherentes a los datos interoperables.	Los datos personales y confidenciales a nivel individual y de grupo están cubiertos por protecciones adecuadas antes de ser integrados, compartidos o procesados mediante herramientas de análisis automatizado. Los riesgos de reidentificación inherentes a los datos interoperables se entienden y mitigan, pero no de manera rutinaria.	Todos los datos se protegen mediante las técnicas adecuadas y se archivan de forma responsable o se eliminan permanentemente al final de su ciclo de vida previsto. El acceso a conjuntos de datos confidenciales se monitorea y documenta para garantizar el rendimiento de cuentas sobre la protección de datos. Antes de su integración, intercambio o procesamiento mediante análisis automatizado, todos los datos se evalúan para detectar riesgos de reidentificación u otros daños potenciales, y solo se utilizan cuando existe un alto grado de certeza de que los datos permanecerán seguros después de su reutilización. La organización ayuda a establecer estándares para la protección de datos dentro del ecosistema de datos más amplio y defiende el uso responsable de datos.

3.5 Protección de datos

En su nivel más básico, la protección de datos es la regulación de cómo se controla el acceso y el uso de los datos almacenados en computadoras, dispositivos digitales y registros en papel.

La protección de datos se puede ver a través de tres lentes. Primero, es una serie de obligaciones legales en jurisdicciones donde se han aprobado leyes de protección de datos que establecen reglas sobre cómo se deben proteger los datos. El Reglamento general de protección de datos de la Unión Europea

(UE 2016) es un buen ejemplo de esto. En segundo lugar, la protección de datos puede relacionarse con la protección de datos personales y está estrechamente relacionada con la preservación de la privacidad y la confidencialidad (sección 2.2). Por último, puede relacionarse con una combinación de consideraciones técnicas y relacionadas con el contenido de los datos.

Desde el punto de vista técnico, la protección de datos está relacionada con la ciberseguridad y la respuesta a incidentes (sección 4.2), mientras que a nivel del contenido de los datos puede relacionarse con cuestiones relativas a “la planificación, el desarrollo y la ejecución de políticas y procedimientos de seguridad para proporcionar una autenticación y autorización adecuadas, acceso y auditoría de datos y activos de información” (DAMA International 2017).

En términos de su vínculo con la interoperabilidad, hay varias consideraciones a tener en cuenta, muchas de ellas relacionadas con la posible identificación de individuos cuando un conjunto de datos se utiliza o se integra con otros datos de manera que se revele la identidad de individuos o grupos (vulnerables) (secciones 1.4, 1.5 y 2.2). A medida que las organizaciones consideran la protección de datos, deben ser conscientes de las obligaciones legales de su jurisdicción, cualquier estándar normativo aplicable (como los Principios Fundamentales de las Estadísticas Oficiales) o códigos éticos, y prácticas consuetudinarias en un sector o dominio en particular. Además, se deben tomar medidas dentro de las organizaciones para

anonimizar y seudonimizar conjuntos de datos que puedan contener observaciones que resulten en la identificación de individuos. La publicación “*Anonymisation and Open Data: An Introduction to Managing the Risk of Re-identification*” del Open Data Institute (ODI 2019b) proporciona una útil descripción general de las cuestiones clave a considerar cuando se busca anonimizar conjuntos de datos.

Los conjuntos de datos que pueden contener información confidencial o de identificación personal deben tener acceso restringido a ellos para defender los derechos de los individuos sobre los datos y cualquier deber de confidencialidad aplicable (sección 2.2). Estos datos deben estar cifrados y almacenados de forma segura y deben establecer una jerarquía de permisos de acceso. El mantenimiento de listas de permisos y el monitoreo de su implementación e integridad debe ser la función de los custodios de datos dentro de los departamentos técnicos de una organización de producción o procesamiento de datos (sección 2.1).

Por último, los vínculos para el acceso, la apertura y el intercambio de datos (sección 3.3) también deben considerarse en la etapa de planificación estratégica (sección 1.1) y se deben agregar un nivel de cifrado a todo conjunto de datos restringidos.



CAPÍTULO 4

INTEROPERABILIDAD TECNOLÓGICA

La interoperabilidad tecnológica requiere una infraestructura de datos adecuada para permitir la interoperabilidad organizacional, humana y de datos de manera adecuada. El marco de evaluación de madurez identifica dos dimensiones principales de interoperabilidad tecnológica de los datos. Estas dimensiones y sus características asociadas para cada nivel de madurez se describen en los Cuadros 4.1 y 4.2.

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
La organización enfrenta escasez de infraestructura digital clave para almacenar, gestionar, intercambiar y procesar datos, incluyendo componentes de hardware y software, suministro confiable de electricidad, o conectividad a Internet.	La organización tiene acceso adecuado a componentes clave de infraestructura digital, pero hay escasez de soluciones de gestión de contenido y almacenamiento de datos específicos de la organización, lo que resulta en una gestión de datos no estandarizada y no alineada.	Todos los miembros del personal apropiados tienen sus propias computadoras y conectividad a Internet. Existen servidores seguros y repositorios de datos, pero el personal y las unidades organizativas los utilizan de manera inconsistente, hay poca supervisión de la infraestructura digital.	Todos los miembros del personal pertinente conocen y están capacitados para utilizar el sistema de gestión de datos de una organización. Los servidores seguros y los repositorios de datos son utilizados habitualmente por los miembros del personal y la supervisión de la infraestructura digital es parte de la estrategia de datos de una organización.	La estrategia de datos de una organización incluye disposiciones para el mantenimiento, la revisión periódica y la actualización de su infraestructura digital, y las líneas presupuestarias se reservan para este propósito. Las políticas de adquisiciones, intercambio de datos y supervisión de la infraestructura están alineadas. La organización aborda los temas de infraestructura digital con un enfoque prospectivo, y desarrolla activamente estrategias y planes para optimizar el uso de tecnologías emergentes y mejorar la interoperabilidad de sus sistemas de datos.

4.1 Infraestructura de datos

Aunque puede ser evidente, vale la pena declarar explícitamente que todas las entidades y organizaciones que deseen participar en el ecosistema de datos más amplio (sección 1.7) y que trabajen para mejorar la accesibilidad y el uso de sus activos de datos deben tener una infraestructura digital básica, como como computadoras portátiles, capacidad de base de datos (servidor o nube) y conectividad a Internet.

Sin embargo, el concepto de infraestructura digital es mucho más amplio de lo que era hace apenas unos años. En la actualidad, existen numerosas categorías de infraestructura que dependen y afectan la interoperabilidad. Algunos ejemplos de componentes de infraestructura interconectados podrían incluir los siguientes:

- Infraestructura de almacenamiento de datos (ya sean servidores locales o basados en la nube)
- Infraestructura de gestión de datos, incluyendo sistemas de gestión de información diseñados a la medida.
- Infraestructura de procesamiento de datos (cada vez más basada en la nube para grandes conjuntos de datos)
- Infraestructura de difusión de datos compuesta por plataformas de datos, centros comunitarios, etc.

Elegir, adquirir y mantener la infraestructura digital es una función clave en las organizaciones basadas en datos. A la hora de decidir qué soluciones adoptar, deben tenerse en cuenta factores tales como el tipo y el volumen de activos de datos que deben almacenarse, accederse y analizarse. Otros factores para considerar son la necesidad

de acceder a los datos y procesarlos en lotes o en tiempo real, o requerimientos de compatibilidad con tecnologías de código abierto.

Todos estos componentes requieren y mejoran la interoperabilidad en los estratos de datos, organizacional y humano. A medida que más funciones de almacenamiento, procesamiento y difusión de datos se basan en la nube, por ejemplo, esto aumenta la necesidad de consideraciones en torno a la ciberseguridad (sección 4.2); el acceso,

la apertura y el intercambio de datos (sección 3.3); cuestiones éticas relacionadas con los derechos otorgados a procesadores de datos de terceros (sección 1.5); y muchas más. En última instancia, lo importante es que la infraestructura se refleje claramente en los objetivos estratégicos de las organizaciones (sección 1.1) y que exista una gestión eficaz (sección 1.2) y supervisión (sección 1.3) de cómo se aplican estos componentes.

Cuadro 4.2: Estrato de interoperabilidad tecnológica: Dimensión de seguridad cibernética y respuesta ante incidentes

Características de cada nivel progresivo de madurez

INDEFINICIÓN	INICIACIÓN	APRENDIZAJE	DESARROLLO	CONSOLIDACIÓN
<p>Se tiene poco o ningún conocimiento de los riesgos de ataques cibernéticos u otras infracciones a los sistemas de datos de la organización, incluidos los riesgos específicos asociados con datos interoperables potencialmente reutilizables.</p> <p>No existe un protocolo o política para enfrentar violaciones de datos.</p>	<p>Personal y unidades diversos de la organización tienen conciencia o muestran preocupación por los riesgos que representaría para sus datos reutilizables un ataque cibernético u otra violación de datos.</p> <p>Surgen campeones que promueven el establecimiento de un protocolo o política para enfrentar violaciones de datos.</p>	<p>Se ha preparado un protocolo para enfrentar violaciones de datos dentro de la organización, pero los riesgos asociados con la posible reutilización de datos interoperables robados siguen siendo vagos y la política se entiende y aplica de forma inconsistente.</p>	<p>Se establece una política clara para enfrentar violaciones de datos que establece una secuencia de procedimientos y responsabilidades.</p> <p>El personal recibe capacitación sobre lo que se debe hacer en caso de una violación de datos y se les enseña sobre los riesgos asociados con la reutilización de datos interoperables robados.</p>	<p>La organización puede hacer frente a las violaciones de datos de manera rápida y efectiva, y toma medidas activas para garantizar que su infraestructura tecnológica sea lo más segura posible.</p> <p>La política para enfrentar violaciones de datos de la organización se revisa y actualiza periódicamente, y esta incluye explícitamente los riesgos asociados con la reutilización de datos interoperables.</p> <p>El personal apropiado recibe capacitación rutinaria sobre cómo responder a una violación de datos.</p>

4.2 Seguridad cibernética y respuesta ante incidentes

A medida que más funciones clave de gestión de datos se mueven en línea y los arreglos de trabajo remoto se convierten en la regla en lugar de la excepción, la ciberseguridad y los protocolos para la respuesta a incidentes se vuelven

más importantes. Tomemos el ejemplo de la computación en la nube. Debido a su dependencia de la tecnología de virtualización independiente del hardware, la computación en la nube permite a las organizaciones realizar rápidamente copias de seguridad de datos, aplicaciones e incluso sistemas operativos en un centro de datos remoto, e implementarlos en múltiples usuarios en muchas ubicaciones diferentes. Sin embargo, toda esta

transmisión de datos a través de Internet lo expone a amenazas de ciberseguridad.

Las amenazas a la seguridad suelen tener éxito en los puntos más débiles de los sistemas de datos de las organizaciones. A medida que los enfoques de interoperabilidad basados en la web comienzan a florecer (sección 3.3), las cuestiones de ciberseguridad y respuesta a incidentes deben tenerse en cuenta en las estrategias de gestión de datos de las organizaciones (sección 1.1). Del mismo modo, el personal debe recibir ayuda para asumir la responsabilidad de la seguridad de los datos de sus propios dispositivos (computadoras portátiles, teléfonos móviles, etc.), fomentando "un enfoque analítico y reflexivo" de la ciberseguridad, en lugar de simplemente brindar capacitación basada en reglas que pueda resultar en una toma de decisiones demasiado rígida o estrecha (HBR

2020). Esta área también está estrechamente alineada con las cuestiones relativas a la protección de datos (sección 3.5); administración de datos (sección 2.1); y acceso, apertura e intercambio de datos (sección 3.3).

La publicación "Data Management Body of Knowledge 2nd Edition" (DAMA International 2017) identifica cinco actividades clave relacionadas con la estrategia de seguridad de datos: identificar los requisitos de seguridad de datos relevantes, definir una política, definir estándares de seguridad, evaluar riesgos e implementar controles y procedimientos. Además, el "Modelo de madurez de datos responsable para organizaciones humanitarias y de desarrollo" de CARE USA (CARE USA 2019) proporciona orientación adicional relacionada con el desarrollo de enfoques para gestionar situaciones de respuesta a incidentes.



APÉNDICE A: EL MARCO DE EVALUACIÓN DE MADUREZ DE DATOS INTEGRADOS

Estratos del gobierno de datos

Dimensiones

INDEFINICIÓN

INICIACIÓN

APRENDIZAJE

DESARROLLO

CONSOLIDACIÓN

Objetivos estratégicos

La interoperabilidad no se reconoce como un objetivo estratégico

Se reconoce la importancia de la capacidad de unir datos, pero ésta no se identifica explícitamente como un objetivo estratégico

La interoperabilidad se identifica como un objetivo estratégico dentro de las unidades técnicas de una organización, pero no fuera de ellas

La necesidad de unir datos entre sistemas se reconoce como un objetivo estratégico dentro de una estrategia de datos organizacionales.
Se reconoce el valor de contar con estándares y un gobierno de datos sólido.

En las estrategias organizacionales se reconoce el valor que la integración de datos puede conferir a la toma de decisiones
La interoperabilidad forma parte de la estrategia de compromiso externo de una organización con otros productores y usuarios de datos

Liderazgo y gestión

No existe un liderazgo definido sobre cuestiones de interoperabilidad

El liderazgo ad hoc en temas de interoperabilidad surge orgánicamente dentro de una organización, pero no está coordinado

El liderazgo en torno a la interoperabilidad emerge a través de varias unidades técnicas pero permanece fragmentado
Los silos persisten

Existe una jerarquía coordinada de liderazgo sobre cuestiones de interoperabilidad
Se han establecido funciones claras relacionadas con la interoperabilidad en toda la organización.

Existe un comité o consejo de gobernanza de datos con un mandato explícito para liderar en asuntos relacionados con la interoperabilidad
Los líderes y gerentes de la organización entienden el valor de los datos unidos, y la integración de datos se identifica claramente como una función en las descripciones de trabajo relevantes

ORGANIZACIÓN

INDEFINICIÓN

INICIACIÓN

APRENDIZAJE

DESARROLLO

CONSOLIDACIÓN

Supervisión y rendimiento de cuentas

No hay supervisión ni rendimiento de cuentas sobre los problemas de interoperabilidad

La organización es consciente de la necesidad tener cadenas de responsabilidad para supervisar cómo se unen los datos, pero aún no se toman medidas para establecerlas.

Las estructuras de supervisión de la interoperabilidad emergen a través de diferentes unidades técnicas pero no están coordinadas o alineadas

La responsabilidad sobre cómo se realizan los esfuerzos de interoperabilidad es fragmentada y poco clara

Las funciones de supervisión y rendición de cuentas sobre interoperabilidad están integradas en la estrategia de una organización y se reflejan en la estructura de liderazgo.

Las cadenas fragmentadas de supervisión y rendimiento de cuentas respecto a la interoperabilidad comienzan a unirse y surgen procedimientos operativos estándar comunes

Una cadena clara de supervisión y rendimiento de cuentas fluye desde el comité o consejo de gobierno de datos de una organización hasta el personal operativo

Las unidades organizacionales tienen claro cuáles son sus funciones en relación con la interoperabilidad y ante quiénes deben rendir cuentas al respecto

Conformidad legal

No hay conciencia sobre las obligaciones legales aplicables relacionadas con la integración de datos interoperables

Existe un entendimiento general de que las acciones que facilita la interoperabilidad (transmisión, intercambio y (re) utilización de datos) podrían estar reguladas, pero no está claro cómo

La organización toma medidas activas para comprender mejor los requisitos de conformidad legal en relación con la retención, la transmisión, el intercambio y la utilización de datos, y los pone a disposición de los usuarios de datos.

El cumplimiento de las leyes aplicables sobre intercambio, transmisión y utilización de datos está integrado en las funciones de supervisión y rendimiento de cuentas, y se refleja en la estrategia de datos de la organización, la cual está publicada en línea

Las actividades de transmisión, intercambio y utilización de datos de la organización cumplen plenamente con las leyes aplicables y, a veces, superan los estándares legales

Ética de datos

No hay conciencia sobre los riesgos éticos que la interoperabilidad de datos podría representar

Existe un entendimiento general de que la unión de datos a veces puede presentar riesgos éticos, pero no está claro cómo

La organización toma medidas activas para comprender las consecuencias éticas que puede ocasionar la integración de datos y para comprender como ésta puede ocasionar daños no intencionales

Se llevan a cabo evaluaciones de impacto ético de manera rudimentaria y ad hoc

Se comprenden y categorizan adecuadamente en la organización los diferentes tipos cuestiones éticas que puede generar la integración de datos

Se toman las medidas adecuadas para minimizar el daños causados por violaciones a los estándares éticos

Se llevan a cabo evaluaciones de impacto ético de manera rutinaria

Se comprenden bien los riesgos que representa la integración de datos y se realizan revisiones éticas a lo largo del ciclo de vida de los datos para monitorear problemas y corregir el curso, según sea necesario.

Las evaluaciones de impacto ético se publican en línea y de forma transparente

La organización solo integra datos una vez que ha emprendido y documentado una revisión de los riesgos potenciales y ha tomado las medidas adecuadas para mitigarlos.

Proceso de adquisiciones de proveedores externos

La organización no es consciente del impacto de la adquisición de soluciones técnicas de interoperabilidad

Existe un nivel emergente de entendimiento sobre la necesidad de integrar datos en todas las actividades relacionadas con adquisiciones de proveedores externos, pero no existe un enfoque coherente

Existe una gran dependencia hacia contratistas externos para solventar carencias, pero no hay coordinación entre diferentes unidades de la organización, lo que a menudo resulta en la adquisición de soluciones para el manejo de datos que son incompatibles.

La coordinación entre diferentes unidades organizacionales en torno a la adquisición de soluciones técnicas de proveedores externos comienza a materializarse, pero aún no está formalizada

Ocasionalmente, diferentes unidades coordinan informalmente la contratación de proveedores externos para garantizar que los nuevos sistemas de datos sean compatibles con la infraestructura de datos existente.

La adquisición de sistemas de datos compatibles e interoperables en una organización se formaliza y coordina

Existe una política común de adquisiciones de proveedores externos en toda la organización, la cual requiere que el personal considere los problemas de interoperabilidad al adquirir nuevos sistemas. La dependencia de proveedores externos es estratégica y coordinada

La organización integra la adquisición de nuevos sistemas y servicios de datos en su estrategia de datos, e incluye planes a futuro

Las diferentes unidades planifican y analizan estratégicamente sus necesidades comunes de adquisición de soluciones, y se aseguran de que cualquier sistema o servicio nuevo de datos sea compatible con la infraestructura existente y satisfaga las probables necesidades futuras

Vínculos con ecosistemas de datos más amplios

No hay conciencia sobre la forma en que se utilizan los datos en el ecosistema de datos y el papel que en ello juega la interoperabilidad de datos

Existe un entendimiento emergente de que la integración de datos entre las diferentes entidades que conforman un ecosistema de datos pueden generar tanto oportunidades como desafíos

La organización interactúa con otros sectores del ecosistema de datos de manera informal y ad hoc

La organización comienza a adjuntar términos de licencia pro forma a los datos que publica, transmite o comparte, los cuales incluyen disposiciones sobre su integración con otros datos, pero no supervisa ni interactúa con los usuarios para garantizar su aplicación.

La organización comienza a documentar los datos que recibe de otras organizaciones.

La organización comienza a documentar y coordinar su involucramiento con otras partes del ecosistema de datos.

La organización clasifica efectivamente sus datos y los licencia para su adecuada utilización

La organización documenta todos los datos que se comparten con ella y tiene entendimiento general de lo que se puede y lo que no se puede hacer con ellos

La organización se involucra con otras partes del ecosistema de datos de manera coordinada, en conformidad con su estrategia de datos

La organización tiene un conjunto de licencias bien establecido y adaptado a sus propias necesidades, el cual define parámetros claros para la utilización de datos, incluida su integración, según el tipo de datos en cuestión

La organización documenta todos los datos que se comparten con ella y cuenta con una guía y procedimientos claros que rigen si esos datos se pueden unir y cómo se pueden unir con otros conjuntos de datos bajo su control.

La organización define como prioridad estratégica el involucramiento con otras partes del ecosistema de datos, y tiene un enfoque bien coordinado y con procesos claros para integrar sus datos con datos externos.

Custodia de datos

No se ha asignado ningún miembro del personal al rol de custodia de datos

La integración de datos forma parte de la función de ciertos miembros del personal, pero no se refleja en sus descripciones de trabajo y se mantiene como una actividad ad hoc

El personal de diferentes unidades tiene explícitamente funciones relacionadas con la interoperabilidad de datos, pero hay poca (o ninguna) coordinación central

Existe un plan coordinado de reclutamiento de personal que refleja las diversas dimensiones y roles relacionados con la custodia de los datos, incluida la interoperabilidad, en toda la organización

Las diferentes unidades se coordinan y comunican entre sí qué datos están uniéndose, cómo y por qué

Un consejo o comité de gobierno de datos de la organización supervisa un plan estratégico para la gestión de datos que incluye un plan claro para la custodia de datos, incluidas las funciones de interoperabilidad

El personal de toda la organización es consciente de cómo se usan, unen y comparten los datos con otras entidades.

Privacidad y preservación de la confidencialidad

Hay poca o ninguna comprensión de los riesgos para la privacidad y la necesidad de preservar la confidencialidad en los conjuntos de datos interoperables

Existe un entendimiento emergente de los riesgos que representa para los individuos o grupos vulnerables el integrar distintos conjuntos de datos.

Existe conocimiento de las leyes (internacionales) de privacidad y confidencialidad aplicables, así como de principios los normativos, buenas prácticas y lineamientos relevantes, pero no se consideran ni se los siguen de manera rutinaria cuando los activos de datos de una organización se integran con otros datos o se utilizan de alguna otra manera.

Se considera de forma rutinaria las leyes, principios, buenas prácticas y lineamientos aplicables

La organización realiza evaluaciones de impacto en la privacidad antes y durante los proyectos relacionados con datos, y esas evaluaciones incluyen consideraciones sobre los riesgos asociados con la interoperabilidad de los datos.

La preservación de la privacidad individual y la confidencialidad de los datos forman parte de las evaluaciones legales y éticas de la organización, y están integrados en todo el ciclo de vida de los datos.

La organización se adhiere a los más altos estándares de privacidad y preservación de la confidencialidad, independientemente de los requerimientos legales locales.

La organización ha integrado la preservación de la privacidad y la confidencialidad en su estrategia de datos y proporciona explícitamente orientación sobre los riesgos de los datos interoperables (incluyendo, por ejemplo, el "efecto mosaico")

La organización tiene una visión de largo plazo, y conoce los riesgos potenciales de privacidad inherentes a la interoperabilidad de datos que plantean tecnologías emergentes tales como la Internet de las cosas, la verificación biométrica de identidad o los procesos automatizados generales

Conocimientos y habilidades del personal

<p>El personal no tiene los conocimientos o habilidades necesarias para integrar datos</p>	<p>Algunos miembros del personal tienen el conocimiento y las habilidades necesarios para integrar datos, pero esto no se refleja en las descripciones de sus roles y es tangencial a sus funciones principales.</p>	<p>El conocimiento y las habilidades relevantes para la interoperabilidad comienzan a reconocerse como parte de las descripciones de puestos de trabajo dentro de algunas unidades, pero el enfoque está fragmentado</p>	<p>Existe un enfoque coordinado para el fortalecimiento del conocimiento y las habilidades en la organización, que reconoce y aborda explícitamente las necesidades de interoperabilidad</p>	<p>Numerosas partes de la organización reconocen el valor de la interoperabilidad de datos, incluidas las unidades no técnicas.</p> <p>Varios cursos de capacitación relacionados con cuestiones de gobierno de datos, incluida la interoperabilidad, están disponibles para todo el personal</p>
--	--	--	--	---

Comunicación interna y externa

<p>No hay comunicación coordinada interna o externamente que refleje el valor de la integración de datos</p>	<p>Los ejemplos de buenas prácticas y el valor generado como resultado de la interoperabilidad comienzan a surgir dentro de una organización, pero no se comunican interna o externamente.</p>	<p>El personal y las unidades comienzan a compartir ejemplos de buenas prácticas entre ellos, pero esto no está coordinado</p> <p>El valor de la interoperabilidad comienza a ser entendido por personal no técnico, pero aún no se comunica externamente</p>	<p>Los mecanismos para facilitar la comunicación interna y el intercambio de mejores prácticas en torno a la interoperabilidad forman parte de la estrategia de datos de una organización</p> <p>La comunicación entre unidades ayuda a traducir las mejores prácticas y ejemplos de generación de valor para audiencias externas.</p>	<p>La organización tiene una variedad de canales coordinados de comunicación interna entre las distintas unidades y su personal, lo que permite compartir buenas prácticas y ejemplos de generación de valor.</p> <p>La organización promueve el valor de la integración en todo el ecosistemas de datos y comunica activamente sus experiencias y ejemplos con otros actores.</p>
--	--	---	--	--

Adaptabilidad y agilidad

<p>Los procesos relacionados con las funciones del personal y la supervisión de la interoperabilidad de los datos son rígidos y difíciles de cambiar.</p>	<p>Existe un entendimiento emergente dentro de la organización del valor de la adaptabilidad en las funciones y la supervisión de la gestión de datos en general, pero no hay un enfoque específico</p>	<p>Las diferentes unidades de la organización comienzan a reconocer formalmente la necesidad de que las funciones del personal y la supervisión de los sistemas de datos sean adaptables para garantizar la continua generación de valor a partir de sus activos de datos.</p>	<p>Una organización reconoce y refleja en su estrategia de datos el valor de capacitar al personal para que sea flexible en el uso de los datos, incluida la forma en que los combinan con otros datos.</p> <p>El personal tiene autoridad para adaptar sus procesos de trabajo y la supervisión de los activos de datos de la organización de manera que mejore su valor, incluso integrándolos.</p>	<p>La organización se ha convertido en líder en gestión adaptativa, el personal se siente empoderado y confía en su capacidad para adaptar su supervisión de los sistemas de datos según sea necesario, incluida la forma en que combinan los datos, para maximizar el valor generado</p>
---	---	--	---	---

Capacidades para modelación de datos y metadatos

Hay poca o ninguna capacidad para modelar datos o metadatos

Existe un entendimiento emergente del valor que la modelación de datos y metadatos puede conferir a los activos de datos, pero la modelación de datos no es una prioridad para las unidades técnicas

Diversas unidades en una organización reconocen el valor de la modelación de datos y metadatos, incluida su importancia para la interoperabilidad de los datos, toman medidas para alinear sus técnicas de modelación, y comienzan a coordinar sus esfuerzos.

Las unidades técnicas coordinan la forma en que abordan la modelación de datos y metadatos, y alinean los esfuerzos para modelar los datos de manera consistente en función de sus necesidades organizacionales

Las necesidades internas aún tienen prioridad sobre los grupos externos, pero los datos se modelan de manera consistente

La organización utiliza rutinariamente modelos canónicos de datos y metadatos que siguen patrones estandarizados, haciéndolos reutilizables y propicios para el intercambio de datos.

La selección y aplicación de modelos canónicos se realiza a través de una planificación cuidadosa, incluso a través del compromiso con los usuarios de datos y otras entidades dentro del ecosistema de datos.

Capacidades para organización y clasificación de datos y metadatos

La organización no es consciente de la importancia de la clasificación de datos para la interoperabilidad y no tiene una idea clara de sus activos de datos

Las unidades de la organización comienzan a inventariar sus datos
Las unidades son conscientes de la necesidad de clasificar sus datos de forma estandarizada, pero solo lo hacen caso por caso.

Existen intentos informales entre algunas unidades para usar clasificaciones comunes, pero estos no están coordinados en todas las unidades relevantes
Se utilizan clasificaciones comunes en toda la organización, pero no de manera consistente

El uso de clasificaciones de datos se aborda en forma coordinada en toda la organización.

Diferentes unidades trabajan juntas para identificar las clasificaciones más apropiadas para sus datos y garantizar que los datos bajo su control se clasifiquen adecuadamente

La organización no solo utiliza clasificaciones de datos de manera rutinaria y adecuada, sino que también produce sus propias clasificaciones para llenar vacíos y garantizar la coherencia

La organización se involucra activamente con otras entidades en el ecosistema de datos para mejorar los sistemas de clasificación comúnmente utilizados y establecer nuevos según sea necesario

La organización comunica efectivamente el valor de una clasificación de datos consistente para la interoperabilidad

Difusión de datos y datos abiertos

La organización tiene poco o ningún conocimiento de las consideraciones de interoperabilidad cuando planea compartir de manera responsable los datos o abrirlos para su utilización

Diversas unidades en la organización son conscientes de las consideraciones de interoperabilidad cuando planean compartir o publicar datos de manera responsable como datos abiertos, pero este conocimiento no es uniforme ni universalmente aplicado

Hay esfuerzos coordinados dentro de las unidades técnicas para garantizar que los datos se compartan de manera responsable, incluidos los permisos o limitaciones de licencia relevantes para la futura integración y la utilización de datos

Algunos datos se publican en una plataforma de datos abiertos de la organización, pero los conjuntos de datos están incompletos, no son oportunos o no se ha garantizado su calidad.

Los datos se comparten responsablemente, protegiendo los derechos que terceros puedan tener sobre ellos.

La publicación de datos abiertos se realiza en formatos que permiten su lectura mecanizada, bajo una licencia clara de datos abiertos, con términos de uso y eliminando los atributos que pueden dar lugar a la reidentificación de individuos o grupos vulnerables.

Los portales de datos abiertos van acompañados de información contextual relevante y se visualizan de manera que promueven su uso por parte de múltiples grupos

La organización aplica una política efectiva de intercambio de datos, la cual brinda orientación sobre las diversas formas en que debe realizarse el intercambio de datos, desde su publicación bajo una licencia abierta, hasta el uso de convenios formales para el procesamiento o intercambio de datos

Se cuenta con asesoramiento legal disponible para el personal que desea compartir datos que serán integrados con otros conjuntos de datos por terceros.

Los datos abiertos no solo se publican en formatos legibles de manera mecanizada y por humanos, sino que también están disponibles como datos vinculados a través de la web semántica

Existen canales de retroalimentación con grupos clave de usuarios, y la organización responde a las necesidades de los usuarios

Análisis de datos y aprendizaje automático (machine learning)

Existe poco o ningún conocimiento sobre cómo habilitar la interoperabilidad entre conjuntos de datos para realizar análisis de datos, o cómo vincular datos para entrenar algoritmos (aprendizaje automático)

Existe un entendimiento desigual entre las unidades organizacionales del papel que juega la interoperabilidad en la realización de análisis automatizado de datos.

Se comprende de manera limitada cómo se debe utilizar los datos interoperables para entrenar algoritmos

Comienza a surgir un enfoque coordinado entre las unidades organizacionales, y algunas unidades comienzan a producir conjuntos de datos depurados, de calidad garantizada y consistentes, que están disponibles para ser integrados y procesados automáticamente.

Hay un esfuerzo coordinado para comprender cómo se pueden combinar los conjuntos de datos para entrenar algoritmos

Las funciones de análisis de datos y aprendizaje automático se reflejan dentro de la estrategia de datos de una organización

Los beneficios y riesgos relativos de ejecutar análisis automatizado de datos interoperables, o usarlos para entrenar algoritmos, generalmente se comprenden bien, pero aún no existe un enfoque coherente en toda la organización.

La estrategia de datos de la organización incluye planes prospectivos sobre cómo las herramientas de análisis de datos se pueden aplicar de manera responsable a múltiples conjuntos de datos interoperables en el futuro

Existe una comprensión matizada y bien establecida de los beneficios y riesgos relativos relacionados con la ejecución de análisis automatizado de datos interoperables, o su uso para entrenar algoritmos; se aplican evaluaciones adecuadas de riesgo y costo-beneficio, según sea necesario.

La organización se involucra de manera proactiva con otras entidades en el ecosistema de datos para compartir sus aprendizajes, y utiliza herramientas de análisis de código fuente abierto siempre que sea posible para permitir un escrutinio transparente

Protección de datos

Existe poca o ninguna comprensión o conocimiento de los vínculos entre la interoperabilidad de datos y las técnicas de protección de datos, que incluyen anonimización, seudonimización y cifrado.

Existe cierto conocimiento y comprensión de la necesidad de proteger los datos que se combinarán con otros datos, incluso mediante el uso de técnicas adecuadas de seudonimización, anonimización y cifrado, según sea necesario, pero este conocimiento no es uniforme, y las técnicas de protección de datos no se aplican de manera consistente.

Diversas unidades aplican rutinariamente técnicas apropiadas de protección de datos a sus conjuntos de datos antes de integrarlos, pero hay poca o ninguna coherencia en la forma en que se aplican esas técnicas. Existe cierta comprensión, pero limitada, de los riesgos de reidentificación inherentes a los datos interoperables.

Los datos personales y confidenciales a nivel individual y de grupo están cubiertos por protecciones adecuadas antes de ser integrados, compartidos o procesados mediante herramientas de análisis automatizado.

Los riesgos de reidentificación inherentes a los datos interoperables se entienden y mitigan, pero no de manera rutinaria.

Todos los datos se protegen mediante las técnicas adecuadas y se archivan de forma responsable o se eliminan permanentemente al final de su ciclo de vida previsto.

El acceso a conjuntos de datos confidenciales se monitorea y documenta para garantizar el rendimiento de cuentas sobre la protección de datos.

Antes de su integración, intercambio o procesamiento mediante análisis automatizado, todos los datos se evalúan para detectar riesgos de reidentificación u otros daños potenciales, y solo se utilizan cuando existe un alto grado de certeza de que los datos permanecerán seguros después de su reutilización.

La organización ayuda a establecer estándares para la protección de datos dentro del ecosistema de datos más amplio y defiende el uso responsable de datos.

Infraestructura digital

La organización enfrenta escasez de infraestructura digital clave para almacenar, gestionar, intercambiar y procesar datos, incluyendo componentes de hardware y software, suministro confiable de electricidad, o conectividad a Internet.

La organización tiene acceso adecuado a componentes clave de infraestructura digital, pero hay escasez de soluciones de gestión de contenido y almacenamiento de datos específicos de la organización, lo que resulta en una gestión de datos no estandarizada y no alineada

Todos los miembros del personal apropiados tienen sus propias computadoras y conectividad a Internet. Existen servidores seguros y repositorios de datos, pero el personal y las unidades organizativas los utilizan de manera inconsistente, hay poca supervisión de la infraestructura digital

Todos los miembros del personal pertinente conocen y están capacitados para utilizar el sistema de gestión de datos de una organización. Los servidores seguros y los repositorios de datos son utilizados habitualmente por los miembros del personal y la supervisión de la infraestructura digital es parte de la estrategia de datos de una organización.

La estrategia de datos de una organización incluye disposiciones para el mantenimiento, la revisión periódica y la actualización de su infraestructura digital, y las líneas presupuestarias se reservan para este propósito. Las políticas de adquisiciones, intercambio de datos y supervisión de la infraestructura están alineadas. La organización aborda los temas de infraestructura digital con un enfoque prospectivo, y desarrolla activamente estrategias y planes para optimizar el uso de tecnologías emergentes y mejorar la interoperabilidad de sus sistemas de datos

Seguridad cibernética y respuesta ante incidentes

Se tiene poco o ningún conocimiento de los riesgos de ataques cibernéticos u otras infracciones a los sistemas de datos de la organización, incluidos los riesgos específicos asociados con datos interoperables potencialmente reutilizables

No existe un protocolo o política para enfrentar violaciones de datos.

Personal y unidades diversos de la organización tienen conciencia o muestran preocupación por los riesgos que representaría para sus datos reutilizables un ataque cibernético u otra violación de datos.

Surgen campeones que promueven el establecimiento de un protocolo o política para enfrentar violaciones de datos

Se ha preparado un protocolo para enfrentar violaciones de datos dentro de la organización, pero los riesgos asociados con la posible reutilización de datos interoperables robados siguen siendo vagos y la política se entiende y aplica de forma inconsistente.

Se establece una política clara para enfrentar violaciones de datos que establece una secuencia de procedimientos y responsabilidades.

El personal recibe capacitación sobre lo que se debe hacer en caso de una violación de datos y se les enseña sobre los riesgos asociados con la reutilización de datos interoperables robados

La organización puede hacer frente a las violaciones de datos de manera rápida y efectiva, y toma medidas activas para garantizar que su infraestructura tecnológica sea lo más segura posible

La política para enfrentar violaciones de datos de la organización se revisa y actualiza periódicamente, y esta incluye explícitamente los riesgos asociados con la reutilización de datos interoperables

El personal apropiado recibe capacitación rutinaria sobre cómo responder a una violación de datos

APÉNDICE B: GLOSARIO

Los **modelos canónicos** de datos son modelos de datos que siguen patrones estandarizados específicos que los hacen altamente reutilizables y aptos para el intercambio de datos (Orrell 2018)

El **gobierno de datos** es la forma en que se ejerce control sobre los datos a lo largo de su ciclo de vida. El gobierno de datos rige todas las demás funciones de gestión de datos y asegura que los datos de manejen adecuadamente, en apego a políticas y buenas prácticas (DAMA International 2017).

La **gestión de datos** es el desarrollo, la ejecución y la supervisión de planes, políticas, programas, y prácticas que entregan, controlan, protegen e incrementan el valor de los activos de datos e información a lo largo de sus ciclos de vida (DAMA International 2017)

La **redundancia de datos** es el fenómeno que ocurre cuando un mismo dato existe en varios lugares al mismo tiempo. La redundancia de datos puede causar confusión e inconsistencia en la categorización de estos, reduciendo su utilidad y su valor.

La **custodia de datos** es la “práctica de gestionar diversos conjuntos de datos en nombre de otros, y buscando el mayor beneficio para la organización.” (DAMA International 2017)

Los **formatos legibles de manera mecanizada** son formatos de datos que pueden ser “leídos” y procesados por computadores. En general, los formatos legibles de manera mecanizada requieren que los datos estén formalmente estructurados.

Los **modelos de madurez** son herramientas que definen criterios y pasos para ayudar a que las organizaciones midan sus capacidades y su mejora continua en áreas o disciplinas específicas.

El **efecto mosaico** ocurre cuando varios conjuntos de datos se vinculan entre sí para revelar nueva información significativa. Mientras que tal información podría utilizarse para obtener mayor conocimiento, también podría ser utilizada por malos actores para causar daño. (OCHA 2020)

Los **datos sensibles de grupo** son observaciones que podrían no representar un riesgo para los individuos en caso de ser revelados, pero que podrían significar un riesgo para grupos vulnerables.

La **dependencia hacia proveedores exclusivos** ocurre cuando diferentes soluciones patentadas de datos y de informática se adquieren de manera aislada como parte de diferentes programas de trabajo o proyectos financiados por donantes externos, sin considerar las arquitecturas o la infraestructura digital existente. Cuando esto ocurre, se crean sistemas de información en silos, lo que produce sistemas o conjuntos de datos que no pueden interoperar o cuyos productos no pueden ser integrados entre sí.

BIBLIOGRAFÍA

- Bower, T. 2020. Boost Your Resistance to Phishing Attacks. *Harvard Business Review*. hbr.org/2020/09/boost-your-resistance-to-phishing-attacks
- CARE USA. 2019. Responsible Data Maturity Model for Development and Humanitarian Organizations. ictworks.org/wp-content/uploads/2019/11/Responsible_Data_Maturity_Model_10-16-19.pdf
- Chang, R. 2018. A Beginner's Guide to Data Engineering – Part II. medium.com/@rchang/a-beginners-guide-to-data-engineering-part-ii-47c4e7cbda71
- DAMA International. (2017). Data Management Body of Knowledge (2nd Edition). New Jersey: Technics Publications.
- Data Leaders. 2017. The Leader's Data Manifesto. dataleaders.org/manifesto/
- Development Gateway. 2019. Understanding National Data Ecosystems. developmentgateway.org/blog/understanding-national-data-ecosystems
- European Court of Justice (ECJ). 2020. The Court of Justice invalidates Decision 2016/1250 on the adequacy of the protection provided by the EU-US Data Protection Shield (press release No 91/20). curia.europa.eu/jcms/upload/docs/application/pdf/2020-07/cp200091en.pdf
- European Union (EU). 2016. General Data Protection Regulation (2016/679). eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32016R0679
- Gartner. 2015. Take Your Analytics Maturity to the Next Level. gartner.com/smarterwithgartner/take-your-analytics-maturity-to-the-next-level/
- Gonzalez, L. and Orrell, T. 2018. Data Interoperability: A Practitioner's Guide to Joining up Data in the Development Sector. data4sdgs.org/sites/default/files/services_files/Interoperability%20-%20A%20practitioner's%20guide%20to%20joining-up%20data%20in%20the%20development%20sector.pdf
- Google. 2020. Dataset Search. datasetsearch.research.google.com
- IBM. 2007. IBM Data Governance Maturity Model. lightsondata.com/data-governance-maturity-models-ibm/

- Jacobson, A. 2019. 5 Useful Things I Learned About Data Storytelling: Lessons from #Visualize2030. data4sdgs.org/news/5-useful-things-i-learned-about-data-storytelling-lessons-visualize2030
- Nissenbaum, H. 2010. Privacy in Context: Technology, Policy, and the Integrity of Social Life. crypto.stanford.edu/portia/papers/privacy_in_context.pdf
- Open Data Institute (ODI). 2015. A Guide to the Open Data Maturity Model: Assessing Your Open Data Publishing and Use. scribd.com/document/260481608/ODI-Maturity-Model-Guide-Assessing-your-open-data-publishing-and-use#download
- Open Data Institute (ODI). 2019a. Data Ethics Canvas. theodi.org/wp-content/uploads/2019/07/ODI-Data-Ethics-Canvas-2019-05.pdf
- Open Data Institute (ODI). 2019b. Anonymisation and Open Data: An Introduction to Managing the Risk of Re-identification. docs.google.com/document/d/1CoXniaTnQL_4ZyQuji9_MA_YCEEIQjx4zISEdB08c2M/edit
- Open Data Watch (ODW). 2016. The State of Development Data Funding. opendatawatch.com/wp-content/uploads/2016/09/development-data-funding-2016.pdf
- Orrell, T. 2018. Using Data to Join Up Development Efforts. data4sdgs.org/sites/default/files/services_files/Data%20Interop%20Brief_0.pdf
- Orrell, T. 2020. Joining-up Data for Universal Healthcare in Kenya: The View from the Ministry of Health. data4sdgs.org/resources/joining-data-universal-healthcare-kenya-view-ministry-health
- Palfrey, J. & Gasser, U. (2012). *Interop: The promise and perils of highly interconnected systems*. New York: Basic Books.
- Partnership in Statistics for Development in the 21st Century (PARIS21). 2020. Guidelines for Developing Statistical Capacity: A Roadmap for Capacity Development 4.0. paris21.org/sites/default/files/inline-files/UNV003_Guidelines%20for%20Capacity%20Development%20WEB_0.pdf
- The GovLab. 2020. Data Stewards. medium.com/data-stewards-network
- Thematic Research Network on Data and Statistics (TReNDS). 2019. Striking the Balance Between “Open by Default” and Targeted Data Sharing. sdsntrends.org/research/2019/3/4/maximizing-access-public-data
- United Kingdom (U.K.) Government. 2018. Open Standards Principles. gov.uk/government/publications/open-standards-principles/open-standards-principles
- United Kingdom (U.K.) Government. 2020. Data Ethics Framework. gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020#overarching-principles
- United Nations (UN). 2020. Principles on Personal Data Protection and Privacy. unsceb.org/principles-personal-data-protection-and-privacy

- UN Conference on Trade and Development (UNCTAD). 2019. Digital Economy Report 2019. Value Creation and Capture: Implications for Developing Countries. unctad.org/en/PublicationsLibrary/der2019_en.pdf
- UN Economic and Social Council (ECOSOC). 2013. Fundamental Principles of Official Statistics (E/RES/2013/21). unstats.un.org/unsd/dnss/gp/FP-Rev2013-E.pdf
- UN Global Pulse. 2020. Risks, Harms and Benefits Assessment. unglobalpulse.org/policy/risk-assessment/
- UN Office for the Coordination of Humanitarian Affairs (OCHA). Centre for Humanitarian Data. 2020. Exploring the Mosaic Effect on HDX Datasets. centre.humdata.org/exploring-the-mosaic-effect-on-hdx-datasets/
- UN Statistics Division (UNSD). 2020a. Data Stewardship: A Solution for Official Statistics' Predicament? unstats.un.org/unsd/statcom/51st-session/side-events/20200302-3A-high-level-forum-on-official-statistics/
- UN Statistics Division (UNSD). 2020b. Guidance on the Implementation of Open Data in National Statistical Offices. unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3v-Guidance_OD-E.pdf
- UN Statistics Division (UNSD). 2020c. UN Global Working Group on Big Data for Official Statistics. unstats.un.org/bigdata/

Characteristics of each increasing level of maturity

1 UNDEFINED

Interoperability is not recognized as a strategic objective

2 EMERGING

The ability to join up data is recognized but it is not explicitly identified as a strategic objective

3 LEARNING

Interoperability is identified as a strategic objective in an organization's technical units, but not outside of them

4 BUILDING

The need to join up data across systems is recognized as a strategic objective in an organizational data strategy

The value of standards and robust data governance is recognized

5 CONSOLIDATING

The strategic value that joined-up data can bring to decision-making is recognized in organizational strategies

Interoperability forms part of an organization's external engagement strategy with other data producers and users

Characteristics of each increasing level of maturity

1 UNDEFINED

Interoperability is not recognized as a strategic objective

2 EMERGING

The ability to join up data is recognized but it is not explicitly identified as a strategic objective

3 LEARNING

Interoperability is identified as a strategic objective in an organization's technical units, but not outside of them

4 BUILDING

The need to join up data across systems is recognized as a strategic objective in an organizational data strategy

The value of standards and robust data governance is recognized

5 CONSOLIDATING

The strategic value that joined-up data can bring to decision-making is recognized in organizational strategies

Interoperability forms part of an organization's external engagement strategy with other data producers and users