

Bloc-note 'Expressions Régulières'

Le 6 décembre 2007, à 0:38 par Ulhume...

Vu que ça me fatigue de rechercher à chaque fois les mêmes regex's, ouverture d'un petit bloc-note dédié ?

URL

Pour parser une URL d'un coup, celle-là marche plutôt bien à un détail près, il englobe la query et le hash (ce qu'il y a après un éventuel #). Pour éviter cela, si cela pose problème, il faut enlever le dernier ? et systématiquement ajouter un # à la fin de l'url avant parsing, s'il n'y est pas déjà évidemment. :

```
sub Parse {
  if ( shift [1]    /^((http[s]?|ftp):\/)?\/?([^:\/\s]+)((\/\w+)*\/)([\w\-\
\.]#[^#\s]+)(.*)?(#[\w-]+)?$/g ) {
    my ( $query $hash ) split [2] ( /#/ $7 )
    if ( defined [3] $hash ) {
      $hash ' ';
    }
    if (!defined $query) {
      $query=' '
    } else {
      $query=substr($query,1);
    }
    my %result = (
      protocol=>$2,
      host=>$3,
      path=>$4,
      file=>$6,
      query=>$query,
      hash=>$hash
    );
    return %result;
  }
  return {};
}
```

Liens:

[1] <http://perldoc.perl.org/functions/shift.html>

[2] <http://perldoc.perl.org/functions/split.html>

[3] <http://perldoc.perl.org/functions/defined.html>