



TDDBUDDY.COM

URL Parts

The Kata

We all know about URLs like <http://www.tddbuddy.com>

Decomposes a given URL into its parts. In the above example, we would get the result

1. The **protocol**: "http"
2. The **subdomain** : "www"
3. The **domain name**: "tddbuddy.com"
4. The **port**: 80 (Default for HTTP)
5. The **path**: an empty string in our example.

Only handle top level domains like .com or .net.

Only handle the protocols specified in the default ports section below.

Be sure to deal with local network hostname only cases. E.g. <http://localhost>

Do not use built-in classes like Uri to solve this.

Do not handle second level domains like .co.uk or co.za

Default Ports

http: 80, https: 443, ftp: 21, sftp: 22

Examples

http://foo.bar.com/foobar.html <ul style="list-style-type: none">● Protocol : http● Subdomain: foo● Domain name: bar.com● Port: 80● Path: foobar.html	ftp://foo.com:9000/files <ul style="list-style-type: none">● Protocol : ftp● Subdomain: empty string● Domain name: foo.com● Port: 9000● Path: files
https://www.foobar.com:8080/download/install.exe <ul style="list-style-type: none">● Protocol : https● Subdomain: www● Domain name: foobar.com● Port: 8080● Path: download/installer.exe	https://localhost/index.html#footer <ul style="list-style-type: none">● Protocol : https● Subdomain:● Domain name: localhost● Port: 442● Path: index.html

Hints

Exclude the leading / when handling path. E.g. /download becomes download.

Be sure to write integration test at the end to dev test your solution.

URL Grammar

Below is a EBNF like grammar for a URL as per this kata.

url = protocol "://" [subdomain] host [top-level-domain] [":" port] [path] ["?" parameters] ["#" anchor]

protocol = "http" | "https" | "ftp" | "sftp"



TDDBUDDY.COM

subdomain = alphanumeric string starting with alpha

host = alphanumeric string

top-level-domain = ".com" | ".net" | ".org" | ".int" | ".edu" | ".gov" | ".mil"

port = numeric

path = alphanumeric string

parameters = alphanumeric string

anchor = alphanumeric string