

TEMA 76. ESTRATEGIA DEL DATO. GOBIERNO DEL DATO, GESTIÓN DEL DATO Y GESTIÓN DE LA CALIDAD DEL DATO. TECNOLOGÍAS Y ESTÁNDARES PARA DESCRIPCIÓN DE CATÁLOGOS Y CONJUNTOS DE DATOS: SEMÁNTICAS Y ONTOLOGÍAS.

Actualizado a 26/06/2023

1. CONTEXTO NORMATIVO

En el ámbito de los datos, desde el punto de vista normativo, se está avanzando en el establecimiento de un marco jurídico que dote de seguridad a las transacciones realizadas dentro denominado mercado de datos.

A continuación, se incluye el conjunto de normas aplicables en España que de una forma u otra se encuentran relacionadas con los datos.

1. **Interoperabilidad:** [Real Decreto 4/2010 del Esquema Nacional de Interoperabilidad](#) y las [Normas técnicas de Interoperabilidad](#)
2. **Seguridad:** [Real Decreto 311/2022 del Esquema Nacional de Seguridad](#) y las [Normas técnicas de seguridad: Esquema Nacional de Seguridad \(Seguridad\)](#)
3. **Protección de datos personales:** [Reglamento \(UE\) 2016/679](#) Reglamento General de Protección de Datos (RGPD)
4. **Reutilización:** [Ley 37/2007 de Reutilización de información del Sector Público](#)
5. **Datos abiertos:** [Directiva \(UE\) 2019/1024](#) conocida como Open Data Act
6. **Datos abiertos:** [Real Decreto Ley 1024/2021](#) por el que se transpone la anterior directiva
7. **Ley de datos:** propuesta de reglamento de datos de la UE: [Data Act](#)
8. **Ley de Gobernanza de datos:** [Reglamento \(UE\) 868/2022](#), del 30 de mayo, denominado Data Governance Act.
9. **Oficina del Dato:** [Orden ETD/803/2020](#) por la que se crea la División Oficina del dato.

2. ESTRATEGIA DEL DATO

Los datos son una pieza fundamental en la sociedad actual y serán aún más relevantes en el futuro. Dada su importancia, en la actualidad, la **economía del dato** se considera uno de los ejes de crecimiento del **Mercado Único Digital** propuesto por la Comisión Europea en 2015, que se integra el Mercado Único Europeo. Esta economía tiene como objetivo fundamental transformar los datos de materia prima a factor de producción que impulse la economía.

Para que se pueda llevar a cabo este impulso se hace necesario, tanto en sector público como en el privado, contar con una estrategia del dato. Esta estrategia persigue conseguir un equilibrio entre la libre circulación de los datos y la protección de la privacidad, la seguridad y las normas éticas.

Teniendo esto presente, desde los diferentes países se está trabajando en desarrollar un marco jurídico que regule el uso que puede realizarse con los datos disponibles. Concretamente, en el **año 2020 la Comisión Europea** aprobó la **Estrategia Europea de Datos**. Esta estrategia persigue crear en la Unión Europea un mercado único para el trato del dato, de forma que todos los países de la Unión Europea empleen el mismo marco normativo, puesto que, según esta estrategia, que se describe más detalladamente a continuación, la economía del dato alcanzaría los **829.000 millones de € en 2025**, suponiendo un **5,8 % del PIB europeo** ese año.

2.1. ESTRATEGIA EUROPEA DEL DATO

La **Estrategia Europea del Dato** fue aprobada en el **año 2020** por la Comisión Europea como parte de la estrategia digital para Europa (Configurar el futuro digital de Europa). Su **objetivo** es convertir a la Unión Europea en líder de una sociedad dirigida por los datos. Busca dotar a los ciudadanos de derechos, herramientas y capacidades para mantener control sobre sus datos. Para ello, propone la **creación** de un **Mercado Único de Datos** que permita que éstos fluyan libremente por la Unión Europea y entre sectores, en beneficio de las empresas, los investigadores y las Administraciones Públicas. Así mismo, la estrategia también persigue que, gracias al uso de los datos, las personas, las empresas y las organizaciones estén más capacitadas para tomar mejores decisiones a partir del conocimiento que aportan esos datos no personales que deben estar a disposición de todos. Debido a todo esto, uno de los aspectos fundamentales que propone regular es la **gobernanza de los datos**.

Esta Estrategia forma parte de un paraguas más amplio de la Unión Europea como es su estrategia digital, denominada **Configurar el futuro digital de Europa** (también se denomina una Europa adaptada a la era digital). Dentro de esta estrategia digital, la Estrategia del dato, se ubica en el 2º pilar: una economía digital justa y competitiva.

Como parte de esta estrategia, desde la Unión Europea se ha aprobado un conjunto de normas que vienen a construir el marco jurídico de este mercado único de datos.

- **Ley de Datos abiertos (ODA):** Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo de 20 junio 2019 relativa a los datos abiertos y la reutilización de la información del sector público (transpuesta mediante el RD-Ley 24/2021 por el que se transponen varias directivas europeas). Se trata de una norma que busca ampliar la normativa vigente incorporando, entre otros aspectos, los denominados **conjuntos de datos de alto valor** y la incorporación de los **metadatos** a las obligaciones de reutilización (ver Tema 31).
- **Ley de Gobernanza de Datos (DGA):** Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos. Su objetivo es sentar las estructuras y los procesos para la creación de los espacios de datos. De esta forma, aumenta la confianza en los servicios de intermediación de datos y fomenta la cesión voluntaria y altruista de datos en la Unión Europea para el bien común.
- **Propuesta de Ley de Datos (DA):** propuesta de Reglamento de datos que se encuentra en fase enmiendas. Esta propuesta busca aclarar quiénes pueden generar valor a partir de los datos y bajo qué condiciones.

2.2. INICIATIVAS EUROPEAS Y NACIONALES

Con el objetivo de impulsar el mercado único de datos dentro de la Unión Europea, además del trabajo desarrollado en la elaboración del marco jurídico descrito anteriormente, se están llevando a cabo diferentes iniciativas, no solamente a nivel de la Unión Europea, sino también en el ámbito nacional, incluyendo tanto al sector público como al privado.

A nivel europeo, se está impulsando la creación de **espacios de datos** comunes e interoperables en sectores trectores que se constituyan como una pieza clave para el desarrollo del mercado de datos. Se trata de ecosistemas donde materializar la compartición voluntaria de los datos dentro de un entorno de soberanía, confianza y seguridad. Estos espacios cuentan con mecanismos de gobernanza, organizativos, normativos y éticos. Son ejemplos de estos espacios de datos, el proyecto **Gaia-X**, que se trata de una iniciativa del sector privado europeo y que cuenta con la participación de España. Nuestro país cuenta con un **Hub de Gaia-X en Talavera de la Reina** (Toledo). Más recientemente, también impulsado desde Europa, esta vez por la propia Unión Europea, se ha apostado por la creación de un **espacio europeo de datos sanitarios**. Este espacio busca poder intercambiar datos sanitarios de los ciudadanos de la Unión Europea permitiendo, por un lado, a los ciudadanos tener el control sobre sus datos y, por otro lado, a los gobiernos, poder trabajar de forma conjunta en la detección de enfermedades, mejorar la atención sanitaria y los tratamientos médicos, entre otros aspectos.

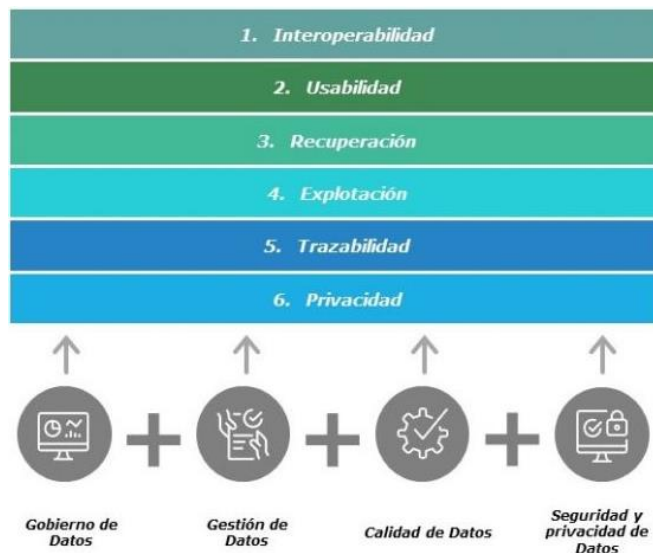
Siguiendo con los espacios de datos, en el ámbito nacional, cabe destacar el **espacio de datos del Turismo**. Se trata de un proyecto en el que se está trabajando actualmente para mejorar este **sector económico clave** en nuestro país. Adicionalmente, se cuentan con iniciativas para la creación de diferentes *Data lakes* nacionales: el **Data Lake de la AGE** o un **Data Lake sanitario** para los datos sanitarios procedentes de las diferentes CCAA y del Estado.

A nivel nacional, la economía del dato se impulsa en la agenda **España Digital 2026**, en concreto, en su “**Eje 4: Economía del Dato e Inteligencia Artificial**”. El desarrollo estratégico de infraestructuras de datos y de información estará liderado por la recién creada **Oficina del Dato**, perteneciente a la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA).

Haciendo referencia al ámbito de las Administraciones Públicas (AAPP), los datos tienen un claro encaje dentro del **Plan de Digitalización de las Administraciones Públicas 2021 - 2025**. Este plan se encuentra enmarcado en el Componente 11 del Plan de Recuperación, Transformación y Resiliencia (PRTR). En concreto, la **medida 6** del mismo impulsa la **gestión e intercambio transparente de datos**, que tiene como objetivo la gestión eficiente y transparente de los datos por parte de ciudadanos, ciudadanas y empresas, así como el desarrollo de servicios públicos basados en los mismos.

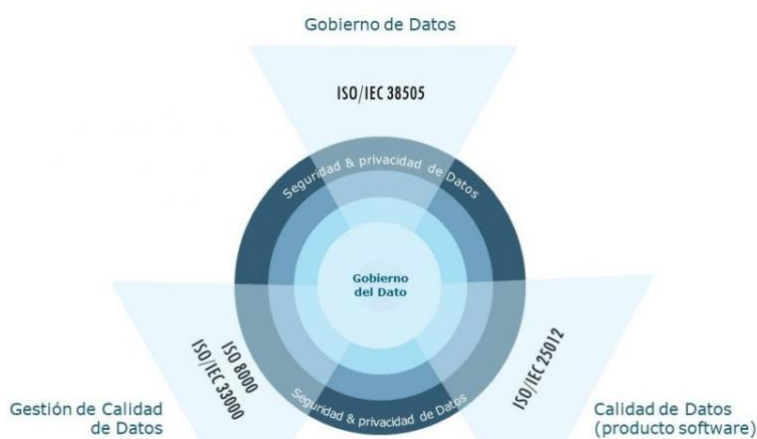
3. GOBIERNO DEL DATO

La implementación de un marco sólido de Gobierno del Dato para el éxito de la transformación digital, necesita el uso de las normas de gobierno, gestión, calidad, seguridad y privacidad de los datos, lo que proporciona 6 importantes beneficios.



La Asociación Española de Normalización UNE así como el resto de entes internacionales trabajan en potenciar la normativa técnica relacionada con los datos centrada en tres áreas además de la **Seguridad y privacidad de datos**.

- **Gobierno de Datos**
- **Gestión de Calidad de Datos**
- **Calidad de producto software (Datos)**



En el siguiente enlace se puede profundizar en el tema:

[Buen Gobierno del Dato gracias a los estándares \(une.org\)](https://www.unenorm.es/buen-gobierno-del-dato-gracias-a-los-estandares)

En el Gobierno del Dato se distinguen tres actividades:

- **Evaluar** el uso actual y el futuro de los datos.

- **Dirigir** la preparación e implementación de las estrategias y políticas de datos que aseguren que los datos satisfacen los objetivos de negocio. Asegurar el nivel adecuado de custodia del dato (**data stewardship**).
- **Monitorizar** el cumplimiento de las políticas y el desempeño con respecto a las estrategias.

Para el gobierno del dato se establecen 6 **principios**:

- **Responsabilidad**
- **La estrategia**
- **Adquisición**
- **Desempeño**
- **Conformidad**
- **Comportamiento humano**

Además de las estructuras existentes para el gobierno de la organización (por ejemplo, Consejo de Administración) y su gestión (por ejemplo, Consejo de Dirección), específicamente para el Gobierno del Dato podemos distinguir el:

Consejo de Gobierno del Dato (Data Governance Board), grupo de responsables de supervisar, soportar y monitorizar las iniciativas de Gobierno del Dato en la organización.

Otras estructuras que soportan a la anterior pueden ser el:

- **Comité de custodios del dato (Data Steward Committee)**, grupo de responsables de asesorar al consejo de Gobierno del Dato, y recopilar y proporcionar la información necesaria sobre los datos de la organización y su valor.
- **Oficina del dato**, que coordina el desarrollo de los diferentes componentes del sistema de gobierno y gestión del dato.

Como roles relacionados con el Gobierno del Dato destaca el **Director del dato (CDO, Chief Data Officer)**, dirige las iniciativas de datos en la organización y gestiona el cambio cultural necesario para que la organización logre optimizar el valor de sus datos.

Y existen muchos otros roles relacionados con la gestión del dato como: propietario del dato (**Data owner**), custodio del dato (**Data steward**), responsable de protección de datos (**DPO, Data Protection Officer**), arquitecto de datos, administrador de datos, administrador de bases de datos, etc.

3.1. NORMATIVA TÉCNICA

1.- GESTIÓN DE LA CALIDAD DE LOS DATOS

ISO 8000: marcos para mejorar la calidad de los datos. Incluye:

- ISO 8000-1, ISO 8000-2 y ISO 8000-8: Conceptos generales de la calidad de los datos
- ISO 8000-6: procesos de gestión de la calidad de los datos:
 - ISO 8000-60: visión general.
 - ISO 8000-61: modelo de referencia de los procesos de gestión.
 - ISO 8000-62: aplicación y evaluación de madurez de procesos organizacionales.
- ISO 800-100 a ISO 800-150: intercambio de datos maestros entre organizacionales.
- ISO 8000-311: guía para la aplicación de la calidad de los datos de la forma del producto.

ISO/IEC 25012: modelo general de calidad aplicable a datos almacenados de forma estructurada en un sistema de información.

2.- MEDICIÓN DE LA CALIDAD

ISO 25024: requisitos y evaluación de la calidad de los sistemas y el software (SQaRE)

3.- GOBIERNO DEL DATO

ISO/IEC 38505-1: marco de gobierno de datos y un mapa de responsabilidad de datos que identifica las áreas de la organización en las que debe aplicarse el gobierno de datos.

ISO/IEC 38505-2: implementación de la Norma ISO/IEC 38505-1 que proporciona orientación sobre el gobierno de datos.

4.- NORMAS TRANSVERSALES PARA LA SEGURIDAD Y PRIVACIDAD DE DATOS

4.1. Seguridad de la Información:

- ISO/IEC 27001: requisitos del SGSI (Sistema de Gestión de Seguridad de la Información).
- ISO/IEC 27002: código de prácticas del SGSI.
- ISO/IEC 27018: PII (información personalmente identificable) en la nube pública.

4.2. Privacidad y Protección de Datos:

- ISO/IEC 27701: requisitos del SGPI (Sistema de Gestión de la Privacidad de la Información)
- ISO/IEC 29100: marco de privacidad.
- ISO/IEC 29151: protección de la información personal. Código de prácticas.
- ISO/IEC 29134: evaluación del impacto de la privacidad.
- ISO/IEC 20889: técnicas de desidentificación de datos para la mejora de la privacidad.

4.3.- Seguridad y Privacidad por diseño/defecto:

- PNE-prEN 17529: protección de los datos y de la privacidad por diseño y por defecto.
- ISO/DIS 31700: protección del consumidor. Privacidad por diseño para bienes y servicios de consumo.

NORMAS UNE PARA UN CORRECTO GOBIERNO DEL DATO

1 GESTIÓN DE LA CALIDAD DE LOS DATOS

ISO 8000

Marcos para mejorar la calidad de los datos. Incluye:

Intercambio de datos maestros entre organizaciones

Guía para la aplicación de la calidad de los datos de la forma del producto

ISO 8000-100 a ISO 8000-150

ISO 8000-311

ISO 8000-1
ISO 8000-2 y
ISO 8000-8

Conceptos generales de la calidad de los datos

Procesos de gestión de la calidad de los datos

- ISO 8000-60: visión general.
- ISO 8000-61: modelo de referencia de los procesos de gestión.
- ISO 8000-62: aplicación y evaluación de madurez de procesos organizacionales.

ISO/IEC 25012

Modelo general de calidad aplicable a datos almacenados de forma estructurada en un sistema de información.

2 MEDICIÓN DE LA CALIDAD

ISO 25024:

requisitos y evaluación de la calidad de los sistemas y el software (SQuaRE).

3 GOBIERNO DEL DATO

- ISO/IEC 38505-1: marco de gobierno de datos y mapa de responsabilidad de datos que identifica las áreas de la organización en las que debe aplicarse el gobierno de datos.
- ISO/IEC 38505-2: implementación de la Norma ISO/IEC 38505-1 que proporciona orientación sobre el gobierno de datos.

4 NORMAS TRANSVERSALES PARA LA SEGURIDAD Y PRIVACIDAD DE DATOS

SEGURIDAD DE LA INFORMACIÓN

ISO/IEC 27001: requisitos del SGSI (Sistema de Gestión de Seguridad de la Información).

ISO/IEC 27002: código de prácticas del SGSI.

ISO/IEC 27018: PII (información personalmente identificable) en la nube pública.

PRIVACIDAD Y PROTECCIÓN DE DATOS

ISO/IEC 27701: requisitos del SGPI (Sistema de Gestión de la Privacidad de la Información).

ISO/IEC 29100: marco de privacidad.

ISO/IEC 29151: protección de la información personal. Código de prácticas.

ISO/IEC 29134: evaluación del impacto de la privacidad.

ISO/IEC 20889: técnicas de desidentificación de datos para la mejora de la privacidad.

SEGURIDAD Y PRIVACIDAD POR DISEÑO/DEFECTO

PNE-prEN 17529: protección de los datos y de la privacidad por diseño y por defecto.

ISO/DIS 31700: protección del consumidor. Privacidad por diseño para bienes y servicios de consumo.

3.2. LA OFICINA DEL DATO

La **Oficina del Dato**, dirigida por la figura de y su dirección, la figura del **Chief Data Officer (CDO)**, dependiente de la Secretaría de Estado de Digitalización e Inteligencia Artificial (SEDIA) nace con la misión de dinamizar la gestión, compartición y el uso de los datos a lo largo de los diferentes sectores productivos de la economía y sociedad española.

La Oficina del Dato es el **organismo competente** en materia de **gobernanza y compartición de datos**, para el desarrollo de una Economía del Dato de manera transversal.

El ámbito de actuación de la Oficina del Dato se alinea firmemente con las directrices y actuaciones definidas en la [Estrategia Europea para los Datos](#) (Comisión Europea, Feb 2020) y trabaja en la prescripción y adopción de diferentes modelos de gobernanza, estándares, buenas prácticas; así como criterios arquitectónicos, y de codificación y semántica para la gestión, compartición y explotación federada de datos, así como aspectos tecnológicos con que asegurar su seguridad y privacidad.

En materia de impulso a una transformación **data-driven** de las industrias sectoriales españolas, la Oficina del Dato trabaja de forma coordinada con los diferentes grupos de interés -tanto nacionales como europeos- que promueven el diseño y despliegue del paradigma de **espacio de datos**.

La Oficina del Dato promueve las siguientes especificaciones para habilitar un marco de referencia.

- **UNE 0077:2023** Gobierno del dato
- **UNE 0078:2023** Gestión del dato
- **UNE 0079:2023** Gestión de la calidad del dato
- **UNE 0080:2023** Guía de evaluación

3.3. METODOLOGÍAS GOBIERNO DEL DATO

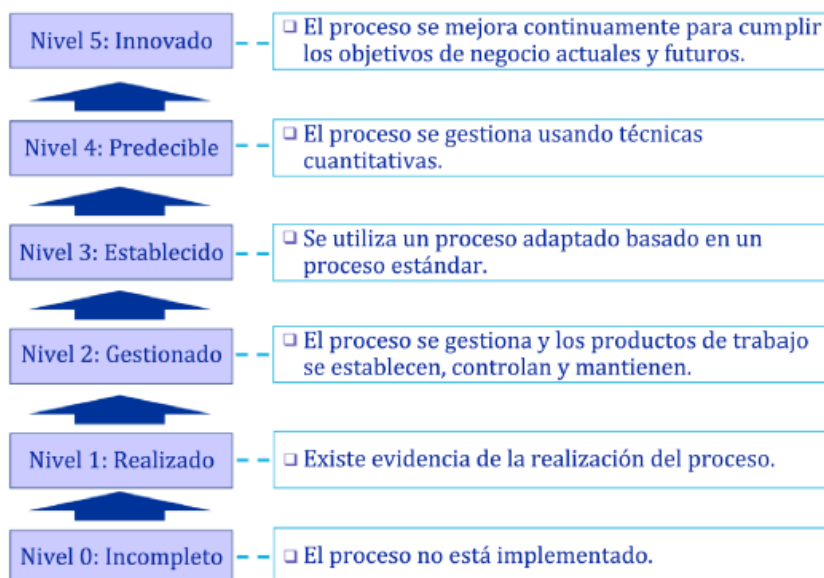
Como apoyo en la gobernanza de los datos, existen diversas metodologías, apoyo y modelos internacionales, como son:

- **DCAM.** Data Management Capability Assessment Model o modelo de evaluación para la capacidad de gestión de datos.
- **MAMD.** [Modelo Alarcos de Mejora de Datos.](#)
- **DGPO.** Data Governance Professionals Organization.
- **CMMI.** Capability Maturity Model Integration. Modelo para la mejora y evaluación de procesos para el desarrollo, mantenimiento y operación de sistemas de software.
- **DAMA.** Data Management Association. Asociación internacional para profesionales de la gestión de datos. **DMBoK.**

DAMA posiciona al Gobierno del Dato como principal actividad alrededor de la cual se gestionan el resto de actividades, como puedan ser arquitectura, interoperabilidad, calidad o metadato, como muestra la siguiente figura:



Con el fin de evaluar el nivel de madurez en la gestión de los datos, CMMI establece una calificación de nivel de 1 al 5:



3.4. GESTIÓN DEL DATO

Especificación UNE 0078:2023 Gestión del dato

La **gestión del dato** es el conjunto de actividades encaminadas a garantizar la entrega exitosa de datos relevantes y con niveles adecuados de calidad a los interesados correspondientes para el correcto desempeño de sus funciones, que consistirán fundamentalmente en el **procesamiento o en análisis del dato** como parte de los procesos de negocio establecidos en la estrategia organizativa.

La realización de una adecuada gestión del dato involucra el desarrollo de **trece procesos**:

- Procesamiento del dato
- Gestión de la infraestructura tecnológica
- Gestión de requisitos del dato
- Gestión de la configuración del dato
- Gestión de datos histórico
- Gestión de seguridad del dato
- Gestión del metadato
- Gestión de la arquitectura y diseño del dato
- Compartición, intermediación e integración del dato
- Gestión del dato maestro
- Gestión de recursos humanos
- Gestión del ciclo de vida del dato
- Análisis del dato



Figura 2 – Componentes de un Sistema de Gestión del Dato [1]

Algunas de las figuras relevantes para la gestión del dato son:

- **Director del dato (CDO, Chief Data Officer)**, dirige las iniciativas de datos en la organización y gestiona el cambio cultural necesario para que la organización logre optimizar el valor de sus datos.
- **Director de tecnología (CTO, Chief Technology Officer)**, que dirigirá las iniciativas relacionadas con la infraestructura tecnológica del dato.
- **Arquitecto del dato**, que estará a cargo de desarrollar y mantener la arquitectura de datos de la organización.
- **Director de seguridad (CISO, Chief Information Security Officer)**, que dirigirá las iniciativas relacionadas con la seguridad del dato.
- **Director o delegado de protección de datos (DPO, Data Protection Officer)**, que estará a cargo de las iniciativas relacionadas con la protección de los datos.

Existen otros muchos roles relacionados:

- Director de operaciones — COÓ, Chief Operating Officer.
- Director financiero – CFO, Chief Financial Officer.
- Roles sobre aspectos específicos de datos y su uso (data stewards, administrador de datos, administrador de bases de datos, desarrolladores de aplicaciones).
- Roles sobre aspectos de calidad del dato (director de calidad del dato).

3.5. GESTIÓN DE LA CALIDAD DEL DATO

Especificación UNE 0078:2023 Gestión del dato

Calidad del dato se define como el grado en que las características de los datos satisfacen necesidades declaradas e implícitas cuando se utilizan en condiciones específicas. Estas necesidades específicas se representan a través de los llamados **requisitos de calidad del dato**; los requisitos de calidad del dato se expresan en términos de las denominadas **características de calidad del dato**.

De acuerdo con el círculo de mejora continua **PDCA de Deming**, la gestión de la calidad del dato involucra **cuatro procesos**:

- **Planificación de calidad del dato**

- Control y monitorización de calidad del dato
- Aseguramiento de calidad del dato
- Mejora de calidad del dato

La calidad o **data quality** implica garantizar que cada dato reúne todos los atributos necesarios: **Exactitud, Completitud, Consistencia, Credibilidad, Actualidad, Cumplimiento, Confidencialidad, Eficiencia, Precisión, Trazabilidad, Comprensibilidad, Disponibilidad, Portabilidad, Recuperabilidad.**

El **aseguramiento de la calidad** de los datos es el proceso de verificación de la fiabilidad y efectividad de los datos, que debe realizarse periódicamente, y que incluye acciones como actualización, normalización o De-duplicación.

Existen una serie de medidas cuantitativas del data quality que son universales:

- **Completitud:** es el grado en el que todos los atributos del dato están presentes.
- **Validez:** representa el ajuste de un valor de datos a su conjunto de valores de referencia.
- **Unicidad:** la medida en que todos los valores distintos de un elemento de datos aparecen sólo una vez.
- **Integridad:** tiene que ver con el grado de conformidad con las reglas de relación de datos definidas.
- **Precisión:** determina en qué medida los datos representan correctamente la verdad sobre un objeto del mundo real o se ajustan a lo establecido por una fuente autorizada.
- **Coherencia:** representa el grado en que una pieza única de datos contiene el mismo valor a través de múltiples conjuntos de datos.
- **Oportunidad:** este atributo de la calidad de datos permite conocer si éstos están disponibles cuando se requiere.
- **Representación:** tiene que ver con el formato, patrón, legibilidad y utilidad de los datos para su uso previsto.

4. TECNOLOGÍAS Y ESTÁNDARES PARA DESCRIPCIÓN DE CATÁLOGOS Y CONJUNTOS DE DATOS

Para garantizar la interoperabilidad entre los distintos catálogos y conjuntos de datos, la UE cuenta con una **especificación para la descripción de catálogos de datos abiertos**. Esta especificación es la denominada **DCAT-AP**, que se basa en el estándar Data Catalog Vocabulary (DCAT) del **W3C**, que fue creado para facilitar la interoperabilidad entre catálogos de datos publicados en la web. El estándar DCAT permite disponer de un vocabulario de propiedades que posibilita describir catálogos y recursos de información.

DCAT-AP surge para resolver la interoperabilidad de los catálogos de datos y, en particular, los publicados mediante los portales de datos abiertos de las Administraciones Públicas en el contexto de la UE. Además, el uso de este vocabulario común también habilita la búsqueda cruzada de datos en diferentes catálogos, lo cual proporciona un claro valor añadido al sector infomediario.

Los tres conceptos clave que constituyen el DCAT-AP son:

- **Catálogo:** como colección de conjuntos de datos.
- **Conjunto de datos (dataset):** como colección de datos.
- **Distribución:** como forma concreta de acceder a un conjunto de datos específico (como un archivo en un formato determinado o una API con una tecnología concretas).

Una de las propiedades que caracteriza a este tipo de especificaciones y estándares es su **capacidad expresiva**, que se refiere al número de términos o propiedades disponibles en el vocabulario. Esta capacidad expresiva delimita la cantidad de características a describir sobre los catálogos y los recursos de datos que alberga.

Uno de los principales objetivos del DCAT-AP es **proporcionar un marco común descriptivo para catálogos "generalistas" de datos que albergan normalmente datos que proceden de catálogos de diferentes dominios de conocimiento**. Por ello, para poder describir con un mayor nivel de detalle un catálogo de datos temático, es mejor utilizar vocabularios especializados –por ejemplo, INSPIRE en el contexto geográfico-. En esta línea, con el objetivo de converger con estándares previamente establecidos, se desarrollaron dos extensiones a partir de DCAT-AP:

- **GeoDCAT-AP:** es una extensión de DCAT-AP que habilita la búsqueda de datos a través de varios portales de datos, como los ofrecidos por el Geoportal INSPIRE, por las Infraestructuras de **datos espaciales** a nivel nacional y por la comunidad geoespacial que utiliza estándares como ISO 19115/19139. Para su desarrollo SEMIC ha colaborado con el Joint Research Center, la DG CONNECT y la Oficina de Publicaciones de la UE.
- **StatDCAT-AP:** es una extensión de DCAT-AP que proporciona un vocabulario para **datos abiertos estadísticos** completamente compatible con los repositorios de datos basados en DCAT-AP. En su elaboración, SEMIC ha trabajado con Eurostat y la DG CONNECT.

Teniendo en cuenta lo anterior, se puede afirmar que el reto para elaborar una especificación que permita describir catálogos de datos abiertos generalistas es doble:

- Por un lado, la especificación debe permitir incorporar la **máxima capacidad descriptiva** para que los usuarios conozcan el máximo detalle posible asociado a los datasets que reutilizarán.
- Por otro lado, se deben **aprovechar los esfuerzos realizados por los editores de conjuntos de datos ya publicados** en catálogos temáticos existentes implementados en base a estándares previamente establecidos.

Para abordar estos retos, DCAT-AP ha evolucionado a una nueva versión, la versión 2.0.1 (publicada en junio de 2020), que aporta un considerable **incremento de propiedades para enriquecer la descripción de catálogos de datos, mejorar su interoperabilidad e incorporar metadatos de uso habitual en el contexto geográfico o estadístico**, como el linaje, la resolución de la información o la descripción de servicios de datos, entre otros. Algunas de las principales novedades que incorpora esta nueva versión de DCAT-AP son:

- **Servicios de datos.** Es una de las aportaciones más relevantes de esta evolución de DCAT-AP. Permite la descripción de servicios de datos, como APIs o Servicios Web, como alternativa a los metadatos utilizados para describir la disponibilidad de recursos descargables. La especificación contempla, además, que un servicio de datos -por ejemplo, un servicio visor de mapas- pueda tener más de un endpoint y, por tanto, más de una descripción para un único dataset.
- **Roles de agentes.** Es otro aspecto que en el contexto geográfico tiene amplio recorrido. De hecho, la extensión GeoDCAT-AP incorpora hasta 11 roles o responsabilidades definidas acorde a la iniciativa INSPIRE. DCAT-AP 2.0.1 admite tres roles de agente que aporta la posibilidad de especificar un punto de contacto o de agente al que se puede contactar para obtener información relativa al recurso, al creador o agente que creó el recurso y al editor o agente que publica y mantiene dicho recurso.
- **Versionado.** Abre la posibilidad a mantener la especificación de una secuencia entre diferentes versiones de un conjunto de datos. Aporta la posibilidad de describir las diferencias entre una versión y su predecesora.
- **Procedencia.** Con la incorporación de metadatos que describen la procedencia, ahora será posible considerar situaciones en las que estas descripciones se importan o exportan desde y hacia otros catálogos de datos -por ejemplo, de una IDE o de un portal de datos estadísticos-. Así se permite que los usuarios comprendan de dónde provienen los datos y cómo pueden haber sido modificados a lo largo del tiempo.
- **Granularidad.** Siempre que sea aplicable, es importante disponer de propiedades de contexto espacial y temporal para indicar resoluciones de datos -en unidades de longitud o de tiempo- para todos los tipos de recursos descritos en un catálogo, independientemente del contexto.
- **Gobernanza de datasets.** DCAT-AP 2.0.1 ofrece un mejor soporte para implementar modelos de gobernanza de datos incorporando, además de los relatados anteriormente, aspectos para identificar recursos agregados de varios publicadores, relaciones entre recursos y otros metadatos que permiten describir con precisión ciclos de vida de los datos, como la planificación de disponibilidad de un dataset o el estado de una determinada distribución en cuanto a su completitud, suspensión de disponibilidad u obsolescencia.
- **Uso de vocabularios controlados.** Se concretan referencias a vocabularios que viene desarrollando JoinUp para, entre otros ajustes de metadatos, modelos de licencia reconocidos internacionalmente, frecuencia de actualización de datasets, formatos de datos e idiomas, así como temáticas y subtemáticas de datos compatibles con la iniciativa INSPIRE.

- **Validación.** La especificación incorpora además una serie de plantillas SHACL para validar y controlar la calidad de grafos RDF de catálogos de datos compatibles con DCAT-AP 2.0.1.

Como se ha indicado en numerosos artículos relacionados con Open Data, **los metadatos son fundamentales para una reutilización eficiente de los datos**. El uso de vocabularios estándar para describir las características de los datos disponibles en los catálogos, -fruto del consenso entre los agentes del ecosistema de datos-, facilitan esa tarea y permiten que éstos sean interoperables entre sí. **La estandarización favorece el descubrimiento de los datos** mediante mecanismos de federación de metadatos entre catálogos.

En resumen, DCAT-AP 2.0.1 es un sólido paso adelante para crear catálogos más eficientes y el camino a recorrer en la evolución de las normas técnicas de referencia para describir recursos de información reutilizables.

4.1. NORMATIVA ESPAÑOLA

El marco normativo en España establece la Norma Técnica de Interoperabilidad de Reutilización de recursos de información (**NTI-RISP**) como la especificación de referencia para la puesta a disposición de recursos de información elaborados o custodiados por el sector público, cumpliendo con lo establecido en la Ley sobre reutilización de la información del sector público (**Ley 37/2007**, modificada posteriormente por la **Ley 18/2015** y la **Ley 9/2017** por la que se trasponen las **Directivas 2014/23/UE** y **2014/24/UE**). Este marco normativo vigente declara que los documentos de las Administraciones y organismos del sector público son reutilizables en los términos previstos en dicha Ley. Es relevante destacar que la NTI-RISP es anterior tanto a la primera versión DCAT (2014) como a la primera versión de DCAT-AP (2013). De este modo, en la hoja de ruta de la política de Datos Abiertos está la **evolución de la NTI-RISP vigente para adaptarla a la especificación DCAT-AP 2.0.1** y, como en el caso de la previa, se recomienda fijar en el marco regulatorio su obligado cumplimiento.