

Table of Contents

[RECAP: Files](#)

[Ricerca delle parole più importanti in un gruppo di documenti](#)

[Ricerca del file più importante dato un gruppo di parole](#)

[Ci serve la frequenza di una parola in ciascun file \(tf = term frequency\)](#)

[e l'inverso della presenza della parola nei file \(idf = inverse document frequency\)](#)

[Altri tipi di file: JSON \(https://docs.python.org/3/library/json.html\)](https://docs.python.org/3/library/json.html)

[Attenzione:](#)

[Abbastanza facili da usare in python](#)

[File YAML \(https://pyyaml.org/wiki/PyYAMLDocumentation\)](https://pyyaml.org/wiki/PyYAMLDocumentation) (un altro modo di rappresentare dati annidati in testo semplice)
[leggere/scrivere file YAML](#)

[Leggere pagine o file da Internet](#)

[le pagine JSON vengono automaticamente decodificate](#)

[Scaricare file binari](#)

[Passare parametri ad una GET](#)

[Passare parametri ad altri metodi HTTP](#)



Fondamenti di Programmazione

Andrea Sterbini

lezione 11 - 10 novembre 2022



RECAP: Files

- **with open(<filename>, mode=<mode>, encoding='utf8') as F:**
 ...
 - **F.read()** legge tutto il contenuto
 - **F.readline()** legge UNA riga (compreso '\n')
 - **F.readlines()** legge TUTTE le righe (compresi '\n')
 - **F.write(<string>)** scrive un testo nel file (dovete aggiungere voi '\n')
 - **print(..., file=F)** stampa nel file (molto comodo)

- `F.seek(0)` torna alla posizione 0 (inizio)

Ricerca delle parole più importanti in un gruppo di documenti

- per ogni documento
 - carichiamo le parole
 - le contiamo
 - ne calcoliamo la frequenza percentuale

Ricerca del file più importante dato un gruppo di parole

Le parole più frequenti potrebbero essere troppo comuni e poco significative (con poco contenuto semantico)

Quindi una parola sarà "interessante" se **appare spesso in un file** ma **appare in pochi file**

Vedi <https://en.wikipedia.org/wiki/tf-idf> (<https://en.wikipedia.org/wiki/tf-idf>)

Ci serve la frequenza di una parola in ciascun file (tf = term frequency)

- per ciascun file
 - contiamo le parole
 - dividiamo per il numero totale di parole

```
In [1]: 1 # ho un elenco di files con il loro encoding, ottenuto unzipando files.zip
        2 files = {
        3     'files/holmes.txt'      : 'utf-8-sig',
        4     'files/alice.txt'      : 'utf-8-sig',
        5     'files/frankenstein.txt': 'utf-8-sig',
        6     'files/alice_it.txt'   : 'latin',
        7     'files/prince.txt'    : 'utf-8-sig'
        8 }
```

```
In [3]: 1 # per estrarre le parole da un file che contiene anche segni di interpunzione
2 def estrai_parole(filename, encoding):
3     # apro il file e ne leggo il contenuto
4     with open(filename, encoding=encoding) as FILE:
5         text = FILE.read()
6         # se voglio lo converto in lowercase (ma potrei preferire di no)
7         text = text.lower()
8         # calcolo l'insieme dei caratteri presenti nel testo
9         caratteri = set(text)
10        # ne estraggo solo i caratteri non-alpha
11        nonalfa = [ c for c in caratteri if not c.isalpha() ]
12        # li rimpiazzo con spazio
13        for c in nonalfa:
14            text = text.replace(c, ' ')
15        # uso split sul testo in cui ho eliminato tutti i non-apha
16        return text.split()
17
18 # esempio: parole del libro Alice nel paese delle meraviglie
19 parole = estrai_parole('files/alice_it.txt', 'latin')
20
21 ### come usare meno memoria? lavorando riga per riga
22 parole[:30]
```

```
Out[3]: ['charles',
'lutwidge',
'dodgson',
'alice',
'nel',
'paese',
'delle',
'meraviglie',
'questo',
'e',
'book',
'è',
'stato',
'realizzato',
'anche',
'grazie',
'al',
'sostegno',
'di',
'-']
```

```
In [10]: 1 # per contare quante sono le parole
2 def conta_parole(parole):
3     # trovo le parole uniche usando un insieme
4     parole_uniche = set(parole)
5     # e costruisco il dizionario parola:conteggio usando count   ### INEFFICIENTE!!!
6     N = len(parole)
7     return { parola: parole.count(parola)/N for parola in parole_uniche }
8
9 # rifaccio i conti sulle parole di Alice
10 conteggi = conta_parole(parole)
11
12 #print(list(conteggi.items())[:20])
13
14 # le 50 parole più presenti in Alice sono:
15 più_presenti = sorted(conteggi, reverse=True,
16                       key=lambda K: conteggi[K])[:50]
17 list(più_presenti)
```

```
Out[10]: ['e', 'il', 'la', 'di', 'che', 'non', 'un', 'alice', 'a', 'si', 'disse', 'in', 'ma', 'con',
'le', 'per', 'è', 'una', 'se', 'era', 'i', 'l', 'più', 'come', 'rispose', 'd', 'da', 'del', 'perchè', 'mi', 'gli', 'al', 'regina', 'della', 'così', 'lo', 'quando', 'poi', 're', 'cosa', 'o', 'aveva', 'io', 'sono', 'questo', 'tu', 'ne', 'tutti', 'qualche', 'cappellaio']
```

```
In [11]: 1 # frequenza delle parole nel file iesimo
2 # - tf_i = # presenze / # parole del file
3 # per tutti i file costruiamo il dizionario del numero di files in cui appaiono
4 frequenze = {} # conteggio delle parole nei file: frequenze[name][parola] -> % d
5
6 # conto le frequenze file per file
7 for filename, encoding in files.items():
8     print(filename)
9     parole = estrai_parole(filename, encoding)
10    conteggio = conta_parole(parole)
11    frequenze[filename] = conteggio
```

```
files/holmes.txt
files/alice.txt
files/frankenstein.txt
files/alice_it.txt
files/prince.txt
```

▼ e l'inverso della presenza della parola nei file (idf = inverse document frequency)

```
In [13]: 1 # conto i file che contengono una parola
2 # presenze[parola] -> # di file che la contengono
3 presenze = {}
4 for filename in files:
5     for parola in frequenze[filename]:
6         presenze[parola] = presenze.get(parola, 0) + 1
7
8 # divido per il numero di file per avere la presenza %
9 Nfile = len(files)
10 for parola in presenze:
11     presenze[parola] /= Nfile
12
13 list(presenze.items())[:30]
```

```
Out[13]: [('acute', 0.2), ('expense', 0.8), ('instant', 0.4), ('definitely', 0.2), ('displaying',
0.8), ('respects', 0.6), ('measure', 0.8), ('jersey', 0.2), ('disadvantages', 0.2), ('best',
0.8), ('put', 0.8), ('foil', 0.2), ('dazed', 0.2), ('wayward', 0.2), ('arrest', 0.6),
('ferret', 0.2), ('clergyman', 0.2), ('agonies', 0.4), ('carriage', 0.4), ('figured', 0.
2), ('feathers', 0.4), ('sherlock', 0.2), ('crest', 0.2), ('gaunt', 0.4), ('introduction',
0.6), ('stage', 0.4), ('imprudence', 0.4), ('gleaming', 0.2), ('fanciful', 0.4), ('crinkled', 0.2)]
```

```
In [15]: 1 from math import log
2
3 # log dell'inverso della frequenza nei documenti
4 # - idf = log( # di file / # di file che contengono la parola )
5 idf = { parola : log(1/quante)
6         for parola,quante in presenze.items() }
7
8 list(idf.items())[:30]
9 # NOTA: ogni parola appare in almeno UN file quindi 1/N è sempre calcolabile
```

```
Out[15]: [('acute', 1.6094379124341003), ('expense', 0.22314355131420976), ('instant', 0.916290731
8741551), ('definitely', 1.6094379124341003), ('displaying', 0.22314355131420976), ('resp
ects', 0.5108256237659907), ('measure', 0.22314355131420976), ('jersey', 1.60943791243410
03), ('disadvantages', 1.6094379124341003), ('best', 0.22314355131420976), ('put', 0.2231
4355131420976), ('foil', 1.6094379124341003), ('dazed', 1.6094379124341003), ('wayward',
1.6094379124341003), ('arrest', 0.5108256237659907), ('ferret', 1.6094379124341003), ('cl
ergyman', 1.6094379124341003), ('agonies', 0.9162907318741551), ('carriage', 0.9162907318
741551), ('figured', 1.6094379124341003), ('feathers', 0.9162907318741551), ('sherlock',
1.6094379124341003), ('crest', 1.6094379124341003), ('gaunt', 0.9162907318741551), ('intr
oduction', 0.5108256237659907), ('stage', 0.9162907318741551), ('imprudence', 0.916290731
8741551), ('gleaming', 1.6094379124341003), ('fanciful', 0.9162907318741551), ('crinkled',
1.6094379124341003)]
```

```
In [18]: ▾ 1 # e finalmente posso vedere quale file "pesa di più" rispetto alle parole della query
2
3 # date le parole di una query
4 # per ciascun file
5     # ne conto la frequenza nel file
6     # la moltiplico per idf
7     # sommando ottendo il "peso" del file per quella query
8
9 query = [ 'turtle', 'alice']
10
11 pesi = {}
12 ▾ for file in frequenze:
13     peso = 0
14     ▾ for p in query:
15         # aggiungo il peso della parola (0 se non presente)
16         peso += frequenze[file].get(p,0) * idf[p]
17     pesi[file] = peso
18     print(f"{file:30} {peso:00.3}")
19
20 # ordino i file per peso decrescente
21 sorted(pesi.items(), reverse=True, key=lambda coppia: coppia[1])
```

```
files/holmes.txt          5.62e-05
files/alice.txt           0.00989
files/frankenstein.txt    0.0
files/alice_it.txt        0.00889
files/prince.txt          0.0
```

```
Out[18]: [('files/alice.txt', 0.009889199619031072), ('files/alice_it.txt', 0.008886242041680276),
('files/holmes.txt', 5.6241811189737675e-05), ('files/frankenstein.txt', 0.0), ('files/prince.txt', 0.0)]
```



Altri tipi di file: [JSON \(https://docs.python.org/3/library/json.html\)](https://docs.python.org/3/library/json.html)

File di testo con sintassi semplificata per rappresentare:

- dizionari
- liste
- interi, stringhe, float, bool, None

Molto usati nelle applicazioni WEB

In [19]:

```
▼ 1 ## Esempio
  2 import json
  3
  ▼ 4 with open('api.github.com.json') as F:
    5     api = json.load(F)
    6     !cat api.github.com.json
```

```
{
  "current_user_url": "https://api.github.com/user",
  "current_user_authorizations_html_url": "https://github.com/settings/connections/applic
ations{/client_id}",
  "authorizations_url": "https://api.github.com/authorizations",
  "code_search_url": "https://api.github.com/search/code?q={query}{&page,per_page,sort,or
der}",
  "commit_search_url": "https://api.github.com/search/commits?q={query}{&page,per_page,so
rt,order}",
  "emails_url": "https://api.github.com/user/emails",
  "emojis_url": "https://api.github.com/emojis",
  "events_url": "https://api.github.com/events",
  "feeds_url": "https://api.github.com/feeds",
  "followers_url": "https://api.github.com/user/followers",
  "following_url": "https://api.github.com/user/following{/target}",
  "gists_url": "https://api.github.com/gists{/gist_id}",
  "hub_url": "https://api.github.com/hub",
  "issue_search_url": "https://api.github.com/search/issues?q={query}{&page,per_page,sor
t,order}",
  "issues_url": "https://api.github.com/issues{/number}"
}
```

Attenzione:

- NON si possono inserire commenti
- NO virgola aggiuntiva in fondo a lista o dizionario
- SOLO chiavi semplici (NO tuple, SI int, float, str, none, bool)
- SOLO doppi apici " (NO singoli apici)

Abbastanza facili da usare in python

Conversione automatica per alcuni tipi di dati

- `json.load(<file>)` -> dict | list | tipo_base
- `json.dump(<obj>, <file>)`

Personalizzabile anche per altri tipi di oggetti (non ovvio)

In [28]:

```
▼ 1 # posso anche leggere da una stringa in formato JSON
 2 XX = json.loads(''
▼ 3 [{"nome": "Minnie", "cognome": "Mouse", "telefono": "555-54321",
 4   "indirizzo": "via di M.me Curie 1", "città": "Topolinia"},
▼ 5 {"nome": "Pippo", "cognome": "de' Pippis", "telefono": "555-33333",
 6   "indirizzo": "via dei Pioppi 1", "città": "Topolinia"}]
 7 ''')
 8 print(XX)
 9 # oppure produrre la stringa in formato JSON da una struttura Python
10 print(json.dumps(XX))
```

```
[{'nome': 'Minnie', 'cognome': 'Mouse', 'telefono': '555-54321', 'indirizzo': 'via di M.m
e Curie 1', 'città': 'Topolinia'}, {'nome': 'Pippo', 'cognome': 'de' Pippis', 'telefono':
'555-33333', 'indirizzo': 'via dei Pioppi 1', 'città': 'Topolinia'}]
[{"nome": "Minnie", "cognome": "Mouse", "telefono": "555-54321", "indirizzo": "via di M.m
e Curie 1", "citt\u00e0": "Topolinia"}, {"nome": "Pippo", "cognome": "de' Pippis", "telef
ono": "555-33333", "indirizzo": "via dei Pioppi 1", "citt\u00e0": "Topolinia"}]
```

In [122]:

```
▼ 1 # o produrre una stringa
 2 json.dumps(XX)
```

Out[122]: '[{"nome": "Minnie", "cognome": "Mouse", "telefono": "555-54321", "indirizzo": "via di M.
me Curie 1", "citt\u00e0": "Topolinia"}, {"nome": "Pippo", "cognome": "de\' Pippis", "te
lefono": "555-33333", "indirizzo": "via dei Pioppi 1", "citt\u00e0": "Topolinia"}]'

In [30]:

```
▼ 1 # Un file json può contenere valori semplici (int, float, str, True=true, False=false)
 2 print(json.dumps(None))
 3
 4 print(json.dumps([False, True]))
 5
 6 ## NOTA: le stringhe DEVONO essere racchiuse da doppi apici
 7 print(json.dumps('Pape"rino'))
```

```
null
[false, true]
"Pape\"rino"
```

```
In [31]: 1 # Esempio: data una tabella come lista di dizionari
2 agenda = [
3     {'nome': 'Paperino', 'cognome': 'Paolino', 'telefono': '555-1313', 'indirizzo': 'via dei Peri 113', 'città': 'Paperopoli'},
4     {'nome': 'Gastone', 'cognome': 'Paperone', 'telefono': '555-1717', 'indirizzo': 'via dei Baobab 42', 'città': 'Paperopoli'},
5     {'nome': 'Paperon', 'cognome': 'de' Paperoni', 'telefono': '555-99999', 'indirizzo': 'colle Papero 1', 'città': 'Paperopoli'},
6     {'nome': 'Archimede', 'cognome': 'Pitagorico', 'telefono': '555-11235', 'indirizzo': 'colle degli Inventori 1', 'città': 'Paperopoli'},
7     {'nome': 'Pietro', 'cognome': 'Gambadilegno', 'telefono': '555-66666', 'indirizzo': 'via dei Ladri 13', 'città': 'Topolinia'},
8     {'nome': 'Trudy', 'cognome': 'Gambadilegno', 'telefono': '555-66666', 'indirizzo': 'via dei Ladri 13', 'città': 'Topolinia'},
9     {'nome': 'Topolino', 'cognome': 'Mouse', 'telefono': '555-12345', 'indirizzo': 'via degli Investigatori 1', 'città': 'Topolinia'},
10    {'nome': 'Minnie', 'cognome': 'Mouse', 'telefono': '555-54321', 'indirizzo': 'via di M.me Curie 1', 'città': 'Topolinia'},
11    {'nome': 'Pippo', 'cognome': 'de' Pippis', 'telefono': '555-33333', 'indirizzo': 'via dei Pioppi 1', 'città': 'Topolinia'}
12 ]
```

```
In [32]: 1 import json
2 # prima la salvo nel file
3 with open('agenda.json', encoding='utf8', mode='w') as F:
4     json.dump(agenda, F)
5
6 !cat agenda.json
```

```
[{"nome": "Paperino", "cognome": "Paolino", "telefono": "555-1313", "indirizzo": "via dei Peri 113", "citt\u00e0": "Paperopoli"}, {"nome": "Gastone", "cognome": "Paperone", "telefono": "555-1717", "indirizzo": "via dei Baobab 42", "citt\u00e0": "Paperopoli"}, {"nome": "Paperon", "cognome": "de' Paperoni", "telefono": "555-99999", "indirizzo": "colle Papero 1", "citt\u00e0": "Paperopoli"}, {"nome": "Archimede", "cognome": "Pitagorico", "telefono": "555-11235", "indirizzo": "colle degli Inventori 1", "citt\u00e0": "Paperopoli"}, {"nome": "Pietro", "cognome": "Gambadilegno", "telefono": "555-66666", "indirizzo": "via dei Ladri 13", "citt\u00e0": "Topolinia"}, {"nome": "Trudy", "cognome": "Gambadilegno", "telefono": "555-66666", "indirizzo": "via dei Ladri 13", "citt\u00e0": "Topolinia"}, {"nome": "Topolino", "cognome": "Mouse", "telefono": "555-12345", "indirizzo": "via degli Investigatori 1", "citt\u00e0": "Topolinia"}, {"nome": "Minnie", "cognome": "Mouse", "telefono": "555-54321", "indirizzo": "via di M.me Curie 1", "citt\u00e0": "Topolinia"}, {"nome": "Pippo", "cognome": "de' Pippis", "telefono": "555-33333", "indirizzo": "via dei Pioppi 1", "citt\u00e0": "Topolinia"}]
```

```
In [33]: 1 # poi la ricarico
2 with open('agenda.json', encoding='utf8') as F:
3     L1 = json.load(F)
4     L1[:2] # e ne mostro i primi 2 elementi
```

```
Out[33]: [{'nome': 'Paperino', 'cognome': 'Paolino', 'telefono': '555-1313', 'indirizzo': 'via dei Peri 113', 'città': 'Paperopoli'}, {'nome': 'Gastone', 'cognome': 'Paperone', 'telefono': '555-1717', 'indirizzo': 'via dei Baobab 42', 'città': 'Paperopoli'}]
```

File [YAML \(https://pyyaml.org/wiki/PyYAMLDocumentation\)](https://pyyaml.org/wiki/PyYAMLDocumentation) (un altro modo di rappresentare dati annidati in testo semplice)

- dizionari (una **chiave** : **valore** per ciascuna riga)
- liste (valori su righe diverse, preceduti da -)
- dati semplici (str senza apici se possibile, True, False, null=None, interi, float)
- documenti multipli
- generici oggetti Python (advanced)

Per annidare le strutture si usa l'**indentazione**

In [34]:

```
1 import yaml
2 with open('agenda.yaml', mode='w', encoding='utf8') as F:
3     yaml.dump(agenda[:2], F)
4 !cat agenda.yaml
```

```
- "città": Paperopoli
  cognome: Paolino
  indirizzo: via dei Peri 113
  nome: Paperino
  telefono: 555-1313
- "città": Paperopoli
  cognome: Paperone
  indirizzo: via dei Baobab 42
  nome: Gastone
  telefono: 555-1717
```

leggere/scrivere file YAML

- `yaml.dump(<oggetto>, FILE)` salva l'oggetto nel file aperto con open
- `yaml.safe_load(FILE)` -> `<oggetto>` legge l'oggetto dal file aperto con open

```
In [35]: 1 ## posso decodificare direttamente del testo
2 testo = """
3     none: [~, null]
4     bool: [true, false, on, off]
5     int: 42
6     float: 3.14159
7     list:
8         - LITE
9         - RES_ACID
10        - SUS_DEXT
11    dict:
12        hp: 13
13        sp: 5
14    """
15    yaml.safe_load(testo)
```

```
Out[35]: {'none': [None, None], 'bool': [True, False, True, False], 'int': 42, 'float': 3.14159, 'list': ['LITE', 'RES_ACID', 'SUS_DEXT'], 'dict': {'hp': 13, 'sp': 5}}
```

```
In [37]: 1 with open('esempio.yaml', mode='w', encoding='utf8') as F:
2         print(testo, file=F)
3
4     !cat esempio.yaml
5     # oppure leggerlo direttamente da file
6     with open('esempio.yaml') as F:
7         EX = yaml.safe_load(F)
8     EX
```

```
none: [~, null]
bool: [true, false, on, off]
int: 42
float: 3.14159
list:
  - LITE
  - RES_ACID
  - SUS_DEXT
dict:
  hp: 13
  sp: 5
```

```
Out[37]: {'none': [None, None], 'bool': [True, False, True, False], 'int': 42, 'float': 3.14159, 'list': ['LITE', 'RES_ACID', 'SUS_DEXT'], 'dict': {'hp': 13, 'sp': 5}}
```

Leggere pagine o file da Internet

Ci sono molte librerie, la più comune è [requests \(https://requests.readthedocs.io/en/latest/user/quickstart/\)](https://requests.readthedocs.io/en/latest/user/quickstart/)

- permette di eseguire sia **GET** che **POST**
- con parametri
- torna un codice di errore/OK
- e decodifica il testo della pagina html o json

Molto potente, permette anche streaming, sessioni, ...

In [39]:

```
1 # importo la libreria
2 import requests
3
4 # leggo la pagina di python.org
5 pagina = requests.get('https://python.org')
6
7 # lo status code ci dice se tutto è andato bene
8 status = pagina.status_code
9
10 # se è un testo nell'attributo text trovo il testo decodificato
11 contenuto = pagina.text
12
13 status, pagina.encoding, contenuto[:100]
```

Out[39]: (200, 'utf-8', '<!doctype html>\n<!--[if lt IE 7]> <html class="no-js ie6 lt-ie7 lt-ie8 lt-ie9"> <![endif]-->\n<!--')

le pagine JSON vengono automaticamente decodificate

```
In [40]: ▾ 1 # leggo da internet una pagina JSON
2 pagina_json = requests.get('https://api.github.com')
3
4 # la decodifica avviene automaticamente col metodo json()
5 risultato = pagina_json.json()
6 risultato
```

```
Out[40]: {'current_user_url': 'https://api.github.com/user', 'current_user_authorizations_html_url': 'https://github.com/settings/connections/applications{/client_id}', 'authorizations_url': 'https://api.github.com/authorizations', 'code_search_url': 'https://api.github.com/search/code?q={query}{&page,per_page,sort,order}', 'commit_search_url': 'https://api.github.com/search/commits?q={query}{&page,per_page,sort,order}', 'emails_url': 'https://api.github.com/user/emails', 'emojis_url': 'https://api.github.com/emojis', 'events_url': 'https://api.github.com/events', 'feeds_url': 'https://api.github.com/feeds', 'followers_url': 'https://api.github.com/user/followers', 'following_url': 'https://api.github.com/user/following{/target}', 'gists_url': 'https://api.github.com/gists{/gist_id}', 'hub_url': 'https://api.github.com/hub', 'issue_search_url': 'https://api.github.com/search/issues?q={query}{&page,per_page,sort,order}', 'issues_url': 'https://api.github.com/issues', 'keys_url': 'https://api.github.com/user/keys', 'label_search_url': 'https://api.github.com/search/labels?q={query}&repository_id={repository_id}{&page,per_page}', 'notifications_url': 'https://api.github.com/notifications', 'organization_url': 'https://api.github.com/orgs/{org}', 'organization_repositories_url': 'https://api.github.com/orgs/{org}/repos{?type,page,per_page,sort}', 'organization_teams_url': 'https://api.github.com/orgs/{org}/teams', 'public_gists_url': 'https://api.github.com/gists/public', 'rate_limit_url': 'https://api.github.com/rate_limit', 'repository_url': 'https://api.github.com/repos/{owner}/{repo}', 'repository_search_url': 'https://api.github.com/search/repositories?q={query}{&page,per_page,sort,order}', 'current_user_repositories_url': 'https://api.github.com/user/repos{?type,page,per_page,sort}', 'starred_url': 'https://api.github.com/user/starred{/owner}/{repo}', 'starred_gists_url': 'https://api.github.com/gists/starred', 'topic_search_url': 'https://api.github.com/search/topics?q={query}{&page,per_page}', 'user_url': 'https://api.github.com/users/{user}', 'user_organizations_url': 'https://api.github.com/user/orgs', 'user_repositories_url': 'https://api.github.com/users/{user}/repos{?type,page,per_page,sort}', 'user_search_url': 'https://api.github.com/search/users?q={query}{&page,per_page,sort,order}'}
```

Scaricare file binari

- il contenuto "raw" è nell'attributo **content** della risposta

```
In [42]: ▼ 1 ## possiamo anche scaricare file binari
2 logo = requests.get('https://www.python.org/static/img/python-logo@2x.png')
3
4 # posso estrarre dagli headers le dimensioni e il tipo del file ()
5 print(logo.headers['Content-Type'], logo.headers['Content-Length'])
6
7 # se si tratta di un file di dati (immagine/audio/film ...)
8 # il contenuto lo trovo nell'attributo content
9 dati = logo.content
10
11 # per salvarlo devo aprire il file in scrittura ('w') e in modo binario ('b')
▼ 12 with open('logo.png', mode='wb') as F:
13     F.write(dati)
```

image/png 15770



▼ Passare parametri ad una GET

Si usa l'argomento **params**

- **parametri={ chiave: valore, ...}**
- **requests.get(<URL>, params=parametri)**

```
In [44]: ▾ 1 # Search GitHub's repositories for requests
        ▾ 2 response = requests.get(
          3     #'https://api.github.com/search/repositories',
          4     'https://google.com',
          5     params={'q': 'requests+language:python'},
          6 )
          7 # response.json()
          8 response.text[:1000]
```

```
Out[44]: '<!DOCTYPE html><html lang="it" dir="ltr"><head><style nonce="7TONh04u7T6gIjcLn2ZHNQ">\n
a, a:link, a:visited, a:active, a:hover {\n  color: #1a73e8;\n  text-decoration: none;\n
}\nbody {\n  font-family: Roboto,RobotoDraft,Helvetica,Arial,sans-serif;\n  text-align:
center;\n  -ms-text-size-adjust: 100%;\n  -moz-text-size-adjust: 100%;\n  -webkit-text-si
ze-adjust: 100%;\n}\n.n.box {\n  border: 1px solid #dadce0;\n  box-sizing: border-box;\n  b
order-radius: 8px;\n  margin: 24px auto 5px auto;\n  max-width: 800px;\n  padding: 24px;\n
}\n.nh1 {\n  color: #2c2c2c;\n  font-size: 24px;\n  hyphens: auto;\n  margin: 24px 0;\n}\n
.n.icaCallout {\n  background-color: #f8f9fa;\n  padding: 12px 16px;\n  border-radius: 10p
x;\n  margin-bottom: 10px;\n}\n.np, .sub, .contentText, .icaCallout {\n  color: #5f6368;;\n
font-size: 14px;\n  line-height: 20px;\n  letter-spacing: 0.2px;\n  text-align: left;\n}\n
.n.signin {\n  text-align: right;\n}\n.n.saveButtonContainer,\n.n.saveButtonContainerNarrowScr
een {\n  width: 100%;\n  margin-top: 12px;\n}\n.n.customButtonContainer {'
```

Passare parametri ad altri metodi HTTP

Si usa l'argomento **data=<dizionario>**

```
requests.post('https://httpbin.org/post',
              data={'key': 'value'})
requests.put('https://httpbin.org/put',
             data={'key': 'value'})
requests.delete('https://httpbin.org/delete')
requests.head('https://httpbin.org/get')
requests.patch('https://httpbin.org/patch',
               data={'key': 'value'})
requests.options('https://httpbin.org/get')
```

