

# GO-PEG. Generación de conjuntos de datos paneuropeos. Caso de uso aplicado a emergencias

*Pan-European dataset generation. Use case for emergencies*

Alejandro Guinea de Salas y Mayte Toscano Domínguez

REVISTA **MAPPING**

Vol.32, 214, 4-8

2023

ISSN: 1131-9100

## Resumen

El proyecto GO-PEG es un proyecto de investigación financiado por la comisión europea, que se ha realizado entre los años 2019 y 2023, y cuyo objetivo es proporcionar acceso a conjuntos de datos temáticos armonizados, y sus correspondientes metadatos en las áreas del medio ambiente, emergencias y gestión.

Se expondrá el caso de uso específico POPIMPACT, dentro del marco de GO-PEG donde se ha generado un conjunto de datos de alto valor para identificar la población en cada edificio de una zona afectada por una emergencia, documentando las lecciones aprendidas y buenas prácticas.

Se establece una metodología en PostGIS mediante distintas funciones que armonizan la información y asigna la población en función de distintos métodos seleccionables (huella del edificio, volumen y área residencial)

Los resultados se validan mediante los principios FAIR y se publican mediante OGC API en el Catálogo europeo de metadatos.

## Abstract

*The GO-PEG project is a research project funded by the European Commission, which has been carried out between 2019 and 2023. Its goal is to provide access to harmonized thematic datasets and their corresponding metadata in the areas of the environment, emergencies, and management.*

*A specific use case, known as POPIMPACT, will be presented within the framework of GO-PEG. In this context, a high-value dataset has been generated to identify the population in each building of an area affected by an emergency, documenting lessons learned and best practices.*

*A methodology is established in PostGIS using various functions to harmonize the information and assign the population based on different selectable methods (building footprint, volume, and residential area). The results are validated using the FAIR principles and are published through the OGC API in the European Metadata Catalog.*

Palabras clave: GO-PEG, Emergencias, Población, PostGIS, Principios FAIR, Datos de alto valor.

Keyword: GO-PEG, Emergencias, Population, PostGIS, FAIR principles, High-value data.

Alejandro Guinea de Sala. CEO de Geograma  
[alejandro.guinea@geograma.com](mailto:alejandro.guinea@geograma.com)  
Mayte Toscano Domínguez  
[mayte.toscano@geograma.com](mailto:mayte.toscano@geograma.com)

Recepción 20/11/2023  
Aprobación 12/12/2023

## 1. INTRODUCCIÓN

GO-PEG (Generation of cross border Pan European Geospatial Datasets and Services) es un proyecto cofinanciado por los fondos Connecting Europe Facility (CEF) de la Comisión Europea dentro de la acción 2018-EU-IA. Conecta a los principales expertos en armonización de datos dentro de un Consorcio formado por empresas de Italia, España, Alemania y Dinamarca, coordinado por la universidad de Leuven, Bélgica.

Los objetivos del proyecto son:

1. Proporcionar acceso a **conjuntos de datos abiertos temáticos armonizados** y a sus **metadatos**.
2. Centrado en las áreas de **Medio Ambiente y Gestión de Emergencias y Catástrofes**.
3. Cobertura geográfica **transfronteriza** en la que participan 5 Estados miembros de la UE: Bélgica, Alemania, Dinamarca, España e Italia.
4. Resultados **accesibles** mediante servicios web que permiten a los usuarios descubrir, visualizar y descargar datos.
5. Mejora y aceleración del valor de la **cultura de los datos abiertos**, impulsada por el Portal Europeo de Datos (EDP).

Cada parte integrante del consorcio elige uno o varios casos de usos para trabajar en las siguientes temáticas:

- **geo-COVID Watch**: Contribuir a una **comprensión de la pandemia COVID-19** utilizando mecanismos de acceso directo a datos de medición desarrollados en el sector medioambiental en los últimos años.
- **FUTUREFORST**: Técnicas que ayude a los propietarios y gestores forestales a **transformar los bosques** existentes como **consecuencia** del rápido **calentamiento del planeta** y las tensiones asociadas, como tormentas, plagas, sequías, etc.
- **POP-WILDFIRE**: **Evaluar la vulnerabilidad de la población al riesgo** potencial de **incendio** en una región determinada.
- **GO-DEPTH**: Metodología para **conceptualizar, organizar y entregar información del subsuelo** fácil de usar, de alta calidad e interoperable para la planificación y el uso sostenibles de los recursos naturales.
- **TRACING**: **Limitar los efectos y rastrear el origen de la contaminación del agua**.
- **FAIRPORT**: Conseguir un conjunto de datos geográficos de las **zonas geográficas**, tal como se indica en el artículo 21 del reglamento alemán, sobre el **espacio aéreo para drones**.
- **POPIMPACT**: Obtener una **capa de edificios paneuropeos** a los que se ha asignado la **población**.
- Nos centraremos en el caso de uso trabajado por Geograma, **POPIMPACT**, cuyo **propósito** es doble:

Por un lado, disponer de un conjunto de datos más precisos de la población afectada por un suceso de emergencia. En la actualidad, los gobiernos conocen la distribución de la población a nivel de edificio. Sin embargo, están obligados a preservar el secreto estadístico, por lo que no pueden distribuir estos datos y para ello lo agrupan en celdillas, secciones censales, municipios, etc. Escala que es demasiado genérica para la gestión de emergencias.

Y por otro, establecer una metodología para saber rápidamente qué población se ve afectada a nivel de edificio en caso de emergencia, con el fin de actuar proporcionalmente.

## 2. MATERIAL Y MÉTODO

Hasta ahora, se han realizado numerosas aproximaciones para intentar conseguir una **metodología** adecuada para estas **desagregaciones**.

- **Huella de los edificios** (Lwin y Murayama 2009)
- **Volumen de los edificios** (Lwin y Murayama 2009)
- **Techo edificable residencial** (Gálvez-Salinas et al. 2013): son necesarios datos urbanísticos.
- **Sistema Dasimétrico Experto de base catastral** (CEDS) (Maantay et al. 2007): Donde haya más viviendas habrá más habitantes.
- **Número de viviendas** (R.T. Mora-García et. 2015): Distribuye proporcionalmente la población en función del número de viviendas.
- **Superficie residencial** (R.T. Mora-García et. 2015): Distribuye la población proporcionalmente según la superficie residencial.

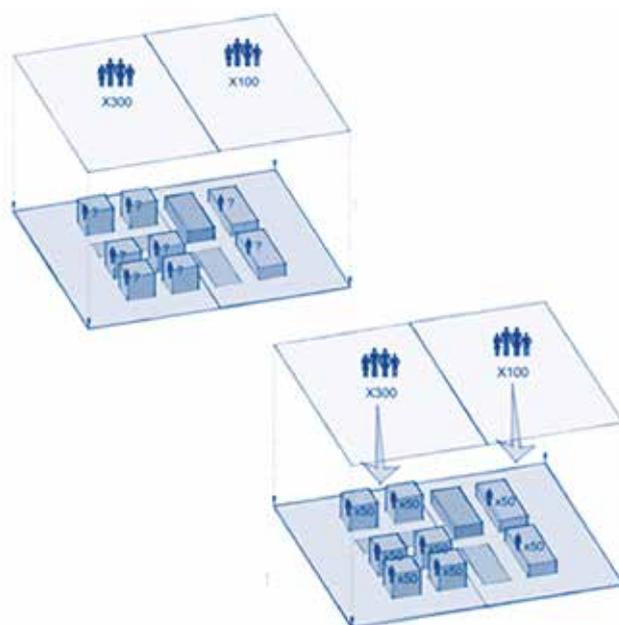


Figura 1. Asignación desde celdillas a edificios

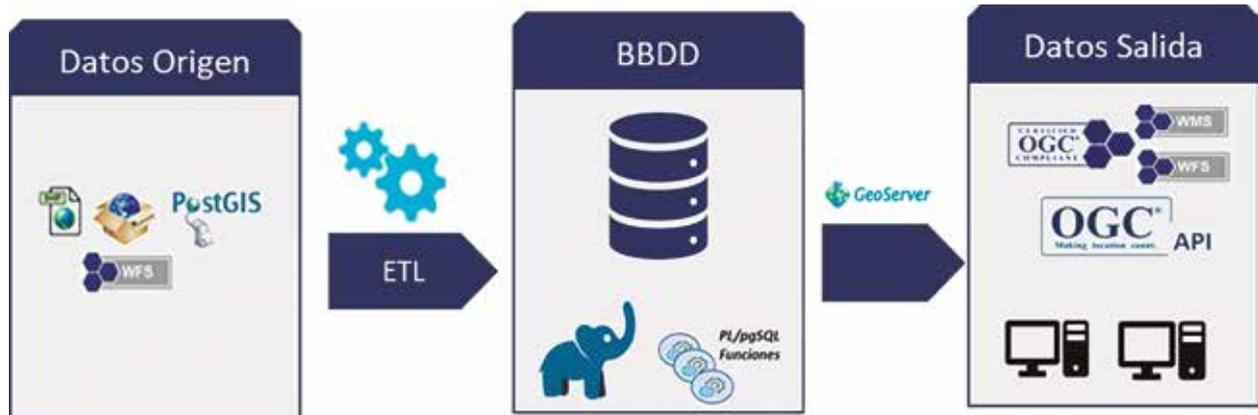


Figura 2. Arquitectura de POPImpact

En **POPImpact** hemos empleado **metodología** basada en **postGIS** donde **diversos métodos de desagregación** pueden ser **seleccionado en función a los datos disponibles** para conocer la población afectada por una emergencia. Los métodos de desagregación son:

- **Huella del edificio:** Reparte la población proporcionalmente al área de cada edificio.
- **Volumen del edificio:** Reparte la población proporcionalmente al volumen, ya que un mismo edificio tendrá más población en función al número de plantas.
- **Número de viviendas:** Reparte la población en función al número de viviendas de cada edificio, siempre que se disponga de este dato.
- **Área residencial:** Reparte la población teniendo en cuenta que un mismo edificio puede tener diversos usos, y seleccionando sólo el uso residencial siempre que se disponga de este dato.

Los datos se vuelcan a postGIS, en donde primero se le pasa una función que **armoniza** los datos y luego se le aplican los **métodos de asignación** seleccionados en una JSON paramétrico donde se le indica a las funciones cuáles son los datos más relevantes: EPSG de los datos, nombre de los atributos

que contiene el área, volumen, etc. De esta forma permitimos que **las funciones se adapten a distintos conjuntos de datos** que no tienen que tener el mismo esquema para ser procesados.

El resultado de estos procesos es la **obtención de un dataset de edificios** con el modelo de datos de **INSPIRE Building** extendido con los atributos de la **población asignada** en función al método o métodos elegidos. Además, a esta salida se incluye un metadato con una serie de indicadores o KPI que permiten tener una visión del resultado y del proceso:

- **Número de edificios:** Recuento del número de edificios en el input.
- **Número de áreas de población:** Recuento del número de áreas de población del input.
- **Número de edificios sin población:** Número de edificios que no tienen población asignada.
- **Número de áreas de población sin población:** Número de áreas con población inferior a 1.
- **Edificios con población mayor de 100:** Edificios que tienen una población mayor de 100.
- **Población sin asignar:** Recuento de población no asignada.



Figura 3. Casos de usos de POPImpact

## 2.1. Casos de usos

La metodología descrita se comprueba con 4 escenarios o casos de usos:

### Escenario 1. Sevilla: Diferentes datos, mismo alcance

Se asigna a los edificios de catastro de la ciudad de Sevilla distinta información de población. La proveniente de las celdillas de 1 kilómetro de Eurostat, del grid de población de 250 metros del IECA y de la población del INE de secciones censales.

Se aplica los siguientes métodos de agregación: huella, volumen y número de viviendas, y se observa que para los distintos datos de entrada los resultados no varían significativamente.

### Escenario 2. Francia: Grande conjunto de datos

Se realiza este escenario para validar cómo se gestiona un gran conjunto de datos, como es el caso de todos los edificios de Francia. Se procesa toda la información con el método de asignación de la huella del edificio, este proceso tiene una duración total de 8 días, lo cual es ampliamente mejorable ya que el sistema debe dar respuesta rápida en caso de emergencias.

### Escenario 3. El volcán de La Palma: Una zona de riesgo real

Partiendo de los edificios catastrales de la zona afectada por el volcán de La Palma y la población de 1 kilómetro de Eurostat, se aplica los siguientes métodos de agregación: huella, volumen y número de viviendas.

Los resultados se comparan con los datos oficiales de las personas afectadas por el volcán del Instituto canario de estadística (ISTAC) recogiendo un porcentaje de acierto de 86,51 en el caso del método de huella, un 98,19% con el método del volumen y un 89,06 con el método de número de viviendas.

### Escenario 4. Suiza-Francia: Datos fronterizos

Se obtiene un conjunto de datos transfronterizo para validar cómo el sistema es capaz de procesar información de

diversas fuentes y distintos modelos de datos. El resultado es un dataset de edificios armonizado con Inspire y al que se le ha asignado la población de 1 kilómetro de Eurostat.

## 2.2 Principios FAIR

Los conjuntos de datos resultantes y sus metadatos de **POPImpact** se publican en formato OGC (WMS; WFS y OGC API) en los portales de datos espaciales de [España](#) y de [Europa](#).

GO-PEG se desarrolla con un enfoque que garantiza el cumplimiento de los principios **FAIR**, que se refieren a la búsqueda (Findable), accesibilidad, interoperabilidad y reutilización. Para lograr esto, se llevan a cabo las siguientes acciones:

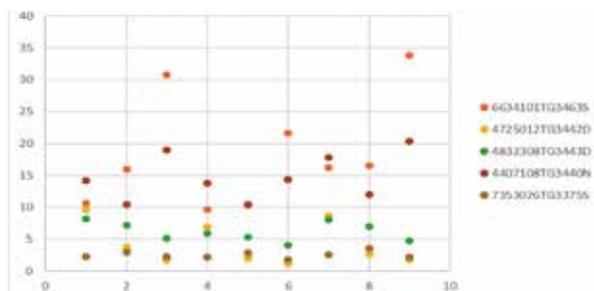
- **Findable:**
  - Crear metadatos legibles por máquina (descripciones detalladas)
  - Seguir estándares (GeoDCAT-AP, ISO19139)
  - Publicar a través de Catálogos Nacionales de Datos en el Portal Europeo de Datos
- **Accesible:**
  - Condiciones de licencia estándar
  - Servicios de datos basados en estándares (OGC GEOAPI, WFS...)
  - Consultas de datos habilitadas, para que los usuarios descarguen "justo lo que necesitan".
- **Interoperable:**
  - Modelos de datos bien documentados (INSPIRE)
  - Registro de listas de códigos para mejorar la interoperabilidad semántica
  - Identificadores únicos para datos y metadatos
- **Reusable:**
  - Metadatos enriquecidos para datos y servicios, liberados con una licencia transparente, adheridos a normas y publicados en EDP
  - Datos disponibles en múltiples formatos
  - Modelos de datos bien documentados: especificaciones de datos INSPIRE
  - Uso de codificaciones alternativas INSPIRE (GeoPackage, GeoJSON).
  - Los servicios API del OGC, facilitan a personas y máquinas el descubrimiento, la comprensión y la interacción con los datos.

Para asegurar que los resultados cumplen los valores FAIR, se valida los metadatos con la herramienta FAIR Enough obteniendo un resultado de una conformidad del 62,5% detectando los siguientes problemas:

El identificador único del recurso de metadatos no es persistente y no está contenido en los propios metadatos.

El recurso no está indexado en los motores de búsqueda: Buscador DuckDuckGo, API de búsqueda personalizada de Google y buscador Bing.

Pérdida de los datos de la licencia: Esta información se



	POC 1			POC 2			POC 3		
	POC1 H. Huella	POC 1 H. Volumen	POC 1 H. Viviendas	POC2 H. Huella	POC 2 H. Volumen	POC 2 H. Viviendas	POC 3 H. Huella	POC 3 H. Volumen	POC 3 H. Viviendas
6634101TG34635	10,64	15,948	30,726	9,69	10,306	21,672	16,2727	16,5369	33,79900
4725012TG3442D	9,658	3,725	1,712	6,848	1,999	1,21	8,69720	2,67817	3
4812308TG3443D	8,179	7,182	5,136	5,908	5,356	3,998	8,04515	6,98751	6
4407108TG3444D	14,15	10,475	19,024	13,78	10,438	14,354	17,7626	11,9597	20,32369
7353026TG33755	3,307	3,004	2,258	2,202	2,905	1,737	2,52141	3,53984	2,174204

Figura 4. Resultado caso de usos de Sevilla

pierde en la transformación entre Metadatos INSPIRE a GeoDCAT-AP, realizada por data.europa.eu.

## CONCLUSIONES

Las lecciones aprendidas de POPIMAPCT se pueden resumir en estas grandes líneas:

**Datos de gran valor:** Se obtiene un conjunto de datos de valor para emergencias, pero esta aproximación se podría mejorar añadiendo más variables de edad tales como edad, situación laboral o escolar, para saber si en un determinado momento la persona está en su casa o no, datos de consumo, etc.

**Fiabilidad:** Se parte de la idea de que el resultado de POPImpact es siempre una aproximación, sólo los estudios que se basan en la posición de la población en tiempo real mediante el análisis de la posición de ellos teléfonos móviles, son los que más se pueden acercar a conocer qué población está afectada en un determinado caso de emergencia, pero son datos privados y no están disponibles en la actualidad.

**Mejorar los datos Inspire:** Los datos de los edificios del modelo de datos Inspire ofrecen todo lo necesario para los cálculos realizados en POPImpact, ya que recogen datos de área, número de plantas y número de viviendas, sin embargo muchos de estos datos no son obligatorios y muchos países no los rellenan. Se debería potenciar que no sólo los atributos obligatorios de los modelos de datos Inspire se complementen, porque si aspiramos a obtener datos de alto valor debemos partir de conjuntos de datos de calidad.

**Identificadores persistentes:** Se debe mejorar el acceso a los catálogos de datos espaciales con identificadores persistentes que mejoren la búsqueda y accesibilidad de los mismos.

**Mejorar el tiempo de procesamiento:** Es necesario mejorar la arquitectura propuesta por POPImpact mediante procesamiento escalable en la nube que de una rápida respuesta de procesamiento en caso de emergencias.

## AGRADECIMIENTOS

GO-PEG (Generation of cross border Pan European Geospatial Datasets and Services) es un proyecto cofinanciado por los fondos Connecting Europe Facility (CEF) de la Comisión Europea dentro de la acción 2018-EU-IA. Se agradece la participación en el mismo del siguiente consorcio:

- Universidad de KU Leuven (<https://ees.kuleuven.be/en/sad/>)
- Epsilon Italia (<https://www.epsilon-italia.it/>)
- Wetransform (<https://www.epsilon-italia.it/>)

- Geokom (<http://www.geokom.dk/om>)
- Billbomática (<https://www.billbomatica.es/es>)

## REFERENCIAS

- Gálvez-Salinas, J.A.; Fischer, J. and Valenzuela-Montes, L.M. (2013): "Methodology for the spatial disaggregation of demographic information in urban areas through the development of modifiable spatial units". *GeoFocus*, Vol. 13 (no. 1), pp. 337-366.
- Mora-García, R. and Martí-Ciriquian, P. (2015): "Desagregación poblacional a partir de datos catastrales", in de la Riva, J., I. de la Riva, J. (2015). *la Riva, J., Ibarra, P., Montorio, R., Rodrigues, M. (Eds.) 2015: Spatial analysis and geographical representation: innovation and application. University of Zaragoza and AGE, pp. 305-314. XIV Congress of the Association of Spanish Geographers, 28-30 October, Zaragoza. Spatial analysis and geographical representation: innovation and application*  
[https://www.academia.edu/1007842/Mapping\\_population\\_distribution\\_in\\_the\\_urban\\_environment\\_the\\_cadastral\\_based\\_expert\\_dasymmetric\\_system\\_CEDS](https://www.academia.edu/1007842/Mapping_population_distribution_in_the_urban_environment_the_cadastral_based_expert_dasymmetric_system_CEDS)  
[https://www.researchgate.net/publication/220605975\\_A\\_GIS\\_Approach\\_to\\_Estimation\\_of\\_Building\\_Population\\_for\\_Micro-spatial\\_Analysis](https://www.researchgate.net/publication/220605975_A_GIS_Approach_to_Estimation_of_Building_Population_for_Micro-spatial_Analysis)

### Sobre los autores

#### Alejandro Guinea de Salas

CEO y director del Departamento de Consultoría de Geograma. Se graduó en Ingeniería Geomática y Topografía en la Universidad del País Vasco. Para seguir aprendiendo, también realizó un máster en Geotecnologías Aplicadas, Urbanismo y Arquitectura en la Universidad de Salamanca. Es miembro de grupos de trabajo españoles de la Directiva INSPIRE. También es el presidente del Comité Ejecutivo de EUROGI (Organización Europea de Coordinación en Información Geográfica).

#### Mayte Toscano Domínguez

Consultora GIS en Geograma. Técnico Superior en Desarrollo de Aplicaciones Web, Licenciada en Historia por la Universidad de Huelva y Máster en Sistemas de Información Geográfica: Planificación, Ordenación del Territorio y Medio Ambiente. Trabaja en la definición de procesos y estándares en Geograma, así como en el análisis funcional de proyectos de información geográfica o el desarrollo de aplicaciones geoespaciales.