

PRODUCT OVERVIEW · LLMSTXTGENERATOR.DE

LLMs.txt Generator

AI Visibility for Online Shops · v4.1 · June 2026

The most comprehensive free tool for generating llms.txt files. Built on tests with 450+ online shops across different CMS systems and industries.

10,000 URLs/analysis (max)	7 analysis phases	20× parallel fetches	450+ shops tested	26+ intent routings	5 output formats
--------------------------------------	-----------------------------	--------------------------------	-----------------------------	-------------------------------	----------------------------

7-Phase Analysis Pipeline

Every URL automatically passes through 7 sequential phases — no manual configuration required.

1
PHASE 1

PHASE 1 · DISCOVERY
robots.txt → Sitemap → Fallback Crawl

Reads robots.txt, loads XML and gzip sitemaps, resolves index sitemaps recursively. Without a sitemap: structured homepage crawl.

✓ gzip .xml.gz · ✓ index recursion · ✓ disallow filter · ✓ up to 10,000 URLs
Gambio · WooCommerce · Shopware · Shopify · JTL · ePages · Magento · TYPO3 · Joomla

robots.txt XML Sitemap gzip Recursion

2
PHASE 2

PHASE 2 · INVENTORY
URL Normalization & Deduplication

Normalizes URLs, removes duplicates, filters out system URLs (admin, checkout, API).

Example shop: 685 URLs → 612 products + 61 categories + 12 info pages
Filter: /admin/ · /checkout/ · /api/ · /login/

Normalization Dedup System filter

3
PHASE 3

PHASE 3 · CLASSIFICATION
5-Layer Classification (autoCat)

Schema.org → URL pattern → manufacturer page → BRAND_WORDS (40+) → structure analysis. Nav-confirmed categories permanently immunized.

1. **Schema.org** Product · CollectionPage · Brand
2. **URL pattern** Gambio .html · WooCommerce /product/
3–5. **Manufacturer** · BRAND_WORDS (40+) · path depth × nav

Schema.org 5 layers Nav immunity 40+ patterns

4

PHASE 4

PHASE 4 · ENRICHMENT

Batch Fetch: 20 URLs in Parallel

Extracts title, description, OG tags, Schema.org JSON-LD, prices, currency, brand, H1-H3 headings. Soft-404 via HTML analysis.

Batch: 20 URLs/call · avg 0.38s/batch

Extraction: title · desc · schema · price · currency · brand · headings

Parallel fetch
JSON-LD
Price data
Soft-404

5

PHASE 5

PHASE 5 · RECLASSIFICATION

Metadata Refinement (reClassify)

Price patterns in titles identify further products. Brands and nav categories stay permanently protected.

Patterns: € · buy · cart · incl. VAT · in stock

WooCommerce: depth≥2 without /product/ = always category

Price detection
Nav protection
Brand immunity

6

PHASE 6

PHASE 6 · FEATURED

Offer Detection & Priority Scoring

specials.php, products_new.php, /sale/ are crawled. Products get featured=true + priority 0.95.

Sources: specials.php · products_new.php · /sale/ · /offers/

Scoring: 0.95 offers · 0.90 new · 0.85 regular

Offers
New products
Priority scoring

7

PHASE 7

PHASE 7 · GENERATION

5 Output Formats Simultaneously

Smart Titles, 26+ industry-specific intent routings, date header. ZIP archive for download.

Smart Titles: Schema.org → URL slug → path fallback

26+ intents: automatically inserted by detected industry

llms.txt
JSON
YAML
robots.txt
ZIP

5 Output Formats

llms.txt	Standard for ChatGPT, Claude & Perplexity. Compact, directly citable.
llms-full.txt	Extended format with full text and structured product data.
llms-data.json	Machine-readable JSON for APIs, cron jobs and further processing.
llms-meta.yaml	YAML for developers, CI/CD pipelines and technical documentation.
robots-llms.txt	Ready-made snippet to add to your existing robots.txt.

Feature Overview

FEATURE	DESCRIPTION	TYPE
Product Drill-Down	Schema.org prices, brands, availability in llms.txt	Unique
26+ Intent Routings	Detect purchase intents & route automatically	Unique
CMS Detection	Gambio, WooCommerce, Shopware, Shopify, JTL, ePages, Magento automatic	Automatic
Recommended Preset	Sale and new products automatically prioritized	Standard
WordPress Plugin	Free on WordPress.org · WordPress 7.0 compatible	Live
AI Visibility Score	0-100 score with 5 sub-scores (Schema, llms.txt, crawlers ...)	New
Shopify App	Native integration · Embedded App (Pro Business)	Soon
Always Free	Core features free, no API key required	Always
GDPR Compliant	German servers, no permanent data storage	Compliant
5 Output Formats	llms.txt, Full, JSON, YAML, robots.txt in one run	Complete
Save Selection	JSON export for repeated generation / cron jobs	Pro

llmstxtgenerator.de/app · Get started free