
Audit SEO

Site analysé : comnplus.fr

Com n' Plus

620, chemin de Pourcate

<https://comnplus.fr/>

N° tél. : 06 89 12 32 75

N° RCS : 8345411013 00019 - N° TVA : FR89 834541013

Score SEO



Vous avez encore du potentiel SEO à exploiter !

Meilleure note possible : 100

1.1 Objectif de ce rapport

Ce document a pour objectif de vous présenter les résultats de l'audit SEO de votre site **comnplus.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 **27 juillet 2020** et a porté sur **44 URL**.

195 images distinctes ont été trouvées dans ces pages (195 images sont hébergées sur comnplus.fr et 0 sur d'autres sites). Ces images n'ont pas été analysées car l'option n'a pas été sélectionnée.

Toutes les analyses qui en découlent doivent donc être considérées en fonction de l'état du site ce jour-là.

Voici les conditions de l'analyse :

- Site analysé : comnplus.fr
- URL de départ : <https://comnplus.fr/>
- Protocoles : aucune restriction (HTTP et HTTPS)
- Taille maximale du code HTML à télécharger : 700 Ko
- Nombre maximal d'URL à explorer : 100
- Motif d'inclusion : aucun
- Motif d'exclusion : .jpg OR .jpeg OR .png OR .gif
- User-Agent pour le site : Mozilla/5.0 (Linux; Android 8.0.0; SM-G960F Build/PPR1.180610.011; wv) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/80.0.3987.149 Mobile Safari/537.36
- User-Agent pour le robots.txt : Googlebot
- Vitesse de l'exploration : 1.0 requête par seconde
- Ignorer le robots.txt : non
- Suivre les liens internes nofollow : non
- Suivre les liens externes nofollow : non
- Limiter l'exploration au domaine comnplus.fr : oui
- CMS déclarés : Joomla
- Analyse des images : 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).

Les données sont disponibles sous forme exhaustive dans des fichiers annexes au format Excel (et sur demande au format TSV Tab-separated values).

Les annexes sont disponibles dans le fichier ZIP annexes-comnplus_fr-2020-07-27.zip.

1.2 Comment exploiter cet audit

Bien entendu, cet audit couvre tous les points essentiels d'un audit essentiellement technique : erreurs dans les liens (404...), URL canonique et balise meta robots noindex, temps de téléchargement, profondeur et maillage, balises title, meta description, H1-H6 ainsi que la taille du contenu dans la zone principale de la page. Le rapport explique tout et les annexes contiennent tout ce qu'il faut pour corriger.

Mais cette audit permet d'être bien plus efficace grâce à des algorithmes exclusifs développés pour tirer le maximum du croisement entre les données de crawl et celles de Google Analytics et Search Console :

- analyse des pages actives et inactives sur 1 an
- calcul de l'indice zombie de chaque page
- identification des pages les plus rentables à travailler
- identification des pages créant le plus d'insatisfaction chez les utilisateurs
- identification des pages pouvant passer de page 2 à page 1 de Google

Voici ce que vous devriez faire, dans l'ordre :

1. s'il y a trop d'erreurs de crawl (3xx, 4xx, 5xx), corrigez-les
2. allez directement voir la répartition de vos pages indexables selon leur performance SEO (nb de sessions générées sur 1 an)
3. si vous avez des pages à fort indice zombie, traitez-les en priorité en suivant les explications
4. exploitez la liste des pages les plus rentables à travailler en 1er et celles à fort potentiel pour passer en page 1 de Google

2 Synthèse des résultats

2.1 Score de performance

Le score pour cet audit est de **94,9** sur 100. Il représente le niveau d'optimisation du référencement naturel en ce qui concerne les éléments analysés sur le site (techniques, contenus, liens internes, qualité...).



Il est recommandé d'atteindre 95 (score en vert) ; entre 70 et 95 le score est en jaune et en-dessous de 70 il est en rouge.

2.2 Pages zombies

Cette analyse n'a pas pu être effectuée en raison de l'absence de couplage avec Google Analytics et Search Console.

2.3 QualityRisk

En utilisant un algorithme d'évaluation de la qualité, appelé QualityRisk, le risque moyen a été évalué à **6,0** sur l'ensemble des pages indexables (le pire possible étant 100).









En outre, **11,36 % des pages indexables sont de trop faible qualité** (elles ont un QualityRisk \geq 20).

Voyez les détails [ici](#).

Élément analysé	Optimisation	Erreurs constatées sur les pages	
		en pourcentage	en nombre de pages
QualityRisk (évaluation de la qualité)		11,36 % ⁽³⁾	5

2.4 Taux d'erreur

Ce tableau regroupe en 8 familles les dizaines d'éléments audités sur le site :

Élément analysé	Optimisation	Erreurs constatées sur les pages	
		en pourcentage	en nombre de pages
URL répondant en erreur		0 % ⁽¹⁾	aucune
Indexabilité des pages HTML		0 % ⁽²⁾	aucune
Temps de téléchargement		4,55 % ⁽²⁾	2
Profondeur et maillage interne		9,09 % ⁽³⁾	4
Balise title		0 % ⁽³⁾	aucune
Balise meta description		0 % ⁽³⁾	aucune
Contenu des pages		11,36 % ⁽³⁾	5
Liens sortants externes		0 % ⁽⁴⁾	aucune

- (1) pourcentage calculé par rapport au nombre d'URL explorées, soit ici 44 URL
- (2) pourcentage calculé par rapport au nombre d'URL de pages HTML sans erreur (code 200), soit ici 44 URL
- (3) pourcentage calculé par rapport au nombre d'URL de pages HTML indexables, soit ici 44 URL
- (4) pourcentage calculé par rapport au nombre d'URL externes distinctes trouvées dans les liens sortants follow, soit ici 10 URL

Explications sur le code couleur :

- vert = succès : moins de 5% des pages sont en erreur
- jaune = avertissement : entre 5% et 30% des pages sont en erreur
- rouge = erreur : plus de 30% des pages sont en erreur

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.

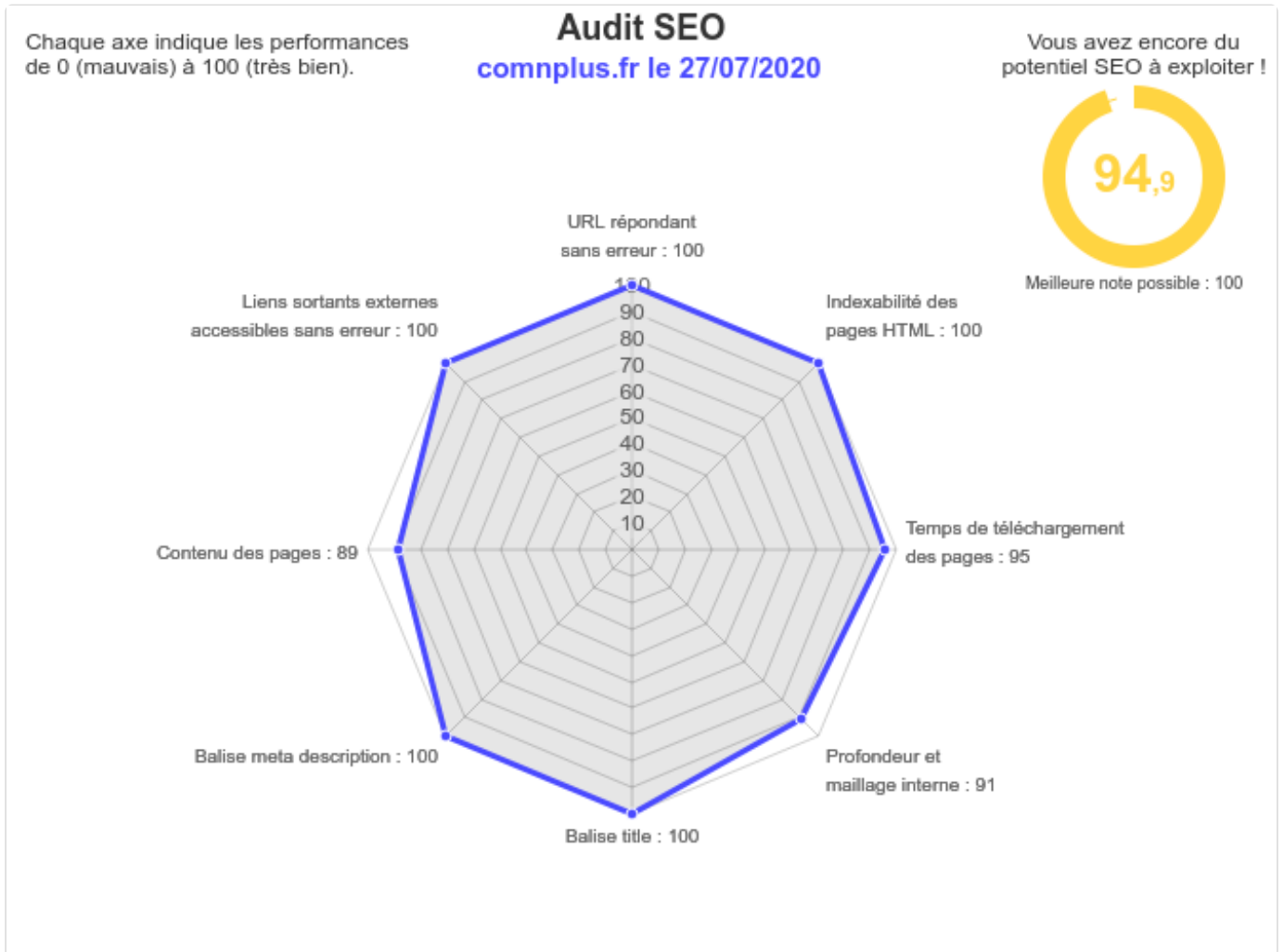
2.5 Performance

Les résultats présentés ci-dessus (et dans le reste du rapport) sont exprimés en termes de taux d'erreur. Retrouvez ci-dessous une autre façon d'analyser les résultats, de façon très visuelle, en calculant des scores de performances (en gros, c'est l'inverse du taux d'erreur).

Pour chaque famille de critères, le score de performances d'un site est égal au nombre de pages (ou de liens) SANS erreur divisé par le nombre de pages (ou de liens) étudié(e)s, exprimé en pourcentage. Par exemple, concernant le temps de téléchargement des pages :

- un taux d'erreur de 12 % signifie que 12 % des pages étudiées sont trop lentes à télécharger
- dans ce cas, le score de performances est de 88 % (= 100 - 12)

Le graphique en radar ci-dessous présente une synthèse des performances du site pour chaque famille de critères.



3 Crawl et indexabilité

Voici la méthode suivie pour l'analyse du crawl et de l'indexabilité :

1. d'abord, le point est fait sur les (éventuelles) URL qui n'ont pas été crawlées en raison d'une interdiction figurant dans un fichier robots.txt
2. ensuite, l'ensemble des URL crawlées sont listées (qu'elles soient accessibles ou pas)
3. enfin, parmi les pages HTML accessibles, celles qui sont indexables sont étudiées (sans vérifier si Google les a indexées, mais aucun critère technique n'empêche qu'elles le soient)

3.1 Bilan des URL bloquées par le(s) fichier(s) robots.txt

Les directives indiquées dans les fichiers robots.txt sont respectées (il doit y en avoir un pour chaque sous-domaine), concernant Googlebot (le crawler de Google).

Il est conseillé de mettre en ligne un fichier robots.txt pour chaque sous-domaine, même s'il est vide. Ceci évite de générer des erreurs 404 à chaque fois qu'un robot cherche à y accéder. C'est la raison pour laquelle le code HTTP du fichier robots.txt de chaque sous-domaine crawlé est indiqué (ce code doit donc être 200).

Voici la répartition des URL bloquées par un fichier robots.txt et qui reçoivent pourtant au moins un lien interne follow (ou qui sont indiquées dans une URL canonique) :

Fichier robots.txt	Code HTTP	Nombre d'URL bloquées
https://comnplus.fr/robots.txt	200	0
Total		0

Bien entendu, aucune URL stratégique pour votre business ne doit être bloquée par un fichier robots.txt ! Cette analyse et surtout l'annexe ci-dessous vous permettent de le vérifier.

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs n'a pas été généré.

Annexe

Le fichier annexe comnplus_fr-2020-07-27-robotstxt.zip contient l'ensemble des fichiers robots.txt récupérés.

3.2 Nombre total d'URL crawlées

A partir d'ici et pour tout le reste de ce rapport, seules sont prises en compte les URL autorisées au crawl (non bloquées dans un fichier robots.txt).

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 **44**.

Voici la répartition selon les protocoles HTTP et HTTPS :

Protocole	Nb d'URL	% des URL crawlées
HTTP	0	0 %
HTTPS	44	100 %

3.3 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 en HTTP	Nb d'URL en HTTPS	Nb d'URL (HTTP et HTTPS)	% des URL crawlées
200	0	44	44	100 %
Total erreurs	0	0	0	0 %

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 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 :

Annexe

Le fichier annexe comnplus_fr-2020-07-27_03-3_codes-http.xlsx contient les données associées.

3.4 URL trouvables uniquement par des liens internes nofollow

Depuis le 01/03/2020, Google peut décider de suivre un lien nofollow, qu'il soit interne ou externe ([source](#)).

L'option de suivi des liens internes nofollow n'a pas été activée pour cet audit.

3.5 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	44	100 %

Annexe

Le fichier annexe comnplus_fr-2020-07-27_03-4_types-mime.xlsx contient les données associées.

3.6 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 comnplus_fr-2020-07-27_03-5_pages-sans-erreur.xlsx contient les données associées.

3.7 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)	44	
Dont pages interdites d'indexation	0	0 %
Dont pages avec URL canonique différente de l'URL crawlée	0	0 %
Nb total de pages non-indexables (B)	0	0 %
Nombre de pages HTML indexables (A - B)	44	100 %

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.

4 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)	2	4,55 %
dont : pages vraiment trop lentes (> 1000 ms)	0	0 %
Nb de pages HTML code 200	44	
Taux d'erreur		4,55 %

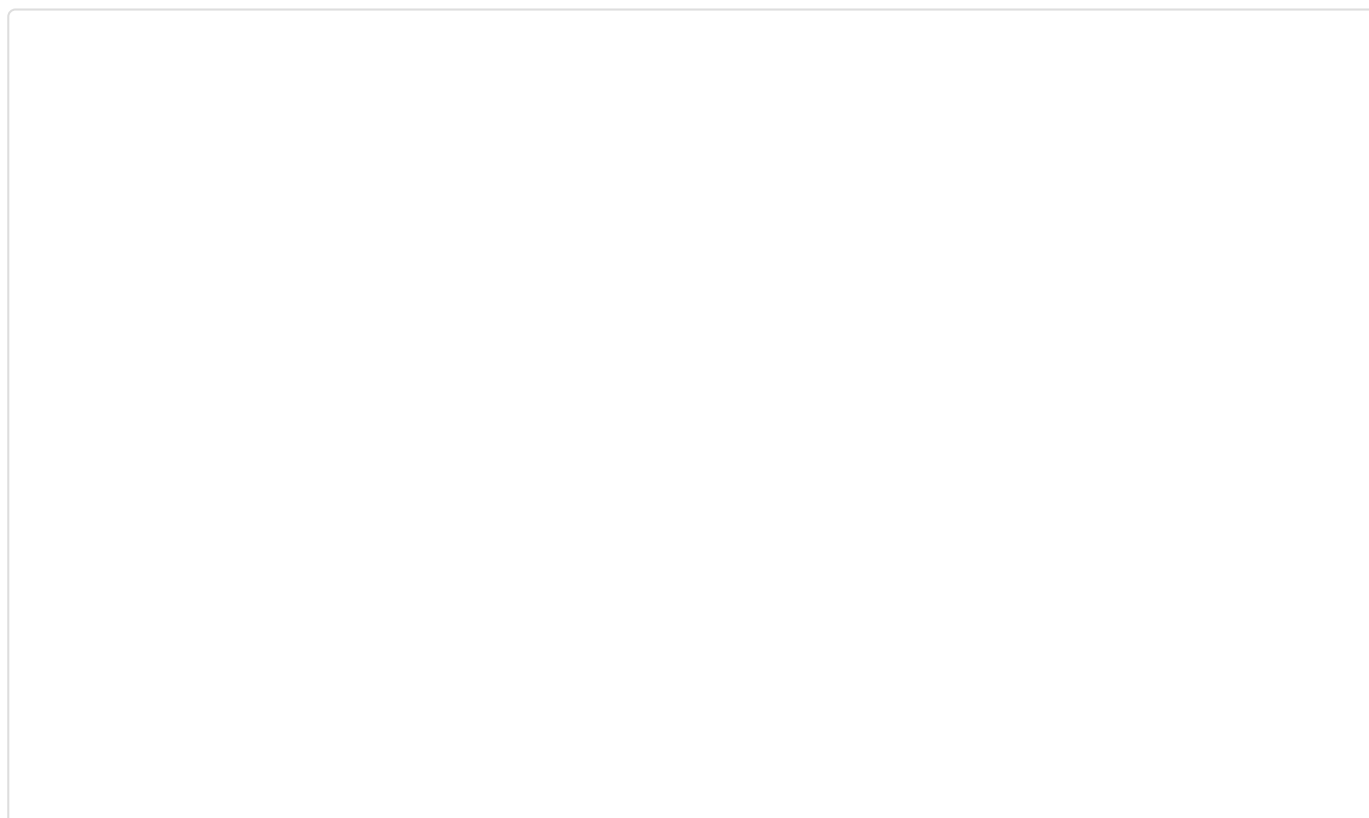
Remarque : **il ne s'agit pas ici du temps de chargement complet de la page**, puisque tous les fichiers externes sont ignorés pour ce calcul (Javascript, CSS, images, etc.). Il s'agit du temps mis par le serveur pour répondre à la requête, générer la page HTML et la télécharger. C'est donc le même principe 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, le 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.

En complément du temps de téléchargement de la page (code HTML), vous devriez surveiller le poids et les dimensions de vos images. 195 images distinctes ont été trouvées dans les pages explorées, mais n'ont pas été analysées car l'option n'a pas été sélectionnée.

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	34	77,27 %
de 200 ms à 399 ms	8	18,18 %
de 400 ms à 599 ms	0	0 %
de 600 ms à 799 ms	0	0 %
de 800 ms à 999 ms	2	4,55 %

Voici ce que cela donne sous forme graphique :



En plus du temps de téléchargement d'une page, il est intéressant d'étudier aussi le poids du code HTML brut (en Ko). Attention, ceci n'inclut aucune ressource de la page : ni les images, ni les javascript, CSS et autres.

Il est déconseillé d'avoir un poids (code HTML brut) trop important car cela augmente le temps de téléchargement. En outre, même si Google crawle des documents très volumineux (plusieurs Mo), Bing s'arrête autour de 700 Ko (ce qui est déjà très gros). C'est un cas très grave puisque le contenu et les liens de la page risquent de ne pas être vus !

La limite pour cet audit est de 700 Ko. Si certaines pages dépassent ce seuil, une partie seulement de leur contenu est prise en compte.

L'annexe ci-dessous indique le poids du code HTML brut de chaque page, ainsi qu'un indicateur qui précise s'il dépasse 700 Ko.

Voici les résultats de notre analyse :

	Erreurs	Avertissements	
Nb de pages dont le poids du code HTML brut dépasse 700 Ko	n/a	0	
Nb de pages HTML en code 200			44
Taux d'erreur		0 %	

Annexe

Le fichier annexe comnplus_fr-2020-07-27_04_temps-telechargement.xlsx contient les données associées.

5 Profondeur des pages et maillage interne

5.1 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.

Remarque importante : attention, il n'est pas conseillé de réduire artificiellement la profondeur des pages uniquement pour ne pas dépasser la profondeur recommandée. Il est préférable de privilégier les liens utiles à l'internaute, c'est-à-dire situés à des endroits faciles à identifier, entre des pages de thématique proche. Par exemple, c'est une mauvaise solution que de mettre des liens vers toutes les pages du site dans une page plan du site, elle-même reliée depuis le pied de page. Il n'est pas toujours possible ou facile d'éviter que certaines pages soient trop profondes, par contre il faut s'assurer que les pages stratégiques sont très faciles d'accès (profondeur 1 si possible).

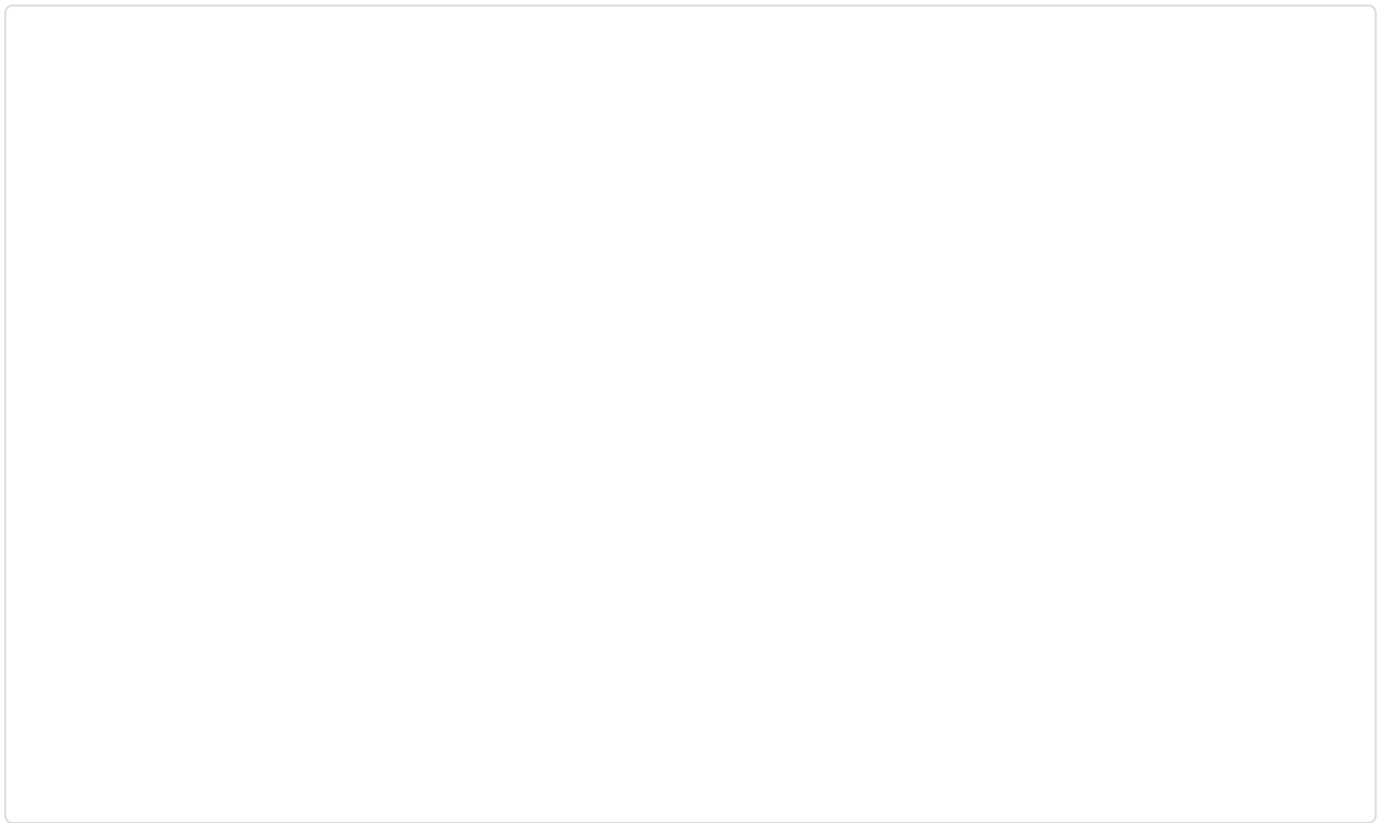
L'étude a porté sur les pages HTML indexables.

Voici les résultats selon les niveaux :

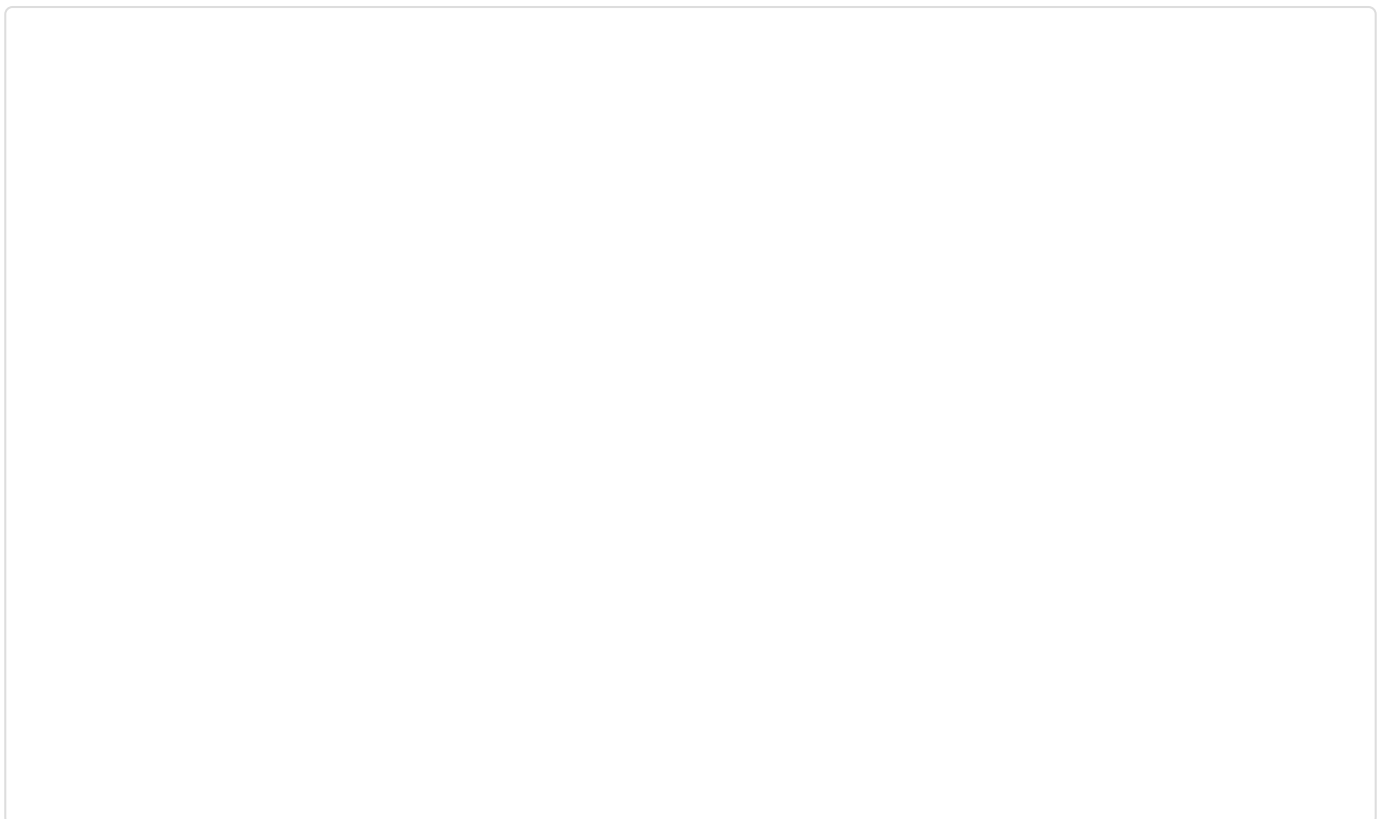
Profondeur	Nb de pages	Nb cumulé de pages	Pourcentage des pages	Pourcentage cumulé
0	1	1	2,27 %	2,27 %
1	30	31	68,18 %	70,45 %
2	13	44	29,55 %	100 %

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

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



Le graphique ci-dessous montre le pourcentage cumulé :



5.2 Analyse du maillage interne

Pour qu'une page soit efficace en référencement naturel, il faut qu'elle reçoive des liens (on appelle ça des backlinks ou liens entrants). Ils peuvent venir d'autres pages du site (backlinks internes) ou pas (backlinks externes).

Le minimum est d'avoir 1 backlink interne (sinon comment Google l'indexerait ? avec les sitemaps ce n'est pas assez efficace pour un bon positionnement Google). Mais c'est tout de même préférable d'en avoir plusieurs, surtout pour les pages stratégiques. Bien sûr, cela dépend de la taille du site, mais il faut éviter d'avoir très peu de backlinks internes.

Toujours concernant les backlinks internes, il est intéressant d'étudier aussi les ancres de liens (c'est-à-dire le texte cliquable qui mène à la page, ou bien l'attribut ALT dans le cas d'un lien sur une image). Là aussi, il est préférable **d'éviter d'avoir toujours exactement le même texte dans les liens qui pointent vers une page donnée**. En variant les ancres de liens, vous rendez les choses plus naturelles ; en plus, vous optimisez pour la longue traîne car la page de destination sera mieux positionnée sur un plus grand nombre de requêtes.

Au sujet des ancres de liens, il faut éviter les ancres vides : il s'agit souvent de liens placés sur des images sans attribut ALT, mais parfois aussi de liens texte vide. C'est déconseillé car cela ne permet pas d'optimiser le référencement de la page recevant le lien.

Remarques importantes :

1. Si une page A fait plusieurs liens fait une même page B, seul le 1er lien (dans l'ordre du code source) est pris en compte dans l'analyse des ancres.
2. Comme pour la profondeur, il n'est pas conseillé d'augmenter artificiellement le nombre d'ancres dans les liens uniquement pour réduire le taux d'erreur. Il est préférable de privilégier les liens utiles à l'internaute, c'est-à-dire situés à des endroits faciles à identifier, entre des pages de thématique proche. Par exemple, c'est une mauvaise solution que de mettre des liens vers (presque) toutes les pages du site dans une page plan du site.
3. Si le site comporte un menu listant une grande partie des pages, il sera sans doute difficile de varier leurs ancres (car seul le 1er lien est pris en compte), mais il n'est pas nécessaire de changer le menu si celui-ci est utile tel quel pour les visiteurs.
4. En résumé, il n'est pas toujours possible ou facile d'éviter que certaines pages n'aient qu'une seule ancre dans leurs backlinks internes, par contre il faut essayer de le faire pour les pages stratégiques.

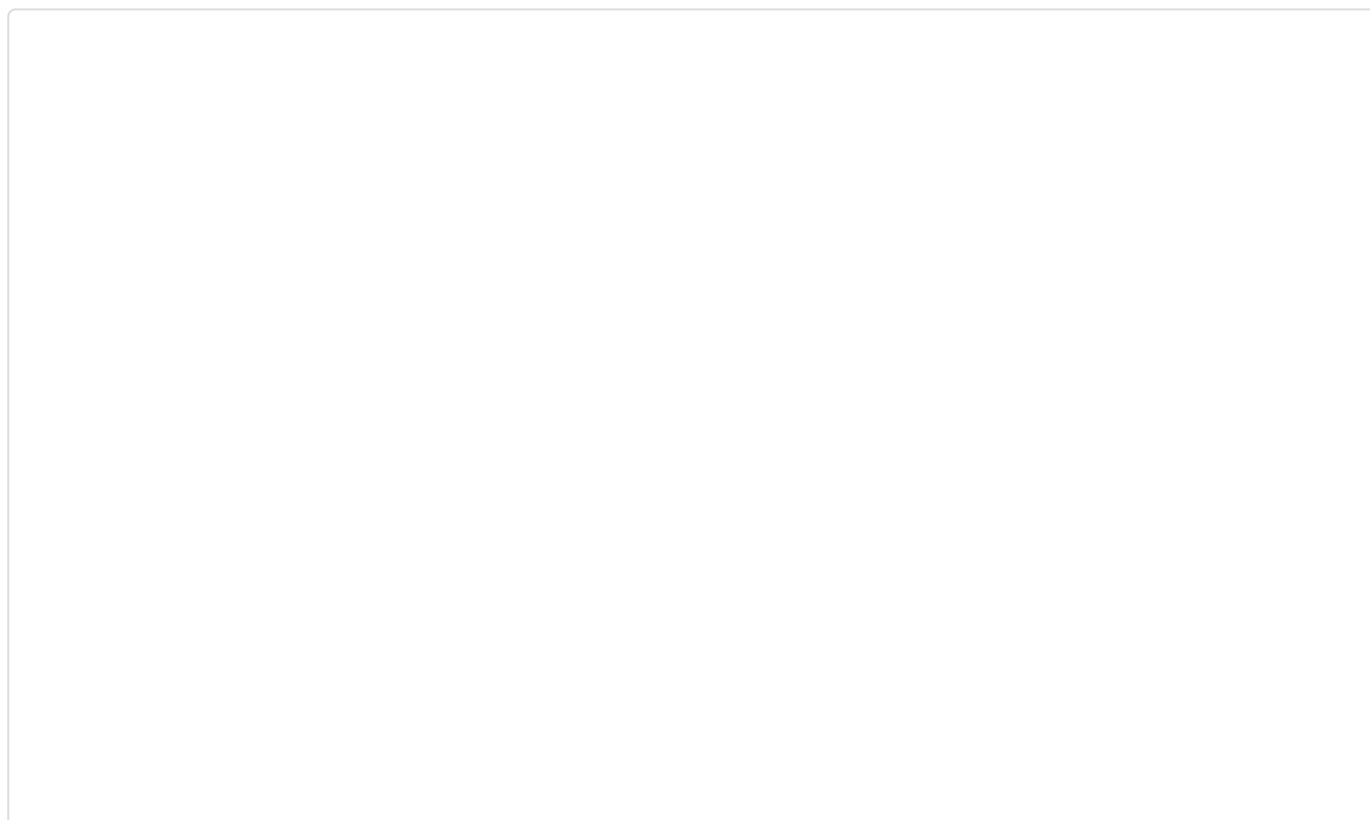
L'étude a porté sur les pages qui reçoivent trop peu de backlinks internes ou qui ont trop peu de variété dans les ancres de liens internes, ou qui ont des ancres vides (liens follow uniquement).

Pour savoir comment les liens sont comptés et identifiés, [cliquez ici](#). L'étude a porté sur les pages HTML indexables.

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages avec seulement 1 backlink interne	0	n/a	
Nb de pages avec seulement 2 backlinks internes	n/a	1	
Nb de pages avec toujours la même ancre (backlinks internes)	4	n/a	
Nb de pages avec seulement 2 ancres distinctes (backlinks internes)	n/a	23	
Nb de pages avec un ou des backlinks internes à ancre vide	n/a	13	
Nb total d'erreurs et d'avertissements	4	28	
Nb de pages indexables			44
Taux d'erreur	9,09 %	63,64 %	

Le graphique ci-dessous illustre l'analyse du nombre de backlinks par page :



Le graphique ci-dessous illustre l'analyse du nombre d'ancres distinctes par page :

5.3 Conclusion sur la profondeur et le maillage interne

Pour optimiser votre référencement, vos pages ne doivent pas être trop profondes (nombre de clics depuis la page d'accueil) et leur maillage doit être efficace (plusieurs liens entrants avec des ancres variées, non vides).

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

Voici les résultats de l'analyse :

	Erreurs	
Nb de pages trop profondes ou avec un maillage interne insuffisant	4	
Nb de pages indexables		44
Taux d'erreur	9,09 %	

Annexe

Le fichier annexe comnplus_fr-2020-07-27_05-3_profondeur-maillage.xlsx contient les statistiques sur la profondeur, le nombre de backlinks internes et d'ancres distinctes.

6 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 ~70 caractères ; au-delà de 80, c'est considéré comme une erreur.

Voici les résultats de l'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	0	n/a	
Nb de pages avec seulement 2 mots dans title	0	n/a	
Nb de pages avec seulement 3 mots dans title	0	n/a	
Nb de pages avec entre 71 et 80 caractères dans title	n/a	0	
Nb de pages avec entre 81 et 100 caractères dans title	0	n/a	
Nb de pages avec + de 100 caractères dans title	0	n/a	
Nb total d'erreurs et d'avertissements	0	0	
Nb de pages indexables			44
Taux d'erreur	0 %	0 %	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs n'a pas été généré.

6.2 Doublons dans la balise title

Pour optimiser la balise title, il faut que chaque page ait une balise title unique (pas utilisée ailleurs sur le site), sinon c'est considéré comme une erreur.

Voici les résultats de l'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			44
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs 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	3	
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	3	
Nb de pages indexables			44
Taux d'erreur	0 %	6,82 %	

Annexe

Le fichier annexe comnplus_fr-2020-07-27_06-3_repetition-mots-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 l'analyse :

	Erreurs	Avertissements	
Nb de pages qui contiennent la marque "comnplus" dans title	n/a	6	
Nb de pages indexables			44
Taux d'erreur		13,64 %	

Annexe

Le fichier annexe comnplus_fr-2020-07-27_06-4_marque-dans-title.xlsx contient les données associées.

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	0	
Nb de pages indexables		44
Taux d'erreur	0 %	

7 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 ~230 caractères disponibles (environ 3 lignes). Il est conseillé de rédiger un texte d'au moins 90 caractères.

Voici les résultats de l'analyse :

	Erreurs	
Nb de pages avec balise meta description vide ou absente	0	
Nb de pages avec une meta description de moins de 90 caractères	0	
Nb total	0	
Nb de pages indexables		44
Taux d'erreur	0 %	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs n'a pas été généré.

7.2 Doublons dans la balise meta description

Pour optimiser la balise meta description, il faut qu'elle soit unique (pas utilisée ailleurs sur le site).

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			44
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs 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	0	
Nb de pages indexables		44
Taux d'erreur	0 %	

8.1 Balises H1-H6

Pour vous donner le plus de chances de réussir votre référencement, il est recommandé d'utiliser les balises de titres H1, H2, H3... H6 (notées "Hn"). Attention, ces balises ne doivent être intégrées en théorie que sur les pages à fort contenu éditorial : sur les autres (accueil, listings, etc.) il est généralement conseillé uniquement une balise H1.

Voici les règles qu'il est recommandé de suivre (considérées sinon comme des erreurs) :

- au minimum, une page doit avoir au moins une H1 (sauf la page d'accueil ou celle de chaque sous-domaine où ce n'est pas indispensable et donc considéré comme un avertissement)
- une balise H1 ne doit pas être vide

Voici d'autres règles qu'il est également recommandé de suivre (considérées sinon comme des avertissements) :

- la 1ère balise d'une page doit être une H1
- il faut éviter les sauts de niveau : quand il s'agit de descendre dans les détails, une balise Hn ne peut être suivie que par une balise H(n+1) dans l'ordre du code source, mais pas H(n+2) ou H(n+3) : on ne doit pas sauter de niveau. Par exemple, après une H1, il ne peut y avoir qu'une H2. Après une H3, il ne peut y avoir qu'une H4 (ou bien une autre H3, ou bien on remonte et c'est une H2), mais pas une H5.
- la taille d'une balise Hn (nb de caractères) ne doit pas dépasser 80 caractères (même si ce chiffre est arbitraire, déterminé par expérience, l'idée est qu'un titre doit rester un titre et non un paragraphe)
- il est recommandé de limiter à 1 le nombre de balises H1 par page
- pour les balises H2-H6 "isolées" : dans un niveau donné, le nombre de titres du niveau inférieur doit être 0 ou au moins 2 (donc pas seulement 1). La raison est que si vous utilisez une balise H2 (ou H3, H4, H5, H6), c'est a priori pour découper en plusieurs sous-parties le bloc de niveau supérieur. Il doit donc y en avoir soit aucune, soit au moins 2 dans une page.
- dans la mesure du possible, il faut éviter d'utiliser un même libellé dans plusieurs balises H1-H6 d'une même page
- pour profiter de la longue traîne (et d'une limitation moins forte en nombre de caractères), essayez d'avoir un titre H1 légèrement différent de celui utilisé pour la balise Title de la (même) page

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de pages sans balise H1	1	0	
Nb de pages avec une balise H1 mais vide	0	n/a	
Nb de pages où la 1ère balise Hn n'est pas une H1	n/a	0	
Nb de pages avec saut de niveau entre les Hn	n/a	2	
Nb de pages avec une balise Hn trop longue	n/a	0	
Nb de pages avec plusieurs balises H1	n/a	0	
Nb de pages avec balise H2-H6 isolée	n/a	2	
Nb de pages avec le même libellé dans plusieurs Hn	n/a	1	
Nb de pages avec le même libellé dans H1 et Title	n/a	0	
Nb de pages indexables			44
Taux d'erreur	2,27 %	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 comnplus_fr-2020-07-27_08-1_balises-h1-h6.xlsx contient les données associées.

8.2 Doublons dans la balise H1

Pour optimiser la balise H1, il faut que chaque page ait une balise H1 unique (pas utilisée ailleurs sur le site), sinon cela est considéré comme une erreur.

Voici les résultats de l'analyse :

	Erreurs	Avertissements	
Nb de balises H1 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			44
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs n'a pas été généré.

8.3 Taille du texte (zone principale de la page)

Comme vous le savez sans doute, l'objectif de l'algorithme Panda de Google est d'identifier les pages à faible valeur ajoutée. Panda est intégré au cœur de l'algorithme principal : même si Google n'indique plus de dates de mises à jour, il fonctionne chaque jour !

Comme vous le verrez plus loin dans ce rapport, un élément essentiel de la qualité est le contenu texte. Même si on trouve des contre-exemples, une page avec un contenu texte court est rarement de haute qualité (de nature à satisfaire tous les internautes). C'est ce qu'on appelle "thin content" dans le jargon SEO en anglais.

Une approche simpliste consiste à compter le nombre de mots dans l'ensemble de la page HTML (incluant l'entête, les menus, les barres latérales et le pied de page). Comme vous vous en doutez, cela ne fonctionne pas car ce qui compte, c'est la taille du contenu dans la zone principale de la page.

Grâce à un algorithme spécifique, ce rapport vous aide à identifier les pages sans doute trop courtes.

Dans l'annexe ci-dessous, sont listées les pages dont le contenu texte est estimé trop court. Pour vous aider à les trier, l'annexe indique un indice de gravité du problème, de 0 (aucun problème) à 100 (gros problème).

L'histogramme ci-dessous montre la répartition des pages de ce site selon leur indice de manque de contenu :

Le tableau ci-dessous donne la synthèse :

	Erreurs	Avertissements	
Nb de pages sans doute trop courtes	5	n/a	
Nb de pages indexables			44
Taux d'erreur	11,36 %	n/a	

Annexe

Le fichier annexe comnplus_fr-2020-07-27_08-3_taille-texte.xlsx contient les données associées.

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, Pinterest, Viadeo, Myspace, Youtube, Dailymotion, Instagram, 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	0	
Nb total	0	0	
Nb de pages indexables			44
Taux d'erreur	0 %	0 %	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs n'a pas été généré.

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			44
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs 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			44
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs 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			44
Taux d'erreur	0 %	n/a	

Annexe

Dans votre cas, comme aucune erreur n'a été détectée, le fichier annexe qui liste les erreurs 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	5	
Nb de pages indexables		44
Taux d'erreur	11,36 %	

9 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, Pinterest, Viadeo, Myspace, Youtube, Dailymotion, Instagram, Vimeo, Flickr.

9.1 Nombre total de liens sortants externes

24 liens externes ont été identifiés, pointant vers 10 URL distinctes.

Annexe

Le fichier annexe `comnplus_fr-2020-07-27_09-1_liens-sortants-externes.xlsx` contient les données associées.

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			10
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 Analyse des images

L'analyse des images consiste à étudier les balises sur l'ensemble des pages HTML qui ont pu être crawlées (code HTTP 200).

L'analyse des images n'a pas été faite car l'option n'était pas activée au lancement. Pour information, sur les 44 pages HTML qui ont pu être analysées dans cet audit, **195 images** ont été trouvées (195 sur le site et 0 sur d'autres sites).

11 Analyse de la qualité

Google améliore sans arrêt son algorithme afin de favoriser les contenus de qualité et pénaliser les autres. C'est un objectif majeur pour l'algorithme Panda, mais aussi pour d'autres (Phantom et dans une certaine mesure Penguin, ainsi que des modules sans nom intégrés au coeur de l'algorithme général).

En se basant sur des aspects techniques mais aussi des éléments du contenu qui nuisent à l'expérience utilisateur, le risque est calculé pour chaque page, noté QR pour "Quality Risk". Plus QR est élevé (le maximum étant 100), plus vous devriez étudier les recommandations fournies en annexe pour améliorer la qualité de la page.

Pour réduire le QualityRisk d'une page, suivez tous les conseils déjà listés précédemment dans ce rapport d'audit. En voici le résumé :

- temps de téléchargement (page brute hors images, JS, CSS...) : ne dépassez pas 600ms (essayez même d'atteindre 200ms)
- pas trop de profondeur (nombre de clics depuis l'accueil)
- un bon maillage interne (si possible plusieurs backlinks internes avec des textes d'ancrage variés)
- balise Title : spécifique (unique) à chaque page, ni trop courte (au moins 3 mots) ni trop longue (moins de 80 caractères), sans répétition de mots
- une balise meta description spécifique (unique) à chaque page, pas trop courte (au moins 90 caractères)
- au moins une balise H1 spécifique (unique) à chaque page, et s'il y a des balises H2-H6, elles doivent décrire l'architecture du contenu principal de la page
- le contenu : spécifique (unique) à chaque page, de taille suffisante
- les liens internes présents dans la page : il vaut mieux éviter d'en faire trop (il est recommandé de ne pas dépasser 300 liens follow)
- les liens sortants externes doivent pointer vers des pages accessibles

Comme vous l'avez compris, il ne s'agit que d'une première piste : à vous de compléter par d'autres analyses !

Pour l'ensemble des pages indexables de ce site, **le QualityRisk moyen est de 6,0** (le pire possible étant 100).

L'histogramme ci-dessous montre la répartition des pages du site en fonction de leur indice QualityRisk :

Il est conseillé de ne pas avoir de pages avec un QualityRisk supérieur à 20. Le tableau ci-dessous donne la synthèse :

	Erreurs	Avertissements	
Nb de pages avec un QualityRisk >= 20	5	n/a	
Nb de pages avec un QualityRisk entre 1 et 19	n/a	4	
Nb de pages indexables			44
Taux d'erreur	11,36 %	9,09 %	

Remarque : avant de supprimer une page identifiée comme étant à risque en termes de qualité, veuillez vérifier qu'elle ne correspond pas à l'un des 3 cas suivants (pour lesquels il ne faut pas supprimer la page mais l'améliorer) :

- si la page génère du trafic via le référencement naturel
- ou si la page a reçu des backlinks de qualité
- ou si après analyse humaine, vous estimez qu'en réalité la page n'est pas en risque

Annexe

Dans l'annexe, sont listées uniquement les pages ayant un QualityRisk supérieur ou égal à 5. Le fichier annexe comnplus_fr-2020-07-27_10_analyse-qualite.xlsx contient les données associées.

12 Conclusion

Avoir toutes les données mesurées et calculées sur les pages indexables est très précieux pour faire des analyses encore plus avancées.

Annexe

Le fichier annexe `comnplus_fr-2020-07-27_conclusion.xlsx` contient les données associées.

L'annexe ci-dessous liste toutes les URL crawlées avec pour chacune la raison de ce crawl (la plupart du temps il s'agit d'un lien mais ça peut aussi être une URL canonique ou une redirection par exemple). Dit autrement, cette annexe liste tous les backlinks internes follow de toutes les URL étudiées, avec l'ancre de chacun.

Annexe

Le fichier annexe `comnplus_fr-2020-07-27_conclusion-maillage-interne.xlsx` contient les données associées.

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 27 juillet 2020 et qu'il faut corriger au plus vite.

12.1 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, aucune erreur de ce type n'a été trouvée ([détails](#)).

Il faut éviter de faire des liens internes vers des URL non indexables. À la date de cette analyse, aucune erreur de ce type n'a été trouvée ([détails](#)).

12.2 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, 4,55 % des URL HTML code 200 ont été identifiées comme étant trop lentes, ce qui n'est pas optimal ([détails](#)).

12.3 Problèmes de profondeur des pages ou de maillage interne

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](#)).

Il faut aussi que les pages reçoivent plusieurs liens entrants internes, avec des ancres distinctes. L'analyse a montré que 9,09 % des pages n'ont pas des liens entrants internes assez efficaces.

12.4 Problèmes 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
contenu (détails)	11,36 %
balise title (détails)	0 %
balise meta description (détails)	0 %

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	5
Nb de pages indexables	44
Taux d'erreur global	11,36 %

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.

12.5 Problèmes sur les liens sortants externes

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

12.6 Quand refaire un audit 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)

13.1 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
- Si un lien autorisé au crawl pointe vers une URL qui n'est pas du HTML (par exemple un PDF mais aussi une image), cette URL sera crawlée. Par contre, l'essentiel de l'analyse se faisant sur les pages HTML, ces URL seront exclues de la plupart des éléments de l'audit.
- Si l'option d'analyse des images est activée, les images étudiées sont celles trouvées dans l'attribut src de la balise , quand cet attribut indique une URL (donc pas sous la forme data:xxx). Les autres attributs de la balise sont ignorés, tout comme les images définies exclusivement par les CSS.
- Pour calculer le nombre de liens entrants internes d'une page (backlinks) :
 - sont recherchées les pages (indexables ou pas) qui ont une balise <a> qui amène au final vers la page qu'on étudie (soit de façon directe, soit via une ou plusieurs redirections)
 - si une page A fait un lien vers elle-même, cela ne compte pas comme un backlink
 - si une page A fait un lien vers B qui redirige vers C, alors A est comptée comme un backlink de C
 - si une page A fait plusieurs liens vers une même page B, alors cela compte comme un seul backlink. Attention : dans ce cas, seul le 1er lien est pris en compte (à condition qu'il soit follow)
 - l'URL canonique n'est pas prise en compte ici, uniquement les liens avec une balise <a href>

13.2 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

13.3 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

13.4 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 SEO	1
1 Introduction	2
1.1 Objectif de ce rapport	2
Annexes	3
1.2 Comment exploiter cet audit	3
2 Synthèse des résultats	4
2.1 Score de performance	4
2.2 Pages zombies	4
2.3 QualityRisk	4
2.4 Taux d'erreur	4
2.5 Performance	5
3 Crawl et indexabilité	7
3.1 Bilan des URL bloquées par le(s) fichier(s) robots.txt	7
Annexe	7
Annexe	7
3.2 Nombre total d'URL crawlées	7
3.3 État des URL crawlées	8
Annexe	9
3.4 URL trouvables uniquement par des liens internes nofollow	9
3.5 Types MIME	9
Annexe	9
3.6 Pages HTML crawlées sans erreur	10
Annexe	10
3.7 Pages HTML indexables par les moteurs	10
4 Temps de téléchargement	12
Annexe	14
5 Profondeur des pages et maillage interne	15
5.1 Analyse de la profondeur des pages	15
5.2 Analyse du maillage interne	16
5.3 Conclusion sur la profondeur et le maillage interne	19
Annexe	19
6 Balise title	20
6.1 Taille de la balise Title	20
Annexe	20
6.2 Doublons dans la balise title	20
Annexe	21
6.3 Répétition de mots dans la balise Title	21
Annexe	22
6.4 Votre marque dans la balise Title	22
Annexe	22
6.5 Conclusion sur la balise Title	22
7 Balise meta description	23
7.1 Taille de la balise meta description	23
Annexe	23
7.2 Doublons dans la balise meta description	23
Annexe	24
7.3 Conclusion sur la balise meta description	24
8 Contenu des pages	25
8.1 Balises H1-H6	25
Annexe	26
8.2 Doublons dans la balise H1	26
Annexe	27
8.3 Taille du texte (zone principale de la page)	27
Annexe	28
8.4 Analyse du nombre de liens par page	28
Annexe	29
8.5 Analyse des contenus dupliqués internes	29
Annexe	30

8.6 Identifiants de session dans les URL	30
Annexe	30
8.7 Paramètres de tracking dans les URL	30
Annexe	31
8.8 Conclusion sur le contenu des pages	31
9 Liens sortants externes	32
9.1 Nombre total de liens sortants externes	32
Annexe	32
9.2 Erreurs sur les liens sortants externes	32
10 Analyse des images	33
11 Analyse de la qualité	34
Annexe	35
12 Conclusion	36
Annexe	36
Annexe	36
12.1 Problèmes de crawl et d'indexation	36
12.2 Problèmes de temps de téléchargement	36
12.3 Problèmes de profondeur des pages ou de maillage interne	36
12.4 Problèmes concernant le contenu	37
12.5 Problèmes sur les liens sortants externes	37
12.6 Quand refaire un audit SEO ?	37
13 Explications complémentaires	39
13.1 Fonctionnement de l'audit	39
13.2 Signification des codes HTTP	39
13.3 Signification des types MIME	40
13.4 Signification des termes SEO	41
Table des matières	42