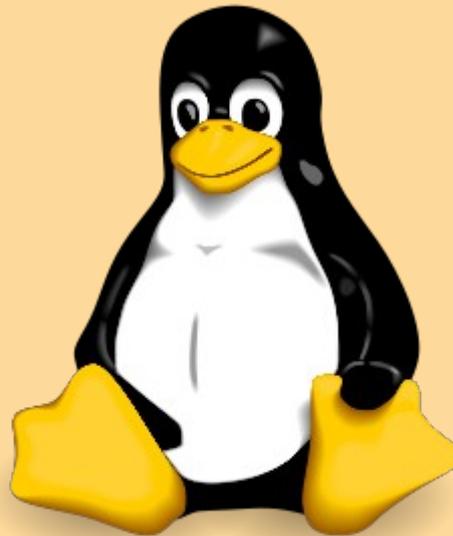


# Le référencement sous Linux



# Ronan CHARDONNEAU

- Responsable webmarketing International chez Jymeo
- Maître de conférences associé à l'Université d'Angers en webmarketing
- Formateur indépendant
- Auteur de Google Analytics, Piwik aux éditions ENI
- En charge de la commission Nord Ouest du SEO Camp
- Utilisateur de Linux Ubuntu depuis 2008
- Alliance Libre à Nantes

# La question du jour

Combien de personnes dans cette salle utilisent  
Linux au quotidien?

# La deuxième question du jour

Combien de personnes dans cette salle ont une personne dans leur entreprise qui utilisent Linux au quotidien?

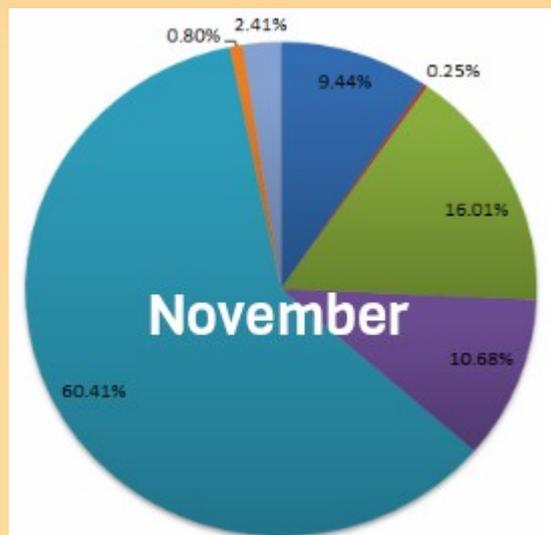
# Parts de marché

- Source: netmarketshare.com
- Linux = 1,21 % du marché en janvier 2013

Operating System	Total Market Share
Windows 7	44.48%
Windows XP	39.51%
Windows Vista	5.24%
Mac OS X 10.8	2.44%
Windows 8	2.26%
Mac OS X 10.6	2.00%
Mac OS X 10.7	1.96%
Linux	1.21%
Mac OS X 10.5	0.52%
Mac OS X 10.4	0.13%
Windows 8 Touch	0.08%
Windows 2000	0.06%
Windows NT	0.05%
Mac OS X (no version reported)	0.02%
Windows 8 RT Touch	0.02%
Windows 98	0.01%
Mac OS X Mach-O	0.00%

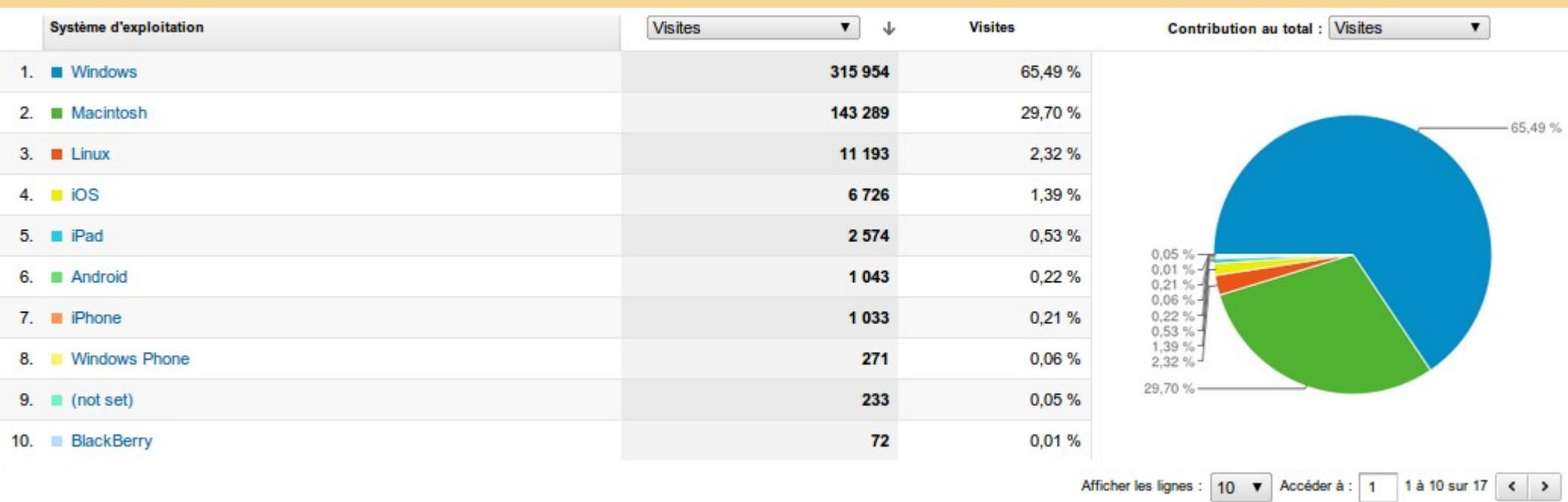
# Parts de marché

- Source: Chitika
- Linux = 0,8% du marché en novembre 2012



# Parts de marché

- Source: autre
- Linux = 2% sur un site recevant 500,000 visites



# Constat

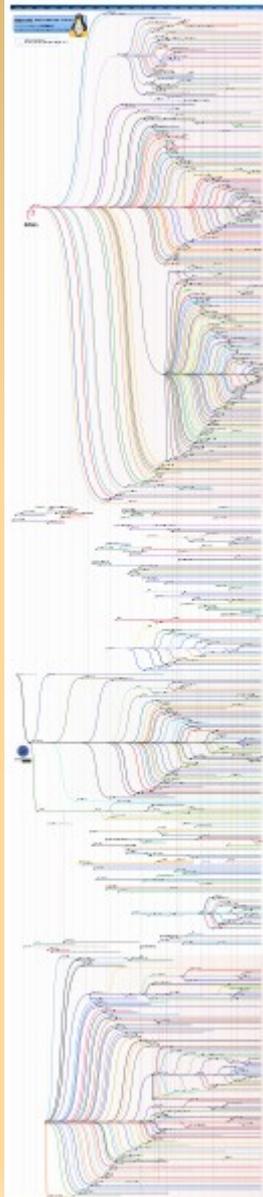
Linux est bien minoritaire en terme de parts de marché comparé à ses deux compatriotes:  
Windows et Mac

# Linux en quelques mots

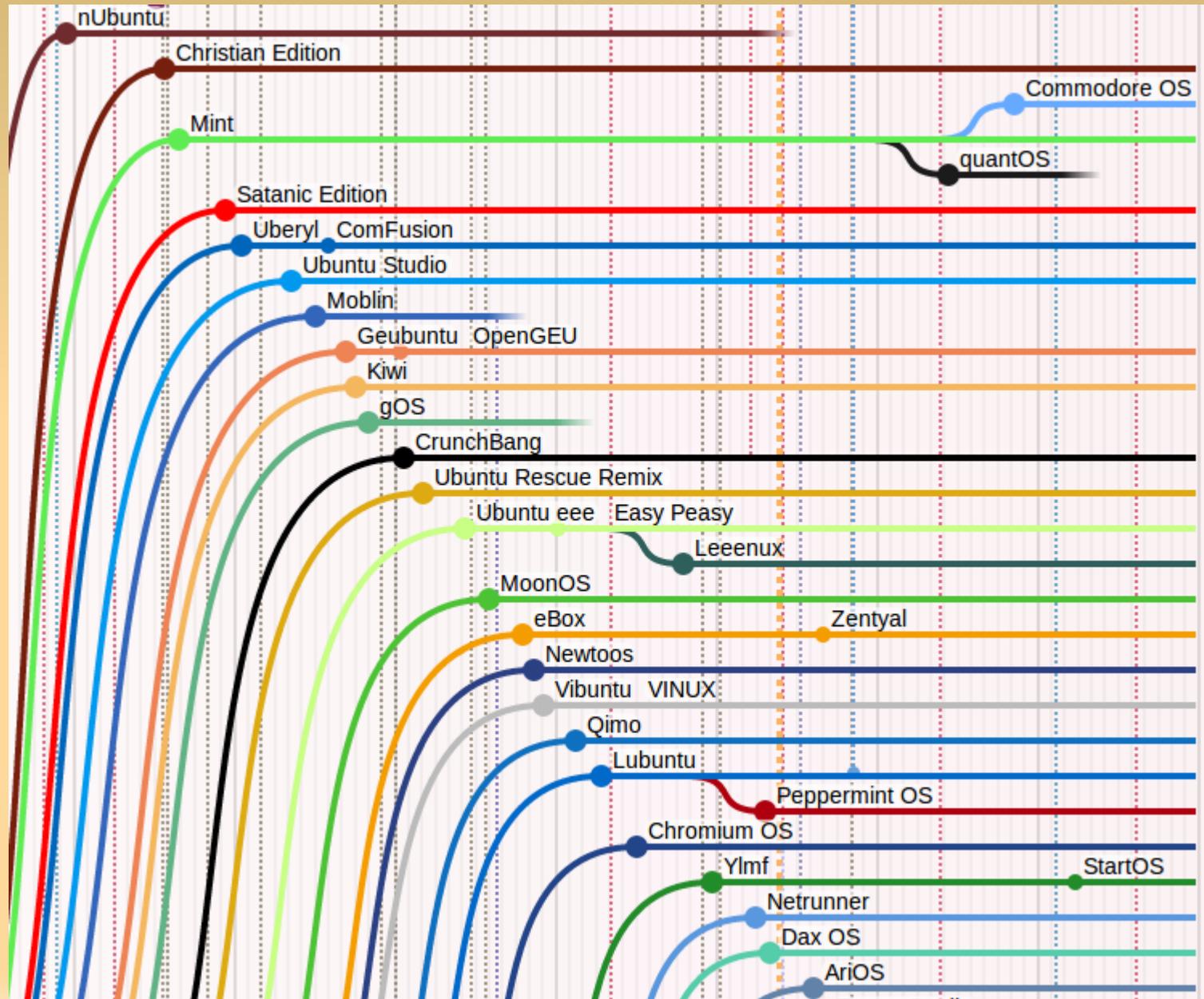
# Linux

- Système d'exploitation principalement utilisé sur les serveurs
- Logiciel libre
- Multiples distributions: Linux Ubuntu, Linux Mint, Debian, Red Hat, Backtrack
- Parts de marché des serveurs web, plus de 50% tournent sous Linux
- Parts de marché sur les super calculateurs: 89,20 % sont sous Linux

# Une multitude de distribution Linux



# Si on zoom



Pour faire simple Linux tourne sur la plupart des ordinateurs surpuissants mais n'arrive pas à percer dans monde des ordinateurs classiques.

# Les raisons

- La vente liée
- Le terminal qui fait peur
- L'aspect technique
- La formation
- Marketing et Logiciel libre ne font généralement pas bon ménage :(

Et Linux pour un référenceur dans tout ça?

# Mes contraintes au quotidien

- Le temps
- L'accès à l'internet
- Software robuste
- Rester maître de ce que l'on fait
- Tâche répétitive
- Vitesse d'exécution plus rapide

# Mes outils

- Lynx
- W3M
- Grep
- Awk
- Sed
- Bash
- Et de nombreux raccourcis claviers et autres commandes de base :)

Lynx

# Lynx

Lynx est un navigateur texte qui se commande par l'intermédiaire de votre clavier. Il est souvent cité en référencement comme étant un navigateur permettant de savoir ce que Google voit lui-même de votre site.

Navigateur Web Firefox

## Sites facilement explorables par Google - Centre d'aide Outils pour les webmasters - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils Aide

Portail x soluce finale ... x Lynx docume... x Modifier l'arti... x SEO Camp D... x Sites facilem... x

support.google.com/webmasters/bin/answer.py?hl=fr&answer=40349 DuckDuckGo

Most Visited Linux Mint Community Forums Blog News

liens non naturels vers votre site sont ajoutés dans le but spécifique de le rendre plus populaire auprès des moteurs de recherche. Certains de ces liens (par exemple, les systèmes de liens et les pages satellite) sont décrits dans nos [Consignes aux webmasters](#).

Seuls les liens naturels sont utiles pour l'indexation et le classement de votre site.

### Facilitez l'accès à votre site

Lors de la conception de votre site, vos liens doivent être définis selon une structure logique. Chaque page doit être accessible à partir d'au moins un lien texte statique.

Utilisez un navigateur texte, tel que [Lynx](#), pour analyser votre site. La plupart des robots d'exploration visualisent votre site de la même manière que Lynx. Si certaines fonctionnalités, notamment les scripts JavaScript, les cookies, les ID de session, les [cadres](#), les balises DHTML ou les contenus [Macromedia Flash](#) empêchent l'affichage de l'intégralité de votre site dans un navigateur texte, les robots risquent de rencontrer des difficultés lors son exploration.

### Actions déconseillées

Ne remplissez pas vos pages de mots clés. Ne faites pas non plus appel à des techniques de dissimulation de contenu (cloaking) ni à des pages spécialement conçues pour être

Menu Sites facile... [Télécharg... linux-seo.o... Terminal Lynx SEO.o... Dossier pe... [titre.txt (~... 165-65-r14... -1- 2 !

Comment ça marche?

# Mode d'emploi

- Installation de lynx (sudo apt-get install lynx)
- Lancement de Lynx (lynx [www.yahoo.fr](http://www.yahoo.fr))
- Bonne navigation :)

# Voici ce que vous voyez

```
#<<< SEO Camp Day Caen - Yahoo! France Résultats de recherche (p2 of 4)
* Recherche : (*) sur tout le Web ( ) en français ( ) en France

FILTRER PAR DATE
* Tous les résultats
* Moins de 24 heures
* Moins d'une semaine
* Moins d'un mois

Résultats de recherche

1.
SEO Camp - L'association des référenceurs
  SEO Camp Day: Mercredi 16 Février à Caen. TunisSEO: Samedi 2 Mars 2013 à Tunis. SEO Camp'us: Jeudi 14 et Vendredi 15 Mars 2013 à L'usine
  (Paris)
  www.seo-camp.org - En cache

2.
» SEO Camp Day de Caen le samedi 16 février 2013 - News ...
  » SEO Camp Day de Caen le samedi 16 février 2013 Veille S2M : actualité quotidienne du référencement naturel, liens sponsorisés, moteurs de
  recherche, réseaux ...
  www.secrets2moteurs.com/seo-camp-day-de-caen-le-samedi... - En cache

3.
L'agenda SEO Camp | SEO Camp
  L'Association de la Communauté du Référencement - Loi 1901 ... Événement: Date: Ville: Apéro SEO à Grenoble: 19/02/2013: Grenoble: SEO Camp
  Day de Caen
  mon.seo-camp.org/agenda?seocamp=00bec3583c26234d250f... - En cache

4.
Nota Bene - Caen, France - Non-Profit Organization | Facebook
  J-24 avant le Seo Camp Day de Caen au Forum digital : http://www.forum-digital.fr/fr/agenda/seo-camp-day-ag103.html - Renseignement :
  willy@leloutre.com
  www.facebook.com/AssociationNotaBene - En cache
  Plus de résultats de facebook.com »
```

Sexy non ? :P

Utiliser Lynx pour parcourir le web en mode clavier est assez prise de tête voir même très prise de tête. Cependant Lynx vous réserve bien des surprises. Que nous allons voir dans un instant.

Mais avant tout comment s'éduquer?

# Que nous dit la littérature?

amazon.fr

Chez Ronan | Promotions | Chèques-cadeaux | Aide

Soldes et promotion  
-40%, -60%, -80%

Parcourir les boutiques

Rechercher

Toutes nos boutiques

lynx browser

Go

Bonjour Ronan  
Votre compte

Adhérer à Premium

0 Panier

Amazon.fr Meilleures ventes Promotions Listes d'envies Chèques-cadeaux Economisez en vous abonnant Options de livraison Vendez !

Boutique

Livres anglais et étrangers

Social Sciences

+ Voir les 2 boutiques

Option d'expédition (Qu'est-ce?)

Livraison gratuite

"lynx browser"

Résultats 1 - 16 sur 548

Choisissez une boutique



**Lynx (Web Browser)** de Jesse Russell et Ronald Cohn (1 mars 2012)

**EUR 11,91** Broché

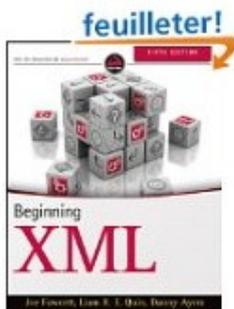
Habituellement expédié sous 2 à 3 semaines

Plus de choix d'achat - Broché

**EUR 9,93** neuf (4 offres)

Livraison gratuite possible (voir fiche produit).

**Livres anglais et étrangers:** Voir l'ensemble des 518 articles



**Beginning XML** de Joe Fawcett, Danny Ayers et Liam R. E. Quin (6 juillet 2012)

**EUR 24,23** Broché

Plus que 7 ex. Commandez vite !

**EUR 20,60** Format Kindle

Disponible pour le téléchargement maintenant

Plus de choix d'achat - Broché

**EUR 21,23** neuf (22 offres)

**EUR 26,39** d'occasion (3 offres)

Livraison gratuite possible (voir fiche produit).

Extrait

Page 655: ... very simple and fast web browser), and Lynx (a text-only web browser ... Voir une page aléatoire dans ce livre.

**Livres anglais et étrangers:** Voir l'ensemble des 518 articles



**Articles on Command-Line Software, Including: Links (Web Browser), Lynx (Web Browser), Emacs/M2, Line Mode** de Harrobalus Books (28 septembre 2011)



Il vaut mieux se tourner vers les sites officiels  
pour un bon guide d'utilisation

# S'éduquer sur Lynx

- [http://lynx.isc.org/lynx2.8.7/lynx2-8-7/lynx\\_help/Lynx\\_users\\_guide.html](http://lynx.isc.org/lynx2.8.7/lynx2-8-7/lynx_help/Lynx_users_guide.html)
- De longues heures à lire de l'anglais et à tester des lignes de commande vous attendent :(

# Quelques must-have techniques sur Lynx

# La fonction crawl

Lynx -crawl -traversal

# La fonction crawl

Copie l'intégralité des pages parcourues dans un fichier. Vous avez désormais tout le contenu textuel dans un joli document que vous pouvez imprimer et montrer à votre client pour lui expliquer que son contenu est à revoir :)

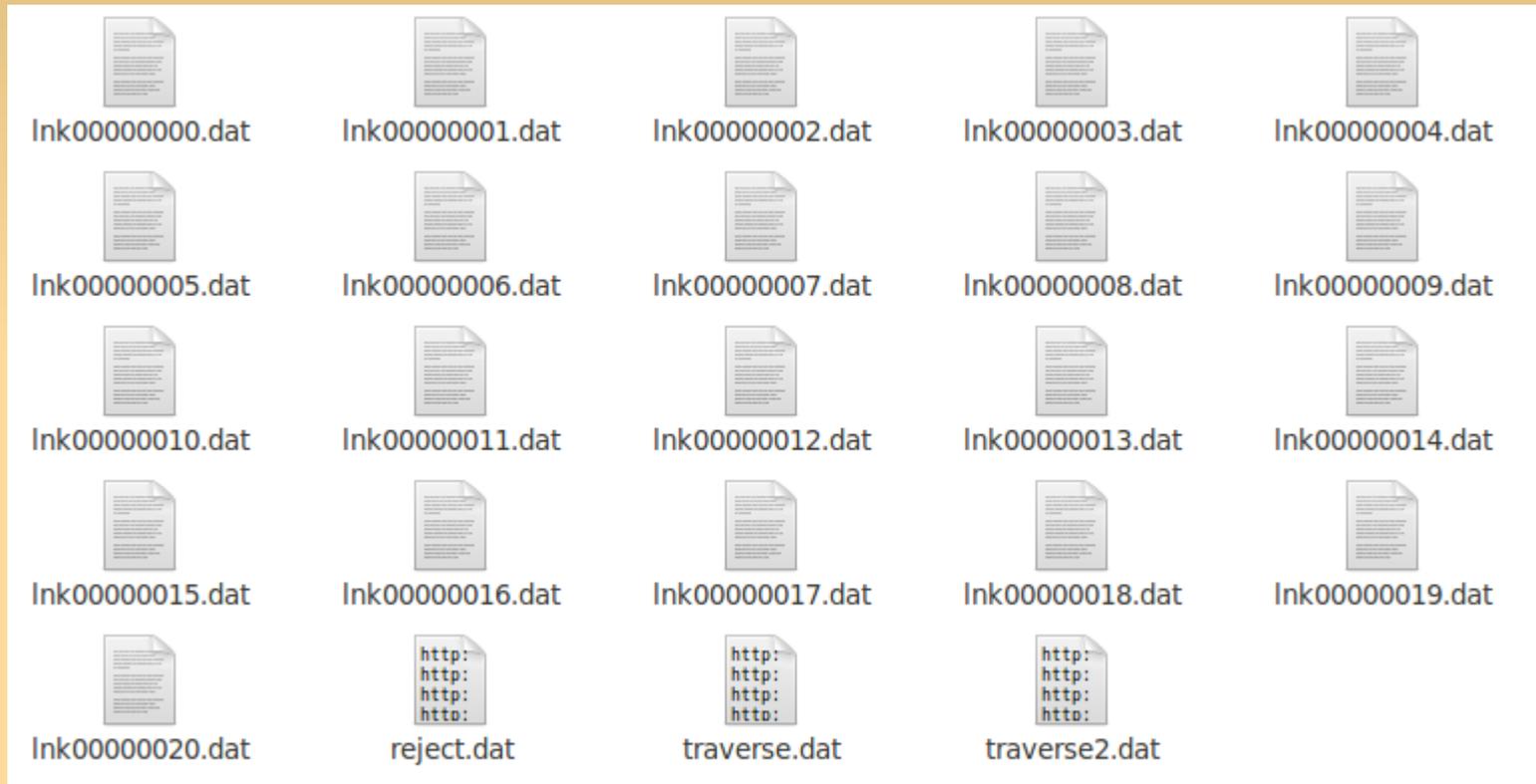
# Crawler les pages

```
lynx -crawl -traversal "http://www.seo-camp.org"
```

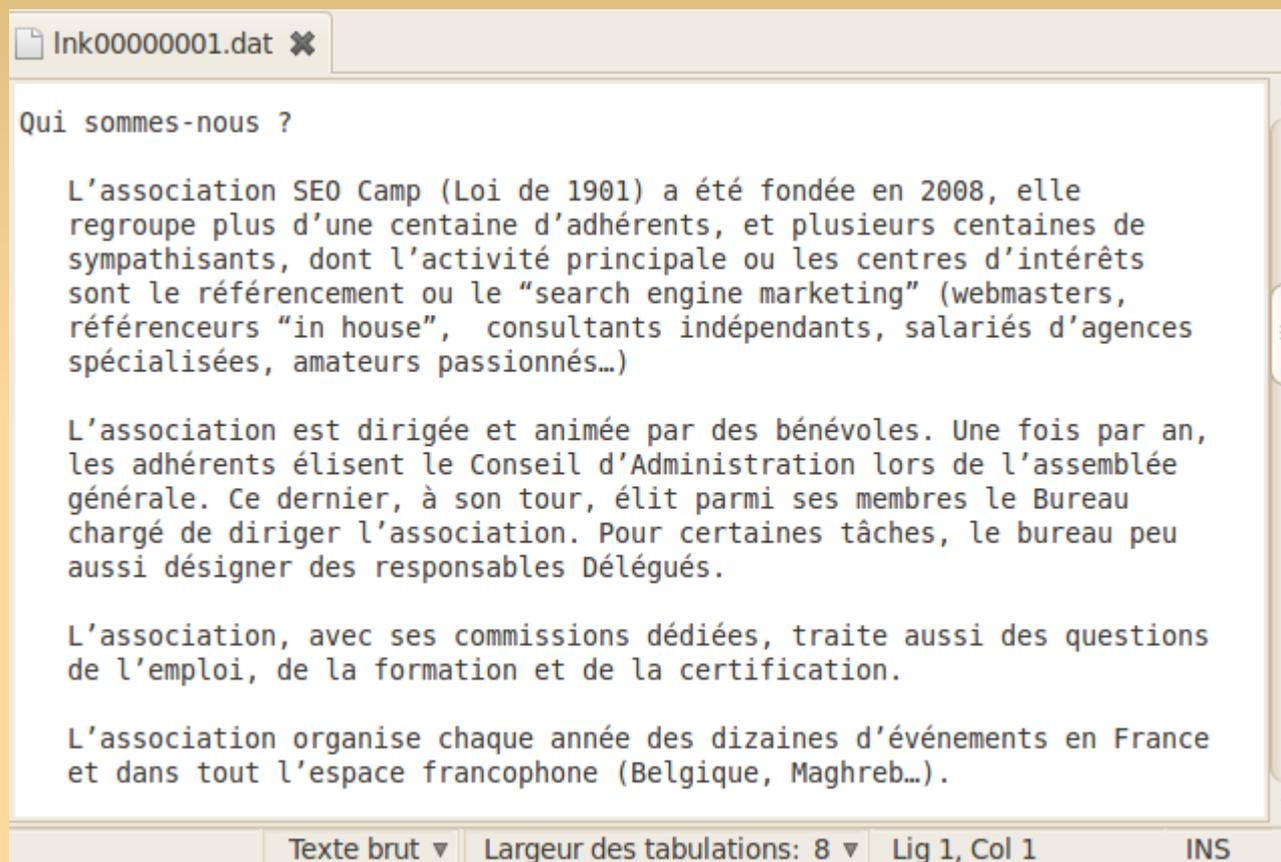
# Les fichiers générés

- `traverse.dat` — Liste des URLs crawlées
- `traverse2.dat` — Liste des urls crawlées avec le titre des pages
- `Ink00000###.dat` — Contenu texte des pages

# Concrètement



# Concrètement



Ink00000001.dat ✕

Qui sommes-nous ?

L'association SEO Camp (Loi de 1901) a été fondée en 2008, elle regroupe plus d'une centaine d'adhérents, et plusieurs centaines de sympathisants, dont l'activité principale ou les centres d'intérêts sont le référencement ou le "search engine marketing" (webmasters, référenceurs "in house", consultants indépendants, salariés d'agences spécialisées, amateurs passionnés...)

L'association est dirigée et animée par des bénévoles. Une fois par an, les adhérents élisent le Conseil d'Administration lors de l'assemblée générale. Ce dernier, à son tour, élit parmi ses membres le Bureau chargé de diriger l'association. Pour certaines tâches, le bureau peut aussi désigner des responsables Délégués.

L'association, avec ses commissions dédiées, traite aussi des questions de l'emploi, de la formation et de la certification.

L'association organise chaque année des dizaines d'événements en France et dans tout l'espace francophone (Belgique, Maghreb...).

Texte brut ▼ Largeur des tabulations: 8 ▼ Lig 1, Col 1 INS

# La fonction source

Vous permet de voir le code source d'une page.

Lynx -source <http://www.seo-camp.org>

Terminal

```
Fichier Fichier Édition Affichage Rechercher Terminal Aide
<div class="blog-description" style="padding:10px;">
L&#8217;Association de la Communauté du Référencement - Loi 1901</div>

<div id="search"><form id="searchform" action="http://www.seo-camp.org" method="
get"><input type="hidden" name="PHPSESSID" value="5c5bd5891838e0d76d17c111416521
d8" />
<label class="hidden" for="s"></label>
33 <input id="s" type="text" onblur="if(this.value=='')this.value=this.defaultValue
;" onfocus="if(this.value==this.defaultValue)this.value='';" name="s" value="Rec
hercher"/>
<input id="searchsubmit" type="submit" value="" />
</form>
</div>

</div><!-- #header-->

34

<div id="leader">
<div id="leader-container">
<h1 class="leader-title">
```

Tâches Afficher

- Pages maîtresses
- Mises en page
  - [Thumbnail]
  - [Thumbnail]
  - [Thumbnail]
  - [Thumbnail]
  - [Thumbnail]
- Ebauches de tableau
- Animation personnalisée
- Transition

Diapo 34 / 68 HumanUbuntu 36%

Menu Lynx documentation... Téléchargements linux-seo.odp - Libre... Terminal

# Extraire une page web sous Lynx et la copier dans un fichier

```
lynx -dump http://www.seo-camp.org > home.txt
```

# Exemple d'utilisation avec vimdiff

```
vimdiff home-ch.txt home-lu.txt
```

# Concrètement

The image shows a LibreOffice Impress presentation window titled "Lynx SEO.odp". The main content is a terminal window displaying a website interface for comparing tire prices. The terminal is split into two panes, each showing a different version of the website.

**Terminal Window Title:** Terminal

**Terminal Menu:** Fichier Édition Affichage Rechercher Terminal Onglets Aide

**Terminal Content (Left Pane):**

```
[1]Belgique Suisse [2]France [3]Luxembourg
| Suisse

Premier comparateur de prix de pneus

Comparateur de Prix de Pneus : Quel Pneu .ch - Le Meilleur Prix pour vos
Pneus Neufs

1
Rechercher un pneu
>>
2
Comparer les prix
>>

+ +-- 45 lignes : 3--

Astuce : vous pouvez monter un pneu dont l'indice de charge est
supérieur à celui d'origine.
[ ]
Runflat
[ ]
Renforcé
Plus d'options de recherche (saison, charge...)
Comparez les prix des pneus neufs vendus sur Internet avec le
comparateur de pneus Quel Pneu .ch .
Sélectionnez les références du pneu qui vous intéresse puis cliquez sur
« Rechercher ».
Pneus auto, pneus 4x4, pneus moto, toutes les références sont sur Quel
Pneu, le premier comparateur de pneus.

Copyright © 2008-2013 - Comparateur de prix de pneus Quelpneu ® -
charte [4]Conditions d'utilisation du site - presse [5]La presse en
parle

Pneus Michelin

[6]Michelin Energy Saver
[7]Michelin Primacy HP
[8]Michelin Pilot Sport 3
[9]Michelin Alpin A4
[10]Michelin Pilot Super Sport

Pneus Hiver
```

**Terminal Content (Right Pane):**

```
[1]Belgique [2]Suisse [3]France Luxembourg
| Luxembourg

+ +-- 45 lignes : 3--

Astuce : vous pouvez monter un pneu dont l'indice de charge est
supérieur à celui d'origine.
[ ]
Runflat
[ ]
Renforcé
Plus d'options de recherche
Comparez les prix des pneus neufs vendus sur Internet avec le
comparateur de pneus Quel Pneu .lu .
Sélectionnez les références du pneu qui vous intéresse puis cliquez sur
« Rechercher ».
Pneus auto, pneus 4x4, pneus moto, toutes les références sont sur Quel
Pneu, le premier comparateur de pneus.

Copyright © 2008-2013 - Comparateur de prix de pneus Quelpneu ® -
charte [4]Conditions d'utilisation du site - presse [5]La presse en
parle
```

W3M

# W3M

Idem que Lynx mais fonctionne avec la souris

# Concrètement

[ Recherche Images Maps Play YouTube Actualités Gmail Drive Plus » ]  
Historique Web | Paramètres | Connexion

France

Search Books w3m browser [Go]

[Achat chaussures] [Recherche Google][J'ai de la chance]

Nouveau ! Explorez le Grand Canyon avec Street View dans Google Maps

Showing 1 - 12 of 40 results

Format Solutions publicitaires Solutions d'entreprise Google+ À propos de Google Google.com

© 2012 - Confidentialité et conditions d'utilisation

Running Linux by Matthias Kalle Dalheimer and Matt Welsh (Dec 29, 2005)

Format	Price	New	Used	Collectible
Paperback	<del>\$49.99</del> \$31.49 <small>off price</small>	\$22.25	\$8.43	\$129.99
Auto-delivered eBook	\$29.48			

Excerpt - Index ... VoIP (Voice over IP), 329  
w3m browser, 142  
watch command (gdb), 706 ... [See a random page in this book.](#)

Sell this back for an Amazon.com Gift Card

Linux Bible 2011 Edition: Boot up to Ubuntu, Fedora, KNOPPIX, Debian, openSUSE, and 13 Other Distributions by Christopher Negus (Jan 4, 2011)

Format	Price	New	Used
Paperback	<del>\$49.99</del> \$31.49 <small>off price</small>	\$2.98	\$8.27



# La littérature ne dit pas grand chose

**amazon** Join Prime Ronan's Amazon.com Today's Deals Gift Cards Help

Free One-Day Shipping On Valentine's Day Jewelry & Watch Gifts [Shop now](#)

Shop by Department  Search   Hello, **Ronan** Your Account

Books Advanced Search Browse Subjects New Releases Best Sellers The New York Times® Best Sellers Children's Books Textbooks Sell Your Books

**Department**  
Any Department  
**Books**  
Computers & Technology (39)  
Education & Reference (9)  
Business & Investing (2)  
Self-Help (1)  
Arts & Photography (1)  
Professional & Technical (2)

**Format**  
Paperback (38)  
Hardcover (2)

**Author**  
Michael H. Jang (1)

**Language**  
 English (33)  
 German (3)

**International Shipping** (What's this?)  
 AmazonGlobal Eligible

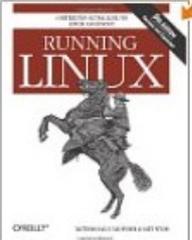
**Shipping Option** (What's this?)  
Free Super Saver Shipping

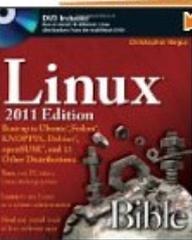
**Promotion**  
 Bargain Books (5)

**Avg. Customer Review**  
★★★★☆ & Up (26)

**Books > "w3m browser"**  
Showing 1 - 12 of 40 Results   Sort by

**Format**

1.  **Running Linux** by Matthias Kalle Dalheimer and Matt Welsh (Dec 29, 2005)  
★★★★☆ (136 customer reviews)  
**Formats** **Price** **New** **Used** **Collectible**  
**Paperback**  
Order in the next **3 hours** to get it by Tuesday, Feb 12.  
~~\$49.99~~ **\$31.49** **\$22.25** **\$8.43** **\$129.99**  
Only 7 left in stock - order soon.  
Eligible for **FREE** Super Saver Shipping.  
Auto-delivered wirelessly **\$29.48**  
**Excerpt - Index:** "... VoIP (Voice over IP), 325 W w3m browser, 142 watch command (gdb), 706 ..." [See a random page](#) in this book.  
Sell this back for an Amazon.com Gift Card

2.  **Linux Bible 2011 Edition: Boot up to Ubuntu, Fedora, KNOPPIX, Debian, openSUSE, and 13 Other Distributions** by Christopher Negus (Jan 4, 2011)  
★★★★☆ (6 customer reviews)  
**Formats** **Price** **New** **Used**  
**Paperback**  
Order in the next **3 hours** to get it by Tuesday, Feb 12.  
~~\$49.99~~ **\$31.49** **\$2.98** **\$8.27**  
Eligible for **FREE** Super Saver Shipping.  
**Kindle Edition** **\$29.99**

# L'idéal restant

- La source officielle pour suivre le manuel d'utilisation
- <http://w3m.sourceforge.net/MANUAL>

Quelques fonctions intéressantes

Grep

# Le Grep

grep est un programme en ligne de commande de **recherche** de chaînes de caractères.

# Signification

Rechercher globalement les correspondances avec l'expression rationnelle (en anglais, regular expression), et imprimer (print) les lignes dans lesquelles elle correspond

Petit exemple d'utilisation

Je cherche à savoir si un site Internet utilise  
Google Analytics

```
lynx -source http://www.seo-camp.org | grep -q  
'_gaq.push' && echo "GA installé" || echo "GA non  
installé"
```

# Résultat

```
ubuntu-laptop:~$ lynx -source http://www.seo-camp.org | grep  
-q '_gaq.push' && echo "GA installé" || echo "GA non installé"  
GA installé
```

Vous pouvez d'ores et déjà imaginer l'ensemble des combinaisons possibles que vous pouvez mettre en place.

Des lignes de commandes un peu plus costaud

Extraire les 5 premiers résultats de la page de  
Google

```
lynx -dump 'http://www.google.fr/search?q=seo
%20&num=5&start=0' | egrep -o "http(s)?://[-
[:alnum:]]+\.[-[:alnum:]]+\.[a-zA-Z]{2,4}" | grep -Ev
'http://webcache.googleusercontent.com|
https://play.google.com|http://www.youtube.com|
https://mail.google.com|https://drive.google.com|
https://accounts.google.com' | grep -v "google.fr" |
egrep -v '[0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}'
| awk '{ print "\'"$1"\'" }' | awk '!( $0 in a )
{ a[$0]; print }' | awk '{ print $0 }'
```

# Concrètement

```
laptop:~$ lynx -dump 'http://www.google.fr/search?q=se
0%20&num=5&start=0 | egrep -o "http(s)?://[-[:alnum:]]+\.[-[:alnum:]]+\.[a-zA-Z
]{2,4}" | grep -Ev 'http://webcache.googleusercontent.com|https://play.google.co
m|http://www.youtube.com|https://mail.google.com|https://drive.google.com|https:
//accounts.google.com' |grep -v "google.fr" | egrep -v '[0-9]{1,3}\.[0-9]{1,3}\.
[0-9]{1,3}\.[0-9]{1,3}' | awk '{ print " \""$1"\""' }' | awk '!($0 in a){a[$0];pri
nt}' | awk '{print $0 }'
"http://fr.wikipedia.org"
"http://www.seo.fr"
"http://www.seo-camp.org"
"http://www.brioude-internet.fr"
"http://support.google.com"
"http://www.journaldunet.com"
```

A partir de là vous pouvez imaginer oh combien vous pouvez aller plus vite avec des petits scripts.

# Autre utilisation

Savoir si un site utilise Google Adsense  
`lynx -source www.justacote.com | grep -q -i  
"google_ad" && echo "adsense accepted"`

Bash

# Bash

Bash est un interpréteur en ligne de commande  
qui sert à faire des scripts

Un petit exemple concret avec `#!/bin/bash`

Je souhaite répéter une même action avec différentes variables.

Exemple concret: savoir pour une liste de sites si Xiti est installé

```
1 #!/bin/bash
2 lynx -source http://agriculture.gouv.fr | grep -q 'xiti' && echo "http://agriculture.gouv.fr xiti installé" || echo "http://a
3 lynx -source http://alimentation.gouv.fr | grep -q 'xiti' && echo "http://alimentation.gouv.fr xiti installé" || echo "http://
4 lynx -source http://alize2.finances.gouv.fr | grep -q 'xiti' && echo "http://alize2.finances.gouv.fr xiti installé" || echo "
5 lynx -source http://basedoc.diplomatie.gouv.fr | grep -q 'xiti' && echo "http://basedoc.diplomatie.gouv.fr xiti installé" ||
6 lynx -source http://bifp.fonction-publique.gouv.fr | grep -q 'xiti' && echo "http://bifp.fonction-publique.gouv.fr xiti insta
7 lynx -source http://bofip.impots.gouv.fr | grep -q 'xiti' && echo "http://bofip.impots.gouv.fr xiti installé" || echo "http://
8 lynx -source http://cache.media.education.gouv.fr | grep -q 'xiti' && echo "http://cache.media.education.gouv.fr xiti installé
9 lynx -source http://carto.observatoire-des-territoires.gouv.fr | grep -q 'xiti' && echo "http://carto.observatoire-des-territo
10 lynx -source http://circulaires.legifrance.gouv.fr | grep -q 'xiti' && echo "http://circulaires.legifrance.gouv.fr xiti insta
11 lynx -source http://competitivite.gouv.fr | grep -q 'xiti' && echo "http://competitivite.gouv.fr xiti installé" || echo "http
12 lynx -source http://directe.gouv.fr | grep -q 'xiti' && echo "http://directe.gouv.fr xiti installé" || echo "http://directe
13 lynx -source http://draaf.auvergne.agriculture.gouv.fr | grep -q 'xiti' && echo "http://draaf.auvergne.agriculture.gouv.fr xi
14 lynx -source http://draaf.centre.agriculture.gouv.fr | grep -q 'xiti' && echo "http://draaf.centre.agriculture.gouv.fr xiti i
15 lynx -source http://draaf.paca.agriculture.gouv.fr | grep -q 'xiti' && echo "http://draaf.paca.agriculture.gouv.fr xiti insta
16 lynx -source http://draaf.rhone-alpes.agriculture.gouv.fr | grep -q 'xiti' && echo "http://draaf.rhone-alpes.agriculture.gouv
17 lynx -source http://en.europe-en-france.gouv.fr | grep -q 'xiti' && echo "http://en.europe-en-france.gouv.fr xiti installé" |
18 lynx -source http://esante.gouv.fr | grep -q 'xiti' && echo "http://esante.gouv.fr xiti installé" || echo "http://esante.gouv
19 lynx -source http://jdb.marine.defense.gouv.fr | grep -q 'xiti' && echo "http://jdb.marine.defense.gouv.fr xiti installé" ||
20 lynx -source http://legifrance.gouv.fr | grep -q 'xiti' && echo "http://legifrance.gouv.fr xiti installé" || echo "http://leg
21 lynx -source http://lesportesdutemps.culture.gouv.fr | grep -q 'xiti' && echo "http://lesportesdutemps.culture.gouv.fr xiti i
22 lynx -source http://m.defense.gouv.fr | grep -q 'xiti' && echo "http://m.defense.gouv.fr xiti installé" || echo "http://m.def
23 lynx -source http://travail-emploi.gouv.fr | grep -q 'xiti' && echo "http://travail-emploi.gouv.fr xiti installé" || echo "ht
24 lynx -source http://urbamet.documentation.developpement-durable.gouv.fr | grep -q 'xiti' && echo "http://urbamet.documentation
25 lynx -source http://www.action-publique.gouv.fr | grep -q 'xiti' && echo "http://www.action-publique.gouv.fr xiti installé" |
26 lynx -source http://www.adoption.gouv.fr | grep -q 'xiti' && echo "http://www.adoption.gouv.fr xiti installé" || echo "http://
27 lynx -source http://www.aft.gouv.fr | grep -q 'xiti' && echo "http://www.aft.gouv.fr xiti installé" || echo "http://www.aft.g
28 lynx -source http://www.ajap.culture.gouv.fr | grep -q 'xiti' && echo "http://www.ajap.culture.gouv.fr xiti installé" || echo
29 lynx -source http://www.ape.minefi.gouv.fr | grep -q 'xiti' && echo "http://www.ape.minefi.gouv.fr xiti installé" || echo "ht"
```

# Lancement du script

```
laptop:~/Bureau$ ./script-xiti  
http://agriculture.gouv.fr xiti installé  
http://alimentation.gouv.fr xiti installé  
http://bifp.fonction-publique.gouv.fr xiti installé
```

# Résultat

	A	B
A1	<a href="http://www.nouvelle-caledonie.gouv.fr">http://www.nouvelle-caledonie.gouv.fr</a> Pas d'installation	
1	<a href="http://www.nouvelle-caledonie.gouv.fr">http://www.nouvelle-caledonie.gouv.fr</a> Pas d'installation	
2	<a href="http://www.nouvelle-caledonie.gouv.fr">http://www.nouvelle-caledonie.gouv.fr</a> Pas d'installation	
3	<a href="http://alize2.finances.gouv.fr">http://alize2.finances.gouv.fr</a> Pas d'installation	
4	<a href="http://basedoc.diplomatie.gouv.fr">http://basedoc.diplomatie.gouv.fr</a> Pas d'installation	
5	<a href="http://bifp.fonction-publique.gouv.fr">http://bifp.fonction-publique.gouv.fr</a> Pas d'installation	
6	<a href="http://bofip.impots.gouv.fr">http://bofip.impots.gouv.fr</a> Pas d'installation	
7	<a href="http://cache.media.education.gouv.fr">http://cache.media.education.gouv.fr</a> Pas d'installation	
8	<a href="http://carto.observatoire-des-territoires.gouv.fr">http://carto.observatoire-des-territoires.gouv.fr</a> Pas d'installation	
9	<a href="http://circulaires.legifrance.gouv.fr">http://circulaires.legifrance.gouv.fr</a> Pas d'installation	
10	<a href="http://competitivite.gouv.fr">http://competitivite.gouv.fr</a> Pas d'installation	
11	<a href="http://draaf.auvergne.agriculture.gouv.fr">http://draaf.auvergne.agriculture.gouv.fr</a> Pas d'installation	
12	<a href="http://draaf.centre.agriculture.gouv.fr">http://draaf.centre.agriculture.gouv.fr</a> Pas d'installation	
13	<a href="http://draaf.paca.agriculture.gouv.fr">http://draaf.paca.agriculture.gouv.fr</a> Pas d'installation	
14	<a href="http://draaf.rhone-alpes.agriculture.gouv.fr">http://draaf.rhone-alpes.agriculture.gouv.fr</a> Pas d'installation	
15	<a href="http://esante.gouv.fr">http://esante.gouv.fr</a> Pas d'installation	
16	<a href="http://legifrance.gouv.fr">http://legifrance.gouv.fr</a> Pas d'installation	
17	<a href="http://m.defense.gouv.fr">http://m.defense.gouv.fr</a> Pas d'installation	
18	<a href="http://urbamet.documentation.developpement-durable.gouv.fr">http://urbamet.documentation.developpement-durable.gouv.fr</a> Pas d'installation	

# Exemple d'utilisation

Modifier l'article < Ronan CHARDONNEAU — WordPress - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils Aide

Portail Boîte de réception (15 1... Lynx documentation.doc Modifier l'article < Rona...

ronan-chardonneau.fr/wp-admin/post.php?post=181&action=edit DuckDuckGo

Most Visited Linux Mint Community Forums Blog News

Ronan CHARDONNEAU 5 12 + Nouveau Afficher l'article SEO Salutations, admin

## Articles

Tous les articles  
Ajouter  
Catégories  
Mots-clés

## Médias

## Liens

## Pages

## Commentaires 12

## Apparence

## Extensions 2

## Utilisateurs

### Modifier l'article

Ajouter

# Quelles solutions de web analytique pour notre gc

Permalien : <http://ronan-chardonneau.fr/quelles-soluti...e-gouvernement/>  
Modifier Obtenir le lien court

Envoyer/Insérer Visuel HTML

**B** *I* ABC [Liste à puces] [Liste à puces inversée] [Citation] [Alignement à gauche] [Alignement au centre] [Alignement à droite] [Lien] [Image] [Table] [Tableau de bord] [Liste à puces] [Liste à puces inversée] [Citation] [Alignement à gauche] [Alignement au centre] [Alignement à droite] [Lien] [Image] [Table] [Tableau de bord]

Paragraphe [Surligné] [Texte en gras] [Texte en italique] [Texte souligné] [Texte barré] [Texte en rouge] [Texte en vert] [Texte en bleu] [Texte en noir] [Texte en blanc] [Texte en gris] [Texte en brun] [Texte en violet] [Texte en rose] [Texte en orange] [Texte en jaune] [Texte en vert] [Texte en bleu] [Texte en noir] [Texte en blanc] [Texte en gris] [Texte en brun] [Texte en violet] [Texte en rose] [Texte en orange] [Texte en jaune]

Cela faisait un petit moment que j'avais envie de me lancer dans cette aventure, c'est chose faite. Je tenais à faire un point sur les solutions de web analytique qui sont installés sur les sites gouvernementaux français.

En effet nos voisins allemands étant très à cheval sur la notion de confidentialité des données, j'ai trouvé intéressant de regarder ce que nous avons de notre côté en terme d'infrastructure.

### Publier

Enregistrer brouillon Aperçu

État : **Brouillon** Modifier

Visibilité : **Public** Modifier

Publier tout de suite Modifier

SEO:  Check

[Déplacer dans la Corbeille](#)

Publier

Menu Modifier l'a... [Télécharg...] linux-seo.o... Terminal Lynx SEO.o... Dossier pe... Dossier pe... titre.txt (~)... -1- 2 !

Awk

# Awk

awk — dont le nom vient des trois créateurs, Alfred Aho, Peter Weinberger et Brian Kernighan — est un langage de traitement de lignes, disponible sur la plupart des systèmes Unix. Il est utilisé pour la manipulation de fichiers textuels pour des opérations de recherches, de remplacement et de transformations complexes. En résumé il affiche les informations telles que vous le souhaitez.

# Awk : Syntaxe

- `awk [-Fs] [-v variable] [-f fichier de commandes] 'program' fichier`
  - F Spécifie les séparateurs de champs
  - v Définit une variable utilisée à l'intérieur du programme.
  - f Les commandes sont lu à partir d'un fichier.

# Exemple concret

```
awk 'BEGIN{IGNORECASE=1;FS="<title>|  
    </title>";RS=EOF} {print $2}'
```

The image shows a screenshot of an Ubuntu desktop environment. A terminal window is open in the foreground, displaying the output of a Lynx command. The terminal output includes statistics for a website scan and the title of the scanned page. A Beamer presentation window is also visible, showing slide 7 of 57. The desktop background is a solid orange color.

```
Terminal
Fichier Fichier Édition Affichage Rechercher Terminal Aide
3 URLs active, 0 URLs queued, 12 URLs checked, runtime 40 seconds
Statistics:
Number of domains: 3
Content types: 0 image, 10 text, 0 video, 0 audio, 5 application, 0 mail and 0 other.
URL lengths: min=22, max=94, avg=47.
That's it. 15 links checked. 0 warnings found. 0 errors found.
Stopped checking at 2013-02-16 09:53:22+002 (44 seconds)
~ $ ^C
~ $ lynx -source www.seo-camp.org | awk 'BEGIN{IGNORECASE=1;
FS="<title>|</title>";RS=EOF} {print $2}'
SEO Camp - L'association des référenceurs
```

Terminal

Édition texte : Paragraphe 1, Ligne 2, Colonne 30 | 1,40 / 4,91 | 24,64 x 12,11 | Diapo 7 / 57 | Standard | 36%

Menu conf - ronanchardon... Téléchargements linux-seo.odp - Libre... Terminal Lynx SEO.odp - Libre... -1 2 !

Tâches Afficher x

- Pages maîtresses
- Mises en page
- Ébauches de tableau
- Animation personnalisée
- Transition

Et si on copiait le tout dans un fichier texte?

The image shows a Linux desktop environment with a terminal window and a Beamer presentation slide. The terminal window is titled "Terminal" and contains the following command:

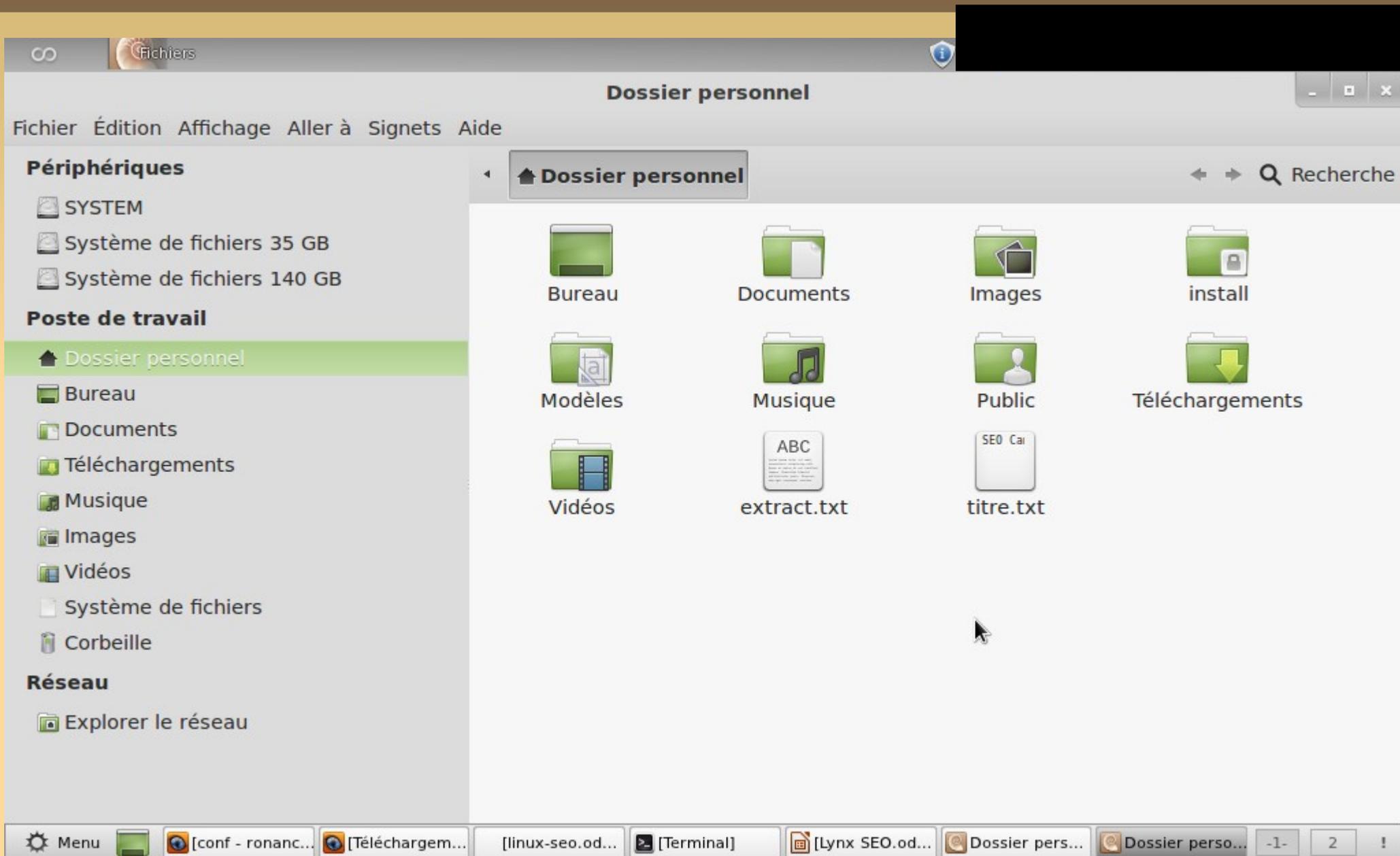
```
cecite@cecite-W250P ~$ lynx -source www.seo-camp.org | awk 'BEGIN{IGNORECASE=1; ES="<title>|</title>".RS=EOF} {print $2}' > titre.txt
```

The Beamer presentation slide is titled "Diapos" and shows a slide number of 69. The slide content is obscured by a large black redaction box. The Beamer navigation pane on the right shows the following structure:

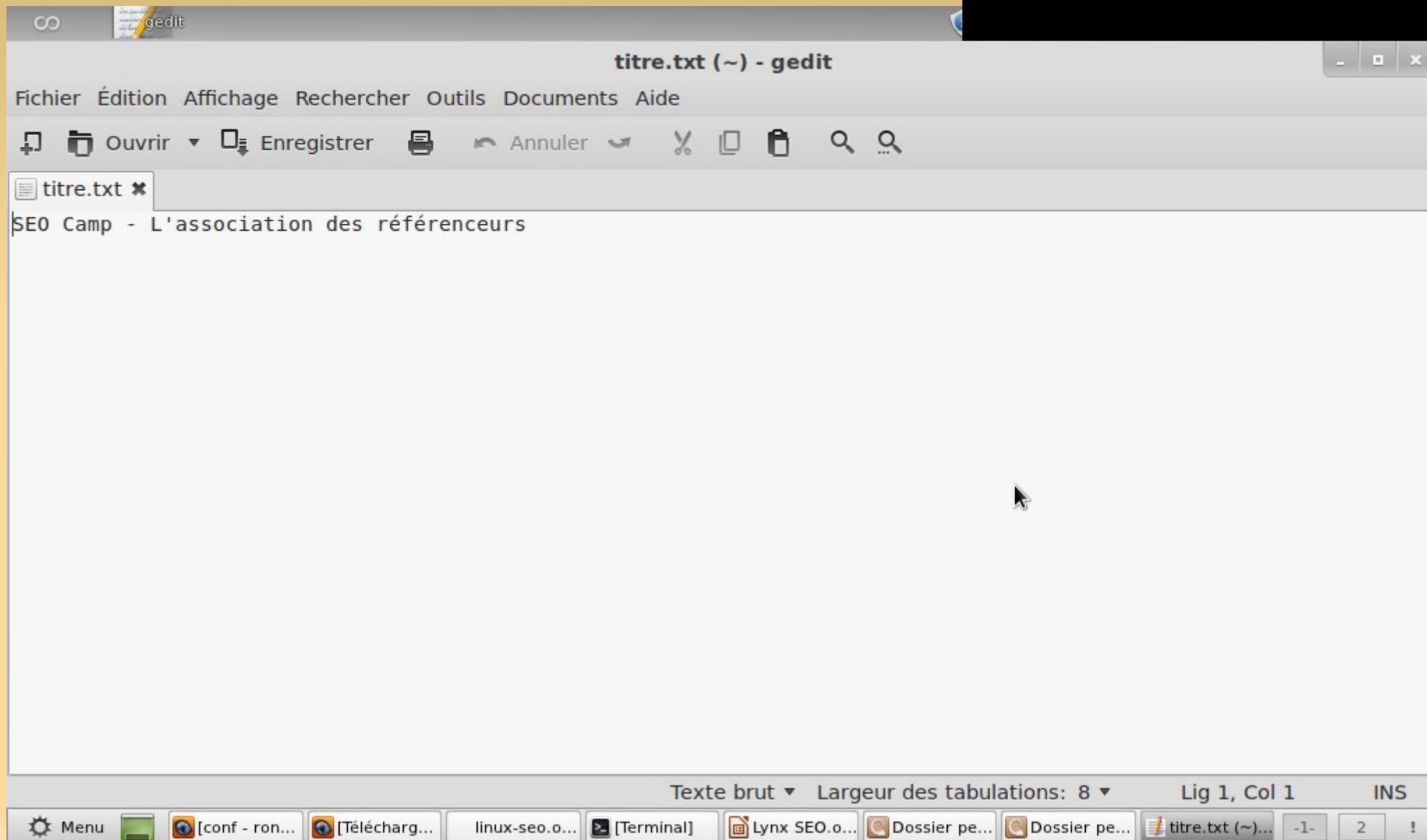
- Pages maîtresses
- Mises en page
  - [Slide thumbnail]
  - [Slide thumbnail]
  - [Slide thumbnail]
  - [Slide thumbnail]
  - [Slide thumbnail]
- Ebauches de tableau
- Animation personnalisée
- Transition

The desktop environment includes a top panel with system icons (network, battery, volume) and the date/time "sam. 10:05 Cécile". The bottom panel shows the taskbar with the following windows: "Menu", "conf - ronanchardon...", "Téléchargements", "linux-seo.odp - Libre...", "Terminal", and "Lynx SEO.odp - Libre...". The system tray shows "Diapo 69 / 82", "HumanUbuntu", and "36%".

# Concrètement



# Concrètement



Sed

# Sed

sed (abréviation de Stream EDitor, « éditeur de flux ») est, comme awk, un programme informatique permettant d'appliquer différentes transformations prédéfinies à un flux séquentiel de données textuelles.

# Ajouter un mot au début de chacune des lignes d'un fichier

```
sed 's/^/mon mot/' infile
```

# Exemple concret

The image shows a Linux terminal window with a file editor open. The terminal displays the following command and output:

```
$ sed 's/^/SEO Camp Day Caen /' titre.txt
SEO Camp Day Caen SEO Camp - L'association des référenceurs
SEO Camp Day Caen
```

The terminal window has a menu bar with "Fichier", "Édition", "Affichage", "Rechercher", "Terminal", and "Aide". The file editor window has a menu bar with "Fichier", "Édition", "Affichage", and "Recherche". The terminal window title is "Terminal". The file editor window title is "titre.txt". The terminal window shows the command being executed and the output of the command. The output shows the first line of the file being replaced with "SEO Camp Day Caen" followed by the original content of the first line. The rest of the file remains unchanged.

# Exemple d'utilisation

Idéal si l'on souhaite ajouter à la volée un tracker HTML de Google Analytics sur plusieurs articles pour un import sur du Wordpress

`sed -i 's/monmot/autremot/g' mon fichier`

# Autres outils

Pdftk

# Convertir plusieurs pdf en un seul

```
pdftk Analytics\ A\ -\ Profil\ général\ Chaines\  
Compare\ Navigateur\ et\ système\ d'exploitation\  
20121202-20130101.pdf Analytics\ A\ -\ Profil\  
général\ Chaines\ Compare\ Réseau\ 20121202-  
20130101.pdf cat output final2.pdf
```

# Image Magick

# Image Magick

ImageMagick est un logiciel en ligne de commande très puissant de manipulation d'images dans pratiquement tous les formats existants.

# Exemple

Je cherche à créer un nombre incalculable d'images en fonction d'une variable.

```
convert -extent 297x108 -background transparent  
-fill white -font Arial -pointsize 30 -distort Arc 120  
-trim +repage -gravity center label:'165 65 R14'  
165-65-r14.png
```



Fichiers



### Dossier personnel

Fichier Édition Affichage Aller à Signets Aide

#### Périphériques

- SYSTEM
- Système de fichiers 35 GB
- Système de fichiers 140 GB

#### Poste de travail

- Dossier personnel**
- Bureau
- Documents
- Téléchargements
- Musique
- Images
- Vidéos
- Système de fichiers
- Corbeille

#### Réseau

- Explorer le réseau

↑ Dossier personnel

Recherche



Bureau



Documents



Images



install



Modèles



Musique



Public



Téléchargements



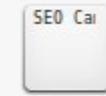
Vidéos



165-65-r14.png



extract.txt

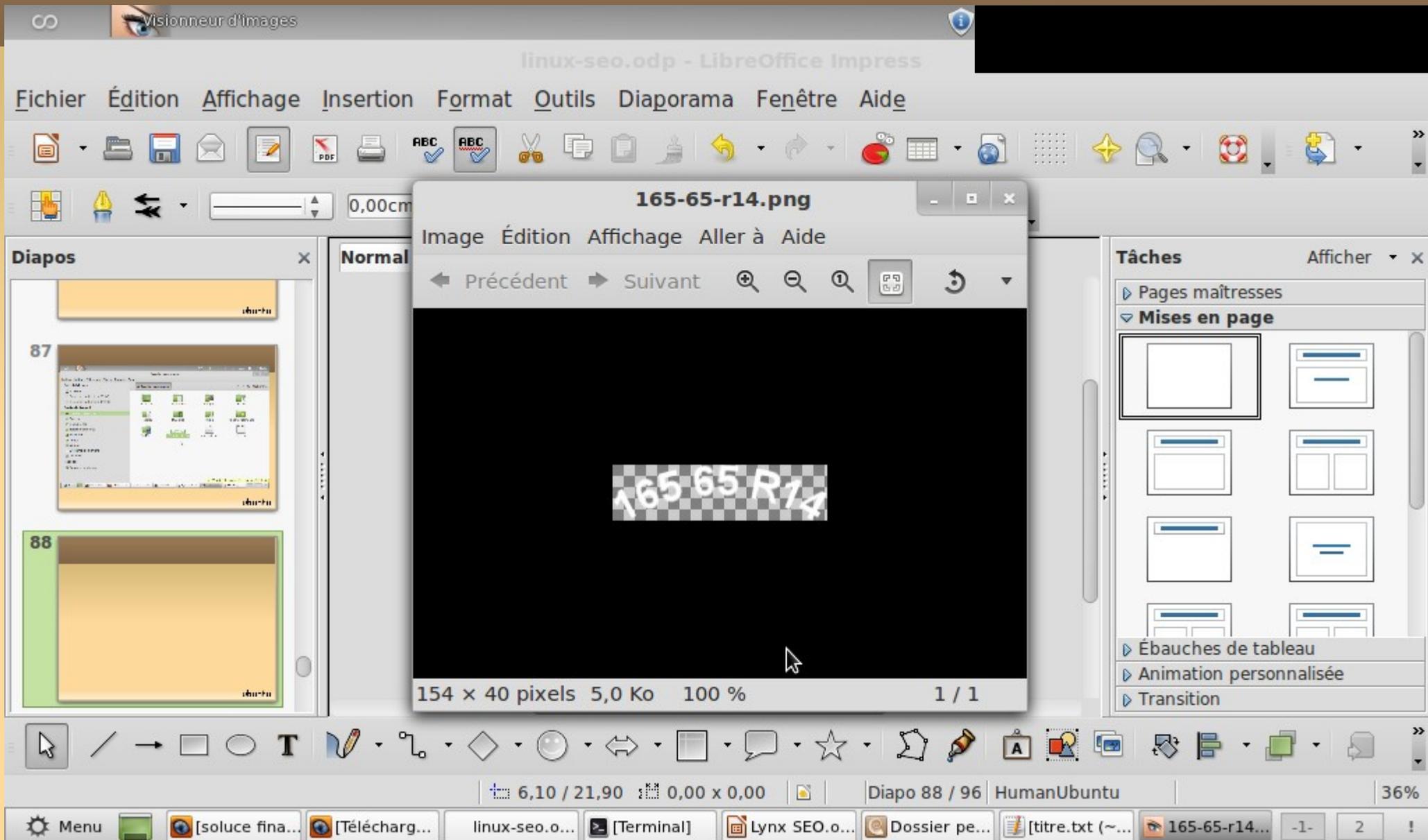


titre.txt



« 165-65-r14.png » sélectionné (5,1 ko)

- Menu
- [soluce finale ...]
- [Téléchargem...]
- [linux-seo.od...]
- [Terminal]
- [Lynx SEO.od...]
- Dossier pers...
- [titre.txt (~) - ...]
- 1-
- 2
- !



# Autre exemple: fusionner des images

```
composite -gravity center 165-65-r14.png modele-  
pneu.png fichier.png
```



### Que signifie 140-60 R 13 sur un pneu auto?

décembre 9, 2012

Dimensions de pneu

Laisser un commentaire



Sur n'importe quel pneu auto de dimensions 140/60 R 13, **140** représente la **largeur du pneumatique**, elle est exprimée **en centimètres**. Vous verrez que cet ordre est rarement inversé, en premier en général il s'agira de la plus grande valeur, séparée par / pour la seconde valeur, puis une lettre, pour une autre valeur après. **60** fait référence à la **hauteur du pneu**, elle est tout comme la largeur exprimée **en centimètres**. **R** fait référence à la manière dont les  **fils qui composent le pneu**  sont organisés, ici R correspond à **radiale**. **13** est le **diamètre de la roue**, attention cette dernière valeur est exprimée **en pouces et non pas en centimètres**. Pour vous en souvenir vous pouvez vous rappelez la combinaison suivante de lettres: LHD pour Largeur Hauteur Diamètre.

Où acheter son pneu auto 140/60 R 13? Pneu Suisse

Bloguez sur WordPress.com. Thème : Blogum.

Follow

# Linkchecker

Vous connaissez Xenu, Linkchecker est son  
cousin sur Linux

Linkchecker [www.seo-camp.org](http://www.seo-camp.org)

The image shows a Linux desktop environment with a terminal window and a Beamer presentation window. The terminal window is titled "Terminal" and shows the output of the LinkChecker 7.0 command-line tool. The Beamer presentation window is titled "HumanUbuntu" and shows a slide with a grid of thumbnail images.

**Terminal Output:**

```
cecite@cecite-N250P ~$ linkchecker www.seo-camp.org
LinkChecker 7.0 Copyright (C) 2000-2011 Bastian Kleineidam
LinkChecker comes with ABSOLUTELY NO WARRANTY!
This is free software, and you are welcome to redistribute it
under certain conditions. Look at the file `LICENSE' within this
distribution.
Get the newest version at http://linkchecker.sourceforge.net/
Write comments and bugs to http://sourceforge.net/projects/linkchecker/support

65 Start checking at 2013-02-16 09:52:38+002
10 URLs active, 114 URLs queued, 4 URLs checked, runtime 5 seconds
10 URLs active, 114 URLs queued, 4 URLs checked, runtime 10 seconds
10 URLs active, 114 URLs queued, 4 URLs checked, runtime 15 seconds
10 URLs active, 114 URLs queued, 4 URLs checked, runtime 20 seconds
^CWARNING MainThread user interrupt; waiting for active threads to finish
WARNING MainThread another interrupt will exit immediately
66 INFO MainThread These URLs are still active:
INFO MainThread http://www.seo-camp.org/wp-content/themes/thematic.0.9/thematic/
library/scripts/superfish.js
INFO MainThread http://www.seo-camp.org/
INFO MainThread http://www.seo-camp.org/xmlrpc.php
INFO MainThread http://www.seo-camp.org/wp-content/themes/thematic.0.9/thematic/
library/scripts/supersubs.js
INFO MainThread http://www.seo-camp.org/xmlrpc.php?rsd
```

**Beamer Presentation:**

The Beamer presentation window shows a slide with a grid of thumbnail images. The slide title is "HumanUbuntu". The slide content includes a grid of thumbnail images and a list of navigation options:

- Pages maîtresses
- Mises en page
- Ebauches de tableau
- Animation personnalisée
- Transition

The desktop environment includes a terminal window, a Beamer presentation window, and a taskbar at the bottom. The taskbar shows the following applications: Menu, Lynx documentation..., Téléchargements, linux-seo.odp - Libre..., and Terminal. The system tray shows the date and time: 1,30 / 0,00, 25,20 x 18,7, Diapo 66 / 73, HumanUbuntu, and 36%.

# Autres petites fonctions de base

Split -b fichier.txt

Ou -b indique la taille de chacun des fichiers

Rassembler deux fichiers en 1

Cat fichier 1 fichier 2 > fichier final

Compter le nombre de mots ou le nombre de caractères:

wc -w ou wc -m

Pour aller plus loin



type here to grep the archive...

## Commands matching w3m

Commands matching w3m from **all time** sorted by **votes** **Go**

Terminal - Commands matching w3m - 15 results

```
$ /usr/bin/man $* || w3m -dump http://google.com/search?q="$*" \&btNI | less
```

### Extended man command

This microscript looks up a man page for each word possible, and if the correct page is not found, uses w3m and Google's "I'm feeling lucky" to output a first possible result. This script was made as a result of an idea on a popular Linux forum, where users often send other people to RTFM by saying something like "man backup" or "man ubuntu one". To make this script replace the usual man command, save it as ".man.sh" in your home folder and add the following string to the end of your .bashrc file:  
alias man='~/man.sh'

9 Up  
Down

2010-10-05 13:51:39  
User: **d1337r**  
Tags: **man google**

Show sample output | Comments (0) | Add to favourites | Report as malicious | Submit alternative | Report as a duplicate

```
$ pdftohtml -i -stdout FILE.pdf | w3m -T text/html
```

### Read PDFs in the command line

Turns a PDF into HTML (without images) and prints it to the standard out which is picked up and interpreted by w3m.

4 Up  
Down

2010-11-03 12:01:01  
User: **Haegin**  
Tags: **pdf w3m**

Comments (0) | Add to favourites | Report as malicious | Submit alternative | Report as a duplicate

```
$ w3m -no-cookie http://m.espn.go.com/nfl/scoreboard?|sed 's/ Final/ : Final/g'|sed 's/ F\\// : F\\//g'|sed 's/, / : /g'|grep -i ':'
```

### Get NFL/MLB Scores/Time

change the nfl in the url to mlb or nba to get those score/times as well

2 Up  
Down

2010-11-15 01:18:19  
User: **SQUIDUX**  
Functions: **grep sed**  
Tags: **sed grep w3m**

Show sample output | Comments (3) | Add to favourites | Report as malicious | Submit alternative | Report as a duplicate

```
$ w3m -dump_source http://www.onelinerz.net/random-one-liners/1/ | awk ' /.*<div id="oneliner_[0-9].*/ {while (! /\>/) { gsub("\n", ""); getline; }; gsub (/<[^>][^>]*>/, "", $0); print $0}'
```

### What's this?

Hide

commandlinefu.com is the place to record those command-line gems that you return to again and again.

Delete that bloated snippets file you've been using and share your personal repository with the world. That way others can gain from your CLI wisdom and you from theirs too. All commands can be commented on, discussed and voted up or down.

If you have a new feature suggestion or find a bug, please get in touch via <http://commandlinefu.uservoice.com/>

### Get involved!

You can sign-in using **OpenID credentials**, or register a traditional username and password.

OpenID

or

Sign in

Register

First-time OpenID users will be automatically assigned a username which can be changed after signing in.

LIVRAISON GRATUITE

# Livres

- Grep O'Reilly
- Awk O'Reilly
- Sed O'Reilly
- Bash O'Reilly

