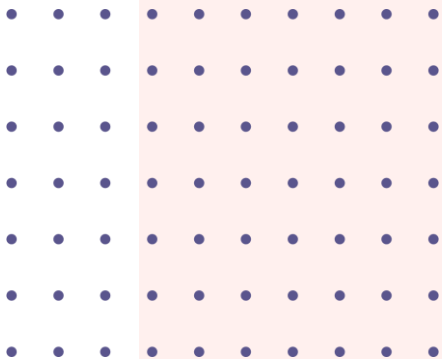


# Урок 14. Основи WEB (HTML, CSS). Xpath і CSS селектори. Character Encoding

# Структура заняття

1. Інтернет (Internet) та WEB
2. Структура HTML
3. Cookies та Cache в браузері
4. Xpath і CSS-селектори
5. Стандарти кодування символів



Що таке Internet?

Як він працює?





# Internet

міжнародна комп'ютерна мережа — всесвітня система сполучених комп'ютерних мереж, що заснована на комплекті Інтернет-протоколів. Інтернет також називають мережею мереж, бо він складається з мільйонів локальних і глобальних приватних, публічних, академічних, ділових і урядових мереж, пов'язаних між собою з використанням різноманітних дротових, оптичних і бездротових технологій.



# Всесвітня мережа (World Wide Web / WWW)

Найбільше всесвітнє багатомовне сховище інформації в електронному вигляді: десятки мільйонів пов'язаних між собою документів, що розташовані на комп'ютерах, розміщених на всій земній кулі. Вважається найпопулярнішою і найцікавішою службою мережі Інтернет, яка дозволяє отримувати доступ до інформації незалежно від місця її розташування.

# Передача даних по Internet

## TCP / IP

Transmission Control Protocol та Internet Protocol – мережева модель передачі даних, яка описує як саме інформація буде оброблятися, прийматись і передаватись

## IP address

Унікальна адреса кожного комп'ютера в мережі (192.168.123.132)

## Domain

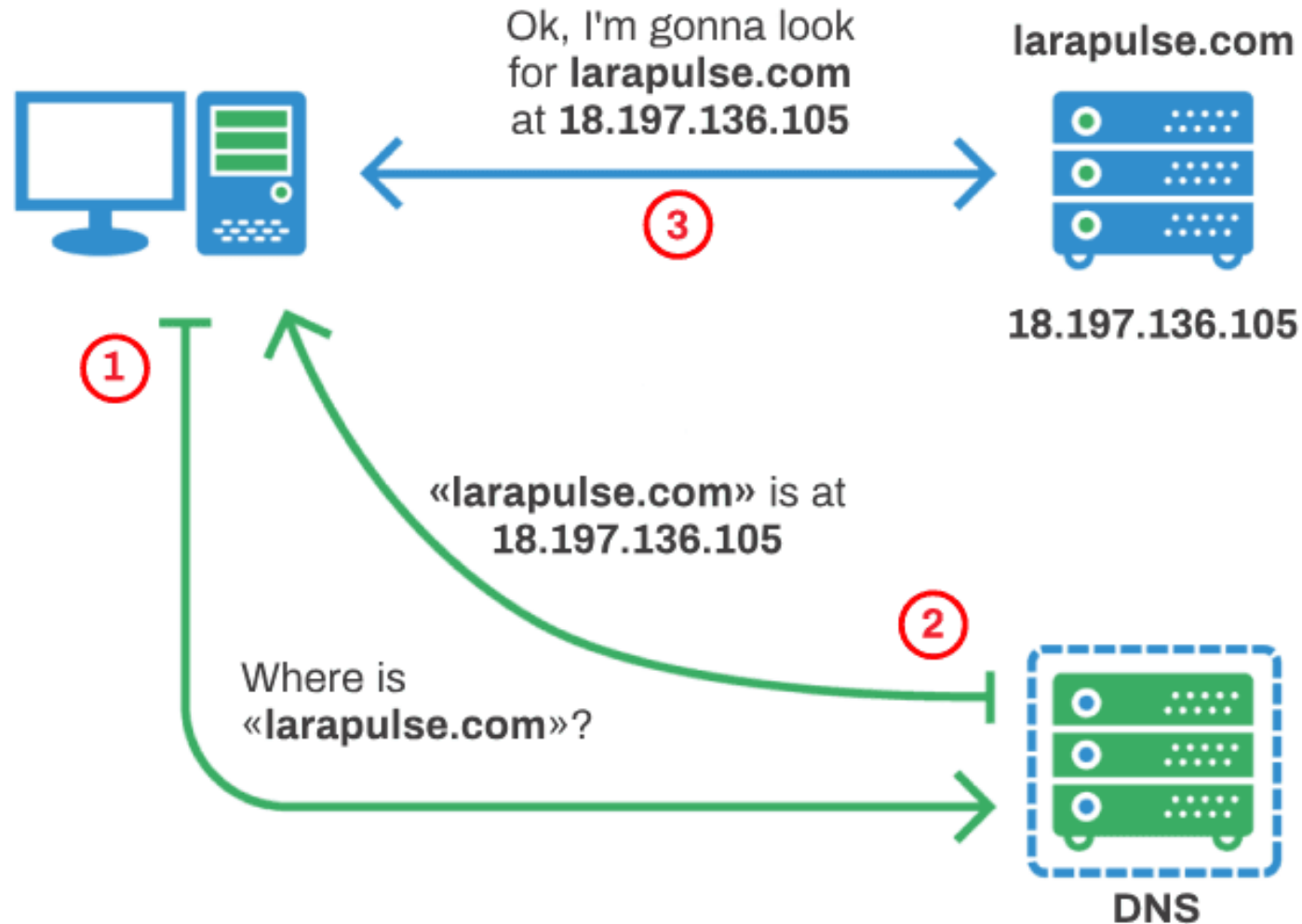
Зрозумілий псевдонім, присвоєний кожній IP адресі (beetroot.academy)

## DNS (Domain Name System)

Сервери, що зв'язують веб-адресу, що ви вводите в браузері з реальною IP адресою сайту



# DNS (Domain Name System)



# Як працює WWW

## URL (Uniform Resource Locator)

Індивідуальна адреса будь-якого об'єкта, розміщеного в Інтернеті: сайту, веб-сторінки, картки товару в онлайн-магазині, відео, документа тощо.

## HTTP (Hyper Text Transfer Protocol )

Протокол передачі даних, що використовується в комп'ютерних мережах, основним призначенням якого, є передача вебсторінок.

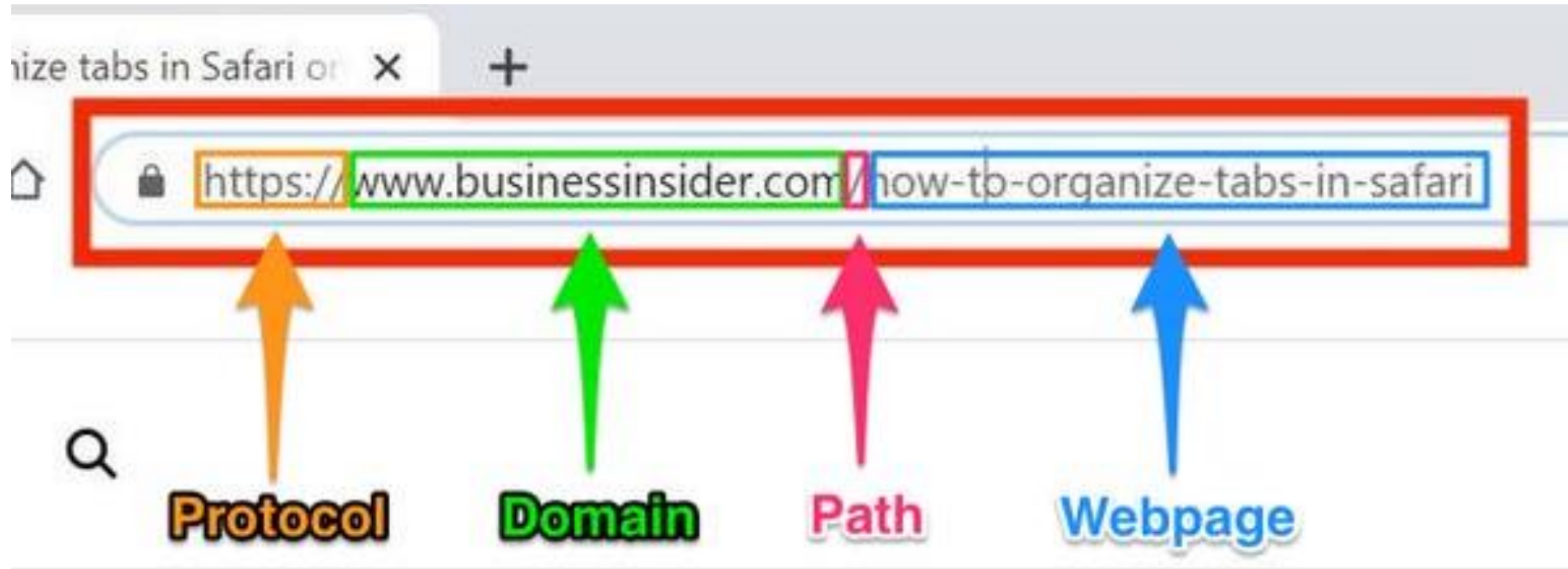
## HTML HyperText Markup Language

Мова розмітки гіпертексту – стандартизована мова розмітки документів для перегляду веб-сторінок у браузері. Веб-браузери отримують HTML документ від сервера за протоколами HTTP/HTTPS або відкривають з локального диска, далі інтерпретують код в інтерфейс, який відображатиметься на екрані монітора.





# Як працює WWW



# З чого складається WWW

Веб-сайт –  
колекція  
сторінок, що  
зв'язані між  
собою

Веб-сервер –  
комп'ютер, що надає  
програмне та  
апаратне  
забезпечення для  
функціонування  
сайсу

Пошукова  
система – сайт,  
що допомагає в  
пошуку інших  
веб-сторінок  
(google, bing,  
yahoo)

Веб-сторінка –  
документ, що  
відображається  
браузерами

Браузер –  
програма для  
перегляду веб –  
сторінок  
(Chrome, Safari,  
IE, FireFox etc)



# Кросбраузерність

Здатність веб-ресурсу відображатися однаково добре у всіх популярних браузерах без перебоїв у функціонуванні та помилок у верстці, з однаково коректною читабельністю контенту. Це дуже важливий показник як для пошукових систем, так і аудиторії користувачів.



Усі операційні системи браузера по різному обробляють графічні та візуальні елементи!

# Практична робота

1. За допомогою запиту до Google за ключовою фразою “check my IP” визначити свою IP-адресу (результат від Google та будь-якого сервісу, що видає пошук).
2. За допомогою ресурсу <https://check-host.net> визначити IP-адреси сайтів [rozetka.com.ua](https://rozetka.com.ua) та [google.com.ua](https://google.com.ua).
3. За допомогою термінальної команди `tracert` (Windows) визначити маршрут від локального комп'ютера до адреси [rozetka.com.ua](https://rozetka.com.ua).  
`[tracert rozetka.com.ua]`

# HTML - Hyper Text Markup Language



- Мова розмітки, що використовується для повідомлення вашому браузеру, як відображати веб-сторінки, які ви відвідуєте.
- HTML описує структуру веб-сторінки
- Елементи HTML (теги) позначають фрагменти вмісту сторінки, наприклад <список>, <таблиця>, <посилання> тощо

[Створимо свою сторінку!](#)



# CSS – Cascading Style Sheets



Спеціальна мова стилю сторінок, що використовується для опису зовнішнього вигляду HTML-документу.

[Покращимо свою сторінку!](#)

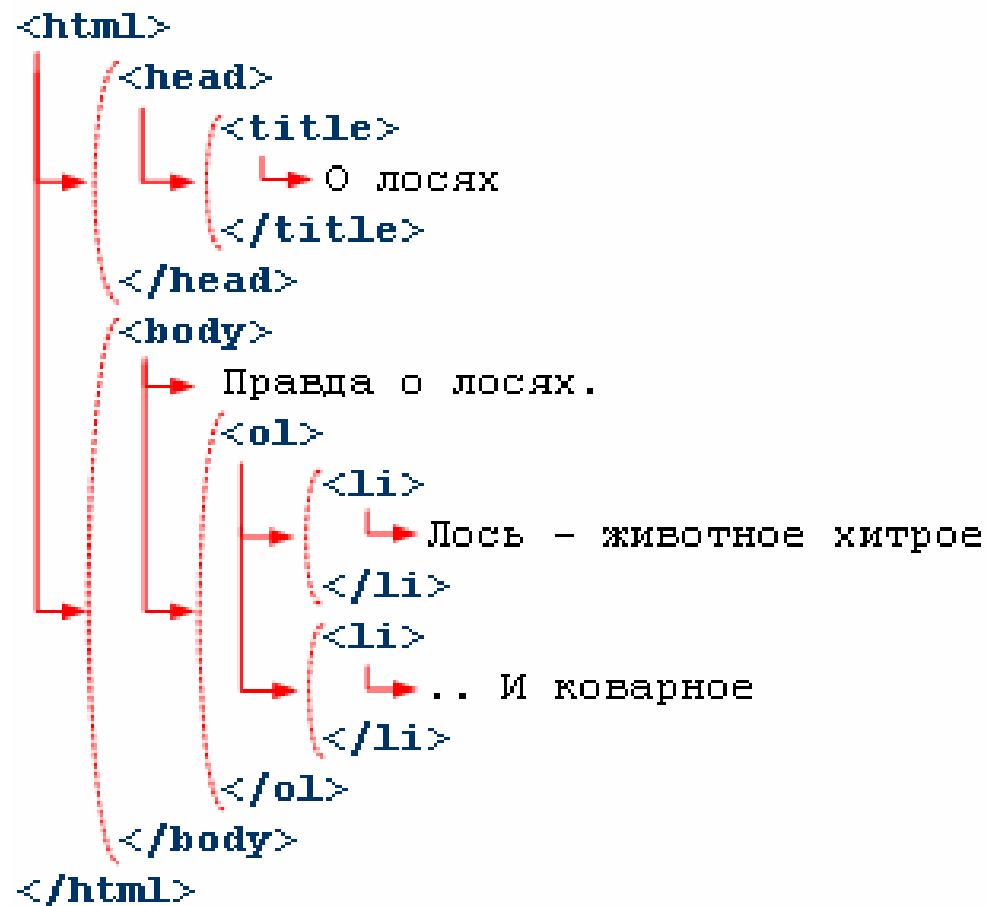
- Вбудований CSS використовується для окремого HTML елемента:  
`<h1 style="color:blue;">A BlueHeading</h1>`
- Внутрішній CSS використовується для визначення стилю для однієї HTML-сторінки: `<style> </style>`
- Зовнішній CSS використовується для визначення стилю для багатьох HTML-сторінок:  
`<link rel="stylesheet" href="styles.css">`

# DOM – Document Object Model

Представлення HTML документа у вигляді дерева тегів.

Браузер створює DOM для того, щоб за допомогою JavaScript та інших скриптів можна було швидко керувати вебдокументом: шукати потрібний елемент, додавати нові елементи, отримати наступний дочірній елемент тощо.

[Переглянемо дерево тегів!](#)





# Кеш (Cache)

Термін «кеш» спочатку походить з французької мови і означає «укриття». Він так називається, тому що прихований від користувача. У більшості випадків цей термін застосовується по відношенню до браузерів. Але в іншого програмного забезпечення також може бути свій кеш.

Наприклад, при відкритті сайтів, в браузері зберігається базове вміст веб-сайту – і все це знаходиться на вашому комп'ютері. Цей процес працює у фоновому режимі і непомітний для користувача. Якщо ви пізніше повернетеся на наш сайт, вміст кешу буде завантажено з ПК. Таке технологічне рішення дозволяє завантажити сайт набагато швидше.





# Чому треба чистити кеш?

## Причина 1.

Він діє як свого роду короткочасна пам'ять браузера. Якщо кеш повинен зберігати все більше і більше інформації, це може уповільнити роботу комп'ютера

## Причина 2.

Якщо ви хочете зберегти певну конфіденційність в інтернеті, необхідно регулярно очищати кеш. Звичайно, після цієї операції вам доведеться знову входити в облікові записи в усіх соціальних мережах, але нічого страшного, зате пам'ять потренуєте.

## Причина 3.

Іноді кешовані файли перешкоджають відображенню актуального вмісту сайтів, так як завантажуються застарілі збережені дані



# Куки (Cookies)

Маленькі файли, які зберігаються на комп'ютері користувача для спрощення роботи в інтернеті.

COOKIES мають свій термін дії. За умовчанням вони зберігаються на період однієї сесії в Інтернеті, а коли користувач закриває браузер, вони видаляються. Але все ж є куки з установленим терміном дії, вони зберігаються постійно, а видаляються або за запитом користувача ПК, або за терміном, який встановлено.

Куки - невід'ємна частина для налаштування реклами за інтересами, тому їх використовують для відстеження переваг користувачів і налаштування тематичної реклами.



# Що зберігають куки файли?

- дані входу на сайти (логін і пароль);
- мову, валюту, яку вважає за краще користувач сайту;
- товари, які вже в кошику, або які ви переглядали на конкретному сайті;
- IP і місце розташування користувача сайту, дату і час відвідин;
- версію ОС і веб-браузера;
- кліки та переходи користувача



## Коли куки просто необхідні для швидкої роботи в інтернеті?

- Для здійснення автоматичної авторизації на сайтах. Коли при повторному заході на сайт, немає необхідності вводити дані облікового запису. Це дуже зручно і економить час користувача, так як браузер вже зберігає його інформацію і здійснює авторизацію без зайвих дій.
- Для зберігання налаштувань пошукових систем, які вже "знають" регіон проживання, мову користувача, і при кожному заході браузеру немає необхідності визначати ці дані.
- Для зручності здійснення покупок в інтернеті. Кошик на сайті використовує куки файли, тому, навіть якщо ви закриєте сторінку сайту, відкладені в кошик товари ще вас чекають.



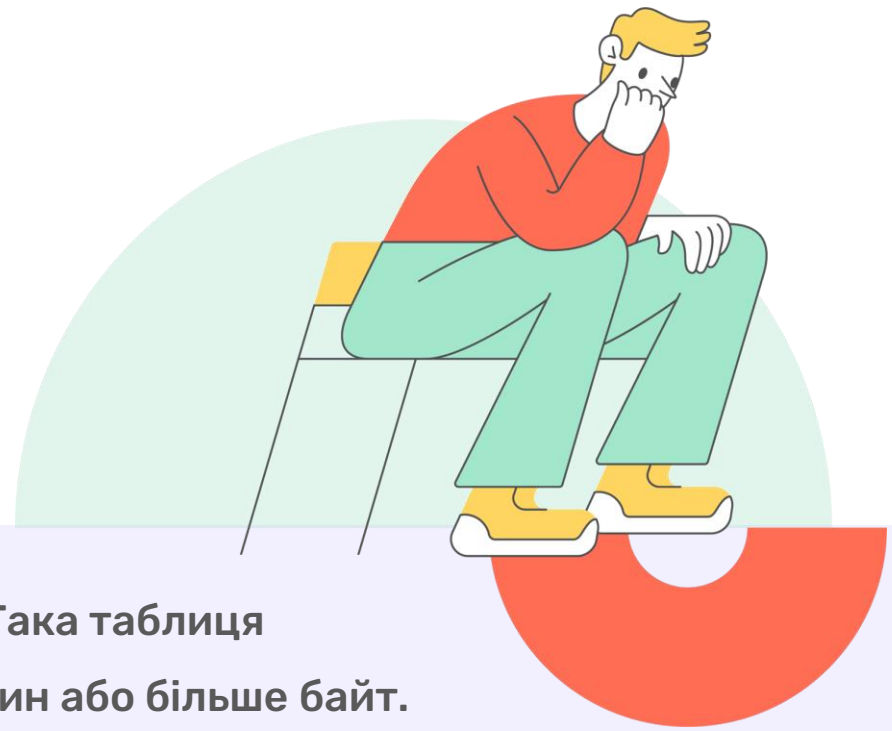
## Основні причини для очищення куки файлів

- для того, щоб інші користувачі вашого комп'ютера не змогли скористатися Вашими логінами при вході на сайти;
- щоб не засмічувати місце на жорсткому диску комп'ютера і не сповільнювати тим самим роботу Вашого ПК;
- з боку конфіденційності за допомогою куки можна простежити Вашу роботу в Мережі, а це вже не дуже добре;
- в разі, коли окремі сторінки сайту або сам сайт відображається некоректно тільки для вашого ПК. У подібній ситуації може допомогти чистка cache і cookie браузера, так як з комп'ютера видаляються тимчасові файли, які і могли викликати неправильне відображення сайту

# Практична робота


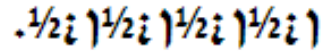
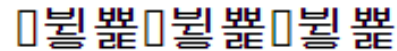
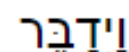
1. Перейти на сайт <https://hotline.ua/> і визначити мову, яка використовується.
2. Перевірити, яка мова записана до Cookies, використовуючи Dev Tools (Вкладка Application -> Cookies).
3. Змінити мову сайту та перевірити зміни в Cookies.
4. Вручну змінити мову в Cookies та перезавантажити вебсторінку. Перевірити Cookies.

# Character encoding (кодування символів)



- Певна таблиця кодування скінченної множини символів. Така таблиця зіставляє кожному символу послідовність довжиною в один або більше байт.
- Використовується для правильного відображення сторінки HTML
- Основні стандарти кодування символів (ASCII, ANSI, 8859, UTF-8)

```
<!DOCTYPE html>
<html class="client-js ve-available" lang="uk" dir="ltr">
  <head>
    <meta charset="UTF-8">
    <title>Кодування символів – Вікіпедія</title>
```

C	D
. 	UTF-8
. 	Windows-1255
	UTF-16
	Passed with function

Enable macro if the data encoding is incorrect

3XVFC,,нlв.№ъ...-6yD-©ХЙЁЌЎъ?™—УЛсйгР>,Г\$|ѓ<%oB||к†д%у}ЙЉR7iK9-йN+‘@Щѓ<ХЇ!  
Cц\$O>-“46[шZSfA‘\$4iТяыђ`°БЖBSЛ zo I Љ\X  
3x°0□`e"]!ЖДcLgІгB9Й1©frгук»X†If6YINKёЄЪVO-uЙАЄЪИѓe©“±μ1V±s\$Lj%цы[щГ`л-  
;c©Вч+б4ыяДSp[;:пцfэХђ'ў«імљ<гFbpкyі/”Гдљ>”Ч©і”oRБ"ЯгCg•ka†ѓ6фCBA8MCK>4...  
k•&тiБЉэLCJXCM”Рљ\*АДЩvГВк’...ђ3иіГы\$O-⟨љIN|sSS3|Ю5Щн?LЎР}™...Р  
в'АтІшv~ђФ6ьBr€>A5'3 мiзЭ&&JF"‡5 @|чђMЉГВ†`ЖкьИ ⟨Be©ђiB€·DjТБ...Ж;цgW-  
3Љ[ђђ[ђѓЎщжО<EPB @™5Hu,8IђμBxм&в№г=нЦ \™©<V™,Ю\,ьQгm g'•HEрш[\\A  
нЩгГМ±3љЛ[Юi]P24<3@ШN4  
ою,3K],ГJхu,4љ» L0Чг`ЖТйCкNPЧЎРЃeb



## Character encoding (кодування символів)



- ASCII був першим стандартом кодування символів. ASCII визначив 128 різних символів, які можна використовувати в Інтернеті: цифри (0-9), англійські літери (AZ) і деякі спеціальні символи, як-от ! \$ + - ( ) @ < > .
- ISO-8859-1 був стандартним набором символів для HTML 4. Цей набір символів підтримував 256 різних кодів символів. HTML 4 також підтримує UTF-8.
- ANSI (Windows-1252) був оригінальним набором символів Windows. ANSI ідентичний ISO-8859-1, за винятком того, що ANSI має 32 додаткові символи.
- Специфікація HTML5 заохочує веб-розробників використовувати набір символів UTF-8, який охоплює майже всі символи та символи у світі!



# XPath (XML path language)

Мова запитів для вибору вузлів з XML-документів у DOM-дереві або для обчислення величин (наприклад, рядкових, числових або булевих) на основі вмісту XML-документа. XPath була розроблена World Wide Web Consortium (W3C). Мова XPath заснована на представленні XML-документа у вигляді дерева і надає можливість навігації деревом та вибору його вузлів за різними критеріями.

Select current node

Select attributes

Value of an attribute

Xpath=//tagname[@Attribute='Value']

Tagname like input,  
div, img etc

Attribute Name



## Приклади побудови XPath

- `tag_name` – обирає всі вузли з ім'ям `tag_name`
- `/` – обирає від кореневого вузла
- `//` – обирає вузли в документі від чинного вузла
- `.` – обирає чинний вузол
- `..` – обирає батька чинного вузла
- `@` – обирає атрибути

# Приклади побудови XPath

- Absolute (абсолютний) XPath

```
/html/body/div[2]/div[1]/div/h4[1]/b/html[1]/body[1]/div[2]/div[1]/div[1]/h4[1]/b[1]
```

- Relative (відносний) XPath

```
Relative XPath: //div[@class='featured-box cloumns1']//h4[1]//b[1]
```

# Практична робота

**1.** Для вебсайту youtube.com (потрібно попередньо вилогінитись) побудувати вручну XPath для таких елементів:

- бургер-меню в лівому верхньому кутку
- поле для пошуку
- кнопка "Sign in".

Знайти ці елементи за створеним XPath.

Порівняти Створений XPath з тим, що генерує сам браузер (RMC на елементі -> Копіювати -> Копіювати повний XPath).



# CSS-селектор

Частина CSS-правила, яка повідомляє браузеру, до якого елемента вебсторінки буде застосований CSS-стиль. Тобто, селектор – це вибір та формальний опис того елемента чи групи елементів, до яких будуть застосовані CSS-стилі.



## Приклади побудови CSS-селектора

- `*` – всі елементи,
- `tag1 > tag2` – обрати тег, вкладений в інший тег
- `h1` – обрати всі теги з таким ім'ям
- `#id` – обрати тег зі вказаним id
- `.class` – обрати тег зі вказаним класом



# Практична робота

1. Для вебсайту `youtube.com` (потрібно попередньо вилогінитись) побудувати вручну CSS-селектор для тих самих елементів:
  - бургер-меню в лівому верхньому кутку
  - поле для пошуку
  - кнопка "Sign in".

Знайти ці елементи за створеним CSS-селектором.

Порівняти Створений XPath з тим, що генерує сам браузер (RMC на елементі -> Копіювати -> Копіювати селектор).



Домашнє завдання