

Dois Projetos de Linguística Computacional

1- Construir um **corpus** do Português

2 - Construir uma **ontologia** do Português

Construir um corpus do Português REPRESENTATIVO



Volumoso

Atualizado

Fonte de informação?

Imprensa Online

Jornais Portugueses online

Destak, Jornal de Noticias, Visão, etc.



Descarregar



Organizar

Fonte, data, autor, temas, etc.



Banco de Dados Textuais

Jornais Portugueses online

Destak, Jornal de Noticias, Visão, etc.



Reproduzir
leitura humana
como processo
informático

Organizar os
dados durante o
descarregamento

Banco de Dados Textuais



Editorial
ISABEL STILWELL
Amanhã é Dia do Pai e de Lua Cheia

LISBOA
17° C

PORTO
14° C

NEWSLETTER E-DESTAK

DESTAK.PT NA SUA HOMEPAGE

pesquisa no site OK

PORTO

Tosta mista perigosa para idoso

18 | 03 | 2011 09.06H

Um idoso que reside sozinho num apartamento da Rua Adolfo Casais Monteiro, no Porto, sofreu na quinta-feira à noite queimaduras de segundo e terceiro grau quando, por acidente, incendiou as próprias roupas, disse hoje à Lusa fonte dos bombeiros.

DESTAK/LUSA | DESTAK@DESTAK.PT

De acordo com a fonte, que citou testemunhos dos vizinhos, o [homem](#) estaria a tentar fazer uma tosta mista e não se terá apercebido de ter queimado a roupa.

"Já com a roupa a arder, o senhor começou a gritar e foram os vizinhos que lhe acudiram", explicou a mesma fonte, referindo que a habitação não sofreu danos.

Quando chegaram ao local, cerca das 21:00, os bombeiros limitaram-se a prestar os primeiros socorros ao idoso que foi transportado para o Hospital de Santo António.

"Foram os vizinhos que, com panos molhados e um balde, apagaram o fogo das roupas do idoso", acrescentou.

MAIS ARTIGOS DE PORTO

- Morreu Raúl de Brito, antigo vice-governador civil do Porto
- Retomadas as buscas para encontrar homem desaparecido no Douro
- Troca de tiros junto a discoteca na rua do Ouro
- 21 anos de prisão para homem que

COMENTAR
ENVIAR
IMPRIMIR

PARTILHAR



0



Uma recomendação. Regista-te para veres o que os teus amigos recomendam.

PUBLICIDADE

Opinião

J.L. PIO ABREU
Progressismo

JOSÉ LUÍS SEIXAS
Instabilidade política

LUISA CASTEL-BRANCO
As notícias do país que nós somos

JOÃO MALHEIRO
Bruxaria



A sua opinião

LIGA
Debate 7 de Março
Depois da derrota do Benfica em Braga, acha que o título de campeão nacional está definitivamente entregue ao FC Porto?
15 comentários

PORTO

Tosta mista perigosa para idoso

18 | 03 | 2011 09.06H

Um idoso que reside sozinho num apartamento da Rua Adolfo Casais Monteiro, no Porto, sofreu na quinta-feira à noite queimaduras de segundo e terceiro graus quando, por acidente, incendiou as próprias roupas, disse hoje à Lusa fonte dos bombeiros.

DESTAK/LUSA | DESTAK@DESTAK.PT

De acordo com a fonte, que citou testemunhos dos vizinhos, o homem estaria a tentar fazer uma tosta mista e não se terá apercebido de ter queimado a roupa.

“Já com a roupa a arder, o senhor começou a gritar e foram os vizinhos que lhe acudiram”, explicou a mesma fonte, referindo que a habitação não sofreu danos.

Quando chegaram ao local, cerca das 21:00, os bombeiros limitaram-se a prestar os primeiros socorros ao idoso que foi transportado para o Hospital de Santo António.

“Foram os vizinhos que, com panos molhados e um balde, apagaram o fogo das roupas do idoso”, acrescentou.

Porto

Tosta mista perigosa para idoso

18 | 03 | 2011 09.06H

Um idoso que reside sozinho num apartamento da Rua Adolfo Casais Monteiro, no Porto, sofreu na quinta-feira à noite queimaduras de segundo e terceiro graus quando, por acidente, incendiou as próprias roupas, disse hoje à Lusa fonte dos bombeiros.

Destak/Lusa | destak@destak.pt

De acordo com a fonte, que citou testemunhos dos vizinhos, o homem estaria a tentar fazer uma tosta mista e não se terá apercebido de ter queimado a roupa. “Já com a roupa a arder, o senhor começou a gritar e foram os vizinhos que lhe acudiram”, explicou a mesma fonte, referindo que a habitação não sofreu danos. Quando chegaram ao local, cerca das 21:00, os bombeiros limitaram-se a prestar os primeiros socorros ao idoso que foi transportado para o Hospital de Santo António. “Foram os vizinhos que, com panos molhados e um balde, apagaram o fogo das roupas do idoso”. acrescentou.

portotostamistaperigosaparaidoso18|03|201109.06humidosoqueresidesozinhonumaparta
mentodaruaadol focasaismonteiro, noporto, sofreunaquinta-feiraànoitequeimadrasdesegun
doeterceirograusquando, poracidente, incendiouasprópriasroupas, dissehojeàlusafontedosb
ombeiros. destak/lusa|destak@destak.pt deacordocomafonte, quecitoutestemunhosdosvizin
hos, ohomemestariaatentarfazerumatostamistaenãose teráapercebidodeterqueimadoaroup
a. “jácomaroupaarder, o senhorcomeçouagritareforamosvizinhosquelheacudiram”, explicou
amesmafonte, referindoqueahabitaçãonãosofreudanos. quandochegaramaalocal, cercadas
21:00, osbombeiroslimitaramseaprestaroprimeiro socorroaoidosoquefoittransportadopara
ohospitaldesantoantónio. “foramosvizinhosque, companosmolhadoseumbalde, apagaramof
ogodasroupasdoidoso”, acrescentou.

HTML & XML

HTML	XML
HyperText Markup Language	Extensible Markup Language
Para EXIBIR dados	Para TRANSPORTAR e ARMAZENAR dados
estilo	significado

** * **

PORTO

Tosta mista perigosa para idoso

18 | 03 | 2011 09.06H

Um idoso que reside sozinho num apartamento da Rua Adolfo Casais Monteiro, no Porto, sofreu na quinta-feira à noite queimaduras de segundo e terceiro graus quando, por acidente, incendiou as próprias roupas, disse hoje à Lusa fonte dos bombeiros.

DESTAK/LUSA | DESTAK@DESTAK

De acordo com a fonte, que citou testemunhos dos vizinhos, o homem estaria a tentar fazer uma tosta mista e não se terá apercebido de ter queimado a roupa.

"Já com a roupa a arder, o senhor começou a gritar e foram os vizinhos que lhe acudiram", explicou a mesma fonte, referindo que a habitação não sofreu danos.

Quando chegaram ao local, cerca das 21:00, os bombeiros limitaram-se a prestar os primeiros socorros ao idoso que foi transportado para o Hospital de Santo António.

"Foram os vizinhos que, com panos molhados e um balde, apagaram o fogo das roupas do idoso", acrescentou.

<i> * </i>

<pretitle> * </pretitle>

<title>

PORTO

Tosta mista perigosa para idoso

<date>

18 | 03 | 2011 09.06H

Um idoso que reside sozinho num apartamento da Rua Adolfo Casais Monteiro, no Porto, sofreu na quinta-feira à noite queimaduras de segundo e terceiro graus quando, por acidente, incendiou as próprias roupas, disse hoje à Lusa fonte dos bombeiros.

<posttitle>

DESTAK/LUSA | DESTAK@DESTAK

<author>

De acordo com a fonte, que citou testemunhos dos vizinhos, o homem estaria a tentar fazer uma tosta mista e não se terá apercebido de ter queimado a roupa.

"Já com a roupa a arder, o senhor começou a gritar e foram os vizinhos que lhe acudiram", explicou a mesma fonte, referindo que a habitação não sofreu danos.

Quando chegaram ao local, cerca das 21:00, os bombeiros limitaram-se a prestar os primeiros socorros ao idoso que foi transportado para o Hospital de Santo António.

"Foram os vizinhos que, com panos molhados e um balde, apagaram o fogo das roupas do idoso", acrescentou.

<text>

XML

DESTAK

```
<meta property="og:title" content="Porto: Tosta mixta perigosa para idoso"...>  
<title>Destak.pt | Tosta mixta perigosa para idoso</title>  
<div class="section">Porto</div><h2>Tosta mista perigosa para idoso</h2>
```

JORNAL DE NOTICIAS

```
<title>Camião carregado com vacas tombou na A7 - JN</title>  
<h1 id="NewsTitle">Camião carregado com vacas tombou na A7</h1>
```

EXPRESSO

```
<h1 class="artigoTit">PEC: Governo tem "total disponibilidade"  
para discutir com PSD</h1>
```


Expressões Regulares

Um método formal para especificar um padrão de texto

- Matches any character except a newline.
- () Groups a series of pattern elements to a single element.
- + Matches the preceding pattern element one or more times.
- ? Matches the preceding pattern element zero or one times.
- | Separates alternate possibilities.
- \b Matches a word boundary.
- ...

Document1 - Microsoft Word

File Edit View Insert Format Tools Table Window Help

100%

Open

Look in: My Documents

- My Recent Documents
- Desktop
- My Documents
- My Computer
- My Network Places

- Downloads
- Mis archivos PSP
- Movie Label Reports
- My Music
- My Pictures
- TopStyle 4
- wsmith5
- WYSIWYG Web Builder

File name:

Files of type: Word Documents (*.doc)

Open Cancel

Expressões Regulares

```
...  
<a href="/horoscopo/dia">horóscopo</a></li></ul></div>  
<div id="main"><div id="article" class="section_red"><div id="col1">  
<div class="section">Porto</div><h2>  
Tosta mista perigosa para idoso</h2><div class="date">  
18 | 03 | 2011 &nbsp; 09.06H</div><div class="lead">  
<p>Um idoso que reside sozinho num apartamento da Rua Adolfo  
Casais Monteiro, no Porto, sofreu na quinta-feira à noite queimaduras  
...
```

```
<h2>(.*?)</h2>
```

Tosta mista perigosa para idoso

```
<div class="section">(.*?)</div>
```

Porto



GOVERNO PORTUGUÊS CAI APÓS CHUMBO DO PEC 4

Sócrates demite-se, Portugal em crise política

- Democracia existe para resolver problemas como estes, diz Paulo Portas
- Jerónimo de Sousa defende que "se devolva a palavra" aos portugueses
- O país atingiu o "limiar da resistência", diz Francisco Louçã
- PSD quer eleições para novo Governo que aplique "estratégia verdadeiramente nacional"
- Filme dos acontecimentos: como um Governo cai em 12 dias
- Cavaco recebe partidos sexta-feira, primeira etapa da crise que deverá levar o país a eleições

PUBLICIDADE

24 Março 2011



Sócrates acusa Oposição de "sofreguidão pelo poder" e reafirma que vai a eleições

GINA PEREIRA

O primeiro-ministro acusou, esta quarta-feira à noite, os partidos da Oposição de "sofreguidão pelo poder" e de "estreiteza de vistas" ao terem-se unido, durante a tarde, no Parlamento para chumbar o Programa de Estabilidade e Crescimento com que o Governo português já se tinha comprometido com

Quinta-feira, 24 Mar

FARMÁCIAS: Lisboa - Farmácia Alameda | TRÂNSITO: Porto - Porto - Rua do Carval

Faça aqui o seu login | registo



Pesquisa Google

pesquisar

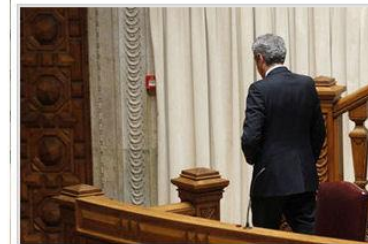
Convite aos Leitores: Deixe aqui a sua Opinião

A a Z Iniciativas Ricardo Araújo Pereira António Lobo Antunes

Japão: Bolsa encerra a perder 0,15 por cento

Parlamento

Sócrates pediu demissão e defende eleições



Últimas

7:20 Medidas do PEC para 2011 podem entrar em vigor mesmo com Governo demissionário

7:15 Trabalho: Mulheres participam mais, mas ganham menos 18% que os homens

6:41 Japão: Bolsa encerra a perder 0,15 por cento

6:29 Líbia: Tiros da defesa antiaérea, explosões em Trípoli

5:44 Japão/Sismo: Pelo menos 9.700 mortos e 16.500 desaparecidos -- Novo balanço

5:21 Nuclear/Japão: Normalizado nível de radioatividade na água de Tóquio

NOTÍCIAS

FUNDADO EM
29 de Dezembro de 1864

pesquisar...

DN Mobile | Iniciativas DN

APÓS CHUMBO DO PEC

Sócrates pede demissão e abre caminho a pedido de ajuda

Hoje
32 comentários



GRANDE INVESTIGAÇÃO DN



Uma viagem ao interior do Sporting

Hoje
Comentar

A dois dias das eleições para a presidência do Sporting, o DN faz uma viagem ao interior do Sporting e a todas as suas empresas associadas.

LEIA HOJE
IMPRESSO e-PAPER

ESPECIAIS



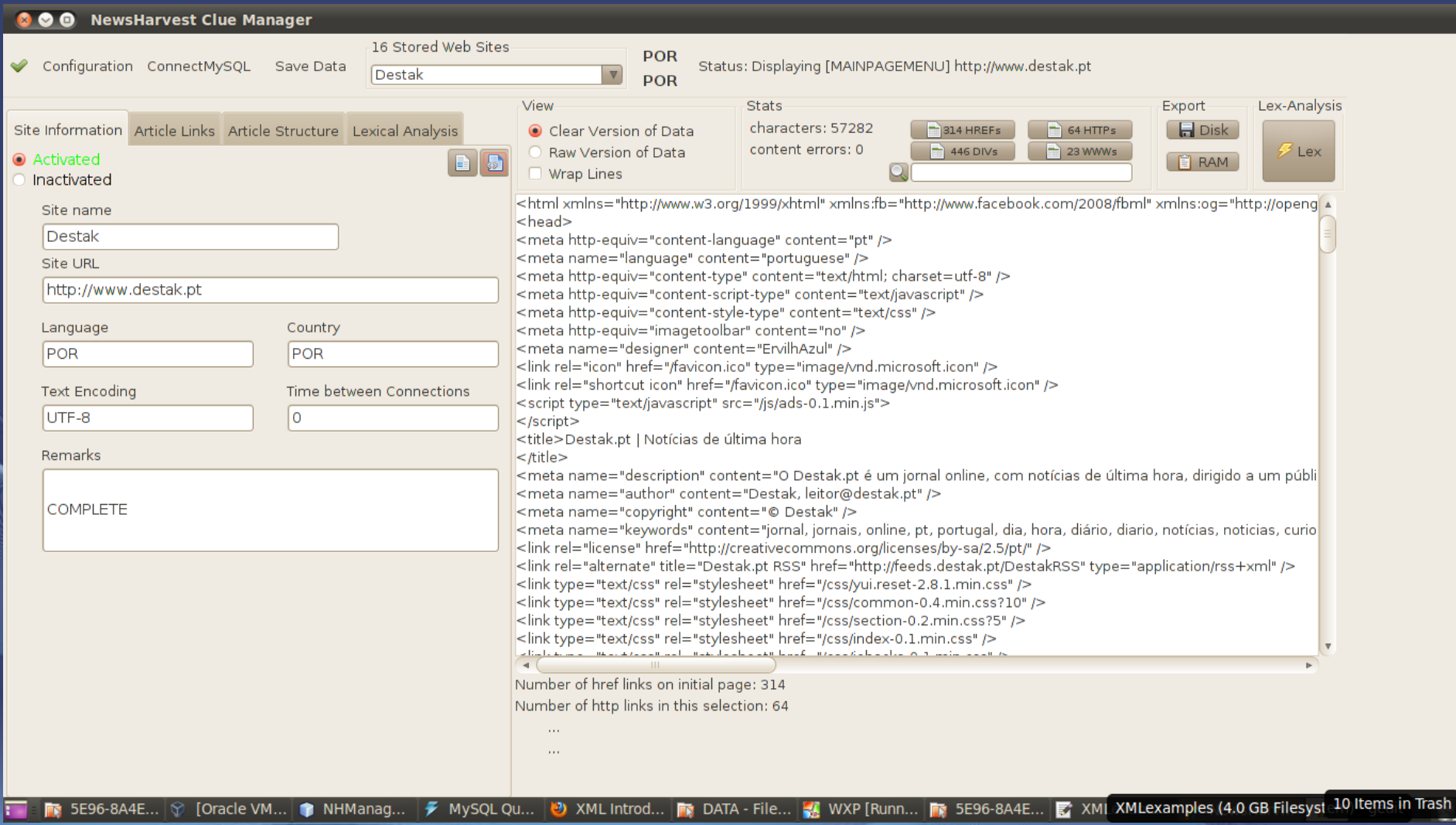
EDITORIAL

Portugal

Hoje

Facto cons demissão

Margem-se o momento para a história: eram exactamente 20h20



NewsHarvest Clue Manager

Configuration ConnectMySQL Save Data

16 Stored Web Sites

Destak

POR
POR

Status: Displaying [MAINPAGEMENU] http://www.destak.pt

Site Information Article Links Article Structure Lexical Analysis

Activated
Inactivated

Site name

Destak

Site URL

http://www.destak.pt

Language

POR

Country

POR

Text Encoding

UTF-8

Time between Connections

0

Remarks

COMPLETE

View

- Clear Version of Data
- Raw Version of Data
- Wrap Lines

Stats

characters: 57282
content errors: 0

314 HREFs

64 HTTPs

446 DIVs

23 WWWs

Export

Disk

RAM

Lex-Analysis

Lex

```
<html xmlns="http://www.w3.org/1999/xhtml" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://openg
<head>
<meta http-equiv="content-language" content="pt" />
<meta name="language" content="portuguese" />
<meta http-equiv="content-type" content="text/html; charset=utf-8" />
<meta http-equiv="content-script-type" content="text/javascript" />
<meta http-equiv="content-style-type" content="text/css" />
<meta http-equiv="imagetoolbar" content="no" />
<meta name="designer" content="ErvilhAzul" />
<link rel="icon" href="/favicon.ico" type="image/vnd.microsoft.icon" />
<link rel="shortcut icon" href="/favicon.ico" type="image/vnd.microsoft.icon" />
<script type="text/javascript" src="/js/ads-0.1.min.js">
</script>
<title>Destak.pt | Notícias de última hora
</title>
<meta name="description" content="O Destak.pt é um jornal online, com notícias de última hora, dirigido a um públi
<meta name="author" content="Destak, leitor@destak.pt" />
<meta name="copyright" content="© Destak" />
<meta name="keywords" content="jornal, jornais, online, pt, portugal, dia, hora, diário, diario, notícias, noticias, curio
<link rel="license" href="http://creativecommons.org/licenses/by-sa/2.5/pt/" />
<link rel="alternate" title="Destak.pt RSS" href="http://feeds.destak.pt/DestakRSS" type="application/rss+xml" />
<link type="text/css" rel="stylesheet" href="/css/yui.reset-2.8.1.min.css" />
<link type="text/css" rel="stylesheet" href="/css/common-0.4.min.css?10" />
<link type="text/css" rel="stylesheet" href="/css/section-0.2.min.css?5" />
<link type="text/css" rel="stylesheet" href="/css/index-0.1.min.css" />
<link type="text/css" rel="stylesheet" href="/css/index-0.1.min.css" />
```

Number of href links on initial page: 314
Number of http links in this selection: 64

...
...

Isolation tag

<div id="menu">(.*?)<div id="main">

Exclusion tag

Extraction tag

href="(.*?)>(.*?)

Correct order <Title-URL>

Add URL basename

Exclusion

ados|1@s*pagina|passatempos|horoscopo)t.*?&

View

Clear Version of Data

Raw Version of Data

Wrap Lines

Stats

characters: 57067

content errors: 0

314 HREFs

64 HTTPs

446 DIVs

23 WWWs

Export

Disk

RAM

Lex-Analysis

Lex

```
<html xmlns="http://www.w3.org/1999/xhtml" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://openg
<head>
<meta http-equiv="content-language" content="pt" />
<meta name="language" content="portuguese" />
<meta http-equiv="content-type" content="text/html; charset=utf-8" />
<meta http-equiv="content-script-type" content="text/javascript" />
<meta http-equiv="content-style-type" content="text/css" />
<meta http-equiv="imagetoolbar" content="no" />
<meta name="designer" content="ErvilhAzul" />
<link rel="icon" href="/favicon.ico" type="image/vnd.microsoft.icon" />
<link rel="shortcut icon" href="/favicon.ico" type="image/vnd.microsoft.icon" />
<script type="text/javascript" src="/js/ads-0.1.min.js">
</script>
<title>Destak.pt | Notícias de última hora
</title>
<meta name="description" content="O Destak.pt é um jornal online, com notícias de última hora, dirigido a um públi
<meta name="author" content="Destak, leitor@destak.pt" />
<meta name="copyright" content="© Destak" />
<meta name="keywords" content="jornal, jornais, online, pt, portugal, dia, hora, diário, diario, notícias, noticias, curio
<link rel="license" href="http://creativecommons.org/licenses/by-sa/2.5/pt/" />
<link rel="alternate" title="Destak.pt RSS" href="http://feeds.destak.pt/DestakRSS" type="application/rss+xml" />
<link type="text/css" rel="stylesheet" href="/css/yui.reset-2.8.1.min.css" />
<link type="text/css" rel="stylesheet" href="/css/common-0.4.min.css?10" />
<link type="text/css" rel="stylesheet" href="/css/section-0.2.min.css?5" />
<link type="text/css" rel="stylesheet" href="/css/index-0.1.min.css" />
<link type="text/css" rel="stylesheet" href="/css/destak-0.2.min.css" />
```

Number of href links on initial page: 314

Number of http links in this selection: 64

...

...

Isolation tag
 <div id="menu">(.*?)<div id="main">

Exclusion tag

Extraction tag
 href="(.*?)\">(.*?)

Correct order <Title-URL>

Add URL basename

Exclusion
 ados|1ª\s*página|passatempos|horóscopo)\t.*?\$

View

Clear Version of Data
 Raw Version of Data
 Wrap Lines

Stats

characters: 1386
content errors: 0

Export

Lex-Analysis

```
<li class="section_red">
<a class="active" href="/">1ª página
</a>
</li>
<li class="section_red">
<a href="/seccao/lisboa">lisboa
</a>
</li>
<li class="section_red">
<a href="/seccao/porto">porto
</a>
</li>
<li class="section_red">
<a href="/seccao/actualidade">actualidade
</a>
</li>
<li class="section_gold">
<a href="/seccao/globo">globo
</a>
</li>
<li class="section_yellow">
<a href="/seccao/desporto">desporto
</a>
</li>
<li class="section_purple">
<a href="/seccao/fama-tv">fama & amp; tv
```

Number of href links on initial page: 314
Number of href links in this selection: 21
...
...

Isolation tag

Exclusion tag

Extraction tag

Correct order <Title-URL>

Add URL basename

Exclusion

View

Clear Version of Data

Raw Version of Data

Wrap Lines

Stats

characters: 954

content errors: 0

0 HREFs 21 HTTPs

0 DIVs 20 WWWs

Export

Disk

RAM

Lex-Analysis

Lex

1ª página	http://www.destak.pt/
lisboa	http://www.destak.pt/seccao/lisboa
porto	http://www.destak.pt/seccao/porto
actualidade	http://www.destak.pt/seccao/actualidade
globo	http://www.destak.pt/seccao/globo
desporto	http://www.destak.pt/seccao/desporto
fama & tv	http://www.destak.pt/seccao/fama-tv
arte & lazer	http://www.destak.pt/seccao/arte-lazer
passatempos	http://www.destak.pt/passatempos
classificados	http://classificados.destak.pt
tecnologia	http://www.destak.pt/seccao/tecnologia
saúde	http://www.destak.pt/seccao/saude
mulher	http://www.destak.pt/seccao/mulher
fugas	http://www.destak.pt/seccao/fugas
automóvel	http://www.destak.pt/seccao/automovel
cinema	http://www.destak.pt/seccao/cinema
música	http://www.destak.pt/seccao/musica
emprego	http://www.destak.pt/seccao/emprego
marcas e empresas	http://www.destak.pt/seccao/marcas-empresas
imobiliário	http://www.destak.pt/seccao/imobiliario
horóscopo	http://www.destak.pt/horoscopo/dia

Number of href links on initial page: 314

Number of http links in this selection: 21

...

...

Isolation tag

Exclusion tag

Extraction tag

Correct order <Title-URL>

Add URL basename

Exclusion

View

Clear Version of Data

Raw Version of Data

Wrap Lines

Stats

characters: 787

content errors: 0

0 HREFs 17 HTTPs

0 DIVs 17 WWWs

Export

Disk

RAM

Lex-Analysis

Lex

lisboa	http://www.destak.pt/seccao/lisboa
porto	http://www.destak.pt/seccao/porto
actualidade	http://www.destak.pt/seccao/actualidade
globo	http://www.destak.pt/seccao/globo
desporto	http://www.destak.pt/seccao/desporto
fama & tv	http://www.destak.pt/seccao/fama-tv
arte & lazer	http://www.destak.pt/seccao/arte-lazer
tecnologia	http://www.destak.pt/seccao/tecnologia
saúde	http://www.destak.pt/seccao/saude
mulher	http://www.destak.pt/seccao/mulher
fugas	http://www.destak.pt/seccao/fugas
automóvel	http://www.destak.pt/seccao/automovel
cinema	http://www.destak.pt/seccao/cinema
música	http://www.destak.pt/seccao/musica
emprego	http://www.destak.pt/seccao/emprego
marcas e empresas	http://www.destak.pt/seccao/marcas-empresas
imobiliário	http://www.destak.pt/seccao/imobiliario

Number of href links on initial page: 314

Number of http links in this selection: 17

0 duplicate links were removed

...

NewsHarvest Clue Manager

Configuration ConnectMySQL Save Data 16 Stored Web Sites Destak POR POR Status: Displaying [MAINPAGEARTinclu] http://www.destak.pt

Site Information Article Links Article Structure Lexical Analysis

Main Page Menu Main Page Articles Topic Page Articles

Isolation tag `<div id="articles"(.*?)<div id="horoscope"`

Exclusion tag `geTracker\._trackPageview\(\this.href\);|`

Extraction tag `(.*?)`

Correct order <Title-URL>

Add URL basename

Exclusion `lerias|edicoes/.*/emprego|edicao-dia|opinioa)$`

Inclusion `?m)^(.*?)\(\http://www.destak.pt/.*/[0-9]{5,}.*/\n`

View:

- Clear Version of Data
- Raw Version of Data
- Wrap Lines

Stats:

- characters: 12529
- content errors: 0
- 0 HREFs
- 83 HTTPs
- 0 DIVs
- 83 WWWs

Export:

- Disk
- RAM

Lex-Analysis:

- Lex

Sócrates demite-se, Portugal em crise política
<http://www.destak.pt/artigo/90931-socrates-demite-se-portugal-em-crise-politica>

Democracia existe para resolver problemas como estes, diz Paulo Portas
<http://www.destak.pt/artigo/90942-democracia-existe-para-resolver-problemas-como-estes-diz-paulo-portas>

Jerónimo de Sousa defende que "se devolva a palavra" aos portugueses
<http://www.destak.pt/artigo/90941-jeronimo-de-sousa-defende-que-se-devolva-a-palavra-aos-portugueses>

O país atingiu o "limiar da resistência", diz Francisco Louçã
<http://www.destak.pt/artigo/90940-o-pais-atingiu-o-limiar-da-resistencia-diz-francisco-louca>

PSD quer eleições para novo Governo que aplique "estratégia verdadeiramente nacional"
<http://www.destak.pt/artigo/90939-psd-quer-eleicoes-para-novo-governo-que-aplique-estrategia-verdadeiramente-nac>

Filme dos acontecimentos: como um Governo cai em 12 dias
<http://www.destak.pt/artigo/90937-filme-dos-acontecimentos-como-um-governo-cai-em-12-dias>

Cavaco recebe partidos sexta-feira, primeira etapa da crise que deverá levar o país a eleições
<http://www.destak.pt/artigo/90935-cavaco-recebe-partidos-sexta-feira-primeira-etapa-da-crise-que-devera-levar-o-pai>

Parlamento chumba as novas medidas de austeridade do PEC. Sócrates fala dentro de minutos
<http://www.destak.pt/artigo/90929-parlamento-chumba-as-novas-medidas-de-austeridade-do-pec-socrates-fala-dentr>

Primeiro-ministro e Presidente reunidos 20 minutos em Belém
<http://www.destak.pt/artigo/90928-primeiro-ministro-e-presidente-reunidos-20-minutos-em-belém>

Number of href links on initial page: 314
 Number of http links in this selection: 83
 0 duplicate links were removed
 0 topic links were removed

5E96-8A4E... [Oracle VM... NHManag... MySQL Qu... XML Introd... DATA - File... WXP [Runn... 5E96-8A4E... XMLexam... NewsHarv...

Site Information Article Links Article Structure Lexical Analysis

URL:

Redirection Notice:

Isolation tag

Exclusion tag

URL topic clue

PRE-Title tag

Title tag

POST-Title tag

Date tag Date pattern

Author tag

Text tag

View

Clear Version of Data

Raw Version of Data

Wrap Lines

Stats

characters: 24187

content errors: 0

124 HREFs 61 HTTPs

150 DIVs 26 WWWs

Export

Disk

RAM

Lex-Analysis

Lex

```
<html xmlns="http://www.w3.org/1999/xhtml" xmlns:fb="http://www.facebook.com/2008/fbml" xmlns:og="http://openg
<head>
<meta http-equiv="content-language" content="pt" />
<meta name="language" content="portuguese" />
<meta http-equiv="content-type" content="text/html; charset=utf-8" />
<meta http-equiv="content-script-type" content="text/javascript" />
<meta http-equiv="content-style-type" content="text/css" />
<meta http-equiv="imagetoolbar" content="no" />
<meta name="designer" content="ErvilhAzul" />
<link rel="icon" href="/favicon.ico" type="image/vnd.microsoft.icon" />
<link rel="shortcut icon" href="/favicon.ico" type="image/vnd.microsoft.icon" />
<script type="text/javascript" src="/js/ads-0.1.min.js">
</script>
<meta property="og:type" content="article"/>
<meta name="created" content="2011-03-18" />
<meta property="og:title" content="Porto: Tosta mista perigosa para idoso" />
<link rel="canonical" href="http://www.destak.pt/artigo/90421" />
<title>Destak.pt | Tosta mista perigosa para idoso
</title>
<meta name="description" content="Um idoso que reside sozinho num apartamento da Rua Adolfo Casais Monteiro
<meta name="author" content="Destak, leitor@destak.pt" />
<meta name="copyright" content="© Destak" />
<meta name="keywords" content="jornal, jornais, online, pt, portugal, dia, hora, diário, diario, noticias, noticias, curio
<link rel="license" href="http://creativecommons.org/licenses/by-sa/2.5/pt/" />
<link rel="alternate" title="Destak.pt RSS" href="http://feeds.destak.pt/DestakRSS" type="application/rss+xml" />
<link type="text/html" href="http://www.destak.pt/artigo/90421" />
</head>
<body>
<div id="main">
<h3 class="section">
<div class="lead">
<div class="text">
</div>
</div>
</div>
</div>
</body>
</html>
```

Number of href links on initial page: 124

Number of http links in this selection: 61

Site Information Article Links Article Structure Lexical Analysis

URL:

Redirection Notice:

Isolation tag

Exclusion tag

URL topic clue

PRE-Title tag

Title tag

POST-Title tag

Date tag Date pattern

Author tag

Text tag

View

Clear Version of Data

Raw Version of Data

Wrap Lines

Stats

characters: 2968

content errors: 0

11 HREFs

0 HTTPs

16 DIVs

0 WWWs

Export

Disk

RAM

Lex-Analysis

Lex

```
<div id="article" class="section_red">
<div id="col1">
<div class="section">Porto
</div>
<h2>Tosta mista perigosa para idoso
</h2>
<div class="date">18 | 03 | 2011 09.06H
</div>
<div class="lead">
<p>Um idoso que reside sozinho num apartamento da Rua Adolfo Casais Monteiro, no Porto, sofreu na quinta-feira à noite queimaduras de segundo e terceiro grau quando, por acidente, incendiou as próprias roupas, disse hoje à Lusa fonte dos bombeiros.
</p>
</div>
<div id="article_author">Destak/Lusa |
<a href="mailto:destak@destak.pt">destak@destak.pt
</a>
<br />
</div>
<div id="HOTWordsTxt" class="text">
<p>De acordo com a fonte, que citou testemunhos dos vizinhos, o homem estaria a tentar fazer uma tosta mista e não se terá apercebido de ter queimado a roupa.
<p>"Já com a roupa a arder, o senhor começou a gritar e foram os vizinhos que lhe acudiram", explicou a mesma fonte, referindo que a habitação não sofreu danos.
<p>Quando chegaram ao local, cerca das 21:00, os bombeiros limitaram-se a prestar os primeiros socorros ao idoso que foi transportado para o Hospital de Santo António.
```

Number of href links on initial page: 124

Number of http links in this selection: 61

...

...

URL:

Redirection Notice:

Isolation tag

Exclusion tag

URL topic clue

PRE-Title tag

Title tag

POST-Title tag

Date tag Date pattern

Author tag

Text tag

View
 Clear Version of Data
 Raw Version of Data
 Wrap Lines

Stats
characters: 1530
content errors: 0
0 HREFs
0 HTTPs
16 DIVs
0 WWWs

Export
Disk
RAM

Lex-Analysis
Lex

Porto

Number of href links on initial page: 124
Number of http links in this selection: 61
...
...

URL:

Redirection Notice:

Isolation tag

Exclusion tag

URL topic clue

PRE-Title tag

Title tag

POST-Title tag

Date tag Date pattern

Author tag

Text tag

View
 Clear Version of Data
 Raw Version of Data
 Wrap Lines

Stats
characters: 1530
content errors: 0
0 HREFs 0 HTTPs
16 DIVs 0 WWWs

Export
Disk
RAM

Lex-Analysis
Lex

Tosta mista perigosa para idosos

Number of href links on initial page: 124
Number of http links in this selection: 61
...
...

URL: http://www.destak.pt/artigo/90421

Redirection Notice:

Isolation tag <div id="main">(.*)<h3 class="section">

Exclusion tag |ais artigos de .*?|<a href.*?>.*?<\/s*\/a*s*?>

URL topic clue

PRE-Title tag <div class="section">(.*)<\/div>

Title tag <h2>(.*)<\/h2>

POST-Title tag <div class="lead">(.*)<\/div>

Date tag Date pattern s="date">(.*)<\/div> | MM | yyyy HH.mm'H'

Author tag "article_author">(.*)<\/>

Text tag class="text">(.*)<\/div>

View Clear Version of Data Raw Version of Data Wrap Lines

Stats characters: 1530 content errors: 0 0 HREFs 0 HTTPs 16 DIVs 0 WWWs

Export Lex-Disk RAM

<p>Um idoso que reside sozinho num apartamento da Rua Adolfo Casais Monteiro, no Porto, sofreu na quinta-feira à noite queimaduras de segundo e terceiro graus quando, por acidente, incendiou as próprias roupas, disse hoje à Lusa fonte dos bombeiros.<\/p>

Number of href links on initial page: 124 Number of http links in this selection: 61

URL:

Redirection Notice:

Isolation tag

Exclusion tag

URL topic clue

PRE-Title tag

Title tag

POST-Title tag

Date tag Date pattern

Author tag

Text tag

View
 Clear Version of Data
 Raw Version of Data
 Wrap Lines

Stats
characters: 1530
content errors: 0
0 HREFs
0 HTTPs
16 DIVs
0 WWWs

Export
Disk
RAM

Lex-Analysis
Lex

18 | 03 | 2011 09.06H

Number of href links on initial page: 124
Number of http links in this selection: 61
...
...

URL:

Redirection Notice:

Isolation tag

Exclusion tag

URL topic clue

PRE-Title tag

Title tag

POST-Title tag

Date tag Date pattern
 |

Author tag

Text tag

View
 Clear Version of Data
 Raw Version of Data
 Wrap Lines

Stats
characters: 1530
content errors: 0
0 HREFs
0 HTTPs
16 DIVs
0 WWWs

Export
Disk
RAM

Lex-Analysis
Lex

18.03.2011.09.06

Number of href links on initial page: 124
Number of http links in this selection: 61
...
...

URL:

Redirection Notice:

Isolation tag

Exclusion tag

URL topic clue

PRE-Title tag

Title tag

POST-Title tag

Date tag Date pattern

Author tag

Text tag

View

Clear Version of Data

Raw Version of Data

Wrap Lines

Stats

characters: 1530

content errors: 0

0 HREFs

0 HTTPs

16 DIVs

0 WWWs

Export

Disk

RAM

Lex-Analysis

Lex

Destak/Lusa

Number of href links on initial page: 124

Number of http links in this selection: 61

...

...

URL:

Redirection Notice:

Isolation tag:

Exclusion tag:

URL topic clue:

PRE-Title tag:

Title tag:

POST-Title tag:

Date tag: Date pattern:

Author tag:

Text tag:

View

Clear Version of Data

Raw Version of Data

Wrap Lines

Stats

characters: 1530

content errors: 0

0 HREFs

0 HTTPs

16 DIVs

0 WWWs

Export

Disk

RAM

Lex-Analysis

Lex

<p>De acordo com a fonte, que citou testemunhos dos vizinhos, o homem estaria a tentar fazer uma tosta mista e não se terá apercebido de ter queimado a roupa.<p>"Já com a roupa a arder, o senhor começou a gritar e foram os vizinhos que lhe acudiram", explicou a mesma fonte, referindo que a habitação não sofreu danos.<p>Quando chegaram ao local, cerca das 21:00, os bombeiros limitaram-se a prestar os primeiros socorros ao idoso que foi transportado para o Hospital de Santo António.<p>"Foram os vizinhos que, com panos molhados e um balde, apagaram o fogo das roupas do idoso", acrescentou.</p>

Number of href links on initial page: 124

Number of http links in this selection: 61

...

...

IDnum	DateAdded	WebSiteName	URL	Lang	Country	Date	Hierarchy	Topic
49061	2011-03-23 13:54:27	Expresso	BLOB	POR	POR	2011-03-23 08:00:00	000	/atualidade
49062	2011-03-23 13:54:28	Expresso	BLOB	POR	POR	2011-03-22 18:20:00	000	/atualidade
49063	2011-03-23 13:54:30	Expresso	BLOB	POR	POR	2011-03-23 08:00:00	000	/atualidade
49064	2011-03-23 13:54:34	Expresso	BLOB	POR	POR	2011-03-22 18:59:00	000	/atualidade
49065	2011-03-23 13:54:47	Expresso	BLOB	POR	POR	2011-03-22 20:10:00	000	/atualidade
49066	2011-03-23 13:55:37	Expresso	BLOB	POR	POR	2011-03-22 15:44:00	000	/economia+tecnologia
49067	2011-03-23 13:55:40	Expresso	BLOB	POR	POR	2011-03-22 15:50:00	000	/economia
49068	2011-03-23 13:55:42	Expresso	BLOB	POR	POR	2011-03-22 18:27:00	000	/economia
49069	2011-03-23 13:56:46	Expresso	BLOB	POR	POR	2011-03-23 09:17:00	000	/life&style
49070	2011-03-23 13:57:16	Expresso	BLOB	POR	POR	2011-03-22 17:41:00	000	/desporto
49071	2011-03-23 13:57:19	Expresso	BLOB	POR	POR	2011-03-22 18:10:00	000	/desporto
49072	2011-03-23 13:57:20	Expresso	BLOB	POR	POR	2011-03-22 18:34:00	000	/desporto
49073	2011-03-23 13:57:23	Expresso	BLOB	POR	POR	2011-03-22 16:34:00	000	/desporto
49074	2011-03-23 13:57:24	Expresso	BLOB	POR	POR	2011-03-22 13:20:00	000	/desporto
49075	2011-03-23 13:57:31	Expresso	BLOB	POR	POR	2011-03-22 15:55:00	000	/desporto
49076	2011-03-23 13:57:33	Expresso	BLOB	POR	POR	2011-03-22 16:45:00	000	/desporto
49077	2011-03-23 13:57:50	Expresso	BLOB	POR	POR	2011-03-22 16:08:00	000	/tecnologia
49078	2011-03-23 13:57:54	Expresso	BLOB	POR	POR	2011-03-21 16:38:00	000	/tecnologia
49079	2011-03-23 13:58:07	Expresso	BLOB	POR	POR	2011-03-23 10:02:00	000	/opinião
49080	2011-03-23 13:59:50	Jornal de Notícias	BLOB	POR	POR	2011-03-23 11:47:00	111	/mundo+início+blog+últimas+país
49081	2011-03-23 13:59:52	Jornal de Notícias	BLOB	POR	POR	2011-03-23 13:18:00	101	/cultura+início+blog+últimas+país
49082	2011-03-23 13:59:54	Jornal de Notícias	BLOB	POR	POR	2011-03-23 13:04:00	101	/país+início+blog+últimas+nacional
49083	2011-03-23 13:59:56	Jornal de Notícias	BLOB	POR	POR	2011-03-23 13:06:00	100	/nacional+início

49243 rows fetched in 0:01.1761

Start Editing Apply Changes First Last Search

Author	Title	Text	TaggedText	Tokens	Types	Dictionary
BLOB	Os "animais" e as "bestas" que apedrejam automóveis	BLOB		854	244	BLOB
	Abusos "repugnantes" no Afeganistão tomados públicos	BLOB		524	152	BLOB
BLOB	Somos deuses frágeis, avisa-nos Fukushima	BLOB		1634	381	BLOB
BLOB	Ah leão!	BLOB		2952	735	BLOB
BLOB	Sócrates vai a Belém depois da votação	BLOB		390	107	BLOB
	iPad 2 chega a Portugal na sexta-feira	BLOB		1114	266	BLOB
BLOB	Concertação social: parceiros chegam a acordo	BLOB		244	67	BLOB
BLOB	Bancos centrais "arrefecem" valorização do iene	BLOB		1018	240	BLOB
BLOB	Vídeo: afinal de que tem medo este bebé?	BLOB		546	183	BLOB
BLOB	Vice-presidente do Benfica pede aos adeptos para não retaliarem (vídeo)	BLOB		272	79	BLOB
BLOB	Dias Ferreira desmente Bruno Carvalho	BLOB		150	51	BLOB
BLOB	Vale e Azevedo aguarda em liberdade decisão sobre extradição	BLOB		720	162	BLOB
BLOB	Sporting: um debate quente que escaldou no final	BLOB		186	71	BLOB
BLOB	F1: cinco campeões juntos pela primeira vez em 41 anos	BLOB		802	211	BLOB
BLOB	FPF: Laurentino Dias só se pronuncia após conhecer resoluções da AG	BLOB		292	82	BLOB
BLOB	Sporting: Hans Van der Zee pode ser o <i>manager</i> de Pedro Baltazar	BLOB		152	45	BLOB
BLOB	Apple, Amazon e Microsoft disputam marca "App Store"	BLOB		550	147	BLOB
BLOB	Twitter faz 5 anos	BLOB		402	123	BLOB
	"Se tivesse sido ouvida, o País não estaria seguramente como está."	BLOB		48	24	BLOB
	Reveladas fotos das orgias de Berlusconi	BLOB		1036	295	BLOB
	Morreu Elizabeth Taylor	BLOB		374	147	BLOB
	Agregação de escolas vai permitir ter "menos um ou outro professor"	BLOB		464	131	BLOB
	Marcha pela Educação no próximo dia 2	BLOB		426	147	BLOB

49243 rows fetched in 0:01.1761

Start Editing Apply Changes First Last Search

Author	Title	Text	TaggedText	Tokens	Types	Dictionary
BLOB	Os "animais" e as "bestas" que apedrejam automóveis	BLOB		854	244	BLOB
	Abusos "repugnantes" no Afeganistão tomados públicos	BLOB		524	152	BLOB
BLOB	Somos deuses frágeis, avisa-nos Fukushima	BLOB		1634	381	BLOB
BLOB	Ah leão!	BLOB		2952	735	BLOB
BLOB	Sócrates vai a Belém depois da votação	BLOB		390	107	BLOB
	iPad 2 chega a Portugal na sexta-feira	BLOB		1114	266	BLOB
BLOB	Concertação social: parceiros chegam a acordo	BLOB		244	67	BLOB
BLOB	Bancos centrais "arrefecem" valorização do iene	BLOB		1018	240	BLOB
BLOB	Vídeo: afinal de que tem medo este bebé?	BLOB		546	183	BLOB
BLOB	Vice-presidente do Benfica pede aos adeptos para não retaliar (vídeo)	BLOB				

Field Viewer

Text Binary

```
<title>Os "animais" e as "bestas" que apedrejam automóveis</title>
<text>Uma pessoa que é capaz de lançar de um viaduto em plena auto-estrada um saco ch
```

Total Length: 2483 bytes

OK

Field Viewer

Text Binary

```
de 24
que 20
a 11
e 11
não 11
uma 11
o 10
é 9
um 8
ss 7
```

Total Length: 2504 bytes

49243 rows fetched in 0:01

Editing Apply Changes First Last Search

IDnum	Period	Articles	Tokens	Types	Dictionary
70	2011-02-00	18352	10746190	85804	BLOB
71	2011-03-00	15428	9255502	75816	BLOB
72	2011-00-00	49243	28407524	132802	BLOB
73	0000-00-00	49243	28407524	132802	BLOB
74	LastDay	700	419210	15340	BLOB
75	LastWeek	4975	2932496	43960	BLOB
76	LastMonth	19991	11954442	86188	BLOB
84	LastDAYOrig	0	0	371	BLOB
77	LastYear	48792	28102128	132345	BLOB
78	2011-03-10	845	529194	18229	BLOB
79	2011-03-11	890	563404	17058	BLOB
80	2011-03-18	914	513020	18247	BLOB
81	2011-03-12	512	310828	13991	BLOB
82	2011-03-13	356	177284	9520	BLOB
83	2011-03-14	768	474778	17559	BLOB
85	2011-03-15	701	437060	15755	BLOB
86	2011-03-16	741	432938	15588	BLOB
87	2011-03-17	812	517046	17256	BLOB
88	2011-03-19	560	298866	14387	BLOB
90	2011-03-21	963	599788	18396	BLOB
89	2011-03-20	285	151628	9419	BLOB
91	2011-03-22	700	419210	15340	BLOB
92	2011-03-23	451	305396	12399	BLOB

92 rows fetched in 0:00.2728

Start Editing
 Apply Changes

Organização cronológica do léxico

Dicionário geral	0000-00-00
Dicionário anual	2011 2012 2013...
Dicionário mensal	1 2 3 4 5 6 7 8 9 10 11 12
Dicionário diário	1 2 3 4 ... 31
Dicionário artigo	0 ... 49526

'Performance' do descarregamento

500 a 900 artigos por dia

300.000 ocorrências por dia

10 milhões de ocorrências por mês

20110323-20110323
ALL --

Period Selection
from 2011-01-01
to 2011-03-31

Country Inclusion
ALL
Country Exclusion

SET PARAMS

Module	Explanation
Corpus	<p>The Corpus menu is where you define global characteristics applied to all ensuing modules (Contexts,Concordances and Co-occurrences):</p> <ul style="list-style-type: none"> • a chronological period • a set of lusophone countries to be included or excluded <p>The chronological period must be defined in a YEAR-MONTH-DAY format. For example 20100101-20111231 restricts corpus exploration to articles published during the years 2010 and 2011. If you want to explore only articles published in 2009, then set your Period Selection to 20090101-20091231. If you want to explore every article stored in the database then define the chronological period as 0-20111231 (or whatever the current date is... or even 20500101 to be safe!).</p> <p>The list of countries to be included / excluded must be defined using THREE-LETTER CODES exactly like those used in the database. These codes define lusophone countries: POR identifies PORTugal, ANG for ANGola, MAC for MACau, etc.</p> <p>The syntax of the country codes you want to include / exclude must be respected! You can enter several country codes to include and several to exclude but these codes must be entered contiguously and in CAPITAL letters!</p> <p>For example, PORANGMOZ in the Country Inclusion textbox and an empty Country Exclusion textbox restricts corpus exploration to articles published in newspaper PORtugal, ANGola and MOZambique.</p> <p>If you don't know which codes to use, then enter code ALL in the Country Inclusion textbox and leave the Country Exclusion textbox empty. This will include ALL countries in your corpus exploration. Then, when your results are displayed, each article extract will be identified with a country code. At this point, take note of the countries you want to include / exclude and re-run the extraction with new parameters.</p> <p>For example, ALL in the Country Inclusion textbox and an empty Country Exclusion textbox includes all countries in textual extraction. If you combine ALL in the Country Inclusion textbox with ANGMOZ in the Country Exclusion textbox then all countries are included in textual exploration except ANGola and MOZambique.</p>
	<p>The Contexts menu is where you define parameters for phrasal context extraction:</p> <ul style="list-style-type: none"> • one or several keywords • a contextual extension

Corpus **Contexts** **Concordances** **Co-occurrences** **Neo-Stats** **Help**

20110323-20110323
ALL --

Keyword(s)

Contextual Extension

RUN QUERY

N	COUNTRY	URL	DATE	CONTEXT
1	MAC	[]	2011-02-23	Como conselheiro de Macau, o que eu quero nestes dois anos que ainda faltam, é transformar a imagem maligna do CCP perante os órgãos do Governo, seja quem for o partido, aliás pelo que sei o PSD não funcionava melhor que o PS, às vezes por culpa própria do CCP, que não incutia uma aproximação colaborante.
2	POR	[]	2011-01-07	"Omaha mudou-me e lixou-me e a escola onde ando agora é ainda pior vão ouvir falar sobre a merda maléfica que fiz mas aquela escola de merda levou-me a fazer isto quero que vocês se lembrem de mim por quem eu era antes disto afetei gravemente as vidas das famílias que arruinei mas peço desculpa.
3	POR	[]	2011-01-09	"A alteração de mentalidades que eu quero que haja por exemplo em relação à tolerância, às cunhas, ao favorecimento, à corrupção é também uma alteração da mentalidade que se deve fazer pelo exemplo" , afirmou.
4	POR	[]	2011-01-11	É meu, e eu quero respeito por esse dinheiro.
5	POR	[]	2011-01-13	"Apesar de me terem apresentado uma proposta para renovar por mais quatro anos, já sabem o que eu quero ", disse o jogador.
6	POR	[]	2011-01-16	"Eu não quero ver o nosso país continuar num beco, é por isso que incomodo, que eles nos temem e nos tentam silenciar", frisou.
7	POR	[]	2011-01-22	Antes quero ser eu assaltada do que eles", desabafa, notando o "afastamento" da clientela por causa de outros assaltos recentes: há 15 dias (um carro) e em Novembro (dois).
8	POR	[]	2011-02-01	O que eu quero dizer é que os mercados só compram situações claras e finais, meias decisões nunca são levadas a sério, mesmo que por algum tempo pareça que sim.
9	POR	[]	2011-02-09	Se eu quero viver abaixo dos meus rendimentos começo por poupar à cabeça: assim que o dinheiro entra na minha conta, ele sai automaticamente para uma conta de poupança.
10	POR	[]	2011-02-14	Quero agradecer a todos os adeptos brasileiros que torceram por mim, que choraram comigo quando chorei, que caíram comigo quando eu caí.
11	POR	[]	2011-02-25	"Quando nós defendemos que é bom que as pessoas possam optar por ter uma segunda opinião, significa que eu hoje quero a opinião de um médico ou de uma equipa para me dizer o que é que eu devo fazer, mas ao mesmo tempo também quero ouvir outra opinião.

Neologismos Estatísticos

Uma operação matemática

Dicionario Geral		Dic. 22 marzo 2011		Neologismo ?
<i>de</i>	675876	<i>de</i>	9995	$G - 22/3 = 665881$
<i>a</i>	519189	<i>a</i>	8009	511180
<i>o</i>	430951	<i>o</i>	6802	424149
...	
<i>intrumento</i>	1	<i>intrumento</i>	1	0
<i>wagasumi</i>	2	<i>wagasumi</i>	2	0
<i>euroárea</i>	1	<i>euroárea</i>	1	0
<i>barcelonistas</i>	2	<i>barcelonistas</i>	2	0



Corpus

Contexts

Concordances

Co-occurrences

Neo-Stats

Help

20110323-20110323

ALL --

Country Selection

PORTugal

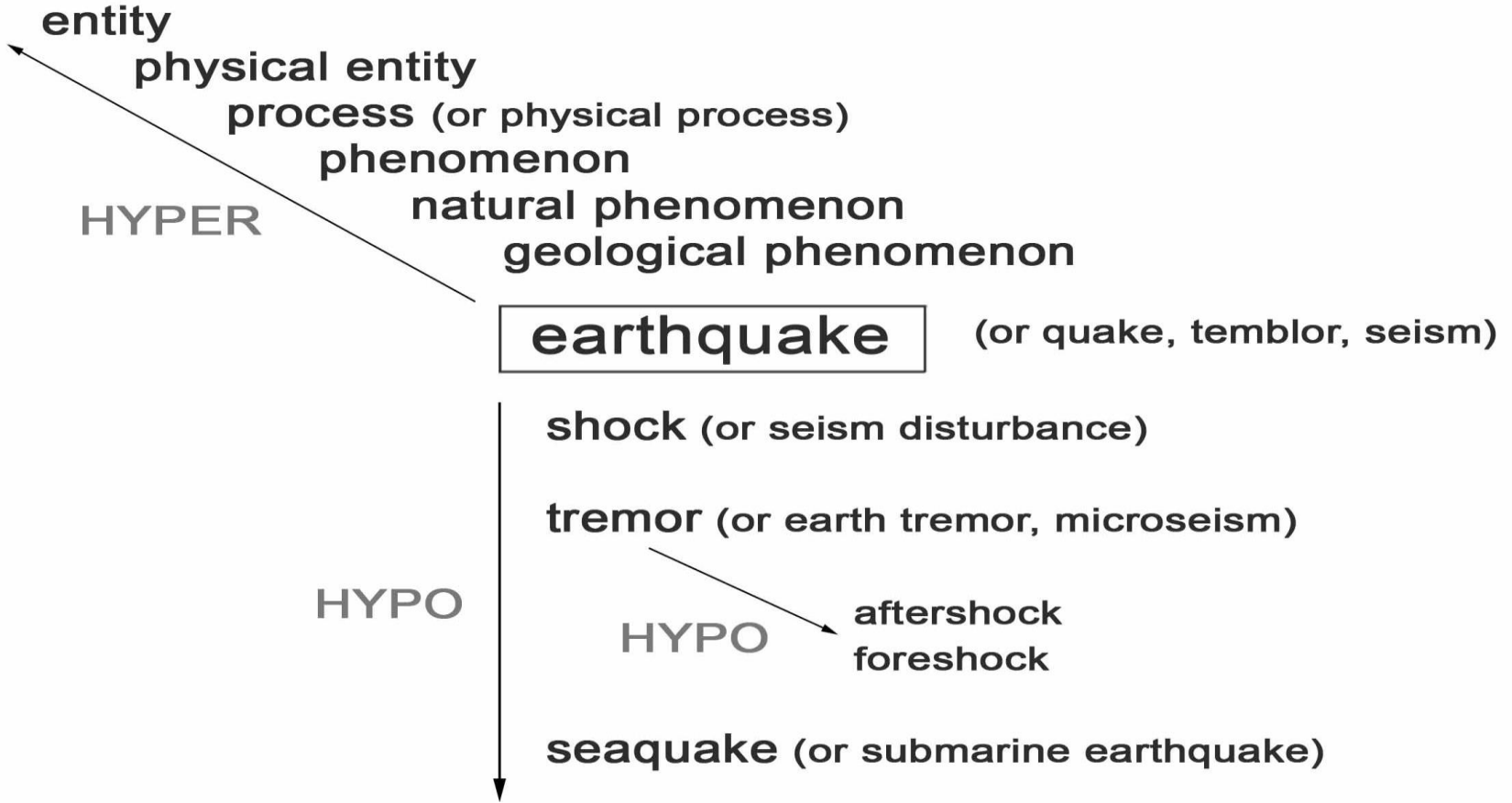
Period Selection

Last Day

EXTRACT

252	oita	2011-03-22	http://dn.sapo.pt/inicio/economia/interior.aspx?content_id=1812036
253	oldemiro	2011-03-22	http://www.ionline.pt/conteudo/112035-portugal-e-os-palop-uma-estrategia-tenaz-incompleta
254	opomo	2011-03-22	http://aeiou.expresso.pt/libia-china-apela-para-cessar-fogo-imediato=f639165
255	ordenará	2011-03-22	http://www.ionline.pt/conteudo/112103-transicao-pacifica-do-poder-negociada-presidente-e-general-no-iyemen
256	orgância	2011-03-22	http://www.jornaldenegocios.pt/home.php?template=SHOWNEWS_V2&id=474773
257	osliar	2011-03-22	http://www.ionline.pt/conteudo/112158-psd-ppc-e-be-entregam-projecto-resolucao-contr-o-pec
258	otpcen	2011-03-22	http://sol.sapo.pt/inicio/Internacional/Interior.aspx?content_id=14847
259	ouahabi	2011-03-22	http://www.abola.pt/nnh/./nnh/ver.aspx?id=253807
260	ouçamo	2011-03-22	http://www.ionline.pt/conteudo/111959-viv-anderson-atiraram-te-fruta-entao-traz-me-duas-peras-e-uma-banana
261	parizes	2011-03-22	http://dn.sapo.pt/inicio/portugal/interior.aspx?content_id=1812335
262	parlamentardo	2011-03-22	http://dn.sapo.pt/inicio/portugal/interior.aspx?content_id=1812485
263	perrota	2011-03-22	http://www.abola.pt/nnh/./nnh/ver.aspx?id=253869
264	peya	2011-03-22	http://www.abola.pt/nnh/ver.aspx?id=253777
265	pineda	2011-03-22	http://www.destak.pt/artigo/90708-ong-acusa-a-policia-de-sequestrar-pessoas-e-vende-las-a-guerrilha-colombiana
266	pinguinhos	2011-03-22	http://www.destak.pt/artigo/90724-nuno-gomes-ja-soma-5-golos...-e-nem-90-minutos-jogou
267	pinguuo	2011-03-22	http://www.abola.pt/nnh/ver.aspx?id=253777
268	piquè	2011-03-22	http://aeiou.expresso.pt/huno-gomes-da-ca-mais-cinco=f639131

Ontologia WordNet



Co-occurrencias

Specific co-occurrences

Nº	Type	Freq.	Co-freq.	Specificity	Contexts
1	the	490004	41936	+50	16720
2	of	268229	33409	+50	19245
3	and	160342	14937	+50	10381
4	1	114129	15835	+50	4190
5	for	78775	7410	+50	4717
6	on	34145	4592	+50	4091
7	from	27034	3939	+50	3466
8	response	25239	5562	+50	5107
9	9	24243	7795	+50	7149
10	10	19039	2167	+50	1836
11	building	18100	4109	+50	3619
12	an	17249	2024	+50	1833
13	no	15676	2042	+50	2021
14	motion	14524	3915	+50	3545
15	ground	13583	2699	+50	2458
16	data	13285	2022	+50	1804
17	frequency	13189	2633	+50	1153
18	acceleration	11851	3183	+50	2720

Ontologia Co-ocurrencial

