

# NODO DE INTEROPERABILIDAD ESTADÍSTICA

En España, se creó en **grupo de trabajo (InterSTAT) dentro del CITE** (Comité Interterritorial de Estadística) para establecer las bases jurídicas y técnicas del Nodo de Interoperabilidad Estadística, conforme al Real Decreto 4/2010, de 8 de enero, que regula el **Esquema Nacional de Interoperabilidad (ENI)** en la Administración Electrónica.

El grupo InterSTAT del CITE destacó que la **burocracia y los intercambios obsoletos** complican el acceso a datos estadísticos, y propuso una **estrategia nacional** para mejorar la interoperabilidad y **simplificar el intercambio de datos**, entre los distintos agentes.

Fruto de lo anterior, surge el presente proyecto que tiene como finalidad **definir un Nodo de Interoperabilidad Estadística**.

Este Nodo debe establecer **políticas de gestión de datos estadísticos** que permitan el intercambio de datos entre todos los **niveles de gestión españoles** (regional y nacional).

Asimismo, debe facilitar el **intercambio de datos con los organismos de la Unión Europea (EU)** y **mejorar la toma de decisiones** en materia de políticas públicas.

Todo esto se logra a través de **herramientas y procedimientos comunes y estandarizados**.

El proyecto ha sido abordado a través de **Instrumento de Asistencia Técnica de la Unión Europea (TSI)**, diseñado para ayudar a los Estados miembros a **implementar reformas que mejoren sus capacidades administrativas y técnicas**. Este apoyo es esencial para diseñar, desarrollar e implementar políticas públicas eficaces y sostenibles que respondan a los desafíos estructurales y coyunturales que enfrentan los países de la Unión Europea (UE).

Ha sido **liderado** por el Instituto de Estadística y Cartografía de Andalucía (**IECA**), Instituto Nacional de Estadística (**INE**) así como el Instituto Canario de Estadística (**ISTAC**) con el apoyo de la **DG-Reform**.



## Objetivos

El presente proyecto tiene como finalidad **definir un Nodo de Interoperabilidad Estadística** que establezca **políticas de gestión de datos estadísticos**.



Definición de **estándares, arquitecturas y metodologías comunes** de datos estadísticos para permitir el intercambio de datos entre agencias españolas y de la UE.



**Aumentar el valor** de los conjuntos de datos estadísticos para mejorar la toma de decisiones gubernamentales en relación con las políticas **públicas en España proporcionando conjuntos de datos de alta calidad**.



Probar la arquitectura del Nodo de **Interoperabilidad Estadística técnica y funcional** definida por los beneficiarios para perfeccionarla con las mejores prácticas y las lecciones aprendidas para futuras adopciones por organismos interesados.

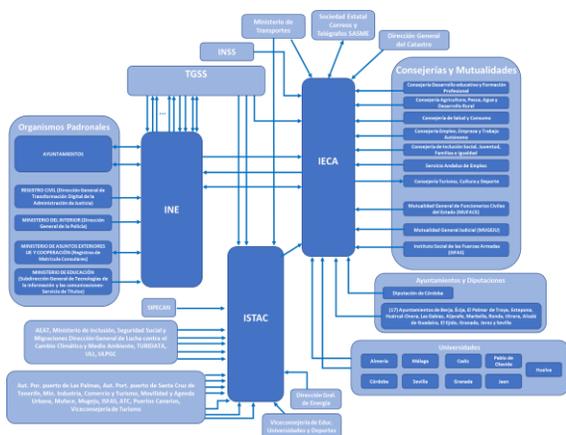
## Diagnóstico

Tras el **diagnóstico de la situación actual del intercambio de datos** estadísticos, entre los diferentes niveles de gestión administrativa en España, se detectó:

**1.- Procesos:** los intercambios de datos estadísticos se establecieron de manera reactiva, creando dependencias manuales y humanas, sin estandarización ni acuerdos formales en muchos de los casos, resultando en estructuras de datos dispares.

**2.- Tecnología:** los intercambios se soportan en sistemas rudimentarios y obsoletos que dificultan la interoperabilidad. Estos sistemas requieren intermediarios y múltiples puntos de acceso, lo que incrementa los costos de mantenimiento y evolución, además de impedir la trazabilidad y estandarización de los datos.

**3.- Marco legal:** la descentralización gubernamental en España provoca una dispersión normativa en la regulación estadística dificultando la homogenización y unificación de criterios. Además, los convenios reguladores del intercambio de datos entre administraciones son burocráticos y costosos, ralentizando el proceso y aumentando los costos administrativos.



A su vez, se detectó una gran **diversidad de flujos y procesos de intercambio** en los que están implicados los beneficiarios. Para una misma entidad, existían **diversos procedimientos y tipologías de archivos**, lo que evidencia la **necesidad de homogeneizar y estandarizar** los distintos intercambios, así como de **crear un Nodo común**.

## Requisitos funcionales

Tras el análisis de la situación actual, se identificaron los requisitos a los que debe dar respuesta el Nodo de Interoperabilidad Estadística, y que, por lo tanto, servirán como base para su diseño:



REQ1 - Gestor de convenios y acuerdos de aprovisionamiento



REQ2 - Catálogo de datos



REQ3 - Modelo de gobierno coordinado



REQ4 - Automatización de los procesos



REQ5 - Aplicación de reglas, validaciones y transformación



REQ6 - Canales seguros de intercambio



REQ7 - Integraciones sincrónicas/asíncronas



REQ8 - Flexibilidad a nuevas necesidades



REQ9 - Acceso de usuarios externos

## Solución seleccionada

Una vez **analizada** la **situación actual**, así como los **requisitos** que debe contener el futuro Nodo, se llevó a cabo un **análisis de mercado** para identificar las **soluciones disponibles** que se ajustaran a las **necesidades del sistema estadístico español**.

La solución que mejor se adapta a este proyecto es la **Plataforma de Datos de la Administración General del Estado**, basada en el producto IBM Cloud Pak for Data. Esta plataforma integra un módulo de compartición de datos con las siguientes funcionalidades:

### Funciones Plataforma del Dato de la Administración General del Estado

Confección del espacio de datos del sector público.

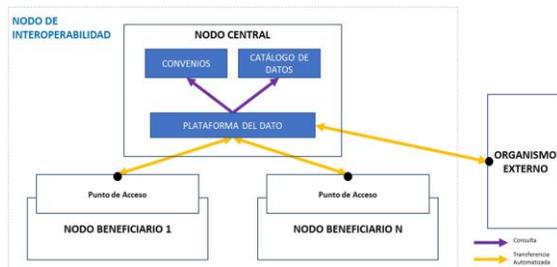
Acceso uniforme a los conjuntos de datos.

Generación de los catálogos de datos (productos de datos) de cada organismo.

Interoperabilidad entre los organismos y con otros espacios de datos sectoriales.

## Futuro Nodo

El futuro Nodo de Interoperabilidad Estadística será el **espacio nacional de datos estadísticos**, garantizando la **estandarización** de los procesos de intercambio y facilitando el acceso a los datos por parte de los diferentes actores involucrados en los procesos.



Se ha definido como un **nodo de nodos**, con el enfoque de un espacio de **datos federados**, donde se respeta la **soberanía digital** de cada uno de los organismos participantes. En este contexto, se ha establecido un **Nodo Central** encargado de **implementar, centralizar y gobernar** los **intercambios** de información estadística entre los diversos organismos participantes, ya sean autonómicos o nacionales.

## Casos de uso

### CU1 – Tipo PUSH: Intercambio INE → IECA

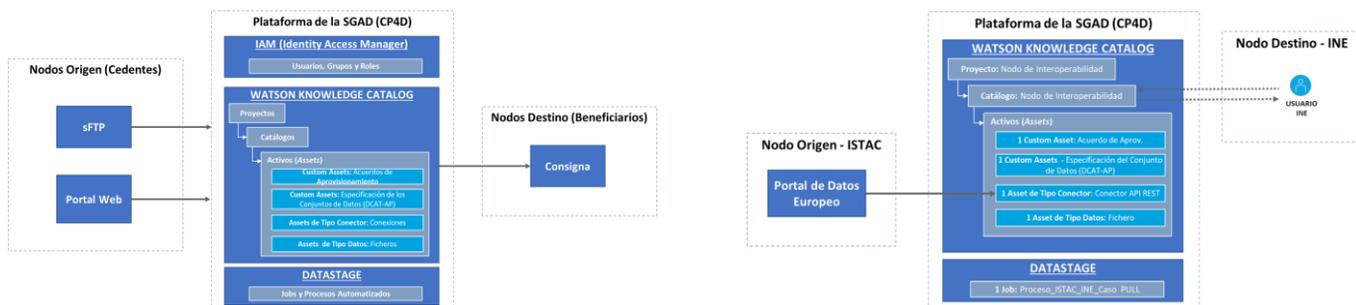
Intercambio en formato fichero y según los acuerdo de aprovisionamiento, con envío automatizado al nodo destino.

### CU2 – Tipo PULL: Intercambio ISTAC → INE

Intercambio con registro y publicación del fichero en el Catálogo, consulta parcial de la información en la plataforma.

### CU2 (bis) – Tipo PULL: Intercambio ISTAC → INE

Mismo intercambio que el CU2 pero la recepción del fichero es automática utilizando como nodo destino un sFTP.



## Conclusiones tras el desarrollo de la PoC



La plataforma **CUMPLE** con los **requisitos principales** detectados y evaluados durante la fase de análisis y diagnóstico.



Existen algunos **condicionantes principales que limitan**, pero **no bloquean**, el uso de la plataforma de la SGAD como solución técnica para la implementación del Nodo de Interoperabilidad Estadística:



La **conectividad** de los usuarios hacia el catálogo se habrá de realizar **siempre dentro de la Red Sara**, no podrá existir un catálogo público en Internet. Sí se podrá acceder a orígenes de datos o destino de datos fuera de la Red Sara, siempre y cuando dicha conectividad sea aprobada y gestionada por el equipo de la SGAD.



Existen **limitaciones en la configuración de la granularidad de la seguridad** de los elementos del catálogo y de las acciones que pueden realizar los usuarios en los mismos. Se espera que en la próxima versión 5 del producto (actual 4.8) haya una evolución significativa en este aspecto.



Para los **ficheros publicados en el catálogo**, **no existe la posibilidad de descarga directa de los mismos sin que estén persistidos en la plataforma**. Para la ejecución de intercambios programados, o que se ejecuten mediante el producto Datastage será necesaria una persistencia temporal de la información durante la ejecución del proceso. Nadie, excepto el propio proceso, tendrá acceso a dicha información.

## Lecciones aprendidas

De cara a la futura implantación del Nodo de Interoperabilidad Estadística, se tienen que tener en cuenta las siguientes lecciones o consideraciones, que se extraen del presente proyecto:

- No existe ningún producto de mercado **que se ajuste al 100% a los requerimientos técnicos del Nodo**. Por lo tanto, para cubrir estos requerimientos íntegramente, serían necesarios desarrollos a medida complementarios.
- La plataforma puede tener **requisitos técnicos específicos** que deben ser cumplidos para su correcta implementación y operación, **limitando las opciones disponibles para el despliegue**.
- Necesidad de **ajustes o actualizaciones** en materia normativa en la plataforma para mantenerse la conformidad.
- Se recomienda **estandarizar y que todos los puntos de acceso** utilicen las mismas tecnologías y protocolos.
- **La gobernanza** de los datos siempre **estará en manos de cada miembro** y el acceso a la información se realizará únicamente mediante los puntos de acceso locales.
- La plataforma deberá definir los **estándares definitivos de tratamiento de la información con los que se deberá trabajar**. Se ha propuesto el uso de DCAT-AP para la definición de metadatos de información, y DDI como definición de metadatos de campos.
- La gestión de identidades en el producto IBM Cloud Pak for Data tiene **limitaciones en la versión evaluada** (4.8). Se debería analizar dicha gestión en las versiones futuras del producto.