Guía 3

¿Qué es una clave API?

Una API es un sistema que permite que dos programas se comuniquen entre sí **(Application Programming Interface**). Cuando haces una solicitud a una API, usualmente recibes una **respuesta en formato JSON** con información estructurada.

Para hacer webscraping con Google News necesitas una clave API. Para esto puedes encontrar servicios como HasData. Esta es una API de extracción de datos web basada en la nube que simplifica las tareas de extracción de estos.

La API de Web Scraping te permite extraer contenido de páginas web sin la molestia de gestionar proxies, navegadores sin interfaz gráfica (headless) y captchas. Simplemente envía la URL y recibe como respuesta el HTML.

1. Webscraping con Scrapingdog’s de Google News API

Puede obtener una Scrapingdog’s de Google News API. Para esto puedes registrarte en Scrapingdog con tú correo Gmail personal y comenzar con las 1000 créditos gratis para probar y hacer tú propio webscrapping (*Scrape Google News Using Python (Export Data in CSV)*, 2024; Scrapingdog, 2025).

Para esto, puedes hacer click en este link: <https://api.scrapingdog.com/register>

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

Luego de realizar tú registro en encontrará una página web como la que se ve en la siguiente imagen. En la esquina superior izquierda encontrarás tú clave API, esta clave que utilizarás en el código Phyton para hacer webcraping.

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

1. Webscraping con HasData

Otra forma de obtener una clave API es con Hasdata. HasData es una API de extracción de datos web basada en la nube que simplifica las tareas de extracción de datos web para empresas de todos los tamaños. Para esto puedes dirigirte a este link: <https://app.hasdata.com/dashboard>

Aquí encontrarás la siguiente información con tú clave API. Esta clave te entrega 200 créditos gratis cada 30 días.

Interfaz de usuario gráfica, Tabla

El contenido generado por IA puede ser incorrecto.

También en el menú settings puedes encontrar tú clave API.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.

1. Elementos del código con tú clave API integrada

La estructura de tú código Phyton con tú clave integrada tendrá estos elementos:

|  |
| --- |
| curl --request POST \  --url https://api.hasdata.com/scrape/web \ ← **URL de la API**  --header 'Content-Type: application/json' \ ← **Formato de los datos enviados**  --header 'x-api-key: <your-api-key>' \ ← **Tu clave de acceso personal a la API**  --data '{  "url": "https://example.com", ← **Página web que quieres scrapear**  "outputFormat": ["html", "text", "markdown"], ← **Formatos de salida que deseas**  "screenshot": true ← **Solicitas también una captura de pantalla**  }'(*Quickstart*, s. f.) |

Posibles respuestas al ejecutar el código:

Estatus del código de la clave API

|  |  |
| --- | --- |
| Código | Detalles |
| 200 | Respuesta satisfactoria |
| 404 | Página requerida no existe |
| 500 | Error en el servidor |
| 429 | Estás enviando requerimientos muy rápido y estás excediendo el límite |
| 403 | Usaste todos tus créditos API |
| 401 | Clave API invalida |

**Referencias**

*Quickstart*. (s. f.). HasData Documentation. Recuperado 29 de julio de 2025, de https://docs.hasdata.com/quickstart

*Scrape Google News Using Python (Export Data in CSV)*. (2024, octubre 16). https://www.scrapingdog.com/blog/scrape-google-news/

Scrapingdog (Director). (2025, marzo 11). *Scraping Google News using Scrapingdog’s News API* [Video recording]. https://www.youtube.com/watch?v=dBGyX1S-kHM