

Jak ograniczyć ruch robotów na[.]

09/19/2020 08:22:48

[FAQ Article Print](#)

Category:	Hosting ITL::Inne	Last update:	06/07/2017 12:50:41
State:	public (all)		
Language:	pl		

Keywords

transfer robots.txt

Symptom (public)

Na mojej stronie internetowej dużo ruchu generują nieznanne roboty internetowe.

Problem (public)

W internecie różne firmy wspierające marketing internetowy, zbierają informacje o stronach WWW, wykorzystując do tego roboty internetowe. Czasem ruch tych robotów mocno obciąża stronę, jednocześnie nie przynosząc właścicielowi strony żadnych korzyści.

Solution (public)

Można ograniczyć ruch robotów internetowych na stronie, odpowiednio przygotowując plik robots.txt. Plik należy umieścić w głównym katalogu każdego serwisu www, tak by był dostępny w następujący sposób:
`http://moja.domena.pl/robots.txt`

Przykładowa treść pliku robots.txt, blokująca wszystkie roboty oprócz robotów wyszukiwarek Google i Microsoft Bing:

```
User-agent: googlebot  
Disallow:
```

```
User-agent: bingbot  
Disallow:
```

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

Uwaga: roboty internetowe traktują zawartość pliku robots.txt jako wskazówki. Większość z robotów do tych wskazówek się stosuje. Jednak może się zdarzyć, że robot celowo i natarczywie ignoruje zawartość pliku robots.txt. Wówczas można takiego robota zablokować mniej przyjazny (dla robota) sposób: [1]w pliku .htaccess

[1] <https://support.intertele.pl/otrs/public.pl?Action=PublicFAQZoom;ItemID=21>