

# Audit SEO complet & plan d'action

---

[www.cleanea.com](http://www.cleanea.com)

---

Diagnostic technique, analyse on-page, étude concurrentielle, stratégie de mots-clés et feuille de route opérationnelle. Un document conçu pour être exécuté, et pour être reproduit.

**Client** : Cleanea (CleanTech B2B, Puteaux)

**Préparé par** : Mo, Chargé de Marketing Digital

**Date** : 7 juin 2026 · **Version** : 1.0

**Périmètre** : 39 URLs publiques, 21 pages analysées en profondeur

**58**/100  
SCORE DE SANTÉ SEO

## DOCUMENT

# Sommaire

---

## SYNTHÈSE &amp; CADRAGE

01	<b>Sommaire exécutif</b>	3
02	<b>Méthodologie &amp; périmètre</b>	5
03	<b>Score de santé SEO</b>	6

## PHASE 1 · DIAGNOSTIC

04	<b>Diagnostic du site</b>	7
	1.1 · Architecture & inventaire	7
	1.2 · Audit technique	8
	1.3 · Audit on-page page par page	8
	1.4 · Données structurées	11
	1.5 · Accessibilité & images	11
	1.6 · Maillage interne	12
05	<b>Analyse concurrentielle</b>	13

## PHASES 2 &amp; 3 · STRATÉGIE &amp; EXÉCUTION

06	<b>Stratégie de mots-clés</b>	15
07	<b>Plan d'optimisation</b>	18
	3.1 · Corrections techniques	18
	3.2 · Réécritures title & meta	19
	3.3 · Plan éditorial (12 articles)	19
	3.4 · Données structurées à déployer	20

## PHASE 4 &amp; PILOTAGE

08	<b>Suivi &amp; KPIs</b>	22
09	<b>Feuille de route priorisée</b>	24

## ANNEXES

A	<b>Protocole reproductible pour Hermes</b>	26
B	<b>Checklist SEO</b>	31

## SYNTHÈSE

# Sommaire exécutif

Le site Cleanea.com repose sur une base technique saine et un travail on-page récent de bonne qualité (balises title et meta descriptions cohérentes, structure par secteurs, contenu dense). Il reste néanmoins bridé par trois leviers : une stratégie de contenu quasi inexistante (un seul article de blog), des données structurées sous-exploitées, et des défauts techniques hérités du template Wix (titres H1 multiples, images sans attribut alt, hreflang incohérent). Le potentiel de progression organique est important et largement à portée.

**39**

URLS PUBLIQUES AU SITEMAP

**1**

ARTICLE DE BLOG PUBLIÉ

**~1 263**

MOTS / PAGE (MOYENNE)

**37 %**

IMAGES SANS ATTRIBUT ALT

## Les 5 constats majeurs

#	Gravité	Constat	Impact
01	CRITIQUE	<b>Stratégie de contenu absente.</b> Un seul article de blog en ligne. Tout le haut de funnel (requêtes problèmes, réglementation, comparatifs) est laissé aux concurrents et aux médias.	Invisibilité sur 80 % des intentions de recherche du secteur.
02	ÉLEVÉ	<b>Données structurées sous-exploitées.</b> 17 pages sur 21 sans aucun balisage Schema.org. La FAQ et les blocs Q/R produits ne sont pas balisés en FAQPage.	Aucun rich result, perte de surface dans la SERP.
03	ÉLEVÉ	<b>Hygiène technique du template.</b> 8 pages sur 21 ont plusieurs balises H1, l'en-tête HTTP renvoie content-language: en-US sur un site français, URLs accentuées encodées.	Signaux de structure et de langue brouillés.
04	MOYEN	<b>Performance et poids des pages.</b> Site Wix, HTML brut de la page d'accueil proche de 2 Mo. Pas de données terrain Core Web Vitals (trafic insuffisant).	Risque sur le LCP mobile, frein à l'expérience.
05	MOYEN	<b>Cannibalisation et incohérences.</b> Deux pages de contact (/contact et /nous-contacter), hreflang manquant sur 3 pages, version anglaise /en active mais non maîtrisée.	Dilution du signal, budget de crawl gaspillé.

## ■ Forces, faiblesses, opportunités

Forces (à préserver)	Faiblesses (à corriger)	Opportunités (à saisir)
<p>Titles et metas récents, bien rédigés et cadrés.</p> <p>Structure claire par secteurs (4 pages dédiées).</p> <p>Contenu dense (1 263 mots/page).</p> <p>Position de pionnier reconnue par les médias.</p> <p>HTTPS, mobile-friendly, CDN rapide (Fastly).</p>	<p>Blog quasi vide, aucun cluster éditorial.</p> <p>Schema.org quasi absent.</p> <p>H1 multiples, alt manquants (37 %).</p> <p>Poids des pages élevé.</p> <p>Doublons et hreflang incohérent.</p>	<p>Capter les requêtes "sans chimie", "EN 14476", "appel d'offres RSE".</p> <p>Rich results FAQ sur tout le site.</p> <p>Pages secteurs déjà prêtes à ranker.</p> <p>Autorité thématique sur l'électrolyse.</p> <p>Données chiffrées RSE valorisables en contenu.</p>

## ■ Les 4 priorités du trimestre

Pri.	Action	Effort	Impact	Horizon
P1	Lancer un calendrier éditorial (12 articles ciblés haut et milieu de funnel)	Élevé	Très élevé	0 à 90 jours
P2	Déployer les données structurées (FAQPage, Organization, BreadcrumbList, Product)	Moyen	Élevé	0 à 30 jours
P3	Corriger l'hygiène technique (H1 uniques, alt, content-language, hreflang)	Faible	Élevé	0 à 30 jours
P4	Connecter Search Console + résoudre les doublons de pages	Faible	Moyen	0 à 15 jours

### LECTURE DU DOCUMENT

Ce rapport suit la méthodologie SEO en 4 phases (Diagnostic, Stratégie, Optimisation, Suivi). L'annexe A fournit un protocole reproductible, pensé pour qu'Hermès puisse régénérer ce type d'audit de façon autonome à l'avenir.

## CADRE

# Méthodologie & périmètre

Cet audit a été réalisé sur données réelles, par crawl direct du site et analyse de la SERP, sans extrapolation. Les outils gratuits et l'analyse manuelle suffisent à produire un diagnostic fiable et actionnable.

## ■ Ce qui a été analysé

Dimension	Méthode appliquée
Architecture & indexation	Lecture du robots.txt, des 4 sous-sitemaps et du sitemap anglais, inventaire complet des URLs.
Audit technique	Récupération du HTML brut de chaque page (équivalent Screaming Frog), en-têtes HTTP, balises canoniques, hreflang, langue, viewport.
On-page	Extraction systématique : title, meta description, hiérarchie Hn, comptage de mots, attributs alt, liens internes.
Données structurées	Détection et typage des blocs JSON-LD Schema.org sur chaque page.
Performance	Mesure du poids réel et du TTFB par page, vérification de la disponibilité des données terrain Core Web Vitals (CrUX).
Concurrence & mots-clés	Analyse des SERP sur les requêtes commerciales et éditoriales clés, identification des concurrents directs (électrolyse) et indirects (ozone, vapeur).

## ■ Outils mobilisés

Crawl HTML (curl + parsing)

Analyse SERP web

Sitemaps XML

En-têtes HTTP

Mesure poids / TTFB

Méthodologie SEO 4 phases

## ■ Limites & prérequis

### DONNÉES PROPRIÉTAIRES NON ENCORE DISPONIBLES

Cet audit n'inclut pas les données de Google Search Console (requêtes réelles, impressions, clics, position moyenne) ni de Google Analytics, qui ne sont pas connectés au moment de l'analyse. La connexion de Search Console est le prérequis n°1 : elle transformera ce diagnostic basé sur l'observation en pilotage basé sur la donnée réelle. Les volumes de recherche cités sont des estimations qualitatives (absence d'outil payant type Ahrefs/Semrush), à confirmer une fois Search Console alimenté (48 à 72 h).

**Note pour la reproduction.** L'ensemble de la démarche, des commandes de crawl jusqu'au barème de notation, est documenté dans l'**Annexe A · Protocole reproductible pour Hermes**. L'objectif est qu'un agent autonome puisse rejouer cet audit sur n'importe quelle URL, en suivant exactement le même protocole.

TABLEAU DE BORD

# Score de santé SEO

Évaluation pondérée sur 7 piliers. Le score global de **58/100** traduit un site "à consolider" : un socle technique correct et un on-page récent de qualité, mais deux angles morts majeurs (contenu et données structurées) qui plafonnent la performance organique.

## 58

SCORE GLOBAL / 100

Échelle : 0-44 critique · 45-59 à consolider · 60-74 correct · 75-100 solide

- Indexation & crawl** **74** Sitemaps OK, mais content-language et URLs à revoir
- Performance / CWV** **55** Pages lourdes (Wix), pas de données terrain
- Architecture & maillage** **70** Bon maillage, doublons de pages à traiter
- On-page (title/meta/Hn)** **68** Titres soignés, mais H1 multiples
- Contenu & sémantique** **48** Un seul article, haut de funnel vide
- Données structurées** **35** Absentes sur 17 pages / 21
- Accessibilité & UX** **52** 37 % d'images sans alt, Hn vides

## ■ Lecture du score

Pilier	Score	Diagnostic en une phrase
Indexation & crawl	74	Le site est correctement crawlable et soumis (sitemaps Wix), mais l'en-tête de langue erroné et les URLs accentuées encodées brouillent les signaux.
Performance / CWV	55	Réponse serveur rapide grâce au CDN, mais le poids des pages (HTML proche de 2 Mo) menace le LCP mobile ; aucune donnée CrUX faute de trafic.
Architecture & maillage	70	Arborescence lisible et bon maillage interne (19 à 29 liens/page), pénalisés par des doublons fonctionnels.
On-page	68	Balises title et meta de bonne facture et bien dimensionnées, mais hiérarchie H1 non maîtrisée et quelques dépassements de longueur.
Contenu & sémantique	48	Pages existantes denses, mais couverture sémantique très partielle : aucune réponse aux requêtes informationnelles et comparatives.
Données structurées	35	Quasi inexistantes hors page d'accueil et fiches produit : la SERP enrichie est hors d'atteinte en l'état.
Accessibilité & UX	52	Manque massif d'attributs alt et présence de titres vides : pénalisant pour l'accessibilité comme pour la compréhension du contenu par Google.

### LE BON RÉFLEXE

Les piliers les plus faibles (contenu, données structurées) sont aussi ceux où l'effort produit le plus de résultat rapidement. C'est exactement là que se concentre le plan d'action des phases 2 et 3.

## PHASE 1

# Diagnostic

État des lieux complet du site : ce qu'il contient, comment il est construit techniquement, et la qualité de chaque page. Cette phase établit la ligne de base sur laquelle s'appuieront les recommandations.

## ■ 1.1 · Architecture & inventaire

Le site est hébergé sur **Wix** (serveur Pepyaka, ressources sur parastorage / wixstatic) et diffusé via le CDN Fastly. Il déclare un index de sitemaps découpé en quatre familles, plus une version anglaise parallèle.

**24**

PAGES ÉDITORIALES

**10**

FICHES PRODUIT

**4**

CATÉGORIES BOUTIQUE

**1**

ARTICLE DE BLOG

L'arborescence fonctionnelle se lit clairement : une page d'accueil, une page technologie, une page machine (Europa V7), une page chapeau "Nos secteurs" déclinée en 4 pages sectorielles (hôtellerie, santé, tertiaire, collectivités), une page produits, des fiches produit, et des pages d'engagement (RSE, filière réemploi, à-propos, nos clients). C'est une bonne fondation. Le point noir est le déséquilibre : la dimension commerciale est riche, la dimension éditoriale (blog) est vide.

### VERSION ANGLAISE NON MAÎTRISÉE

Un sitemap en `_en-sitemap.xml` et des pages `/en/...` sont actifs (l'URL `/en` répond bien en 200). Si cette version n'est pas réellement traduite et entretenue, elle génère des pages dupliquées ou pauvres qui diluent le signal. À auditer puis trancher : entretenir sérieusement, ou désindexer.

## 1.2 · Audit technique

Élément	État	Détail & recommandation
HTTPS	OK	HSTS actif, certificat valide.
robots.txt	OK	Bien configuré (blocage des paramètres ?lightbox=, crawl-delay sur bots agressifs, sitemaps déclarés).
Sitemaps XML	OK	Index + 4 sous-sitemaps à jour. Soumettre dans Search Console une fois connecté.
Mobile-friendly	OK	Balise viewport présente sur 100 % des pages.
En-tête de langue	À corriger	L'en-tête HTTP renvoie content-language : en-US alors que <html lang="fr"> et le contenu sont en français. Incohérence à régler dans les réglages Wix de langue par défaut.
Balises canoniques	OK	Présentes et auto-référencées sur les pages testées.
Hreflang	Incohérent	Implémenté sur la majorité des pages (x-default, fr-fr, en-gb) mais absent sur /blog, l'article de blog et la fiche /solution-rouge-désinfectante (alors que la Solution Bleue l'a). À homogénéiser.
URLs	À nettoyer	URLs accentuées encodées en sortie (/à-propos vers /%C3%A0-propos) et incohérence (/proprete-sante sans accent vs /galerie-santé avec accent). Préférer des slugs ASCII sans accent.
Performance	À surveiller	HTML brut très lourd (page d'accueil proche de 2,0 Mo, page produits proche de 2,1 Mo) avant JS et images ; 20 à 50 images par page. TTFB bon (0,14 à 0,64 s, CDN).
Core Web Vitals (terrain)	Indispo.	Aucune donnée CrUX : le trafic réel est insuffisant pour que Google agrège les métriques. C'est en soi un indicateur de faible visibilité actuelle.

### ANOMALIE DE LANGUE À TRAITER EN PRIORITÉ

Le couple content-language : en-US (en-tête HTTP) + version /en active peut amener Google à hésiter sur la langue/cible du site. Pour une entreprise qui vend en France, le signal doit être 100 % cohérent : langue française affirmée partout, version anglaise soit assumée et balisée proprement, soit retirée de l'index.

## 1.3 · Audit on-page page par page

Extraction réelle des balises sur les 21 pages analysées. Lecture des colonnes : Ti = longueur du titre (cible : 60 caractères maximum), Me = longueur de la meta description (cible 150 à 160), H1 = nombre de balises H1 (cible : 1), Mots = volume de texte.

Page / URL	Ti	Me	H1	Mots	Signaux à corriger
<b>Accueil</b> /	43	157	2	1298	<b>2 H1</b>
<b>Technologie</b> /technologie	41	153	2	1104	<b>2 H1 0 schema</b>
<b>Nos secteurs</b> /nos-secteurs	37	156	1	962	<b>1 H1 0 schema</b>
<b>Professionnels</b> /professionnels	54	160	1	1576	<b>1 H1 0 schema</b>
<b>Produit (detergent-desinfectant)</b> /produit-detergent-desinfectant	58	225	1	1248	<b>1 H1 meta longue 0 schema 20/23 sans alt</b>
<b>Machine Europa V7</b> /machine-europa-v7	54	156	4	1634	<b>4 H1 0 schema</b>
<b>Proprete hotellerie</b> /proprete-hotellerie	38	157	2	1302	<b>2 H1 0 schema</b>
<b>Proprete sante</b> /proprete-sante	47	158	1	1125	<b>1 H1 0 schema</b>
<b>Proprete tertiaire</b> /proprete-tertiaire	52	154	1	1220	<b>1 H1 0 schema</b>
<b>Proprete collectivites</b> /proprete-collectivites	54	162	1	1203	<b>1 H1 0 schema</b>
<b>RSE</b> /rse	51	164	2	1515	<b>2 H1 0 schema</b>
<b>Filiere reemploi</b> /filieres-reemploi	26	106	1	984	<b>1 H1 0 schema</b>
<b>A propos</b> /à-propos	62	148	3	1373	<b>3 H1 title long 0 schema</b>
<b>Nos clients</b> /nos-clients	21	121	1	935	<b>1 H1 0 schema 45/50 sans alt</b>
<b>FAQ</b> /faq	13	154	1	2001	<b>1 H1 0 schema</b>
<b>Blog</b> /blog	14	0	1	894	<b>1 H1 meta absente 0 schema 0 hreflang</b>
<b>Nous contacter</b> /nous-contacter	24	157	1	891	<b>1 H1 0 schema</b>
<b>Contact</b> /contact	31	165	1	806	<b>1 H1 0 schema</b>
<b>Produit Solution Rouge</b> /solution-rouge-desinfectante	38	126	3	1524	<b>3 H1 0 hreflang 22/34 sans alt</b>
<b>Produit Solution Bleue</b> /solution-bleu-nettoyante	35	127	3	1407	<b>3 H1 22/34 sans alt</b>
<b>Article blog (1000 solutions)</b> /post/ils-ont-selectionné-1000-solutions-pour-changer-le-monde-et-cleanea-en-fait-partie	83	0	1	1515	<b>1 H1 title long meta absente 0 hreflang 5/9 sans alt</b>

Lecture : **vert** = conforme · **rouge** = à corriger. La densité de signaux rouges sur les fiches produit et les pages à fort contenu vient surtout des H1 multiples et des images sans alt, deux corrections rapides et systémiques.

## Ce que révèle le tableau

- **Titres et metas : du bon travail.** La grande majorité des balises sont bien rédigées, dans les bonnes longueurs, avec le mot-clé en tête. C'est le fruit d'une optimisation récente, à préserver.
- **H1 multiples : défaut systémique.** 8 pages sur 21 affichent 2 à 4 H1 (jusqu'à 4 sur la page Machine Europa V7). C'est un travers du template Wix où des éléments décoratifs sont balisés en H1. Un seul H1 par page, descriptif et porteur du mot-clé principal.
- **3 dépassements à corriger.** Le titre "À propos" (62 c) et le titre de l'article (83 c) seront tronqués ; la meta de la page produits (225 c) est presque le double de la cible.
- **2 pages sans meta description.** La page /blog et l'article unique n'ont pas de meta description : Google composera un extrait au hasard.
- **Coquilles visibles dans les titres Hn.** "interrogations" (Machine Europa V7), "électricité" (À propos) : à relire, ces fautes apparaissent dans des titres indexés.

## PHASE 1 (SUITE)

## ■ 1.4 · Données structurées (Schema.org)

Les données structurées disent explicitement à Google ce qu'est chaque page (entreprise, produit, FAQ, article). Elles débloquent les résultats enrichis. Sur Cleanea, c'est le pilier le plus en retard.

Page	Balisage présent	Balisage manquant (à ajouter)
Accueil	<b>WebSite</b> , <b>LocalBusiness</b> (WebSite dupliqué à nettoyer)	Organization consolidée
Fiches Solution Rouge / Bleue	<b>Product</b>	AggregateRating, Offer, FAQPage (la fiche a une FAQ produit non balisée)
Article de blog	<b>BlogPosting</b> , <b>Organization</b>	Cohérence (éviter NewsArticle + BlogPosting concurrents)
FAQ (page dédiée)	<b>Aucun</b>	<b>FAQPage</b> (gain de rich result immédiat)
Page produits, Technologie, Secteurs, 4 pages sectorielles, RSE, À-propos, Contact...	<b>Aucun (17 pages)</b>	Organization, BreadcrumbList, et FAQPage là où il y a des Q/R

### LE GISEMENT LE PLUS RENTABLE DU SITE

La page FAQ et plusieurs pages (produits, machine, secteurs) contiennent déjà des blocs questions/réponses rédigés, mais aucun n'est balisé en FAQPage. Ajouter ce balisage est rapide et peut générer des résultats enrichis (questions dépliées dans la SERP) qui augmentent la surface et le taux de clic, sans créer une ligne de contenu supplémentaire.

## ■ 1.5 · Accessibilité & images

Sur les 21 pages analysées : **388 images au total, dont 143 sans attribut alt (37 %)**. L'attribut alt sert à la fois l'accessibilité (lecteurs d'écran), le référencement (Google Images, compréhension du contexte) et la robustesse (texte de repli).

Page	Images	Sans alt	Priorité
Nos clients	50	<b>45</b>	Les logos clients ne sont pas décrits : double perte (preuve sociale non indexée + accessibilité).
Fiches Solution Rouge / Bleue	34	<b>22</b>	Visuels produit non décrits : opportunité manquée sur Google Images.
Page produits	23	<b>20</b>	Vignettes produit à décrire avec le nom + usage.
Accueil, À-propos, Pro, Hôtellerie...	20-27	2-4	Quelques visuels décoratifs à compléter.

On note aussi des **titres Hn vides** (caractères invisibles) sur "Nos clients" et "Filière réemploi", et des **H2 trop longs** sur "Filière réemploi" (des paragraphes entiers balisés en H2). À reprendre : un H2 est un intertitre court, pas un paragraphe.

## ■ 1.6 · Maillage interne

Le maillage est un point fort relatif : chaque page pointe vers 19 à 29 autres pages internes, ce qui fait largement circuler le "jus" de lien. Deux corrections néanmoins :

- **Doublon de pages de contact.** /contact et /nous-contacter coexistent, avec le même H1 "Contactez-nous". Cela dilue le signal et brouille le parcours. Choisir une page maîtresse, rediriger l'autre (301).
- **Ancre à enrichir.** De nombreux liens internes utilisent "EN SAVOIR PLUS" comme ancre. Préférer des ancres descriptives porteuses de mots-clés ("Découvrir la machine Europa V7", "Nettoyage écologique en hôtellerie").
- **Pages stratégiques à renforcer.** Les 4 pages sectorielles et la page technologie devraient recevoir davantage de liens internes contextuels depuis le futur contenu de blog.

## PHASE 1 · CONCURRENCE

# Analyse concurrentielle

Le secteur du nettoyage "sans chimie" est fragmenté entre plusieurs technologies (électrolyse, eau ozonée, vapeur, microfibre). Comprendre qui occupe la SERP révèle l'espace que Cleanea peut conquérir.

## ■ Qui ranke réellement sur les requêtes du secteur

Sur les requêtes catégorielles ("nettoyage écologique professionnel", "nettoyage sans produit chimique entreprise"), la première page de Google est dominée non pas par des fabricants, mais par du **contenu éditorial** : guides comparatifs, sources institutionnelles (mission transition écologique de l'État, ADEME), et données d'autorité (Anses, INRS). C'est l'enseignement central de cet audit.

### L'INSIGHT STRATÉGIQUE

La bataille du haut de funnel se gagne avec du contenu, pas avec des pages produit. Les concurrents et les médias captent les internautes au stade « je m'informe », puis les orientent. Cleanea, avec un seul article, est absente de cette conversation. C'est précisément le levier le plus sous-exploité et le plus accessible.

## ■ Concurrents directs (électrolyse sur site)

Acteur	Positionnement	Note concurrentielle
everyclean®	Eau électrolysée, solution détergente + désinfectante	Le plus proche du modèle Cleanea (deux fonctions, électrolyse). Concurrent frontal sur le discours "écologique + biodégradable".
Olimpe Pro (Olimpe Technology)	Électrolyse en ligne, gamme fixe/mobile	Vient du traitement de l'eau (élevage) et monte vers l'industrie. Fort sur la désinfection de l'eau et des canalisations.
Effiterr Eau (Purealyse)	Station produisant de l'acide hypochloreux (HOCl)	Très orienté "traitement de l'eau / biocide". Vocabulaire technique proche.
EcoGénèse	Désinfectant électrolytique HOCl, flacons	Cible plus large (grand public + pro), depuis 2009 comme Cleanea.
ProMinent (CHLORINSITU)	Installations industrielles de dosage / électrolyse	Acteur industriel international, plutôt sur l'eau de procédé et l'industrie lourde.

## ■ Concurrents indirects (autres technologies "sans chimie")

Technologie	Acteurs repérés	Ce que ça implique pour Cleanea
Eau ozonée	Aquaflow Sanzonate (uni-vert-partners)	Argumentaire quasi identique (sans rinçage, sans stockage, RSE). Cleanea doit clarifier ce qui distingue l'électrolyse de l'ozone (rémanence, double fonction détergent + désinfectant).
Vapeur	Polti, centrales vapeur	Forte notoriété grand public, bon contenu SEO. Concurrent sur "ménage écologique".
Éditeurs de contenu	ruedelhygiene.fr, france-clean.fr, solservice.fr	Ne vendent pas la même chose mais captent l'audience via des guides très complets citant les normes EN. Modèle à répliquer.

## ■ Atouts différenciants de Cleanea à marteler

- **Pionnier reconnu (depuis 2009)** et cité comme tel par la presse spécialisée (La Gazette des Communes). Une légitimité que les nouveaux entrants n'ont pas.
- **Double production sur site** (un détergent + un désinfectant via cellule 3 chambres brevetée), là où beaucoup ne produisent qu'un biocide désinfectant.
- **Preuves chiffrées RSE** (réduction des déchets plastiques et des émissions, label Solar Impulse, ACV) : un capital de contenu énorme, aujourd'hui peu exploité côté SEO.
- **Normes & conformité (EN 14476)** : un terrain où le contenu fait autorité et rassure les acheteurs B2B.

### TERRITOIRE À CONQUÉRIR

Cleanea doit posséder, dans l'ordre : (1) son nom et ses produits, (2) « détergent ET désinfectant écologique produit sur site par électrolyse », (3) les requêtes sectorielles (hôtellerie, santé, tertiaire, collectivités), (4) les requêtes problèmes/réglementation (sans chimie, EN 14476, appels d'offres RSE, sécurité des agents).

## PHASE 2

# Stratégie de mots-clés

Organiser la cible en 4 clusters d'intention, du plus chaud (la marque) au plus froid (les problèmes). Pour un site de cette taille, la longue traîne et les requêtes sectorielles priment sur les mots-clés génériques très concurrentiels.

**Note sur les volumes.** En l'absence d'outil de volumétrie (Ahrefs/Semrush) et de Search Console, les niveaux de volume et de difficulté ci-dessous sont des estimations qualitatives fondées sur l'analyse des SERP. À confirmer et affiner dès que Search Console aura accumulé 48 à 72 h de données réelles.

## ■ Cluster 1 · Notoriété & marque

Requêtes sur le nom de l'entreprise et ses produits. Objectif : être en position 1 sur son propre univers (souvent déjà acquis, à verrouiller).

cleanea · cleanea avis · machine cleanea · cleanea europa · solution rouge cleanea · solution bleue cleanea · cleanea électrolyse · cleanea puteaux

## ■ Cluster 2 · Technologie & solution

Requêtes sur le fonctionnement. Objectif : éduquer les curieux et capter ceux qui découvrent l'électrolyse. Pages cibles : Technologie, Machine Europa V7, fiches produit.

nettoyage par électrolyse · eau électrolysée nettoyage · détergent désinfectant sur site · désinfectant acide hypochloreux · produit nettoyage eau sel électricité · machine électrolyse nettoyage · acide hypochloreux HOCl désinfectant · hydroxyde de sodium nettoyant

## ■ Cluster 3 · Secteurs & cas d'usage

Requêtes par type de client. Objectif : capter l'intention d'achat B2B. Pages cibles : les 4 pages sectorielles, déjà en place et prêtes à performer.

nettoyage écologique hôtellerie · désinfection écologique hôpital · nettoyage écologique EHPAD · nettoyage bureaux écologique · nettoyage écologique collectivités · nettoyage écologique école · produit nettoyage site HQE · désinfection clinique sans produit chimique

## ■ Cluster 4 · Problèmes & réglementation

Requêtes sur les problèmes que Cleanea résout. Objectif : attirer en haut de funnel et asseoir l'autorité. Pages cibles : **articles de blog à créer**. C'est le cluster le plus vide aujourd'hui.

nettoyage sans produit chimique entreprise · alternative eau de javel professionnelle · désinfectant norme EN 14476 · sécurité agents nettoyage produits chimiques · réduire déchets plastiques nettoyage · critères RSE appel d'offres propreté · danger produits d'entretien professionnels · COV produits nettoyage

## ■ Top 12 mots-clés prioritaires

Priorisation selon le principe : **volume moyen + difficulté faible + pertinence haute = priorité maximale**. Ce sont les 12 requêtes à attaquer en premier.

#	Mot-clé	Volume	Difficulté	Intention	Page cible	Prio.
1	nettoyage sans produit chimique entreprise	Moyen	Moyenne	Info	Blog (article pilier)	Haute
2	alternative eau de javel professionnelle	Moyen	Faible	Info	Blog	Haute
3	désinfectant EN 14476	Faible	Faible	Info/Achat	Blog + Technologie	Haute
4	nettoyage écologique hôtellerie	Faible	Faible	Achat	/proprete-hotellerie	Haute
5	désinfection écologique EHPAD / hôpital	Faible	Faible	Achat	/proprete-sante	Haute
6	nettoyage par électrolyse	Moyen	Moyenne	Info	/technologie	Haute
7	eau électrolysée nettoyage professionnel	Faible	Faible	Info	/technologie	Moyenne
8	produit nettoyage eau sel électricité	Faible	Faible	Info	/professionnels	Moyenne
9	nettoyage bureaux écologique	Moyen	Moyenne	Achat	/proprete-tertiaire	Moyenne
10	critères RSE appel d'offres propreté	Faible	Faible	Info	Blog	Haute
11	danger produits d'entretien agents	Moyen	Faible	Info	Blog	Haute
12	désinfectant acide hypochloreux	Moyen	Moyenne	Info	Fiche Solution Rouge	Moyenne

## ■ Mapping page / mot-clé principal

Chaque page doit cibler UN mot-clé principal et 2 à 3 secondaires. État cible :

Page	Mot-clé principal	Secondaires
Accueil	cleanea / produits de nettoyage écologiques	nettoyage écologique professionnel, désinfectant écologique
/technologie	nettoyage par électrolyse	eau électrolysée, acide hypochloreux, membrane protonique
/machine-europa-v7	machine électrolyse nettoyage	produire détergent sur site, Europa V7
/professionnels	solution nettoyage écologique professionnels	produit eau sel électricité, sans chimie lourde
/proprete-hotellerie	nettoyage écologique hôtellerie	hygiène hôtel sans odeur, désinfection chambres
/proprete-sante	désinfection écologique santé / EHPAD	nettoyage hôpital sans produit chimique, sécurité patients
/proprete-tertiaire	nettoyage bureaux écologique	nettoyage immeuble HQE, propreté tertiaire RSE
/proprete-collectivites	nettoyage écologique collectivités	nettoyage école / mairie, appel d'offres RSE
/rse	propreté écoresponsable / RSE nettoyage	réduction déchets plastiques, empreinte carbone nettoyage
Fiche Solution Rouge	désinfectant naturel acide hypochloreux	virucide sans alcool, désinfectant sans rinçage
Fiche Solution Bleue	nettoyant écologique professionnel	dégraissant naturel, bidon rechargeable
Blog (12 articles)	cluster 4 (problèmes / réglementation)	longue traîne déclinée par article

## PHASE 3

# Plan d'optimisation

Du diagnostic à l'exécution. Cette phase liste les corrections concrètes, dans l'ordre, avec le "avant / après" prêt à appliquer.

## ■ 3.1 · Corrections techniques prioritaires

#	Correction	Gravité	Où
1	Réduire chaque page à un <b>seul H1</b> descriptif (reclasser les autres en H2/H3)	ÉLEVÉ	8 pages (Accueil, Techno, Hôtellerie, Europa V7, RSE, À-propos, 2 fiches)
2	Corriger l'en-tête content - Language en fr-FR	ÉLEVÉ	Réglages Wix (langue par défaut)
3	Homogénéiser le <b>hreflang</b> (l'ajouter sur /blog, l'article, la fiche Solution Rouge)	MOYEN	3 pages
4	Trancher la <b>version /en</b> : entretenir et baliser, ou désindexer	MOYEN	Tout le sous-dossier /en
5	Résoudre le <b>doublon contact</b> (301 de l'une vers l'autre)	MOYEN	/contact + /nous-contacter
6	Renseigner les <b>attributs alt</b> (priorité : logos clients, visuels produit)	MOYEN	143 images
7	Slugs en <b>ASCII sans accent</b> pour les nouvelles pages ; garder les anciennes en 301 si modifiées	FAIBLE	URLs accentuées
8	Corriger les <b>coquilles</b> dans les titres ("interrogations", "électricité")	QUICK WIN	Europa V7, À-propos
9	Nettoyer les <b>H2 vides / paragraphes en H2</b>	QUICK WIN	Nos clients, Filière réemploi

## ■ 3.2 · Réécritures title & meta (avant / après)

Page	Avant	Après (proposé)
<b>À-propos</b> title 62 c	Notre histoire   Cleanea, secteur du nettoyage par électrolyse	<b>Notre histoire : le nettoyage par électrolyse depuis 2009   Cleanea (59 c)</b>
<b>Page produits</b> meta 225 c	Découvrez les produits nettoyants, détergents et désinfectants écologiques de Cleanea. Nos produits de nettoyage professionnels sont efficaces pour les sols, la salle de bain et toutes les surfaces de vos locaux d'entreprise.	<b>Détergents et désinfectants écologiques pour professionnels : efficaces sur toutes vos surfaces, sans chimie lourde. Découvrez la gamme Cleanea. (150 c)</b>
<b>/blog</b> meta absente	(aucune meta description)	<b>Le blog Cleanea : nettoyage écologique, réglementation, RSE et innovations qui réinventent la propreté professionnelle. (120 c)</b>
<b>Article (Solar Impulse)</b> title 83 c, meta absente	Ils ont sélectionné 1000 solutions pour changer le monde et Cleanea en fait partie.	<b>Title</b> : Cleanea labellisée Solar Impulse : 1000 solutions pour la planète <b>(58 c)</b> <b>Meta</b> : Cleanea rejoint les 1000 solutions Solar Impulse. Ce que signifie ce label et pourquoi il rassure les acheteurs engagés.
<b>Collectivités</b> title finit par un point	Nettoyage écologique pour les collectivités   Cleanea.	<b>Nettoyage écologique pour les collectivités et écoles   Cleanea</b>

Les autres titres et metas sont conformes : ne pas y toucher (principe : on ne réécrit que ce qui pose problème).

## ■ 3.3 · Plan éditorial : 12 articles pour exister sur le haut de funnel

C'est le chantier à plus fort impact. Objectif : un article toutes les 2 semaines pendant 6 mois, chacun ciblant une requête du cluster 4, avec maillage vers les pages secteurs/technologie.

#	Titre d'article (angle)	Mot-clé visé	Cluster
1	Nettoyage sans produit chimique en entreprise : le guide complet 2026	nettoyage sans produit chimique entreprise	Pilier
2	Eau de javel en milieu professionnel : 4 alternatives plus sûres et aussi efficaces	alternative eau de javel professionnelle	Problème
3	Norme EN 14476 : ce que "virucide" veut vraiment dire pour vos achats	désinfectant EN 14476	Réglementation
4	Agents d'entretien et produits chimiques : 91 % exposés, comment réduire le risque	danger produits d'entretien agents	Problème
5	Critères RSE dans les appels d'offres de propreté : comment marquer des points	critères RSE appel d'offres propreté	Réglementation
6	Électrolyse, ozone, vapeur : quelle techno "sans chimie" pour quel usage ?	nettoyage écologique comparatif	Comparatif
7	Acide hypochloreux (HOCl) : le désinfectant que votre corps fabrique déjà	désinfectant acide hypochloreux	Techno
8	Réduire les déchets plastiques de nettoyage : la méthode du "produit sur site"	réduire déchets plastiques nettoyage	Problème
9	Hôtellerie : une hygiène irréprochable sans odeur chimique, c'est possible	nettoyage écologique hôtellerie	Secteur
10	EHPAD et cliniques : désinfecter sans exposer patients et soignants	désinfection écologique EHPAD	Secteur
11	COV et qualité de l'air intérieur : l'angle mort du nettoyage de bureaux	COV produits nettoyage bureaux	Problème
12	Coût réel du nettoyage : pourquoi produire sur site revient moins cher	coût produits nettoyage professionnel	ROI

### RÈGLE DE PRODUCTION

Chaque article : 800 à 1 500 mots, un H1 unique avec le mot-clé, des H2 en questions, au moins 2 liens internes vers les pages secteurs ou technologie, une image avec alt, une meta description rédigée, et un balisage Article. Réutiliser les données chiffrées RSE déjà disponibles (ACV, label Solar Impulse) comme matière première.

## 3.4 · Données structurées à déployer

Exemple de balisage FAQPage à injecter sur la page FAQ et sur les pages contenant des Q/R (produits, machine). Adapter les questions au contenu réel :

```
// À placer dans un bloc HTML/Embed Wix sur la page FAQ
<script type="application/ld+json">
{
  "@context": "https://schema.org",
  "@type": "FAQPage",
  "mainEntity": [{
    "@type": "Question",
    "name": "Le produit est-il prêt à l'emploi ?",
    "acceptedAnswer": { "@type": "Answer",
      "text": "Oui. Les solutions Cleanea sortent de la machine prêtes à l'emploi, sans dilution ni rinçage." }
  },{
    "@type": "Question",
    "name": "Pourquoi ça sent le chlore comme l'eau de javel ?",
    "acceptedAnswer": { "@type": "Answer",
      "text": "L'odeur vient de l'acide hypochloreux, le principe actif, présent à très faible concentration et sans danger." }
  }]
}
</script>
```

À déployer également :

- **Organization** (sur tout le site, via le footer) : nom, logo, URL, profils sociaux, coordonnées de Puteaux. Consolide l'entité "Cleanea" aux yeux de Google.
- **BreadcrumbList** : fil d'Ariane balisé sur les pages profondes (secteurs, produits) pour des SERP plus riches.
- **Product** complété sur les fiches : ajouter **Offer** (prix, disponibilité) et, à terme, **AggregateRating** à partir des avis clients.
- **Nettoyer le doublon WebSite** sur la page d'accueil (déclaré deux fois).

## PHASE 4 · PILOTAGE

# Suivi & KPIs

Un audit ne vaut que par ce qu'on en fait, et par la capacité à mesurer l'effet des actions. Cette phase installe le tableau de bord, la cadence de revue et les signaux d'alerte qui transforment le plan en boucle d'amélioration continue. Tout repose sur un prérequis non négociable : disposer d'une mesure fiable.

## PRÉREQUIS N°1 : LA MESURE

Avant toute optimisation, connecter **Google Search Console** (propriété de domaine, vérification DNS) et **Google Analytics 4** (avec un événement de conversion sur les demandes de devis et le formulaire de contact). Sans ces deux sources, aucune des cibles ci-dessous n'est mesurable. La baseline doit être figée au mois 0 : c'est le point de départ contre lequel toute progression se compare.

## Tableau de bord des indicateurs

Onze indicateurs, regroupés en trois familles : visibilité (ce que Google montre), engagement (ce que les visiteurs font), et production (ce que l'on publie). Les cibles sont volontairement progressives et tiennent compte de l'absence de baseline historique. Elles seront recalées une fois la baseline mesurée.

Indicateur	Ce qu'il mesure	Source	Fréq.	3 mois	6 mois	12 mois
Impressions organiques	Surface de visibilité dans la SERP	GSC	Hebdo	+40 %	x2	x3 à x4
Clics organiques	Trafic réel capté depuis Google	GSC	Hebdo	+30 %	+80 %	x2,5
CTR moyen	Attractivité des titres et metas	GSC	Mensuel	Stable	+0,5 pt	+1 pt
Position moyenne	Classement agrégé des requêtes	GSC	Mensuel	-2 places	-4 places	-6 places
Mots-clés en top 10	Ampleur de la présence en 1re page	GSC	Mensuel	+15	+40	+80
Mots-clés en top 3	Positions à fort trafic	GSC	Mensuel	+3	+10	+25
Pages indexées	Couverture réelle du site	GSC	Mensuel	45	55	70
Sessions organiques	Volume d'audience SEO	GA4	Mensuel	+35 %	+90 %	x2,5
Demandes (devis/contact)	Conversions d'origine organique	GA4	Mensuel	+2/mois	+5/mois	+10/mois
Articles publiés (cumul)	Indicateur d'effort éditorial	Interne	Mensuel	4	9	16
Domaines référents	Autorité et notoriété (off-site)	Outil SEO	Trimestriel	+5	+15	+30

GSC = Google Search Console · GA4 = Google Analytics 4. Les cibles exprimées en pourcentage s'appliquent à la baseline figée au mois 0. Les valeurs absolues (pages indexées, articles) partent de l'état constaté dans cet audit (39 URLs, 1 article).

## Cadence de suivi

Cinq rituels, du plus court au plus profond. L'objectif est de détecter vite (hebdo), de piloter régulièrement (mensuel) et de réauditer en profondeur (trimestriel).

Rituel	Fréquence	Ce que l'on regarde	Outils
Revue express	Hebdomadaire	Nouvelles impressions, requêtes émergentes, erreurs d'indexation, chutes ou pics anormaux.	GSC
Reporting de performance	Mensuel	Clics, positions, top pages et requêtes, conversions, comparaison aux cibles.	GSC + GA4 + Looker Studio
Revue éditoriale	Mensuel	Performance des articles publiés, sujets à fort potentiel, ajustement du calendrier.	GSC + skill SEO
Audit technique	Trimestriel	Core Web Vitals, couverture, données structurées, liens cassés, nouveau crawl complet.	Crawl + PageSpeed + GSC
Veille concurrentielle	Trimestriel	Nouveaux contenus des concurrents, fonctionnalités SERP, déplacements de positions.	Recherche manuelle + skill

### SIGNAUX D'ALERTE À SURVEILLER

Déclencher une investigation immédiate si l'un de ces signaux apparaît : chute des impressions ou des clics de plus de 20 % sur une semaine, hausse soudaine des pages exclues dans la couverture GSC, désindexation d'une page stratégique, recul des positions sur les requêtes de marque (cleanea), pic d'erreurs 4xx ou 5xx, dégradation mesurée des Core Web Vitals après une mise à jour Wix. Ces alertes priment sur le calendrier de revue habituel.

**Reporting léger et reproductible.** Un tableau de bord Looker Studio branché sur GSC et GA4 suffit à automatiser 90 % du reporting mensuel. C'est précisément le type de livrable récurrent qu'Hermès peut générer et commenter chaque mois, en s'appuyant sur le barème et la structure définis en Annexe A.

SYNTHÈSE OPÉRATIONNELLE

# Feuille de route priorisée

Toutes les recommandations du rapport, ordonnées par le rapport entre impact et effort, puis séquencées en quatre sprints. La logique : sécuriser la mesure et récolter les gains rapides d'abord, bâtir l'actif de contenu ensuite, installer l'autorité enfin.

## ■ Matrice impact / effort

	Effort faible	Effort moyen	Effort élevé
Impact élevé	FAQPage + Organization H1 uniques · alt images content-langage FR	BreadcrumbList + Product Optimiser templates lourds Pages secteurs enrichies	Programme éditorial (12 articles)
Impact moyen	Connecter GSC + GA4 Résoudre doublon contact Corriger coquilles titres	Maîtriser ou désindexer /en Ancres de liens explicites hreflang cohérent	Refonte performance profonde du thème
Impact faible	Nettoyer doublon WebSite Harmoniser casse des URLs	Micro-ajustements cosmétiques	Chantiers à reporter

Vert : à faire en priorité (fort levier, faible coût). Ambre : à planifier. Rouge : à reporter ou à ne déclencher qu'après les deux premières vagues.

## ■ Séquencement en 4 sprints

Sprint	Période	Actions clés	Livrable
<b>Sprint 0</b> Fondations de mesure <b>QUICK WIN</b>	0 à 15 j	Connecter GSC et GA4, figer la baseline, résoudre le doublon de contact, corriger les coquilles de titres (interrogations, électricité), basculer l'en-tête content-langage en français.	Mesure fiable + quick wins techniques
<b>Sprint 1</b> Site propre & balisé	0 à 30 j	Forcer un H1 unique par page, ajouter les attributs alt manquants (priorité aux pages clients et produits), harmoniser le hreflang, déployer FAQPage, Organization et BreadcrumbList.	Hygiène technique + premiers rich results
<b>Sprint 2</b> Présence haut de funnel	30 à 90 j	Publier les 4 à 6 premiers articles prioritaires (sans chimie, EN 14476, appels d'offres RSE, comparatif électrolyse), enrichir les pages secteurs, mailler les nouveaux contenus vers les pages de conversion.	Cluster de contenu initial actif
<b>Sprint 3</b> Autorité thématique	3 à 6 mois	Compléter le calendrier (6 à 10 articles), collecter et baliser les avis clients (AggregateRating), engager des actions de relations presse et de netlinking ciblé, suivre et itérer via les revues mensuelles.	Autorité reconnue + signaux off-site

**PRINCIPE DIRECTEUR**

Ne pas attendre que tout soit parfait pour publier. Les Sprints 0 et 1 (techniques, à fort levier et faible coût) se mènent en parallèle des premiers articles du Sprint 2. La vitesse de mise en ligne du contenu est, ici, le facteur le plus déterminant pour la trajectoire à 12 mois.

## ANNEXE A

# Protocole reproductible pour Hermes

Cette annexe est le coeur opérationnel du document. Elle décrit, étape par étape, comment régénérer un audit SEO de cette qualité sur n'importe quelle URL. Un agent autonome qui suit ce protocole produit le même type de livrable, avec la même rigueur, sans intervention humaine au-delà de la validation finale.

## OBJET DE CETTE ANNEXE

Transformer une démarche d'expert en une procédure exécutable. Chaque étape fournit la commande ou le script réellement utilisé, les signaux à extraire, et la règle de notation associée. Hermes prend un domaine en entrée et rend un audit structuré en sortie. Les blocs de code ci-dessous sont directement réutilisables (Python 3 + requests + BeautifulSoup4).

## ■ A.1 · Vue d'ensemble du pipeline

1. **Découverte** : robots.txt et parcours récursif des sitemaps pour obtenir la liste complète des URLs.
2. **Crawl on-page** : pour chaque URL, extraction des signaux (title, meta, H1, H2, images, alt, mots, liens internes).
3. **Signaux techniques** : en-têtes HTTP, langue déclarée, HSTS, poids du HTML, temps de réponse serveur.
4. **Données structurées** : détection des types JSON-LD présents sur chaque page.
5. **Concurrence** : requêtes catégorie, identification des dominants et des fonctionnalités SERP.
6. **Notation** : application du barème des 7 piliers et calcul du score pondéré global.
7. **Stratégie** : clusters de mots-clés, priorités, calendrier éditorial.
8. **Livrable** : assemblage du PDF selon la structure normalisée (section A.8).

## ■ A.2 · Étape 1 et 2, découverte et crawl

La découverte part toujours du sitemap déclaré dans robots.txt. Un sitemap d'index pointe vers des sous-sitemaps : le parcours doit être récursif pour ne rien manquer (pages, produits, catégories, blog).

```
# Découverte : robots.txt + parcours récursif des sitemaps
import requests, xml.etree.ElementTree as ET
NS = {'s': 'http://www.sitemaps.org/schemas/sitemap/0.9'}

def urls_from(sitemap_url):
    r = requests.get(sitemap_url, timeout=20)
    root = ET.fromstring(r.content)
    subs = [e.text for e in root.findall('.//s:sitemap/s:loc', NS)]
    if subs:
        # c'est un index : on descend
        out = []
        for u in subs:
            out += urls_from(u)
        return out
    return [e.text for e in root.findall('.//s:url/s:loc', NS)]

all_urls = urls_from('https://www.exemple.com/sitemap.xml')
print(len(all_urls), 'URLs trouvées')
```

Le crawl on-page extrait, pour chaque URL, les signaux qui alimentent la notation. La sortie est un dictionnaire par page, sérialisé en JSON pour être réexploité dans le rapport et le tableau page par page.

```
# Crawl on-page : un dictionnaire de signaux par URL
import requests, json
from bs4 import BeautifulSoup
HEAD = {'User-Agent': 'Mozilla/5.0 (audit SEO)'}

def audit_page(url):
    r = requests.get(url, headers=HEAD, timeout=30)
    soup = BeautifulSoup(r.text, 'html.parser')
    title = (soup.title.get_text().strip() if soup.title else '')
    md = soup.find('meta', attrs={'name': 'description'})
    meta = (md.get('content', '').strip() if md else '')
    imgs = soup.find_all('img')
    no_alt = [i for i in imgs if not (i.get('alt') or '').strip()]
    text = soup.get_text(' ', strip=True)
    links = soup.find_all('a', href=True)
    internes = [a for a in links if a['href'].startswith('/') or 'exemple.com' in a['href']]
    return {
        'url': url,
        'title': title, 'title_len': len(title),
        'meta_desc': meta, 'meta_desc_len': len(meta),
        'h1_count': len(soup.find_all('h1')),
        'h2_count': len(soup.find_all('h2')),
        'img_total': len(imgs), 'img_no_alt': len(no_alt),
        'word_count': len(text.split()),
        'internal_links': len(internes),
    }

results = {u: audit_page(u) for u in all_urls}
json.dump(results, open('crawl_results.json', 'w'), ensure_ascii=False, indent=2)
```

## ■ A.3 · Étape 3, signaux techniques

Trois vérifications en ligne de commande suffisent à qualifier la santé technique de surface. Elles révèlent les anomalies les plus courantes (langue mal déclarée, absence de HSTS, pages trop lourdes, serveur lent).

```
# En-têtes HTTP : langue déclarée, sécurité, cache
curl -sI https://www.exemple.com/ | grep -iE 'content-language|strict-transport|cache-control'

# Poids du HTML brut, en octets (alerte au-dessus de 1,5 Mo)
curl -s https://www.exemple.com/ | wc -c

# Temps de réponse serveur (TTFB)
curl -s -o /dev/null -w 'TTFB: %{time_starttransfer}s\n' https://www.exemple.com/
```

Pour les Core Web Vitals, interroger l'API PageSpeed Insights. En l'absence de données terrain (CrUX), se rabattre sur les métriques de laboratoire (Lighthouse), en notant clairement qu'il s'agit d'une estimation et non d'un vécu utilisateur.

## ■ A.4 · Étape 4, données structurées

La détection des données structurées se fait en lisant les blocs JSON-LD de chaque page. On relève les types présents (WebSite, Organization, Product, FAQPage, BreadcrumbList) pour mesurer la couverture et repérer les gisements de rich results.

```
# Détection JSON-LD + hreflang par page
import json

def structured(soup):
    types = []
    for b in soup.find_all('script', attrs={'type': 'application/ld+json'}):
        try:
            data = json.loads(b.string or '{}')
        except Exception:
            continue
        items = data if isinstance(data, list) else [data]
        for it in items:
            t = it.get('@type')
            if t:
                types.append(t)
    hreflang = [l.get('hreflang')
                for l in soup.find_all('link', attrs={'rel': 'alternate'})
                if l.get('hreflang')]
    return {'schema_types': types, 'hreflang': hreflang}
```

## ■ A.5 · Étape 5, analyse concurrentielle

L'objectif est de comprendre qui occupe la SERP sur les intentions du secteur, et avec quel type de contenu. On exécute une poignée de requêtes catégorie et l'on classe les résultats en trois familles.

Requête type (à adapter au secteur)	Ce que l'on relève
nettoyage écologique professionnel sans produit chimique	Médias et guides comparatifs vs sites de marque
désinfectant sans danger norme EN 14476	Présence de contenu réglementaire et technique
électrolyse nettoyage entreprise	Concurrents directs sur la technologie
produit entretien appel d'offres RSE	Intentions d'achat B2B et marchés publics
alternative eau de javel collectivité	Requêtes problème et substitution

Pour chaque requête, classer les 10 premiers résultats en : **éditorial** (guides, médias), **institutionnel** (ADEME, Anses, INRS, ministères), **concurrents directs**, **concurrents indirects**. Noter aussi les fonctionnalités SERP (extrait optimisé, People Also Ask, pack local). La conclusion stratégique se déduit de la famille dominante.

## ■ A.6 · Barème de notation des 7 piliers

Le score global est une moyenne pondérée de 7 piliers, chacun noté de 0 à 100. Les poids reflètent l'effet d'un pilier sur la performance organique d'un site B2B. La notation de chaque pilier suit des sous-critères explicites pour rester reproductible d'un audit à l'autre.

Pilier	Poids	Méthode de notation (sur 100)
Indexation & exploration	12 %	robots.txt valide (20) + sitemaps valides (20) + ratio pages indexées / connues (30) + absence d'erreurs de couverture (30).
Performance / CWV	15 %	LCP (40) + INP (30) + CLS (20) + poids de page (10), notés sur les seuils PageSpeed.
Architecture & maillage	13 %	Profondeur de 3 clics maximum (30) + densité de liens internes (20) + cohérence des URLs (25) + absence de doublons (25).
On-page	20 %	Titles conformes (25) + metas conformes (20) + H1 unique par page (25) + densité de contenu (15) + hiérarchie Hn propre (15).
Contenu & sémantique	18 %	Couverture des clusters d'intentions (40) + fraîcheur et cadence de publication (30) + profondeur des pages (30).
Données structurées	12 %	Présence de schema par page (50) + pertinence des types (30) + validité technique (20).
Accessibilité & UX	10 %	Ratio d'images avec alt (60) + labels, contrastes et titres non vides (40).

Le score global s'obtient par la somme des scores de pilier multipliés par leur poids. Exemple, sur cet audit Cleanea :

Pilier	Score	Poids	Contribution
Indexation & exploration	74	12 %	8,9
Performance / CWV	55	15 %	8,3
Architecture & maillage	70	13 %	9,1
On-page	68	20 %	13,6
Contenu & sémantique	48	18 %	8,6
Données structurées	35	12 %	4,2
Accessibilité & UX	52	10 %	5,2
		100 %	

Échelle de lecture : 0 à 44 critique, 45 à 59 à consolider, 60 à 74 correct, 75 à 100 solide.

## ■ A.7 · Structure normalisée du livrable

Tout audit régénéré par Hermes doit suivre cette ossature, dans cet ordre. C'est ce qui garantit la cohérence d'un rapport à l'autre.

1. Couverture (domaine, date, score global).

2. Sommaire exécutif (5 constats majeurs, forces/faiblesses/opportunités, priorités du trimestre).
3. Méthodologie et périmètre.
4. Score de santé SEO (dashboard des 7 piliers).
5. Phase 1, diagnostic (architecture, technique, on-page page par page, données structurées, accessibilité, maillage).
6. Analyse concurrentielle (directs, indirects, différenciants).
7. Phase 2, stratégie de mots-clés (clusters, priorisation, mapping page / mot-clé).
8. Phase 3, plan d'optimisation (corrections techniques, réécritures, calendrier éditorial, données structurées).
9. Phase 4, suivi et KPIs (tableau de bord, cadence, signaux d'alerte).
10. Feuille de route priorisée (matrice impact/effort, sprints).
11. Annexes (ce protocole, checklist).

## ■ A.8 · Règles de rédaction et prompt de référence

Le prompt ci-dessous cadre le comportement attendu d'Hermes. Il fixe la procédure, le format de sortie et les règles de style non négociables.

```
RÔLE : consultant SEO senior. Tu produis un audit complet et reproductible.  
ENTRÉE : un domaine (ex. www.exemple.com).
```

```
PROCÉDURE :
```

1. Découverte : robots.txt + parcours récursif des sitemaps -> liste d'URLs.
2. Crawl on-page : par URL, extraire title, meta, nb H1, nb H2, nb images, images sans alt, nb de mots, liens internes.
3. Signaux techniques : en-têtes HTTP (content-language, HSTS, cache), poids du HTML, TTFB, compatibilité mobile.
4. Données structurées : détecter les types JSON-LD par page.
5. Concurrence : 5 à 8 requêtes catégorie ; classer les résultats en éditorial / institutionnel / concurrents directs / indirects ; relever les fonctionnalités SERP.
6. Notation : appliquer le barème des 7 piliers, calculer le score pondéré.
7. Plan : priorités P1 à P4, calendrier éditorial, corrections techniques, données structurées, KPIs et cadence de suivi.

```
SORTIE : un PDF structuré selon "Structure normalisée du livrable" (A.7).
```

```
RÈGLES DE STYLE :
```

- Langue : français.
- Interdiction absolue du tiret cadratin. Utiliser virgule, parenthèse ou point médian.
- Ton expert, factuel, actionnable : chaque constat est suivi d'une action.
- Toute valeur chiffrée provient d'une donnée réellement collectée, jamais inventée. En cas d'absence de donnée (ex. CrUX), le dire explicitement.

**Garde-fou qualité.** Avant de livrer, Hermes vérifie trois choses : (1) aucun tiret cadratin dans le document, (2) chaque score de pilier est justifié par au moins une donnée du crawl, (3) chaque recommandation est rattachée à un constat. Si l'une de ces conditions n'est pas remplie, le rapport est repris avant envoi.

## ANNEXE B

# Checklist SEO

Récapitulatif actionnable de tous les points de contrôle, avec l'état constaté sur cleanea.com au moment de l'audit. À utiliser comme liste de suivi : chaque ligne "À faire" ou "Partiel" correspond à une action du plan.

## ■ Technique & indexation

Point de contrôle	État
robots.txt présent et non bloquant	Fait
Sitemap XML soumis et à jour	Fait
HTTPS actif avec HSTS	Fait
En-tête content-language cohérent (français)	À faire
hreflang cohérent sur toutes les pages	Partiel
URLs propres, sans accent encodé, casse homogène	Partiel
Aucune page en doublon fonctionnel	À faire
Poids des pages maîtrisé (sous 1,5 Mo)	À faire
Core Web Vitals dans le vert (à mesurer)	À faire
Google Search Console connectée	À faire

## ■ On-page

Point de contrôle	État
Title unique et sous 60 caractères	Partiel
Meta description présente et sous 165 caractères	Partiel
Une seule balise H1 par page	À faire
Hiérarchie Hn logique, sans titre vide	Partiel
Aucune coquille dans les titres	À faire
Balise canonique correcte par page	Partiel
Contenu dense et utile	Fait

## ■ Contenu & sémantique

Point de contrôle	État
Blog actif avec une cadence régulière	À faire
Couverture des 4 clusters d'intentions	À faire
Pages secteurs optimisées par mot-clé	Partiel
Maillage interne vers les pages de conversion	Partiel
Ancres de liens explicites (pas de "en savoir plus")	À faire

## ■ Données structurées

Point de contrôle	État
Organization déployé sur tout le site	À faire
LocalBusiness sur la page d'accueil	Fait
Product sur toutes les fiches produit	Partiel
FAQPage sur la FAQ et les blocs Q/R	À faire
BreadcrumbList sur les pages profondes	À faire
Aucun doublon de balisage (WebSite déclaré une fois)	À faire

## ■ Accessibilité & suivi

Point de contrôle	État
Toutes les images ont un attribut alt pertinent	À faire
Titres de page descriptifs	Fait
Contrastes et navigation cohérents	Partiel
GA4 configuré avec événements de conversion	À faire
Tableau de bord Looker Studio en place	À faire
Revue mensuelle et veille trimestrielle planifiées	À faire

### FIN DU PROTOCOLE

Ce document, de la couverture à cette checklist, constitue à la fois le livrable client et le gabarit de référence. Hermes peut le rejouer sur n'importe quel domaine en suivant l'Annexe A, puis produire le même PDF structuré, dans le même style, avec les mêmes garde-fous de qualité.