

# PHP - Créer un moteur de recherche

Juin 2014

- [Idée générale](#)
- [Création de la base de données](#)
- [Remplissage de la base de données](#)
- [Exploitation de la base de données](#)
- [Remarques](#)

## Idée générale

Le [moteur de recherche](#) ci-dessous

ne correspond qu'à une idée possible de moteur de recherche simple, ne gérant qu'un seul mot clé.

Le concept du fonctionnement de ce moteur est de créer une base de donnée contenant les mots clés

de chaque page du site stockés dans des balises spéciales appelées [méta tags](#). Un premier script permet donc de "scanner" l'ensemble des répertoires (et sous-répertoires) d'un site à la recherche de fichiers dont l'extension est *.htm* ou *.html*, puis de scruter leur contenu à la recherche des balises méta *keywords* (les mots-cés), *title* (le titre de la page), puis de stocker ces données, ainsi que l'emplacement du fichier sur le serveur, dans une base de données. Chaque fichier du site devant être indexé devra comporter les balises méta suivantes :

```
<head>
<meta name="ROBOTS" content="all">
<meta name="keywords" content="mot clé 1,mot cle 2,Mot clé 3, Mots clefs">
<meta name="title" content="titre de votre page">
<title>titre de votre page</title>
</head>
```

La balise *ROBOTS* permet de spécifier si le fichier doit être indexé ou non. Si vous ne voulez pas qu'une page de votre site soit indexée, il suffit alors de lui incorporer la balise méta suivante :

```
<head>
<meta name="ROBOTS" content="none">
</head>
```

Les balises précédentes ont l'avantage d'être utilisées aussi par les principaux moteurs du marché (Google,Yahoo!,Altavista,Voilà,...)

Dans un second temps, un script PHP permet d'exploiter la base de données ainsi créée en cherchant dans celle-ci le mot clé saisi par l'utilisateur par l'intermédiaire d'un formulaire.

## Création de la base de données

Le but est de créer une table comportant les champs nécessaires au stockage des enregistrements, c'est-à-dire contenant les champs suivants :

- **lien** : URL (absolue ou relative) de la page
- **keywords** : mots clés
- **titre** : titre de la page

Il suffit donc de créer un petit script PHP dont le seul but sera de créer cette table, que l'on appellera *search*

```
<html>

<head>

<title>Creation de la table</title>

</head>

<body>

<?php
$host = "Votre serveur de base de données";

$user = "Votre nom d'utilisateur";

$password = "Votre mot de passe";

$bdd = "Votre base de données sur le serveur";

mysql_connect($host, $user, $password) or die ("Connexion au serveur impossible");

// on choisit la bonne base
mysql_select_db($bdd) or die ("Connexion a la base impossible");

$query = "CREATE TABLE search (
    lien varchar(128) NOT NULL,
    keyword text,
    titre varchar(128),
    id INT(11),
    PRIMARY KEY (id)
)";

mysql_query($query) or die ("Erreur de modification de la table");

// on ferme la base
mysql_close();

?>

</body>

</html>
```

## Remplissage de la base de données

Le script suivant est le script principal du moteur, car c'est lui qui permet de parcourir

l'arborescence

du site à la recherche des fichiers HTML, de les scruter afin de déterminer s'ils doivent être ajoutés à la base et, le cas échéant, stocker les données nécessaires à leur exploitation.

Le script va dans un premier temps supprimer les anciens enregistrements de la table, puis lancer une procédure ayant pour rôle de parcourir l'arborescence à la recherche de fichier HTML.

Cette procédure est récursive, c'est-à-dire qu'elle s'appelle elle-même. Elle admet en paramètre le répertoire dans lequel elle doit rechercher des fichiers, ainsi que le chemin d'accès relatif à ce répertoire. La procédure examine chaque enregistrement du répertoire.

- S'il s'agit d'un répertoire (autre que le répertoire courant (.) ou parent (..)), alors la fonction examine le sous-répertoire
- S'il s'agit d'un fichier dont l'extension est *.htm* ou *.html*, la fonction récupère l'ensemble des méta

tags dans un tableau associatif

- Si le méta tag *ROBOTS* contient la valeur *NONE*, la fonction passe son chemin
- Sinon, la fonction récupère les méta tags relatifs aux mots clés et au titre et les stocke dans la base

```
<?php
echo "
<p>\n
<table BGCOLOR=\"#EFF2FB\" BORDER=\"0\"
    CELLSPACING=\"0\"
    CELLPADDING=\"1\"
    WIDTH=\"100%\">\n
<tr><td>\n
<a name=\"#index\"><h2>Indexation du site en cours</h2></a>\n
</td></tr>\n
</table>\n
<p>\n";

$host = "Le serveur de base de données";

$bdd = "Votre base de données";

/* Connexion avec MySQL */
mysql_connect($host,$user,$password) or die ("Impossible de se connecter
au serveur de base de donnees");

mysql_select_db($bdd) or die ("Impossible d'accéder à la base $bdd");

$query = "DELETE FROM search";

mysql_query($query) or die ("Erreur de modification de la table");
```

```
function ScanDir($Directory){
$MyDirectory = opendir($Directory);

while($Entry = readdir($MyDirectory)) {
echo "<br>entry= $Entry<br>";

echo "repertoire= $Directory<br>";

echo "chemin= $Directory/$Entry<br>";

if(is_dir($Entry)&& $Entry != "." && $Entry != "..") {
echo "<b><font color=\"red\">$Entry</font>
</b> est un repertoire<br>";

ScanDir("$Entry/$Directory");

}
else {
if (eregi(".htm",$Entry)) {
$MetaTags = get_meta_tags($Directory."/".$Entry);

if ($MetaTags["robots"] == "all") {
$MetaKey = $MetaTags["keywords"];

$MetaKey = strtoupper($MetaKey);

echo "Meta($Directory/$Entry): $MetaKey
\n";

$MetaTitre = $MetaTags["title"];

echo "Meta($Directory/$Entry): $MetaTitre
\n";

$query = "INSERT INTO search (lien,keyword,titre)
VALUES(\"$Directory/$Entry\", \"$MetaKey\", \"$MetaTitre\")";

$mysql_result = mysql_query($query) or die ("Erreur
de modification de la table par la requete \"$query\");

}
}
}
}
closedir($MyDirectory);
}
```

```
$open_basedir=".";

ScanDir(".");

mysql_close();

?>
```

## Exploitation de la base de données

Pour exploiter la base de données, il faut dans un premier temps créer un formulaire permettant à vos utilisateurs de saisir un mot clé à chercher. Voici un exemple de formulaire simple leur permettant d'effectuer leur recherche :

```
<form method="post" action="search.php3">

Entrez un mot clé:<br>

<input type="text" name="Mot" size="15">

<input type="submit" value="Rechercher" alt="Lancer la recherche!">

</form>
```

Voici le résultat de ce code HTML :

```
<form method="post" action="search.php3"> Entrez un mot clé: <input type="text" name="Mot"
size="15"> <input type="submit" value="Rechercher" alt="Lancer la recherche!"> </form>
```

Puis il s'agit de créer le script PHP (le dernier) recherchant l'ensemble des enregistrements de la base de données contenant la chaîne entrée par l'utilisateur. Ce script est simple, il effectue dans un premier temps une requête SQL sélectionnant les enregistrements contenant la chaîne. Puis il affiche le nombre d'enregistrements retournés, et une boucle while exploite ces enregistrements et les affiche les uns à la suite des autres.

```
<?php
$host = "Votre serveur de base de données";

$user = "Votre nom d'utilisateur";

$password = "Votre mot de passe";

$bdd = "Votre base de données sur le serveur";

mysql_connect($host, $user, $password) or die ("Connexion au serveur impossible");
```

```
// on choisit la bonne base
mysql_select_db($bdd) or die ("Connexion a la base impossible");

echo "
<html>

<head>

<title>Résultat de la recherche</title>

</head>

<body>";

if (($Mot == "") || ($Mot == "%")) {
// Si aucun mot clé n'a été saisi,
// le script demande à l'utilisateur
// de bien vouloir préciser un mot clé

echo "
Veuillez entrer un mot clé s'il vous plaît!
<p>";

}

else {
// On selectionne les enregistrements contenant le mot clé
// dans les keywords ou le titre
$query = "SELECT distinct count(lien) FROM search
WHERE keyword LIKE \"%$Mot%\"
OR titre LIKE \"%$Mot%\"
";

$result = mysql_query($query);

$row = mysql_fetch_row($result);

$Nombre = $row[0];

// Si aucun enregistrement n'est retourné,
// on affiche un message adéquat
if ($Nombre == "0") {
echo "
<h2>Aucun résultat ne correspond à votre recherche</h2>

<p>

";
}
```

```

}

// Sinon, on affiche le nombre d'enregistrements correspondant
// et les résultats eux-mêmes
else {
$query = "SELECT distinct lien,keyword,titre FROM search
WHERE keyword LIKE \"%$Mot%\"
OR titre LIKE \"%$Mot%\" ORDER by titre ASC";

$result = mysql_query($query);

// Si un seul enregistrement est trouvé, on affiche un message au singulier
if ($Nombre == "1") {
echo "
<a name=\"#resultat\"><h2>Résultat: Un article trouvé</h2></a>

<p>";

}
// Dans le cas contraire le message est au pluriel...
else {
echo "
<a name=\"#resultat\"><h2>Résultat: $Nombre articles trouvés</h2></a>

<p>";

}
while($row = mysql_fetch_row($result))
{
echo "
<p>\n
<b>$row[2]</b>\n
<br><a href=\"../$row[0]\">Visualiser l'article</a>\n
<p>&gt;\n
";
}
}

// on ferme la base
mysql_close();

?>

</body>

```



## Remarques

Le moteur présenté ci-dessus permet de faire une recherche basique, il est donc possible de lui ajouter des fonctionnalités, permettant par exemple de faire une recherche à partir de plusieurs mots-clés (comme les moteurs de recherche courants). D'autre part, les résultats sont affichées de manière basique les uns à la suite des autres. Il vous revient d'adapter la présentation des résultats à votre site, ainsi que celle du formulaire ou des messages d'avertissement...

Ce document intitulé « PHP - Créer un moteur de recherche » issu de **CommentCaMarche** ([www.commentcamarche.net](http://www.commentcamarche.net)) est mis à disposition sous les termes de la licence Creative Commons. Vous pouvez copier, modifier des copies de cette page, dans les conditions fixées par la licence, tant que cette note apparaît clairement.