

Server-side search engines

Table des matières

1	Intro	1
2	Swish-e	1

1 Intro

- comparatif
- bechmark
- API

2 Swish-e

- Article
- Linux Mag
- installation

```
# apt-get install swish-e catdoc
```

```
doc
```

-

- `/var/www/narvali.conf`

```
IndexDir /var/www/narvali
IndexFile /var/www/narvali.index
```

```
FileFilter .doc      /usr/bin/catdoc "-s8859-1 -d8859-1 %p"
FileFilter .pdf      pdftotext      "%p -"
```

```
IndexContents TXT .html .pdf .doc
IndexOnly      .html .pdf .doc
```

```
# reduce multiple inflections of the same word (Eats, Eating... -> Eat)
FuzzyIndexingMode Stemming_fr
```

- Construction de l'index

```
// construction de l'index
$ swish-e -c /var/www/narvali.conf
```

- Recherche (cette configuration fonctionne sur les accents) :

```
$ swish-e -f /var/www/narvali.index -w 'repertoire'
# Number of hits: 7
```

```
$ swish-e -f /var/www/narvali.index -w 'répertoire'
# Number of hits: 27
```

- *API*

```
#!/usr/bin/perl -w
#search4.pl
use SWISH::API;
my ($max, $cnt) = (10,0);
my $index = "/var/www/narvali.index";
my $query = join(" ", @ARGV);
my $handle = SWISH::API->new($index);
my $results = $handle->Query($query);
while ( ($cnt++ < $max) &&
(my $res = $results->NextResult)) {
printf "%d '%s' %d\n",
    $res->Property("swishrank"),
    $res->Property("swishdocpath"),
    $res->Property("swishdocsize");
}
```