

## Ficha de reto

### EUSTAT

¿Cómo podríamos **automatizar la normalización y codificación de las direcciones postales de las fuentes externas con el objetivo de integrar la información en nuestro sistema?**

#### Sub-retos:

- ¿Cómo podríamos incorporar un indicador de fiabilidad de los resultados obtenidos por cada elemento enlazado?
- ¿Cómo podríamos visualizar para las uniones con más de una posibilidad, opciones con mayor probabilidad, para su posible selección final de forma manual?
- ¿Cómo podríamos utilizar información adicional que, consultando con otras bases de datos de EUSTAT, (registro estadístico de población, directorio de actividades económicas etc.) permitan mejorar la identificación y los indicadores de fiabilidad?

#### Contexto:

EUSTAT es el organismo autónomo administrativo de la C.A. de Euskadi que desarrolla, produce y difunde información estadística oficial de calidad, objetiva y contrastada científicamente destinada a las administraciones públicas vascas, a los agentes sociales y a la sociedad.

Dentro de este marco de actuación, EUSTAT es el encargado de facilitar las tareas de codificación geográfica a las diferentes operaciones estadísticas del propio instituto y de otros Organismos, así como la implantación de una homogeneización para dichos códigos. Para ello, actualmente se reciben datos y ficheros de fuentes externas diversas con información de direcciones postales no codificadas. Con el objetivo de poder integrar la información de estas fuentes externas, se necesita una herramienta que codifique las direcciones postales en base a las codificaciones estandarizadas generadas en EUSTAT.

#### Objetivos:

Buscamos soluciones que al menos cumplan uno de estos objetivos:

- Codificar cada parte de una dirección postal hasta el nivel más bajo posible de la jerarquía de los elementos que conforman una unidad territorial estadística (UTE), teniendo en cuenta los componentes de dirección postal (territorio histórico, municipio, entidad/núcleo, calle, portal, planta, puerta) con la mayor fiabilidad.
- Ofrecer fiabilidad de los datos y en los casos en que la fiabilidad esté por debajo de un baremo acordado, tener la posibilidad de visualizar un subconjunto de candidatos con la mayor posibilidad y tener la opción de seleccionar uno de ellos.

- Facilitar la utilización de esta solución para los proyectos de las distintas áreas de EUSTAT (como directorios, territorio, población).

### ¿Qué buscamos?:

Se valorarán los siguientes aspectos:

- El éxito y fiabilidad de la codificación debe ser alto.
- La solución debe ser capaz de trabajar con otras bases de datos Oracle, donde se encuentra la información.
- No debe consumir excesivos recursos del sistema.
- Debe poder integrarse fácilmente en nuestro sistema.
- Debe ser apta para ser utilizada por cualquier tipo de usuario/a.

### Proceso y fechas clave:

Para participar en el reto, inscríbete en este enlace [pendiente de recibir].

- La fecha límite para recibir la respuesta a la propuesta es: 02/05/2024 a las 23:59h.
- Se comunicarán las startups semifinalistas el 10/05/2024.
- Se comunicarán las startups finalistas el 24/05/2024.
- Se comunicará la startup ganadora el 13/06/2024.

### El proceso de selección consta de 3 fases:

1. **Fase 1:** pitches de semifinalistas con la sociedad. Se seleccionarán tres startups como finalistas. Tendrán lugar del 16 al 24 de mayo de 2024.
2. **Fase 2:** ronda de entrevistas de finalistas con la sociedad. Tendrán lugar la semana del 27 de mayo de 2024.
3. **Fase 3:** reunión final con finalistas para la selección de la startup ganadora para el piloto con la sociedad. Tendrá lugar entre el 11 y 12 de junio de 2024, durante el programa de aceleración.

### A qué accedes:

#### Startups finalistas (3)

- **Programa de aceleración** govtech de 2 días. Se llevará a cabo el 11 y 12 de junio de 2024 de forma presencial en Bilbao.

#### Startup ganadora

- **Presentación de la solución en el Demo Day de BIND GovTech 2024**, de forma presencial en Bilbao.
- **Piloto remunerado** con un máximo de 15.000 euros. El piloto para implementar la solución ganadora comenzará en Julio de 2024 y tendrá una duración de 5 meses.