
Audit Technique SEO

Site analysé : vins-independance.com

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









Ce document a pour objectif de vous présenter les résultats de notre pré-audit technique (sous l'angle du référencement naturel) de votre site **vins-independance.com**.

L'audit de votre site a eu lieu le **29 janvier 2016** et a porté sur **2 297 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 : vins-independance.com
- URL de départ : <http://www.vins-independance.com/>
- Protocoles : aucune restriction (HTTP et HTTPS)
- 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.vins-independance.com : non

2 Synthèse des résultats obtenus

Les critères techniques SEO analysés pendant le pré-audit ont été regroupés en 8 indices dont voici les résultats pour votre site :

Indice	Optimisation	Erreurs constatées sur les pages	
		en pourcentage	en nombre de pages
Pages répondant en erreur		15,32 % ⁽¹⁾	352
Pages non indexables		0,31 % ⁽²⁾	6
Pages trop lentes		71 % ⁽²⁾	1 381
Pages trop profondes		5,98 % ⁽³⁾	116
Titres non optimisés		45,90 % ⁽³⁾	890
Description non optimisée		39,76 % ⁽³⁾	771
Contenus non optimisés		14,29 % ⁽³⁾	277
Liens sortants externes en erreur		20 % ⁽⁴⁾	1

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

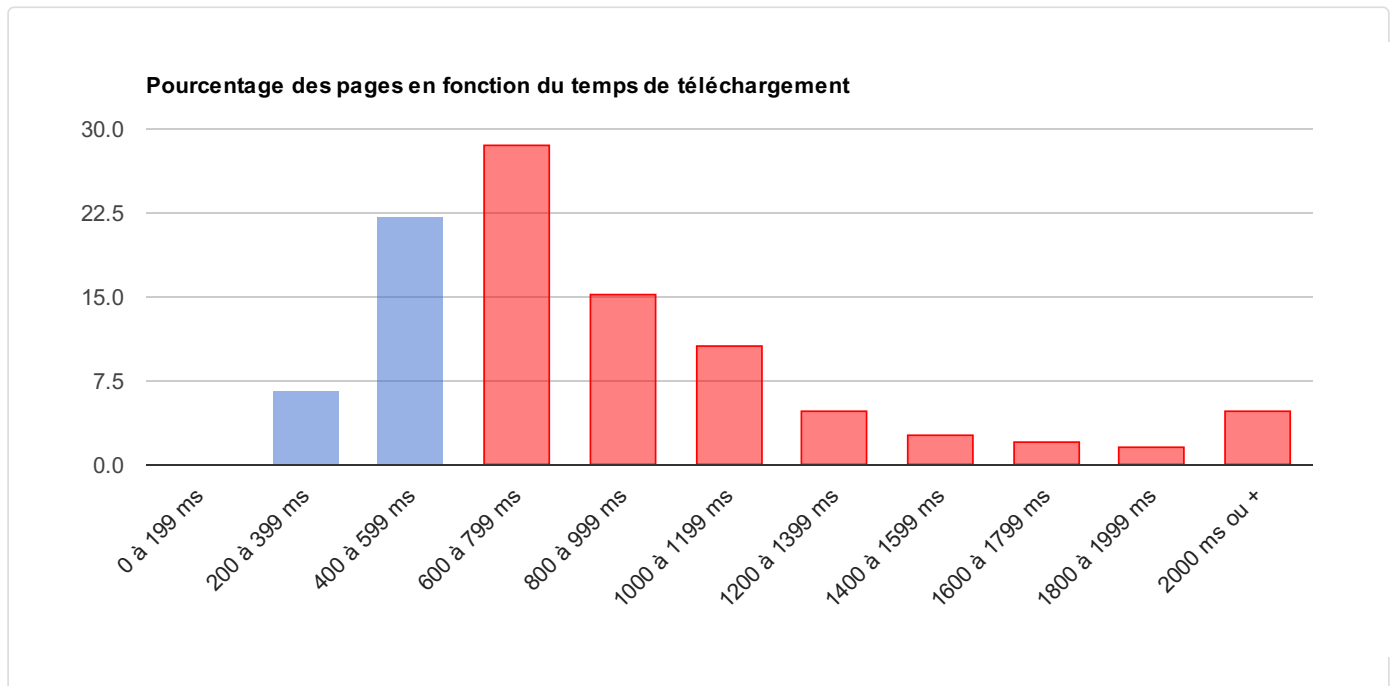
Comme vous pouvez le constater, ce pré-audit a révélé un certain nombre de problèmes techniques. Le bon côté des choses, c'est que vous disposez d'une marge de progression !

Pour avoir un bon référencement sur Google, vous devez absolument corriger tous ces points : c'est une première étape indispensable. Par la suite, il faudra produire du contenu de qualité, optimiser l'expérience utilisateur et acquérir des bons backlinks (liens vers votre site).

3 Quelques exemples des problèmes identifiés

3.1 Pages trop lentes

Au cours de notre analyse, nous avons identifié trop de pages trop lentes à télécharger. Attention, ceci peut dégrader le crawl de Google ainsi que l'expérience utilisateur et donc au final votre référencement. Pour information, voici la répartition de vos pages par temps de téléchargement (HTML brut, hors fichiers externes tels que les images, CSS et Javascript) :



Les pages suivantes étaient vraiment trop lentes à télécharger :

URL	Temps de téléchargement
http://www.vins-independance.com/	2019 ms
http://www.vins-independance.com/2-vins	2157 ms
http://www.vins-independance.com/4-couleurs	3226 ms

Il ne s'agit ici que de 3 exemples, mais le rapport complet détaille les **1 381 pages trop lentes** avec le temps de téléchargement pour chacune (ce que ne fournit pas Google Search Console).

3.2 Pages trop profondes

Pour qu'une page soit efficace en référencement naturel, il faut éviter qu'elle soit trop profonde dans le site. Notre algorithme a évalué **la profondeur maximale recommandée pour votre site**, valeur que vous trouverez dans le rapport complet. Par exemple, les 3 pages suivantes sont trop profondes :

URL	Profondeur
http://vins-independance.com/blog/tag/cave-orvault/page/3/	5
http://www.vins-independance.com/2-vins/categories-selection_du_mois/2013-1997?p=8	4
http://vins-independance.com/blog/tag/degustation-oenologique-nantes/page/3/	5

Il ne s'agit ici que de 3 exemples, mais le rapport complet détaille les **116 pages trop profondes** dans votre site, compte-tenu de la profondeur optimale que nous avons déterminée spécifiquement pour votre site.

3.3 Pages avec titres non optimisés

La balise TITLE est un élément prioritaire d'optimisation du référencement. Nous avons trouvé sur votre site un trop grand nombre de pages comportant des erreurs à ce niveau. Voici 3 exemples :

URL	Balise title	Problème identifié
http://vins-independance.com/blog/do...		balise title absente ou vide
http://www.vins-independance.com/18...	Roussillon Roc des Anges, Mas Amie...	balise title sur-optimisée
http://www.vins-independance.com/9-v...	Vins de Bordeaux Paul Barre, Châtea...	balise title sur-optimisée

Il ne s'agit ici que de 3 exemples, mais le rapport complet détaille les **890 pages pouvant être optimisées au niveau de la balise TITLE**.

3.4 Pages avec une description non optimisée

La balise meta description joue un rôle majeur dans l'incitation au clic dans les pages de résultats de Google, c'est pourquoi il est préférable de bien les rédiger. Nous avons trouvé sur votre site un trop grand nombre de pages comportant des erreurs à ce niveau. Voici 3 exemples :

URL	Balise meta description	Problème identifié
http://www.vins-independance.com/5-c...		balise meta description absente ou vide
http://www.vins-independance.com/acc...	Commande	balise meta description trop courte
http://www.vins-independance.com/30...	Domaine Ganevat	balise meta description trop courte

Il ne s'agit ici que de 3 exemples, mais le rapport complet détaille les **771 pages pouvant être optimisées au niveau de la balise meta description**.

3.5 Pages avec un contenu non optimisé

Sans contenu, une page a peu de chances d'avoir des bonnes performances en termes de référencement. Sans aller jusqu'à une analyse sémantique ou linguistique, il est possible d'identifier des problèmes graves à ce niveau. Nous avons trouvé sur votre site un trop grand nombre de pages comportant des erreurs à ce niveau. Voici 3 exemples :

URL	Problème identifié
http://vins-independance.com/blog/tag/sancerre/	page sans doute trop courte
http://www.vins-independance.com/our-domains.php	page sans aucune balise H1
http://vins-independance.com/blog/du-vin-de-copains/	page sans aucune balise H1

Il ne s'agit ici que de 3 exemples, mais le rapport complet détaille les **277 pages pouvant être optimisées au niveau du contenu**.

4.1 Avantages de la version complète

Ce rapport de pré-audit fournit uniquement une synthèse de l'ensemble des critères analysés, ce qui vous permet de savoir si vous avez besoin de la version complète. Cette dernière vous permettra de **comprendre et corriger tous les problèmes** techniques SEO, ainsi que d'**optimiser** votre site.

La version complète comprend :

- une méthodologie complète avec des explications détaillées sur les problèmes rencontrés,
- la totalité des critères les plus efficaces (plusieurs dizaines), avec les erreurs rencontrées sur chacun,
- la liste exhaustive des erreurs (fichiers Excel contenant des dizaines de lignes, ou des dizaines de milliers selon les cas),

4.2 Fonctionnalités de la version complète

Voici les points audités faisant l'objet d'explications détaillées et d'annexes exhaustives (données aux formats Excel et CSV) :

- Analyse du **crawl** :
 - liste des URL en erreur (codes 301, 302, 404, 410, 500, etc.)
 - liste des URL faisant un lien vers une URL en erreur
 - répartition des documents rencontrés par type MIME (HTML, PDF, XML, images, etc.)
 - tableau de toutes les pages du site sans erreur, avec pour chacune URL, balise title, balise meta description, balise d'URL canonique, URL canonique via en-tête HTTP, balise meta robots, en-tête HTTP X-Robots-Tag
- Analyse de l'**indexabilité** :
 - liste des URL interdites d'indexation
 - liste des URL ayant une URL canonique différente de l'URL crawlée
 - liste des URL faisant un lien vers une URL non indexable
- Analyse du **temps de téléchargement** :
 - tableau et graphiques récapitulatifs des temps de téléchargement des pages
 - tableau exhaustif du temps de téléchargement de chaque page HTML code 200
- Analyse de la **profondeur** des pages :
 - détermination du niveau de profondeur maximale recommandé pour votre site
 - tableau et graphiques récapitulatifs des niveaux de profondeur des pages indexables
 - tableau exhaustif du niveau de profondeur de chaque page HTML indexable
- Analyse de la **balise title** :
 - liste des URL sans balise title ou avec une balise title vide
 - liste des URL avec une balise title trop courte (3 types)
 - liste des URL avec une balise title trop longue
 - liste des URL utilisant une balise title non unique (dupliquée ailleurs sur le site)
 - liste des URL avec une balise title sur-optimisée (3 types)
 - liste des URL incluant votre marque dans la balise title
- Analyse de la **balise meta description** :
 - liste des URL sans balise meta description ou avec une balise meta description vide
 - liste des URL avec une balise meta description trop courte (3 types)
 - liste des URL utilisant une balise meta description non unique (dupliquée ailleurs sur le site)
- Analyse du **contenu** des pages :
 - liste des URL sans aucune balise H1
 - détermination de la taille minimale des pages (en nombre de mots) recommandée pour votre site, en fonction de votre template
 - liste des URL avec un contenu vraiment trop court (a priori des erreurs)
 - liste des URL avec un contenu trop court
 - liste des URL sans aucune mise en exergue
 - liste des URL non optimales en termes de liens sortants internes et externes (4 analyses spécifiques)
 - liste des URL incluant un identifiant de session
 - liste des URL incluant des paramètres de tracking
 - liste des URL ayant un contenu dupliqué (strictement le même qu'une autre page interne)
- Analyse des **liens sortants externes** :
 - identification du nombre d'URL externes (hors réseaux sociaux) recevant un lien follow depuis votre site
 - identification du nombre d'URL du site faisant un lien sortant externe follow (hors réseaux sociaux)
 - liste de tous les liens sortants externes follow, avec URL source, URL cible et anchor text (texte cliquable)
 - chaque URL externe est vérifiée afin de lister toutes les erreurs trouvées (codes 301, 302, 404, 410, 500, etc.)

5 Explications complémentaires

5.1 Fonctionnement de l'audit

Pour chaque audit, une analyse exhaustive de votre 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

5.2 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

5.3 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

5.4 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
2 Synthèse des résultats obtenus	3
3 Quelques exemples des problèmes identifiés	4
3.1 Pages trop lentes	4
3.2 Pages trop profondes	4
3.3 Pages avec titres non optimisés	5
3.4 Pages avec une description non optimisée	5
3.5 Pages avec un contenu non optimisé	5
4 La version complète de l'audit	6
4.1 Avantages de la version complète	6
4.2 Fonctionnalités de la version complète	6
5 Explications complémentaires	7
5.1 Fonctionnement de l'audit	7
5.2 Signification des codes HTTP	7
5.3 Signification des types MIME	7
5.4 Signification des termes SEO	8
Table des matières	9