

Hobo - Development #20263

middleware robots.txt

24 novembre 2017 13:48 - Frédéric Péters

Statut:	Fermé	Début:	24 novembre 2017
Priorité:	Normal	Echéance:	
Assigné à:	Nicolas Roche	% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:		Planning:	Non
Patch proposed:	Oui		

Description

L'idée étant de pouvoir depuis hobo contrôler l'apparition d'un robots.txt avec Disallow: / comme contenu.

Ça pourrait être une interface où on choisirait dans des boutons radio "autoriser l'indexation", "ne pas autoriser l'indexation", "fichier personnalisé"; "autoriser" produirait un fichier vide, "ne pas autoriser" Disallow: /, et personnalisé laisserait taper n'importe quoi.

Ça serait enregistré en variable et ça se retrouverait du coup distribué partout.

Et le middleware attraperait ça pour servir quelque chose à l'URL /robots.txt.

Révisions associées

Révision 52ec23f0 - 14 avril 2020 09:07 - Nicolas Roche

seo: add middleware for /robots.txt url (#20263)

Révision e0a3f766 - 14 avril 2020 09:07 - Nicolas Roche

seo: add robots_txt views (#20263)

Révision 02e67cf0 - 14 avril 2020 09:07 - Frédéric Péters

seo: add settings for meta tags (#20263)

Historique

#1 - 06 avril 2020 09:31 - Frédéric Péters

- Description mis à jour

#2 - 06 avril 2020 18:44 - Nicolas Roche

- Assigné à mis à Nicolas Roche

#3 - 07 avril 2020 18:09 - Frédéric Péters

C'est d'un grand ridicule mais je dirais de plutôt taper ça dans un répertoire "seo", qui accueillera plus facilement d'autres trucs en rapport avec l'indexation par les moteurs de recherche qu'un répertoire nommé "robots", genre le paramétrage de meta_description et meta_keywords.

#4 - 07 avril 2020 19:18 - Nicolas Roche

- Fichier 0002-seo-add-robots-views-20263.patch ajouté

- Fichier 0001-seo-add-middleware-for-robots.txt-url-20263.patch ajouté

- Statut changé de Nouveau à Solution proposée

- Patch proposed changé de Non à Oui

(j'ai renommé, mais j'ai laissé quelques 'robots' là où ça me semblait spécifique)

des boutons radio ... et personnalisé laisserait taper n'importe quoi.

J'ai mis des boutons pour simplifier le rendu, et rester homogène avec FC et matomo.

contrôler l'apparition d'un robots.txt

Je retourne un fichier vide plutôt qu'une 404 si la variable (automatiquement créée) est vide.

#5 - 07 avril 2020 19:23 - Frédéric Péters

J'écrivais « "autoriser" produirait un fichier vide » donc oui pas de 404.

#6 - 13 avril 2020 10:14 - Frédéric Péters

- Fichier 0001-misc-add-custom-robots.txt-view-20263.patch ajouté

J'ai poussé une branche -review avec des modifications :

Sur le middleware, 0001, problème, il doit être utilisé par les autres modules, qui n'ont pas accès au modèle "Variable" de Hobo, il faut donc utiliser settings.TEMPLATE_VARS qui contiendra les infos distribuées par l'agent hobo.

À côté de ça, adaptations mineures, notamment pour que ça se trouve activé dans les modules, via debian/debian_config_common.py (sauf pour w.c.s., qui ne peut pas utiliser le middleware parce qu'on n'est pas dans la mécanique de multitenant classique, patch séparé nécessaire).

Sur l'UI, 0002, des modifications mineures (genre affichage en vert quand l'indexation est autorisée, réécriture du petit texte d'info pour être plus court, et utilisation unique du bouton pour modifier une valeur custom. (et la vue principale pourrait ne plus du tout être FormView).

Là-dessus, commit supplémentaire pour ajouter l'édition du contenu des <meta> "description" et "keywords".

Sur 0001 et 0002 j'ai laissé mes commits intermédiaires pour que tu puisses bien voir les changements, ils seront à squasher avant d'être poussés.

#7 - 13 avril 2020 16:01 - Nicolas Roche

- Fichier 0003-seo-add-settings-for-meta-tags-20263.patch ajouté

- Fichier 0002-seo-add-robots_txt-views-20263.patch ajouté

- Fichier 0001-seo-add-middleware-for-robots.txt-url-20263.patch ajouté

il doit être utilisé par les autres modules, qui n'ont pas accès au modèle "Variable" de Hobo

Oups. Et je réalise que ce middleware n'a pas besoin de s'appliquer à hobo lui même, qui n'a pas de visibilité publique (si besoin on pourra toujours lui ajouter un fichier robots.txt en dur).

A propos du "bouton" pour refuser l'indexation, j'avais un doute mais, <https://support.google.com/webmasters/answer/7424835#h17> me semble confirmer que si l'on place un fichier robots.txt qui refuse l'indexation, alors les balises meta ne seront pas lues.

you should not disallow the page in robots.txt, because the page must be crawled in order for the tag to be seen and obeyed.

J'ai intégré les modifications à 0001 et 0002 y compris aussi certaines de 0003 (*pardon*) afin de simplifier les patches.

#8 - 14 avril 2020 08:31 - Frédéric Péters

- Statut changé de Solution proposée à Solution validée

y compris aussi certaines de 0003 (*pardon*)

Il n'y a pas de soucis à ça mais c'est utile de préciser lesquelles, ça évite de devoir chercher.

#9 - 14 avril 2020 09:09 - Nicolas Roche

- Statut changé de Solution validée à Résolu (à déployer)

```
commit 02e67cf0fc33174bcd00247330b8b0f58b4f53ef
Author: Frédéric Péters <fpeters@entrouvert.com>
Date: Mon Apr 13 09:07:46 2020 +0200
```

```
seo: add settings for meta tags (#20263)
```

```
commit e0a3f76626ca629d6f1726b4249e2285b387edaa
Author: Nicolas ROCHE <nroche@entrouvert.com>
Date: Tue Apr 7 18:46:48 2020 +0200
```

```
seo: add robots_txt views (#20263)
```

```
commit 52ec23f04650753059acea0947fdaa540a045831
Author: Nicolas ROCHE <nroche@entrouvert.com>
Date: Tue Apr 7 18:38:16 2020 +0200
```

```
seo: add middleware for /robots.txt url (#20263)
```

#10 - 14 avril 2020 13:16 - Frédéric Péters

- Statut changé de Résolu (à déployer) à Solution déployée

#11 - 14 avril 2020 19:34 - Frédéric Péters

+

```
commit 4ed9079fd6ebdc6accf33100793c1b4ba869c38d
Author: Frédéric Péters <fpeters@entrouvert.com>
Date: Tue Apr 14 19:33:21 2020 +0200
```

```
build: ship seo templates
```

Fichiers

0001-seo-add-middleware-for-robots.txt-url-20263.patch	4,18 ko	07 avril 2020	Nicolas Roche
0002-seo-add-robots-views-20263.patch	16,1 ko	07 avril 2020	Nicolas Roche
0001-misc-add-custom-robots.txt-view-20263.patch	1,36 ko	13 avril 2020	Frédéric Péters
0003-seo-add-settings-for-meta-tags-20263.patch	4,92 ko	13 avril 2020	Nicolas Roche
0002-seo-add-robots_txt-views-20263.patch	15,4 ko	13 avril 2020	Nicolas Roche
0001-seo-add-middleware-for-robots.txt-url-20263.patch	5,66 ko	13 avril 2020	Nicolas Roche