SISTEMAS WEB CURSO 2023/2024

HTTP - HyperText Transfer Protocol

Solicitud y respuesta. Ejemplo utilizando Python: librería *requests*Códigos de Error: 400 y 404.

Redirecciones (códigod 3XX)

Envió de datos en formato formulario



Web Sistemak by <u>Oskar Casquero</u> & <u>María Luz Álvarez</u> is licensed under a <u>Creative Commons Reconocimiento 4.0 Internacional License</u>.

EJEMPLO UTILIZANDO PYTHON: LIBRERÍA REQUESTS

Utilizando la librería *requests* de Python solicita el siguiente recurso:
 http://www.httpwatch.com/httpgallery/chunked/chunkedimage.aspx

```
import requests
# La petición tiene 4 partes: metodo, uri, cabecera y cuerpo
metodo = 'GET'
uri = "http://www.httpwatch.com/httpgallery/chunked/chunkedimage.aspx"
cabeceras= {'Host': 'www.httpwatch.com'}
cuerpo= ''
respuesta = requests.request(metodo, uri, headers = cabeceras, data = cuerpo)
# La respuesta tiene también 4 apartados: codigo, descripción, cabeceras eta
cuerpo
codigo= respuesta.status code
descripcion= respuesta.reason
print(str(codigo) + " " + descripcion)
for cabecera in respuesta.headers:
    print(cabecera + ": " + respuesta.headers[cabecera])
cuerpo= respuesta.content
print(cuerpo)
fichero= open("imagen.jpg", 'wb')
                                                       Documentación de requests:
fichero.write(cuerpo)
fichero.close()
                                                    https://pypi.org/proect/requests/
```

FUNCIONAMIENTO DE HTTP: ERRORES

- ¿Qué ocurre si una petición no puede ser satisfecha?
 Códigos de respuesta 4xx y 5xx.
 - Cuando el cliente o el servidor no puede completar una solicitud, se devuelven los siguientes códigos de estado (STATUS) en la respuesta.

Errores provocados por el **cliente**:

4xx (RFC 2616, Sección 10.4)

- 400 Bad Request
- 401 Unauthorized
- 403 Forbidden
- 404 Not Found

• ...

Errores provocados por el **servidor**:

5xx (<u>RFC 2616, Sección 10.5</u>):

- 500 Internal Server Error
- 503 Service Unavailable
- ...

FUNCIONAMIENTO DE HTTP ERROR 400

400 Bad Request

- El servidor web no entiende la petición, su sintaxis/semántica es incorrecta.
- Supongamos que se solicita el recurso /html/main page.html

Petición de ejemplo

GET /html/main page.html HTTP/1.1

Host: sw2024.com Accept: text/html

Accept-Encoding: gzip,identity;q=0.5 Accept-Language: en-US,es-ES;q=0.8 User-Agent: Mozilla Windows Escritorio

Lo que interpreta el servidor al procesar la petición:

- Método: GET
- RequestURI: /html/main
- Versión del protocolo: page.html -> SEMÁNTICA ERRONEA (La cadena debe ser HTTP/1.1)
- Cadena adicional: HTTP/1.1 -> SINTAXIS ERROEA (la primera línea tiene que tener 3 elementos)

Solución: Codificar el espacio en la RequestURI: /html/main%20page.html

Respuesta de ejemplo

HTTP/1.1 400 Bad Request

Date: Wed, 25 Nov 2015 08:07:43 GMT

Content-Length: 138

Content-Type: text/html; charset=UTF-8

<html><head><title>Error 400 (Bad Request)</title></head><body>Your client has issued a malformed or illegal request.</body></html>

EJEMPLO: ERROR 400 (SEMÁNTICA ERRÓNEA)

 En el código del ejemplo anterior, añadir al final del nombre de la cabecera Host un espacio.

```
import requests
# La petición tiene 4 partes metodo, uri, cabecera y cuerpo
metodo = 'GET'
uri = "http://www.httpwatch.com/httpgallery/chunked/chunkedimage.aspx"
cabeceras= {'Host'} 'www.httpwatch.com'}
cuerpo= ''
respuesta = requests.request(metodo, uri, headers = cabeceras, data = cuerpo)
# La respuesta tiene también 4 apartados: codigo, descripción, cabeceras eta
cuerpo
codigo= respuesta.status code
descripcion= respuesta.reason
print(str(codigo) + " " + descripcion)
for cabecera in respuesta.headers:
    print(cabecera + ": " + respuesta.headers[cabecera])
cuerpo= respuesta.content
print(cuerpo)
fichero= open("imagen.jpg", 'wb')
fichero.write(cuerpo)
fichero.close()
```

FUNCIONAMIENTO DE HTTP ERROR 404

404 Not Found

- La petición esta bien formada, pero el servidor web no encuentra el recurso.
- Supongamos que se solicita al servidor sw2024.com el recurso /login/index.php, que no existe en el servidor

Petición de ejemplo

GET /login/index.php HTTP/1.1

Host: sw2024.com Accept: text/html

Accept-Encoding: gzip,identity;q=0.5 Accept-Language: en-US,es-ES;q=0.8 User-Agent: Mozilla Windows Escritorio

Respuesta de ejemplo

HTTP/1.1 404 Not Found

Date: Wed, 25 Nov 2015 08:07:43 GMT

Content-Length: 134

Content-Type: text/html; charset=UTF-8

<html><head><title>Error 404 (Not Found)</title></head><body>The requested resource was not found in this server.</body></html>

EJEMPLO: ERROR 404

• En el código del ejemplo anterior, cambiar el nombre del recurso:

/httpgallery/chunked/chunkedtext.aspx

```
import requests
# La petición tiene 4 partes: metodo, uri, cabecera y cuerpo
metodo = 'GET'
uri = "http://www.httpwatch.com/httpgallery/chunked/chunkedtext.aspx"
cabeceras= {'Host': 'www.httpwatch.com'}
cuerpo= ''
respuesta = requests.request(metodo, uri, headers = cabeceras, data = cuerpo)
# La respuesta tiene también 4 apartados: codigo, descripción, cabeceras eta
cuerpo
codigo= respuesta.status code
descripcion= respuesta.reason
print(str(codigo) + " " + descripcion)
for cabecera in respuesta.headers:
    print(cabecera + ": " + respuesta.headers[cabecera])
cuerpo= respuesta.content
print(cuerpo)
fichero= open("imagen.jpg", 'wb')
fichero.write(cuerpo)
fichero.close()
```

FUNCIONAMIENTO DE HTTP: REDIRECCIONES

- ¿Qué son las redirecciones?
 - Algunas veces, el URI de un recurso puede cambiarse o adecuarse. Por ejemplo, en el año académico 2017-18, el URI http://egela.ehu.eus/ fue dirigido al URI http://egela2017-18.ehu.eus/; en cambio, durante el año académico 2018-19, se dirigió al URI http://egela2018-19.ehu.eus/.
 - En ambos casos, el usuario no necesitaba conocer el URI real, porque al solicitar http://egela.ehu.eus/ el navegador redirige automáticamente al usuario al URI.
- ¿Cómo se hacen las redirecciones?
 - El protocolo HTTP ofrece una forma de redirigir a otro URI. Para realizar las redirecciones utiliza:
 - Los códigos de respuesta 301, 302 o 303
 - y el encabezado "Location".

FUNCIONAMIENTO DE HTTP: REDIRECCIONES

- Supongamos que un cliente ubicado en Madrid solicita un recurso cuya URI es: http://www.google.com/
- El servidor web que atiende la petición HTTP detecta que dirección IP de origen del paquete es de España, por lo que devuelve una respuesta al cliente indicándole que realice una redirección a http://www.google.es/
- El cliente, al detectar un código de respuesta 302, extrae el contenido de la cabecera "Location" y realiza una nueva petición a esta URI. En un navegador, este proceso ocurre de forma transparente para el usuario.

Ejemplo de Solicitud

GET / HTTP/1.1

Host: www.google.com

Accept: text/html

Accept-Encoding: identity

Accept-Language: en-US,es-ES;q=0.8 User-Agent: Mozilla Windows Escritorio

Ejemplo de Respuesta

HTTP/1.1 302 Found

Content-Length: 137

Content-Type: text/html; charset=UTF-8

Location: http://www.google.es/

<html><head><title>Redirection

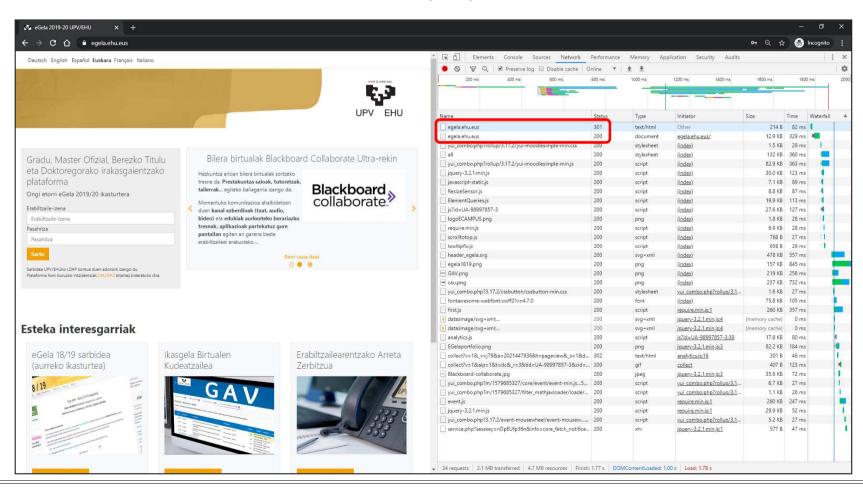
302</title></head><body>Redirect to http://www.google.es/</body></htm>

FUNCIONAMIENTO DE HTTP: REDIRECCIONES

- https://tools.ietf.org/html/rfc7231#section-6.4
 - 301 Moved Permanently: Este código de respuesta significa que el URI del recurso solicitado ha sido cambiado. Probablemente la nueva URI sea devuelta en la respuesta.
 - 302 Found: Este código de respuesta significa que el recurso de la URI solicitada ha sido cambiado temporalmente. En el futuro se realizarán nuevos cambios en la URI. Por lo tanto, la misma URI debe ser usada por el cliente en futuras solicitudes.
 - 303 See Other: El servidor envía esta respuesta para dirigir al cliente a un recurso diferente, indicado en la cabecera. El cliente deberá solicitar el nuevo recurso.

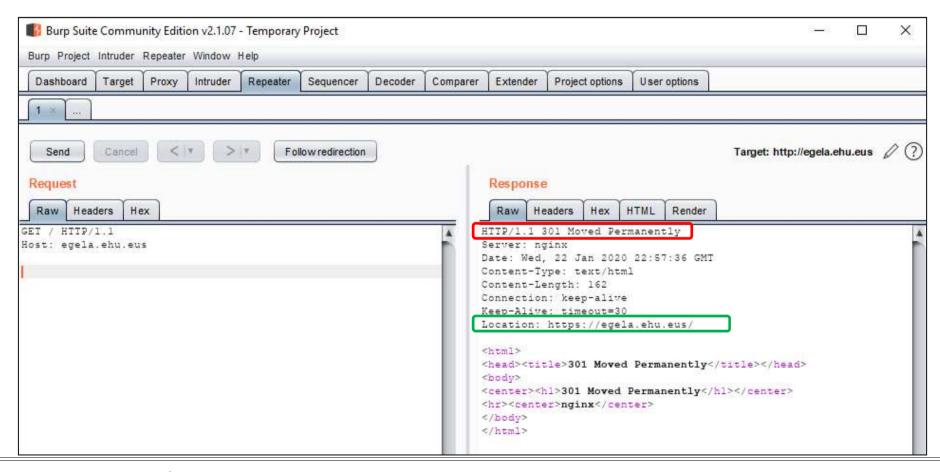
EJEMPLO USANDO LAS HERRAMIENTAS DE DESARROLLO DEL NAVEGADOR

Abre una nueva pestaña en su navegador, solicita el recurso http://egela.ehu.eus/
 usando las herramientas de desarrollador (F12).



EJEMPLO UTILIZANDO BURP

 Utilizando la pestaña Repeater de Burp, solicita el siguiente recurso: http://egela.ehu.eus/



EJEMPLO UTILIZANDO PYTHON

- Utilizando la librería requests en Python-en, realiza el siguiente ejemplo

```
import requests
met.odo = 'GET'
uri = "http://egela.ehu.eus/"
cabeceras = {'Host': 'egela.ehu.eus'}
cuerpo = ''
respuesta = requests.request(metodo, uri, headers=cabeceras, data=cuerpo, allow redirects=False)
codigo = respuesta.status code
descripcion = respuesta.reason
print(str(codigo) + " " + descripcion)
for cabecera in respuesta.headers:
    print(cabecera + ": " + respuesta.headers[cabecera])
cuerpo = respuesta.content
print(cuerpo)
met.odo = 'GET'
                                           Tener en cuenta que en la nueva URI el host puede
uri = respuesta.headers['Location']
cabeceras = {'Host': uri.split('/')[2]}
                                           cambiar!!
cuerpo = ''
respuesta = requests.request(metodo, uri, headers=cabeceras, data=cuerpo, allow_redirects=False)
codigo = respuesta.status_code
descripcion = respuesta.reason
print(str(codigo) + " " + descripcion)
for cabecera in respuesta.headers:
    print(cabecera + ": " + respuesta.headers[cabecera])
cuerpo = respuesta.content
print(cuerpo)
```

FUNCIONAMIENTO DE HTTP: FORMULARIOS

- ¿Cómo se envían información desde el cliente al servidor?
 - Un cliente HTTP puede enviar diferentes tipos de datos a un servidor web:
 - datos en formato binario (ejemplo: imágenes)
 - datos en formato JSON o XML
 - datos en formato formulario
- En esta sesión vamos a estudiar el envío de datos en formato formulario, es decir, cadenas cortas de datos que se envían como pares de nombre-valor, por ejemplo:
 - Modelo genérico: name1=value1&name2=value2&name3=value3
 - Ejemplo del modelo: dni=12345678N&nombre=marta&apellido=perez

FUNCIONAMIENTO DE HTTP: ENVÍO DE DATOS EN FORMATO FORMULARIO (POST)

- Supongamos que en un servidor web reside un recurso que calcula la letra del DNI. Si se envía una solicitud HTTP con el número DNI en un parámetro denominado dni, este recurso devuelve una respuesta HTTP que contiene la letra del DNI
- Para enviar el dato en formato formulario, el cliente define la siguiente petición HTTP:

Método: POST

- Cabeceras para caracterizar el cuerpo del mensaje:
 - Content-Type: application/x-www-form-urlencoded
 - Content-Length: la longitud del cuerpo del mensaje
- Los pares nombre-valor deben codificarse de una manera particular (codificación UTF-8 + por ciento) y formatearse (como un nombre = valor y un separador entre pares "&").

Solicitud de ejemplo

POST /recurso HTTP/1.1

Host: sw2020.com Accept: text/plain

Content-Type: application/x-www-form-urlencoded

Content-Length: 12

User-Agent: Cliente Python

dni=12345678

Respuesta de ejemplo

HTTP/1.1 200 OK

Date: Wed, 25 Nov 2015 08:07:43 GMT

Content-Length: 1

Content-Type: text/plain

Z

FUNCIONAMIENTO DE HTTP: ENVÍO DE DATOS EN FORMATO FORMULARIO (GET)

- En lugar de enviar los pares de valores en el cuerpo del mensaje, se puede enviar al URI utilizando el carácter "?"-> query string
 - Método: GET
 - No se deben usar los encabezados Content-Type y Content-Length
 - En *query string* solo puede enviar datos formateados.
 - La estructura de la solicitud HTTP limita la *query string*: los pares nombre- valor se encuentra entre el carácter "?" y la cadena "HTTP / 1.1".
 - La codificación y el formato de los datos se realizan como antes.
 - Desventaja: la cantidad de datos que se pueden enviar está limitada a 1024 octetos.

Petición de ejemplo

GET /recurso?dni=12345678 HTTP/1.1

Host: sw2016.com Accept: text/plain

User-Agent: Cliente Python

Respuesta de ejemplo

HTTP/1.1 200 OK

Date: Wed, 25 Nov 2015 08:07:43 GMT

Content-Length: 1

Content-Type: text/plain

Ζ

FUNCIONAMIENTO DE HTTP: RESUMEN

GET /app/servlet/contactUs?name=XXX&email=YYY&message=ZZZ HTTP/1.1

Host: ws2017.com Accept: text/html

Accept-Language: en-US,en Connection: keep-alive

POST /app/servlet/contactUs HTTP/1.1

Host: ws2017.com Accept: text/html

Accept-Language: en-US,en Connection: keep-alive

Content-Type: application/x-www-form-urlencoded

Content-Length: 30

name=XXX&email=YYY&message=ZZZ

FUNCIONAMIENTO DE HTTP: ENVÍO DE DATOS EN FORMATO FORMULARIO

- En el formato de formulario, se especifica para el contenido (las cadenas de nombre y valor de los pares nombre-valor) el siguiente procedimiento de codificación:
 - Si aparece un carácter especial (caracteres que son **delimitadores** en el URI o en la cadena de nombre-valor) o un carácter no recogido en el alfabeto <u>US-ASCII</u> (por ejemplo, un carácter con tilde), éste debe codificarse en dos pasos antes formar la URI:
 - 1. Codificar el carácter en <u>UTF-8</u>. (valores en hexadecimal)
 - Ejemplo 1: $\& \rightarrow$ 26
 - Ejemplo2 2: á → C3 A1
 - 2. Codificar el resultado anterior en "porciento" (percent encoding)
 - Ejemplo $1:26 \rightarrow \%26$
 - Ejemplo 1 : C3 A1 → %C3 %A1

NOTA: En los ejemplos anteriores, solo se codifica un carácter; pero debido al procedimiento de codificación, ¿cuántos caracteres se enviarán?

- Ejemplo 1: $\%26 \rightarrow 3$ caracteres: %, 2, 6
- Ejemplo 2: %C3 %A1 \rightarrow 6 caracteres: %, C, 3, %, A, 1

FUNCIONAMIENTO DE HTTP: ENVÍO DE DATOS EN FORMATO FORMULARIO

- Ejemplo:
 - datos a enviar: nombre=Iñaki y apellido=Pérez
 - datos que se codificarán en el cuerpo del mensaje:

nombre=I%C3%B1aki&apellido=P%C3%A9rez

- Si en la cadena del nombre o del valor de los parea nombre-valor aparece un espacio, éste se codifica con "+":
 - datos a enviar: nombre_apellidos=Iñaki Pérez
 - datos que se codificarán en el cuerpo del mensaje:

nombre apellidos=I%C3%B1aki+P%C3%A9rez

- Pero si se desea utilizar el "+" como dato en lugar de como separador, éste debe codificarse según lo indicado arriba (primero en UTF-8 y luego en porciento):
 - datos a enviar: num1=2, num=2 y operator=+
 - datos que se codificarán en el cuerpo del mensaje:

num1=2&num2=2&operato=%2B

EJEMPLO:

ENVÍO DE DATOS EN FORMATO FORMULARIO -- BURP

Envío de datos a un servidor para obtener la letra del DNI:

http://gae-sw-2017.appspot.com/

FUNCIONAMIENTO DE HTTP: FORMULARIOS (HTML)



```
<html>
  <head>
    <title>Formulario</title>
  </head>
  <body>
       <form method="GET" action="/app/servlet/contactUs">
          <b>Name: </b>
          <input type="text" name="name">
          <br/>
          <b>Email: </b>
          <input type="text" name="email">
          <br/>
          <b>Message: </b>
          <input type="text" name="message">
          <br/><br/>
          <input type="submit">
       </form>
    </body>
</html>
```

EJEMPLO:

ENVÍO DE DATOS EN FORMATO FORMULARIO -- POST

Envío de datos a un servidor para obtener la letra del DNI:

http://gae-sw-2017.appspot.com/

Codificación de datos

Petición de Letra DNI POST

POST /processForm HTTP/1.1

Host: gae-sw-2017.appspot.com

Content-Type: application/x-www-form-urlencoded

Content-Length:12

dni=12345678

Respuesta del servidor

HTTP/1.1 200 OK

Date: Wed, 25 Nov 2015 08:07:43 GMT

Content-Length: 9

Server: Google Frontend

Content-type: text/html; charset=utf-8

12345678Z

Generamos la petición

```
codigo = respuesta.status_code
descripcion = respuesta.reason
print(str(codigo) + " " + descripcion)
for cabecera in respuesta.headers:
    print(cabecera + ": " + respuesta.headers[cabecera])
cuerpo = respuesta.content
print(cuerpo)
```

Procesamos la respuesta

EJEMPLO:

ENVÍO DE DATOS EN FORMATO FORMULARIO -- GET

Envío de datos a un servidor para obtener la letra del DNI:

http://gae-sw-2017.appspot.com/

Petición de Letra DNI GET

GET /processForm?dni=12345678 HTTP/1.1 Host:gae-sw-2017.appspot.com

Respuesta del servidor

HTTP/1.1 200 OK

Date: Wed, 25 Nov 2015 08:07:43 GMT

Content-Length: 9

Server: Google Frontend

Content-type: text/html; charset=utf-8

12345678Z

```
codigo = respuesta.status_code
descripcion = respuesta.reason
print(str(codigo) + " " + descripcion)
for cabecera in respuesta.headers:
    print(cabecera + ": " + respuesta.headers[cabecera])
cuerpo = respuesta.content
print(cuerpo)

Procesamos la respuesta
```

Datos en la URI

EJERCICIO: CONSULTA EN EL *DIRECTORIO DE EHU*UTILIZANDO BURP Y EN PYTHON

https://www.ehu.eus/bilatu/buscar/bilatu.php?lang=es

