

## Evitando conteúdo duplicado

### Transcrição

Temos nosso site criado, assim como o primeiro post, ajustado de acordo com todas as sugestões do **Yoast**. Ainda há alguns detalhes avançados e "escondidos" que muitas vezes passam despercebidos. Por conta disto, muitos sites acabam sendo penalizados ou não tendo um *ranking* tão bom.

Estamos na página inicial do blog, com um *template* padrão do Wordpress. No lado direito da página, há um link com todas as categorias ("Categorias") e, ao clicarmos nele, temos exatamente o mesmo conteúdo da home.

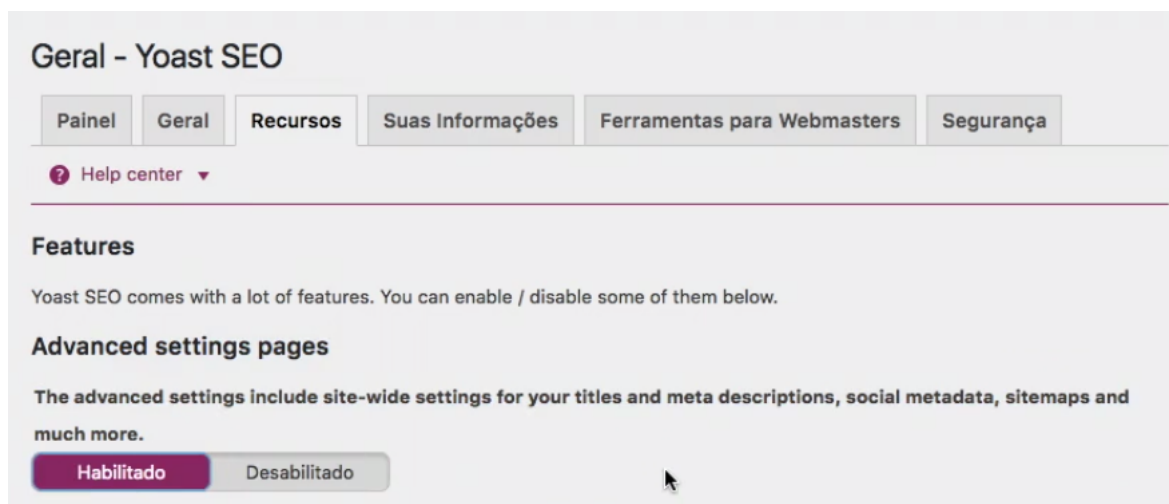
Se clicarmos em "Arquivos", o endereço para onde somos levados ( <http://desbravamundo.com/2017/02/> (<http://desbravamundo.com/2017/02/>) ) é outro, trata-se de outra página que será indexada pelo Google. O mesmo ocorre com o link das categorias ( <http://desbravamundo.com/category/sem-categoria> (<http://desbravamundo.com/category/sem-categoria>) ).

Sendo assim, os sites de busca indexarão o mesmo conteúdo várias vezes. No início do post, por exemplo, encontramos a informação de que ele foi criado em 8 de fevereiro de 2017, por mim, AdrianoAlmeida7. Clicando neste nome, o link também é outro ( <http://desbravamundo.com/author/adrianoalmeida7> (<http://desbravamundo.com/author/adrianoalmeida7>) ). Ou seja, o mesmo post aparece em quatro páginas com *slugs* distintos: na home, na página do autor, por categoria e por data.

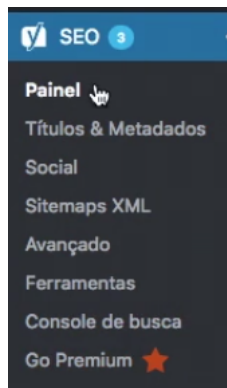
Vamos ao **Admin** do Wordpress para edição da postagem, à qual atribuiremos uma tag ("restaurantes", por exemplo), salvando-a em seguida. Veremos que acabamos gerando mais um link diferente para o mesmo post ( <http://desbravamundo.com/tag/restaurantes> (<http://desbravamundo.com/tag/restaurantes>) ). Isso precisa ser evitado.

Dentro do mesmo painel do Wordpress, vemos um item novo no menu, chamado "SEO", e é ele que administra todas as configurações do *plugin*, todas as especificidades do site, que envolvem não apenas um determinado post e, sim, o site todo.

Nesta página, portanto, clicaremos na aba "Recursos" e habilitaremos as configurações avançadas, clicando em "Salvar alterações" em seguida:



Feito isto, outras opções surgiram no menu:



"Títulos & Metadados > Taxonomias" lida com estes "links alternativos" que servem para acessar o mesmo conteúdo. O que precisamos fazer é decidir se queremos que a página específica seja indexada ou não.

Na taxonomia referente às categorias, alteraremos a parte de "Meta Robots" de " index " para " noindex ". Abrindo-se a página de categorias, clicando com o lado direito do mouse e selecionando a opção "Exibir código fonte da página", vê-se que houve inclusão do código HTML `<meta name = "robots" content="noindex,follow" />` para que a página de categorias não seja indexada. Faremos o mesmo para as opções de tags e formato.

Na aba de "Arquivo", repetiremos este procedimento, para as páginas classificadas por autor e data. Com o *Yoast*, pode-se desabilitar esta página por completo, ou apenas dizer se você quer que ela seja indexada ou não. A diferença entre estas duas abordagens é que, ao desabilitarmos a opção de exibir a página classificada por autor, ao clicarmos no nome do autor do post, somos redirecionados à home. O arquivo de autor deixa de existir.

Particularmente, não gosto de desabilitar a página, pois dependendo do template usado, o link para a página do autor possivelmente continuará existindo. Há pessoas que defendem o uso desta abordagem em blogs mantidos por apenas um autor, por exemplo. Neste caso, é preciso tomar o cuidado para escolher um layout (ou alterá-lo) para que não haja links para autores em lugar algum.

A fim de evitarmos este cenário, a recomendação é de simplesmente alterar para " noindex ". E a mesma situação vale para o arquivo de datas: deixaremos a página habilitada, porém com o " noindex " ativado. Desta forma, garantimos que não haja conteúdos duplicados dentro do site.