

Technické otázky: error_log súbor mám plný záznamov typu "File does not exist: /home/users/userid/www.domena.sk/htdocs/robots.txt". Čo to znamená?

Roboti, ktorý prehľadáva web, indexujú vaše stránky a následne sprístupňujú zindexovaný obsah cez prehľadávač (ako napr. google.com alebo yahoo.com). Použitím súboru robots.txt v koreňovom adresári vášho webu, môžete zdefinovať adresáre, ktoré si neprajete aby boli zindexované. Napr., ak do súboru robots.txt napíšete nasledovné:

```
User-agent: *  
Disallow: /stat/
```

nemal by žiaden robot indexovať adresár /stat/ ani žiadne jeho podadresáre. Je možné obmedziť prístup aj pre explicitne zadaného robota, napr.:

```
User-agent: Google  
Disallow: /
```

vyjadruje, že si neprajete aby Google indexoval hociktorú stránku z vášho webu. Naopak, nasledovným obsahom robots.txt povolíte indexovať všetky stránky aj podadresáre vášho webu všetkým robotom:

```
User-agent: *  
Disallow:
```

Roboti sa snažia pri začatí indexovania vášho webu najprv prečítať súbor robots.txt. To spôsobuje, že ak súbor robots.txt vytvorený nemáte, tak si v súbore chybových hlásení nájdete chybovú hlášku typu:

```
File does not exist: /home/users/userid/www.domena.sk/htdocs/robots.  
txt
```

Vytvorením súboru robots.txt (so správnym obsahom) sa vyhnete chybovým hláseniam vo vašom error_log súbore. Treba však pripomenúť, že údaje zapísané v súbore robots.txt sú len vyjadrením vášho želania. Súbor robots.txt zohľadňujú len "slušné" roboty.

Existencia súboru robots.txt v žiadnom prípade neslúži na obmedzenie prístupu k časti webu.

Technické otázky: error_log súbor mám plný záznamov typu "File does not exist: /home/users/userid/www.domena.sk/htdocs/robots.txt". Čo to znamená?

Adresáre uvedené v súbore robots.txt sú naďalej voľné čítať, akékoľvek z internetu, pokiaľ nezavediete iné mechanizmy na obmedzenie prístupu (zaheslovanie alebo použitím .htaccess súboru).

Unikátne riešenie ID: #1023

Autor: Admin

Posledná aktualizácia: 2009-11-05 13:54