

# HTTP

---

## HTTP

---

HTTP steht für "*Hypertext Transfer Protocol*" und ist das grundlegende Protokoll zur Datenübertragung im Internet zwischen Clients (Nutzern) und Servern.

Zwischen dem Client und dem Server werden Messages ausgetauscht. Zunächst schickt der Client eine Anfrage an den Server, etwa wenn eine HTML-Datei angefragt wird. Diese Anfrage wird vom Webserver bearbeitet und anschließend erfolgt die Antwort-Nachricht vom Webserver an den Client (HTTP-Response).

Jede Message besteht aus dem Nachrichtenheader (Meta-Informationen) sowie dem Body (Inhalt der Nachricht).

## HTTP-Anfragemethoden

---

Um zu beschreiben, welche Art von Operation durch einen Request auf dem Server ausgeführt werden soll, besitzt jeder Request eine bestimmte Anfragemethode. Die zwei wichtigsten Methoden sind:

- **GET**: Die gebräuchlichste Methode. Sie wird genutzt, um Ressourcen anzufragen. Es können zusätzlich Parameter angehängen werden, um die Anfrage näher zu beschreiben, z.B. `GET /users/MIEDERJ`. Die GET-Methode wird auch genutzt, wenn im Browser eine URL aufgerufen wird, da hierdurch lediglich Dokumente (HTML-Dateien, Bilder, ...) angefordert werden.
- **POST**: Wird genutzt, um Daten an den Server zu übertragen, etwa wenn ein Formular ausgefüllt wurde.

Weitere Methoden sind **HEAD** (Server antwortet lediglich mit dem Response-Header), **PUT** (Upload einer Datei), **PATCH** (Änderung eines Dokuments), **DELETE** (Löschen einer Ressource)

## Webserver

---

Ein Webserver wird benötigt, um Inhalte (HTML-Dateien, Bilder, Videos, ...) zu speichern und an Clients (Nutzer) auszuliefern.

Inhalte können statisch (z.B. HTML-Dateien, Bilder, Videos, ...) oder dynamisch erzeugt sein. Für jede benötigte Datei schickt der Client (Browser) eine eigene Anfrage an den Webserver (HTTP-Request), der daraufhin die entsprechende Datei an den Client ausliefert und zusätzlich mit einem Statuscode antwortet (Response).

Daher muss der Webserver multithreadingfähig sein, um Anfragen parallel abarbeiten zu können.

Der erste Webserver mit dem Namen "*CERN httpd*" wurde von 1990 - 1996 von Tim Berners-Lee am CERN entwickelt.

Die bekanntesten Webserver-Programme sind *Apache HTTP Server*, *nginx* und *Microsoft Internet Information Services*.

## Statuscodes

---

Um dem Client mitzuteilen, ob eine Anfrage erfolgreich war oder fehlgeschlagen ist, wird vom HTTP-Webserver ein Statuscode zurückgegeben.

Diese sind jeweils dreistellig und anhand der ersten Ziffer gruppiert. Die wichtigsten Statuscodes sind 200 ("OK", Anfrage erfolgreich verarbeitet und Ergebnis zurückgegeben) und 404 ("Not Found", Ressource wurde nicht gefunden).

Die Gruppen sind 1xx (Informationen - Die Bearbeitung der Anfrage dauert noch an), 2xx (Erfolgreiche Operation), 3xx (Umleitung / Redirect, für eine erfolgreiche Bearbeitung der Anfrage sind weitere Schritte des Clients erforderlich), 4xx (Client Fehler, Fehler liegt im Verantwortungsbereich des Clients), 5xx (Server Fehler, Fehler liegt im Verantwortungsbereich des Servers).

## Aufgabe:

---

1. Erstelle einen simplen HTTP-Server mit dem `http.server`-Package. Der Server soll unter der URL `"http://localhost:5000/hello"` das HTML-Dokument `helloworld.html` zurückgeben.

*Tip: Erzeuge eine Klasse, die von `http.server.SimpleHTTPRequestHandler` erbt, überschreibe die entsprechende Methode zum Verarbeiten der GET-Methode.*