

## Top 3 APIs for n8n + Scraping/Search

### 1. Apify

- Integration with n8n: n8n has a built-in “Apify” node. You can run Apify Actors (scrapers), fetch datasets, or scrape a single URL.
- What it's good at:
  - Full browser-based scraping (Playwright / Puppeteer) or lightweight HTTP client → handle complex, JS-heavy e-commerce sites.
  - Use “Actors” (pre-built scrapers) or build custom ones to extract product details, reviews, listings, etc.
  - Automate workflows: schedule scrapes, chain scrapers, trigger other n8n nodes when data is ready.
- Why choose it: Very flexible, scalable, and well-integrated into n8n for more advanced / large-scale scraping tasks.

### 2. ScraperAPI

- Integration with n8n: There is an official / community node for ScraperAPI in n8n.
- What it's good at:
  - Simple “send URL, get back HTML (or parsed response)” model.
  - Handles proxy rotation, CAPTCHAs, rendering of JS automatically.
  - You can configure device type (desktop/mobile), geo-targeting, and level of proxy / rendering.
  - Supports integration with n8n's MCP client for real-time AI agent use.
- Why choose it: Lightweight and reliable for “on-demand scraping” from workflows. If your agent just needs to fetch product pages and extract info, this is easy to plug in.

### 3. ScrapingBee

- Integration with n8n: There is a verified ScrapingBee integration in n8n.
- What it's good at:
  - Web scraping API that supports JavaScript rendering, rotating proxies, and CAPTCHAs.
  - Good for extracting structured data (e.g., product details from Amazon, Alibaba) because you can reliably fetch fully rendered HTML.
  - Their Amazon-specific endpoint helps with product APIs (ScrapingBee has Amazon Search / Amazon Product API).
- Why choose it: Balanced between power and simplicity: more robust than plain HTTP + proxies, but lighter than full-blown browser automation.

## **Bonus Option to Consider**

- Scrapfly: They have an official n8n integration. Their Web Scraping API supports browser rendering, anti-bot, and they also have an Extraction API for parsing product data directly.
- Bright Data: Also supported in n8n. You can use their HTTP API or MCP integration to do large-scale scraping.
- ScrapeNinja: Has an n8n community node for browser + JS scraping, extracting content via custom JS.

## **My Recommendation (for AI Agents that Search + Scrape)**

- Use Apify for your core scraping: run actors to get product pages, crawl, extract structured product data — fits well in an “agent picks a URL → scrape → return JSON → agent reasons.”
- Use ScraperAPI for lighter “on-the-fly” scraping from agents when they need a quick look at a page or search result.

- Use ScrapingBee when you specifically want product data from e-commerce sites and need JS rendering + stable proxy but don't need full actor control.