



elektor **BONUS**

design > share > earn

Freigeschaltet

606B
SEIT 1961

Woche 50

Jede Woche frische, kostenlose Bonus-Artikel!

Special
Gast-Ausgabe

Verfügbar in Woche 51

Verfügbar in Woche 51

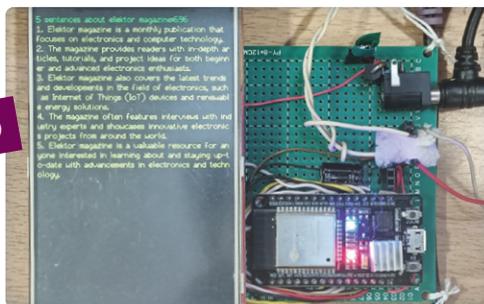
Verfügbar in Woche 1 (2025)



Leistungsstarke Edge-KI

Der neue STM32N6

Verfügbar in Woche 1 (2025)



Ein persönliches KI-Terminal
ChatGPT bekommt eine Stimme

Verfügbar in Woche 52

ASUS IoT PE8000G

REVOLUTIONÄRE EDGE-AI-LEISTUNG

Robust und zuverlässig

Der PE8000G erfüllt den MIL-STD-810H-Standard und ist damit gegen extreme Temperaturen, Vibrationen und Spannungsschwankungen geschützt – ideal für den Einsatz in rauen Industrieumgebungen.

Dual-GPU-Unterstützung

Verarbeiten Sie mehrere neuronale Netzwerke gleichzeitig dank der Unterstützung von zwei 450-Watt-GPUs – nahtlos und in Echtzeit. So können Sie auch komplexe Aufgaben effizient und schnell erledigen.



Vielfältige Einsatzmöglichkeiten

Von KI-getriebener Fabrikautomatisierung bis hin zu intelligenten Videoanalysen – der PE8000G liefert die Rechenleistung und Flexibilität, die Sie brauchen.

Flexible Stromversorgung

Mit einem DC-Eingang von 8–48V und integrierter Zündsteuerung ist der PE8000G für vielseitige industrielle Anforderungen bestens ausgestattet.

FORTEC
INTEGRATED

Interesse geweckt?





55. Jahrgang, Nr. 25001 (606)
 Bonus-Gastausgabe KI 2024
 ISSN 0932-5468

Das Elektor Magazin wird 8 Mal im Jahr herausgegeben von
Elektor Verlag GmbH
 Lukasstraße 1, 52070 Aachen (Deutschland)
 Tel. +49 (0)241 95509190
 www.elektor.de | www.elektormagazine.de

Chefredakteur: Jens Nickel (v.i.S.d.P.)

Für alle Ihre Fragen: service@elektor.de

Mitglied werden: www.elektormagazine.de/abo

Anzeigen: Büsra Kas
 Tel. +49 (0)241 95509178 – busra.kas@elektor.com
 www.elektormagazine.de/mediadaten

Urheberrecht
 © Elektor International Media b.v. 2024

Die in dieser Zeitschrift veröffentlichten Beiträge, insbesondere alle Aufsätze und Artikel sowie alle Entwürfe, Pläne, Zeichnungen einschließlich Platinen sind urheberrechtlich geschützt. Ihre auch teilweise Vervielfältigung und Verbreitung ist grundsätzlich nur mit vorheriger schriftlicher Zustimmung des Herausgebers gestattet. Die veröffentlichten Schaltungen können unter Patent- oder Gebrauchsmusterschutz stehen. Herstellen, Feilhalten, Inverkehrbringen und gewerblicher Gebrauch der Beiträge sind nur mit Zustimmung des Verlages und ggf. des Schutzrechtsinhabers zulässig. Nur der private Gebrauch ist frei. Bei den benutzten Warenbezeichnungen kann es sich um geschützte Warenzeichen handeln, die nur mit Zustimmung ihrer Inhaber warenzeichengemäß benutzt werden dürfen. Die geltenden gesetzlichen Bestimmungen hinsichtlich Bau, Erwerb und Betrieb von Sende- und Empfangseinrichtungen und der elektrischen Sicherheit sind unbedingt zu beachten. Eine Haftung des Herausgebers für die Richtigkeit und Brauchbarkeit der veröffentlichten Schaltungen und sonstigen Anordnungen sowie für die Richtigkeit des technischen Inhalts der veröffentlichten Aufsätze und sonstigen Beiträge ist ausgeschlossen.

DEUTSCHE
 FACHPRESSE

Elektor ist Mitglied des 1929 gegründeten VDZ (Verband Deutscher Zeitschriftenverleger), der „die gemeinsamen Interessen von 500 deutschen Consumer- und B2B-Verlagen vertritt.“



Mehr KI-Projekte und Artikel

Haben Sie wirklich geglaubt, dass wir mit KI-bezogenen Inhalten fertig sind, nachdem wir die Gast-Ausgabe KI 2024 von Elektor angekündigt hatten? Weit gefehlt! Wie regelmäßige Elektor-Leser wissen, haben wir immer neue Projekte, Tutorials und Hintergrundartikel in der Pipeline. Mit dieser Bonus-Ausgabe, die wir in den folgenden vier Wochen enthüllen werden, wollen wir Sie dazu anregen, in den kommenden Wochen und Monaten neue KI-bezogene Anwendungen zu entwickeln.

Wenn Sie auf der Suche nach einem Projekt sind, mit dem Sie mit KI experimentieren können, ist der Artikel „Universal-IR-Fernbedienung mit KI“ ein hervorragender Startpunkt für Sie! Mit einem Raspberry Pi 5, der Plattform *MediaPipe Studio* von Google und einer winzigen Plug-in-Schnittstellenplatte können Sie ein vielseitiges Gestenerkennungssystem zur Steuerung Ihres Fernsehers oder anderer Geräte mit IR-Fernbedienungen entwickeln.

Kann eine moderne Konstruktion den Geist von Alan Turings ursprünglichem Experiment einfangen? In „Ein persönliches KI-Terminal“ erfahren Sie, wie ein ESP32-Modul zusammen mit einem TFT-Display und einem I²S-Verstärker zu einer modernen Version des Turing-Tests des

21. Jahrhunderts wird, die direkt mit ChatGPT kommuniziert.

In „KI für das Produktkonzept-Design“ sehen wir uns an, wie KI die Bilderstellung revolutioniert hat. Entwickler und Designer können Skizzen und Ideen mit einer gut formulierten Textanweisung in fotorealistische Renderings verwandeln. Tauchen Sie ein in diese Erkundung von KI-gestützten Design-Tools.

Sind Sie neugierig auf die Geschichte der künstlichen Intelligenz auf den Seiten von Elektor? In dieser Ausgabe nehmen wir Sie mit auf eine Reise durch die Elektor-Archive und zeigen Ihnen, wie sich die künstliche Intelligenz in der Elektronikbranche entwickelt hat. Tauchen Sie ein in die Empfehlungen unserer Redakteure aus vergangenen Ausgaben Ihrer Lieblingszeitschrift! Und noch viel mehr erwartet Sie in dieser kostenlosen Bonusausgabe!

Nachdem Sie diese Ausgabe studiert und Ihre eigenen KI-Projekte in Angriff genommen haben, teilen Sie bitte online Ihre Erfahrungen auf der Plattform Elektor Labs: www.elektormagazine.com/labs. Wir freuen uns darauf, mehr über Ihre Kreationen zu erfahren!

C. J. Abate (Content Director, Elektor)

INHALT

4 Ein persönliches KI-Terminal
 ChatGPT bekommt eine Stimme



8 Leistungsstarke Edge-KI: Der neue STM32N6
 Der neue STM32N6 ist Maßstab für Leistungs- und KI-Fähigkeit



12 **Verfügbar in Woche 51**

16

22 **Verfügbar in Woche 52**

26

32

34 **Verfügbar in Woche 1 (2025)**

42



Die Gast-Ausgabe KI 2024 von Elektor ist am Kiosk und im Elektor-Store erhältlich.
www.elektor.de/dp-0606



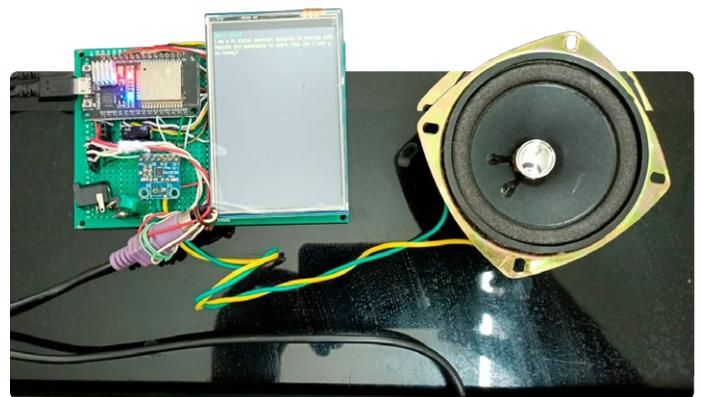
Ein persönliches KI-Terminal

ChatGPT bekommt eine Stimme

Von Somnath Bera (Indien)

Noch immer wird darüber diskutiert, ob KI-Lösungen wie ChatGPT den Turing-Test bestehen können. Alan Turing stellte sich einen menschlichen Operator vor, der die Antworten auf Fragen bewertet, die mit einem elektromechanischen Fernschreiber gesendet und empfangen werden. Hier bauen wir eine Version von Turings ursprünglichem experimentellem Konzept für das 21. Jahrhundert, indem wir ein ESP32-Board mit Tastatur und TFT-Bildschirm ausstatten, um ausschließlich mit ChatGPT über das Internet zu kommunizieren. Zusätzlich ermöglicht Google Text-to-Speech in Verbindung mit einem winzigen I²S-Verstärkermodul und einem Lautsprecher das Mithören der Konversation. In unserem Fall ist es natürlich von Anfang an klar, dass wir mit einer Maschine kommunizieren. Oder?

Es besteht kein Zweifel daran, dass KI-Tools wie ChatGPT von OpenAI und Gemini von Google in vielen Situationen echte Killerapplikationen sein können. Ich habe zum Beispiel mit ChatGPT ziemlich komplexe Steuerungslösungen entwickelt, wobei ich die anfängliche Idee lieferte, und wenn ich weitere Eingaben machte, verfeinerte ChatGPT den Code und machte ihn mit jeder Iteration besser. ChatGPT kann sogar Python-Code in MicroPython oder einen Arduino-Sketch umwandeln. Der Knackpunkt liegt darin, den Prozess sorgfältig zu steuern, damit er nicht zu weit vom Kurs abweicht. Manchmal kommt dies vor und ChatGPT wiederholt dieselben Fehler, aber es macht mir Spaß, diese Fehler zu finden und die Ausgabe von ChatGPT so zu steuern, dass sie meinem Ziel näher kommt.



Hardware

Herzstück dieses Projekts ist das ESP32-Entwicklungsmodul von Espressif. Seine Dual-Core-Architektur ist leistungsfähig genug, um die WLAN-Kommunikation zu übernehmen, die serielle Eingabe einer PS2-Tastatur zu verarbeiten, Daten an das 3,5-Zoll-TFT-Display zu senden und digitale Audiodaten an das I²S-Modul auszugeben. Der vollständige Schaltplan des ChatGPT-Terminals ist in **Bild 1** dargestellt. Sie werden sich vielleicht über die alte PS2-Tastatur wundern; der einzige Grund dafür ist, dass es mir nicht gelungen ist, für diesen Zweck einen USB-Anschluss auf dem ESP32 zu implementieren. In der begrenzten Zeit, die mir zur Verfügung stand, habe ich mich für die einfachere PS2-Schnittstelle entschieden, die weit weniger Ressourcen verbraucht. Einen PS2-Adapter für eine USB-Tastatur kann man problemlos im nächsten Elektronikladen kaufen. Die PS2-Pinbelegung lässt sich dem Schaltplan entnehmen. Sicherlich wird es Entwickler geben, die das System auf USB aufrüsten wollen und werden und ich freue mich über deren Beiträge.

Das verwendete 3,5-Zoll-TFT-Touch-Display hat statt eines SPI-Anschlusses eine parallele Schnittstelle. Dadurch werden zwar zwangsläufig mehr Verbindungen benötigt, aber in dieser Anwendung haben wir genügend freie GPIOs und eine parallele Schnittstelle bietet eine deutlich schnellere Reaktionszeit. Die Touchscreen-Funktionen des Displays werden in dieser Anwendung nicht genutzt.

Ein I²S-Verstärkermodul MAX98357A (**Bild 2**) verwandelt das vom ESP32 erzeugte digitale I²S-Audiosignal in analoges Audio. Der integrierte Klasse-D-Verstärker verstärkt das Signal auf einen ausreichenden Pegel von 3 W, um einen 4-Ω-Lautsprecher anzuweisen. Der GAIN-Eingang des Verstärkermoduls ist mit Masse verbunden, so dass die maximale Verstärkung (= Lautstärke) erreicht wird.



Bild 2. Das 3-W-Audioverstärkermodul wandelt das digitale I²S- in ein kräftiges analoges Audiosignal um.

Verlieren Sie den Schlüssel nicht!

ChatGPT von OpenAI erzeugt textbasierte Antworten auf Aufforderungen oder Fragen, die vom Benutzer eingegeben werden. Normalerweise funktioniert dies über eine Browserschnittstelle. Für Entwickler und Unternehmen wurde die OpenAI-API entwickelt, damit die KI-Funktionen in eigene Anwendungen, Software oder Websites integriert werden können. Es handelt sich nicht nur um eine Chatbot-Schnittstelle, sondern um eine flexible Programmierschnittstelle, die es Unternehmen

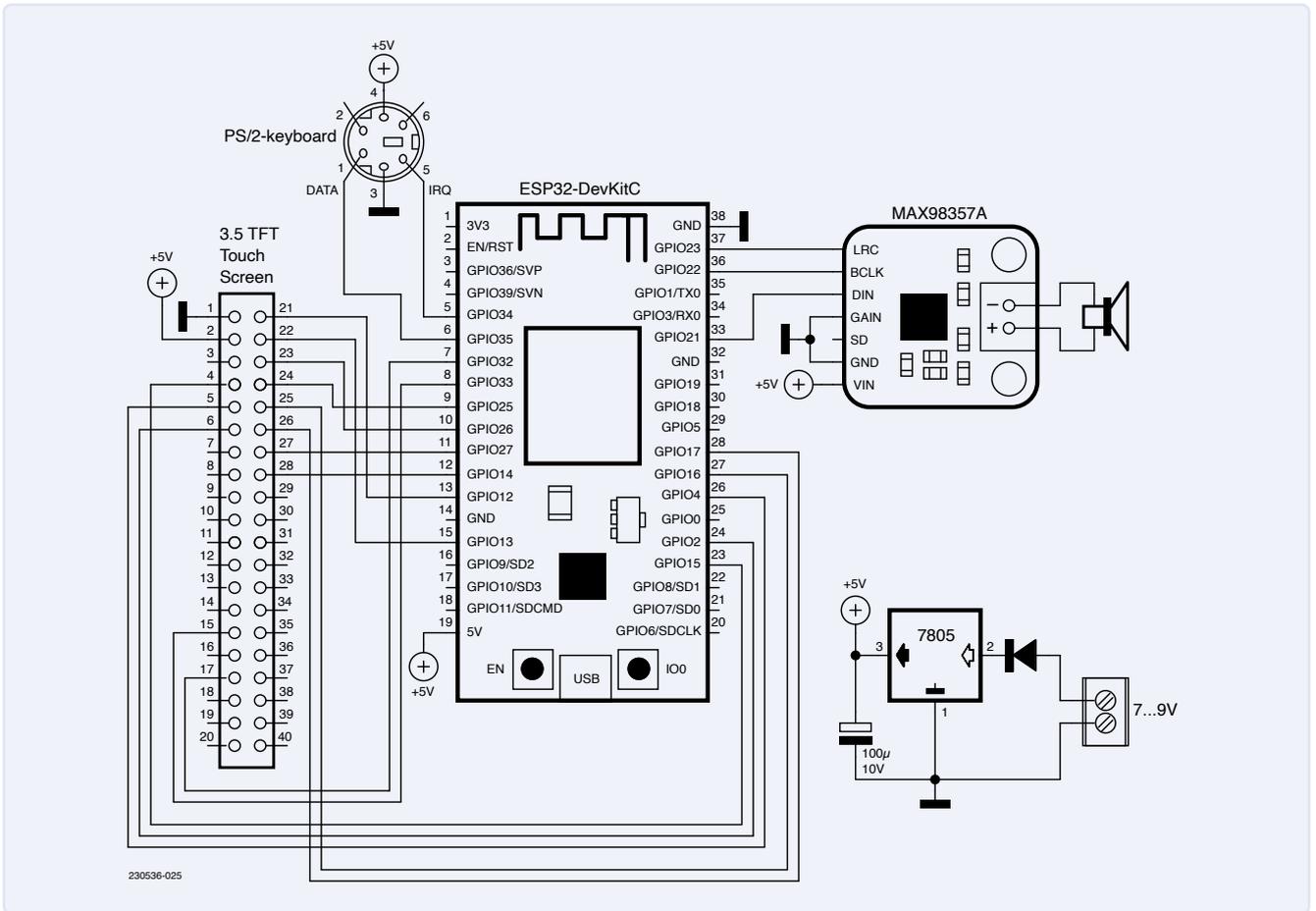


Bild 1. Schaltbild des sprechenden ChatGPT-Terminals.

ermöglicht, die Fähigkeiten von ChatGPT in ihre Produkte einzubinden. Entwickler senden programmgesteuert Daten an die API und erhalten Antworten zurück. Um die OpenAI-API zu nutzen, benötigt man einen eindeutigen Schlüssel, der für Ihr Konto generiert wird und mit dem Sie sich anmelden. Rufen Sie dazu die OpenAI-Website [1] auf und klicken Sie auf *Sign up*.

Füllen Sie das Anmeldeformular mit Ihrer E-Mail-Adresse, Ihrem Passwort und anderen erforderlichen Informationen aus. Sobald Sie diese Informationen eingegeben haben, navigieren Sie zum Dashboard Ihres OpenAI-Kontos und klicken Sie auf die *New Project*. Geben Sie Ihrem Projekt einen Namen und (optional) eine Beschreibung. Klicken Sie in Ihren Projekteinstellungen auf den Tab *API Keys*. Es wird eine Liste der vorhandenen geheimen Schlüssel angezeigt (Bild 3). Klicken Sie auf die Schaltfläche *Create new secret key*, um einen neuen

Schlüssel zu erzeugen. Erstellen Sie eine Kopie dieses API-Schlüssels und speichern Sie ihn sicher (ganz altmodisch auch auf einem Zettel), da Sie ihn aus Sicherheitsgründen nicht wieder abrufen können. Sie benötigen diesen Schlüssel, um Ihre Anwendungen bei den Diensten von OpenAI zu authentifizieren. Auf der Website von OpenAI gibt es ein Schnellstart-Tutorial für Entwickler, das Sie durch den Prozess der Generierung einer Antwort über die API unter Verwendung des generierten Schlüssels zur Autorisierung Ihres Zugangs führt. Als dieser Artikel geschrieben wurde, stellte OpenAI bei der ersten Anmeldung kostenlose „Credits“ zur Verfügung, mit denen Sie die API ausprobieren können. Wenn diese Credits aufgebraucht sind, müssen Sie für die Nutzung bezahlen. Wenn Sie noch keine Zahlungsmethode eingerichtet haben, müssen Sie dies tun. Lesen Sie dazu die OpenAI-Nutzungsrichtlinien und -bedingungen.

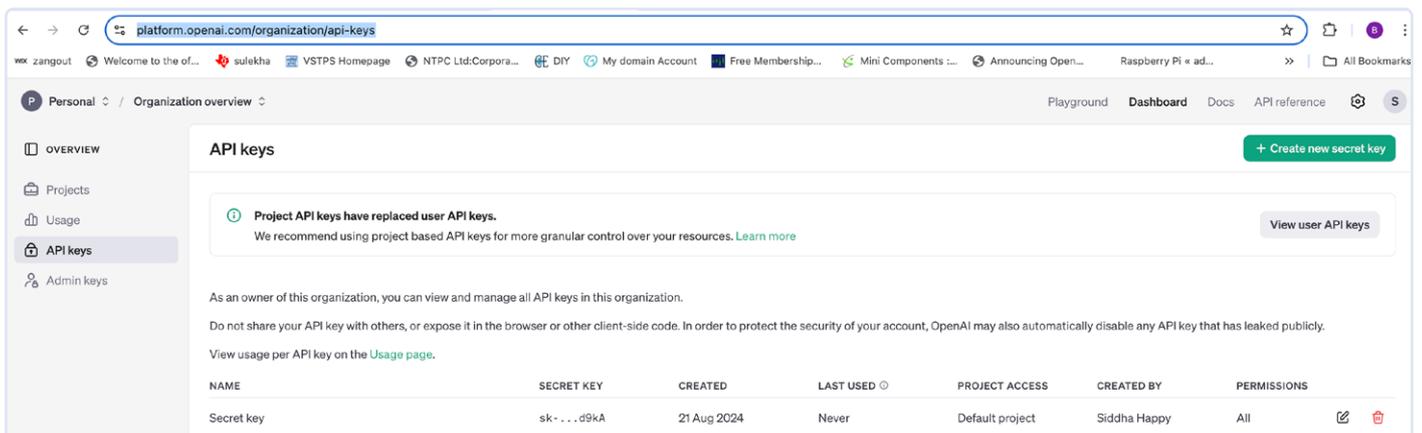


Bild 3. Die OpenAI-Webseite zur Registrierung für die OpenAI-API.

Text to Speech

Eine Text-to-Speech-API (TTS) wird verwendet, um die Textantwort von OpenAI in einen digitalen Audiodatenstrom zu konvertieren. Es gibt eine Reihe von Text-to-Speech-APIs, die wir dafür verwenden können. OpenAI verfügt über eine eigene TTS-API, die eine Reihe verschiedener Sprachalternativen bietet, die recht natürlich klingen. Um auf diese API zuzugreifen, verwenden Sie denselben Schlüssel, den Sie für die Nutzung der OpenAI-API erhalten haben.

Für dieses Projekt nutzen wir die Text-to-Speech-API der Google Cloud. Sie bietet eine breite Palette von Stimmen in verschiedenen Sprachen und Dialekten. Im Vergleich zu OpenAI-TTS klingen die Stimmen zwar ein wenig mechanischer, und bei langen Textstrings bricht die Ausgabe ab, aber die Nutzung dieser API ist zur Zeit kostenlos, ganz im Gegensatz zur API von OpenAI.

Um Google-TTS zu verwenden, müssen wir zunächst ein Projekt auf Google Cloud erstellen und die Google-TTS-API aktivieren, um einen API-Schlüssel zu erhalten. Die Textstrings können nun mit einer HTTP-POST-Anfrage zusammen mit dem Schlüssel an die API gesendet werden. Der resultierende digitale I²S-Audiostrom wird dann gespeichert und wiedergegeben, um über den MAX98357A ein analoges Audiosignal zu erzeugen.

Software: Bibliotheken

Der Arduino-Sketch ist unter [2] abrufbar. Überprüfen Sie, ob alle Bibliotheken in Ihrer Entwicklungsumgebung installiert sind, auf die im Header des Sketches verwiesen wird. Falls dies nicht der Fall sein sollte, installieren Sie sie jetzt mit dem Bibliotheksmanager der Arduino-IDE. Mit den Methoden der Bibliothek *Audio.h* war es eine einfache Aufgabe, die Audioausgabe zu erzeugen, die die Textausgabe auf dem TFT-Display begleitet. Es war nur notwendig, ein paar Zeilen in der Schleife hinzuzufügen, um den Ton zu erzeugen. Schauen Sie sich das im Arduino-Sketch an:

```
#include <PS2Keyboard.h> //Keyboard specific
#include <WiFi.h>
#include <HTTPClient.h>
#include <ArduinoJson.h>
#include <SPI.h>
#include <TFT_eSPI.h> // Hardware-specific library
#include "Audio.h" //Audio header file

// GPIO connections to I2S board
#define I2S_D0UT 21
#define I2S_BCLK 22
#define I2S_LRC 23
```

```
Audio audio; //audio instance creation
```

ArduinoJSON.h parst die JSON-formatierten Antwortdaten der OpenAI-API in ein Format, das vom Arduino-Code verwendet werden kann.

Software: ChatGPT-API-Anfrage

Die Interaktion mit der OpenAI-API ist in der Funktion `makeApiRequest(String prompt1)` enthalten:

1. Zuerst richten wir einen `HTTP client` ein:

```
HTTPClient http;
```

```
http.setTimeout(24000); // 24 seconds timeout
http.begin("https://api.openai.com/v1/chat/completions");
http.addHeader("Content-Type", "application/json");
http.addHeader("Authorization", "Bearer " +
                String(api_key));
```

2. Als Nächstes bereiten wir die Nutzdaten (Payload) im JSON-Format vor:

```
StaticJsonDocument<1024> jsonPayload;
// setting a maximum size of 1KB
jsonPayload["model"] = "gpt-3.5-turbo";
// model is gpt-3.5-turbo,
// change if you have access to 4 or 4.5
jsonPayload["temperature"] = 0.8;
// randomness of response, the higher
// the value, the higher the randomness
jsonPayload["max_tokens"] = 2000;
// maximum words & punctuations
// limit to be generated by response
```

3. Bereiten Sie die Nachricht (ein verschachteltes Array von Nachrichten) vor, serialisieren Sie sie im JSON-Format und senden Sie die Anfrage:

```
JSONArray messages = jsonPayload.
    createNestedArray("messages");
//put in a nested format not random
JsonObject userMessage = messages.createNestedObject();
userMessage["role"] = "user";
userMessage["content"] = prompt1;
String payloadString;
serializeJson(jsonPayload, payloadString);
```

4. Senden Sie die Anfrage und empfangen Sie die Antwort:

```
int httpResponseCode = http.POST(payloadString);
//transfer to the open page
if (httpResponseCode == HTTP_CODE_OK) {
    String response = http.getString();
    // if everything goes OK get the
    // reply = output and put in a string
    ...
}
```

5. Die Webseite enthält in der Regel einen Dschungel von überflüssigen Informationen (im HTML-Code), die für unsere Anwendung nicht nützlich sind und entfernt werden können. Wir normalisieren die Antwort, indem wir JSON-Tags und Ähnliches entfernen, damit sie auf dem TFT-Display gelesen werden kann:

```
StaticJsonDocument<1024> jsonResponse;
//parse the tags etc of the response string.
deserializeJson(jsonResponse, response);
// and put in a simple stripped string & return
String assistantReply = jsonResponse
    ["choices"][0]["message"]["content"].
    as<String>();
// select the first part which contains our reply
```

```
return assistantReply;
```

Software: Einrichtung und Schleife

In der `setup`-Funktion werden das TFT-Display, das I²S-Board und die WLAN-Anmeldedaten (Verbindung mit dem Internet) initialisiert.

```
void setup() {  
  delay(300);  
  ...  
  audio.setPinout(I2S_BCLK, I2S_LRC, I2S_DOUT);  
  //I2S board initialised  
  audio.setVolume(50);  
}
```

In der `loop`-Funktion haben wir Fragen an ChatGPT gesendet. Wenn die Sitzung beendet ist, bleiben wir in der `loop`:

```
String response = makeApiRequest(msg);  
  // sent to Chatgpt  
  ...  
  if (l1>200) {  
    response = response.substring(0, 200);  
    // truncate first 200 characters  
  }  
  audio.connecttospeech(response.c_str(), "en");  
  //speak up the 200 characters
```

Google-TTS hat ein Limit von 200 Zeichen und gibt, wenn dieses Limit überschritten wird, gar keinen Text mehr aus. Um dieses Problem zu umgehen, wird der Antwortstring für Google-TTS auf 200 Zeichen gekürzt. So wird stets die vollständige Antwort auf dem Bildschirm angezeigt, aber nur die ersten 200 Zeichen gesprochen. Bei längeren Antworten scrollt der Bildschirm, was aber durch kleine Änderungen im Sketch geändert werden kann.

Projekt-Tests

Die Verzögerungen, die ich in den Software-Schleifen verwendet habe, sind ziemlich spezifisch. Sie können sie anpassen, aber ich empfehle, mit den im Code verwendeten Standardwerten zu beginnen. Wenn Sie mit den Antworten zufrieden sind, können Sie die Werte ändern. Ich habe mit einfachen Fragen wie „Who r u?“ begonnen. ChatGPT

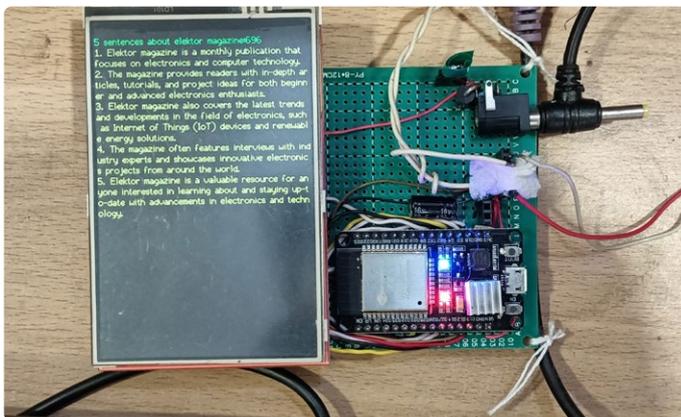


Bild 4. Testen des Systems mit Aufforderungen zur Eingabe von Blink-Sketches für Arduino UNO, ESP32 und Raspberry Pi Pico.

reagierte angemessen, indem es die Einleitung auf dem Bildschirm anzeigte, während der Sprecher sie deutlich artikulierte.

Dann testete ich das System mit Fragen wie „Write 5 sentences about Elektor magazine“ (Bild 4) und beauftragte ChatGPT anschließend, Blink-Sketches für Arduino UNO, ESP32 und Raspberry Pi Pico zu schreiben. In jedem Fall funktionierte ChatGPT einwandfrei, verstand den Kontext perfekt und antwortete präzise, wobei die Sprache laut und deutlich wiedergegeben wurde.

Zum Abschluss

Alle Dateien zu diesem Projekt finden Sie unter [2]. Das Internet und Webbrowser wie Google haben unseren Zugang zu Informationen völlig revolutioniert und die sperrigen Enzyklopädien und Lexika ersetzt, die früher unsere Bücherregale zu Hause füllten. Jetzt erleben wir den Aufstieg von KI-Software und Maschinen auf der Grundlage von ChatGPT-API, TensorFlow Lite Micro, Edge Impulse, OpenMV, TinyML, die sich anschicken, bestehende Geschäftsmodelle und traditionellere Methoden der Problemlösung abzulösen. Wir leben in spannenden Zeiten! ◀

SE – 230536-02

Über den Autor

Somnath Bera, ein Maschinenbauingenieur am indischen *Jalpaiguri Government Engineering College*, arbeitet als General Manager bei NTPC, dem größten Stromerzeuger des Landes. Er hat eine tiefe Leidenschaft für Elektronik, was durch seine mehr als 60 innovativen Projekte bei Elektor-Labs bewiesen wird, von denen mehr als zehn auch in der Zeitschrift Elektor veröffentlicht worden sind. Seine Projekte beschäftigen sich mit Problemlösungen in der Abfallwirtschaft und der Bewirtschaftung natürlicher Ressourcen. Somnath verwendet gerne innovative Ansätze und Plattformen wie Arduino, Raspberry Pi und ESP32 in Verbindung mit verschiedenen Arten von Sensoren und drahtlosen Systemen, um effiziente und kostengünstige Lösungen zu schaffen.

Sie haben Fragen oder Kommentare?

Wenn Sie technische Fragen oder Anmerkungen zu diesem Artikel haben, wenden Sie sich bitte per E-Mail an den Autor unter berasomnath@gmail.com oder an die Elektor-Redaktion unter redaktion@elektor.de.



Passendes Produkt

➤ **Espressif ESP32-DevKitC-32E**
www.elektor.de/20518

WEBLINKS

- [1] OpenAI-Webseite: <https://platform.openai.com>
- [2] Elektor-Webseite für diesen Artikel (Downloads):
<https://www.elektormagazine.de/230536-02>



Leistungsstarke Edge-KI: Der neue STM32N6

Ein Wendepunkt in Sachen Leistungs- und KI-Fähigkeit

Ein Beitrag von STMicroelectronics

Der neue STM32N6 soll den wachsenden Anforderungen an hochleistungsfähiger Edge-KI gerecht werden und neue Applikationen in den Bereichen Computer Vision, Audioverarbeitung und anderen hoch entwickelten Anwendungen ermöglichen. Der Mikrocontroller hebt den Maßstab für Video- und Multimedia-Anwendungen weiter an und verbessert das Benutzererlebnis durch seine umfangreichen Multimedia-Funktionen.

Die Mikrocontroller STM32 - basierend auf den Arm-Cortex-CPUs von M0 bis M7 - sind in der Elektroniker-Community wohlbekannt, da sie mit einer Vielzahl von Peripheriefunktionen ausgestattet sind und ein umfangreiches Ökosystem von Entwicklungsboards, Bibliotheken und Softwaretools für alle Arten von Anwendungen zur Verfügung steht. Der neue STM32N6 ist speziell ausgerichtet an der Vorverarbeitung von Daten „on the Edge“, mit Hilfe von KI in Form einer neuronalen Recheneinheit (Neural Processing Unit, NPU). Damit eröffnen sich neue Anwendungsfelder für Arm-Cortex-M-Prozessoren (**Bild 1**).

Merkmale und Vorteile des STM32N6

- Die eingebettete NPU (Neural-ART-Beschleuniger von ST) ist das Herzstück des neuen STM32N6. Die NPU bietet beeindruckende 600 GOPS bei sehr geringer Stromaufnahme. Sie ist ein entscheidender Faktor, um High-End-KI-Leistung (**Bild 2**) auf einem Mikrocontroller mit geringem Platzbedarf und niedriger Stromaufnahme bei geringen Kosten zu ermöglichen.
- Der STM32N6 wird von einem Arm-Cortex®-M55-Kern, der bei 800 MHz arbeitet, angetrieben und bietet 1280 DMIPS und 3360 CoreMark, was die bisher höchste Leistung für eine STM32-MCU darstellt. Dies macht ihn ideal für hoch entwickelte Bildverarbeitungs- und Grafikanwendungen, die von einem Chrom-ART-Beschleuniger, einem NeoChrom-Beschleuniger, einem H264-Encoder und JPEG-Encoder/Decoder unterstützt werden.
- Die MCU verfügt über 4,2 MB eingebetteten Arbeitsspeicher, der die Datenverarbeitung in Echtzeit, Multitasking und eine effiziente Ausführung unterstützt, ohne dass (in den meisten Fällen) ein externer Speicher erforderlich wäre. Darüber hinaus bietet eine umfang-

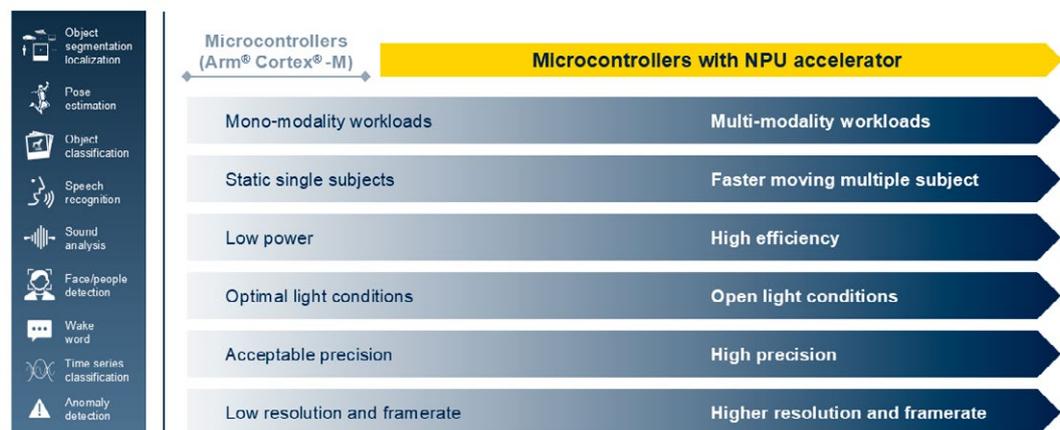


Bild 1. Die Neural Processing Unit (NPU) erschließt neue Anwendungsbereiche.

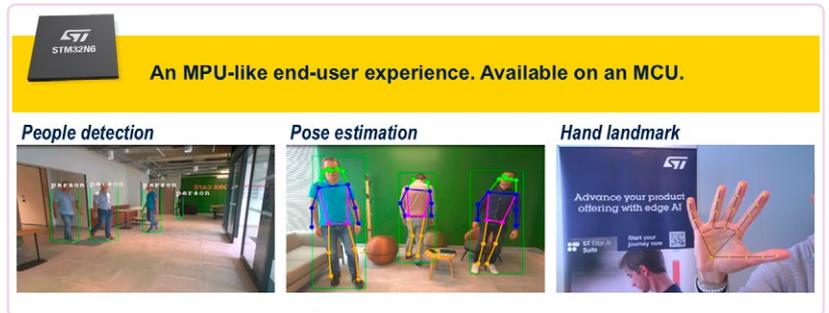
reiche Peripherie wie Dual-USB, Gbit-Ethernet, I3C- und sieben UART-Schnittstellen vielseitige Anschlussmöglichkeiten für verschiedenste Anwendungen.

- Ein dedizierter Bildsignalprozessor (Image Signal Processor, ISP) ist für eine 5-MP-Kamera mit einer Framerate von 30 Bildern pro Sekunde dimensioniert. Eingebettete Firmware für den Arm-Cortex-Prozessor ermöglicht automatischen Weißabgleich und automatische Belichtung. Mit einem Software-Tool kann der ISP fein abgestimmt werden (**Bild 3**).
- Der STM32N6 ist vollständig in das STM32-Ökosystem integriert und wird von der *ST Edge AI Suite* unterstützt, die Tools wie *STM32Cube.AI*, eine KI-Modellsammlung, verschiedene Ressourcen und Anwendungsstudien umfasst. Der Controller wird zudem von zahlreichen Partnern unterstützt, darunter *Edge Impulse*, *Nota.AI* und *EmbedUR*, die eine umfassende Entwicklungsumgebung bereitstellen.

Umfassendes Software-Ökosystem

STMicroelectronics steht an der Spitze der Edge-KI-Bewegung und bietet umfassende Lösungen, die den unterschiedlichen Anforderungen von Ingenieuren und Entwicklern verschiedener Bereiche gerecht werden. Mit der *Edge AI Suite* bietet ST eine breite Palette von Software-Tools, die auf unterschiedliche Nutzerprofile und Wissensstände zugeschnitten sind, sowie eine große Sammlung von Fallstudien, die Entwickler anregen sollen.

STM32Cube.AI und *ST Edge AI Developer Cloud* sind für Anwender konzipiert, die ein hohes Maß an Flexibilität und Anpassungsfähigkeit benötigen. Diese Tools bieten Zugang zu vielen Ressourcen für die Optimierung und den Einsatz von benutzerdefinierten Modellen (**Bild 4**).



▲ Bild 2. Der Neural-ART-Beschleuniger von ST ermöglicht die Ausführung von hoch entwickelten KI-Anwendungen.



◀ Bild 3. Ein spezieller Bildsignalprozessor (ISP) ist für eine 5-MP-Kamera mit 30 Bildern pro Sekunde ausgelegt und kann über ein Softwaretool feinabgestimmt werden.

Der STM32 Edge AI Contest

Möchten Sie die Fähigkeiten des neuen STM32N6 in einem eigenen Projekt nutzen? Nehmen Sie am **STM32 Edge AI Contest** teil, der von STMicroelectronics und Elektor unterstützt wird! Wenn Sie ein professioneller Ingenieur, Maker oder Student sind, bewerben Sie sich für ein kostenloses STM32N6-Entwicklungsboard und entwickeln Sie Ihre eigene KI-Anwendung bis Ende April 2025. Weitere Informationen über den Wettbewerb finden Sie unter www.elektormagazine.com/stm32ai und in der Januar/Februar-Ausgabe 2025 von Elektor.

Model and code examples	ST Edge AI model zoo Online resources
Model benchmark	ST Edge AI Developer Cloud Online tool
Model optimization	STM32Cube.AI Plug-in for STM32CubeMX ST Edge AI Core Command line interface (CLI)

ST Edge AI Suite

st.com/st-edge-ai-suite

◀ Bild 4. *STM32Cube.AI* und *ST Edge AI Developer Cloud* sind für Anwender konzipiert, die auf eine große Flexibilität Wert legen. Die umfassende Modellsammlung mit Applikations-Beispielen ist ein hervorragender Ausgangspunkt für Entwickler, die schnell KI-Lösungen entwickeln möchten.



▲ **Bild 5.** Damit Sie Ihr Projekt schnell starten können, steht eine Reihe von Code-Beispielen für den STM32N6 zur Verfügung.

ST hat außerdem eine umfassende Modellsammlung entwickelt, die zusätzliche Ressourcen wie Retraining-Skripte und Anwendungscode-Beispiele enthält. Dies ist ein hervorragender Ausgangspunkt für Entwickler, die schnell KI-Lösungen implementieren möchten, ohne bei null anfangen zu müssen. Die Modellsammlung bietet vortrainierte Modelle für verschiedene Anwendungen wie Objekterkennung, Bildklassifizierung und Erkennung von Audioszenen, die an die individuellen Anforderungen angepasst werden können. Die Skripte zum Nachtrainieren und die Beispiele für den Anwendungscode erleichtern den Entwicklern die Anpassung dieser Modelle an ihre speziellen Anwendungen, wodurch der Entwicklungsprozess beschleunigt und die Markteinführungszeit verkürzt wird (**Bild 5**).

Das ist der STM32

Entwickler können von dem großartigen Ökosystem und den Erfahrungen mit den STM32-Produkten profitieren, um ihr eigenes Design zu entwickeln. Der STM32N6 bietet eine Vielzahl von Vorteilen für Entwickler, eröffnet neue Möglichkeiten und vereinfacht den Entwicklungsprozess:

Nahtlose KI-Integration

Der STM32N6 ist vollständig in das STM32-Ökosystem integriert, was den Einsatz von KI sehr einfach

macht. Entwickler können die *ST Edge AI Suite* mit Tools wie *STM32Cube.AI* nutzen, um ihre neuronalen Netzwerk-Modelle einfach zu optimieren und einzusetzen. Dies verringert die Schwierigkeiten der KI-Entwicklung und verkürzt die Markteinführungszeit für neue Produkte.

Kosten- und Leistungseffizienz

Der STM32N6 bietet KI-Benutzererfahrung bei kleinerem Platzbedarf, niedrigem Strombedarf und reduzierten Kosten. Dadurch können Entwickler innovative KI-fähige Anwendungen ohne die hohen Kosten, die normalerweise mit MPU- und GPU-Lösungen verbunden sind, erstellen.

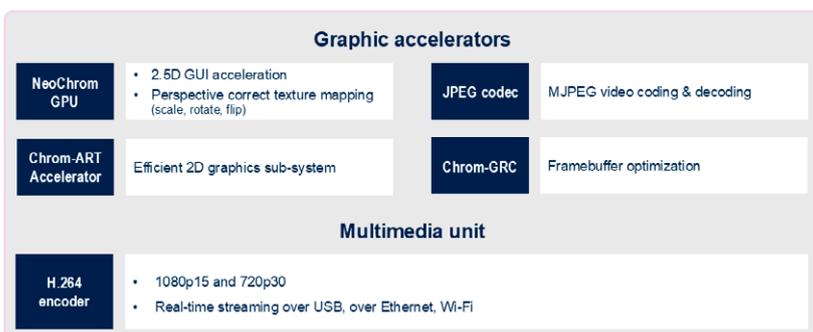
Verbesserte Leistung

Der leistungsstarke Arm-Cortex®-M55-Kern in Verbindung mit der eingebetteten NPU und reichlich RAM sorgt dafür, dass Entwickler komplexe KI-Algorithmen und Multimedia-Anwendungen effizient ausführen können. Damit lassen sich innovative Anwendungen entwickeln, die sich auf dem Markt von der Konkurrenz abheben. Die Leistungsfähigkeit des STM32N6 ermöglicht es Entwicklern, die Grenzen dessen, was mit MCUs möglich ist, zu erweitern.

Reichhaltige Grafikfunktionen

Der STM32N6 ist die leistungsstärkste STM32-MCU in Hinsicht auf die Grafikfunktionen. Er verfügt über den Chrom-ART™-Beschleuniger für 2D- und den NeoChrom™-Beschleuniger für 2,5D-Beschleunigung, einschließlich erweitertem Zeichnen, perspektivischer Korrektur und Textur-Mappings (**Bild 6**). Darüber hinaus dient der Chrom-GRC™ als Formatierer für Grafikressourcen für nicht-quadratische Displays. Mit 4,2 MB integriertem SRAM bietet der STM32N6 ausreichend Speicher, um große Grafiksammlungen und komplexe Animationen zu verarbeiten, ohne dass ein externer Speicher erforderlich wäre, was (reaktions-)schnellere Benutzeroberflächen gewährleistet.

▼ **Bild 6.** Der STM32N6, den man auch ohne neuronale Recheneinheit erwerben kann, ist die leistungsstärkste STM32-MCU in Bezug auf die Grafikfähigkeiten.



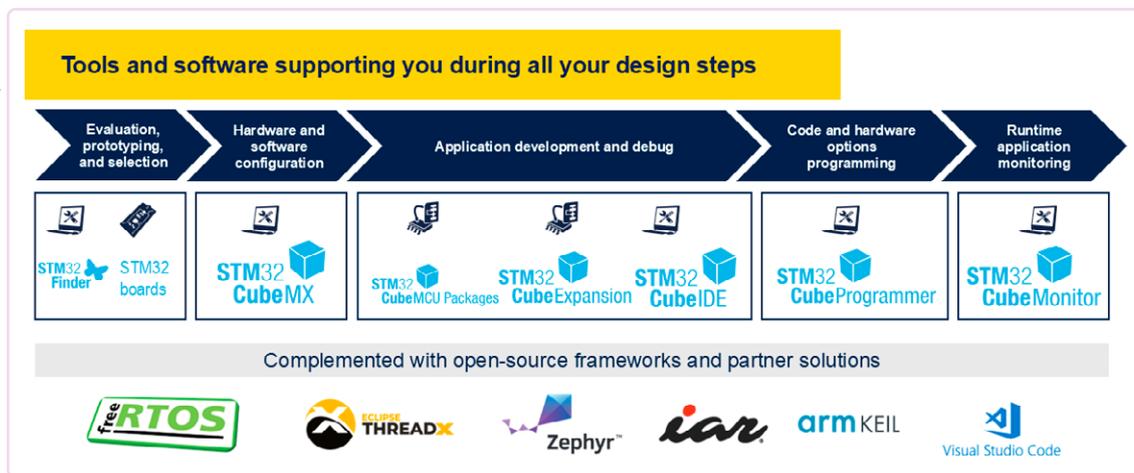


Bild 7. Der STM32N6 wird von einem robusten Ökosystem und zahlreichen Partnern unterstützt.

Umfassende Entwicklungsumgebung

Der STM32N6 wird von einem robusten Ökosystem und zahlreichen Partnern unterstützt und bietet Entwicklern eine umfassende Entwicklungsumgebung. Er nutzt die Software-Suite STM32Cube und bietet eine Reihe von Software-Tools für Konfiguration, Entwicklung, Programmierung und Überwachung. Für eingebettete Software umfasst es Pakete, Middleware und Treiberunterstützung, die durch RTOS-Optionen wie FreeRTOS, Zephyr und ThreadX von Eclipse ergänzt werden (**Bild 7**).

Im Bereich der Künstlichen Intelligenz haben Entwickler Zugang zu Tools, Ressourcen, Applikationsbeispielen und Support von Partnern wie Edge Impulse, Nota AI und EmbedUR. Für die Grafikentwicklung wird der STM32N6 von TouchGFX unterstützt, was die Erstellung von anspruchsvollen grafischen Benutzeroberflächen (GUIs) vereinfacht. Durch dieses umfangreiche Support-Netzwerk erhalten die Entwickler in jeder Phase des Entwicklungsprozesses die nötige Unterstützung, um optisch ansprechende und reaktions-schnelle Anwendungen zu erstellen.

Flexibilität und Vielseitigkeit

Mit seiner reichhaltigen Peripherieausstattung und den umfangreichen Anschlussmöglichkeiten bietet der STM32N6 Entwicklern die Flexibilität für eine Vielzahl von Anwendungen. Ob Computer Vision, Audioverarbeitung oder hoch entwickelte Multimedia-Anwendungen, der STM32N6 hält die notwendigen Werkzeuge und Fähigkeiten bereit, um innovative Ideen zu verwirklichen. Darüber hinaus ist der STM32N6 auch in einer Version ohne NPU für Anwendungen erhältlich, die Hochleistungsrechnen, Grafik und Multimedia, aber keine KI-Funktionen benötigen. Durch diese Vielseitigkeit können Entwickler die richtige Konfiguration für ihre spezifischen Anforderungen wählen. Und so eignet sich der STM32N6 für eine Vielzahl von Branchen, darunter Automotive, Gesundheitswesen, Unterhaltungselektronik und industrielle Automatisierung.

Spannende Möglichkeiten für Entwickler

Der STM32N6 eröffnet neue Horizonte für eingebettete KI-Anwendungen und bietet unübertroffene KI-Leistung auf einer MCU. Sein geringer Platz-, der niedrige Strombedarf und die geringen Kosten machen den Controller zur idealen Wahl für eine Vielzahl von Anwendungen, von der Audioverarbeitung bis zur Computer Vision. Der STM32N6 ermöglicht es Entwicklern, intelligentere, effizientere und sicherere Geräte zu entwickeln, die an der „Edge“ des Netzwerks arbeiten.

Entwickler können mit dem STM32N6 intelligente Hausautomatisierungssysteme realisieren, die verschiedene Befehle und Situationen erkennen und darauf reagieren und so den Komfort und die Sicherheit erhöhen. Im Gesundheitswesen kann der STM32 für die Entwicklung fortschrittlicher medizinischer Geräte verwendet werden, die den Gesundheitszustand von Patienten in Echtzeit überwachen und genaue Diagnosen sowie rechtzeitige Warnmeldungen liefern. In der Industrieautomation ermöglicht der STM32N6 die Entwicklung von robusten Systemen, die komplexe Aufgaben mit hoher Präzision und Effizienz ausführen und so die Produktivität und Sicherheit verbessern. Im Bereich der Unterhaltungselektronik kann dieser Mikrocontroller die Entwicklung neuer Gadgets fördern, die ein verbessertes Multimedia-Erlebnis bieten, von der hochauflösenden Videowiedergabe bis hin zu fesselnden Spielen.

Der STM32N6 ist ein Beweis für das Engagement von STMicroelectronics für Innovation und Spitzenleistungen. Durch die neue Definition der Mikrocontroller-Leistung und die Integration hoch entwickelter KI-Funktionen ermöglicht der STM32N6 Entwicklern die Erstellung innovativer Anwendungen, die sich im Wettbewerb abheben. Mit seiner nahtlosen Integration in das STM32-Ökosystem und der Unterstützung durch zahlreiche Partner ist der STM32N6 bereit, die Welt der Mikrocontroller und der eingebetteten KI zu revolutionieren. ◀

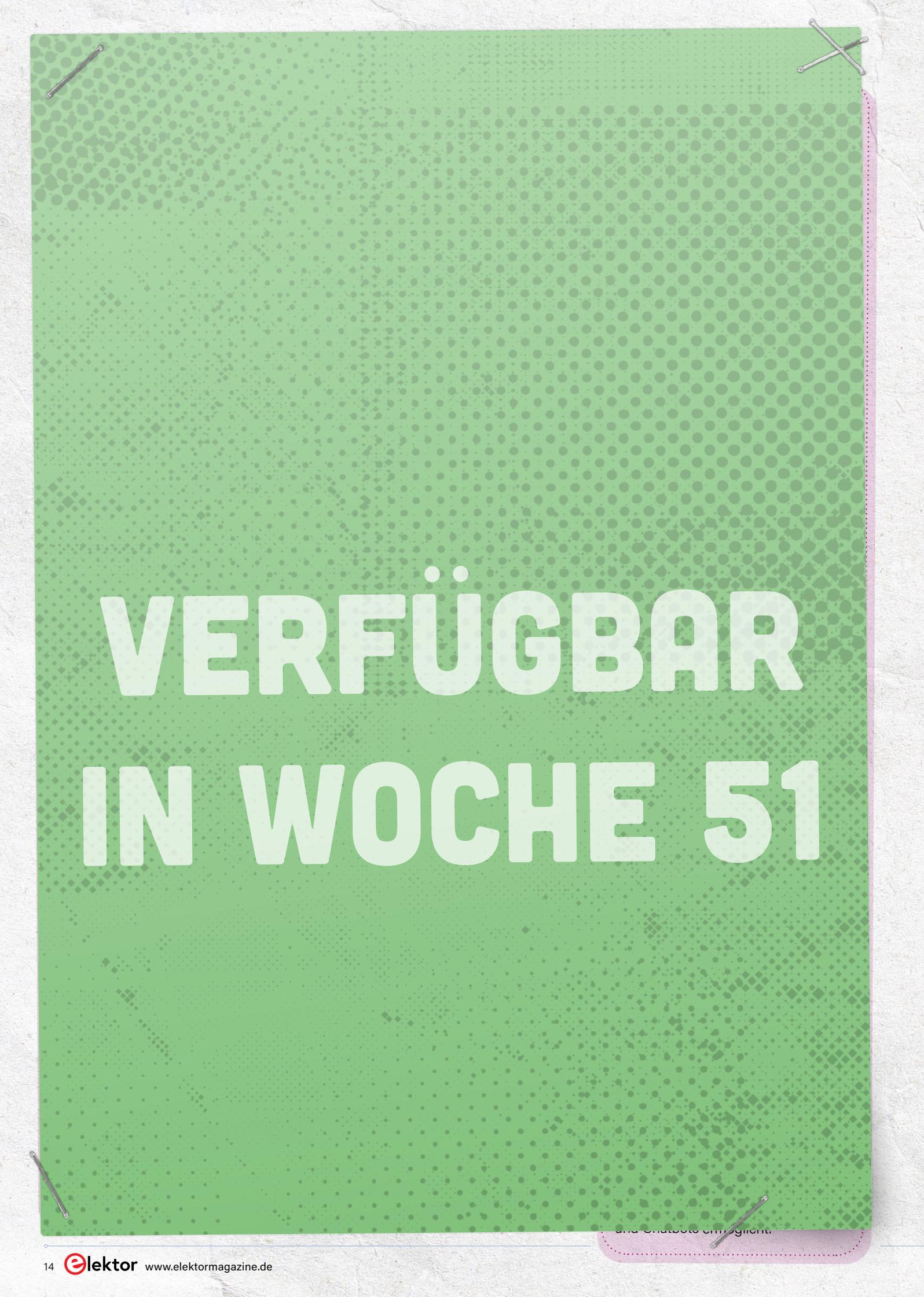
Übersetzung: Rolf Gerstendorf – 240676-02

WEBLINKS

Weitere Informationen finden Sie unter: www.st.com/stm32n6

VERFÜGBAR IN WOCHE 51

**VERFÜGBAR
IN WOCHE 51**



VERFÜGBAR IN WOCHE 51

VERFÜGBAR IN WOCHE 51

...erkennung, Spracherkennung und
Zeitreihenvorhersage verwendet.

VERFÜGBAR IN WOCHE 51

**VERFÜGBAR
IN WOCHE 51**



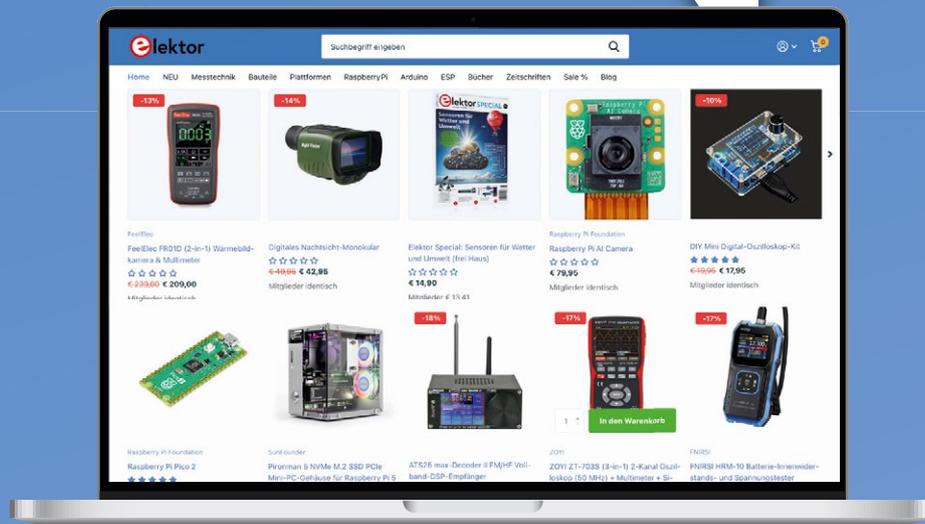
**VERFÜGBAR
IN WOCHE 51**

**VERFÜGBAR
IN WOCHE 51**



VERFÜGBAR IN WOCHE 51

Was ist Ihre Meinung?



Bei Elektor bieten wir mehr als nur Elektronik - wir schaffen ein Erlebnis, von Qualitätsprodukten und außergewöhnlicher Kundenbetreuung, unterstützt durch die Leidenschaft unserer Community.

Teilen Sie Ihre Meinung auf
www.elektor.de/pages/customer-reviews



Günther A.

Das Meßgerät wurde kurzfristig geliefert. Solide Ausführung, gut ablesbares Display, zuverlässige Kontaktierung der 4-Leiter Kelvin Meßleitungen, hohe Genauigkeit der Meßwerte (Spannung, Innenwiderstand)



Matthias D.

Ich hatte die Elektor schon in der 80ern und dann etwas aus den Augen verloren. Es ist wie damals: prima Schaltungsideen, anwendbar, sehr gut erklärt, alles bestens. Danke und bitte weiter so.



Hartmut W.

Sehr guter Service!

Schnelle Lieferung, gute Bezahlmöglichkeit, Gerät top. Sehr zu Empfehlen! Werde weiter Bestellen.

VERFÜGBAR IN WOCHE 52

**VERFÜGBAR
IN WOCHE 52**

VERFÜGBAR IN WOCHE 52

**VERFÜGBAR
IN WOCHE 52**

VERFÜGBAR IN WOCHE 52

**VERFÜGBAR
IN WOCHE 52**



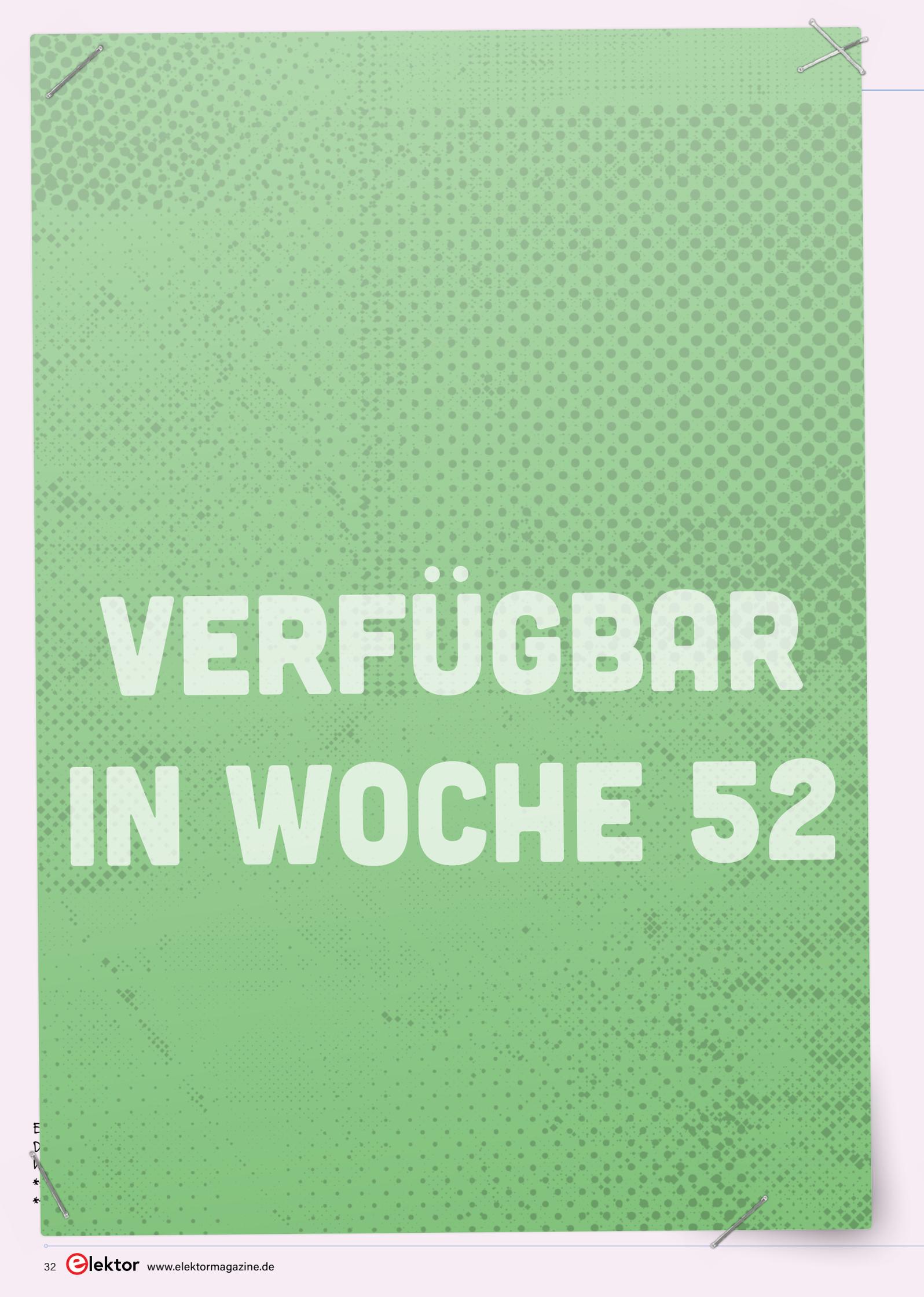
**VERFÜGBAR
IN WOCHE 52**

**VERFÜGBAR
IN WOCHE 52**



**VERFÜGBAR
IN WOCHE 52**

**VERFÜGBAR
IN WOCHE 52**



VERFÜGBAR IN WOCHE 52

**VERFÜGBAR
IN WOCHE 52**

VERFÜGBAR IN WOCHE 1

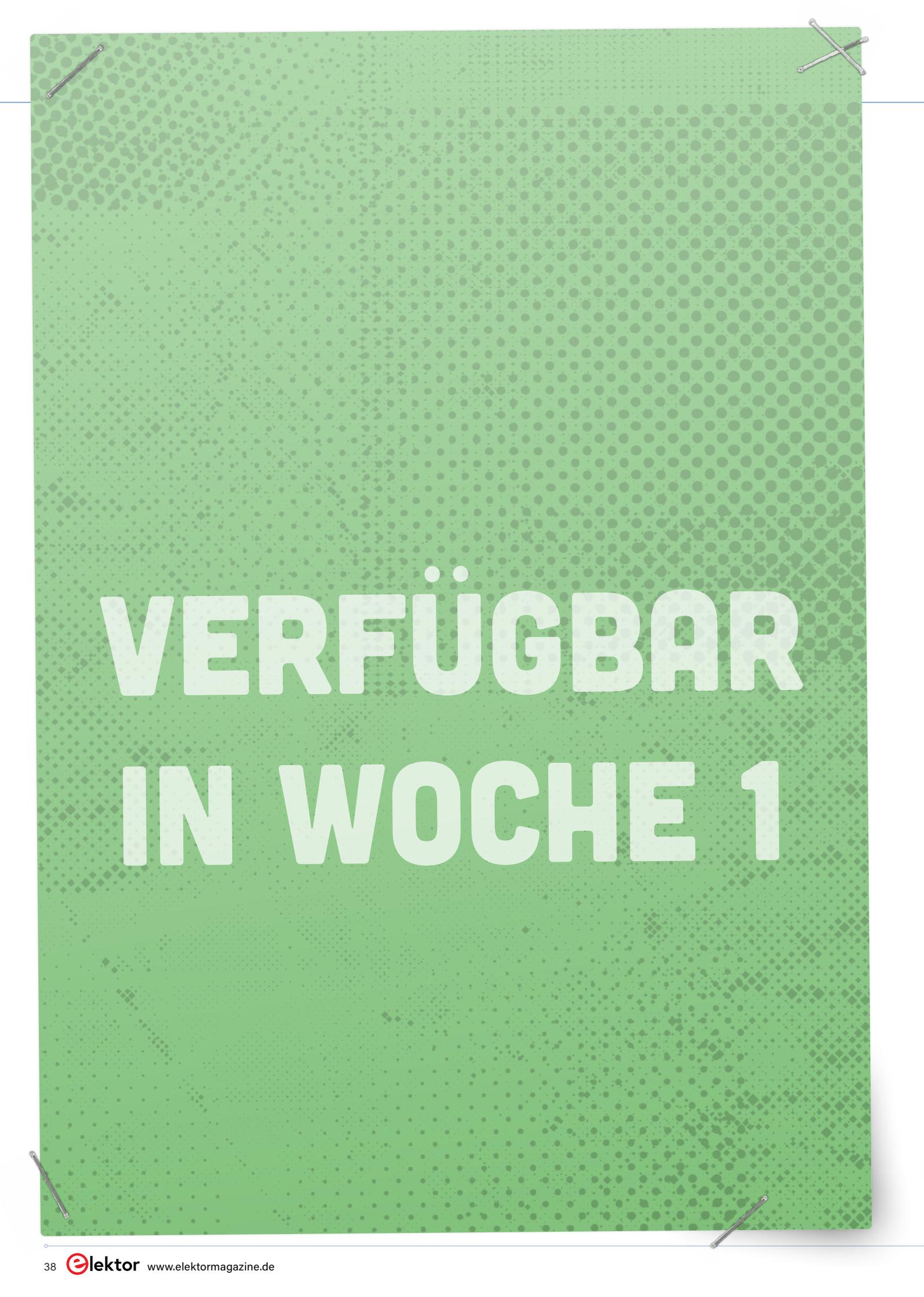
VERFÜGBAR IN WOCHE 1



VERFÜGBAR IN WOCHE 1

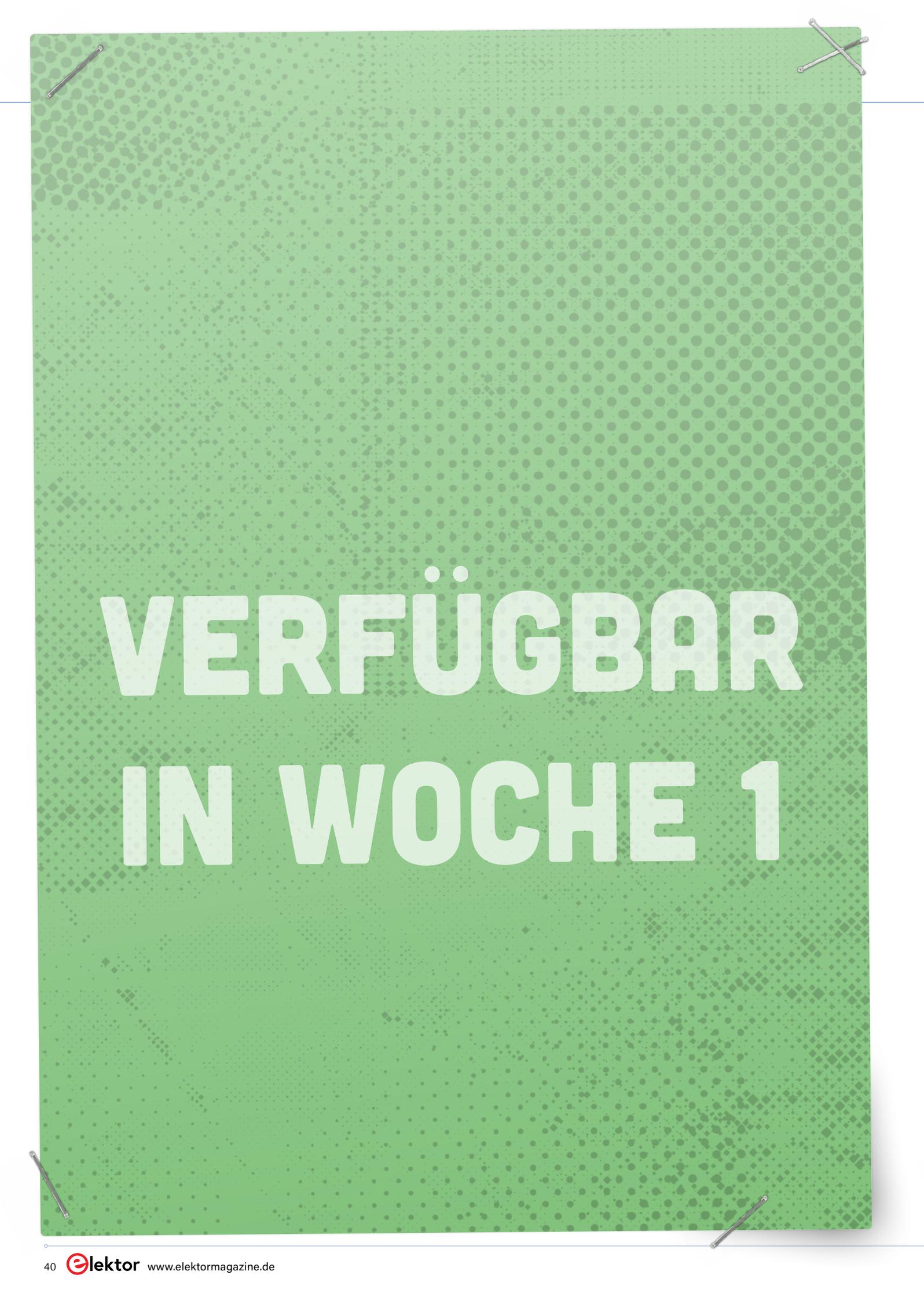


VERFÜGBAR IN WOCHE 1

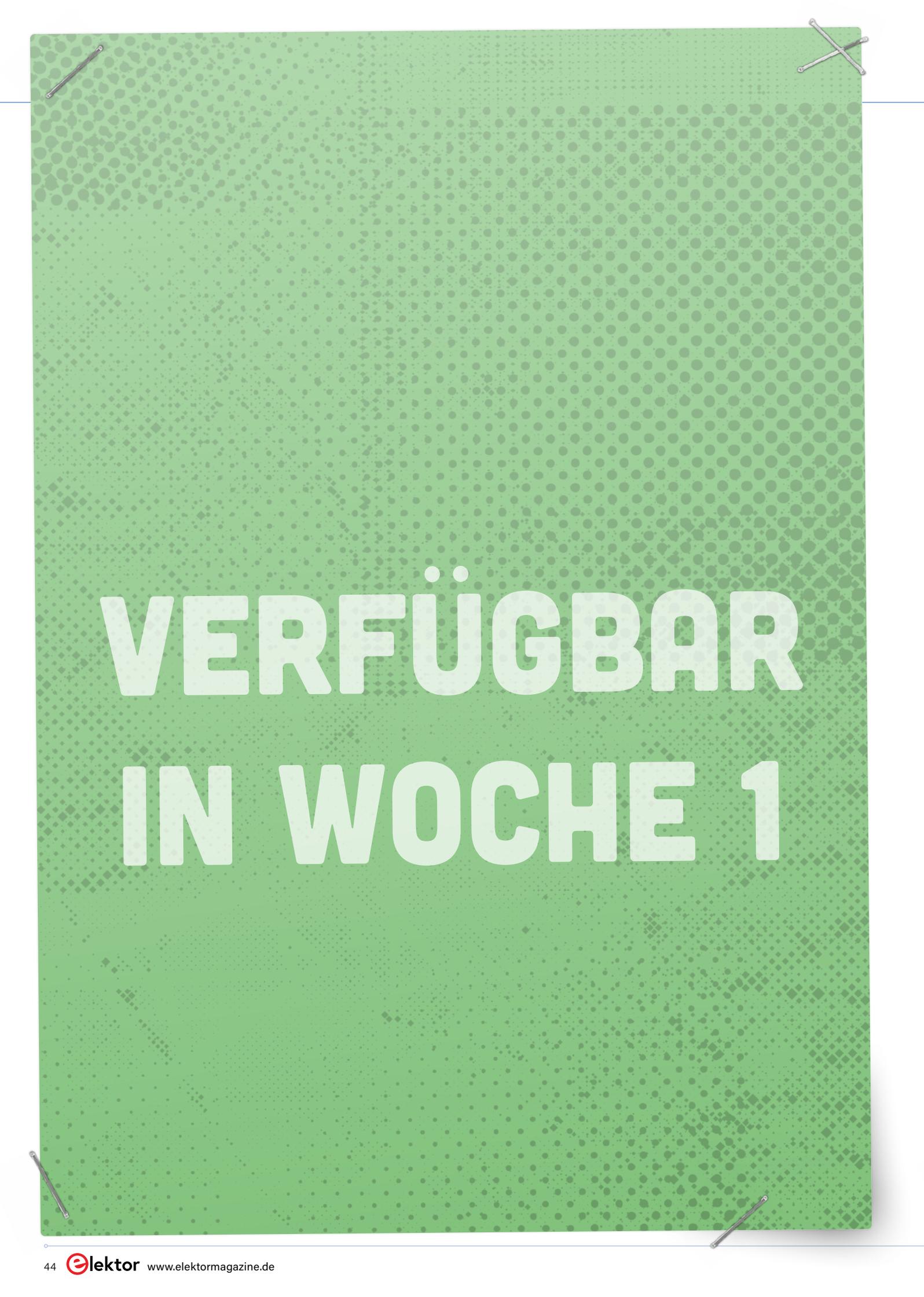


VERFÜGBAR IN WOCHE 1

VERFÜGBAR IN WOCHE 1

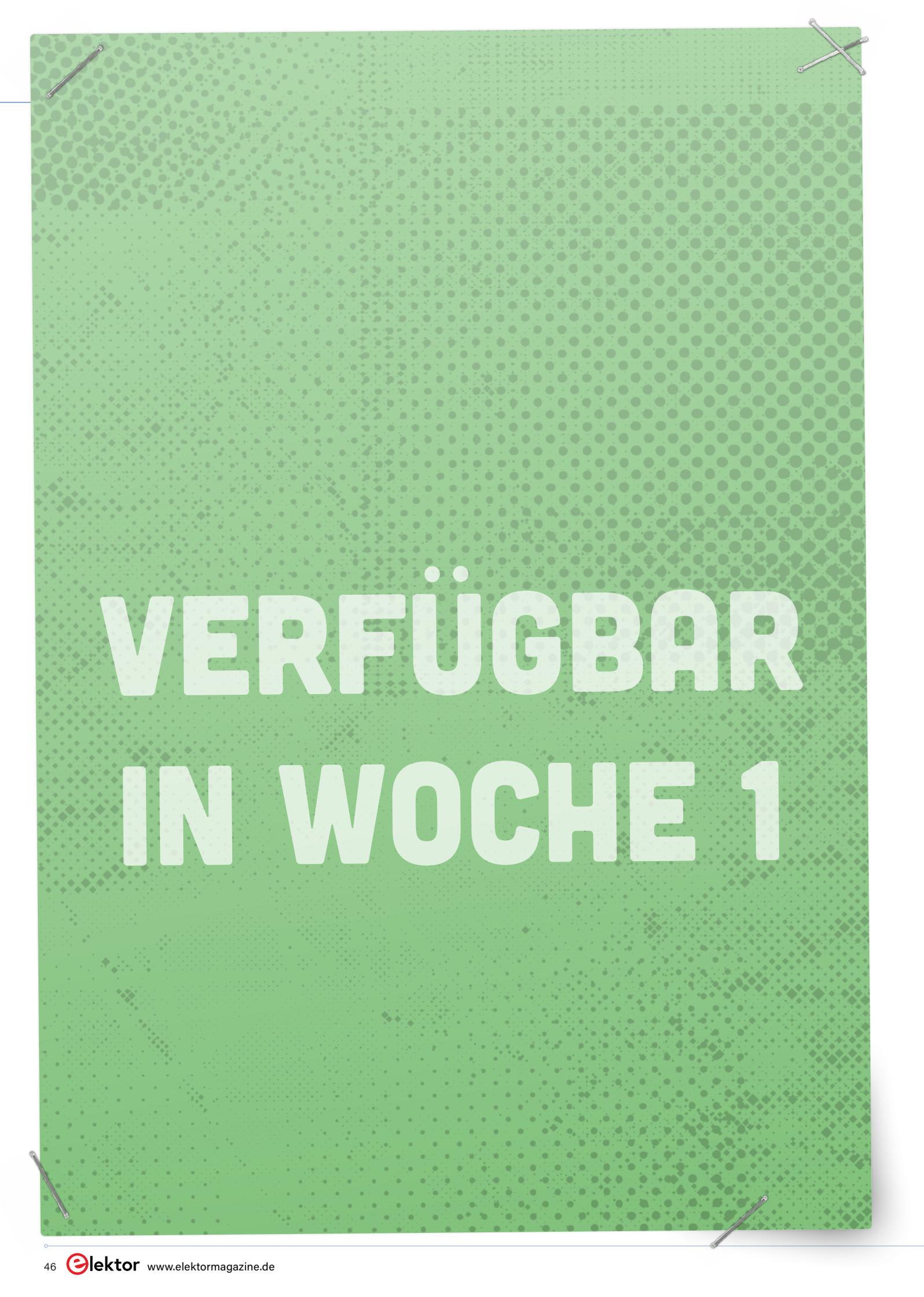


VERFÜGBAR IN WOCHE 1



VERFÜGBAR IN WOCHE 1

VERFÜGBAR IN WOCHE 1



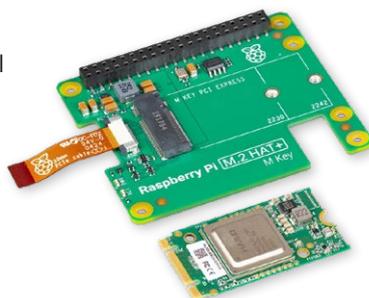
VERFÜGBAR IN WOCHE 1

Holen Sie sich modernste KI-Hardware ins Haus!

Raspberry Pi AI Kit

Das Raspberry Pi AI Kit enthält die Raspberry Pi M.2 HAT+ und ein Hailo-KI-Beschleunigungsmodul zur Nutzung mit dem Raspberry Pi 5. Es bietet eine zugängliche, kostengünstige und energieeffiziente Möglichkeit zur Integration leistungsstarker KI-Anwendungen. Entdecken Sie Anwendungen in den Bereichen Prozesssteuerung, Sicherheit, Heimautomatisierung und Robotik!

www.elektor.de/20879



Unitree Go2 Pro Vierbeiniger Roboter

Der Unitree Go2 ist ein vierbeiniger Roboter, der für die Forschung und Entwicklung autonomer Systeme in den Bereichen Mensch-Roboter-Interaktion (HRI), SLAM und Transport entwickelt wurde. Dank seiner vier Beine und den 12 Freiheitsgraden (12DOF) kann dieser Roboter eine Vielzahl unterschiedlicher Gelände bewältigen.

www.elektor.de/20357



BeagleY-AI SBC mit GPU, DSP und KI-Beschleunigern

BeagleY-AI ist ein kostengünstiger, quelloffener und leistungsstarker 64-Bit-Quad-Core-Single-Board-Computer, der mit einer GPU, DSP und Beschleunigern für Vision/Deep Learning ausgestattet ist und für Entwickler und Macher konzipiert wurde. Benutzer können die von BeagleBoard.org bereitgestellten Debian-Linux-Software-Images nutzen, die eine integrierte Entwicklungsumgebung enthalten.

www.elektor.de/20991

Raspberry Pi AI Camera

Die Raspberry Pi AI Camera ist ein kompaktes Kameramodul, das auf dem Sony IMX500 Intelligent Vision Sensor basiert. Der IMX500 kombiniert einen 12-MP-CMOS-Bildsensor mit integriertem Inferenzbeschleuniger für verschiedene gängige neuronale Netzwerkmodelle. Dies ermöglicht es Anwendern, anspruchsvolle KI-Anwendungen auf der Basis visueller Erkennung zu entwickeln, ohne einen separaten Beschleuniger zu benötigen.

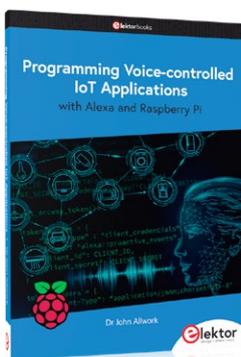
www.elektor.de/20953



Waveshare Jetson Orin Nano AI Development Kit

Dieses KI-Edge-Computing-Entwicklungskit basiert auf dem Jetson Orin Nano-Modul und bietet zahlreiche Peripherie-Schnittstellen wie M.2, DP, USB usw. Das Kit wird außerdem mit einer vorinstallierten AW-CB375NF-Wireless-Netzwerkkarte geliefert, die Bluetooth 5.0 und Dualband-WLAN unterstützt und zwei zusätzliche PCB-Antennen zur Verfügung stellt, um eine schnelle und zuverlässige drahtlose Netzwerkverbindung und Bluetooth-Kommunikation zu gewährleisten.

www.elektor.de/20762



Programming Voice-controlled IoT Applications with Alexa and Raspberry Pi

Dieses Buch ist in zwei Teile gegliedert: die Erstellung von Alexa Skills und die Gestaltung von IoT-Geräten mit einem Raspberry Pi. Es behandelt Themen wie die Entwicklung von Alexa Skills, In-Skill-Käufe und die Nutzung von AWS Lambda sowie den Bau von Smart-Home-Geräten, die von Alexa gesteuert werden. Die Leser lernen auch MQTT-Messaging, das Erstellen proaktiver Benachrichtigungen und die Umwandlung eines Raspberry Pi in ein eigenständiges Alexa-Gerät.

www.elektor.de/20400

Treten Sie jetzt der Elektor Community bei!



Jetzt



Mitglied werden!



- ✓ Zugang zum kompletten Online-Archiv (1970-heute)!
- ✓ 8x Elektor Magazin (gedruckt)
- ✓ 8x Elektor Magazin (PDF)
- ✓ 10% Rabatt im Elektor Store und exklusive Angebote
- ✓ Zugriff auf über 5.000 Gerber-Dateien u.v.m. aus der Projektplattform Elektor Labs



Auch erhältlich

Die digitale
Mitgliedschaft!



- ✓ Zugang zum kompletten Online-Archiv
- ✓ 8x Elektor Magazin (PDF)
- ✓ 10% Rabatt im Elektor Store und exklusive Angebote
- ✓ Zugriff auf über 5.000 Gerber-Dateien u.v.m. aus der Projektplattform Elektor Labs



www.elektormagazine.de/abo