

Internet- технології

Мова розмітки HTML

Версія слайдів 2023.02.20

Рекомендовані джерела

- Documenting web technologies, including CSS, HTML, and JavaScript (MDN) <https://developer.mozilla.org/>
- Офіційний сайт W3C <https://www.w3.org>
- W3schools <http://www.w3schools.com/>
- Керівництво по HTML5 и CSS3 <https://metanit.com/web/html5/>
- Стислий довідник (з використанням MDN) по тегам мови розмітки <https://overapi.com/html>
- Типові приклади використання елементів HTML <https://htmlcheatsheet.com>
- Сервіс перевірки вихідного коду сайту на відповідність стандартам <https://validator.w3.org/>
- SEO (пошукова оптимізація) <https://moz.com/learn/seo>

Мова розмітки HTML

HTML (HyperText Markup Language – «мова гіпертекстової розмітки») – мова розмітки документів, що використовується у мережі Internet. Документ на мові HTML інтерпретується браузером, отриманий після інтерпретації результат відображається користувачеві. HTML-сторінки зазвичай передаються браузеру від сервера за протоколами HTTP або HTTPS, у вигляді простого тексту або з використанням шифрування.

Мова розмітки – набір символів або послідовностей, які вставляються в текст для передачі інформації про його виведення або структуру. Таким чином, текстовий документ, написаний з використанням мови розмітки, містить не тільки сам текст (як послідовність слів і розділових знаків), а й додаткову інформацію про різні його фрагменти – наприклад, вказівку на заголовки, форматування, списки, інтерактивні елементи тощо.

Мова розмітки **не є мовою програмування!**

Приклади мов розмітки: HTML, XML, XHTML, MathML, TeX, LaTeX.

Мова розмітки HTML. Теги

Будь-який документ на мові HTML є набором елементів, початок і кінець яких позначається **тегами**. Крім вмісту, елементи можуть мати атрибути, що визначають якісь властивості елементу.

Загальна структура тегів:

`<ім'я_тега атрибути>вміст елемента</ім'я_тега>`

Теги можуть бути парними та непарними – коли не потрібно закривати тег (тобто елемент створюється одним тегом, без завершення). Такі елементи, створені непарними тегамі, не містять тексту або вкладених елементів, наприклад, тег розриву рядка **
**.

Приклади:

```
<form action="handler.php">
```

```
  <input type="radio" name="browser" value="ie"> Internet Explorer <br>
```

```
  <input type="radio" name="browser" value="opera"> Opera <br>
```

```
</form>
```

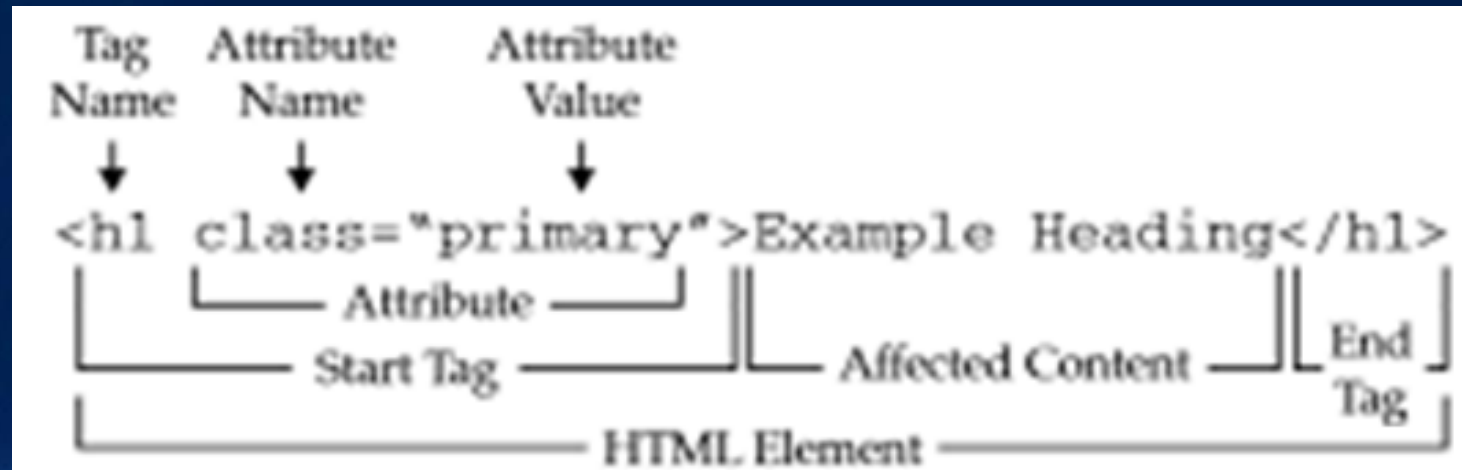
```
<hr><br>
```

- ☐ Internet Explorer
- ☐ Opera

Мова розмітки HTML. Теги

Елементи можна вкладити один в одного. При цьому необхідно дотримуватися правильної вкладеності тегів – не можна вкладений тег закривати після закриття батьківського.

Атрибути – елементи мови розмітки, які містять додаткову інформацію про елемент. У загальному випадку, атрибут має значення.



Мова розмітки HTML. Теги, приклади

- для визначення структури html-документа(<!DOCTYPE>, <html>, <head>, <body>);
- які описують структуру тіла документа(<header> и т.п., заголовки від <h1> до <h6>, параграфи <p>, блоки <div>, розділювачі (
 и <hr/>), теги створення списків (,), таблиць);
- теги фізичного та логічного форматування (виділити курсивом (<i>), використати жирне накреслення (), виділити як цитату (<cite>) и т.і.);
- зображення та карти зображень <map>;
- гіперпосилання <a>;
- форми для організації діалогу з користувачем <form>, <input>, <select>;
- тег для опису скриптів <script>.
- тег для опису стилів <style>.

Мова розмітки HTML

Регістр, в якому вказано ім'я елемента та імена атрибутів, у HTML значення не має (на відміну від XHTML). Але на практиці, написання тегів та їх атрибутів малими літерами (тобто в нижньому регістрі) є де-факто стандартом.

Значення будь-яких атрибутів необхідно укласти у лапки.

Крім елементів, у HTML-документах є і «сутності» (англ. entities) для використання спеціальних символів. Сутності починаються з символу **амперсанду** (закінчуються **крапкою з комою**) і мають вигляд `&ім'я;` або `&#NNNN;`, де NNNN – код символу в Юнікодi в десятковій системі числення.

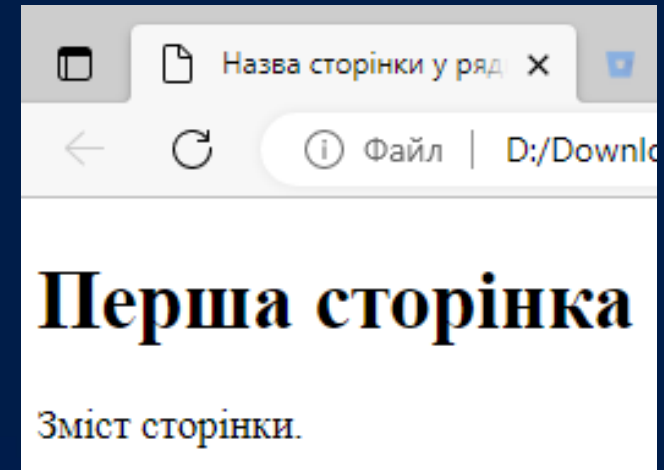
Наприклад, `&сору;` – знак авторського права (©). Як правило, сутності використовуються для представлення символів, відсутніх у поточному кодуванні документа, або для представлення спеціальних символів : `&амп;` – амперсанду (&), `<` – символу «менше» (<) та `>` – символу «більше» (>), які некоректно записувати «звичайним» чином, через їхнє особливе значення в HTML.

Мова HTML розмітки. Представлення спецсимволів

Числовий код	Іменна заміна	Символ	Опис
"	"	"	Лапки
&	&	&	Амперсанд
<	<	<	Менше
>	>	>	Більше
 	 		Нерозривний пробіл
©	©	©	Копірайт
«	«	«	Ліва кутова лапка
®	®	®	Зареєстрована торгова марка
±	±	±	Плюс або мінус
»	»	»	Права кутова лапка

Загальна структура HTML-документа

```
<!DOCTYPE html>
<html>
  <!-- секція заголовок документа -->
  <head>
    <title>Назва сторінки у рядку заголовок браузера</title>
  </head>
  <!-- тіло документа -->
  <body>
    <header>
      <h1>Перша сторінка</h1>
    </header>
    Зміст сторінки.
  </body>
</html>
```



Загальна структура HTML-документа

`<!DOCTYPE html>` – вказівка браузеру, що використовується мова розмітки HTML п'ятої версії.

Тег `<html>` є контейнером, який містить у собі весь вміст веб-сторінки, включаючи теги `<head>` та `<body>`.

Тег `<head>` містить у собі інформацію про сторінку – заголовок, метатеги, посилання на стилі та скрипти та іншу службову інформацію.

Тег `<body>` є контейнером для всього вмісту, який буде відображено користувачеві.

Якщо структура сторінки порушена – браузер все одно спробує відобразити сторінку, додавши необхідні структурні теги, за своїм розумінням. Зловживати цим не варто.

Елемент `<!DOCTYPE>`

`<!DOCTYPE>` – DTD (document type definition, опис типу документа) – призначений для вказівки браузеру, як саме слід інтерпретувати веб-сторінку, оскільки стандартів мови HTML існує кілька версій.

Формат DTD (загальний, його можливо зустріти для документів, які створені з використанням **застарілих** версій HTML):

```
<!DOCTYPE [Елемент верхнього рівня] [Публічність]  
"[Реєстрація]//[Організація]//[Тип] [Ім'я]//[Мова]" "[URL]">
```

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"  
"http://www.w3.org/TR/html4/strict.dtd">
```

Такий DOCTYPE задає тип кореневого елемента документа (`<html>`), публічний ідентифікатор (`-//W3C//DTD HTML 4.01//EN`) показує, хто є автором DTD (W3C); назва DTD (`DTD HTML 4.01`); і мова, якою написаний DTD (`EN`, тобто, англійська). DTD не вказує на мову самої Web-сторінки; там вказана мова, якою складено DTD. Для коректного відображення Web-сторінки у браузері бажано вказати, з яким стандартом вона узгоджена.

Елемент `<!DOCTYPE>`. Поширені DTD для HTML

HTML 4.01 (застарілі! Але інколи все ще зустрічаються)

Суворий синтаксис HTML:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
```

Перехідний синтаксис HTML:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"  
"http://www.w3.org/TR/html4/loose.dtd">
```

В HTML-документі використовуються фрейми.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN"  
"http://www.w3.org/TR/html4/frameset.dtd">
```

HTML 5

```
<!DOCTYPE html>
```


Контрольні запитання

- З яких структурних одиниць складається HTML-документ?
- Для чого використовується `<!DOCTYPE>`?
- Яка версія мови HTML використовується зараз?
- Що таке атрибути? Навести приклади.
- У якому регістрі записуються HTML-теги?
- Чи можна використовувати такі символи, як '>', '<' або '&' у тексті HTML-сторінки?
- У чому різниця між тегами `<head>` та `<header>`?