

CATÁLOGO DE DATOS ABIERTOS DEL ISTAC

Jesús Alberto González Yanes

Instituto Canario de Estadística, ISTAC

jgonyanp@gobiernodecanarias.org

Rita Díaz Adán

Arte Consultores Tecnológicos S.L., ARTE

rdiaada@arte-consultores.com

José Joaquín Escobar Gómez

Arte Consultores Tecnológicos S.L., ARTE

jescgom@arte-consultores.com

Francisco Javier Rodríguez García

Arte Consultores Tecnológicos S.L., ARTE

frodgar@arte-consultores.com

Miguel Jonás Bravo García

Arte Consultores Tecnológicos S.L., ARTE

mbragar@arte-consultores.com

PALABRAS CLAVE

Catálogo Datos, Open Data, Open Source, Datos Abiertos, CKAN

RESUMEN

La Directiva de la Unión Europea relativa a la reutilización de la información del sector público especifica que los documentos elaborados por los organismos del sector público de los Estados miembros constituyen un conjunto amplio, diverso y valioso de recursos que pueden beneficiar a la economía del conocimiento. Esta Directiva se transpone en la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.

Recientemente el Consejo de Europa ha adoptado una nueva normativa sobre los datos abiertos y la reutilización de la información del sector público. En esta nueva Directiva se introduce el concepto de conjuntos de **datos de alto valor**, que deberán facilitarse gratuitamente a través de una [interfaz de programación de aplicaciones \(API\)](#). El texto define seis categorías principales de conjuntos de datos de alto valor, e incluye a los **datos estadísticos** dentro de esta categoría. Asimismo, establece que los datos de investigación deben ser fáciles de encontrar, accesibles, interoperables y reutilizables (principios FAIR) y que los datos del sector público estarán disponibles de forma gratuita o a un coste muy bajo.

El Decreto por el que se aprueba el Plan Estadístico de Canarias 2019-2022 (PEC-22), establece que durante su ejecución se impulsará la reutilización de los datos estadísticos de conformidad con la Ley sobre reutilización de la información del sector público. A su vez indica que la Infraestructura de Datos y Metadatos Estadísticos (**eDatos**) será el

soporte para la difusión abierta e interoperable de los datos publicados por las operaciones del PEC-22, constituyéndose como el canal único para la difusión descentralizada de las estadísticas en las web corporativas del Gobierno de Canarias.

Para cumplir con las orientaciones señaladas, desde el Instituto Canario de Estadística (ISTAC) se pone a disposición de la ciudadanía el [portal de datos abiertos de la estadística pública en Canarias](#).

Entre las funcionalidades ofrecidas por dicho portal se sitúa el [catálogo de datos abiertos](#). Este sistema agrupa los conjuntos de datos reutilizables generados por la actividad del ISTAC, permitiendo así el acceso a recursos que ofrecen información pública de alto impacto y valor.

El uso de este tipo de sistemas resulta especialmente interesante de cara a la difusión de su contenido puesto que a través de federación se consigue que la información que figura en el catálogo de datos abierto también puede consultarse a través del [catálogo de datos abiertos de la Administración General del Estado](#) o del [portal oficial de datos europeos](#).

Cabe destacar que el [catálogo de datos abiertos de la estadística pública canaria](#) es actualmente el catálogo de datos autonómico que más conjuntos de datos reúne (más de 7.500 datasets) y lo que es más importante, conjuntos de datos actualizados de forma puntual y sincronizada con el resto de sistemas de difusión del ISTAC.

Manejar tal cantidad de datos ha supuesto un gran reto a nivel de organización y gestión de la información que han desembocado en diversas líneas de trabajo como son:

- Incorporar los conjuntos de datos al catálogo de forma automática y sincronizada con la publicación en el resto de sistemas.
- Meta-documentación y categorización adecuada de los conjuntos de datos de manera que se simplifique la búsqueda sobre los mismos.
- Organización y gestión de los conjuntos de datos de .
- Cumplimiento de la [NTI de Reutilización de recursos de información](#) y su [guía de aplicación](#).
- Optimización del sistema de almacenamiento para la gestión de grandes volúmenes de datos.

A nivel tecnológico, el desarrollo del catálogo de datos abiertos se ha basado en la reutilización del software CKAN. Sin embargo, han sido numerosas las modificaciones a realizar para poder cumplir con dos requisitos fundamentales:

1. Optimización del sistema para el manejo de grandes volúmenes de datos.
2. Mejora de la infraestructura para dar cumplimiento a la normativa de seguridad definida por el Gobierno de Canarias para el alojamiento de aplicaciones en servidores corporativos.