

# amazon



IHK-Netzwerk Künstliche Intelligenz

## Agenten, Models und Mr. Q – Die Amazon Tools Bedrock, Q und CodeWhisperer

Workshop im Amazon Research and Development Center,  
Tübingen

# Agenten, Models und Mr. Q – Die Amazon Tools Q, CodeWhisperer und Bedrock

Ihr Ansprechpartner

Dr. Stefan Engelhard  
[engelhard@reutlingen.ihk.de](mailto:engelhard@reutlingen.ihk.de)

## Einleitung und Zusammenfassung

„GenAI ist the biggest technical transformation in our lifetime“, ist sich Amazon Chef Andy Jassy sicher. Beim Netzwerk KI Treffen im Amazon Research and Development Center auf dem Campus des Cyber Valleys in Tübingen ging es um die neuesten Tools von Amazon in diesem Bereich Generativer Künstlicher Intelligenz. Statt Modelle mittels Machine Learning selbst zu erstellen, gibt es Foundation Models, also Basis Modelle. „Diese reduzieren die Kosten enorm, denn das Training und Labeling von Daten ist sehr kostenintensiv“, so Sambit Tripathy, Solutions Architect bei Amazon. Setzt man diese Foundation Models ein, so ist zumeist nur noch Fine-Tuning für eine voll funktionsfähige Lösung erforderlich.

GenAI kann nicht nur Texte mittels Large Language Models (LLM), also ChatGPT und co., sondern auch Bilder, Videos und Audio erzeugen. Die Einsatzmöglichkeiten sind enorm. „Jedes Unternehmen sollte heute eine Datenstrategie haben, um GenAI erfolgreich fürs Unternehmen einzusetzen“, so Christian Müller, Principal Solutions Architect bei Amazon. Es geht um Content. „Wenn zu wenig Content in den Models enthalten ist, halluzinieren die GenAI-Lösungen“.

Ein LLM ist auch Amazon CodeWhisperer, das nachweislich den Programmieraufwand zwischen 30 und 57% reduziert hat. Amazon Q, als weiteres Tool, kombiniert die Models mit Realtime-Infos und Amazon Bedrock bietet den Zugang zu vielen unterschiedlichen Modellen, die alle so ihre Besonderheiten haben.

Um neben Q und Models die wichtigsten James Bond Protagonisten komplett zu machen, bedarf es natürlich noch Agenten. Diese integriert man in kostengünstigen kleinen LLMs, beispielsweise ganz individuellen firmenspezifischen lokalen LLMs. Sie kommunizieren mit anderen großen LLMs und der Außenwelt. Fertig ist die „perfekte Eierlegende-Wollmichsau-LLM“ für alle im Unternehmen.

Bei der bereits vierten Veranstaltung des IHK-IWW mit dem Amazon ging es um diese drei Amazon Tools Q, CodeWhisperer und Bedrock. Den Workshop kann ein 20-zeiliges Python Programm mittels Q zukünftig besser zusammen, als es es dieser Artikel tut.

## Über das IHK-IWW

Das Institut für Wissensmanagement und Wissenstransfer der IHK Reutlingen (IHK-IWW) führt den Wissenstransfer ausgehend vom Bedarf der Unternehmen der Region Neckar-Alb durch. In ausgewählten Bereichen organisiert das Institut das Wissensmanagement. Dazu zählen Technische Textilien, Innovationsmethode TRIZ, KI-Checks, Regionales Innovationsmanagement für Neckar-Alb und das Business Incubation Centre der Weltraumagentur ESA in Baden-Württemberg (ESA BIC).

Diverse Netzwerke des IHK-IWW unterstützen die Unternehmen bei Innovationsthemen, darunter die Netzwerke Forschung & Entwicklung, Innovation, Wasserstoff, Sicherheit, Qualität/ISO 13485, Künstliche Intelligenz, Produktion sowie der Cluster Technische Textilien Neckar-Alb.

## Amazon Q

*Vorab Achtung: Amazon Q hat PreView Status, sollte also noch nicht in der Produktion eingesetzt werden.*

Amazon Q ist ein GenAI basierter Assistent. Prinzipiell ist es ein ganz normales Foundation Model. Ein zusätzlicher Aspekt ist allerdings, dass Daten von außen hinzukommen. Die weitere wichtige Besonderheit sind die Zugriffsrechte. So kann die Personal-Abteilung auf ihre Personaldaten zugreifen, jedoch nicht auf die ihrer Kollegen aus der Finanzabteilung. Beide verwenden aber die gleiche Lösung und greifen gemeinsam auf alle anderen Daten zu. Die Inhalte in beispielsweise pdf-Dateien mit besonderen Zugriffsrechten bleiben auf den richtigen Personenkreis beschränkt. Durch diese beiden



*Firmengebäude von Amazon in Tübingen*

Ergänzungen – Außenwelt und Zugriffsrechte – wird die Verwendung von unterschiedlichen LLMs obsolet. Weiterer großer Vorteil: Insbesondere da große Modelle teuer sind, kann man mit den schlanken lokalen LLMs auf Q enorme Kosten sparen, denn es reicht zumeist aus, dass nur einige wenige Dokumente und Infos aktuell zur Verfügung stehen. Beispielsweise Modelle die auch für die Erstellung von Unternehmensbilanzen eingesetzt werden, können so sehr klein gehalten werden. Die für eine Bilanz wichtigen Informationen lassen sich in wenigen Dokumenten zusammenfassen.

Beeindruckend bei der Demo war, dass das Mini-Python-Programm nicht nur eine Konversation perfekt zusammenfassen konnte, sondern dass man im Dialog mit dem LLM auch erfahren konnte welche Meinung die einzelnen Teilnehmer bei dem Gespräch geäußert hatten.

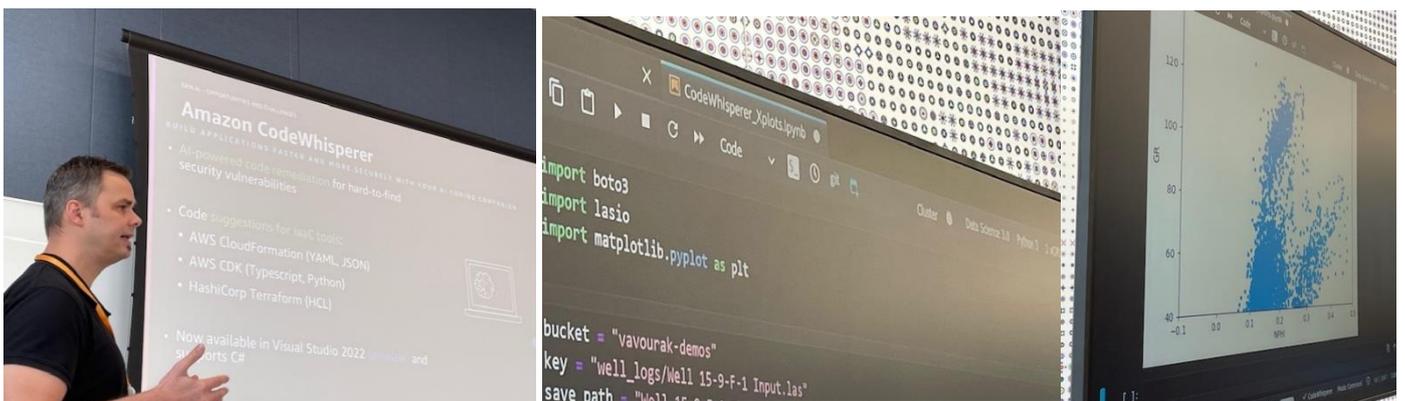


Interessant bei Amazon Q ist auch dessen Verbindung zu Amazon QuickSide. Dieses Tool kann optisch Inhalte ansprechend visualisieren. Perfekt. Selbst die Aufarbeitungen der Infos für die (Bilanz)-Präsentation wird via GenAI erledigt.

## Amazon CodeWhisperer

Bei der Live-Demo von Christian Müller ging es um die Programmierung einer Visualisierung von Daten. Hier werden oft stundenlang Parameter hin und hergeschoben bis die unterschiedlichen Aufbereitungstools

die Daten so bereitstellen, dass sie grafisch aussagekräftig sind. CodeWhisperer übernimmt den Job. Es greift auf zigtausende Erfahrungswerte zurück. Das Ausprobieren entfällt. Einsparungen an Programmieraufwand von 30% soll so nach Einschätzung von AWS im Schnitt realisiert werden.

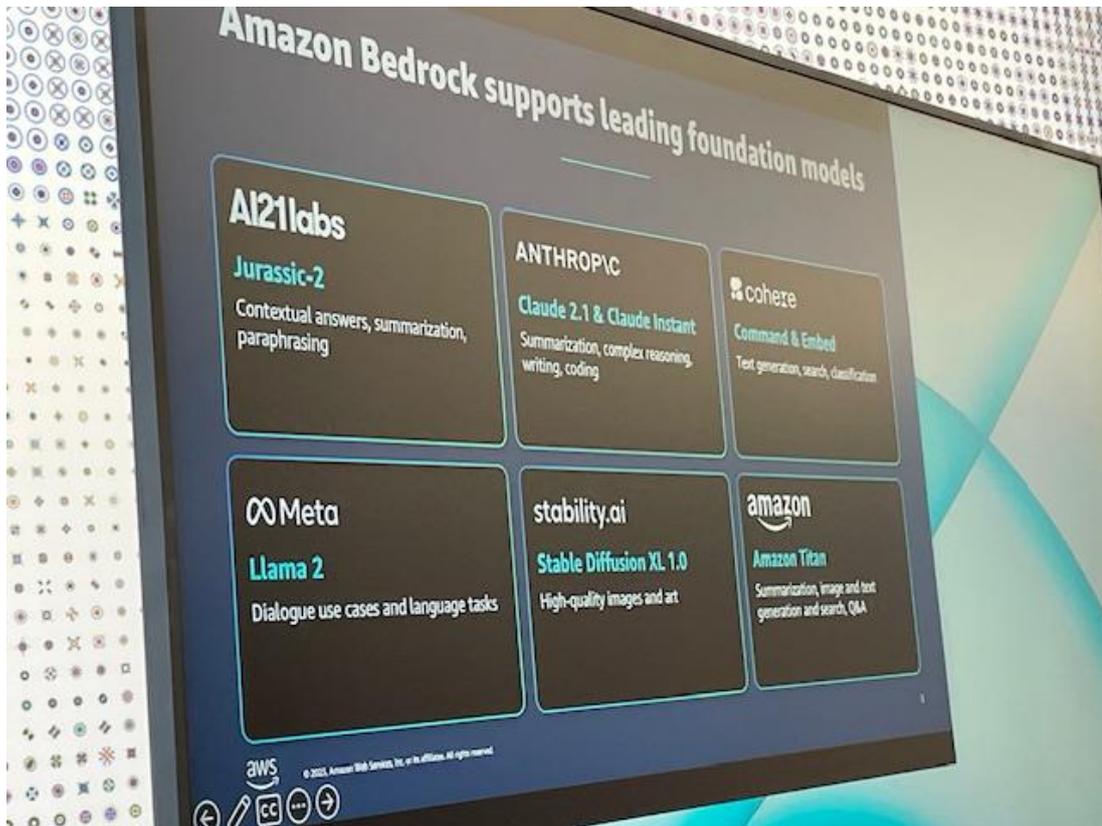


Christian Müllers Live Demo von Amazon CodeWhisperer zur grafischen Darstellung von Daten

## Amazon Bedrock

Bedrock steht für Grundstein. Es geht beim Einsatz von GenAI-Lösungen darum, die Basis und damit das richtige Foundation Model auszuwählen. Auch bringt es einen Mehrwert, wenn man die Ergebnisse von

unterschiedlichen Modellen vergleicht. Die von Bedrock bereitgestellten Modelle stammen beispielsweise von erfolgreichen Startups, wie AI21 Labs, andere wie Amazon Titan von Global Playern.



Die benutzerindividuelle „Knowledge Base“ ist die Ausgangsbasis bei Bedrock. Jedes Model hat das Problem des Cut-Datums, bis zu dem Content eintrainiert wurde. Der aktuelle Euro-Kurs ist damit nicht zu ermitteln und firmeninternes Wissen ebenso nicht. Auf der Knowledge Base sind neben den präferiert zu benutzenden Modellen diese weiterführenden Infos hinterlegt, die eigentlich nachtrainiert werden sollten



Sambit Tripathy ist Experte für Amazon Bedrock

Eine Wissenschaft für sich ist das stellen der passenden Anfrage an das Modell. Eine Anfrage ist ein Prompt. Prompt-Engineering nennt man diese Disziplin der Suche nach der richtigen Frage. Die beiden Amazon-Experten empfehlen den [„Prompt-Engineering-Guide“](#) sich anzuschauen. Viele Anfragen laufen im Übrigen unbemerkt vom Nutzer im Hintergrund ab.

Besonders relevant sind bei Bedrock die zusätzlichen Agenten. Eine einfache Frage „Ich suche Schuhe“ würde ohne Knowledge Base ein vernünftiges LLM einfach ablehnen. Es ist zu unspezifisch. Die Knowledge Base kennt aber die Schuhgröße, die Präferenz des Nutzers für beispielsweise qualitativ hochwertige Lederschuhe oder auch die Leidenschaft Wandern. Dies alleine würde aber auch nicht ausreichen. Eine einfache Vorinitialisierung würde diese Eingrenzung ebenfalls leisten. Die Knowledge Base verfügt daher zudem über Agenten die Real-Time-Daten heranziehen, um damit beispielsweise das aktuell verfügbare Angebot im Internet einzubeziehen.

Entsprechend kann die Frage nun doch mit dem Content der Knowledge Base beantwortet werden. Fast wie eine Schuhverkäuferin, die auf die Frage „Ich suche Schuhe“ beim Kunden die Schuhgröße und die gewünschte Schuhart erahnt, werden nun vom mit Bedrock erweiterten LLM Lederschuhe und Wanderschuhe der Größe 42 angeboten.

## Ausblick

Am Workshop, dem genau 22. Treffen des Netzwerks KI des IHK-IWW, nahmen 32 Teilnehmerinnen und Teilnehmer teil, überwiegend aus den mittelständischen Unternehmen der Region Neckar-Alb. Es werden nun in der Folge vertiefende Workshops zu Q, CodeWhisperer und Bedrock im Rahmen des Netzwerks KI angeboten, in denen die Teilnehmer selbst aktiv werden müssen. Eigene Daten und Laptops sind mitzubringen.

## Impressum

**Herausgeber:** Institut für Wissensmanagement und Wissenstransfer der IHK Reutlingen (IHK-IWW), Hindenburgstr. 54, 72762 Reutlingen  
**Autoren:** Gabriele Beese und Dr. Stefan Engelhard (Ansprechpartner)

Januar 2024

Alle Rechte vorbehalten. Nachdruck und Vervielfältigung auf Papier und elektronischen Datenträgern sowie Einspeisung in Datennetze nur mit Genehmigung des Herausgebers