

Osnove rada sa tekстом

Osnova svakog HTML dokumenta jeste tekst. Ipak, tekst unutar HTML dokumenta može imati različite uloge, baš kao i unutar bilo kojeg drugog tekstualnog dokumenta gde se mogu razlikovati različite tekstualne sekcije. Tako se pod pojmom teksta mogu podvesti naslovi, paragrafi, prelasci u novi red, prazna mesta, vertikalne linije, kao i razni drugi oblici teksta koji posebnim formatiranjem dobijaju specijalno značenje. Modul koji je pred vama obradiće sve relevantne elemente koji omogućavaju dodavanje teksta HTML dokumentima.

Naslovi (headings)

Svaki dokument počinje naslovom. Različiti delovi dokumenta mogu biti identifikovani podnaslovima različitog nivoa. HTML jezik poznaje 6 nivoa naslova, koji se kreiraju korišćenjem sledećih HTML elemenata:

- `<h1>`
- `<h2>`
- `<h3>`
- `<h4>`
- `<h5>`
- `<h6>`

Primer različitih HTML naslova:

```
<h1>Lorem ipsum dolor sit amet</h1>
<h2>Lorem ipsum dolor sit amet</h2>
<h3>Lorem ipsum dolor sit amet</h3>
<h4>Lorem ipsum dolor sit amet</h4>
<h5>Lorem ipsum dolor sit amet</h5>
<h6>Lorem ipsum dolor sit amet</h6>
```

Prikazani kod u browseru proizvešće rezultat kao na slici 3.1.

Lorem ipsum dolor sit amet

Lorem ipsum dolor sit amet

Lorem ipsum dolor sit amet

Lorem ipsum dolor sit amet

Lorem ipsum dolor sit amet

Lorem ipsum dolor sit amet

Slika 3.1. Različiti tipovi naslova HTML dokumenta

Kao što se sa slike 3.1. može videti, naslovi različitog nivoa prikazuju se korišćenjem različite veličine slova i formatiranja. Takođe, prilikom prikaza browser dodaje po jednu praznu liniju pre i posle naslova.

Naslovi se u jednom HTML dokumentu navode po hijerarhiji naslova. Tako se h1 koristi za glavne naslove, h2 za podnaslove i tako dalje. Ovu hijerarhiju potrebno je poštovati, tako da nije ispravno koristiti, na primer, h4 naslov odmah nakon h1 naslova.

Hijerarhija naslova zasebno se primenjuje na različitim regionima HTML dokumenta, naravno, ukoliko je HTML dokument sačinjen iz više regiona. Tako je moguće imati više <h1> naslova unutar jednog dokumenta koji bi se odnosili na različite sekcije dokumenta. O različitim sekcijama HTML dokumenta biće više reči u jednoj od narednih lekcija.

Paragrafi (paragraphs)

U HTML dokumentima, paragrafi se kreiraju postavljanjem teksta između otvarajućeg <p> i zatvarajućeg </p> taga.

Primer paragrafa u HTML dokumentu:

```
<p>Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do  
eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad  
minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip  
ex ea commodo consequat.</p>  
<p>Duis aute irure dolor in reprehenderit in voluptate velit esse  
cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat  
cupidatat non proident, sunt in culpa qui officia deserunt mollit anim  
id est laborum.</p>
```

Prikazani kod proizvešće rezultat kao na slici 3.2.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do
eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut
enim ad minim veniam, quis nostrud exercitation ullamco laboris
nisi ut aliquip ex ea commodo consequat.

Duis aute irure dolor in reprehenderit in voluptate velit esse cillum
dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat
non proident, sunt in culpa qui officia deserunt mollit anim id est
laborum.

Slika 3.2. Dva paragrafa u jednom HTML dokumentu, nakon parsiranja

HTML elementi za formatiranje teksta

HTML jezik poznaje specijalne elemente za definisanje teksta korišćenjem različitih formatiranja. Ti specijalni elementi prikazani su tabelom 3.1.

Element	Opis	Primer	Efekat
	bold tekst	Lorem ipsum	Lorem ipsum
	posebno važan tekst	Lorem ipsum	Lorem ipsum
<i>	italic tekst	<i>Lorem ipsum</i>	<i>Lorem ipsum</i>
	istaknut tekst	Lorem ipsum	<i>Lorem ipsum</i>
<mark>	markiran tekst	<mark>Lorem ipsum</mark>	Lorem ipsum
<small>	sitan tekst	<p>Normal <small>Small</small></p>	Normal Small
	obrisan tekst	Lorem ipsum	Lorem ipsum
<ins>	umetnut tekst	<ins>Lorem ipsum</ins>	<u>Lorem ipsum</u>
<sub>	subscript	<p>This is _{subscripted} text.</p>	This is subscripted text.
<sup>	superscript	<p>This is ^{superscripted} text.</p>	This is superscripted text.

Tabela 3.1. Elementi za formatiranje teksta

Kao što se iz tabele 3.1. može videti, kako bi se nad tekstom postigao bold efekat, moguće je koristiti elemente i . Ipak, između dva pomenuta elementa postoji razlika. Iako oba elementa omogućavaju kreiranje bold teksta, tekst unutar elementa ima semantičko značenje, čime mu se daje na značaju. Ista situacija je i sa <i> i elementima. Iako oba proizvode italic tekst, tekstu daje posebno značenje.

Posebno semantičko značenje teksta koristi se od strane raznih čitača HTML dokumenata, koji nailaskom na ovakve tagove znaju da je takav deo teksta naglašen, odnosno bitniji od ostatka.

Pitanje

Italic tekst se u HTML jeziku kreira:

- a) tagovima ili <i>
- b) tagom
- c) tagom <i>
- d) tagom <italic>

Objašnjenje

Za kreiranje italic teksta mogu se koristiti ili <i> tagovi. Tag <i> postavlja italic varijantu fonta, dok tag postavlja i semantičko značenje. Pomenuti tag označava da se tekst unutar njega ističe od ostatka rečenice. Em je skraćeno od engleske reči *emphasize*, što znači naglasiti. I je skraćeno od *italic*.

Novi red

Prilikom korišćenja paragraf elementa, automatski se prelazi u novi red iznad i ispod teksta paragrafa. Ipak, ukoliko je potrebno samostalno uticati na to, gde će doći do preloma reda teksta, moguće je koristiti tag `
`.

Element `
` nema svoj sadržaj i jedan je od elemenata koji se kreiraju korišćenjem samozatvarajućeg taga, s obzirom na to da ne poseduje zatvarajući tag. Primer njegove upotrebe je sledeći:

```
<p>Excepteur sint occaecat,<br>cupidatat non proident, sunt in culpa  
qui officia deserunt mollit anim id est laborum</p>
```

Prikazani kod proizvodi efekat kao na slici 3.3.

Excepteur sint occaecat,
cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum

Slika 3.3. Prelazak u novi red

Hr

`<Hr>` je element koji omogućava tematski prelom teksta. Ipak, ovaj element dobio je naziv po pojmu *horizontal rule*, što je linija koju je moguće koristiti kako bi se razdvojile dve sekcije teksta. Većina browsera upravo tako i renderuje ovaj element, koji, pored svog vizuelnog efekta, poseduje i semantičko značenje, koje omogućava tematsko razdvajanje teksta.

Baš kao i `
` element, `<hr>` ne poseduje sadržaj i kreira se korišćenjem samozatvarajućeg taga.

Primer upotrebe `<hr>` elementa.

```
<p>Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do  
eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim  
ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut  
aliquip  
ex ea commodo consequat.</p>  
<hr>  
<p>Duis aute irure dolor in reprehenderit in voluptate velit esse  
cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat  
cupidatat non proident, sunt in culpa qui officia deserunt mollit anim  
id est laborum.</p>
```

Prikazani kod proizvodi efekat kao na slici 3.4.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Slika 3.4. Efekat koji proizvodi hr element

Razmaci i redovi u tekstu

U ovom trenutku veoma je bitno napomenuti nekoliko stvari o razmacima i novim redovima u tekstu. Za primer će poslužiti sledeći HTML kod:

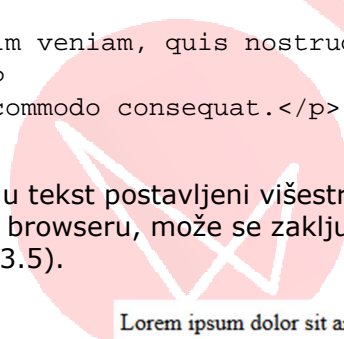
```
<p>Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.</p>
```

Prikazani kod proizvodi jedan paragraf (pasus) teksta. Tekst paragrafa može se, radi demonstracije, napisati i ovako:

```
<p>Lorem ipsum dolor sit amet,                consectetur adipisicing
elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.
Ut enim

ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut
aliquip
ex ea commodo consequat.</p>
```

Ovoga puta su u tekst postavljeni višestruki razmaci, ali i prelasci u novi red. Kada se ovakav tekst prikaže u browseru, može se zaključiti da browser ignoriše dodate razmake i prelaske u novi red (slika 3.5).



>Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Slika 3.5. Jedan paragraf

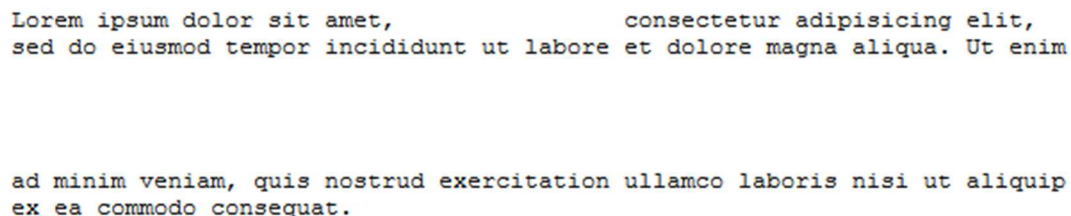
Razlog za upravo ilustrovano ponašanje veoma je jednostavan. Browser ne uzima u obzir višestruke razmake, kao i prelaske u novi red unutar HTML koda. Višestruki razmaci se uvek tretiraju kao jedan razmak, dok se prelasci u novi red ignorišu. Ukoliko je potrebno preći u novi red, koristi se element `
`, kao što je to prikazano nešto ranije u ovoj lekciji.

Ukoliko se nekada ipak javi potreba za očuvanjem višestrukih razmaka, kao i prelazaka u novi red bez upotrebe elementa `
`, moguće je koristiti element `<pre>`. Kompletan sadržaj unutar `<pre>` elementa zadržava sve razmake i prelasku u novi red. Njegova upotreba ilustrovana je primerom:

```
<pre>
<p>Lorem ipsum dolor sit amet, consectetur adipisicing elit,
sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut
enim

ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut
aliquip
ex ea commodo consequat.</p>
</pre>
```

Ovoga puta efekat prikazanog koda jeste kao na slici 3.6.



```

Lorem ipsum dolor sit amet,                consectetur adipisicing elit,
sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim

ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip
ex ea commodo consequat.
```

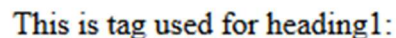
Slika 3.6. Jedan paragraf sa očuvanim razmacima i prelascima u novi red

Specijalni znaci

Određeni karakteri u HTML jeziku poseduju specijalno značenje. Idealni predstavnici su svakako karakteri veće i manje (`<>`), pomoću kojih se grade HTML tagovi. Već više puta je prikazano da HTML tagovi unutar dokumenta služe kao uputstvo za izgradnju stranice upućeno browseru. Jednostavno, u procesu parsiranja browser izostavlja tagove i na stranici prikazuje samo sadržaj HTML elemenata. Na primer, takva situacija postoji u sledećem primeru:

```
<p>
This is tag used for heading1: <h1>
</p>
```

Efekat koji prikazani kod ima jeste kao na slici 3.7.



This is tag used for heading1:

Slika 3.7. Sadržaj paragrafa ne sadrži specijalne znake, zato što ih browser interpretira na specijalan način

Kao što se sa slike 3.7. može videti, tag `<h1>`, koji je napisan unutar sadržaja paragrafa izgubio se s obzirom na to da ga je browser interpretirao na specifičan način. Sada se postavlja pitanje šta ukoliko želimo da unutar HTML dokumenta ispišemo neki HTML tag u njegovom izvornom, tekstualnom obliku.

Odgovor na ovo pitanje leži u upotrebi znamenskih karaktera. Znamenski karakteri ili HTML entiteti, kako se drugačije nazivaju, korisni su u nekoliko različitih situacija. Tako se HTML entiteti mogu koristiti da reše problem rezervisanih karaktera. HTML entiteti takođe se mogu koristiti i za pisanje nekih specijalnih simbola za koje ne postoji taster na tastaturi. Neki od najkorišćenijih HTML entiteta prikazani su tabelom 3.2.

HTML entitet	Izgled na HTML strani
 	(blanko, tj. non breaking space)
<	<
>	>
&	&
©	©
»	»

Tabela 3.2. HTML entiteti

Sada se prethodni primer može realizovati korišćenjem zamenskih karaktera na sledeći način:

```
<p>
This is tag used for heading1: &lt;h1&gt;
</p>
```

Ovoga puta efekat će biti kao na slici 3.8.

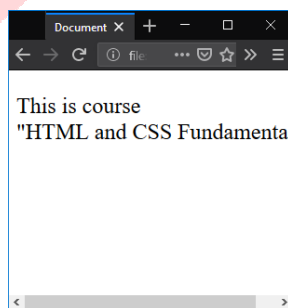
This is tag used for heading1: <h1>

Slika 3.8. Prikaz specijalnih karaktera kao sadržaja elemenata upotrebom HTML entiteta

U tabeli 3.2. prikazan je i *non breaking space* HTML entitet. On se može koristiti ukoliko je bitno naglasiti browseru da određene reči ne smeju biti razbijene na redove. To ilustruje sledeći primer HTML koda:

```
<p>This is course "HTML&nbsp;and&nbsp;CSS&nbsp;Fundamentals"</p>
```

U prikazanom kodu između svake reči naziva kursa *HTML and CSS Fundamentals* postavljen je *non breaking space* HTML entitet i na taj način naglašeno je da se prelamanje teksta na tom mestu ne sme sprovoditi. Tako će efekat prikazanog koda biti kao na slici 3.9.



Slika 3.9. Efekat non breaking space HTML entiteta

Bez korišćenja HTML entiteta browser bi u situaciji, kao na slici 3.9, izvršio prelamanje teksta u novi red. S obzirom na to da se koristi *non breaking space* HTML entitet, do prelamanja ne dolazi, tako da tekst izlazi iz okvira browsera.

HTML entiteti mogu se koristiti i za kreiranje specijalnih karaktera. To ilustruje sledeći kod:

```
<p>Copyright &copy; LINKgroup</p>
```

Prikazani kod ima efekat kako na slici 3.10.

Copyright © LINKgroup

Slika 3.10. Specijalni Copyright karakter

Usmerenje teksta

HTML jezik omogućava definisanje teksta različitih pisanih sistema. Do sada smo unutar HTML dokumenata definisali tekst napisan engleskim jezikom, što je jedan od jezika pisanim sistemom koji poseduje usmerenje sleva nadesno. Jednostavno, tekst engleskog jezika piše se sa leva na desno, pa se tako može reći da poseduje usmerenje *left-to-right* (**LTR** skraćeno).

Pored LTR pisanih sistema, postoje i oni kod kojih se pisanje obavlja sa zdesna nalevo, *right-to-left* ili skraćeno **RTL**. HTML jezik omogućava definisanje usmerenja teksta i to korišćenjem atributa **dir**.

Atribut `dir` može se naći na bilo kojem HTML elementu, ali njegovo korišćenje najviše smisla ima na elementima kojima se markira tekst, a koji će biti prikazan na stranici. Na primer:

```
<p dir="ltr">
  This is some paragraph text.
</p>
```

Atribut `dir` može imati dve vrednosti:

- `ltr`
- `rtl`

Podrazumevana vrednost jeste `ltr`, pa u slučaju pisanja teksta *left-to-right* usmerenja nije obavezno definisati usmerenje. Ipak, ukoliko je unutar HTML dokumenta potrebno postaviti neki tekst *right-to-left* usmerenja, neophodno je iskoristiti `dir` atribut sa `rtl` vrednošću:

```
<p dir="rtl">
  عقلاً وهبوا وقد .والحقوق الكرامة في متساوين أحراراً الناس جميع يولد.
  الإخاء بروح بعضاً بعضهم يعامل أن وعليهم وضميراً
</p>
```


Veoma često se `dir` atribut definiše globalno, na nivou celog HTML dokumenta, naravno, ukoliko kompletan tekst dokumenta poseduje identično usmerenje. To se postiže definisanjem `dir` atributa na `html` elementu:

```
<html dir="ltr">
...
</html>
```

Jezik teksta

Autor HTML dokumenta poseduje mogućnost definisanja jezika koji se koriste za tekstualni sadržaj. To se postiže atributom **lang**, koji je, kao i prethodno opisani `dir`, jedan od globalnih atributa.

Primer upotrebe `lang` atributa:

```
<p lang="es">Este es el párrafo.</p>
```

Atribut `lang` moguće je definisati i globalno, na `html` elementu. U takvom slučaju definisani jezik odnosi se na ceo dokument:

```
<html lang="en">
...
</html>
```

Kao vrednost atributa `lang` definišu se ISO jezički kodovi. Neki od najznačajnijih prikazani su tabelom 3.3.

Jezik	ISO Code
Engleski	en
Francuski	fr
Nemački	de
Grčki	el
Japanski	ja
Rumunski	ro
Ruski	ru
Srpski	sr
Španski	es
Turski	tr
Ukrajinski	uk

Tabela 3.3. ISO jezički kodovi

Bitno je naglasiti da definisanje jezika tekstualnog sadržaja unutar HTML dokumenta ne proizvodi apsolutno nikakav vizuelni efekat. Drugim rečima, ukoliko atribut `lang` postavite na neki paragraf njegov sadržaj, ostaće identičan onom pre definisanja takvog atributa.

Definisanje jezika HTML dokumenta korišćenjem atributa `lang` namenjeno je programima koji obavljaju čitanje (parsiranje) sadržaja takvih dokumenata. Tako vrednost ovog atributa mogu koristiti browseri i crawleri, za bolje razumevanje tekstualnog sadržaja.

HTML komentari

Na kraju ove lekcije biće reči i o jednoj specifičnoj vrsti teksta koja se može naći unutar HTML dokumenata. Reč je o komentarima.

Komentari su pojam koji postoji u gotovo svim kompjuterskim jezicima. Oni omogućavaju autoru dokumenta definisanje određenih beleški unutar samog HTML dokumenta, koje pritom neće biti vidljive na samoj stranici nakon njenog parsiranja.

HTML komentari se formiraju korišćenjem oznaka koje su nalik na otvarajuće i zatvarajuće tagove. Početak HTML komentara označava se oznakom `<!--`, dok se završetak komentara obeležava oznakom `-->`.

Primer komentara unutar HTML dokumenta:

```
<!-- This is HTML comment. -->
```

HTML komentari su intenzivno korišćeni unutar primera koji se nalaze u prilogima lekcija ovoga programa. Oni su tamo korišćeni kako bi opisali različite jezičke elemente koji su korišćeni prilikom realizacije primera.

Rezime

- Osnova svakog HTML dokumenta jeste tekst koji se unutar dokumenta može pojaviti u različitim oblicima.
- HTML jezik poznaje 6 nivoa naslova, koji se kreiraju korišćenjem elemenata: `h1`, `h2`, `h3`, `h4`, `h5` i `h6`.
- U HTML dokumentima paragrafi se kreiraju postavljanjem teksta između otvarajućeg `<p>` i zatvarajućeg `</p>` taga.
- HTML jezik poznaje specijalne elemente za definisanje teksta korišćenjem različitih formatiranja.
- Ukoliko je potrebno preći u novi red, moguće je koristiti tag `
`.
- HTML ne tretira višestruke razmake, kao i prelaske u novi red.
- HTML entiteti omogućavaju da se na stranici ispišu rezervisani i specijalni karakteri.