

Optimale Einstellungen für den LLMs.txt Generator

Erweiterte Einstellungen - Detaillierte Empfehlungen

Die erweiterten Einstellungen im LLMs.txt Generator ermöglichen es Ihnen, die Qualität und den Umfang der generierten Dateien präzise zu steuern. Hier sind die optimalen Einstellungen für verschiedene Anwendungsfälle:

Die drei Haupteinstellungen im Detail

1. Inhalte extrahieren

Verfügbare Optionen:

-  "Ja - Echte Inhalte extrahieren" (Empfohlen)
-  "Nein - Nur Metadaten verwenden"

Empfehlung: "Ja - Echte Inhalte extrahieren"

Warum diese Einstellung optimal ist:



- **Vollständige KI-Optimierung:** LLMs erhalten den tatsächlichen Seiteninhalt, nicht nur Titel und Beschreibungen
- **Bessere Kontextverständnis:** ChatGPT, Claude & Co. können präzisere Antworten über Ihre Website geben
- **Höhere Relevanz:** Echte Produktbeschreibungen, Preise und Details werden erfasst
- **Professionelle Ergebnisse:** Die generierten Markdown-Dateien enthalten verwertbare Inhalte

Wann "Nein" wählen:

- Bei sehr großen Websites (>100 URLs) für schnellere Verarbeitung
 - Wenn nur eine grundlegende Seitenstruktur-Übersicht benötigt wird
 - Bei Websites mit viel dynamischem Content (JavaScript-basiert)
-

2. Bilder einbeziehen

Verfügbare Optionen:

-  **"Ja"** (Empfohlen für E-Commerce)
-  **"Nein"** (Empfohlen für Text-basierte Sites)

Empfehlung: "Ja" für E-Commerce, "Nein" für reine Informationsseiten

"Ja" wählen bei:

- **Online-Shops:** Produktbilder sind essentiell für KI-Verständnis
- **Portfolio-Websites:** Visuelle Inhalte sind zentral
- **Restaurants/Hotels:** Bilder vermitteln wichtige Informationen
- **Handwerks-/Dienstleistungsseiten:** Vorher-Nachher-Bilder, Referenzen

"Nein" wählen bei:

- **Blog/News-Websites:** Text steht im Vordergrund
- **Dokumentations-Seiten:** Technische Inhalte ohne visuelle Komponente
- **Rechtliche/Beratungsseiten:** Fokus auf Textinhalte
- **Performance-kritischen Anwendungen:** Reduziert Dateigröße erheblich

3. Max. Inhaltslänge (Zeichen)

Standard-Wert: 2000 Zeichen

Empfohlene Bereiche:

Optimale Werte nach Anwendungsfall:

Kleine bis mittlere Websites (1-20 Seiten):

- **Empfehlung:** 5000-8000 Zeichen
- **Begründung:** Ausführliche Beschreibungen möglich, überschaubare Gesamtgröße

Große E-Commerce-Sites (20-100 Seiten):

- **Empfehlung:** 3000-4000 Zeichen
- **Begründung:** Balance zwischen Detail und Übersichtlichkeit

Sehr große Websites (100+ Seiten):

- **Empfehlung:** 1500-2500 Zeichen
- **Begründung:** Verhindert übermäßig große Dateien, behält Kerninfos

Blog/News-Websites:

- **Empfehlung:** 4000-6000 Zeichen

- **Begründung:** Artikel-Abstracts und Einleitungen vollständig erfassen

Empfohlene Konfigurationen nach Website-Typ

E-Commerce/Online-Shop (wie Solar-Autark.com)

Plain Text

- ✓ Inhalte extrahieren: Ja - Echte Inhalte extrahieren
- ✓ Bilder einbeziehen: Ja
- 📏 Max. Inhaltslänge: 3500 Zeichen

Begründung: Produktdetails, Preise und Bilder sind essentiell für KI-Verständnis

Blog/News-Website

Plain Text

- ✓ Inhalte extrahieren: Ja - Echte Inhalte extrahieren
- ✗ Bilder einbeziehen: Nein
- 📏 Max. Inhaltslänge: 5000 Zeichen

Begründung: Fokus auf Textinhalte, längere Artikel-Abstracts

Unternehmens-/Dienstleistungswebsite

Plain Text

- ✓ Inhalte extrahieren: Ja - Echte Inhalte extrahieren
- ✓ Bilder einbeziehen: Ja
- 📏 Max. Inhaltslänge: 4000 Zeichen

Begründung: Ausgewogene Darstellung von Services und visuellen Referenzen

Dokumentations-/Wissensseite

Plain Text

- ✓ Inhalte extrahieren: Ja - Echte Inhalte extrahieren
- ✗ Bilder einbeziehen: Nein
- 📏 Max. Inhaltslänge: 6000 Zeichen

Begründung: Detaillierte Textinformationen ohne visuelle Ablenkung

Portfolio/Kreativ-Website

Plain Text

- ✓ Inhalte extrahieren: Ja - Echte Inhalte extrahieren
- ✓ Bilder einbeziehen: Ja
- 📏 Max. Inhaltslänge: 2500 Zeichen

Begründung: Bilder sind zentral, kurze aber prägnante Beschreibungen

Profi-Tipps für optimale Ergebnisse

Performance-Optimierung:

- **Große Websites:** Starten Sie mit niedrigeren Werten und erhöhen Sie schrittweise
- **Test-Durchläufe:** Generieren Sie zunächst 5-10 URLs zum Testen der Einstellungen
- **Monitoring:** Beobachten Sie die Dateigröße der generierten llms-ctx-full.txt

Qualitäts-Optimierung:

- **Konsistenz:** Verwenden Sie einheitliche Einstellungen für alle URLs einer Website
- **Relevanz:** Höhere Zeichenlimits für wichtige Seiten (Startseite, Hauptkategorien)
- **Balance:** Zu kurze Limits (unter 1000) können wichtige Informationen abschneiden

Workflow-Optimierung:

- **Speichern:** Nutzen Sie die "Speichern"-Funktion für wiederkehrende Projekte
- **Templates:** Erstellen Sie Vorlagen für verschiedene Website-Typen
- **Versionierung:** Dokumentieren Sie erfolgreiche Einstellungen für zukünftige Projekte

Auswirkungen der Einstellungen auf die Dateigröße

Beispiel-Kalkulation für 50 URLs:

Minimale Einstellungen:

- Inhalte: Nein, Bilder: Nein, Länge: 1000

- **Geschätzte Größe:** ~15-25 KB

Empfohlene Einstellungen:

- Inhalte: Ja, Bilder: Ja, Länge: 3500
- **Geschätzte Größe:** ~150-300 KB

Maximale Einstellungen:

- Inhalte: Ja, Bilder: Ja, Länge: 8000
 - **Geschätzte Größe:** ~400-800 KB
-

Häufige Fehler vermeiden

Zu vermeiden:

- **Zu hohe Limits bei großen Sites:** Kann zu mehreren MB großen Dateien führen
- **Bilder bei reinen Text-Sites:** Unnötig aufgeblähte Dateien
- **Zu niedrige Limits:** Wichtige Informationen werden abgeschnitten
- **Inkonsistente Einstellungen:** Verschiedene Limits für ähnliche Seiten




Best Practices:

- **Schrittweise Anpassung:** Beginnen Sie konservativ und erhöhen Sie bei Bedarf
 - **Zielgruppen-orientiert:** Denken Sie an die KI-Assistenten, die die Dateien nutzen werden
 - **Regelmäßige Überprüfung:** Testen Sie die generierten Dateien mit ChatGPT/Claude
 - **Dokumentation:** Notieren Sie erfolgreiche Einstellungen für verschiedene Projekte
-

Fazit: Die goldene Regel

Für die meisten Websites ist diese Konfiguration optimal:

Plain Text

-  Inhalte extrahieren: Ja - Echte Inhalte extrahieren
-  Bilder einbeziehen: Ja (E-Commerce) / Nein (Text-Sites)
-  Max. Inhaltslänge: 3500 Zeichen

Diese Einstellungen bieten die beste Balance zwischen Informationsgehalt, Dateigröße und

KI-Verständlichkeit für die meisten Anwendungsfälle.