

FAQ - Schnellhilfe für LibreChat

← Zurück zur LibreChat-Menü



Ziel dieser Seite: Diese FAQ beantwortet typische Nutzerfragen. Für Schritt-für-Schritt-Anleitungen siehe [Erweiterte Funktionen nutzen](#), für UI-Details [Benutzeroberfläche verstehen](#), für Modellauswahl und Parameter [Modellauswahl & Optimierung](#), für Sprachthemen [Sprach-Ein- und Ausgabe](#), für temporäre Chats [Temporäre Chats nutzen](#). Detaillierte Modelllisten finden Sie im Endpoints-Menü der LibreChat-Oberfläche oder im [GWDG-Katalog](<https://docs.hpc.gwdg.de/services/chat-ai/models/index.html>).

1. Schnellstart & Modellwahl

Welche Option ist mein Standard?

- **Text, E-Mails, schnelle Fragen** → `gpt-4.1-mini` (OpenAI) oder `meta-llama-3.1-8b-instruct` (GWDG, kostenfrei).
- **Technische & wissenschaftliche Inhalte** → `qwen3-30b-a3b-instruct-2507` (GWDG).
- **Ich bin unsicher** → wählen Sie einen passenden Agenten (z.B. **DeepWiki Helper**, **GPT-4.1 web search**), dort sind Modelle und Parameter bereits eingestellt.

Wann brauche ich GPT-5 oder Reasoning-Modelle?

Nur bei mehrstufigen Analysen, komplexen Logikproblemen, tiefem Debugging oder wenn Sie explizit Reasoning-Presets wählen. Für Routineaufgaben kostet GPT-5 nur Zeit und Ressourcen. Die wichtigsten Reasoning-Fragen finden Sie weiter unten.

Wo finde ich vollständige Modelltabellen?

Im Endpoints-Menü der LibreChat-Oberfläche (Modellauswahl → Endpoints) oder im [GWDG-Katalog](#). Diese Listen werden automatisch aktualisiert und müssen nicht manuell gepflegt werden.

2. Agenten & Ressourcen

Welche Agenten sollte ich mir merken?

- **DeepWiki Helper** – technische Dokumentationen.
- **GPT-4.1/GPT-5-mini web search** – aktuelle Informationen, Live-Websuche.
- **Math & Logic / Code Debugging / Deep Analysis** – bereitgestellte Reasoning-Presets für komplexe Fälle.

Alle Agenten erreichen Sie über die Symbolleiste → Vier-Quadrate-Symbol, und zwar Agenten-Übersicht in [Benutzeroberfläche](#).

Wer trägt die Kosten?

- **GWDG-Modelle**: kostenfrei und in der EU gehostet.
- **OpenAI-Modelle**: Kosten übernimmt die TU Clausthal; bitte dennoch verantwortungsvoll einsetzen.
- **Empfehlung**: einfache Aufgaben → GWDG oder `gpt-4.1-mini`, komplexe Projekte → GPT-5/o3/o4.

Muss ich Parameter selbst einstellen?

Nein. Standardnutzer bleiben bei den Voreinstellungen oder wählen einen Agenten. Nur wer bewusst Feintuning möchte, öffnet das rechte Seitenmenü → **KI-Einstellungen** (siehe [Modellauswahl & Optimierung](#)).

3. Reasoning-Modelle: Häufige Fragen

Warum antworten Reasoning-Modelle so langsam?

Sie erzeugen vor der Ausgabe eine interne Gedankenkette. Das kostet Zeit und Tokens, liefert aber verlässliche, nachvollziehbare Ergebnisse. Für kurze Fragen ist ein leichtes Modell besser geeignet.

Muss ich vor komplexen Aufgaben viele Parameter setzen?

In der Regel nein. GPT-5/GPT-5.1 entscheiden automatisch, wie viel Reasoning nötig ist, und unsere Agenten/Preset-Buttons stellen `Reasoning Effort`, `Reasoning Summary` und `Verbosity` bereits ein. Nur bei Spezialfällen lohnt sich ein Blick in [Modellauswahl & Optimierung](#).

Welches Modell oder welcher Agent ist die beste Wahl?

- **Schnelle Hilfe / Routine** → `gpt-4.1-mini`, Llama/Qwen (GWDG).
- **Komplexe Logik, Mathe, Debugging** → Reasoning-Agenten oder o4-mini, qwen3-30b-a3b-thinking-2507, deepseek-r1.
- **Unsicher?** → Agent starten; wenn zu langsam, Effort auf **medium/low** senken oder zurück zu einem Standardmodell wechseln.

Wie gehe ich mit langen Wartezeiten um?

- Prüfen, ob wirklich Reasoning nötig ist.
- `Reasoning Effort` auf **medium** oder **low** setzen, `Reasoning Summary` auf **Concise**.
- Aufgabe klar strukturieren (Dateien hochladen, Schritt-für-Schritt formulieren).
- Erst leichtes Modell fragen, danach Reasoning-Agent für die finale Analyse verwenden.

4. Dateien & erweiterte Funktionen

Welche Dateitypen kann ich hochladen?

Bilder (JPG, PNG, GIF, WebP), Dokumente (PDF, TXT, DOC/DOCX), Code (py, js, html, css). Bei Unsicherheit einfach ausprobieren.

Wie groß dürfen Dateien sein?

Größenlimits werden über das **File Token Limit** gesteuert (siehe [Modellauswahl & Optimierung](#)). Große Dokumente am besten aufteilen.

Werden meine Dateien dauerhaft gespeichert?

- **Text mit OCR**: nur innerhalb des Chats.
- **Dateisuche**: bleibt gespeichert, bis Sie sie löschen.
- **Bild hochladen**: temporär.

Verwaltung und Löschung: rechtes Seitenmenü → Dateien anhängen; Details in [Nutzungsanleitung](#).

Kann ich mehrere Dateien gleichzeitig hochladen?

Ja, alle Methoden unterstützen Mehrfachauswahl.

Funktionieren Datei-Uploads in temporären Chats?

Ja, allerdings werden Dateien nicht dauerhaft gespeichert, sobald der temporäre Chat endet (siehe [temporäre Chats](#)).

Mehr Schritt-für-Schritt-Anleitungen finden Sie in der [Nutzungsanleitung](#).

Temporäre Chats - häufige Fragen

- **Was passiert nach einem Refresh?** Der Chat ist unwiederbringlich verloren; exportieren oder kopieren Sie wichtige Inhalte vorher.
- **Kann ich mehrere temporäre Chats parallel führen?** Ja, in getrennten Browser-Tabs/Fenstern.
- **Wie exportiere ich einen temporären Chat?** Über den Export/Share-Button oben rechts (z. B. HTML oder Markdown).
- **Kann ich einen normalen Chat in einen temporären Chat umwandeln?** Nein. Stattdessen normalen Chat löschen.

5. Sprachfunktionen: Häufige Probleme

Mikrofon funktioniert nicht

- Browser-Berechtigungen prüfen, Mikrofon-Anschluss checken, ggf. Browser aktualisieren oder externe Engine (OpenAI Whisper) wählen.

Spracherkennung ist ungenau

- Langsamer sprechen, Hintergrundgeräusche minimieren, passendes Eingabegerät wählen oder zur externen Engine wechseln.

Keine Sprachausgabe / Stimme klingt künstlich

- Prüfen, ob TTS aktiviert ist und der Ton nicht stumm geschaltet wurde.
- Andere Stimme oder Engine (OpenAI TTS) testen.
- Bei roboterhaftem Klang hilft meist der Wechsel zur externen Engine.

Browser fragt erneut nach Berechtigungen?

- „Erlauben“ anklicken; ohne Zustimmung funktioniert STT nicht.

- Sprachfunktionen laufen auch in [temporären Chats](#).

Mehr Details: [Sprach-Ein- und Ausgabe](#).

6. Ressourcen & Support

- **Kosten:** GWDG-Modelle sind kostenfrei; OpenAI-Kosten trägt die TU Clausthal. Nutzen Sie dennoch ressourcenschonend.
- **Modellwechsel:** jederzeit über die Modellauswahl möglich (siehe [Benutzeroberfläche](#)).
- **Kontakt:** Bei technischen Problemen wenden Sie sich an support@rz.tu-clausthal.de.

Weitere Informationen

- [Anwendungsbeispiele - Erste Schritte mit LibreChat](#)
- [Benutzeroberfläche von LibreChat](#)
- [FAQ – Schnellhilfe für LibreChat](#)
- [Modellauswahl & Optimierung](#)
- [Erste Schnitte](#)
- [Erweiterte Nutzungsanleitung](#)
- [Sprach-Ein- und Ausgabe in LibreChat](#)
- [Temporäre Chats](#)

[← Zurück zur Startseite](#)

Direkt-Link:

https://doku.tu-clausthal.de/doku.php?id=sonstige_dienste:ki-dienste:librechat:fag

Letzte Aktualisierung: **09:40 03. December 2025**

