

Ejecución en red

Requests

Python 2020

¿Qué son?

Biblioteca HTTP con licencia Apache2, escrita en Python

Está diseñado para ser utilizado por humanos para interactuar con el lenguaje.

Esto significa que no tiene que agregar manualmente cadenas de consulta a las URL ni codificar de forma sus datos POST.

Permiten enviar solicitudes HTTP / 1.1 utilizando Python.

GET y POST

Se puede agregar contenido como encabezados, datos de formulario, archivos multiparte y parámetros a través de bibliotecas Python simples.

También le permite acceder a los datos de respuesta de Python de la misma manera.

La solicitud HTTP devuelve un Objeto de respuesta con todos los datos de respuesta (contenido, codificación, estado, etc.).

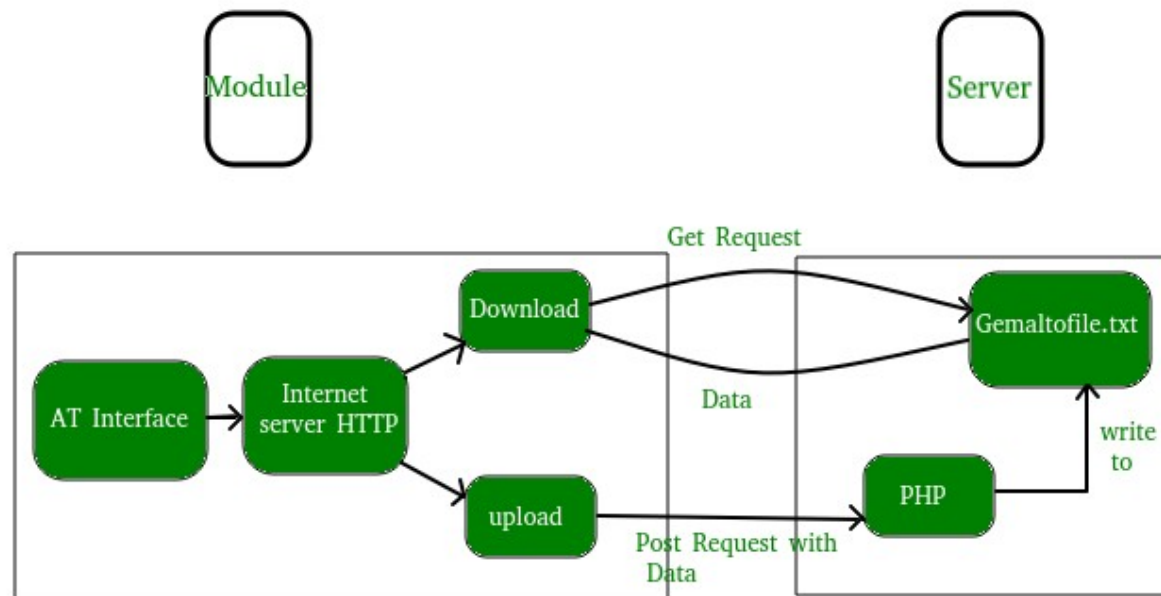


GET y POST

Entonces, para solicitar una respuesta del servidor, existen principalmente dos métodos:

GET: para solicitar datos del servidor.

POST: para enviar datos para ser procesados en el servidor.



Beautiful soup

Beautiful Soup es una biblioteca de Python para analizar documentos HTML

Esta biblioteca crea un árbol con todos los elementos del documento y puede ser utilizado para extraer información

Útil para realizar web scraping — extraer información de sitios web



Objetos en BeautifulSoup

Hay cuatro objetos que debemos conocer: Tag, NavigableString, BeautifulSoup, y Comment

.

Tag: se refiere a una etiqueta XML o HTML real en el documento

NavigableString: Donde se va a almacenar el texto de una etiqueta

Comment: Comentarios de una página web.



Beautiful soup

Beautiful Soup es una biblioteca de Python para analizar documentos HTML

Esta biblioteca crea un árbol con todos los elementos del documento y puede ser utilizado para extraer información

Útil para realizar web scraping — extraer información de sitios web

