

# Annexe technique au cas d'étude Valberg.

## *Méthodologie, architecture, standards.*

Ce document accompagne la note commerciale (Doc A) transmise à l'Office de Tourisme de Valberg en mai 2026. Il documente la méthodologie de mesure des citations IA, l'architecture du pipeline d'agrégation, les standards exposés et les limites méthodologiques. Pour la proposition, les engagements et les conditions, se référer au Doc A.

# Trois questions. *Trois LLMs. Neuf réponses.*

## Outils mobilisés (DataForSEO API)

- `ai_optimization_llm_response` — 9 requêtes (3 prompts × ChatGPT gpt-5-mini, Claude Sonnet 4.5, Perplexity sonar-pro), recherche web active sur les trois
- `ai_optimization_chat_gpt_scraper` — 3 requêtes en France (fr-FR, force\_web\_search=true) pour reproduire l'expérience d'un utilisateur réel via l'interface ChatGPT
- `ai_optimization_keyword_data_search_volume` — volumétrie sur 10 mots-clés Valberg
- `serp_organic_live_advanced` — Google FR (depth=10) sur les 3 prompts pour disposer du sol de vérité

Coût total des mesures : 0,30 \$. Toutes les requêtes sont reproductibles à l'identique.

## Prompts mesurés

- **Q1 · Découverte** — « Que faire à Valberg ce week-end ? »
- **Q2 · Quantitatif** — « Combien d'hôtels à Valberg ? »
- **Q3 · Localisation** — « Office de tourisme de Valberg, c'est où ? »

## Limites de la mesure

- **Date unique** — snapshot du 4 mai 2026. La cadence en exploitation est mensuelle (cf. § T6).
- **3 LLMs sur 4** — Gemini volontairement exclu pour aligner strictement la mesure sur l'offre commerciale `web-mcp.fr/territoires/`. Sa surface utilisateur en France reste marginale au moment de la mesure ; nous concentrons les ressources de mesure sur les trois LLMs dont les volumes sont significatifs.
- **Échantillon de 3 prompts** — couvre trois types d'intention (découverte, quantitatif, localisation) mais ne saisit pas les requêtes réservation ou comparatives multi-stations. Étendu à 20-50 prompts dans le cycle mensuel.

## Trois questions. *Neuf transcriptions.*

Les réponses brutes des trois assistants, reproduites sans embellissement. Sources citées en fin de chaque verbatim. Date de capture : 4 mai 2026.

### Q3 · « Office de tourisme de Valberg, c'est où ? »

CHATGPT (GPT-5-MINI, WEB SEARCH ACTIVÉE) – *correct*

« Place Charles Ginésy — 06470 VALBERG. Téléphone : +33 (0)4 93 23 24 25. »

Sources citées : valberg.com, fr.wikipedia.org

CLAUDE (SONNET 4.5) – *correct*

« Place Charles Ginésy, 06470 VALBERG » + 04 93 23 24 25.

Sources citées : valberg.com, demarchesadministratives.fr, alpesdazur-tourisme.fr

PERPLEXITY (SONAR-PRO) – *correct, le plus riche*

« Place Charles Ginésy, 06470 Valberg » + horaires détaillés par saison + services (WiFi, boutique, espace enfants).

Sources citées : 10 (valberg.com × 2, demarchesadministratives.fr, provence-alpes-cotedazur.com, cotedazurfrance.fr, mappy, petitfute.com, alpesdazur-tourisme.fr, pagesjaunes, france-voyage.com)

### Q2 · « Combien d'hôtels à Valberg ? »

CHATGPT (GPT-5-MINI, WEB SEARCH) – *le plus honnête*

Recense honnêtement la divergence : valberg.com officiel = 4 hôtels, dossier de presse de la station = 3, hotelhotel.com = 8.

Sources citées : 3.

CLAUDE (SONNET 4.5) – *raisonnable, sans valberg.com*

« 7 hôtels selon certaines sources, 8 selon d'autres ». Cite Le Chalet Suisse, Le Chastellan, La Vallée Blanche, Le Blanche Neige, Adonis Valberg Grand Hotel Residence.

Sources citées : cybevasion.fr, hotelhotel.com (ne cite pas valberg.com).

PERPLEXITY (SONAR-PRO) – *le plus exhaustif*

« Environ 10 à 11 hôtels » selon Tripadvisor (11) et Booking (10).

Sources citées : 10 (valberg.com, tripadvisor.fr, booking.com, petitfute.com, adonis-valberg-lechastellan.com, trivago.fr, sowell.fr, kayak.fr).

### Q1 · « Que faire à Valberg ce week-end ? »

CHATGPT (GPT-5-MINI, WEB SEARCH) – *aucune réponse*

« Pour quel week-end parlez-vous exactement ? Voulez-vous dire le week-end qui arrive... ? »

Sources citées : 0.

CLAUDE (SONNET 4.5) – *renvoi générique*

*Mentionne DOME, astronomie, Val'Pass, événement « 90 ans en 2026 ». Pas de programme du week-end concret.*

Sources citées : valberg.com (× 3), tripadvisor.com.

PERPLEXITY (SONAR-PRO) – *le plus honnête sur ses limites*

*« Les recherches ne listent pas d'animations confirmées à cette date exacte ». Liste les activités récurrentes (luge d'été, accrobranche, Bike Park, planétarium, golf).*

Sources citées : 9.

## Sol de vérité Google FR (Q1)

Le 1<sup>er</sup> résultat est [valberg.com/agenda/](https://valberg.com/agenda/) (site officiel). L'agenda affiche : Week-end Passion Rallye, Printemps des séniors, Festival Astro'Valberg, Championnat. Aucun de ces événements n'est nommé par les LLMs. La donnée existe en HTML, mais n'est pas exposée en `Event` JSON-LD ni en flux ICS public — invisible pour les agents.

# Trois sources publiques croisées. *Une fiche unifiée.*

## Sources

- **DATAtourisme** — base nationale touristique, licence ouverte Etalab. API REST géo-filtrée. Ontologie propre dérivée de Schema.org.
- **Google Places** — API officielle (Nearby Search), 10 catégories ciblées ( `lodging` , `restaurant` , `tourist_attraction` , `campground` , `spa` , `health` , `gym` , `museum` , `art_gallery` , `park` ).
- **Cloudflare Browser Run** ( `/crawl` ) — crawl JavaScript-rendu des sites communaux pour extraire entités structurées et liens externes.

Optionnel : surcharges manuelles via `overrides.yaml` (ajouts, patches, exclusions explicites).

## Filtrage géographique

- **Centre station** — 44.075 N, 6.938 E (Valberg)
- **Rayon principal** — 5 km (cœur station, configuration `dar.config.yaml` )
- **Communes croisées** — Péone (06094), Beuil (06016), Guillaumes (06078)
- **Tiers** — `core` ( $\leq 3$  km), `nearby` ( $\leq 15$  km), `excluded` (au-delà)

## Dédoublonnage

Fusion par identifiant (Place ID Google + slug DATAtourisme) et proximité géographique ( $\leq 50$  m + similarité de nom  $> 0,8$ ). Priorité de merge : `manual` > `Google Places` > `crawl` > `DATAtourisme` .

## Classification d'éligibilité

VERDICT	CRITÈRE
<code>eligible</code>	Type accepté ( <code>HotelEstablishment</code> , <code>Restaurant</code> , <code>TouristAttraction</code> , <code>SportsAndLeisurePlace</code> , <code>TouristInformationCenter</code> , <code>NaturalHeritage</code> , <code>CulturalSite</code> ) sans caractère excluant
<code>eligible_if_registered</code>	Meublé de tourisme (article L324-1-1) — éligible si déclaré en mairie. Borne haute.
<code>ineligible_*</code>	Service public, retail générique, service personnel sans angle touristique
<code>duplicate</code>	Doublon non fusionné (rare après dédup)
<code>ambiguous</code>	Typage Google Places imprécis ( <code>LocalBusiness</code> appliqué à des restaurants/bars ) — à reclasser après revue manuelle

## Boucle de retour open data

Chaque entité enrichie ou corrigée par notre pipeline (téléphone manquant, géolocalisation imprécise, doublon fusionné, description complétée) est restituée aux datasets publics d'origine. Le travail bénéficie autant à votre OT qu'à l'écosystème open data français — contribution, pas captation.

## Trois communes. *109 entités uniques.*

### Stats par commune

COMMUNE	ENTITÉS TOTALES	TAGÉES VALBERG	CŒUR ( $\leq 3$ KM)	PROCHES ( $\leq 15$ KM)
<b>Peone</b>	108	74	2	72
<b>Beuil</b>	385	57	52	5
<b>Guillaumes</b>	168	80	3	77

*Après dédoublonnage cross-communes : 109 entités uniques.*

### Verdicts de classification

<p><b>55</b></p> <p>strictement éligibles 50.5 % du total</p>	<p><b>27</b></p> <p>meublés conditionnels 24.8 % du total</p>	<p><b>19</b></p> <p>hors champ 17.4 % du total</p>	<p><b>8</b></p> <p>cas ambigus 7.3 % du total</p>
-------------------------------------------------------------------	-------------------------------------------------------------------	--------------------------------------------------------	-------------------------------------------------------

*Le classeur **xlsx** joint contient les 109 entités détaillées en cinq feuilles (Synthèse + A1 éligibles + A2 conditionnels + A3 hors champ + A3 bis ambigus), avec colonnes éditables pour le suivi opérationnel.*

## Trois formats. *Une seule fiche.*

La fiche destination Valberg est servie depuis `web-mcp.fr/destinations/valberg/` dans trois formats consommables sans plugin et sans intégration côté Valberg.

### SCHEMA.ORG / JSON-LD

Vocabulaire structuré standardisé, maintenu conjointement par Google, Microsoft, Yahoo. Reconnu nativement par les LLMs depuis leur entraînement. Embarqué dans une balise `<script type="application/ld+json">` de la page. Une entité racine `TouristDestination` nommée « Valberg », avec son périmètre géographique, sa description officielle, ses coordonnées, et la liste explicite des 55 entités enfant éligibles.

### LLMS.TXT

Fichier texte placé à la racine indiquant aux crawlers IA quelles ressources lire en priorité, dans quel ordre, dans quel format. Émergent depuis 2024, supporté par Anthropic, OpenAI, Cloudflare. Notre fiche pointe explicitement vers les listes structurées hébergements, restaurants, activités, événements.

### HTML STATIQUE EN EDGE

Servi via Cloudflare Pages, accessible sans JavaScript, sans cookies, sans authentification. Citable par n'importe quel crawler. Durable, archivable.

### À VENIR : ENDPOINT WEBMCP

WebMCP (Web Model Context Protocol) est le standard W3C qui permet aux agents IA d'**exécuter des actions** au nom de l'utilisateur (réserver, vérifier disponibilités, consulter agenda). Supporté par Chrome 146 bêta, généralisation attendue à partir de 2027. La fiche actuelle évoluera vers un endpoint MCP **sans modification côté Valberg** — la couche infrastructure est déjà en place.

### DISTRIBUTION MULTI-CANALE (EN COURS)

Au-delà de l'exposition côté serveur, l'atlas est en cours de mise à disposition sur les principales surfaces de découverte agent : registre Anthropic MCP, Custom GPT ChatGPT, Zapier MCP marketplace, intégrations Cursor / Continue. Échéance cible : été 2026.

# Mensuel. *Évolutif.*

## Cadence et maintenance

La fiche est régénérée **chaque mois** à partir des sources publiques. Les corrections que vous nous signalez par email sont intégrées dans le cycle mensuel suivant ; les corrections urgentes hors-cycle restent possibles sur demande.

Les rapports de citation par assistant IA (ChatGPT, Claude, Perplexity) sont produits **mensuellement** à partir d'un échantillon de **20 à 50 requêtes types**. Le premier rapport est disponible à **partir du mois 2** suivant la publication. Quand l'endpoint WebMCP sera activé, vous serez la première destination prévenue.

## Limites méthodologiques

**Typage Google Places.** Les entités issues de Google Places sont labelisées telles quelles par l'API, sans garantie de typage rigoureux. Le filtre d'éligibilité corrige ces approximations en s'appuyant sur les classes de l'ontologie DATAtourisme. Reste 7.3 % d'entités classées « ambiguës » qui demandent une revue manuelle.

**Meublés de tourisme.** Non éligibles à DATAtourisme sauf déclaration en mairie au titre de l'article L324-1-1 du Code du tourisme et, le cas échéant, classement par la préfecture. Nous n'avons pas vérifié ce statut au cas par cas — la liste de 27 meublés conditionnels est une borne haute.

**Crawl des sites communaux.** Sur ce périmètre, le crawl n'a produit aucune entité structurée exploitable — les sites communaux concernés ne déclarent pas Schema.org. Le crawl reste utile pour découvrir des liens externes (annuaires partenaires, sites prestataires) qui alimenteront les itérations futures.

**Pipeline en évolution continue** selon la veille technologique et l'apparition de nouvelles sources publiques utiles. Les changements de méthode sont versionnés et communiqués dans les rapports mensuels.

## Dix termes. *Définitions.*

TERME	DÉFINITION
<b>MCP</b> (Model Context Protocol)	Protocole standardisé qui permet à un agent IA de découvrir et invoquer des outils exposés par un serveur (lecture de données, exécution d'actions). Origine : Anthropic, 2024. En cours de standardisation au W3C.
<b>WebMCP</b>	Spécification W3C qui adapte MCP au web : permet à une page web d'exposer des outils consommables par les agents IA du navigateur. Supporté par Chrome 146 bêta. Généralisation attendue : 2027.
<b>llms.txt</b>	Fichier texte placé à la racine d'un site qui indique aux crawlers IA quelles ressources lire en priorité, dans quel ordre, dans quel format. Émergent depuis 2024, supporté par Anthropic, OpenAI, Cloudflare.
<b>Schema.org</b>	Vocabulaire structuré standardisé pour décrire des entités sur le web. Maintenu conjointement par Google, Microsoft, Yahoo. Reconnu nativement par les LLMs.
<b>JSON-LD</b>	Format de sérialisation de Schema.org en JSON. Embarqué dans une balise <code>&lt;script type="application/ld+json"&gt;</code> . Lisible par tous les crawlers, invisible pour l'utilisateur humain.
<b>DATAtourisme</b>	Base nationale française d'informations touristiques publiques, alimentée par les CRT/CDT/OT régionaux. Licence ouverte Etalab. Ontologie propre dérivée de Schema.org.
<b>Meublé L324-1-1</b>	Article du Code du tourisme imposant qu'un meublé de tourisme soit déclaré en mairie pour être commercialement actif. Critère d'éligibilité à DATAtourisme.
<b>Endpoint</b>	Point d'accès URL d'une API ou d'un service. Pour MCP : l'URL où un agent IA envoie ses requêtes pour découvrir et invoquer les outils.
<b>Domain authority</b>	Mesure agrégée de l'autorité d'un domaine vis-à-vis des moteurs de recherche, fonction du profil de backlinks et de l'historique. Indicateur indirect de crédibilité auprès des LLMs.
<b>Web agentique</b>	Couche émergente du web où les agents IA, plutôt que les humains, sont les principaux consommateurs et acteurs des sites. Lecture (RAG, citations), écriture (réservations, transactions) via des protocoles type MCP/WebMCP.

# Tout est vérifiable.

SOURCE	DESCRIPTION	LICENCE
DATAtourisme	Base nationale touristique — datatourisme.fr	Etalab (ouverte)
Google Places	Données d'établissements et avis utilisateurs (API officielle)	Conditions d'usage Google
Cloudflare Browser Run	Crawl JavaScript-rendu des sites communaux	Conditions Cloudflare
DataForSEO	Mesures de citation IA — dataforseo.com	Compte payant
INSEE / IGN / SIRENE	Référentiels publics français	Etalab (ouverte)

## Ontologie de référence

<https://www.datatourisme.fr/ontology/core/> — classes acceptées : `HotelEstablishment` , `Restaurant` , `TouristAttraction` , `SportsAndLeisurePlace` , `TouristInformationCenter` , `NaturalHeritage` , `CulturalSite` . Critères de rejet : utilités publiques, retail générique, services personnels sans angle touristique, meublés sans déclaration L324-1-1.

## Pages web-mcp.fr utiles

PAGE	URL
Présentation de l'offre territoires	<a href="http://web-mcp.fr/territoires/">web-mcp.fr/territoires/</a>
Fiche commune Beuil (déployée, périmètre Valberg)	<a href="http://web-mcp.fr/communes/06016-beuil/">web-mcp.fr/communes/06016-beuil/</a>
Fichier llms.txt de Beuil	<a href="http://web-mcp.fr/communes/06016-beuil/llms.txt">web-mcp.fr/communes/06016-beuil/llms.txt</a>
Schema.org JSON-LD de Beuil	<a href="http://web-mcp.fr/communes/06016-beuil/schema/commune.jsonld">web-mcp.fr/communes/06016-beuil/schema/commune.jsonld</a>
Atlas civique alpha (163 communes 06)	<a href="http://web-mcp.fr/atlas/">web-mcp.fr/atlas/</a>

## Avant publication. *Trois vérifications, par email.*

Avant que la fiche destination Valberg passe en publication, nous vous invitons à effectuer trois lectures, chacune renvoyant un email court à [contact@web-mcp.fr](mailto:contact@web-mcp.fr). Ces vérifications restent informelles — il ne s'agit pas d'une procédure de validation contractuelle, mais d'une opportunité de capter ce que les sources publiques ne nous permettent pas de détecter.

### 1. Classes d'éligibilité.

Vérifier que les classes DATAtourisme retenues (cf. § T3 et § T8) correspondent à votre lecture de l'ontologie pour le périmètre station. Si vous identifiez une classe pertinente que nous n'aurions pas mobilisée, ou inversement une classe que nous mobilisons et qui ne devrait pas l'être pour Valberg, nous corrigeons.

### 2. Exclusions opérationnelles.

Signaler les entités qui figurent dans les listes mais devraient être exclues pour des raisons que nos sources publiques ne permettent pas de détecter : litige privé en cours, fermeture non encore déclarée, restructuration prévue, demande explicite d'un prestataire de ne pas figurer.

### 3. Périmètre 3 communes.

Confirmer que le périmètre Péone + Beuil + Guillaumes correspond bien à votre lecture institutionnelle de la station, ou suggérer une délimitation différente (ajout de Daluis ? restriction au cœur Péone+Beuil ?).

Toute remarque par email est intégrée au cycle mensuel suivant. Pas de validation formelle requise — votre silence vaut accord tacite, mais votre retour vaut beaucoup plus.

#### CONTACT TECHNIQUE

*équipe web-mcp.fr · [contact@web-mcp.fr](mailto:contact@web-mcp.fr)*

*Pour la proposition commerciale et les engagements, voir la note commerciale jointe (Doc A).*