

Università degli Studi di Roma "Tor Vergata"
Laurea in Informatica

Sistemi Operativi e Reti
(modulo Reti)
a.a. 2024/2025

Approfondimento su HTTP: content negotiation

dr. Manuel Fiorelli

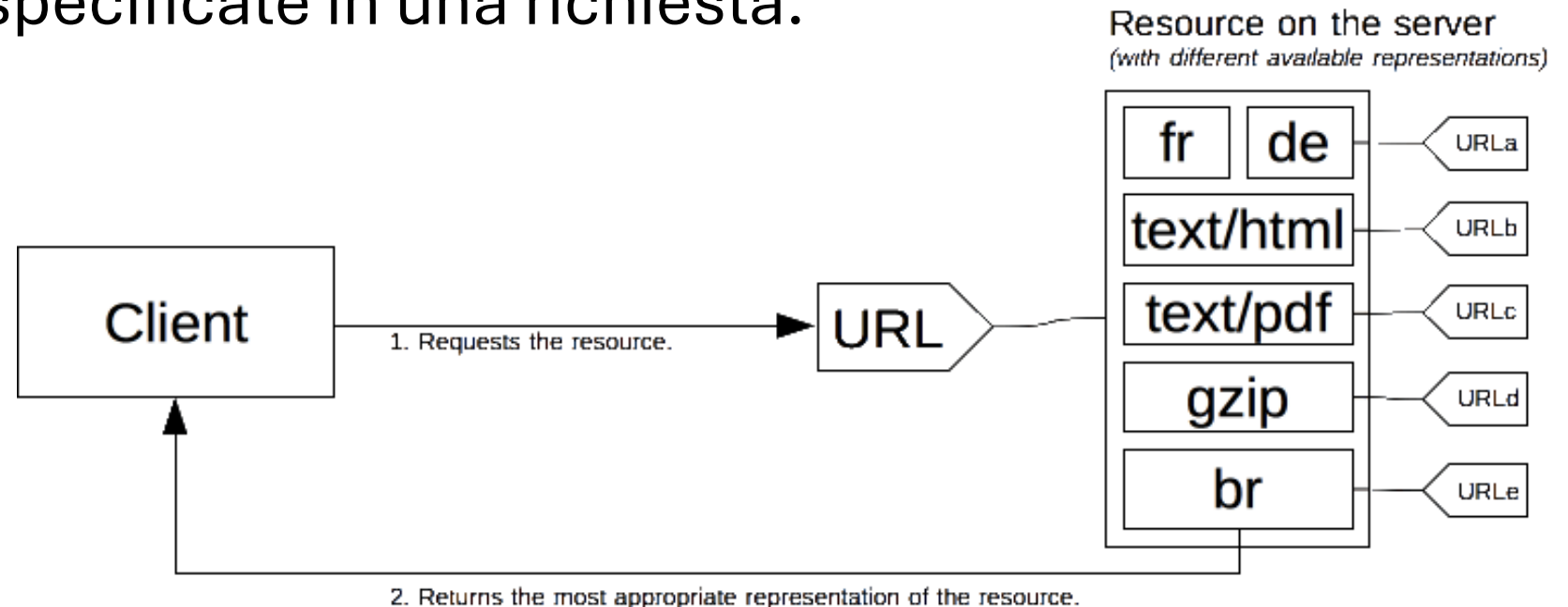
manuel.fiorelli@uniroma2.it

<https://art.uniroma2.it/fiorelli>

Content negotiation

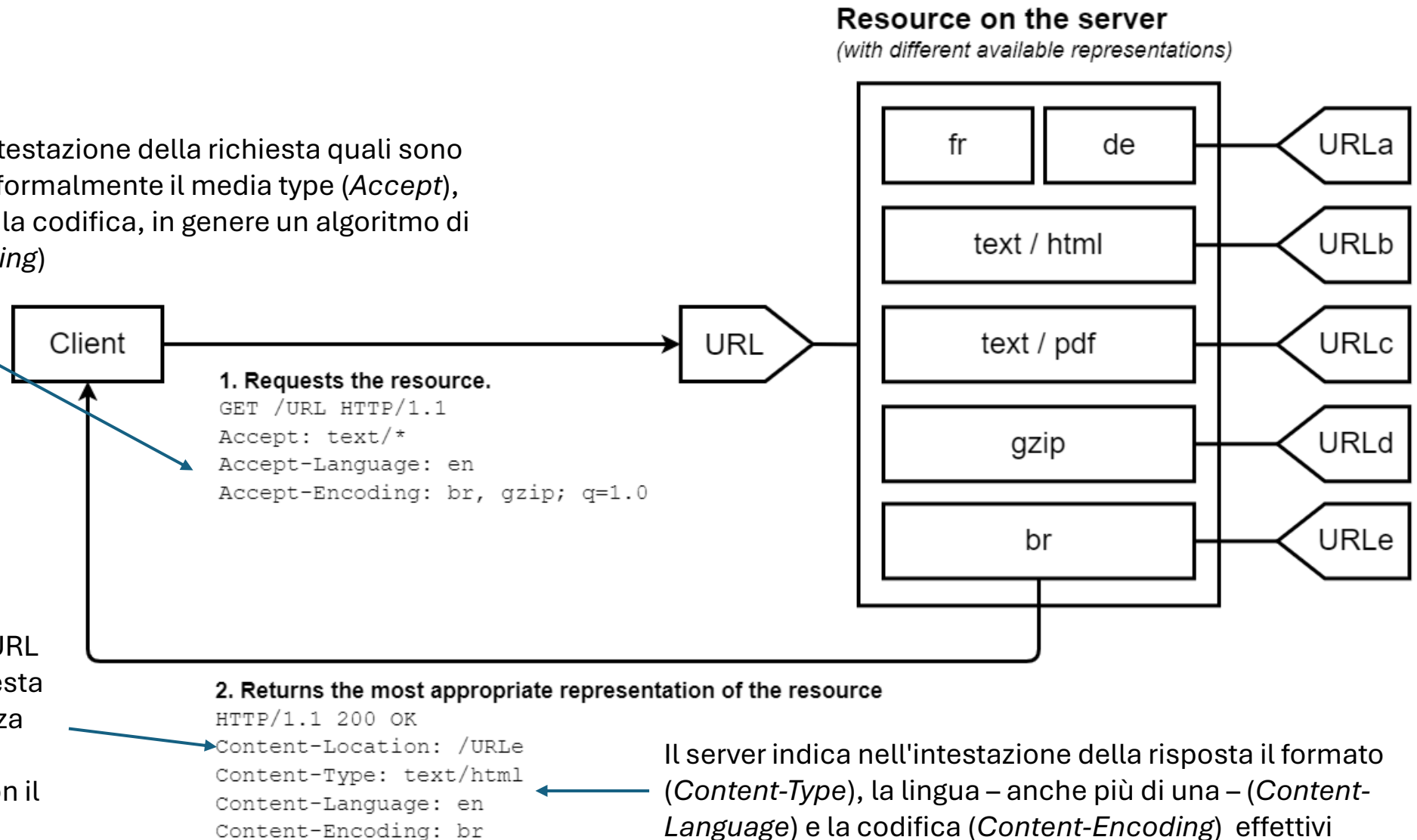
Si riferisce al fatto di servire diverse rappresentazioni di un oggetto (più precisamente, una risorsa) allo stesso URL.

Nel seguito faremo riferimento alla *server-driven content negotiation*: il server restituisce la rappresentazione più appropriata alle preferenze specificate in una richiesta.



Server-driven content negotiation

Lo *user agent* specifica nell'intestazione della richiesta quali sono le preferenze circa il formato, formalmente il media type (*Accept*), la lingua (*Accept-Language*) e la codifica, in genere un algoritmo di compressione (*Accept-Encoding*)



Content-Location fornisce una URL per recuperare direttamente questa specifica rappresentazione, senza ricorrere in futuro alla content negotiation: **NON** confondere con il campo di intestazione *Location*!

Il server indica nell'intestazione della risposta il formato (*Content-Type*), la lingua – anche più di una – (*Content-Language*) e la codifica (*Content-Encoding*) effettivi

Selezione della rappresentazione

Il server dovrebbe restituire la rappresentazione che meglio soddisfa le preferenze dell'utente indicate dal suo *user agent* nella richiesta.

I dettagli della selezione possono variare da implementazione a implementazione, ma di base c'è l'idea di confrontare il formato, la lingua e la codifica delle diverse rappresentazioni disponibili con le preferenze indicate nella richiesta.

Se il server non ha una rappresentazione adeguata può inviare una risposta con codice *406 Not Acceptable* oppure può restituire una rappresentazione predefinita.

Infatti, il client dovrebbe sempre verificare tramite l'intestazione della risposta le caratteristiche di ciò che ha ricevuto.

Accept

Indica uno o più media type* (separati da virgole) che il client comprende, ciascuno dei quali può essere espresso come:

- <MIME_type>/<MIME_subtype>
- <MIME_type>/*
- */*

Esempi di *Accept*:

- image/png
- image/*
- */*

media type di una rappresentazione

image/png

image/jpeg

application/pdf

corrispondenze

*noti in passato come Multipurpose Internet Mail Extensions o MIME type

Accept-Language

Indica una o più lingue preferite (separate da virgole) ciascuna espressa come:

- *language tag*, che identifica una lingua di base combinata a altri *subtag* separati da trattini. Il loro uso più comune è la rappresentazione di varianti geografiche

- *

Esempi di Accept-Language:

- en

- en-US

- *

Lingua di una rappresentazione

en

en-US

en-GB

it

corrispondenze: si noti che la lingua *en-US* soddisfa la richiesta per *en*, mentre la lingua *en* non soddisfa la richiesta *en-US*.

Accept-Encoding

Indica una o più codifiche (in genere algoritmi di compressione), tipo gzip, deflate e br, oppure * (come wildcard) che il client comprende.

Note su Accept, Accept-Language e Accept-Encoding

Questi campi di intestazione possono avere più valori separati da virgole. Ciascun valore può avere un *quality value* (*più grande è meglio*), espresso come numero decimale tra 0 e 1 (default, se omesso) con non più di 3 cifre decimali.

Accept: */*;q=0.1,text/*;q=0.7,application/pdf;q=0.5

Dato il media type *application/pdf*, il suo quality value è chiaramente 0.5. Nel caso di *text/html*, vediamo che sono possibili due match con quality value diversi, ma viene disambiguato considerando il match più specifico: in questo caso, il match con *text/** che produce un quality value di 0.7.

Vary

Campo di intestazione della risposta che indica quali campi di intestazione della richiesta (*Accept*, *Accept-**) hanno determinato la scelta di questa specifica rappresentazione.

Ciò serve a istruire una cache web a non restituire questa stessa rappresentazione per una richiesta per la stessa URL ma con diversi valori nei campi di intestazione indicati nell'header *Vary*.

Alcuni media type

- application/pdf (documento PDF)
- application/json (formato JSON)
- application/xml oppure text/xml (formato XML)
- text/html (formato HTML)
- text/css (foglio di stile CSS)
- text/javascript (file JavaScript)
- text/plain (testo semplice)
- image/png (immagine PNG)
- image/jpeg (immagine JPEG)
- image/gif (immagine GIF)

Per i media type testuali (*text/**) è possibile indicare un charset (con il suo nome, in modo case insensitive): nell'esempio seguente indichiamo un document HTML che usa il charset UTF-8

```
text/html;charset=utf-8
```

Alcuni language tag

- it (italiano)
- en (inglese)
- en-US (inglese americano)
- en-GB (inglese britannico)
- fr (francese)
- de (tedesco)
- es (spagnolo)
- pt (portoghese)

Comandi

- `curl -H "Accept: image/png" -o /dev/null --verbose http://httpbin.org/image`
- `curl -H "Accept: image/jpeg" -o /dev/null --verbose http://httpbin.org/image`
- `curl -H "Accept: image/jpeg, image/png" -o /dev/null --verbose http://httpbin.org/image`
- `curl -H "Accept: application/pdf" -o /dev/null --verbose http://httpbin.org/image`

Riferimenti

- https://developer.mozilla.org/en-US/docs/Web/HTTP/Content_negotiation
- <https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Accept>
- <https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Accept-Language>
- <https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Accept-Encoding>