

# Google Maps POI & Review Scraping

Complete guide to scraping Google Maps POI data and reviews — 70+ tools across 20 categories. Compiled June 2026.

- [Official Google APIs](#)
  - [Places API \(New\) — Overview & Pricing](#)
  - [Google Business Profile API](#)
- [Reverse-Engineered Internal Endpoints](#)
  - [Protobuf Endpoints & Internal APIs](#)
  - [Protobuf Tools & Decoders](#)
- [Browser Automation & DIY Techniques](#)
  - [Headless Browser Automation \(Playwright, Puppeteer, Selenium\)](#)
  - [DIY Scraping Techniques \(Reddit & Community\)](#)
- [Commercial Platforms & Search API Proxies](#)
  - [Commercial Scraping Platforms](#)
  - [Search API Proxies \(SerpAPI, Serper, etc.\)](#)
- [Open-Source Scrapers & Python Libraries](#)
  - [Open-Source GitHub Scrapers](#)
  - [Python Libraries & Popular Times Tools](#)
- [No-Code, Extensions & Automation Workflows](#)
  - [No-Code Platforms](#)

- [Browser Extensions & Userscripts](#)
- [Automation Workflows \(n8n, Make, Zapier\)](#)
  
- [AI-Powered Scrapers & Alternative POI Sources](#)
  - [AI/LLM-Powered Scrapers](#)
  - [Alternative POI Data Sources \(No Scraping\)](#)
  
- [Anti-Bot Countermeasures & Legal Landscape](#)
  - [Google's Anti-Bot Countermeasures \(2025-2026\)](#)
  - [Legal & TOS Landscape](#)

# Official Google APIs

Places API (New), Business Profile API, pricing, field masks

# Places API (New) — Overview & Pricing

## Places API (New)

The current-generation API replacing the legacy Places API. Uses a mandatory **field mask** in every request — omitting it returns an error. Billing is determined by the highest-tier field requested.

### Three-Tier SKU Billing

Tier	What's Included	Cost
<b>Essentials</b>	Place ID, name, address, location, types, business status	Lowest
<b>Pro</b>	Phone, website, opening hours, price level, photos	Medium
<b>Enterprise</b>	Reviews, ratings, userRatingCount, editorial summary, atmosphere data	Highest (~\$17-35/1K requests)

### Subscription Plans

Plan	Cost	Calls	Savings
Pay-as-you-go	Variable	Unlimited	Baseline
Starter	\$100/mo	50K	Moderate
Pro	\$1,200/mo	250K	Up to \$5,000 vs PAYG
Enterprise	Custom	Custom	Negotiated

### Key API Operations

- **Place Details** — full data for a known place\_id

- **Text Search** — search by query string (e.g., "restaurants in Berlin")
- **Nearby Search** — search within a radius of coordinates
- **Autocomplete** — predictive place suggestions
- **Place Photos** — retrieve photos by reference

## Rate Limits

Enforced **per API method per project** (unlike legacy API which aggregated all methods). Default ~600 requests/minute; increases require contacting Google support.

## Free Tier Status (2026)

**Warning:** The previously documented "\$200 monthly credit" and "10,000 free calls" claims were both refuted during adversarial verification. Verify directly at [mapsplatform.google.com/pricing](https://mapsplatform.google.com/pricing).

## Data Fields (50+)

Name, formatted address, address components, phone, international phone, website, opening hours (regular + secondary), business status, price level, rating, user rating count, reviews (text/author/rating/time/language — **max 5 per place**), photos, types, coordinates, viewport, place ID, URL, UTC offset, plus 20+ boolean amenities (delivery, dineIn, takeout, outdoorSeating, liveMusic, allowsDogs, etc.).

## New-Only Fields (Not in Legacy API)

`regularSecondaryOpeningHours`, `paymentOptions`, `parkingOptions`, `subDestinations`, `fuelOptions`, `evChargeOptions`, `shortFormattedAddress`, `primaryType`, `primaryTypeDisplayName`

## Sources

- [Usage and Billing](#)
- [Data Fields Reference](#)
- [Operations Overview](#)
- [Maps Platform Pricing](#)

# Google Business Profile API

## Google Business Profile API

Access **all reviews** (not limited to 5) — but only for locations you own or manage. Requires OAuth2 authentication.

### Key Endpoints

Endpoint	Purpose
<code>accounts.locations.reviews.list</code>	All reviews for one location
<code>accounts.locations.batchGetReviews</code>	Batch review extraction across locations

Returns: `reviewId`, `starRating`, `comment text`, `creation/update timestamps`. Free for business owners.

### Google Takeout Alternative

Business owners can download all reviews from their GBP as JSON via Google Takeout (free, official).

Source: [Business Profile API — Review Data](#)

# Reverse-Engineered Internal Endpoints

Protobuf APIs, tbm=map, APP\_INITIALIZATION\_STATE, pagination

# Protobuf Endpoints & Internal APIs

## Google Maps Internal Protobuf API

Google Maps communicates with its backend using **Protocol Buffer (protobuf) encoding** over HTTP. Several internal endpoints have been documented.

### tbm=map Search Endpoint

Adding `tbm=map` to a Google search returns protobuf-over-JSON instead of HTML. Single HTTP request, no browser rendering needed.

```
https://www.google.com/search?tbm=map&hl=en&q=restaurants+in+berlin&tch=1&ech=1
```

Returns: nested arrays parseable with JSON indexing — business names, ratings, review counts, categories, addresses, coordinates, thumbnails.

### /maps/preview/place — Place Details

Returns protobuf-over-JSON place details via a `pb` parameter with feature ID embedded. More stable than HTML scraping — follows internal schema, not frontend layout.

### /maps/vt/pb — Tile Endpoint

Map tile requests containing ~730 fields across ~125 protobuf messages. Uses `!`-separated text-based protobuf with type characters (s, i, d, f, j, u, v, x, y, g, h, n, o, e, z, B, b, m).

### window.APP\_INITIALIZATION\_STATE

Google Maps embeds data in `window.APP_INITIALIZATION_STATE` in the HTML response. No headless browser needed:

```
JSON.parse(window.APP_INITIALIZATION_STATE[3][2].split("\n")[1])
```

Source: [r/webscraping](#)

## Protobuf Pagination Details

- **PSI parameter:** stored in `window.APP_OPTIONS[11]`, changes after each page reload — no static URL construction possible
- **Altitude formula:**  $\text{altitude} = (27.3611 * 6371010 * 768 * \cos(\text{lat})) / (2^{\text{zoom}} * 256)$
- **Pagination:** `!7i20` = 20 results/page, `!8i[offset]` = offset (increments of 20)
- **Responses:** served through `f.txt` endpoint with "XHR1" signature

## Data Available

Potentially *everything* visible in the Maps UI: full review text with timestamps, all photos, Q&A, popular times histograms, wait times, related places, owner responses — far exceeding the official API's 5-review limit.

## Fragility Warning

Google's cryptographic constants rotate continuously. SearchGuard makes "reverse-engineered bypasses obsolete within minutes." One Reddit user reported successfully building a pure HTTP/protobuf scraper that reduced costs from \$50K/mo to \$1K/mo — but maintenance is constant.

# Protobuf Tools & Decoders

## Protobuf Tools

Tool	Language	Purpose	Link
<b>pbtk</b>	Python	Extract and fuzz protobuf structures from Chrome	<a href="#">GitHub</a>
<b>google-maps-pb-decoder</b>	Ruby	Decode pb URL params to JSON	<a href="#">GitHub</a>
<b>deproto</b>	Python	Decode, manipulate, re-encode protobuf strings (round-trip conversion)	<a href="#">GitHub</a>

## XHR/Network Interception

Intercept browser XHR requests during Google Maps navigation; responses contain protobuf-encoded data with "XHR1" signature. Use CDP `Network.requestWillBeSent` event + `history.replaceState()` hooks. Supported by Puppeteer request interception and Playwright route handlers.

## Sources

- [SerpAPI — Reverse Engineering Maps Pagination](#)
- [\\$5K Google Maps XSS via Protobuf](#)
- [Scrape.do — Maps Scraping Guide](#)

# Browser Automation & DIY Techniques

Playwright, Puppeteer, Selenium, community tricks from Reddit

# Headless Browser Automation (Playwright, Puppeteer, Selenium)

## Browser Automation Frameworks

The foundation technique used by most scrapers. A browser renders Maps, executes JS, and extracts from the DOM.

### Playwright (Most Popular)

Microsoft's library. Used by gosom/google-maps-scraper, HasData, many commercial tools. Supports Chromium, Firefox, WebKit.

- Languages: Python, Node.js, Java, .NET
- Anti-detect: stealth plugins available
- Parallel tabs: 10+ tabs in one browser, ~1.7s/URL
- Codegen: record browser interactions, auto-generate scraper code

### Puppeteer (Node.js)

Google's Node.js browser automation. `puppeteer-extra-plugin-stealth` provides 17 evasion modules.

- Best stealth ecosystem (17 modules)
- XHR interception via `Network.requestWillBeSent`
- Caveat: anti-bot companies study the stealth package

### Selenium (Legacy)

Original framework. `undetected-chromedriver` patches for detection evasion.

- Languages: Python, Java, C#, Ruby, JS
- Larger fingerprint, easier to detect
- Still used by HasData and Zubdata scrapers

## Extraction Strategies

Strategy	Token Cost	When to Use
CSS selectors ( <code>querySelectorAll</code> )	~52/item	Known structure — default choice
<code>aria-label</code> attributes	~52/item	More resilient — accessibility attrs are stabler than CSS classes
<code>body.innerText</code>	~5K/page	Discovery — learn structure once, then switch
Network/XHR interception	Minimal	Capture protobuf responses directly — best approach
Accessibility tree (filtered)	~28K/page	Find buttons, forms, interactive elements
Screenshot	~132K	CAPTCHA solving, visual debugging only

# DIY Scraping Techniques (Reddit & Community)

## Community-Sourced Techniques

Practical techniques from Reddit, HN, and forums — the stuff not in vendor blog posts.

### Grid-Scanning / Map Splitting (Essential)

Overcome the **120-result limit** per search by dividing areas into smaller grid cells. Auto-adjust zoom to ~16 per cell. Can use QGIS + Python to generate coordinate grids. Tune tightness to urban density.

```
https://www.google.com/maps/search/{query}/@{lat},{long},16z
```

Used by every serious scraper (gosom, Apify actors, Octoparse). Hexagonal grid sampling is the academic version.

### "Search in This Area" Automation

Moving the map and clicking "search in this area" reveals different/more results than initial search. Automate by systematically panning across coordinates.

### HTTP-Only / No Browser Approach (50x Cheaper)

Replicate Google Maps' internal protobuf HTTP calls directly — no headless browser. One HN user reduced costs from **\$50K/mo to \$1K/mo**. "Built entirely with cURL, avoiding headless browsers since they're really slow." Google rarely updates their map API structure.

# Dual-Architecture Engine

Use high-performance HTTP extraction for bulk data, only spin up headless browsers for dynamic elements. Extracts Booking URL, FID, CID from metadata. Faster than pure Playwright/Puppeteer.

## Place ID Pre-Extraction

Get Place IDs first (cheap/fast), then pipe them for direct detail extraction without browser search overhead — data comes back "almost instantly."

## CID-Based Direct Access

CID (Customer ID) is a 64-bit decimal that **never changes**, even through rebrand/relocation. More permanent than Place ID (which expires after 12 months).

```
https://www.google.com/maps/place/?cid=<CID_NUMBER>
```

Find CID: URL `?cid=` parameter, or inspect Knowledge Panel for "ludocid".

Reference: [Scrap.io CID Guide](#), [CID Converter](#)

## LLM-Powered Query Expansion

Use LLMs to generate search term synonyms to overcome per-query result limits. "dentist" becomes "dental clinic", "oral surgeon", "dental practice", "orthodontist" etc. Multiplies coverage without geographic splitting.

## Other Community Tricks

- **Ctrl+S page saving**: Bot performs search, saves entire page, parses offline later
- **Google Search local pack mining**: Harvest Maps data from regular Google SERPs (lower detection risk)
- `udm=1` **parameter**: Returns Google Places tab results — but needs browser context for location
- **Geocoding API grid trick**: Grid area, geocode lat/lng to addresses, dedupe. 80%+ coverage
- **"No website" filter**: Gold mine for web dev agencies — scrape businesses without websites

# Sources

- [r/webscraping — Scraping GMaps at Scale](#)
- [r/SaaS — Dual-Architecture GMaps Scraper](#)
- [r/webscraping — Google Maps Data Extraction](#)

# Commercial Platforms & Search API Proxies

Outscraper, Apify, Bright Data, SerpAPI, Serper.dev, etc.

# Commercial Scraping Platforms

## Commercial Scraping Platforms

### Outscraper (Best Value)

Pricing	<b>\$1-3/1K records</b> (basic); ~\$14/1K full enrichment
Free Tier	500 records (first 500 reviews free)
Fields	74 fields incl. name, address, phone, hours, rating, reviews, photos, emails
Reviews	Full text, not limited to 5
Real-world	Reddit user: \$50 for 20K listings, ~6,500 usable after dedup

### Apify (Marketplace)

Pricing	<b>\$1.50-4/1K places</b> (varies by actor)
Free	\$5 monthly platform credit
Fields	25+ core + enrichment actors for emails/social
Tips	Set <code>maxCrawledPlaces</code> to 9999999 for all results. Niche searches have 75-80% junk rate
Actors	compass/crawler-google-places, beatanalytics/google-maps-reviews-scraper

### Bright Data (Enterprise)

Dataset	<b>200.7M+ pre-scraped records</b> at \$0.0025/record (\$250/100K)
Scraper API	From \$500/mo + proxy costs
Proxy Network	72M+ residential IPs — industry's largest

## Other Commercial Services

Service	Pricing	Key Feature
Scrap.io	From EUR 35/mo (10K leads)	Simplest UX, radius/polygon search, count preview, 200M+ places
ScrapFly	~\$3.37/1K requests	General-purpose, credit-based, 99% success rate
LocalAPI	Free tier: 50 searches	Only API exposing Popular Times histograms
HasData	API plans	Cloud-powered, built-in email finder
Leads-Extractor.com	Various	Dedicated Maps scraper
MapSHunt.com	Various	Prioritizes businesses with websites, crawls for emails/social
LocalProspects	Various	Auto-crawls for owner name, emails, mobile phones
GeoScraper	Unlimited/mo	Reviews scraper + email/social extraction
Minelead	Various	Email finder for Maps businesses
Estrattore Dati	Various	Email and phone extraction

## Cost Comparison (per 1,000 POI records)

Platform	Basic POI	With Reviews	With Enrichment
Outscraper	\$1-3	\$5-8	~\$14
Apify	\$1.50-4	\$3-6	\$5-10
Bright Data (dataset)	\$2.50	Included	Included
Scrap.io	~\$3.50	Extra	Extra
ScrapFly	\$3.37+	Same	Same
PhantomBuster	\$50-80	N/A	N/A
Google Places API	\$17-35	Enterprise tier	Enterprise tier



# Search API Proxies (SerpAPI, Serper, etc.)

## Search API Proxies

Services that handle scraping/anti-bot, returning clean JSON from Google Maps queries.

Service	Pricing	Speed	Notes
<b>SerpAPI</b>	\$50/5K queries	Standard	Market leader. Active DMCA lawsuit from Google (Dec 2025)
<b>Serper.dev</b>	~\$1/1K queries; \$0.30/1K at volume	~2.87s	2,500 free queries. SERP-based Maps results
<b>Scrapingdog</b>	\$0.33/1K	3.05s	100% success rate (benchmarked), fastest
<b>Scrape.do</b>	\$1.16/1K	1.73s	AI Overview detection
<b>ScraperAPI</b>	\$50/4K searches	Standard	Maps in all plans
<b>ScrapeBadger</b>	Competitive	Standard	Review support
<b>DataForSEO</b>	Pay-per-task	Standard	Maps SERP + reviews endpoints
<b>Oxylabs SERP</b>	From \$49/mo	Standard	Enterprise SERP API
<b>ValueSERP</b>	Various	~40s (slow)	Some user complaints
<b>Zenserp</b>	From \$30/mo	Standard	Multi-engine support

## SerpAPI Details

Accepts standard Google Maps query syntax (`q=restaurants+in+Berlin`), returns structured JSON with: title, address, phone, rating, reviews, hours, service\_options, GPS coordinates, place\_id, data\_cid, price, type, photos, thumbnails.

Separate endpoint for individual place reviews.

**Legal warning:** Google v. SerpAPI (Dec 2025) — active DMCA lawsuit with ruling pending.

# Open-Source Scrapers & Python Libraries

GitHub repos, populartimes, gomaps, review scrapers

# Open-Source GitHub Scrapers

## Open-Source Scrapers

Repository	Language	Stars	Fields	Key Feature
<a href="#">gosom/google-maps-scraper</a>	Go	4,200	33+	120 places/min, CLI/Web/REST, K8s-ready, email extraction, SOCKS5/HTTP/HTTPS proxy rotation
<a href="#">omkarcloud/google-maps-scraper</a>	Python	2,700+	50+	Social profiles (LinkedIn/Twitter/FB/IG), ad spend detection, \$28 Pro
<a href="#">HasData/google-maps-scraper</a>	Python	-	7	3 backends: Selenium, Playwright stealth, HasData API
<a href="#">google-reviews-scraper-pro</a>	Python	-	Reviews	Multi-language reviews + images, MongoDB, bypassed Feb 2026 "limited view"
<a href="#">conor-is-my-name</a>	Docker	-	Various	Designed for n8n + Postgres. Multi-container, 1 req/min/IP = ~144K/day
<a href="#">webAutomationLover</a>	Userscript	-	Various	Tampermonkey, auto-scrolls, exports .xlsx. No server needed
<a href="#">alltheplaces</a>	Python/Scrapy	-	POI	20M+ POIs from 4,100+ spiders. CC-0 license, weekly updates

<a href="#">tripadvisor-scraper</a>	Python	-	Reviews	Alternative review source via TripAdvisor GraphQL
-------------------------------------	--------	---	---------	---

## Feb 2026 "Limited View" Lockdown

Google introduced a "limited view" restriction on review pages. **google-reviews-scraper-pro** bypassed it via search-based navigation (no login needed). Review pagination uses offset increments of 10, supports 4 sort modes (relevant/newest/highest/lowest).

# Python Libraries & Popular Times Tools

## Python Libraries

Library	Status	API Key?	Fields	Notes
<code>googlemaps</code> (official)	Active	Yes	50+ (API)	Official wrapper. Same pricing/limits as API
<code>gomaps</code>	Dormant (Sep 2023)	No	9	Name, URL, address, coords, website, phone, rating, hours, Popular Times. Likely broken

## Popular Times & Foot Traffic Tools

**Popular Times is NOT available via any official Google API.** Must be scraped or obtained from third-party services. Historical data is also inaccessible.

### populartimes (m-wrZR) — 900 Stars

```
pip install --upgrade git+https://github.com/m-wrZR/populartimes
```

Returns: weekly hourly popularity (0-100 scale), wait times, time spent, current popularity. Requires API key with billing. MIT license. Unstable as Maps updates.

### LivePopularTimes (GrocerCheck)

Extension adding live busyness data to populartimes. Functions: `get_populartimes_by_address` (no API call), `get_populartimes_by_PlaceID` (API + scrape), `get_places_by_search`.

# BestTime.app (Commercial)

Endpoints	Forecasts (hourly 0-100%), Live Busyness, Venue Filter (by traffic/day/hour/ratings/type), Venue Search
Coverage	150+ countries
Unique	Dwell time analysis, surge detection, filter by foot traffic intensity
Free Tier	Test account with limited credits

Website: [besttime.app](https://besttime.app)

## LocalAPI

REST API returning 20+ fields per business in 2-3 seconds including Popular Times histograms (peak hours by day of week). Claims to be the only API exposing Popular Times via endpoint. Free tier: 50 searches. [local-api.com](https://local-api.com)

## Workarounds for the 5-Review API Limit

Method	Reviews Available	Cost
Official Places API	5 max	Enterprise tier
Google Business Profile API (own locations)	All	Free
Google Takeout (own GBP)	All	Free
Outscraper	All	First 500 free
Apify reviews actors	All	\$0.25-0.40/1K
google-reviews-scraper-pro	All	Free (+ proxies)
Tampermonkey userscript	All (manual)	Free
DIY Playwright/Puppeteer scroll	All	Free (+ proxies)
SerpAPI reviews endpoint	All	\$50+/mo

# No-Code, Extensions & Automation Workflows

Octoparse, browser extensions, userscripts, n8n, Make.com

# No-Code Platforms

## No-Code / Low-Code Platforms

Platform	Pricing	Maps Feature	Notes
<b>Octoparse</b>	Free + \$69/mo	Auto search splitting to overcome 120-result limit	20K results/hour. Best no-code option
<b>ParseHub</b>	Free (5 projects) / \$189/mo / \$599/mo	JS rendering handles Maps SPA well	Steep price jump from free
<b>WebAutomation.io</b>	\$99-999/mo or \$1/1K PAYG	Pre-built Google Maps extractor	14-day unlimited trial
<b>PhantomBuster</b>	\$69-439/mo	Maps phantoms available	\$50-80/1K leads — 15-50x more expensive. NOT recommended for Maps

# Browser Extensions & Userscripts

## Browser Extensions & Userscripts

Tool	Type	Price	Notes
<b>DataMiner</b>	Chrome ext.	Free (500 pages/mo) to \$200/mo	Custom recipes, pagination, scheduled runs. No email extraction
<b>Instant Data Scraper</b>	Chrome ext.	Free	One-click, ~120 results cap. Unmaintained
<b>Web Scraper (webscraper.io)</b>	Chrome ext.	Free + cloud \$50/mo	No pre-built Maps template
<b>GMaps Extractor</b>	Chrome ext.	Free/Freemium	gmapsextractor.com. No coding needed
<b>G Maps Extractor</b>	Chrome ext.	Freemium	"Probably the cheapest" per Reddit
<b>AI Web Scraper</b>	Chrome ext.	Various	Multi-purpose: Reddit, Maps, LinkedIn, Amazon
<b>Tampermonkey Reviews Exporter</b>	Userscript	Free	<a href="#">Greasyfork</a> . Adds "Scrape Reviews" button, auto-expands, exports JSON
<b>webAutomationLover</b>	Userscript	Free	Tampermonkey. Accumulates results as you browse, exports .xlsx
<b>Firefox GPX Exporter</b>	Firefox addon	Free	Export saved Maps lists as GPX with lat/lon

**120-Result UI Limit:** All browser extensions are capped by Google Maps' sidebar pagination (~120 results per search). Must split searches into smaller areas for more.

# Automation Workflows (n8n, Make, Zapier)

## Automation Workflows

### n8n (Most Flexible)

- Official **Outscraper node** available
- Templates: Google Maps scraper, lead gen with Google Sheets export, email scraping
- **conor-is-my-name/google-maps-scraper**: Docker-based, designed for n8n + Postgres. Multi-container, IP rotation, 1 req/min/IP = ~144K/day
- **Gemini AI + n8n pipeline**: Use Google AI Studio to vibe-code scraping, pipe via webhook to Sheets
- Example pipeline: Apify scrape + Anymail Finder emails + Airtable storage
- Also works with official Google Maps API via HTTP Request node

### Make.com

Scrap.io integration (200M businesses). Visual workflow builder, 3,000+ connectors. Automated lead gen tutorials available.

### Zapier

Outscraper integration available. Simpler but less flexible than Make.com.

### Clay

Data enrichment platform. Pull basic Maps data then enrich with Clearbit/Dropcontact. Filter by review count/recency for active businesses.

# Google Sheets & Apps Script

- **Apps Script Maps Service:** built-in geocoding/reverse-geocoding without API keys.  
Cannot do Places search
- **IMPORTFROMWEB add-on:** handles dynamic JS pages in Sheets (IMPORTXML fails on Maps)
- **Full cold outreach pipeline** (Apps Script v3.0): AI email writing, website discovery, CEO identification, Firecrawl scraping, email enrichment, dedup — all in one sheet

# AI-Powered Scrapers & Alternative POI Sources

ScrapeGraphAI, Crawl4AI, Overture Maps, AllThePlaces, OSM

# AI/LLM-Powered Scrapers

## AI/LLM-Powered Scrapers

### ScrapeGraphAI (MIT / Free)

Describe what you want in plain English, it extracts structured JSON without CSS/XPath selectors. Auto-adapts to layout changes. Supports GPT, Gemini, Groq, Azure, Hugging Face, local Ollama models.

[GitHub](#)

### Crawl4AI (58K Stars, Apache 2.0)

Open-source LLM-friendly web crawler. Outputs clean Markdown for RAG/agents. Heuristic noise filtering, CSS/XPath/LLM extraction. Local-first (no API costs).

[GitHub](#)

### Firecrawl

Developer API that outputs clean Markdown from any URL, optimized for LLM ingestion. Used in enrichment pipelines alongside Maps scrapers.

[firecrawl.dev](https://firecrawl.dev)

## LLM Query Expansion

Not a scraper per se, but a technique: use LLMs to generate category synonyms and related search terms to multiply coverage per geographic area. "dentist" becomes "dental clinic", "oral surgeon", "dental practice", etc. Combined with any scraper for 3-5x more results.

# Alternative POI Data Sources (No Scraping)

## Alternative POI Data Sources

Free or commercial POI databases that don't require scraping Google Maps.

### Overture Maps Foundation (64.8M POIs Free)

Backed by Meta, Microsoft, Amazon, TomTom, Foursquare. The strongest free alternative to Google Maps data.

Size	<b>64.8M places</b> — Meta (~59.2M), Foursquare (~6.7M), Microsoft (~7.4M), AllThePlaces (~1.7M)
Fields	Names, categories (64+), phones, emails, websites, socials, addresses, brand, operating_status, confidence score, coordinates
Format	GeoParquet on S3 and Azure Blob. Python CLI, DuckDB SQL, browser Explorer
License	CDLA-Permissive-2.0 / ODbL (commercial use OK)
Limits	Monthly updates (not real-time). No reviews/photos. Thinner outside Western countries

**Warning:** Reddit reports Places layer stopped updating as of Sept 2024 release. Verify current state.

[Overture Places Guide](#) | [Downloads](#)

### Overture-Based API (Community-Built, 200x Cheaper)

A developer built a Places API using Overture data + Rust/Axum + PostGIS. Free 5K/mo, \$10/100K, \$30/500K, \$80/2M — vs Google's ~\$1,700 for 100K.

## Other Alternative Sources

Source	POIs	Free Tier	Key Strength	Reviews?
<b>OpenStreetMap (Overpass API)</b>	Varies	Unlimited, no key	Free, query any tag combo. <a href="#">Overpass Turbo</a>	No
<b>Foursquare Places</b>	100M+	Commercial	Richest venue data, behavioral insights, check-ins	Tips only
<b>HERE Technologies</b>	Global, 400+ cats	250K tx/mo	TripAdvisor ratings, EV/fuel data, chain ID	Via TripAdvisor
<b>TomTom</b>	~100M, 180+ countries	50K daily tx	Navigation-optimized, relevance scoring	No
<b>Mapbox</b>	Global	100K req/mo	Polished SDKs	No
<b>Geoapify</b>	OSM-based, 400+ cats	3K credits/day	Transparent pricing, can cache/store	No
<b>Yelp Fusion API</b>	Millions	5K calls/day	3 reviews max via API. Open Dataset: 8.6M reviews (academic)	3 (API), 8.6M (dataset)
<b>AllThePlaces</b>	20M+	Unlimited (CC-0)	4,100+ Scrapy spiders, weekly updates	No
<b>MapQuest</b>	Various	Limited	Search API v5: radius/rect/polygon/corridor	No
<b>Nominatim</b>	OSM geocoding	Free	Forward/reverse geocoding	No
<b>Photon (Komoot)</b>	OSM geocoding	Free	Typo-tolerant, multilingual	No

**OSM Reality Check (from Reddit):** OpenStreetMap is poor for business/POI data. Business coverage is ~75% at best, biased toward bigger/popular locations.

## Data Providers & Marketplaces

Provider	Dataset	Pricing	Google Maps?
----------	---------	---------	--------------

Bright Data	200.7M+ records	\$0.0025/record	Yes
Datarade	60M+ US (varies)	By provider	Yes
Veridion	134M+ businesses	~\$99/user/mo	Partial
SafeGraph	52M+ POIs	Commercial	Partial (foot traffic)
Dataplor	Various	Commercial	Partial (LatAm strong)
Xtract.io	6M+ locations	Commercial	Partial
Coresignal	N/A	N/A	No (employee data only)
Data.world	N/A	N/A	No (data governance)

## Export & Data Liberation Tools

- **Google Takeout:** Export saved/starred places (CSV/JSON, no coordinates — needs geocoding)
- [Takeout Tools](#): Adds coordinates, converts to GeoJSON/KML/GPX
- [json2kml](#): Python converter for saved places to KML
- [Export-Google-Maps-Saved-Places](#)
- **Firefox GPX Exporter:** Export saved lists as GPX with lat/lon

# Anti-Bot Countermeasures & Legal Landscape

BotGuard/SearchGuard, proxy strategies, Google v. SerpAPI, GDPR

# Google's Anti-Bot Countermeasures (2025-2026)

## Anti-Bot Countermeasures

Google Maps difficulty score: **90/100** — one of the toughest platforms to scrape.

## BotGuard / SearchGuard Architecture

Built on **BotGuard** (internally "Web Application Attestation"), deployed across YouTube, reCAPTCHA v3, and Maps. **SearchGuard** (January 2025) is the Search-specific evolution — "tens of thousands of person hours and millions of dollars."

## Detection: Behavioral Analysis (4 Signal Categories)

Signal	Bot Threshold	Human Range
Mouse movement (trajectory, velocity, acceleration, micro-tremors)	Velocity variance <10	50-500
Keyboard rhythm (inter-key intervals, duration, errors)	Variance <5ms	20-50ms
Scroll behavior (amplitude, direction, timing)	Delta variance <5px	20-100px
Timing jitter (Welford's algorithm)	>200 events/sec	10-50

# Detection: Browser Fingerprinting (100+ Signals)

Navigator, screen, performance metrics, WebRTC leaks, TLS fingerprinting. Explicit checks for `navigator.webdriver`, ChromeDriver, Puppeteer, Selenium, PhantomJS.

## reCAPTCHA v3 (Invisible)

No visible challenge — assigns 0.0-1.0 score based on session behavior. On Maps, appears **inconsistently**. Only bypass: never trigger it.

## Cryptographic Protection

ARX cipher (similar to NSA's Speck). Magic constants rotate per script update. Scripts served with integrity hashes. **Bypasses become obsolete within minutes.**

## Blocking Behavior

Trigger	Risk
Datacenter IPs	Blocked immediately — non-viable
Uniform request timing	High — humans pause, bots don't
Direct navigation to data pages	Medium — humans wander first
Default/missing headers	Medium

**Stealth detection:** Google returns **poisoned/incomplete data** to detected scrapers rather than blocking outright. Soft bans aren't clean HTTP codes — "some weird JSON response or just incomplete page loads." DOM class names change every few months.

## Proxy Effectiveness

Type	Effectiveness	Cost	Notes
Datacenter	Very Low	\$	Non-viable for Maps
Residential	Good	\$\$	~30-50 searches/hour/IP before CAPTCHAs
Mobile (4G/5G)	Best	\$\$\$	Most reliable for large-scale

ISP	OK (low volume)	\$\$	Burns fast. Once flagged, stays flagged for days
Hybrid	Cost-optimized	\$\$	Datacenter for non-Maps, mobile for Maps. Saves 40-60%

Practical rate limit: **max 1 req/min/IP** = ~144K results/day per IP.

## Anti-Detection Toolkit

- `playwright-stealth` / `puppeteer-extra-plugin-stealth` (17 evasion modules) — near-essential
- `undetected-chromedriver` for Selenium
- CapMonster Cloud for automated CAPTCHA solving
- FlareProx — Cloudflare Workers proxy for IP masking (100K free daily requests)
- Cookie banner handling: `document.querySelector('#sp-cc-accept')?.click()`
- CAPTCHA detection: `document.querySelector('form[action*="validateCaptcha"]')`
- Image/font blocking via `Network.setBlockedURLs`
- Search-based navigation to bypass Feb 2026 "limited view" lockdown

# Legal & TOS Landscape

## Legal & TOS Landscape

### Google Maps TOS

**"Customer will not export, extract, or otherwise scrape Google Maps Content for use outside the Services."**

This is a contractual prohibition, not criminal statute. Breach of contract, not a crime.

## Google v. SerpAPI (Dec 2025 — Ruling Pending)

**Filed:** December 19, 2025, N.D. California (Case No. 4:25-cv-10826)

### Google's DMCA Claims

1. **Access circumvention** (17 U.S.C. 1201(a)(1)(A)): SerpAPI circumvented SearchGuard "on billions of separate occasions." \$200-\$2,500 per violation.
2. **Trafficking in circumvention tools** (17 U.S.C. 1201(a)(2)): Marketing services to bypass SearchGuard.

SerpAPI's requests increased "25,000%" over two years — hundreds of millions daily.

### SerpAPI's Defense (Motion to Dismiss, Feb 2026)

- Google doesn't own copyright to third-party search content
- SearchGuard protects business model, not copyrighted works
- "Google's entire business began with a web crawler that copied the content"

Hearing: May 19, 2026 before Judge Yvonne Gonzalez Rogers — **ruling not yet published.**

**Industry impact:** If Google prevails, rank tracking, competitive intelligence, and SEO analytics could become legally untenable.

# Key Legal Precedents

Case	Year	Impact
Van Buren v. US	2021	CFAA limited to insiders. ToS violations are not computer crime
hiQ v. LinkedIn	2022	Public data scraping doesn't violate CFAA (Ninth Circuit, reaffirmed)
X Corp v. Bright Data	2023	Platforms can't claim copyright on user-generated content
Meta v. Bright Data	2024	Logged-out users haven't accepted ToS — no contract breach

**Key shift:** Google abandoned CFAA arguments (neutered for public data) for **DMCA anti-circumvention claims** — targeting SearchGuard bypass specifically.

## US vs. EU

### United States

- Public data scraping generally legal (CFAA precedent)
- ToS violations = contract, not criminal
- CAN-SPAM for outreach from scraped data

### European Union

- GDPR: public availability does NOT equal lawful basis (Article 6)
- Names, phones, reviewer profiles = personal data
- Need legitimate interest (Art. 6(1)(f)) for B2B
- Must provide opt-out, honor right to be forgotten
- Penalties: up to 20M EUR or 4% global turnover
- EU AI Act enforcement: August 2026

## Risk by Method

Method	Legal Risk	TOS Violation	DMCA Exposure
Official Places API	None	No	No
Data marketplace purchase	Low	No (you didn't scrape)	No
Commercial platforms	Medium	Yes	Indirect
Open-source scrapers	Medium	Yes	Low

SERP API proxies	High	Yes	Active lawsuit
Reverse-engineered APIs	Highest	Yes	Circumvention

# Enforcement Reality

Survey of 40-50 agencies scraping Maps at scale: **zero cease-and-desist letters**. Google relies on technical countermeasures for most scrapers, reserving legal action for large commercial operations (SerpAPI).

## Sources

- [IPWatchdog — Google Sues SerpAPI](#)
- [SerpAPI Motion to Dismiss](#)
- [Is Scraping Google Maps Legal?](#)