

Masó, J. & Pons, X. (2024). Green Deal Data Space: continuing with the INSPIRE open data, helping to create a digital economy, or both? / Espacio de Datos del Pacto Verde: continuar con los datos abiertos de INSPIRE, ayudar a crear una economía digital, o ambas cosas?, *GeoFocus, Revista Internacional de Ciencia y Tecnología de la Información Geográfica* (Editorial), 34, 1-4. <http://dx.doi.org/10.21138/GF.894>

GREEN DEAL DATA SPACE: CONTINUING WITH THE INSPIRE OPEN DATA, HELPING TO CREATE A DIGITAL ECONOMY, OR BOTH?

ESPACIO DE DATOS DEL PACTO VERDE: CONTINUAR CON LOS DATOS ABIERTOS DE INSPIRE, AYUDAR A CREAR UNA ECONOMÍA DIGITAL, O AMBAS COSAS?

¹JOAN MASÓ , ^{2,1}XAVIER PONS 

^{1,2}Grumets Research Group

¹CREAF. Universitat Autònoma de Barcelona. 08193, Bellaterra, Catalonia, Spain

²Departament de Geografia, Edifici B. Universitat Autònoma de Barcelona. 08193, Bellaterra, Catalonia, Spain

¹joan.maso@ieec.org ²xavier.pons@uab.cat

The INSPIRE (Infrastructure for Spatial Information in Europe) Directive set a legal framework to create a European Union Spatial Data Infrastructure (SDI) to support EU environmental policies and activities. The main objective was to enable the sharing of environmental spatial data by public sector organizations and facilitate open access throughout Europe. In order to assist in policy-making across boundaries, standard data models and protocols were defined to ensure compatibility and interoperability of spatial data across EU member states.

After 17 years of implementation, INSPIRE has successfully integrated the member states data catalogues but still faces some data accessibility issues: while metadata is progressively growing, the data is not always available (the complexities in the technical requirements have resulted in an irregular level of adoption among member states). As a consequence, the dream of having homogenous pan-European datasets has not materialized yet. There are concerns about the capacity of making more progress in an economic climate that makes more challenging to invest in new infrastructures and ways of working. Some people argue that the INSPIRE directive was too precise in the type of technologies needed to build the European SDI. INSPIRE should align with other horizontal data legislation now that technologies have evolved and spatial data is no longer considered in isolation. In addition, other people say that the technical specifications were too strict and demand a simplification of the technical requirements. Finally, the political focus has shifted more to the sustainability aspect of the environmental protection. Meanwhile, the European Commission has set new priorities for creating the foundations for a strong digital economy. While one of the pillars continues to be to make public sector data open (reinforced now with the definition of high value datasets), the target group for the resulting data applications has partially moved from policy monitoring to the commercial sector that can take advantage of this data accessibility for the development of new products, services, and business models in the digital economy (e.g., using open data to offer services on top of AI models). In this context, the European commission has launched the creation of several thematic data spaces including the Green Deal Data Space (GDSD). Data Spaces are more flexible than data infrastructures, allowing for a variety of data models and formats.

The data space is a concept that comes from the private sector. However, the European Commission is reinterpreting it and adapting it for the public sector to participate. This configures a mixture of a secure environment for data transactions among private companies and an open environment for the sharing of public data. The GDSD is focusing the interest in the European Green

Deal topics: climate change, circular economy, zero pollution, biodiversity, deforestation and compliance assurance. In practice, the GDDS interests are even broader than pure environmental data. The GDDS will provide the right infrastructure for a digital economy, for example around carbon stocks and circular products, but at the same time will incorporate open data from the public sector participating in INSPIRE. Apart from the public and commercial sector, the GDDS is an opportunity to integrate the academic sector and even the society at large, represented by the citizen science initiatives. This cross-pollination of ideas and information should stimulate economic growth and digital transformation.

From a technical point of view, the concept of data spaces comes from the general information technology. The International Data Space Association (IDSA) has proposed a standard Data Space Protocol that enables the exchange of data assets among any data space participants. This standardization reduces barriers to entry for new players and facilitates the creation of cross-sector applications. A catalogue based on DCAT (Data Catalog Vocabulary) metadata allows for the discovery of assets. The architecture of the data space fosters trust in the digital ecosystem by authenticating participants. Then a component called "connector" manages the access to data by negotiating "contracts" between the parties involved. The connectors should ensure data sovereignty allowing data owners to maintain control over their data while enabling sharing. However, none of the proposed technical solutions fully ensure sovereignty once the data has been transferred to the user. It is questionable that the adoption of the data space protocol will simplify data consumption for users and will allow for an easier integration with other data sources as the Data Spaces Protocol mandates the adoption of data connectors for the data consumers. Furthermore, the Data Space Protocol has no special provisions for the geospatial domain. Adapting the data space concept to the specificities of the geospatial data and geospatial services (*e.g.*, big volumes, dynamic data, long time series and geospatial filters) is the main challenge for the GDDS. **The geospatial community should provide answers to this challenge soon or there is a risk that geospatial data will not be fully adopted by the data spaces.** The new OGC Web APIs and the improvement of the semantic annotation of data are two elements that can be used to embrace the data spaces, but they only help in such a complex goal.

This editorial was possible in part thanks to the European Union projects AD4GD (HORIZON-CL6-2021-101061001) and MoRe4nature (HORIZON-CL6-2023-101133983).

La Directiva INSPIRE (Infraestructura para la Información Espacial en Europa) estableció un marco legal para crear una Infraestructura de Datos Espaciales (IDE) de la Unión Europea (UE) con el fin de apoyar las políticas y actividades medioambientales de la UE. El objetivo principal era permitir el intercambio de datos espaciales medioambientales por parte de organizaciones del sector público y facilitar el acceso abierto en toda Europa. Para ayudar en la formulación de políticas transfronterizas, se definieron modelos y protocolos de datos estándar y así garantizar la compatibilidad e interoperabilidad de los datos espaciales en todos los estados miembros de la UE.

Tras 17 años de aplicación, INSPIRE ha logrado integrar con éxito los catálogos de datos de los Estados miembros, pero aún se enfrenta a algunos problemas de accesibilidad a los datos: si bien los metadatos están aumentando progresivamente, los datos no siempre están disponibles (la complejidad de los requisitos técnicos ha dado como resultado un nivel irregular de adopción entre los Estados miembros). Como consecuencia, el sueño de tener conjuntos de datos paneuropeos homogéneos aún no se ha materializado. Existen preocupaciones sobre la capacidad de lograr más avances en un clima económico que dificulta la inversión en nuevas infraestructuras y formas de trabajo. Algunas personas sostienen que la directiva INSPIRE era demasiado precisa en el tipo de tecnologías necesarias para construir la IDE europea. INSPIRE debería alinearse con otras legislaciones horizontales sobre datos ahora que las tecnologías han evolucionado y los datos espaciales ya no se consideran de forma aislada. Además, otras personas afirman que las especificaciones técnicas eran demasiado estrictas y demandan una simplificación de los requisitos técnicos. Finalmente, el enfoque político se ha desplazado más hacia el aspecto de sostenibilidad de la protección del medio ambiente. Mientras tanto, la Comisión Europea ha establecido nuevas prioridades para crear las bases de una economía digital fuerte. Si bien uno de los pilares sigue siendo la apertura de los datos del sector público (reforzado ahora con la definición de conjuntos de datos de alto valor), el grupo objetivo de las aplicaciones de datos resultantes se ha

desplazado parcialmente del seguimiento de políticas al sector comercial, el cual puede aprovechar esta accesibilidad de los datos para el desarrollo de nuevos productos, servicios y modelos de negocio en la economía digital (por ejemplo, utilizando datos abiertos para ofrecer servicios basados en modelos de IA). En este contexto, la Comisión Europea ha puesto en marcha la creación de varios espacios de datos temáticos, incluido el Espacio de Datos del Pacto Verde (Green Deal Data Space, GDDS). Los espacios de datos son más flexibles que las infraestructuras de datos, lo que permite una mayor variedad de modelos y formatos de datos.

El espacio de datos es un concepto que proviene del sector privado, pero la Comisión Europea lo está reinterpretando y adaptando para que el sector público pueda participar en él. Esto configura una mezcla de un entorno seguro para las transacciones de datos entre empresas privadas y un entorno abierto para el intercambio de datos públicos. El GDDS centra el interés en los temas del Pacto Verde Europeo: cambio climático, economía circular, contaminación cero, biodiversidad, deforestación y garantía de cumplimiento. En la práctica, los intereses del GDDS son incluso más amplios que los datos medioambientales puros. El GDDS proporcionará la infraestructura adecuada para una economía digital, por ejemplo en torno a las reservas de carbono y los productos circulares, pero al mismo tiempo incorporará datos abiertos del sector público que participa en INSPIRE. Además del sector público y comercial, el GDDS es una oportunidad para integrar al sector académico e incluso a la sociedad en general, representada por las iniciativas de ciencia ciudadana. Esta fertilización cruzada de ideas e información debería estimular el crecimiento económico y la transformación digital.

Desde un punto de vista técnico, el concepto de espacio de datos proviene de la tecnología de la información en general. La Asociación Internacional de Espacios de Datos (IDSA) ha propuesto un protocolo estándar de espacio de datos que permite el intercambio de activos de datos entre todos los participantes del espacio de datos. Esta estandarización reduce las barreras de entrada para nuevos actores y facilita la creación de aplicaciones intersectoriales. Un catálogo basado en metadatos DCAT (Vocabulario de Catálogo de Datos) permite el descubrimiento de estos activos. La arquitectura del espacio de datos fomenta la confianza en el ecosistema digital al autenticar a los participantes. Luego, un componente llamado "conector" administra el acceso a los datos mediante la negociación de "contratos" entre las partes involucradas. Los conectores deben garantizar la soberanía de los datos, lo que permite a los propietarios de los datos mantener el control sobre ellos y, al mismo tiempo, permitir el intercambio. Sin embargo, ninguna de las soluciones técnicas propuestas garantiza por completo la soberanía una vez que los datos se han transferido al usuario. Es cuestionable que la adopción del protocolo de espacio de datos simplifique el consumo de datos para los usuarios y permita una integración más fácil con otras fuentes de datos, ya que el protocolo de espacios de datos exige la adopción de los conectores de datos también para los consumidores de datos. Además, el protocolo de espacio de datos no tiene disposiciones especiales para el dominio geoespacial. El principal desafío del GDDS es adaptar el concepto de espacio de datos a las especificidades de los datos y los servicios geoespaciales (por ejemplo, grandes volúmenes, datos dinámicos, series temporales largas y filtros geoespaciales). **La comunidad geoespacial debe proporcionar pronto respuestas a este desafío, o existe el riesgo de que los datos geoespaciales no sean adoptados por completo por los espacios de datos.** Las nuevas API web de OGC y la mejora de la anotación semántica de los datos son dos elementos que se pueden utilizar para abarcar los espacios de datos, pero pueden ayudar en este complejo propósito.

Este editorial ha sido posible en parte gracias a los proyectos de la Unión Europea AD4GD (HORIZON-CL6-2021-101061001) y MoRe4nature (HORIZON-CL6-2023-101133983).

References

Kotsev A, Minghini M, Tomas R, Cetl V, Lutz M (2020) From spatial data infrastructures to data spaces—A technological perspective. *ISPRS Int. J. Geo-Inf.* 9(3), 176. <https://doi.org/10.3390/ijgi9030176>.

Ponti, M., Portela, M., Pierri, P., Daly, A., Milan, S., Kaukonen Lindholm, R., Maccani, G., Peter De Souza, S. and Thabit Gonzalez, S. (2024) Unlocking Green Deal Data: Innovative Approaches for Data Governance and Sharing in Europe. *Publications Office of the European Union, Luxembourg*, 2024, <https://data.europa.eu/doi/10.2760/0517622>.

Conde, J., Pozo, A., Munoz-Arcentales, A., Choque, J. and Alonso, Á. (2024) Fostering the integration of European Open Data into Data Spaces through High-Quality Metadata, arXiv preprint arXiv:2402.06693

Kiourtis, A., Mavrogiorgou, A., & Kyriazis, D. (2024). A Cross-Sector Data Space for Correlating Environmental Risks with Human Health. *In* Papadaki, M., Themistocleous, M., Al Marri, K., Al Zarouni, M. (eds) Information Systems. EMCIS 2023. *Lecture Notes in Business Information Processing*, vol 501. Springer, Cham. https://doi.org/10.1007/978-3-031-56478-9_17.