

## PR Sculpting

SEO Camp'us -4 et 5 février 2009

**Philippe YONNET**

Directeur du pôle métiers – Aposition  
Président de l'association SEOCamp



# L'importance des CIR

CIR : Critères indépendants de la requête  
(query independent criteria)

Critères utilisés pour améliorer la pertinence « perçue » par les utilisateurs une fois que les autres critères ont permis :

- d'identifier les pages constituant de bonnes réponses
- de classer ces pages en fonction de scores de proximité avec la requête

Parmi ces critères on trouve tous les « ranks » : Pagerank, Trustrank, PR thématiques, Badrank, Localrank etc...

On peut aussi utiliser des critères comme le trafic sur les pages ou le taux de bounce (critères d'usage).

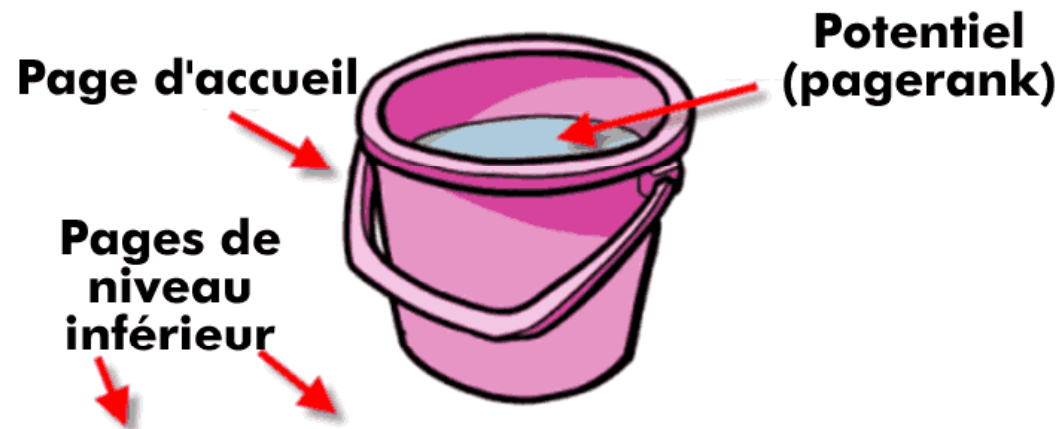
# La structure hypertexte influence fortement la plupart des CIR

Liens  
externes



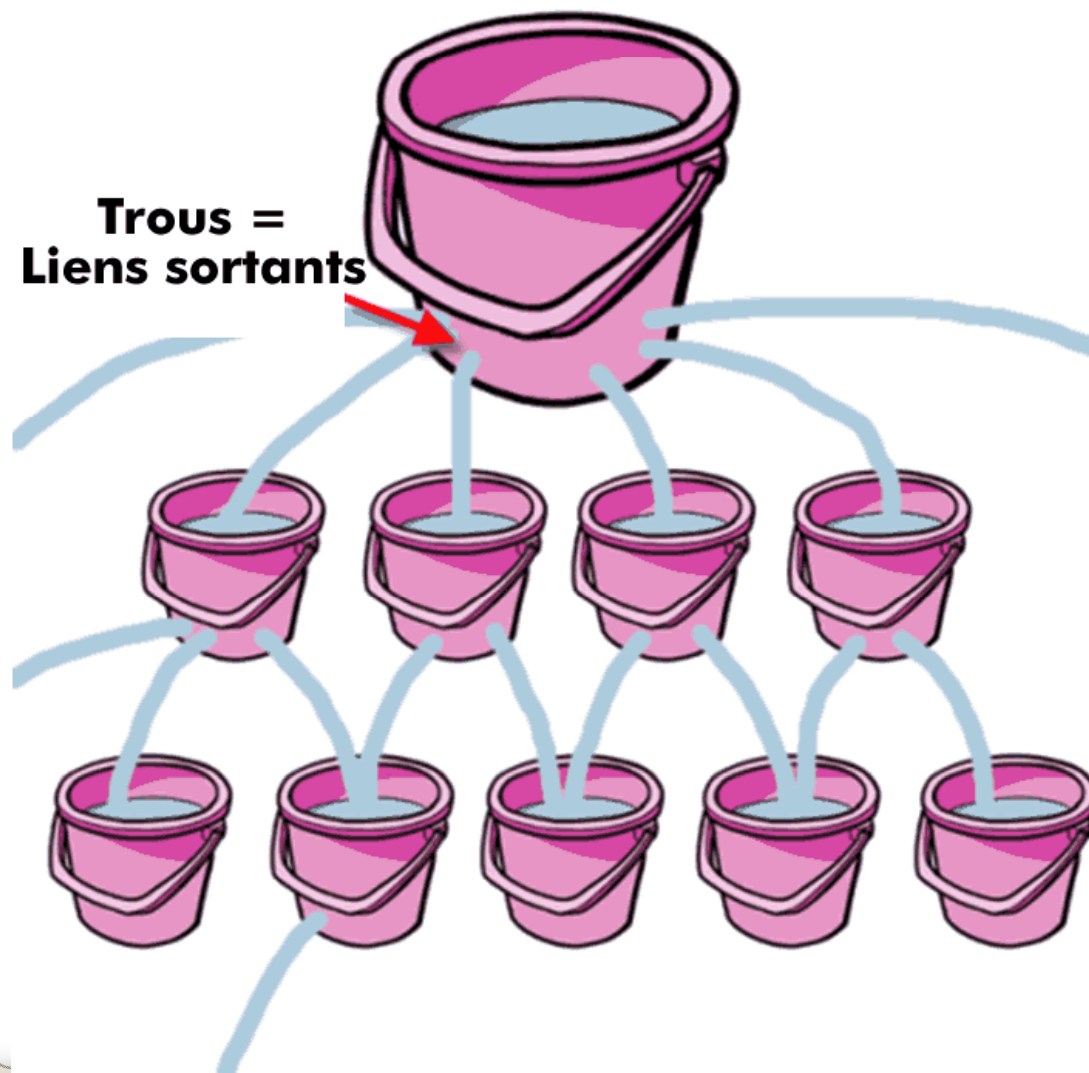
Liens  
internes

## Les principes du PR sculpting



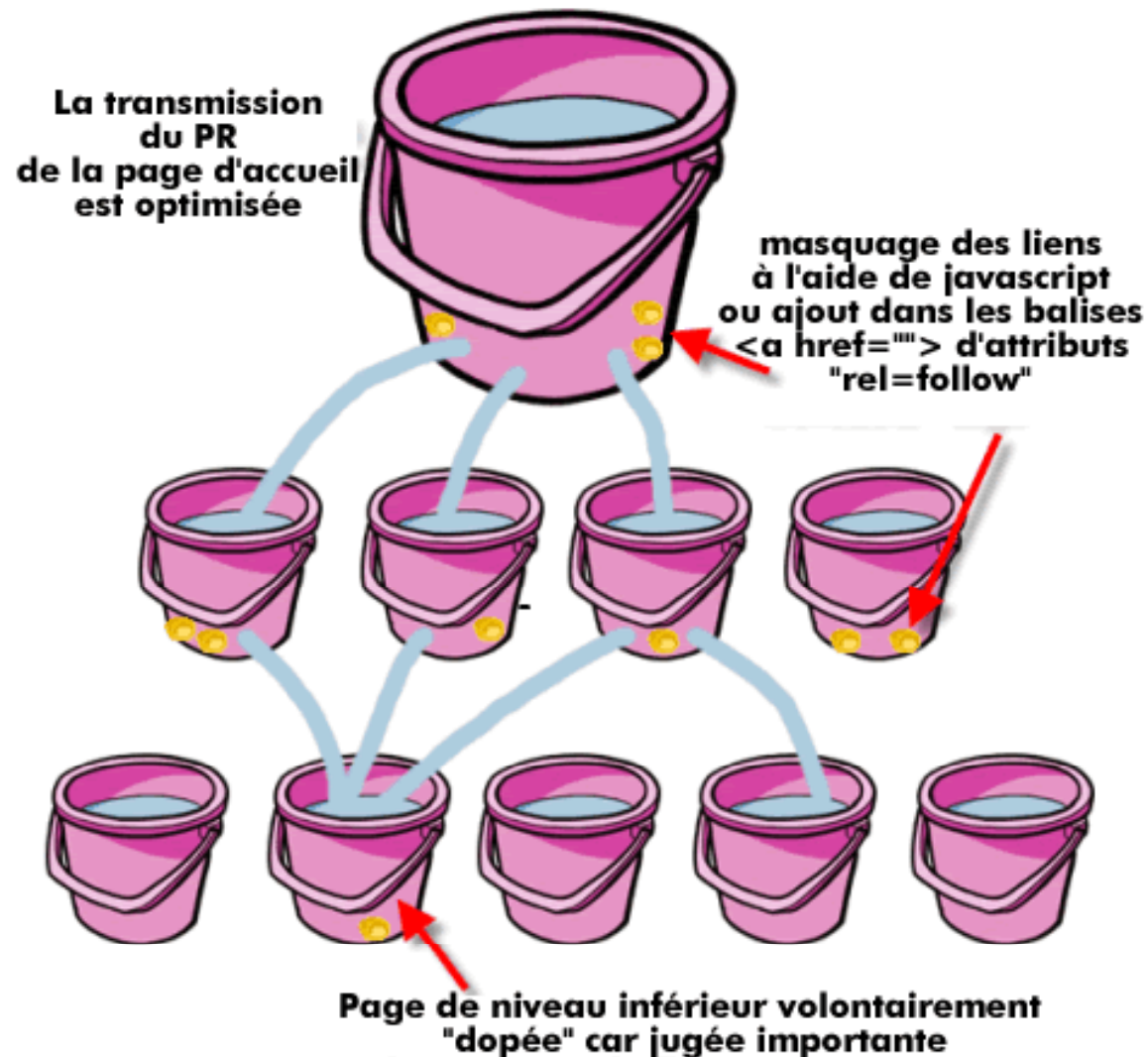
Illustrations  
adaptées du blog de  
l'agence de  
netmarketing :  
[www.evisibility.com](http://www.evisibility.com)

## Les principes du PR sculpting



Illustrations  
adaptées du blog de  
l'agence de  
netmarketing :  
[www.evisibility.com](http://www.evisibility.com)

## Les principes du PR sculpting



Illustrations  
adaptées du blog de  
l'agence de  
netmarketing :  
[www.evisibility.com](http://www.evisibility.com)

## Qu'est-ce que le "link juice" La notion de potentiel de pagerank

- La somme de tous les PR des pages d'un site constitue le "potentiel de PR" de ce site

$$P_{pr} = \sum \text{PR des pages}$$

## La "formule" du potentiel

- Le potentiel de PR a été "théorisé" à partir d'un travail mathématique sur la formule du PR par des chercheurs Italiens :

"Inside Pagerank"

Monica Bianchini, Marco Gori, Franco Scarselli

Université de Sienne (publié en 2005)

$$P = |I| + P_{in} + P_{out} - P_{Dp}$$

## La "signification" de la formule

$$P = | I | + P_{in} - P_{out} - P_{Dp}$$

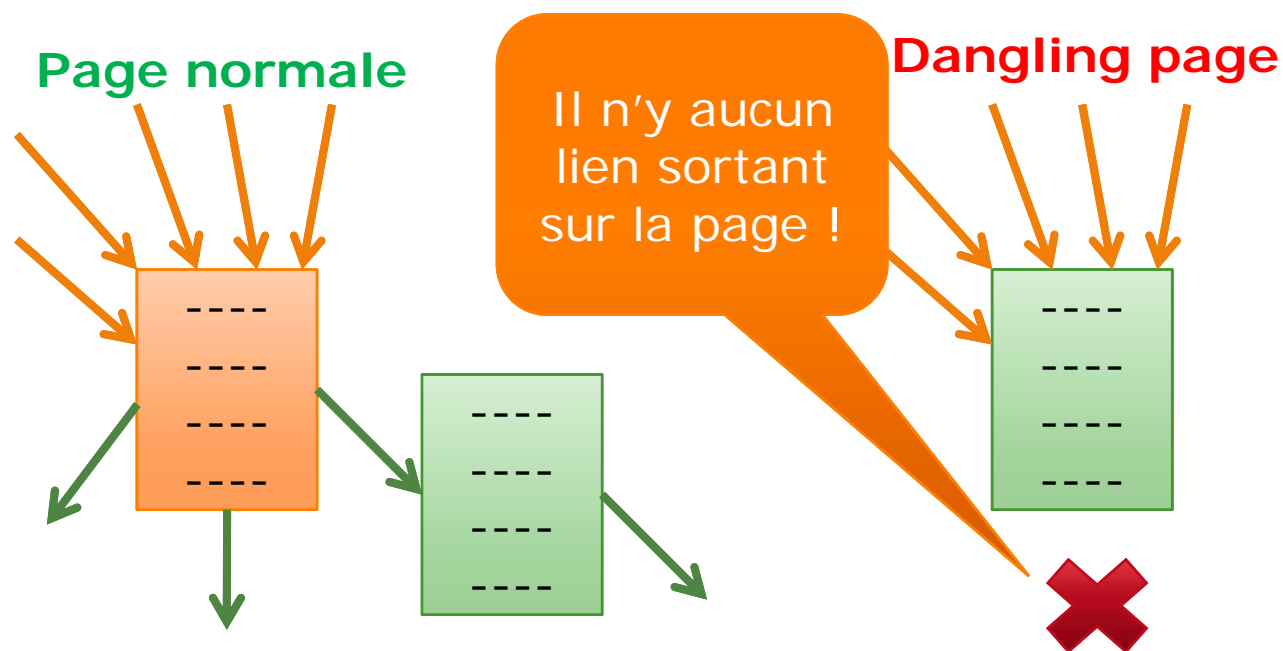
- $P$  = Potentiel de PR total
- $I$  = Potentiel "interne" du site – dépend uniquement du nombre de pages
- $P_{in}$  = Potentiel provenant des backlinks
- $P_{out}$  = Potentiel perdu par les liens sortants
- $P_{Dp}$  = Potentiel perdu par les pages pendantes

## Euh... c'est quoi une « dangling page » ?

- Traduction : page "ballante", "en suspens"
- Transposition : page cul de sac
- Définition : page qui ne contient aucun lien vers d'autres pages, mais qui reçoit des liens

## Euh... c'est quoi une « dangling page » ?

- Traduction : page "ballante", "en suspens"
- Transposition : page cul de sac
- Définition : page qui ne contient aucun lien vers d'autres pages, mais qui reçoit des liens



## De l'utilité de calculer le PR interne...

Path	Outbound...	PageRank	Approximate ...	Backlink P...	Increment
/blog/		0(0.0)			
/blog/2006/12/13/why-google-will-not-move-away-from-pagerank.html		0(0.0)			
/blog/2006/11/24/pagerank-doesnt-matter-is-now-officially-anthony-lewis.html		0(0.0)			
/blog/2007/02/19/why-duplicate-content-causes-supplemental-links.html		0(0.0)			
/blog/2007/06/07/got-supplementals-accepting-pagerank-is-optional.html		0(0.0)			
/blog/category/google/		0(0.0)			
/blog/2007/04/29/if-you-lack-motivation-read-this.html		0(0.0)			
/blog/2007/04/16/21-reasons-why-anti-nofollow-seos-cant-think.html		0(0.0)			
/blog/2007/03/07/how-does-6-seos-miss-the-obvious.html		0(0.0)			
/blog/2007/01/05/seo-myth-there-is-no-duplicate-content-penalty.html		0(0.0)			
/blog/2007/05/22/how-will-people-find-your-site-if-search-engines-cant.html		0(0.0)			
/blog/category/seo/		0(0.0)			
/blog/category/adult-webmastering/		0(0.0)			
/blog/category/duplicate-content/		0(0.0)			
/blog/category/google/google-sitemaps/		0(0.0)			
/blog/category/googlebot/		0(0.0)			
/blog/2007/06/22/hey-it-just-sounds-to-me-like-you-need-to-use-it.html		0(0.0)			
/blog/category/un-categorized/		0(0.0)			
/blog/category/seo/mod-rewrite/		0(0.0)			
/blog/category/on-page-optimization/		0(0.0)			
/blog/category/supplemental-bot/		0(0.0)			
/blog/category/google/supplementals/		0(0.0)			
/blog/2007/06/16/dbc-classnotfoundexception-netbeans-class-not-found.html		0(0.0)			
/blog/2007/05/14/stop-blowing-money-on-text-link-ads.html		0(0.0)			
/blog/2007/03/20/free-seo-course-offering-expert-training-over-100-days.html		0(0.0)			
/notes/supplementals.html		0(0.0)			
/		0(0.0)			
/blog/contact-me/		0(0.0)			
/blog/googlers/		0(0.0)			
/blog/seo-blogs/		0(0.0)			
/blog/seo-sem-internet-marketing-forums/		0(0.0)			
/blog/about/		0(0.0)			
/blog/2006/05/25/googlebot-refreshing-supplementals.html		0(0.0)			
/blog/2006/09/27/google-sitemap-reporting-404s-under-sun.html		0(0.0)			
/blog/2007/02/25/youtube-taking-down-more-music-videos.html		0(0.0)			
/blog/2007/01/23/long-tail-de-jour-whats-the-point-in-adding-more.html		0(0.0)			
/blog/2007/04/17/googles-motives-are-selfish-so-are-yours-are.html		0(0.0)			
/blog/2006/10/12/matt-cutts-pagerank-primary-factor-determined.html		0(0.0)			
/blog/2007/02/03/results-dont-always-say-theyre-personalized.html		0(0.0)			
/blog/2006/09/17/supplemental-edges-fuzzier-than-ever.html		0(0.0)			

<http://www.seo4fun.com/php/pagerankbot.php>

Indicateurs :

Nb et poids des liens entrants

Nb de liens sortants

Ratio liens entrants / liens sortants

Ratio liens internes / liens externes

## Le bot herding

### **Définition :**

Ensemble des actions visant à diriger les crawlers des moteurs de recherche vers les pages que l'on souhaite indexer, et uniquement les pages que l'on souhaite indexer.

## Le bot herding

# Etymologie

Herd : troupeau en anglais

Herding : conduite d'un troupeau (pour la transhumance)

### Le parallèle avec les crawlers ?

On ne peut pas les diriger précisément, mais leur indiquer une direction à suivre et les empêcher d'aller dans certaines zones.

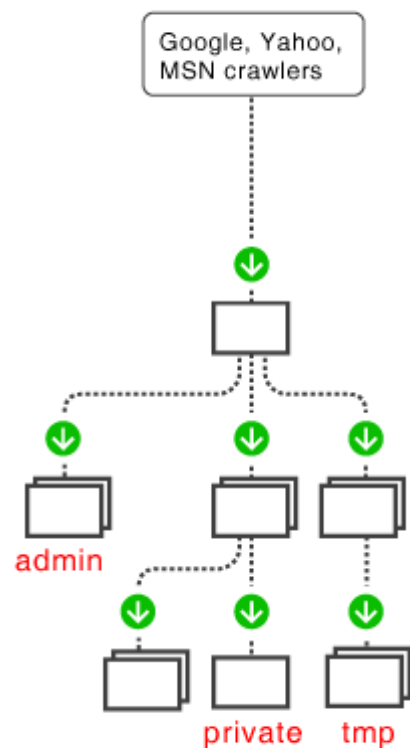
Référenceur = berger (ou border colley)

Crawler = mouton



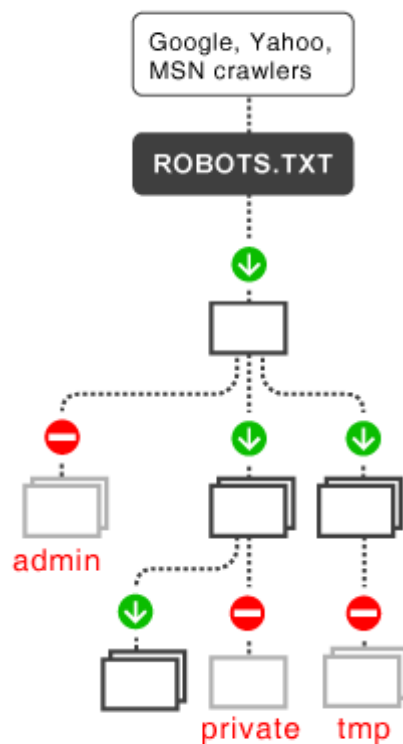
# Le bot herding

Site without Robots.txt



© 2007 Eilliance, Inc.

Site with Robots.txt

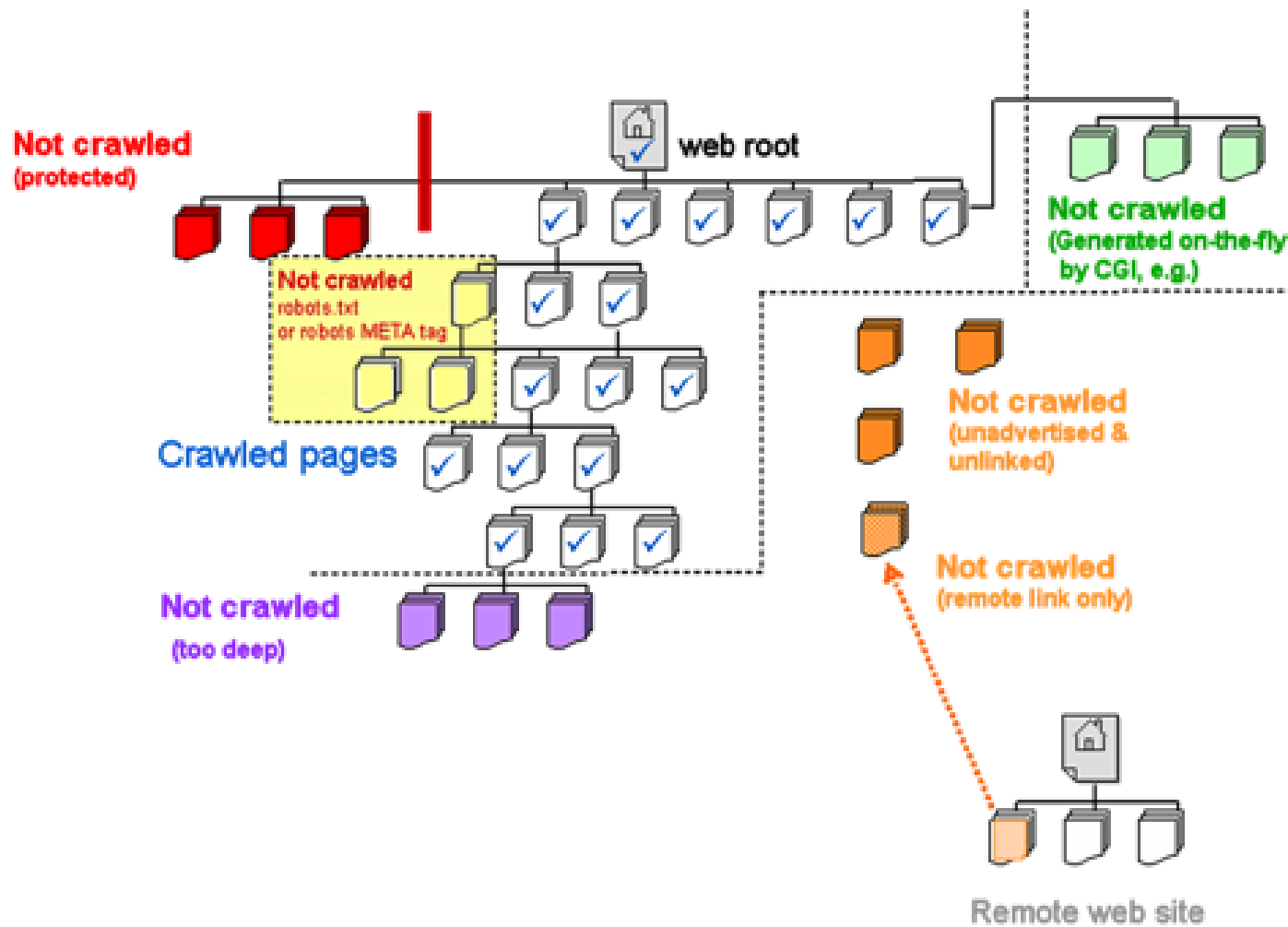


La structure « vue par le moteur » n'est pas la structure réelle.

Parfois c'est volontaire :

Rôle du fichier robots.txt

# Parfois, ce n'est pas volontaire !



## Le bot herding : les leviers principaux

### Crawlabilité

- Supprimer toute barrière technique empêchant un crawl optimal
- Gérer l'accessibilité par un lien, et la notion de profondeur

### Duplicates

- Eliminer les sources de duplication de contenu et les problèmes d'urls
- Gérer les « near duplicates »

### Canalisation

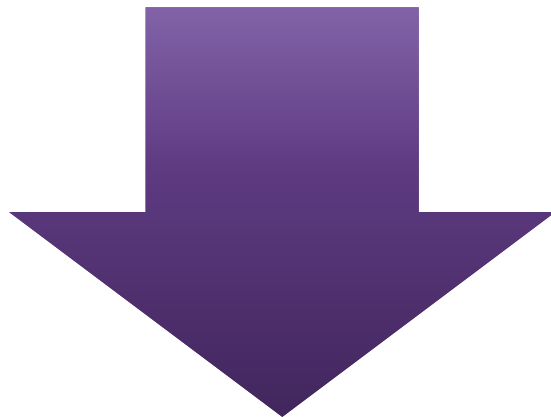
- Blocage du crawl de certaines urls : robots.txt, javascriptage, nofollow, noindex, retrait des urls ...
- Action sur les critères d'importance priorisant le crawl

## Une notion à ne pas confondre avec le « PR sculpting »



### Bot herding

- 100% TOS compliant
- Valable pour tous les moteurs
- Toujours utile, toujours efficace
- Ne règle pas que des problèmes SEO
- Pas d'effet de bord si travaillé avec rigueur



### PR sculpting

- Manipulation de l'algo
- Dirigé vers Google principalement
- Cible uniquement le SEO, et un seul critère (la gestion du flux de pagerank )
- **Dangereux et antiproductif si mal maîtrisé !!**

## Le « siloing »

# Définition :

Méthodologie visant à créer des structures hypertextes « verticalisées » sur des sites où cette structure est fondamentalement absente :

Ex : Blogs, Forums

## Le « siloing »

# Etymologie

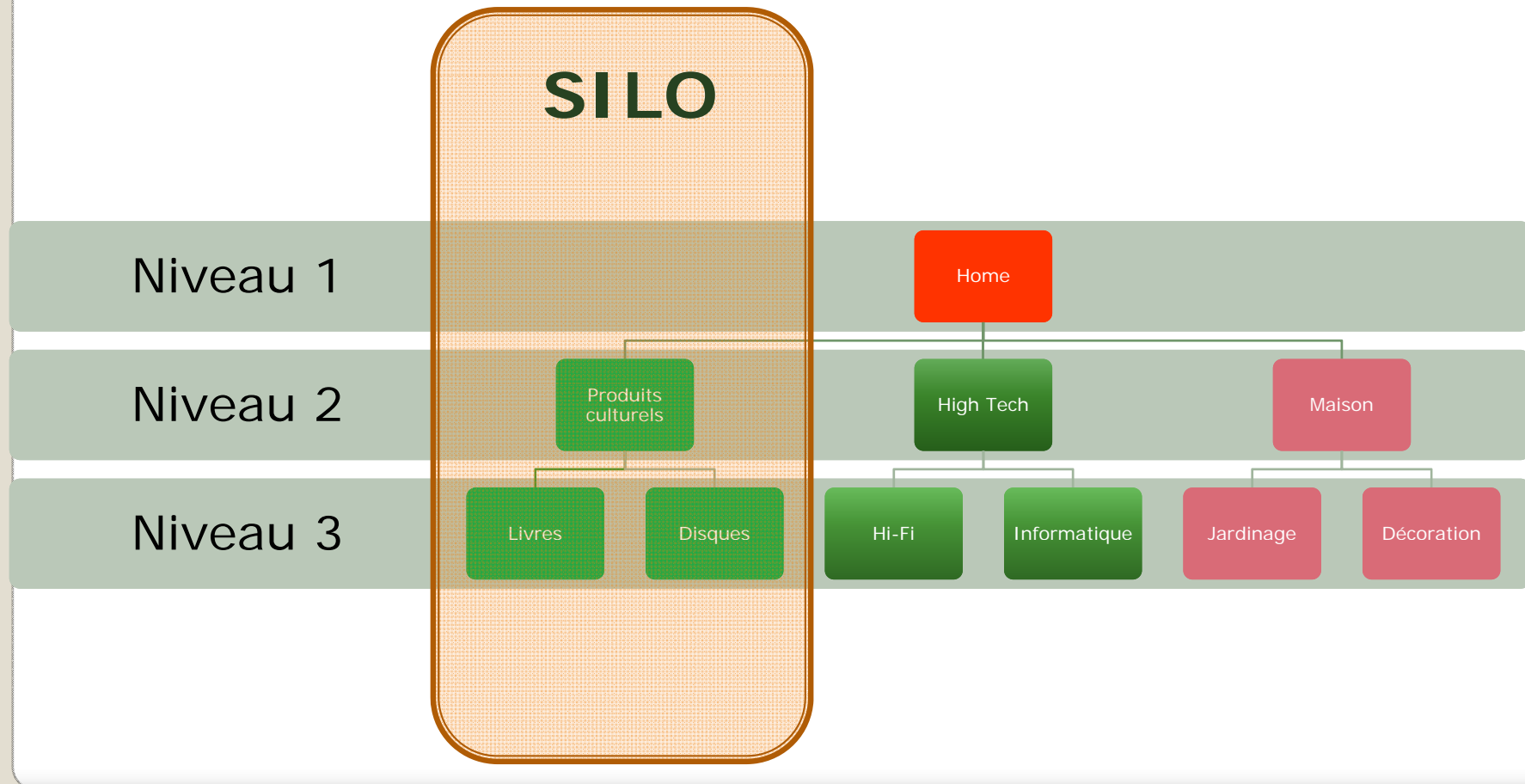


Les silos sont des structures verticales qui se remplissent par le haut...

Par analogie, le terme « silo » est souvent utilisé en jargon marketing US pour décrire un « secteur d'activité »

## Le « siloing »

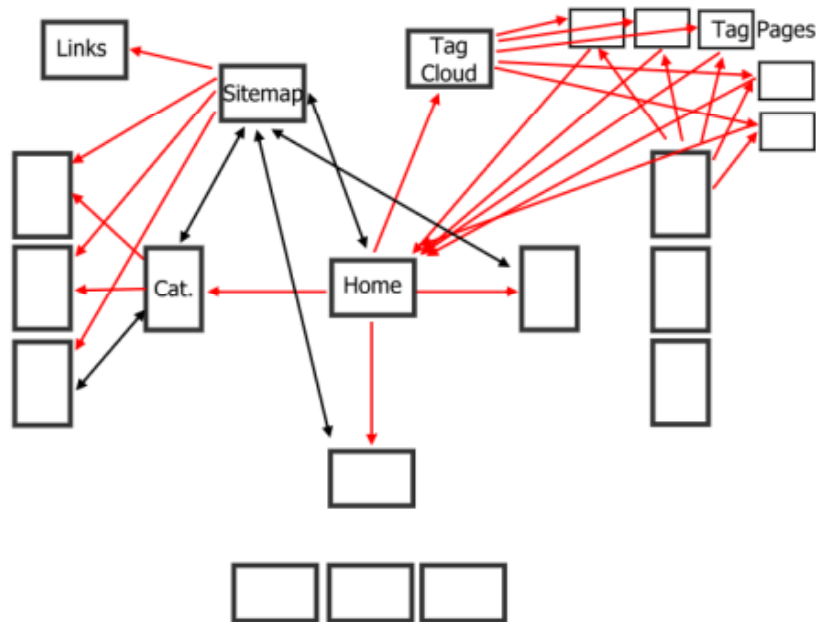
# Structure en silo sur un site web



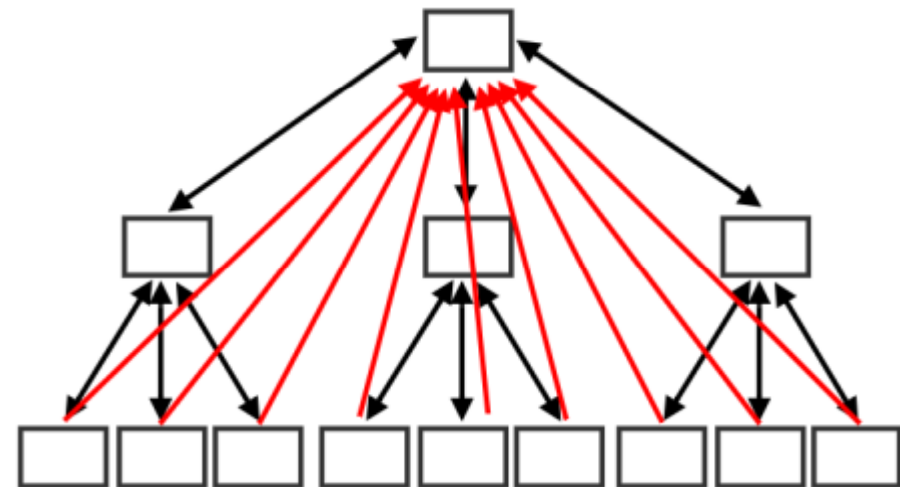
# Le constat : la structure hypertexte de certains sites n'est pas optimale

## Cas d'école : les blogs

AVANT

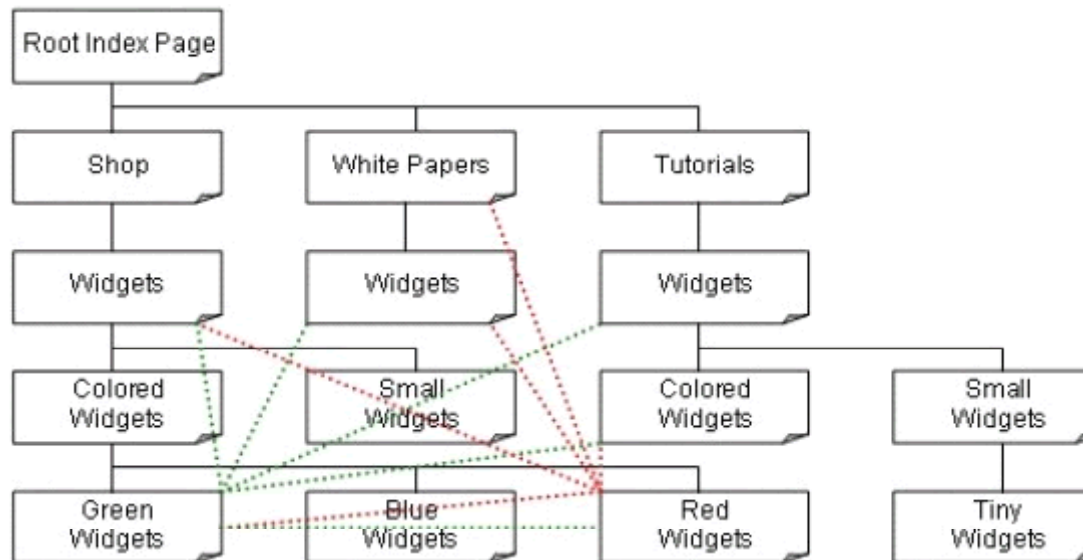


APRES



## Le constat : la structure hypertexte de certains sites n'est pas optimale

### Cas d'un site organisé mais pas en structure verticale





## Structure de linking décrite par Bruce Clay Inc.

### Principe

La home pointe vers les sous homes qui définissent les silos.  
Des liens relient tous les silos entre eux  
Chaque sous home tête de silo pointe vers les catégories enfant  
Les catégories sont toutes reliées entre elles  
Les catégories pointent vers les produits  
Les produits sont tous reliés entre eux

Chaque page est reliée à la home  
Chaque page dispose de liens vers les pages du niveau inférieur et  
les pages de même niveau du silo  
Chaque page est reliée au niveau supérieur

Les silos sont faiblement reliés entre eux

Euh... ça ne vous rappelle rien ?

## La méthode de Bruce Clay

Elle s'appuie sur l'idée (partiellement vraie, partiellement fausse) que l'efficacité du silo provient du recentrage sémantique du silo : tous les contenus du silo sont « on topic » et les anchor texts aussi.

Bruce Clay travaille beaucoup le wording, les mots clés, et la « thématisation »

Mais elle fait fi des effets sur le PR interne, et la structure préconisée ne marche pas toujours, par méconnaissance de certains mécanismes liés à la diffusion du PR.

## Pourquoi ça marche, le siloing ?

La structure en silo maximise le pagerank localement à l'intérieur des silos (le promeneur aléatoire « reste » dans le silo une fois qu'il est tombé dedans)

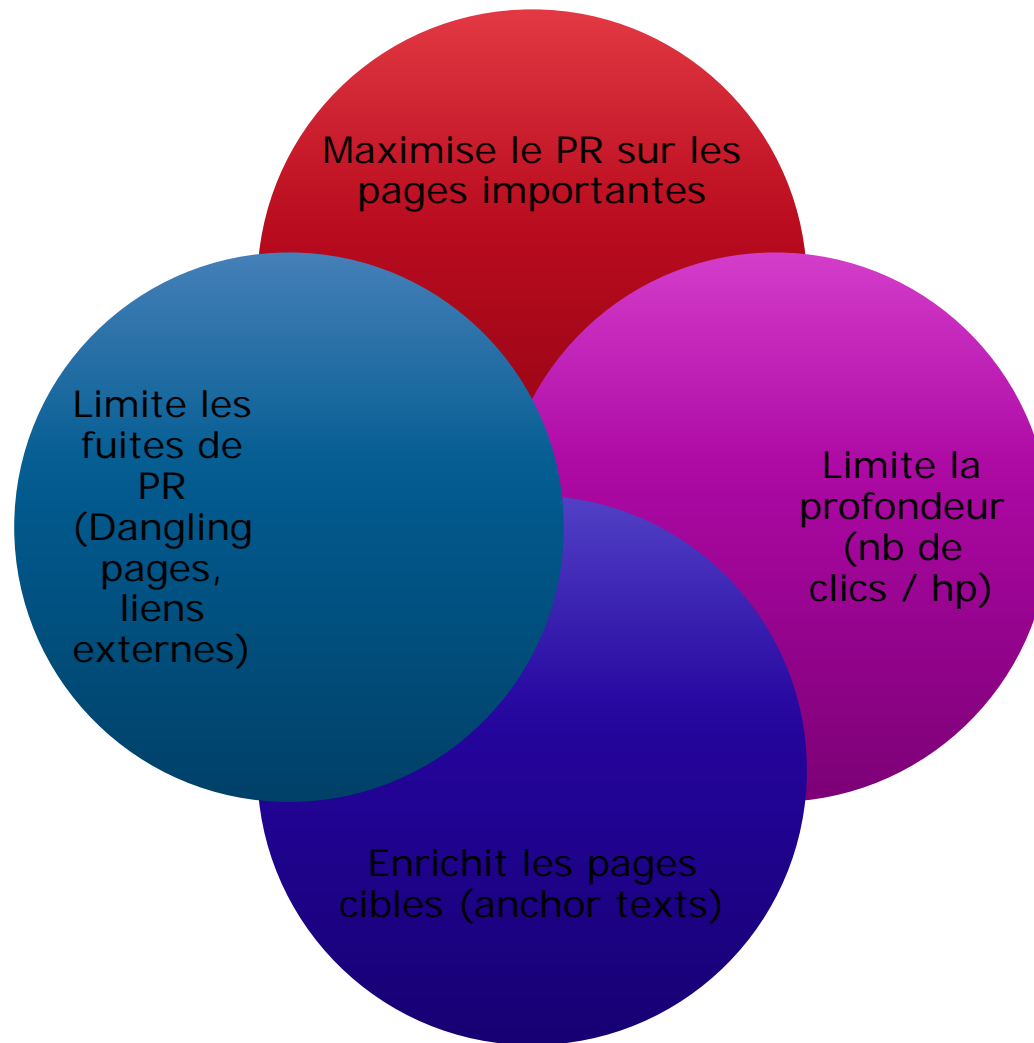
Par contre il faut choisir entre des structures qui maximisent le pagerank :

- en haut du silo (les rubriques, les catégories principales)
- en bas du silo (les produits, les articles)
- au milieu du silo (les sous rubriques, les sous catégories)

La plupart des structures intuitives maximisent le PR en haut des silos, l'effet de la verticalisation n'est donc pas toujours celui recherché quand on ne maîtrise pas la technique à fond.

=> Il faut tester les structures !

## Quelle est la meilleure structure hypertexte ?



**Merci !**

**Et maintenant, si vous avez des questions,  
n'hésitez pas !**