

API Dokumentation

Übersicht

Diese API bietet Funktionen zur Verwaltung von Modellen und zur Interaktion mit ihnen, einschließlich der Generierung von Antworten, des Ladens von Modellen und der Anzeige von Modellinformationen.

Inhaltsverzeichnis

- 1. [Endpunkte](#)
 - [Serverstatus abrufen](#)
 - [Antwort generieren](#)
 - [Modelle auflisten](#)
 - [Laufende Modelle anzeigen](#)
 - [Modell laden](#)
- 2. [Antwortformate](#)
- 3. [Fehlerbehandlung](#)

Endpunkte

Serverstatus abrufen

Beschreibung: Überprüft, ob der Server aktiv ist und gibt grundlegende Informationen zurück.

- **Endpunkt:** `/api/ping`
- **Methode:** GET
- **Antwort:**

```
{
  "serverAddress": "<serverIp>:<serverPort>",
  "clientIp": "<clientIp>",
  "currentTime": "<currentTime>"
}
```

Antwort generieren

Beschreibung: Generiert eine Antwort basierend auf einem gegebenen Text-Prompt.

- **Endpunkt:** `/api/generateResponse`
- **Methode:** POST
- **Request-Body:**

```
{
  "prompt": "<Ihr Prompt>"
}
```

- **Antwort:**

```
{
  "response": "<Generierte Antwort>"
}
```

Modelle auflisten

Beschreibung: Gibt eine Liste aller verfügbaren Modelle zurück.

- **Endpunkt:** `/api/listModels`
- **Methode:** GET
- **Antwort:**

```
{
  "models": ["model1", "model2", "model3"]
}
```

Laufende Modelle anzeigen

Beschreibung: Zeigt alle derzeit auf dem Server laufenden Modelle an.

- **Endpunkt:** /api/listRunningModels
- **Methode:** GET
- **Antwort:**

```
{
  "runningModels": [
    { "modelName": "model1" },
    { "modelName": "model2" }
  ]
}
```

Modell laden

Beschreibung: Lädt ein bestimmtes Modell auf den Server.

- **Endpunkt:** /api/loadModel
- **Methode:** POST
- **Request-Body:**

```
{
  "modelName": "<Modellname>"
}
```

- **Antwort:**

```
{
  "modelName": "<Modellname>",
  "loaded": true
}
```