



Actualización selectiva de la red de carreteras: desarrollo de nuevos Bots dentro del sistema Cartobot

N.º del tema de las jornadas: 1. Herramientas y tecnologías

Resumen:

En la presente publicación se presentan los avances más significativos en el desarrollo de nuevos BOTS, que tienen por objeto automatizar la búsqueda de nuevas actuaciones en carreteras, y su contribución al incremento de la agilidad en la actualización de la capa vectorial de la red viaria de la Información Geográfica de Referencia de Redes de Transporte (IGR-RT).

Como ya se presentó en pasadas ediciones de este foro, el sistema CARTOBOT, creado mediante un convenio de colaboración entre el Instituto Geográfico Nacional (IGN) y la Universidad Politécnica de Madrid (UPM), logró, mediante técnicas de Web Scraping, el rastreo con éxito de páginas Web de organismos oficiales. Dicha extracción de datos permitió obtener un conjunto de incidencias o indicadores de cambios geométricos y semánticos en los viales.

En 2021 el equipo de Redes de Transportes del IGN ha desarrollado dos nuevos bots, ampliándose así los 15 ya existentes en el sistema CARTOBOT. El primero extrae incidencias geográficas de la página web de la Consejería de Infraestructuras y Movilidad de Galicia y el segundo extrae incidencias de la web de la Conselleria de Infraestructuras de Catalunya.

Dentro de la disciplina de Web Scraping, la técnica utilizada en el desarrollo de estos dos nuevos bots es la de rastreo sobre página Web estática con dimensión horizontal y vertical. Para ello, se ha utilizado la librería scrapy de Python, que provee una clase CrawlSpider que posee los atributos y métodos necesarios para realizar con éxito y de manera automatizada la extracción de incidencias geográficas sobre los viales.

Así, el bot hace una petición a la URL de ambas páginas Web, accediendo primero a cada una de las páginas horizontales en las que se estructuran y, después, a cada una de las incidencias en las vías, o lo que en la terminología Web Scraping se denomina "la paginación vertical". A continuación, de cada incidencia ligada a la vía se extraen los seis pares clave-valor que la definen



unívocamente: nombre de la vía, descripción, fecha de la obra, puntos kilométricos, municipio y geometría. Siguiendo con algunas de las técnicas ya utilizadas en los bots existentes, también se han empleado las expresiones regulares que mejoran el filtro de búsqueda.

El porcentaje de éxito de las incidencias obtenidas es del 60% para incidencias de tipo BBOX, de total de 42 incidencias reportadas por ambos bots. En el caso de incidencias de tipo PK o puntual el porcentaje de éxito es del 55%.

Mediante las incidencias reportadas por los bots se han podido realizar actualizaciones selectivas de la red de viaria en las cuatro provincias gallegas y en las respectivas catalanas, aplicándose así una metodología de mantenimiento de la red más eficiente y rápida gracias a la automatización en la detección de cambios.

Palabras claves

Bots, web scraping, red viaria, actualización, Cartobot.

Autores

Daniel Jesús Rolanía Soto

djrolania@mitma.es

Instituto Geográfico Nacional (IGN)

Cristina Calvo Guinea

mccalvo@mitma.es

Instituto Geográfico Nacional (IGN)

Alicia González Jiménez

agjimenez@mitma.es

Instituto Geográfico Nacional (IGN)

Pablo de la Presa Rodríguez

pdelapresa@mitma.es

Instituto Geográfico Nacional (IGN)