

上海众森计算机科技有限公司



eZFind Extension 1.0 用户手册

--

提交至:
所有者:
作者: Michael Lee
编辑者: Michael Lee
版本: 1.0
日期: 2008-10-01

文件修改履历

版本	修改日	修改人	注释
1.0	2008/10/01	Michael Lee	初稿

目录

1 简介.....	4
1.1 目标读者.....	4
1.2 约定.....	4
1.3 更多资源.....	4
1.4 联络 eZ.....	4
1.5 联络众森.....	5
1.6 版权与商标.....	5
2 eZ Find 的工作原理.....	5
3 安装.....	5
3.1 启动 Solr.....	5
4 配置.....	6
4.1 建立多站点索引.....	6
4.2 搜索多站点.....	6
5 启用 eZ Find.....	6
5.1 更新搜索索引.....	6
6 定制.....	7
6.1 定制模板.....	7
7 使用 eZ Find.....	8
7.1 基本搜索.....	8
7.2 搜索选项.....	9
7.2.1 短语搜索.....	9
7.2.2 排除或包含关键字.....	10
7.2.3 搜索多关键字.....	10
7.3 高级搜索.....	10
7.3.1 内容类限制.....	11
7.3.2 其他限制.....	11
8 附录.....	11
8.1 词汇表.....	11
8.2 引用.....	12

1 简介

eZ Find 是一个为 eZ publish 开发的搜索扩展。相比 eZ Publish 默认搜索，它提供更多的搜索功能和更好的搜索结果。本文档是一份介绍如何安装，配置以及使用 eZ Find 的手册。

相比默认搜索，eZ Find 最大的优势是它支持相关度排序(relevancy ranking)和结果页面中的关键字高亮显示。搜索引擎用启发算法来分析信息结构，从而确定相关度。例如，对于某个关键字，内容对象标题中的匹配要比文本块中的匹配有更高的相关度。

另外，eZ Find 也改善了更新索引的算法。搜索结果由索引的一份副本驱动，因而，在搜索的同时，索引也可以得到持续的更新。

索引更新完成后，当前索引的副本也会被更新。搜索引擎也可以记住所有之前搜索所建立起的缓存结构，在更新索引的同时，这些缓存也会被更新。

1.1 目标读者

本手册介绍了如何安装，配置以及使用 eZ Find 搜索扩展，因而适用于系统管理员与最终用户。

1.2 约定

- 代码示例，函数，变量名，命令行使用灰色背景使用灰色背景，freemono 字体
- 在示例 URL 中，用您的域名替换 example.com

1.3 更多资源

可以使用以下资源获得更多支持：

- eZ Publish 文档：<http://www.ez.no/doc>
- eZ Publish 论坛：<http://ez.no/community/forum>
- eZ Publish 合作伙伴：联络 sales@ez.no 帮助您寻找一个 eZ Publish 合作伙伴
- 其他 eZ 解决方案：<http://ez.no/products/solutions>
- 培训及认证：联络 sales@ez.no 或访问 <http://ez.no/services/training>

1.4 联络 eZ

对于非技术性的问题，请通过以下联络方式联络 eZ：

- <http://ez.no/company/contact>
- info@ez.no

1.5 联络众森

您也可以通过以下方式联络我们获得中文帮助:

- <http://www.zerustech.com>
- info@zerustech.com
- michael.lee@zerustech.com

1.6 版权与商标

eZ Systems As 2007 版权所有。您有权在 GNU Free Document License 版本 1.2 以后所规定的条款范围内, 复制, 分发并/或修改本文档。本文档提及的其他的的产品与公司名称有可能是其所有者的注册商标。我们不对这些名称加注商标符号。

2 eZ Find 的工作原理

eZ find 是 eZ Publish 的一个搜索扩展, 它提供增强的搜索功能。它使用 Solr 搜索服务器。Solr 是一个基于 Lucene Java search library 的开源的企业级搜索服务器。eZ 默认的搜索功能在数据库中建立内容索引, Solr 则使用高度优化的文件来保存索引。eZ Find 使用 Solr 来处理索引与搜索。同时 eZ Publish 通过一组配置文件来控制 Solr 的行为。

当建立索引的时候, eZ Find 根据配置来完成索引建立, 查询时间分析。当用户搜索内容时, eZ Find 根据用户提交的搜索条件生成一个 Solr 查询。

搜有与 Solr 的通讯都是基于 Web service (使用 REST 模块)。

3 安装

eZ Find 需要安装 Java Runtime Environment (JRE)版本 1.5 以上。

您可以从 <http://java.sun.com> 下载 JRE。

JRE 安装之后, 从 http://ez.no/developer/contribs/applications/ez_find 下载 eZ Find。

把下载的文件保存到 eZ Publish 安装的根目录, 然后解压缩。

3.1 启动 Solr

进入 extension/ezfind, 用下面的命令行启动 Solr

```
java -jar start.jar
```

这会启动 Solr 应用。索引与搜索功能需要这一应用, 因而它必须启动。您可以访问 <http://localhost:8983/solr/admin/> 来测试 Solr 是否启动。

如果在搜索或建立索引的过程中, Java 进程异常退出, 您需要增加 Java 进程的内存配额。用以下的命令行来指定 Java 进程的内存配额:

```
java -Xms512M -Xm512M -jar start.jar
```

这会分配 512MB 的内存堆空间。

4 配置

4.1 建立多站点索引

eZ Fine 可以通过一个 Solr 运行实例为多个站点建立索引，这个选项是默认启用的。所有被索引的内容对象都与一个唯一的站点标识符关联。这个站点标识符可用于标识内容的原始安装位置。如果您希望一个站点中的内容被其他站点编入索引，在 `extension/ezfind/settings/ezfind.ini.append.php` 中启用 `IndexPubliclyAvailable` 选项。

```
IndexPubliclyAvailable=enabled
```

`IndexpubliclyAvailable` 选项只对匿名用户可以访问的内容有效。

`IndexPubliclyAvailalbe` 选项默认为启用状态

4.2 搜索多站点

如上所述，同时搜索多个站点的内容是可行的。如果需要搜索不同站点的内容，需要在 `extension/ezfind/settings/ezfind.ini.append.php` 中启用 `SearchOtherInstallations` 选项。

```
[SiteSettings]
```

```
SearchOtherInstallations=enabled
```

5 启用 eZ Find

在 `settings/override/site.ini.append.php` 中启用 `ezfind`

```
[ExtensionSettings]
```

```
...
```

```
ActiveExtensions[]=ezfind
```

如果与 `ezwebin`(Webiste interface)扩展同时使用，`ezfind` 必须在 `ezwebin` 之前启用。

5.1 更新搜索索引

运行 `updatesearchindexsolr.php` 将内容编入索引。

```
php extension/ezfind/bin/php/updatesearchindexsolr.php -s <admin  
siteaccess>
```

建立索引的速度通常为每秒 5-15 个对象。但是，如果您使用外部的过滤器转换二进制文件到文本文件，这可能需要更多时间。

6 定制

6.1 定制模板

定制 eZ Find 的模板与定制普通搜索模板的方法一样，但是 eZ Find 还提供一些额外的选项。eZ Find 返回一些附加的信息，包括：

- 相关度
- 语言信息
- 完整的 URL

默认的 eZ Publish 搜索返回 eZContentObjectTreeNode 对象列表。eZ Find 返回 eZFindResultTreeNode 对象列表。eZFindResultTreeNode 继承 eZContentObjectTreeNode 类，并增加了以下的新属性：

- is_local_installation: 布尔类型；标识结果是否来自于当前站点
- global_url_alias: 搜索结果的 URL，包括协议，域名
- published: 发布时间戳
- language_code: 搜索结果语言代码
- highlight: 包含搜索关键字的内容摘要
- score_percent: 相对值用于标识该结果与搜索条件的匹配程度

这些值可以在模板中作为变量使用。eZ Find 搜索与结果基于 eZ Publish 模板系统。默认的模板在 extension/ezfind/design/standard/templates/content。eZ Find 中也包含了一个定制模板的例子

(extension/ezfind/design/ezwebin/templates/node/view/ezfind_line.tpl)。\$node 是一个 eZFindResultTree 的对象。

Grenland in Telemark



Grenland in Telemark ... The **Grenland** region comprises four municipalities: Skien, Porsgrunn, Bamble, and Siljan. Skien **Grenland** is the hometown of the eZ Systems headquarters. ... and varied scenery has made Siljan a popular excursion area for the **Grenland** region.

74% - [/-hovik/ezfind/ezp3_trunk/index.p...land-in-Telemark/\(language\)/eng-GB](#) - 28/08/2007 2:39 pm

Illustration 1: line view 图例

ezfind_line.tpl 代码如下：

```
<div class="content-view-line">
<div class="class-article float-break">
<div class="attribute-title">
<h2 style="margin-top: 0.5em; margin-bottom: 0.25em"><a
href="{ $node.global_url_alias }">{ $node.name | wash }</a></h2>
</div>
{if is_set( $node.data_map.image )}
{if $node.data_map.image.has_content}
```

```

<div class="attribute-image">
{attribute_view_gui image_class=small href=concat( ' ',
$node.global_url_alias, ' ' ) attribute=$node.data_map.image}
</div>
{/if}
{/if}
<div class="attribute-short">
{$node.highlight}
</div>
<div class="attribute-short">
<i>{$node.score_percent}% - <a
href="{ $node.global_url_alias }">{$node.global_url_alias|shorten(70, '...',
'middle')|wash}</a> - {$node.object.published|l10n(shortdatetime)}</i>
</div>
</div>
</div>

```

7 使用 eZ Find

7.1 基本搜索

输入一些关键字，点击搜索按钮开始搜索。



Illustration 2: 搜索表单

搜索结束之后，搜索结果页面显示。结果页面列出搜索的结果并显示与每一条记录有关的附加信息。搜索结果总数显示在关键字文本框下面。搜索时间为完成该搜索所耗时。


Search

For more options try the [Advanced search](#)

Search for "Telemark" returned 8 matches

Search time: 49 msec

[Grenland in Telemark](#)



Grenland in **Telemark** ... of **Telemark** met the waterway from the sea. As one of the oldest towns in Norway and as the county capital of the **Telemark** Canal – the first of which is in the heart of town! Porsgrunn Porsgrunn grew from being a coastal


100% - [/~hovik/ezfind/ezp3_trunk/index.p...land-in-Telemark/\(language\)/eng-GB](#) - 28/08/2007 2:39 pm

[Grenland in Telemark](#)

Grenland in **Telemark** ... of **Telemark** met the waterway from the sea. As one of the oldest towns in Norway and as the county capital of the **Telemark** Canal – the first of which is in the heart of town! Porsgrunn Porsgrunn grew from being a coastal

64% - [http://localhost/~hovik/ezfind/ez...land-in-Telemark/\(language\)/eng-GB](#) - 30/08/2007 2:52 pm

[Hiking in Grenland](#)



expanses of the **Telemark** coast. The rolling Bamble countryside dominates to the west and north , incidentally, is part of the **Telemark** Canal. The path continues into picturesque forest areas and along

37% - [/~hovik/ezfind/ezp3_trunk/index.p...king-in-Grenland/\(language\)/eng-GB](#) - 28/08/2007 2:39 pm

[Case studies](#)

on their corporate multilingual website. Download PDF See reference **Telemark** University College **Telemark** University

2% - [http://ez.no/customers/case_studies/\(language\)/eng-GB](#) - 24/03/2003 9:34 am

[Hiking in Grenland](#)

expanses of the **Telemark** coast. The rolling Bamble countryside dominates to the west and north , incidentally, is part of the **Telemark** Canal. The path continues into picturesque forest areas and along

2% - [http://localhost/~hovik/ezfind/ez...king-in-Grenland/\(language\)/eng-GB](#) - 30/08/2007 2:52 pm

Illustration 3: 搜索结果页面

每一条记录包含一段关于内容对象的简介。结果也可能包含标题，图片，摘要，相关度，位置和创建时间。摘要包含一些包含关键字的段落。相关度显示记录与关键字的匹配程度。标题，图片与位置链接到内容对象。对于其他站点内的内容，不支持图片。

7.2 搜索选项

7.2.1 短语搜索

To search for a phrase, surround the phrase with quotation marks (for example, "Grenland in Telemark").

要搜索短语，用双引号环绕短语。(例如,"Grenland in Telemark")

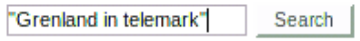


Illustration 4: 搜索短语

您可以同时按照短语与独立的关键字搜索。例如，输入：

"Grenland in Telemark" boat

搜索结果页面会返回所有文档同时包含短语"Grenland in Telemark"与单词"boat"

7.2.2 排除或包含关键字

您可以从搜索结果中排除含有特定关键字的文档。这可以通过在关键字前加 '-' 前缀来完成。

例如：

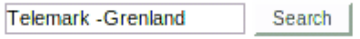


Illustration 5: 排除关键字

将会返回所有包含 "Telemark" 关键字但是不包含 "Grenland" 关键字的文档。同理，在关键字前加 '+' 前缀，意味着所有的文档必须包含这一关键字。

7.2.3 搜索多关键字

当指定多个关键字时，一篇文章必须匹配一定数量（依赖于关键字的数量）的关键字才会被包含在搜索结果中。例如，指定 2 个关键字的时候，一篇文章必须至少匹配 1 个关键字才能被包含在搜索结果。当指定 3 或 4 关键字时，至少匹配 2 个；超过 4 个关键字的时候，至少需要匹配 30% 的关键字。这些启发式的规则可以提供更精简，更有意义的搜索结果。

7.3 高级搜索

点击高级搜索链接启动高级搜索。

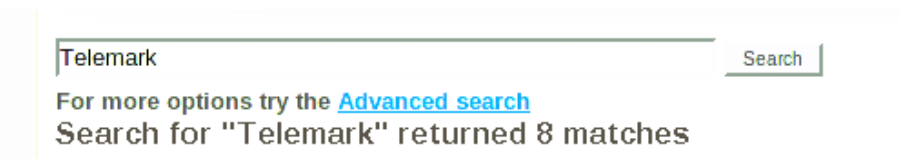


Illustration 6: 高级搜索

高级搜索界面允许您进一步定制搜索的选项。

Illustration 7: 高级搜索

"Search the exact phrase" (精确匹配短语) 文本框实际上提供了与短语搜索相同的功能。

7.3.1 内容类限制

如果需要限制搜索的对象类型，从"class"下拉框中选择类，然后点击"Update attributes" (更新属性)按钮来显示所选类的属性。

Illustration 8: 选择类

如果选择了类的属性，关键字必须在属性中存在，对象才会被包含到搜索结果。

7.3.2 其他限制

您还可以进一步用分区(section)或发布时间(published)限制搜索结果。

8 附录

8.1 词汇表

词汇	解释

附录 1: 词汇表

8.2 引用

引用	注释
上海众森计算机科技有限公司	上海众森计算机科技有限公司 (ZerusTech Ltd) 专注于视觉传达，内容管理系统与网站开发。我们是 eZ Publish 在亚洲的首个官方合作伙伴，也是其在亚洲的首个银牌官方合作伙伴 http://www.zerustech.com michael.lee@zerustech.com

附录 2: 引用