
Audit Technique SEO

Site analysé : pausebento.fr

Johann Corbel Consulting

26 rue George Sand 75016 Paris

<http://www.johanncorbel.fr>

N° tél. : 06.83.85.82.30

N° RCS : 520 744 434 00033

1 Introduction

1.1 Objectif de ce rapport

Ce document a pour objectif de vous présenter les résultats de l'audit technique (sous l'angle du référencement naturel) du site **pausebento.fr**. Vous y trouverez une série de recommandations avec de nombreuses explications, ainsi que des annexes contenant la liste exhaustive de tous les éléments à améliorer sur le site.

L'audit du site a eu lieu le **22 décembre 2015** et a porté sur **261 URL**. Toutes les analyses qui en découlent doivent donc être considérées en fonction de l'état du site ce jour-là. Par ailleurs, les conditions de l'analyse sont rappelées ci-dessous :

- Site analysé : pausebento.fr
- URL de départ : <http://www.pausebento.fr/>
- Protocoles : aucune restriction (HTTP et HTTPS)
- Nombre maximal d'URL à explorer : 5 000
- Motif d'inclusion : aucun
- Motif d'exclusion : .jpg OR .jpeg OR .png OR .gif
- User-Agent : Mozilla/5.0 (Windows NT 6.3; Win64; x64; rv:45.0) Gecko/20100101 Firefox/45.0
- Ignorer le robots.txt : non
- Prendre en compte les liens externes ayant un attribut rel="nofollow" : non
- Limiter l'exploration au sous-domaine www.pausebento.fr : non

Vous trouverez à la fin de ce rapport des explications techniques sur le fonctionnement de l'analyse effectuée ainsi que des définitions des termes utilisés dans le jargon du référencement naturel (SEO).

1.2 Intérêt de cette analyse

Ce rapport de synthèse ainsi que ses annexes vous permettent :

- D'avoir le détail des erreurs techniques encore présentes sur votre site ;
- D'avoir la liste détaillée de toutes vos URL indexables par Google (à comparer au nombre indiqué par Google dans Search Console) ;
- De vous aider à établir un plan de migration / refonte de site en cas de changements d'URL ;
- De disposer d'une liste exhaustive de vos URL, avec pour chacune :
 - la balise title
 - la balise meta description
 - l'éventuelle URL canonique
 - l'éventuelle balise meta robots
- D'étudier les niveaux de profondeur des URL (c'est un point stratégique, surtout pour les gros sites). Ce rapport vous aide à comprendre quelles URL sont mises en avant dans le maillage interne (notion de hiérarchie). Rappelez-vous que dans l'idéal, vous devriez adapter votre maillage en fonction de la vie du site (effets saisons, produits phares, etc.) ;
- De vous aider à créer un ou plusieurs fichiers sitemaps en fonction de vos besoins.

1.3 Rapport d'audit et annexes contenant toutes les données









Ce rapport d'audit détaille tous les éléments analysés. Les données sont disponibles sous forme exhaustive dans des fichiers annexes au format Excel (et sur demande au format TSV Tab-separated values).

Annexes

Les annexes sont disponibles dans le fichier ZIP [annexes-pausebento_fr-2015-12-22.zip](#).

2 Synthèse des résultats obtenus

Voici une synthèse des principaux éléments audités sur le site (code couleur vert=succès, jaune=avertissement, rouge=erreur), détaillés dans le reste de ce rapport :

Élément analysé	Optimisation	Erreurs constatées sur les pages	
		en pourcentage	en nombre de pages
URL répondant en erreur		7,66 % ⁽¹⁾	20
Pages HTML non indexables		59,75 % ⁽²⁾	144
Pages HTML trop longues à télécharger		78,42 % ⁽²⁾	189
Pages HTML trop profondes		0 % ⁽³⁾	aucune
Pages avec balise Title non optimale		35,05 % ⁽³⁾	34
Pages avec balise meta description non optimale		2,06 % ⁽³⁾	2
Pages HTML avec un contenu non optimal		3,09 % ⁽³⁾	3
Liens sortants externes en erreur		0 % ⁽⁴⁾	aucune

- (1) pourcentage calculé par rapport au nombre d'URL explorées, soit ici 261 URL
- (2) pourcentage calculé par rapport au nombre d'URL de pages HTML sans erreur (code 200), soit ici 241 URL
- (3) pourcentage calculé par rapport au nombre d'URL de pages HTML indexables, soit ici 97 URL
- (4) pourcentage calculé par rapport au nombre d'URL externes distinctes trouvées dans les liens sortants follow, soit ici 0 URL

Ce tableau de synthèse est basé uniquement sur ce qui est considéré comme des **erreurs**, mais vous verrez que par endroits dans ce rapport sont également listés des **avertissements** : il s'agit de points à vérifier ou d'optimisations moins importantes à mettre en place. Les taux d'erreurs sont en gras (mais pas ceux concernant les avertissements).

Certaines cases des tableaux contiennent le libellé "n/a" ce qui signifie "non applicable" pour indiquer qu'il est normal que la case ne contienne pas de valeur.

3 Analyse du crawl et de l'indexabilité

3.1 Nombre total d'URL crawlées

Dans votre cas, 100% des URL du site (trouvables par des liens depuis la page de départ) ont pu être analysées. Le nombre d'URL crawlées a été de **261**.

3.2 État des URL crawlées

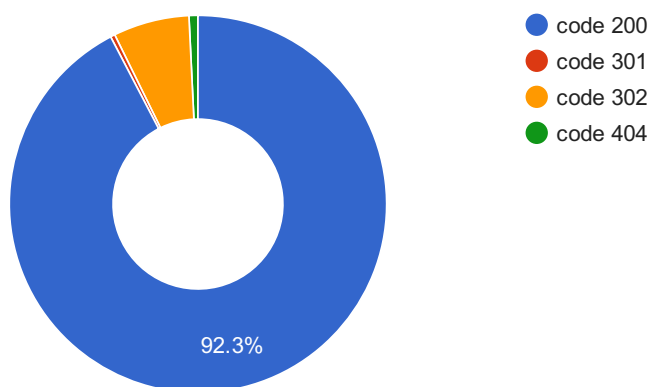
Voici le bilan des codes HTTP rencontrés sur le site (lisez leur signification [en fin de rapport](#)) :

Code HTTP	Nb d'URL	% des URL crawlées
200	241	92,34 %
301	1	0,38 %
302	17	6,51 %
404	2	0,77 %
Total erreurs 3xx	18	6,90 %
Total erreurs 4xx	2	0,77 %
Total erreurs 5xx	0	0 %
Total erreurs 3xx + 4xx + 5xx	20	7,66 %

Remarque : même si les redirections ne sont pas à proprement parler des erreurs techniques, c'est pénalisant pour le référencement. C'est pourquoi elles ont été les avons incluses dans le calcul des erreurs. Il est recommandé de faire des liens directs (directement vers la page finale) au lieu de passer par une ou des redirection(s) qui sont des étapes intermédiaires.

Voici ce que cela donne sous forme graphique :

Répartition des codes HTTP (% des URL crawlées)



Annexe

Le fichier annexe `pausebento_fr-2015-12-22-codes-http.xlsx` contient les données associées.

Annexe

Afin de vous permettre de localiser les erreurs 4xx et 5xx sur le site, le fichier annexe `pausebento_fr-2015-12-22-liens-vers-erreurs-4xx-5xx.xlsx` contient les URL des pages contenant les liens en erreur.

Annexe

Afin de vous permettre de localiser les erreurs 3xx sur le site, le fichier annexe `pausebento_fr-2015-12-22-liens-vers-erreurs-3xx.xlsx` contient les URL des pages contenant les liens en erreur.

3.3 Types MIME

Voici la liste des types MIME (lisez leur signification [en fin de rapport](#)) rencontrés pour les pages répondant correctement (code HTTP 200) :

Type MIME	Nb d'URL	%des URL crawlées en code 200
text/html	241	100 %

Annexe

Le fichier annexe `pausebento_fr-2015-12-22-types-mime.xlsx` contient les données associées.

3.4 Pages HTML crawlées sans erreur

Le référencement naturel se travaille essentiellement avec les pages HTML. Pour vous aider à avoir une vue d'ensemble, les principales informations disponibles pour l'ensemble de vos pages HTML crawlées sans erreur (code HTTP 200) ont été regroupées dans les annexes :

- URL
- Title
- Meta description
- Balise d'URL canonique
- URL canonique via en-tête HTTP
- Meta robots
- En-tête HTTP X-Robots-Tag

Annexe

Le fichier annexe `pausebento_fr-2015-12-22-pages-sans-erreur.xlsx` contient les données associées.

3.5 Pages HTML indexables par les moteurs

Le référencement naturel se travaillant essentiellement pour les pages HTML, l'analyse porte ici exclusivement sur les pages HTML.

Il s'agit ici d'étudier l'indexabilité de vos pages, c'est-à-dire la possibilité qu'elles soient indexées. On parle donc de pages « indexables ».

Les pages HTML indexables sont toutes celles :

- qui renvoient un code 200
- et qui ne sont pas interdites d'indexation (via une balise meta robots noindex ou none, ou bien un en-tête HTTP X-Robots-Tag)
- et qui ne contiennent pas d'URL canonique différente de l'URL crawlée

Voici le décompte pour le site :

	Nb d'URL	% des URL HTML code 200
Pages HTML avec code HTTP 200 (A)	241	
Dont pages interdites d'indexation	122	50,62 %
Dont pages avec URL canonique différente de l'URL crawlée	118	48,96 %
Nb total de pages non-indexables (B)	144	59,75 %
Nombre de pages HTML indexables (A - B)	97	40,25 %

Sauf cas particuliers dont vous connaissez bien l'existence sur ce site, le pourcentage de pages indexables devrait être élevé (proche de 100%).

En effet, il n'est pas efficace de laisser de nombreuses pages se faire crawler alors que vous exigez qu'elles ne soient pas indexées.

Dans votre cas, 122 pages HTML interdites d'indexation ont été identifiées, représentant 50,62 % des pages HTML en code 200. Il faudrait vérifier si c'est normal d'interdire l'indexation de ces pages. Vous trouverez la liste des URL concernées dans les annexes.

Dans votre cas, 118 pages HTML avec une URL canonique différente de l'URL crawlée ont été identifiées, représentant 48,96 % des pages HTML en code 200. Il faudrait vérifier si c'est normal pour ces pages d'indiquer une URL canonique différente de l'URL crawlée. Vous trouverez la liste des URL concernées dans les annexes.

Annexe

Le fichier annexe pausebento_fr-2015-12-22-indexabilite.xlsx contient les données associées.

Annexe

Afin de vous permettre de localiser les erreurs liées à une directive noindex (balise meta ou X-Robots-Tag) sur le site, le fichier annexe pausebento_fr-2015-12-22-liens-vers-noindex.xlsx contient les URL des pages contenant les liens en erreur.

Annexe

Afin de vous permettre de localiser les erreurs potentielles liées à l'URL canonique sur le site, le fichier annexe pausebento_fr-2015-12-22-liens-vers-erreurs-balise-canonique.xlsx contient les URL des pages contenant les liens vers ces URL.

4 Analyse du temps de téléchargement

Plus vos pages mettent de temps à être téléchargées, plus c'est pénalisant :

- pour les internautes, car le temps de chargement complet de la page risque d'être long ;
- pour les robots comme Googlebot, car ils crawlent moins de pages par jour si elles sont lentes

Il est recommandé d'optimiser la conception du site, le contenu de vos pages ainsi que le serveur d'hébergement afin d'avoir des pages qui se téléchargent rapidement. Dans cette partie, l'analyse concerne le temps de téléchargement de chaque page HTML répondant sans erreur (code HTTP 200). Il est recommandé que chaque page soit téléchargée en 600 ms maximum.

Voici les résultats de l'analyse :

	Nb d'URL	% des URL HTML code 200
Pages trop lentes (> 600 ms)	189	78,42 %
dont : pages vraiment trop lentes (> 1000 ms)	93	38,59 %
Nb de pages HTML code 200	241	
Taux d'erreur		78,42 %

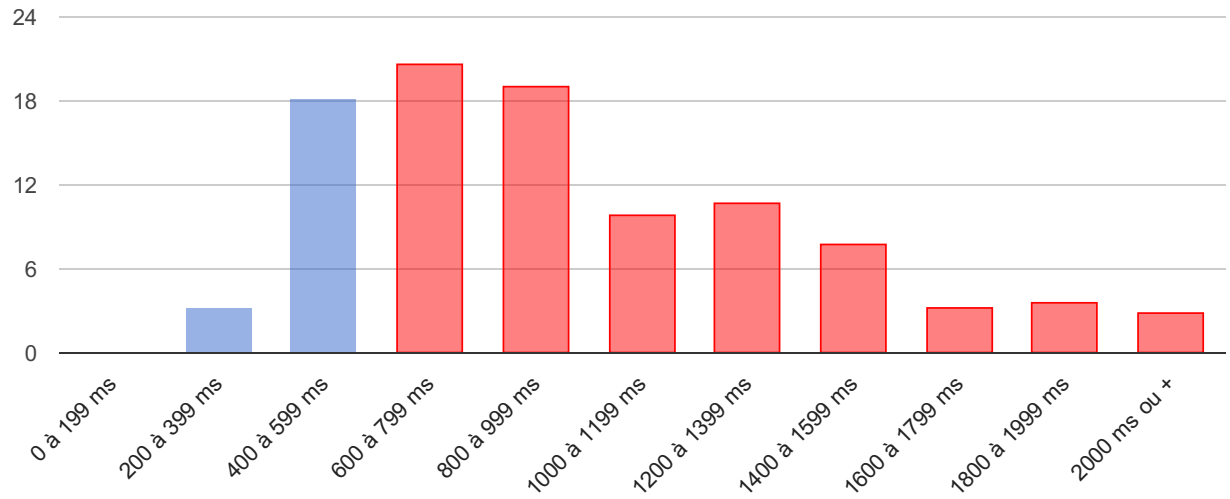
Remarque : il ne s'agit pas ici du temps de chargement complet de la page, puisque tous les fichiers externes sont ignorés (Javascript, CSS, images, etc.). Il s'agit du temps mis par le serveur pour répondre à la requête, générer la page HTML et la télécharger. C'est donc le même principe que le temps indiqué dans Google Search Console (ex Webmaster Tools) rubrique Exploration > Statistiques sur l'exploration > Temps de téléchargement d'une page (en millisecondes). Cependant, notre calcul ne tient compte que des fichiers HTML alors que les données fournies par la Search Console incluent également les CSS, images, PDF et autres documents crawlés par Googlebot.

Voici la répartition par tranches de 200 ms (pour les pages HTML code 200) :

Temps de téléchargement (ms)	Nb de pages	Pourcentage des pages
de 0 ms à 199 ms	0	0 %
de 200 ms à 399 ms	8	3,32 %
de 400 ms à 599 ms	44	18,26 %
de 600 ms à 799 ms	50	20,75 %
de 800 ms à 999 ms	46	19,09 %
de 1000 ms à 1199 ms	24	9,96 %
de 1200 ms à 1399 ms	26	10,79 %
de 1400 ms à 1599 ms	19	7,88 %
de 1600 ms à 1799 ms	8	3,32 %
de 1800 ms à 1999 ms	9	3,73 %
2000 ms ou +	7	2,90 %

Voici ce que cela donne sous forme graphique :

Pourcentage des pages en fonction du temps de téléchargement



Annexe

Le fichier annexe pausebento_fr-2015-12-22-temps-telechargement.xlsx contient les données associées.

5 Analyse de la profondeur des pages

Pour qu'une page soit efficace en référencement naturel, il faut éviter qu'elle soit trop profonde dans le site, la profondeur étant mesurée par le nombre de clics nécessaires (sur des balises <a href>) pour l'atteindre depuis la page d'accueil.

Dans votre cas, **la profondeur maximale recommandée est de 2**. Concrètement, cela signifie qu'il est préférable que la majorité des URL soit à une profondeur inférieure ou égale à 2. Plus vous aurez d'URL situées à des profondeurs supérieures, plus vous aurez des problèmes de crawl, d'indexation et même de positionnement.

L'étude a porté sur les pages HTML indexables.

5.1 Répartition des résultats pour chaque profondeur

Voici les résultats selon les niveaux :

Profondeur	Nb de pages	Nb cumulé de pages	Pourcentage des pages	Pourcentage cumulé
0	1	1	1,03 %	1,03 %
1	63	64	64,95 %	65,98 %
2	33	97	34,02 %	100 %

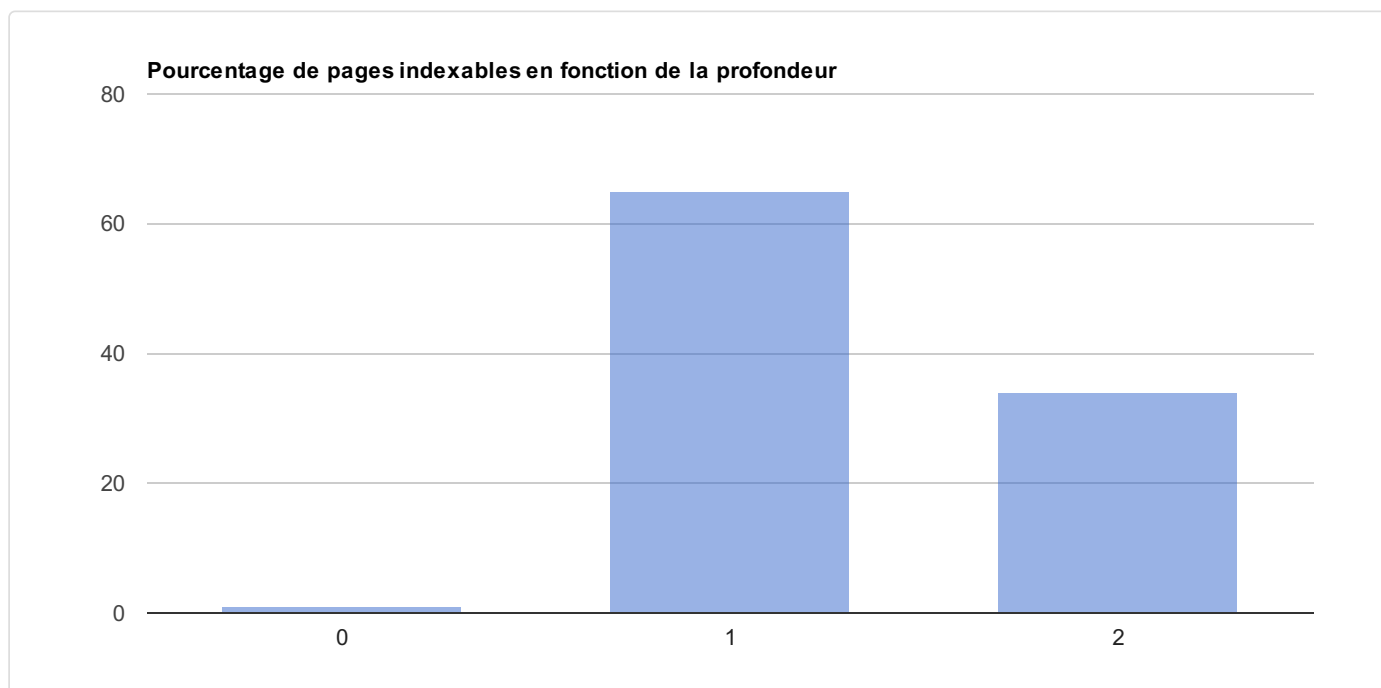
Dans votre cas, aucune page n'est trop profonde.

Annexe

Le fichier annexe pausebento_fr-2015-12-22-profondeur.xlsx contient les données associées.

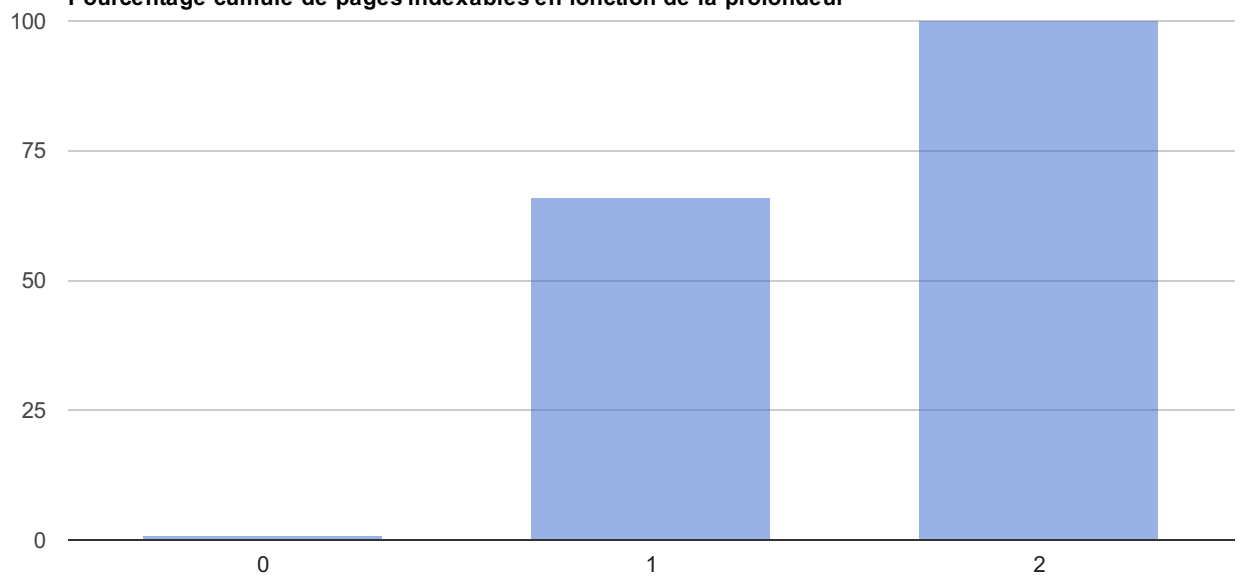
5.2 Analyse graphique

Le graphique ci-dessous reprend les éléments du tableau précédent :



Le graphique ci-dessous montre le pourcentage cumulé :

Pourcentage cumulé de pages indexables en fonction de la profondeur



6 Analyse de la balise title

Toutes les analyses qui suivent portent sur les pages HTML indexables.

6.1 Taille de la balise Title

Pour optimiser la balise title, il faut exploiter son potentiel, c'est-à-dire utiliser au mieux la place disponible. Mais il faut également éviter les titres trop longs qui risquent de ne pas s'afficher en entier dans les résultats de recherche. L'idéal est de ne pas dépasser ~55 caractères ; au-delà de 70, c'est considéré comme une erreur.

Voici les résultats de notre analyse :

	Erreurs	Avertissements	
Nb de pages sans balise title ou avec une balise title vide	0	n/a	
Nb de pages avec seulement 1 mot dans title	4	n/a	
Nb de pages avec seulement 2 mots dans title	11	n/a	
Nb de pages avec seulement 3 mots dans title	18	n/a	
Nb de pages avec entre 56 et 70 caractères dans title	n/a	0	
Nb de pages avec entre 71 et 100 caractères dans title	1	n/a	
Nb de pages avec + de 100 caractères dans title	0	n/a	
Nb total d'erreurs et d'avertissements	34	0	
Nb de pages indexables			97
Taux d'erreur	35,05 %	0 %	

Annexe

Le fichier annexe pausebento_fr-2015-12-22-balise-title.xlsx contient les données associées.

6.2 Doublons dans la balise title

Pour optimiser la balise title, il faut que chaque page ait une balise title unique, sinon c'est considéré comme une erreur.

Voici les résultats de notre analyse :

	Erreurs	Avertissements	
Nb de balises title utilisées plus d'une fois sur le site (hors balise vide)	0	n/a	
Nb de pages indexables par Google concernées (balise vide incluse)	0	n/a	
Nb de pages indexables			97
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, aucune donnée n'étant disponible, le fichier annexe n'a pas été généré.

6.3 Répétition de mots dans la balise Title

Pour éviter la sur-optimisation de la balise title, il faut éviter de répéter plus de 2 fois le même mot (et encore, c'est souvent trop de répéter 2 fois un même mot, c'est pourquoi ce cas est considéré comme un avertissement). Néanmoins, il peut y avoir des cas particuliers, notamment avec les articles et autres mots très courants. Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages avec 2 fois un même mot dans title	n/a	7	
Nb de pages avec 3 fois un même mot dans title	0	n/a	
Nb de pages avec 4 fois un même mot dans title	0	n/a	
Nb de pages avec 5 fois ou plus un même mot dans title	0	n/a	
Nb total d'erreurs et d'avertissements	0	7	
Nb de pages indexables			97
Taux d'erreur	0 %	7,22 %	

Annexe

Le fichier annexe pausebento_fr-2015-12-22-meme-mot-dans-title.xlsx contient les données associées.

6.4 Votre marque dans la balise Title

Pour optimiser la balise title, même s'il peut y avoir des exceptions, il est recommandé de ne pas inclure votre marque (faites-le dans la meta description).

Voici les résultats de notre analyse :

	Erreurs	Avertissements	
Nb de pages qui contiennent la marque "pausebento" dans title	n/a	0	
Nb de pages indexables			97
Taux d'erreur		0 %	

Annexe

Dans votre cas, aucune donnée n'étant disponible, le fichier annexe n'a pas été généré.

6.5 Conclusion sur la balise Title

Pour optimiser votre référencement, aucune balise title ne doit comporter une ou plusieurs des erreurs décrites précédemment. Ce rapport a listé tous les cas où une page HTML contient au moins une de ces erreurs dans la balise Title.

Pour cette conclusion, seules les erreurs sont prises en compte (pas les avertissements).

Voici les résultats de l'analyse :

	Erreurs	
Nb de pages avec une ou plusieurs erreurs dans la balise Title	34	
Nb de pages indexables		97
Taux d'erreur	35,05 %	

7 Analyse de la balise meta description

7.1 Taille de la balise meta description

Pour optimiser la balise meta description, il faut exploiter son potentiel, c'est-à-dire utiliser au mieux les ~160 caractères disponibles (environ 2 lignes).

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages avec balise meta description vide ou absente	2	n/a	
Nb de pages avec 1 mot dans meta description	0	n/a	
Nb de pages avec 2 mots dans meta description	0	n/a	
Nb de pages avec 3 mots dans meta description	0	n/a	
Nb total	2	n/a	
Nb de pages indexables			97
Taux d'erreur	2,06 %	n/a	

Annexe

Le fichier annexe pausebento_fr-2015-12-22-meta-desc.xlsx contient les données associées.

7.2 Doublons dans la balise meta description

Pour optimiser la balise meta description, il faut qu'elle soit unique.

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de balises meta description utilisées plus d'une fois (hors balise vide)	0	n/a	
Nb de pages concernées	0	n/a	
Nb de pages indexables			97
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, aucune donnée n'étant disponible, le fichier annexe n'a pas été généré.

7.3 Conclusion sur la balise meta description

Pour optimiser votre référencement, aucune balise meta description ne doit comporter une ou plusieurs des erreurs décrites précédemment. Ce rapport a listé les cas où une page HTML contient au moins une de ces erreurs dans la balise meta description.

Pour cette conclusion, seules les erreurs sont prises en compte (pas les avertissements).

Voici les résultats de l'analyse :

	Erreurs	
Nb de pages avec une ou plusieurs erreurs dans la balise meta description	2	
Nb de pages indexables		97
Taux d'erreur	2,06 %	

8 Analyse du contenu des pages

8.1 Balises H1

Pour vous donner le plus de chances de réussir votre référencement, il est recommandé d'utiliser les balises de titres h1, h2, h3... h6. Attention, ces balises ne doivent être intégrées en théorie que sur les pages à fort contenu éditorial. Étant donné qu'il n'y a pas toujours forcément besoin d'une balise h2 ou h3, l'analyse a porté uniquement sur les cas où aucune balise h1 n'a été utilisée (et considéré que dans ce cas il s'agit bien d'une erreur, la page d'accueil pouvant faire exception).

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages avec balise H1 absente ou vide	1	n/a	
Nb de pages indexables			97
Taux d'erreur	1,03 %	n/a	

Vous trouverez dans l'annexe non seulement la liste des pages en erreur (sans balise H1), avec leur balise TITLE, mais aussi le libellé de toutes les balises H1 trouvées sur les autres pages.

Annexe

Le fichier annexe pausebento_fr-2015-12-22-baliseh1.xlsx contient les données associées.

8.2 Taille du texte (nb de mots)

2 types de problèmes ont été étudiés :

- Les pages trop courtes, ayant un contenu estimé trop faible
- Les pages sans doute trop courtes, ayant un contenu estimé un peu trop faible

Rappelez-vous qu'avec son algorithme Panda, Google cherche à identifier les pages à faible valeur ajoutée. En laissant se faire indexer des pages à trop faible contenu, vous prenez donc des risques... C'est pourquoi pour les 2 cas indiqués, il est indiqué qu'il s'agit d'une erreur.

Les analyses ont porté sur l'ensemble de la page, en incluant le haut de page, les éventuelles marges et le pied de page. L'algorithme utilisé ici sait s'adapter à chaque site pour se concentrer sur la zone utile de chaque page (celle qui généralement change d'une page à l'autre), c'est-à-dire le contenu qui intéresse vraiment l'internaute. Dans les annexes, vous trouverez le nombre total de mots situés dans la page (body) même si l'analyse ne se base pas que sur celui-ci.

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages vraiment très courtes	2	n/a	
Nb de pages sans doute trop courtes	0	n/a	
Nb total	2	n/a	
Nb de pages indexables			97
Taux d'erreur	2,06 %	n/a	

Annexe

Le fichier annexe pausebento_fr-2015-12-22-nb-mots-body.xlsx contient les données associées.

8.3 Mise en exergue (strong ou bold)

Pour vous donner le plus de chances de réussir votre référencement, c'est en général mieux si chaque page contient quelques mots mis en exergue (gras ou bold). C'est néanmoins un point mineur dans l'optimisation du référencement (c'est pourquoi c'est noté ici uniquement comme un avertissement). Utilisez la mise en exergue (balise strong) pour mettre en évidence certains mots ou expressions dans la page, avant tout pour capter l'attention de l'internaute.

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages qui ne contiennent aucun mot en gras (strong ou bold)	n/a	0	
Nb de pages indexables			97
Taux d'erreur	n/a	0 %	

Annexe

Dans votre cas, aucune donnée n'étant disponible, le fichier annexe n'a pas été généré.

8.4 Analyse du nombre de liens par page

Pour vous donner le plus de chances de réussir votre référencement, il faut que chaque page fasse suffisamment de liens vers les autres pages du site, mais pas trop.

4 types de problèmes ont été identifiés :

- Les pages avec très peu de liens internes follow, a priori des erreurs : elles font moins de 5 liens
- Les pages avec trop de liens internes follow, sans doute également des erreurs : elles en font plus de 300
- Les pages avec des liens internes en nofollow (sauf cas particulier, cela n'a pas d'intérêt)
- Les pages qui font plus de 10 liens follow vers d'autres sites (à vous de valider que c'est bien votre choix). Les réseaux sociaux suivants ont été exclus des analyses : LinkedIn, Twitter, Facebook, Google +, Pinterest, Viadeo, Myspace, FriendFeed, Youtube, Dailymotion, Instagram, Vine, Vimeo, Flickr.

Les analyses ont porté sur l'ensemble de la page, en incluant le haut de page, les éventuelles marges et le pied de page.

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages avec en tout moins de 5 liens	0	n/a	
Nb de pages avec en tout plus de 300 liens	n/a	0	
Nb de pages avec plus de 10 liens externes	n/a	0	
Nb de pages avec 1 ou plusieurs liens internes nofollow	n/a	97	
Nb total	0	97	
Nb de pages indexables			97
Taux d'erreur	0 %	100 %	

Annexe

Le fichier annexe pausebento_fr-2015-12-22-nb-liens.xlsx contient les données associées.

8.5 Analyse des contenus dupliqués internes

Pour optimiser votre référencement, aucun contenu ne doit être disponible en interne à plusieurs URL différentes. En effet, si cela se produit, vous êtes confrontés aux problèmes de contenus dupliqués internes. Même si Google indique qu'il ne pénalise pas pour ça, avoir des contenus dupliqués internes dégrade les performances de votre référencement :

- le crawl est moins efficace car Google doit crawler des URL en trop. S'il y en a peu ce n'est pas bien grave, mais parfois cela impose à Google de crawler bien plus d'URL inutiles que d'URL importantes correspondant aux vrais contenus à faire indexer.
- vous ajoutez de la concurrence entre vos propres pages là où il ne devrait pas y en avoir. Laquelle des différentes URL Google va-t-il choisir de mettre en avant dans ses résultats ?
- vous éparpillez vos efforts, car les internautes qui souhaitent partager vos contenus ou vous faire des liens n'utiliseront pas une unique URL pour un contenu donné, mais un peu de chaque. Conséquence : vous aurez un score faible sur plusieurs URL (donc très peu visibles dans Google) au lieu d'avoir un bon score sur une seule URL (ce qui donne des chances d'être bien visible).

L'ensemble des pages HTML indexables ont été analysées afin d'identifier les cas de contenus dupliqués à 100%. Il s'agit donc des cas où l'ensemble du contenu HTML est strictement identique pour des URL différentes. Les cas de contenus presque dupliqués n'ont pas été étudiés (il s'agit des cas où la plupart du contenu est identique, mais pas la totalité). Les cas listés sont donc tous considérés comme des erreurs.

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb d'URL de pages HTML ayant un contenu strictement identique	0	n/a	
Nb de pages indexables			97
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, aucune donnée n'étant disponible, le fichier annexe n'a pas été généré.

8.6 Identifiants de session dans les URL

Il ne faut surtout pas que les robots trouvent des identifiants de session dans vos URL, car cela génère de nombreux problèmes.

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages avec un ID de session	0	n/a	
Nb de pages indexables			97
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, aucune donnée n'étant disponible, le fichier annexe n'a pas été généré.

8.7 Paramètres de tracking dans les URL

En interne, il ne devrait pas y avoir de liens incluant des paramètres de tracking dans les URL (passés derrière le signe ?).

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages avec une variable de tracking	0	n/a	
Nb de pages indexables			97
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, aucune donnée n'étant disponible, le fichier annexe n'a pas été généré.

8.8 Conclusion sur le contenu des pages

Pour optimiser votre référencement, aucune page ne doit comporter une ou plusieurs des erreurs décrites précédemment. Ce rapport a listé les cas où une page HTML contient au moins une de ces erreurs.

Pour cette conclusion, seules les erreurs sont prises en compte (pas les avertissements).

Voici les résultats de l'analyse :

	Erreurs	
Nb de pages avec une ou plusieurs erreurs dans le contenu	3	
Nb de pages indexables		97
Taux d'erreur	3,09 %	

9 Analyse des liens sortants externes

Dans cette rubrique, l'analyse porte sur tous les liens qui sortent du site (ils pointent donc vers un autre nom de domaine). Ils sont appelés « liens externes », qu'il ne faut pas confondre avec des backlinks (qui peuvent être eux aussi internes ou externes).

Les liens nofollow ont été ignorés, ainsi que les liens vers les réseaux sociaux suivants : LinkedIn, Twitter, Facebook, Google +, Pinterest, Viadeo, Myspace, FriendFeed, Youtube, Dailymotion, Instagram, Vine, Vimeo, Flickr.

9.1 Nombre total de liens sortants externes

Aucun lien externe n'a été identifié.

Annexe
Dans votre cas, aucune donnée n'étant disponible, le fichier annexe n'a pas été généré.

9.2 Erreurs sur les liens sortants externes

L'analyse concerne ici uniquement les cas où les liens sortants externes pointent vers un document en erreur. Comme vous le savez, il vaut mieux éviter ce type de situation, aussi bien pour les internautes que pour le référencement.

Dans certains cas, l'URL de destination se fait rediriger : si l'URL finale est accessible, ce n'est pas un problème pour l'internaute mais dans l'idéal il ne devrait y avoir que des liens directs. C'est pourquoi ce cas de figure est considéré simplement comme un avertissement.

	Erreurs	Avertissements	
Nombre total de liens sortants externes redirigés	n/a	0	
Nombre total de liens sortants externes en erreur	0	n/a	
Nombre total de liens sortants externes			0
Taux d'erreur	0 %	0 %	

Tous les détails (liste des codes HTTP pour chacune des URL de destination externe en erreur ainsi que l'URL du site qui fait un lien vers une page en erreur) sont dans le fichier annexe.

10 Conclusion

Vous trouverez ci-après une liste des problèmes techniques du référencement en cours sur le site (étudié dans les conditions précisées dans l'introduction) à la date du 22 décembre 2015 et qu'il faut corriger au plus vite.

10.1 Problèmes de crawl et d'indexation

Il faut éviter de faire des liens internes vers des URL qui répondent en erreur. À la date de cette analyse, votre taux d'erreur était de 7,66 % (détails).

Il faut éviter de faire des liens internes vers des URL non indexables. À la date de cette analyse, votre taux d'erreur était de 59,75 % (détails).

10.2 Problèmes de temps de téléchargement

Il est conseillé de faire en sorte que chaque page soit téléchargée en 600 ms maximum. À la date de cette analyse, 78,42 % des URL HTML code 200 ont été identifiées comme étant trop lentes, ce qui n'est pas optimal (détails).

10.3 Problèmes de profondeur des pages

Pour ce site (et ce nombre d'URL analysées), il faut que la majorité des URL soit à une profondeur inférieure ou égale à 2. À la date de l'analyse, aucune erreur de ce type n'a été identifiée (détails).

10.4 Problèmes techniques concernant le contenu

Dans un premier temps, voici le point sur le plus important, à savoir l'optimisation du contenu des pages HTML indexables. Le référencement étant tellement concurrentiel, il faut se débrouiller pour n'avoir aucun frein sur son propre site. Voici les éléments techniques que vous devez corriger au plus vite, classés par taux d'erreur décroissant :

Élément d'analyse	Taux d'erreur
balise title (détails)	35,05 %
contenu (détails)	3,09 %
balise meta description (détails)	2,06 %

Pour réussir votre référencement, il faudrait donc dans l'idéal qu'aucune des pages n'ait de sous-optimisation. Un taux d'erreur global a été calculé en identifiant toutes les pages HTML indexables qui contiennent une ou plusieurs erreurs (concernant les éléments techniques du tableau précédent).

Nb de pages ayant au moins une erreur technique	35
Nb de pages indexables	97
Taux d'erreur global	36,08 %

Plus vous réussirez à réduire ce taux d'erreur global, plus vous vous donnerez de chances de réussir un bon référencement. Évidemment, cela ne suffit pas, il faut par exemple :

- obtenir régulièrement des liens de qualité ;
- avoir une bonne présence dans les réseaux sociaux ;
- que le comportement des internautes sur votre site montre qu'ils en sont satisfaits.

10.5 Problèmes sur les liens sortants externes

À la date de cette analyse, aucune erreur de ce type n'a été identifiée (détails).

10.6 Quand refaire un audit technique SEO ?

Ce type de rapport d'audit vous aide à identifier et à corriger les problèmes techniques qui freinent votre référencement naturel. Afin de détecter au plus vite les prochains problèmes techniques, il est conseillé de faire une nouvelle analyse :

- après avoir modifié ou mis à jour votre CMS
- après avoir ajouté de nouvelles rubriques ou pages importantes dans le site
- après avoir modifié le maillage interne
- après avoir ajouté des balises dans vos modèles de pages (exemple : URL canoniques)
- après avoir modifié les règles de crawl (robots.txt) ou d'indexation (balise meta robots)
- après des mises à jour de Google (algorithmiques ou d'interface de ses pages de résultats, par exemple la taille du titre)
- après avoir reçu une pénalité Google pour liens **sortants** externes artificiels (ou avant, pour l'anticiper)

11.1 Fonctionnement de l'audit

Pour chaque audit, une analyse exhaustive du site est effectuée, page par page, en partant d'une page de départ sélectionnée, comme Google peut le faire.

Ce fonctionnement reproduit de manière assez fidèle celui de Googlebot (le robot de Google). Par exemple :

- Respect du fichier robots.txt de la même façon que Google (avec donc quelques particularités par rapport au standard)
- Prise en compte :
 - des balises meta robots
 - des directives passées dans l'entête HTTP (pour l'URL canonique ou pour les directives d'indexation et de suivi des liens)
- Ne sont pas pris en compte :
 - les cookies
 - le javascript (et ajax ou équivalent)
 - le CSS
- Ne sont pas suivis :
 - les formulaires
 - les redirections meta
 - les liens nofollow
- Les images et les vidéos sont exclues de l'analyse

11.2 Signification des codes HTTP

Voici les principaux codes HTTP :

Les principales familles de codes :

- Codes 2XX (commençant par "2") : succès
- Codes 3XX (commençant par "3") : redirection
- Codes 4XX (commençant par "4") : erreur du client
- Codes 5XX (commençant par "5") : erreur du serveur

Les codes les plus courants :

- 200 = *OK* : l'URL a été trouvée
- 301 = *Moved Permanently* : l'URL a été déplacée de façon permanente (redirection 301)
- 302 = *Moved Temporarily* : l'URL a été déplacée de façon temporaire (redirection 302)
- 400 = *Bad Request* : la syntaxe de la requête est erronée
- 401 = *Unauthorized* : une authentification est nécessaire pour accéder à la ressource
- 403 = *Forbidden* : accès interdit (droits d'accès insuffisants)
- 404 = *Not Found* : l'URL est introuvable
- 410 = *Gone* : la ressource est indisponible et aucune adresse de redirection n'est connue (en général, c'est le cas des documents supprimés définitivement)
- 500 = *Internal Server Error* : erreur interne du serveur
- 503 = *Service Unavailable* : serveur temporairement indisponible ou en maintenance
- 504 = *Gateway Time-out* : temps d'attente d'une réponse d'un serveur à un serveur intermédiaire écoulé

11.3 Signification des types MIME

Voici les principaux types MIME, c'est-à-dire les formats de données des documents situés sur un site Internet :

- text/css : CSS (feuilles de styles)
- text/csv : CSV (données séparées par une virgule)
- text/html : HTML
- text/plain : texte brut
- text/xml : XML
- image/bitmap : image au format bitmap
- image/gif : image au format GIF
- image/jpeg : image au format JPEG
- image/png : image au format PNG
- image/svg+xml : image au format SVG

- image/tiff : image au format TIFF
- application/javascript : JavaScript
- application/json : JSON (JavaScript Object Notation)
- application/msword : Microsoft Word
- application/pdf : PDF
- application/xhtml+xml : XHTML
- application/xml : XML
- application/x-shockwave-flash : Flash
- application/zip : ZIP

11.4 Signification des termes SEO

Voici les principaux termes utilisés dans le jargon du référencement naturel (SEO) et notamment dans ce rapport d'audit. Vous trouverez des liens vers des articles qui détaillent les notions correspondantes :

- **HTTP et HTTPS** : HTTP est le protocole standard d'accès à des documents sur le web, tandis que HTTPS est la version sécurisée (transmission cryptée). Une URL en HTTP et la même en HTTPS sont considérées comme 2 URL distinctes.
- **Google Search Console** (ex- Webmaster Tools) : interface proposée par Google pour fournir au webmaster plusieurs informations concernant son site dans Google (problèmes de crawl ou d'indexation, pénalités, etc.). En savoir plus : [site officiel](#). Bing propose la même chose (voir [ici](#)).
- **URL indexable** : il s'agit d'une URL qui ne présente aucun obstacle technique pour que Google l'indexe. Il faut qu'elle ne soit pas bloquée au crawl (dans le fichier robots.txt) et qu'elle soit accessible (donc que le code HTTP renvoyé par le serveur soit 200) et qu'elle ne soit pas interdite d'indexation (via la balise meta robots noindex ou un entête HTTP X-Robots-Tag) et qu'elle ne contienne pas d'URL canonique différente. En général, une URL indexable est indexée par Google (mais ce rapport ne le vérifie pas).
- **URL canonique** : il s'agit de l'URL officielle d'une page. On l'utilise surtout dans le cas où un même contenu est disponible à plusieurs URL (dans ce cas pour ces différentes URL on indique quelle est l'URL officielle). En savoir plus : [aide Google](#).
- **Fichier robots.txt** : situé à la racine de chaque sous-domaine, il liste les éventuels endroits du site pour lesquels tous ou certains robots n'ont pas le droit de venir crawler. En savoir plus : [aide Google](#).
- **Balise meta robots** : cette balise permet de restreindre les robots tels que Googlebot pour leur interdire par exemple d'indexer la page (noindex) ou de tenir compte des liens présents sur la page (nofollow). En savoir plus : [aide Google](#).
- **En-tête HTTP X-Robots-Tag** : pour les documents non HTML, il n'est pas possible d'utiliser une balise meta robots. On peut dans ce cas passer des instructions par l'entête HTTP. En savoir plus : [aide Google](#).
- **Redirection (301 ou 302)** : on parle de redirection d'une URL A vers une URL B quand, en cherchant à accéder à A, le serveur nous renvoie vers B. En référencement, il faut privilégier la redirection permanente (code HTTP 301) dans la majorité des cas. Il existe de nombreux moyens de gérer les redirections. En savoir plus : [aide Google](#).
- **Sitemap** : un fichier sitemap liste des URL que l'on souhaite faire crawler par les moteurs de recherche (en général pour qu'ils indexent les URL concernées). Ce n'est pas obligatoire mais ça peut aider à accélérer le crawl ou dans le suivi du référencement. En savoir plus : [aide Google](#).
- **Lien nofollow** : si une page A fait un lien nofollow vers une page B, alors les moteurs de recherche feront comme si ce lien n'existe pas. En d'autres termes, la page B ne bénéficie d'aucun avantage (ou inconvénient) de recevoir un lien depuis A. Google impose que les liens commerciaux soient en nofollow, par le biais d'un attribut rel=nofollow sur la balise A (liens vendus, liens affiliés). En savoir plus : [aide Google](#).

Table des matières

Audit Technique SEO	1
1 Introduction	2
1.1 Objectif de ce rapport	2
1.2 Intérêt de cette analyse	2
1.3 Rapport d'audit et annexes contenant toutes les données	2
Annexes	2
2 Synthèse des résultats obtenus	3
3 Analyse du crawl et de l'indexabilité	4
3.1 Nombre total d'URL crawlées	4
3.2 État des URL crawlées	4
Annexe	4
Annexe	5
Annexe	5
3.3 Types MIME	5
Annexe	5
3.4 Pages HTML crawlées sans erreur	5
Annexe	5
3.5 Pages HTML indexables par les moteurs	5
Annexe	6
Annexe	6
Annexe	6
4 Analyse du temps de téléchargement	7
Annexe	8
5 Analyse de la profondeur des pages	9
5.1 Répartition des résultats pour chaque profondeur	9
Annexe	9
5.2 Analyse graphique	9
6 Analyse de la balise title	11
6.1 Taille de la balise Title	11
Annexe	11
6.2 Doublons dans la balise title	11
Annexe	11
6.3 Répétition de mots dans la balise Title	12
Annexe	12
6.4 Votre marque dans la balise Title	12
Annexe	12
6.5 Conclusion sur la balise Title	12
7 Analyse de la balise meta description	14
7.1 Taille de la balise meta description	14
Annexe	14
7.2 Doublons dans la balise meta description	14
Annexe	14
7.3 Conclusion sur la balise meta description	14
8 Analyse du contenu des pages	16
8.1 Balises H1	16
Annexe	16
8.2 Taille du texte (nb de mots)	16
Annexe	16
8.3 Mise en exergue (strong ou bold)	17
Annexe	17
8.4 Analyse du nombre de liens par page	17
Annexe	17
8.5 Analyse des contenus dupliqués internes	18
Annexe	18
8.6 Identifiants de session dans les URL	18
Annexe	18
8.7 Paramètres de tracking dans les URL	18
Annexe	19

8.8 Conclusion sur le contenu des pages	19
9 Analyse des liens sortants externes	20
9.1 Nombre total de liens sortants externes	20
Annexe	20
9.2 Erreurs sur les liens sortants externes	20
10 Conclusion	21
10.1 Problèmes de crawl et d'indexation	21
10.2 Problèmes de temps de téléchargement	21
10.3 Problèmes de profondeur des pages	21
10.4 Problèmes techniques concernant le contenu	21
10.5 Problèmes sur les liens sortants externes	21
10.6 Quand refaire un audit technique SEO ?	21
11 Explications complémentaires	23
11.1 Fonctionnement de l'audit	23
11.2 Signification des codes HTTP	23
11.3 Signification des types MIME	23
11.4 Signification des termes SEO	24
Table des matières	25