

KI & LLM

Künstliche Intelligenz und Large Language Models

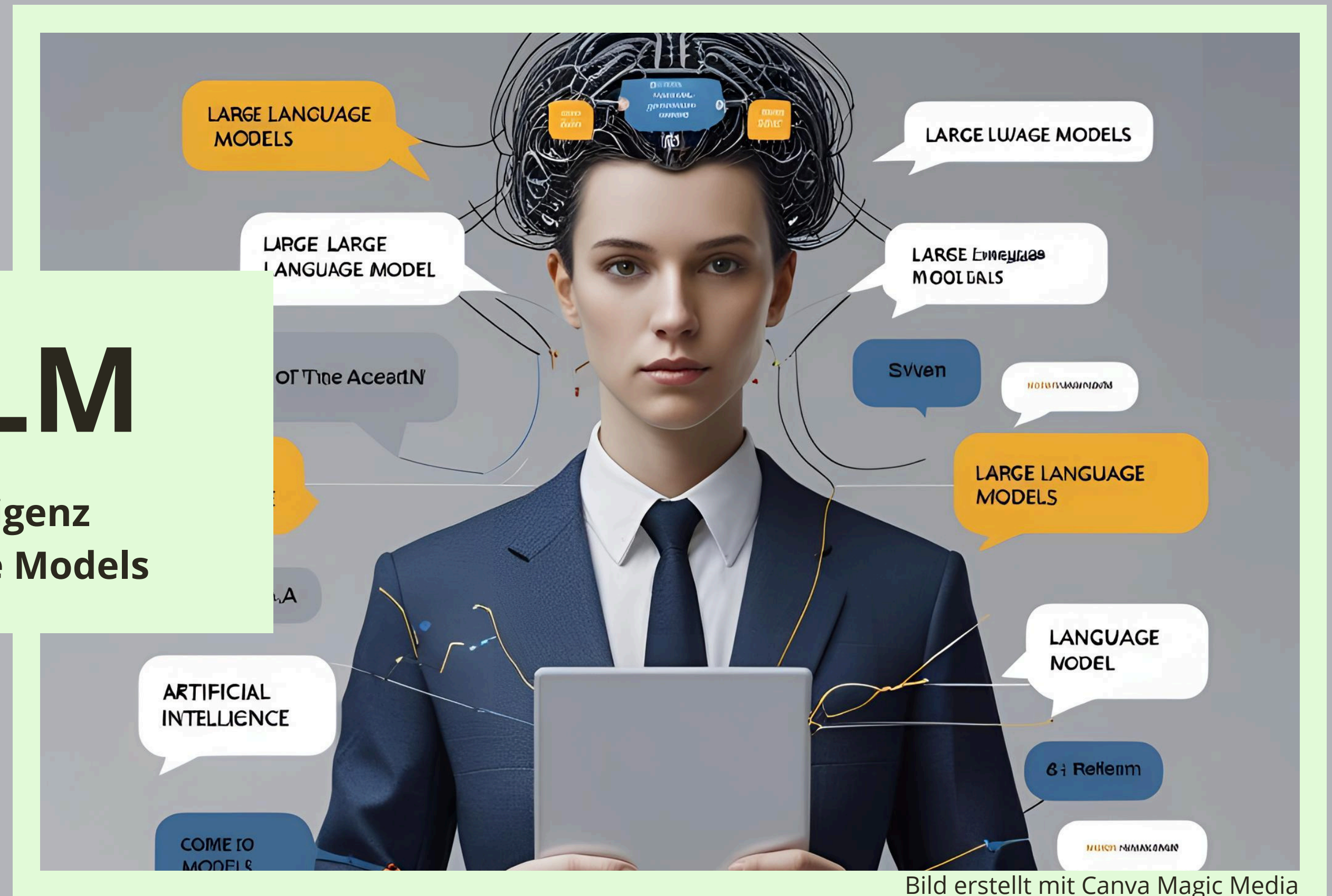


Bild erstellt mit Canva Magic Media

Was ist KI?

Was ist ein LLM?



Was ist KI?

- künstliche Intelligenz = Fähigkeit einer Maschine, menschliche Fähigkeiten wie logisches Denken, Planen und Kreativität zu imitieren
- basiert auf programmierten Abläufen oder maschinellem Lernen

Was ist ein LLM?

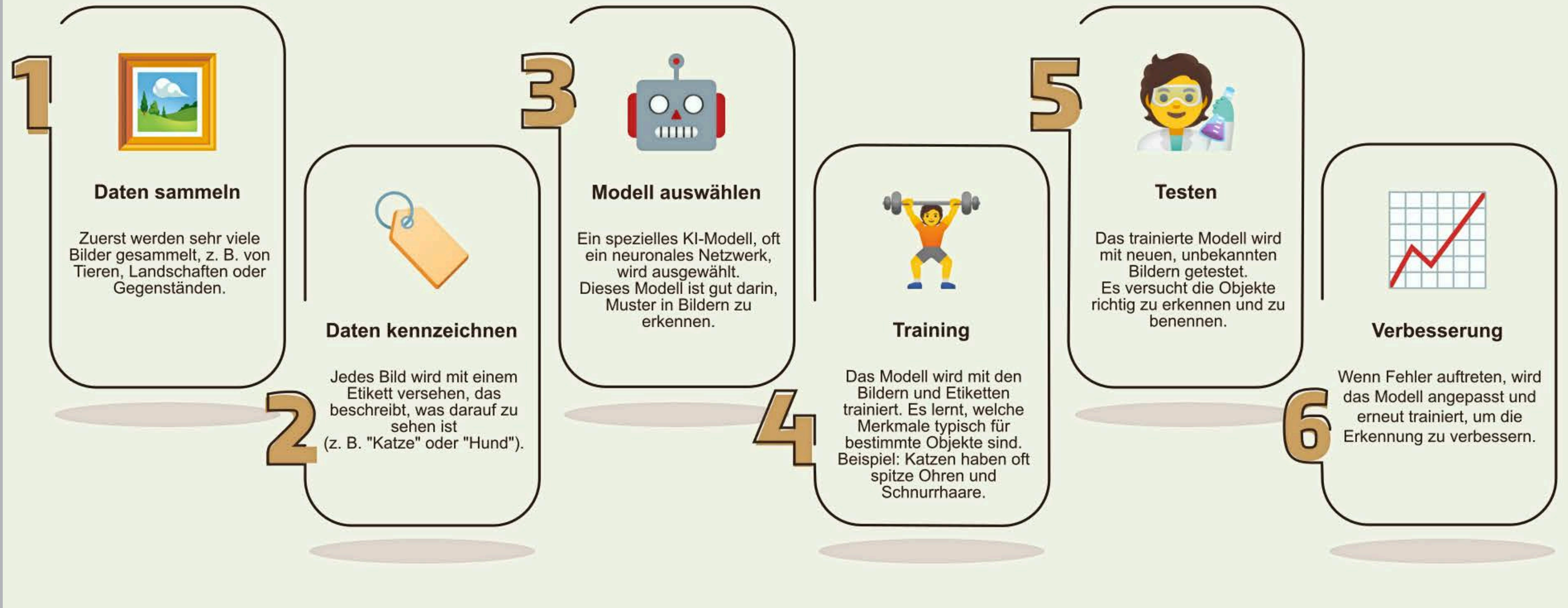
- LLM = Large Language Model (auf Deutsch: Großes Sprachmodell)
- leistungsstarke KI-Modelle, die darauf trainiert sind, menschliche Sprache zu “verstehen” und zu erzeugen.
- Beispiele: **GPT-4** von OpenAI, **Llama 2** von Meta, **Mistral Large** von Mistral AI



Quellen:

1. [https://www.iks.fraunhofer.de/de/themen/kuenstliche-intelligenz.html#:~:text=K%C3%BCnstliche%20Intelligenz%20\(KI\)%20ist%20ein,durch%20maschinelles%20Lernen%20erzeugt%20werden.](https://www.iks.fraunhofer.de/de/themen/kuenstliche-intelligenz.html#:~:text=K%C3%BCnstliche%20Intelligenz%20(KI)%20ist%20ein,durch%20maschinelles%20Lernen%20erzeugt%20werden.)
2. <https://www.europarl.europa.eu/topics/de/article/20200827STO85804/was-ist-kunstliche-intelligenz-und-wie-wird-sie-genutzt>
3. <https://www.iese.fraunhofer.de/blog/large-language-models-ki-sprachmodelle/>

Wie KI lernt - am Beispiel der Bildererkennung:



Übung: Gute Äffchen, böse Äffchen

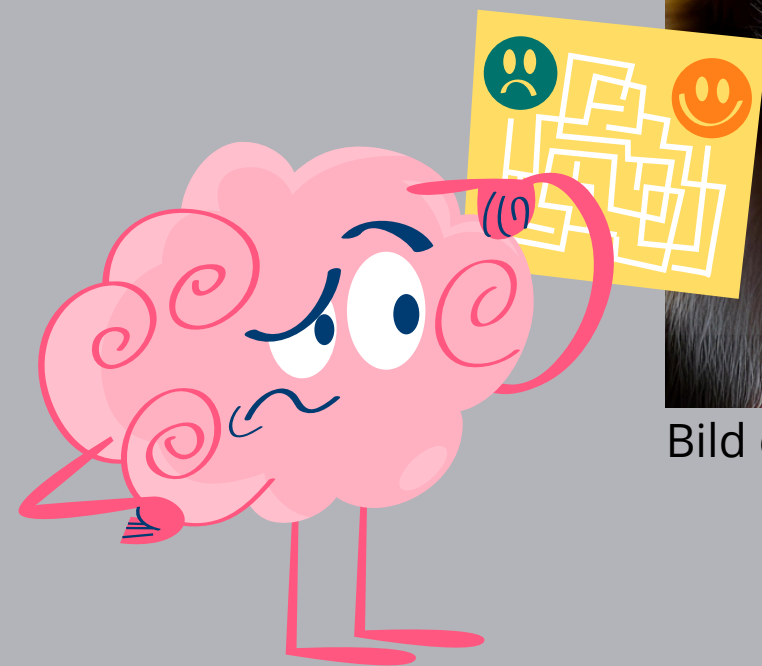


Bild erstellt mit Canva Magic Media

Wir trifft ein Computer selbstständig Entscheidungen? Wie entscheidet er zum Beispiel, welche Person einen Kredit erhalten sollte oder ob eine Person sportlich ist?

Solche Einteilungsvorgänge sind eine häufige Anwendung von KI.

In der Übung **Äffchenspiel** habt ihr die Möglichkeit, selbst ein Klassifikationsmodell zu erstellen. Dabei durchläuft ihr gleichzeitig die Trainings- und die Testphase einer lernenden KI.

Quelle "Äffchenspiel": AI Unplugged - Wir ziehen künstlicher Intelligenz den Stecker - Aktivitäten und Unterrichtsmaterial zu künstlicher Intelligenz ohne Strom

Urheber: Professur für Didaktik der Informatik / Friedrich-Alexander-Universität Erlangen-Nürnberg / Annabel Lindner, Stefan Seegerer

Link: <https://www.aiunplugged.org/german.pdf>

Lizenz: CC BY NC 3.0

Das ursprüngliche Material wurde auf den folgenden Folien (Seiten 6 - 13) adaptiert

Arbeitsauftrag Äffchenspiel - Durchgang 1

Wir sind Tierpflegerin in einem Zoo. Alle Äffchen sehen sehr süß aus, aber manche beißen. Von den Äffchen im Zoo wissen wir bereits, welche das sind.

Bald kommen aber neue Äffchen zu uns. Wir müssen überlegen wie wir herausfinden können, welche davon beißen und welche nicht.

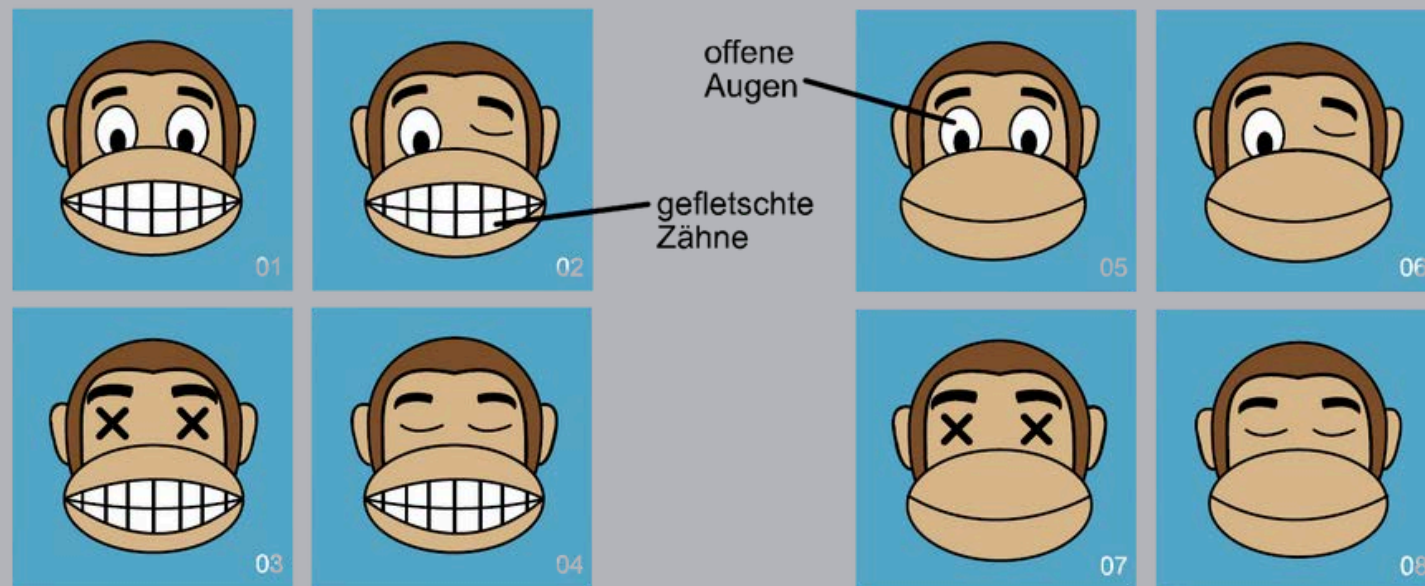


Abb. 1. In diesem einfachen Beispiel beißen alle Äffchen mit gefletschten Zähnen

1 Gruppenbildung

Bildet Teams aus 2 bis 3 Personen.

2 Trainingsphase: Kriterien festlegen (5 Min.)

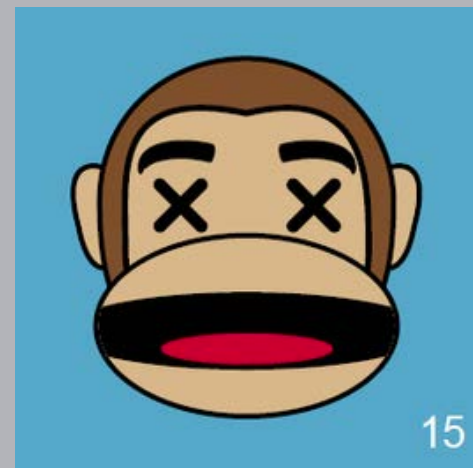
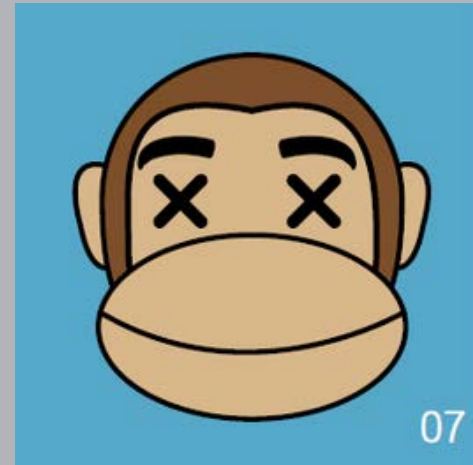
- Überlegt euch anhand der Trainingsdaten Kriterien, wie ihr beißende von nicht beißenden Äffchen unterscheiden könnt.
- Notiert die Kriterien so eindeutig, dass auch die anderen Gruppen sie verstehen.

3 Testphase: neue Äffchen zuordnen (5 Min.)

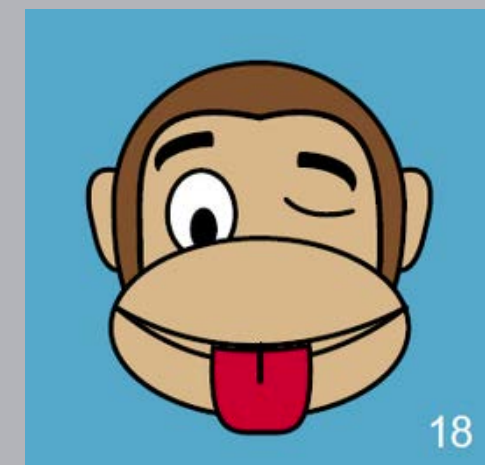
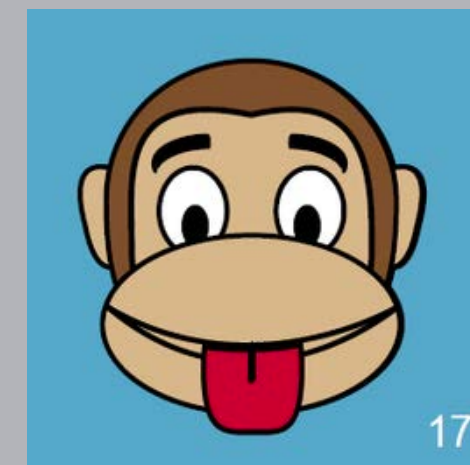
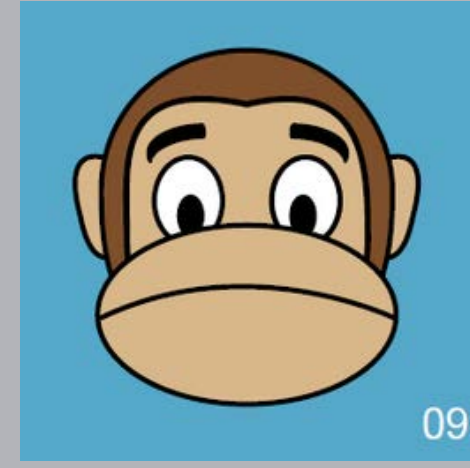
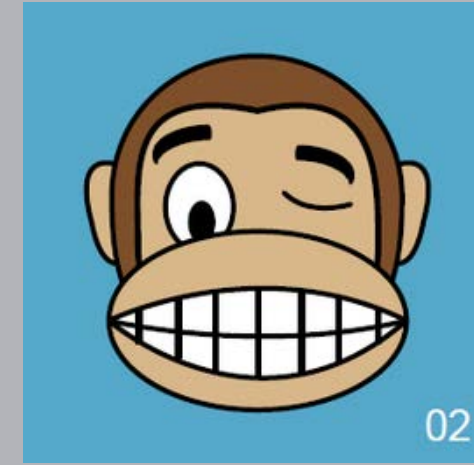
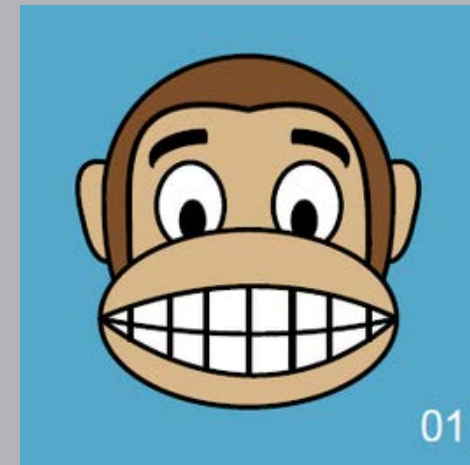
- Tauscht Eure Kriterien mit einer anderen Gruppe.
- Entscheidet anhand der Kriterien der anderen Gruppe, welche der Äffchen aus den Testdaten beißen.
- Notiert eure Entscheidungen.

Trainingsdaten 1:

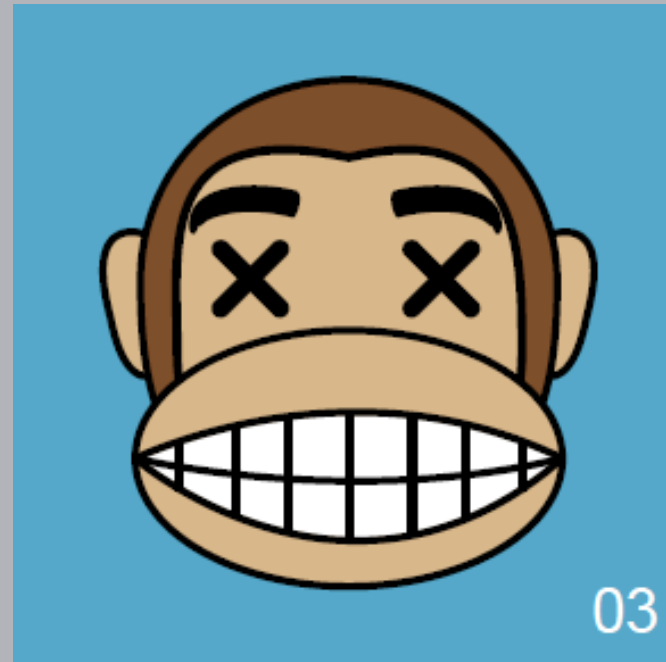
Beißt



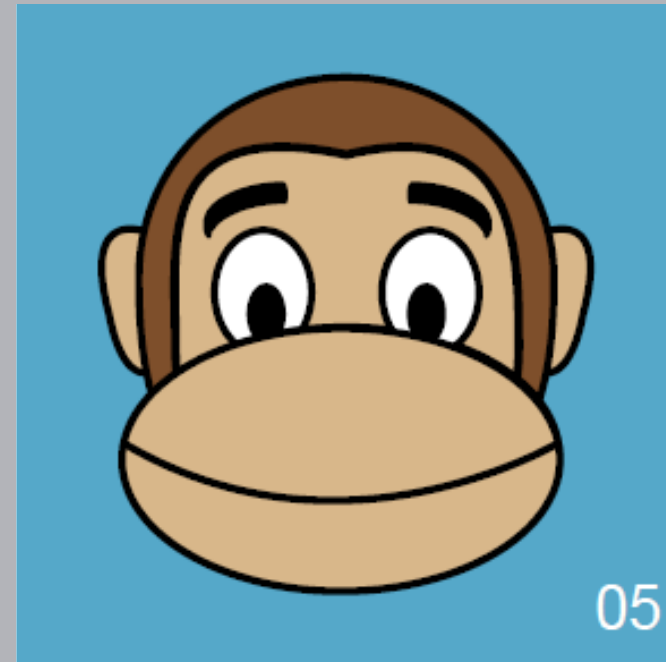
Beißt nicht



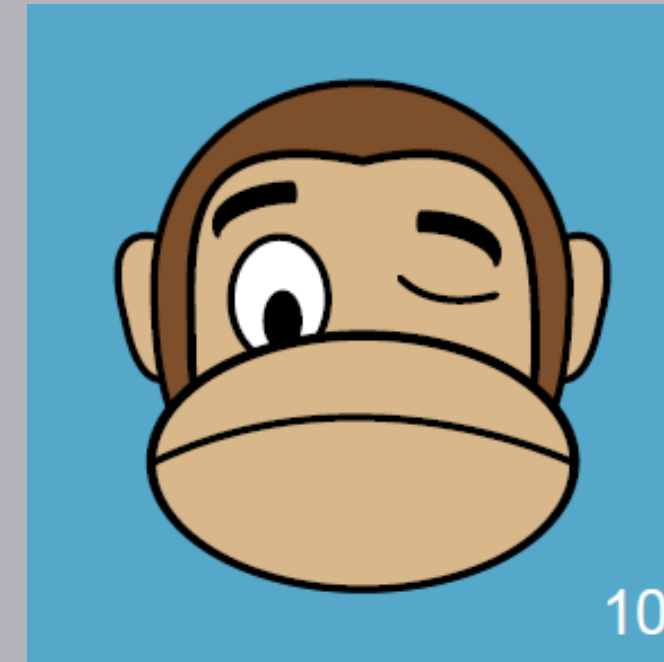
Testdaten 1: Welcher Affe beißt?



beißt



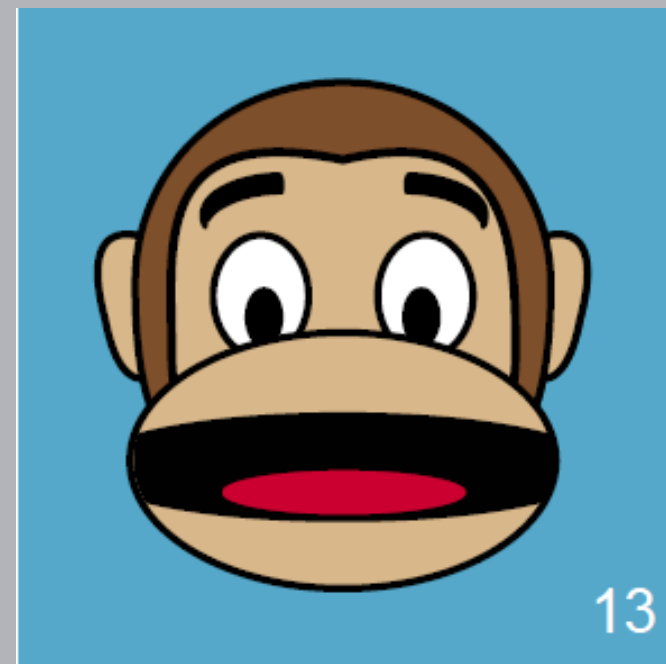
beißt



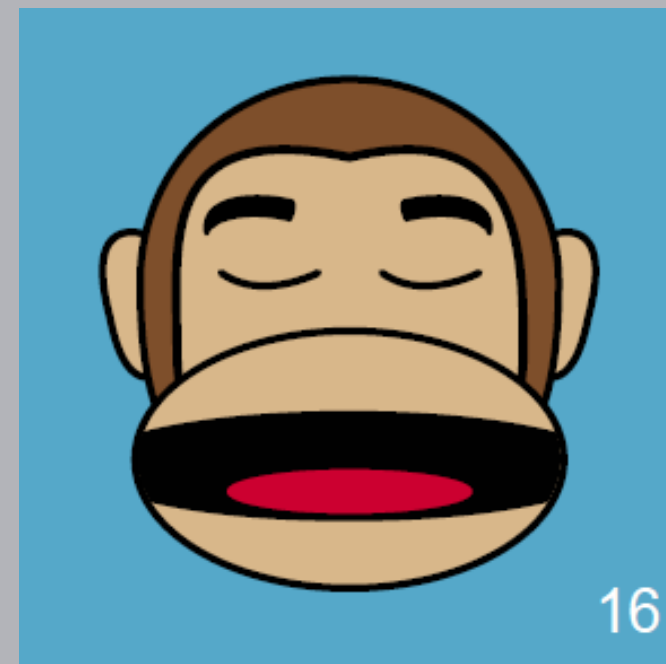
beißt nicht



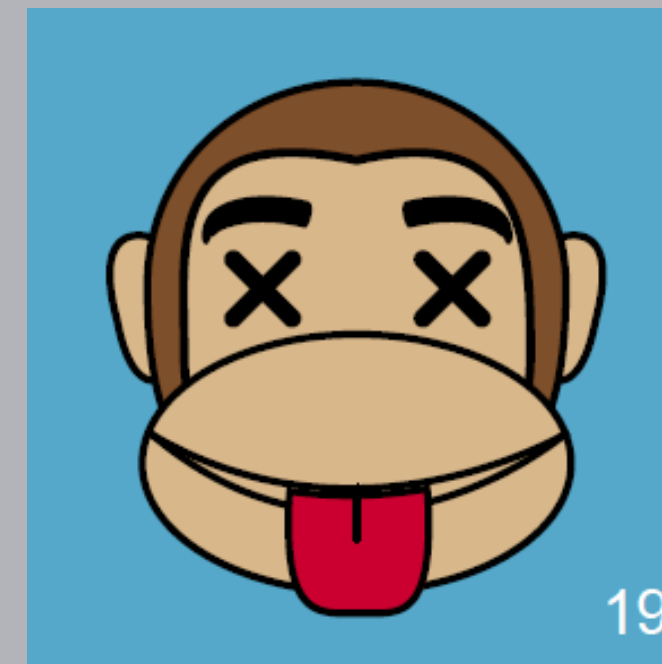
beißt



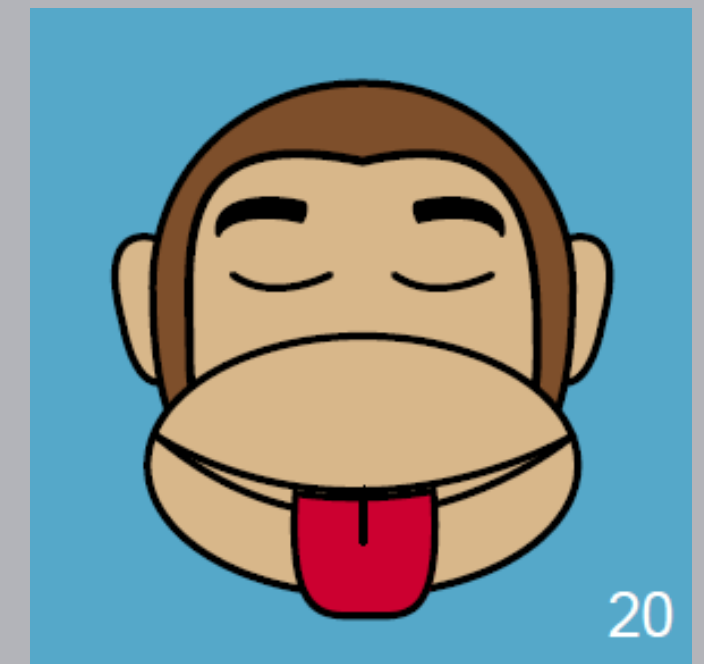
beißt nicht



beißt nicht

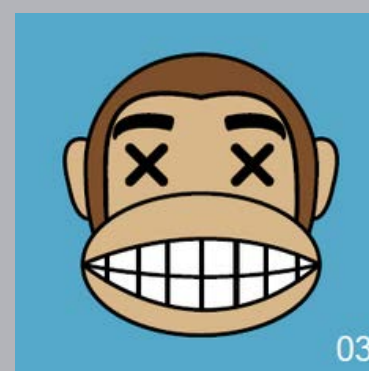
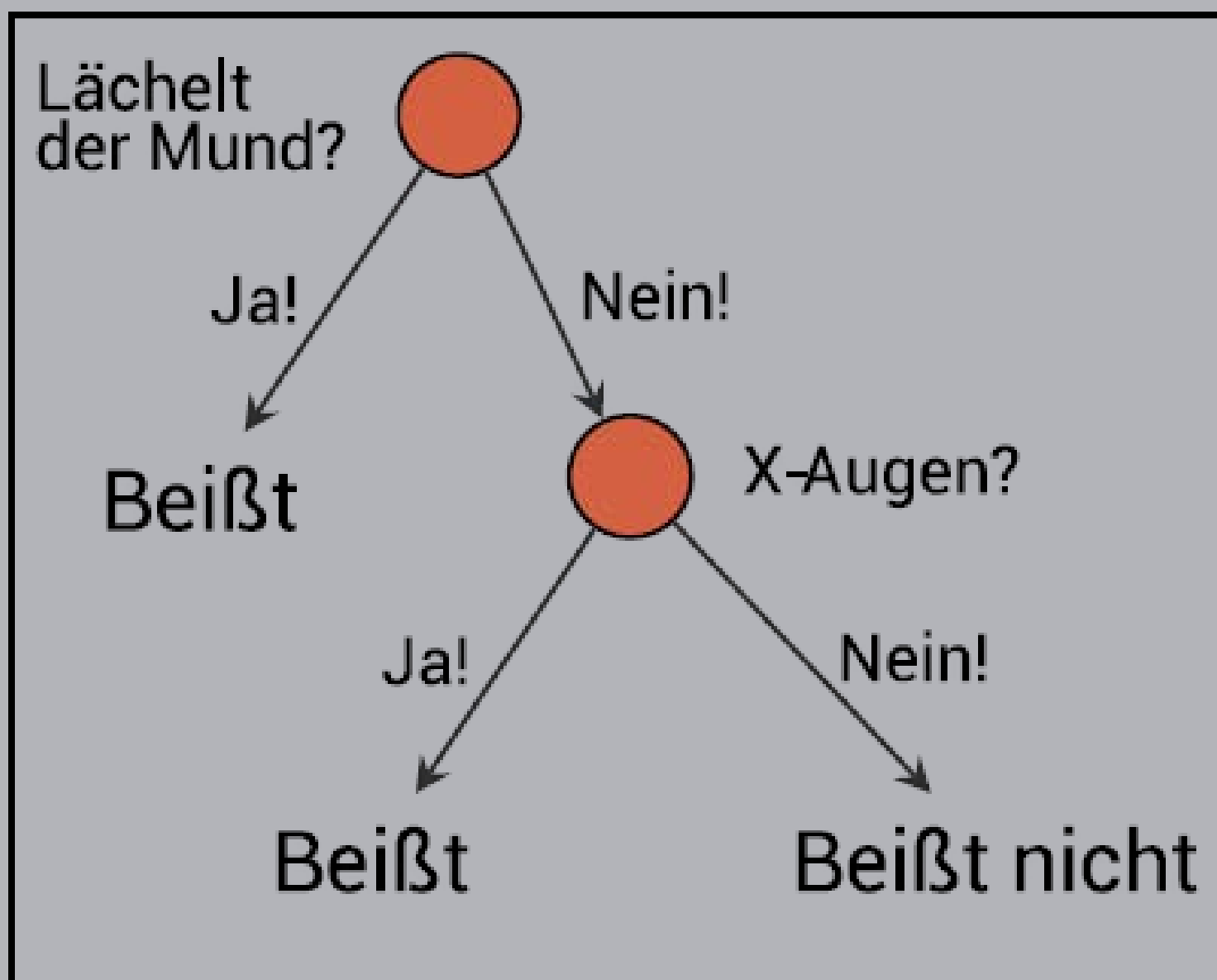


beißt

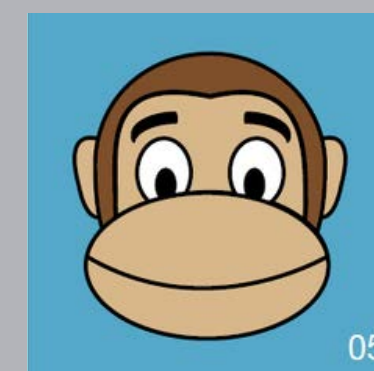


beißt nicht

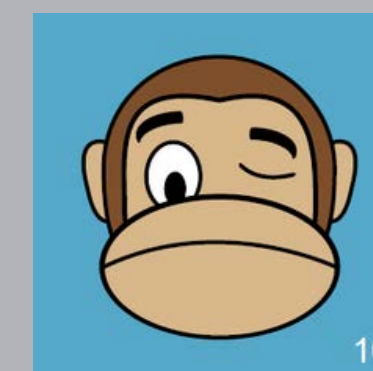
Auflösung: Kriterien für die Testdaten 1



beißt



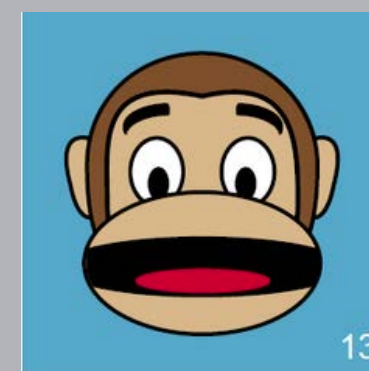
beißt



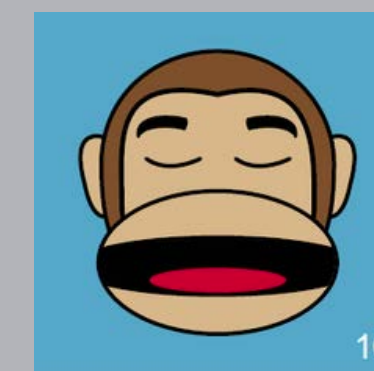
beißt nicht



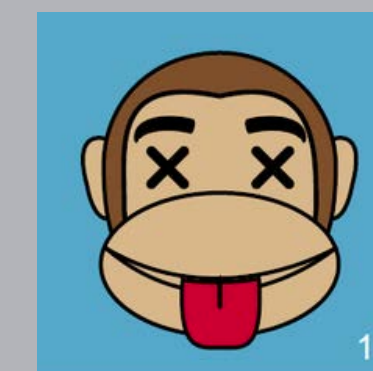
beißt



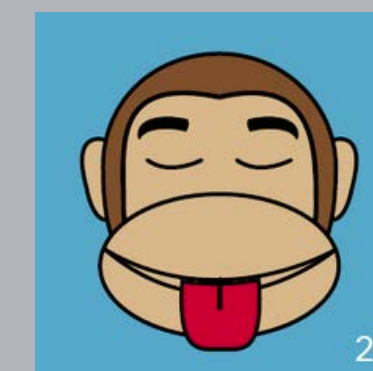
beißt nicht



beißt nicht



beißt



beißt nicht

Arbeitsauftrag Äffchenspiel - Durchgang 2

Wir sind Tierpflegerin in einem Zoo. Alle Äffchen sehen sehr süß aus, aber manche beißen. Von den Äffchen im Zoo wissen wir bereits, welche das sind.

Bald kommen aber neue Äffchen zu uns. Wir müssen überlegen wie wir herausfinden können, welche davon beißen und welche nicht.

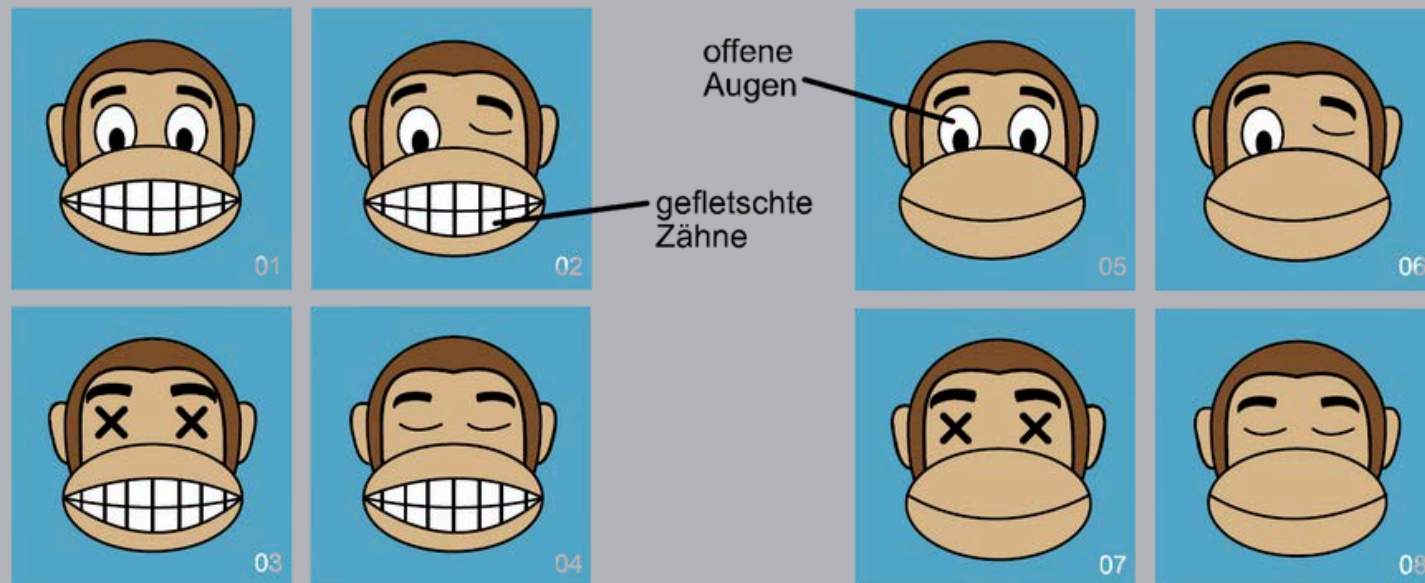


Abb. 1. In diesem einfachen Beispiel beißen alle Äffchen mit gefleitschten Zähnen

1 Gruppenbildung

Bildet Teams aus 2 bis 3 Personen.

2 Trainingsphase: Kriterien festlegen (5 Min.)

- Überlegt euch anhand der Trainingsdaten Kriterien, wie ihr beißende von nicht beißenden Äffchen unterscheiden könnt.
- Notiert die Kriterien so eindeutig, dass auch die anderen Gruppen sie verstehen.

3 Testphase: neue Äffchen zuordnen (5 Min.)

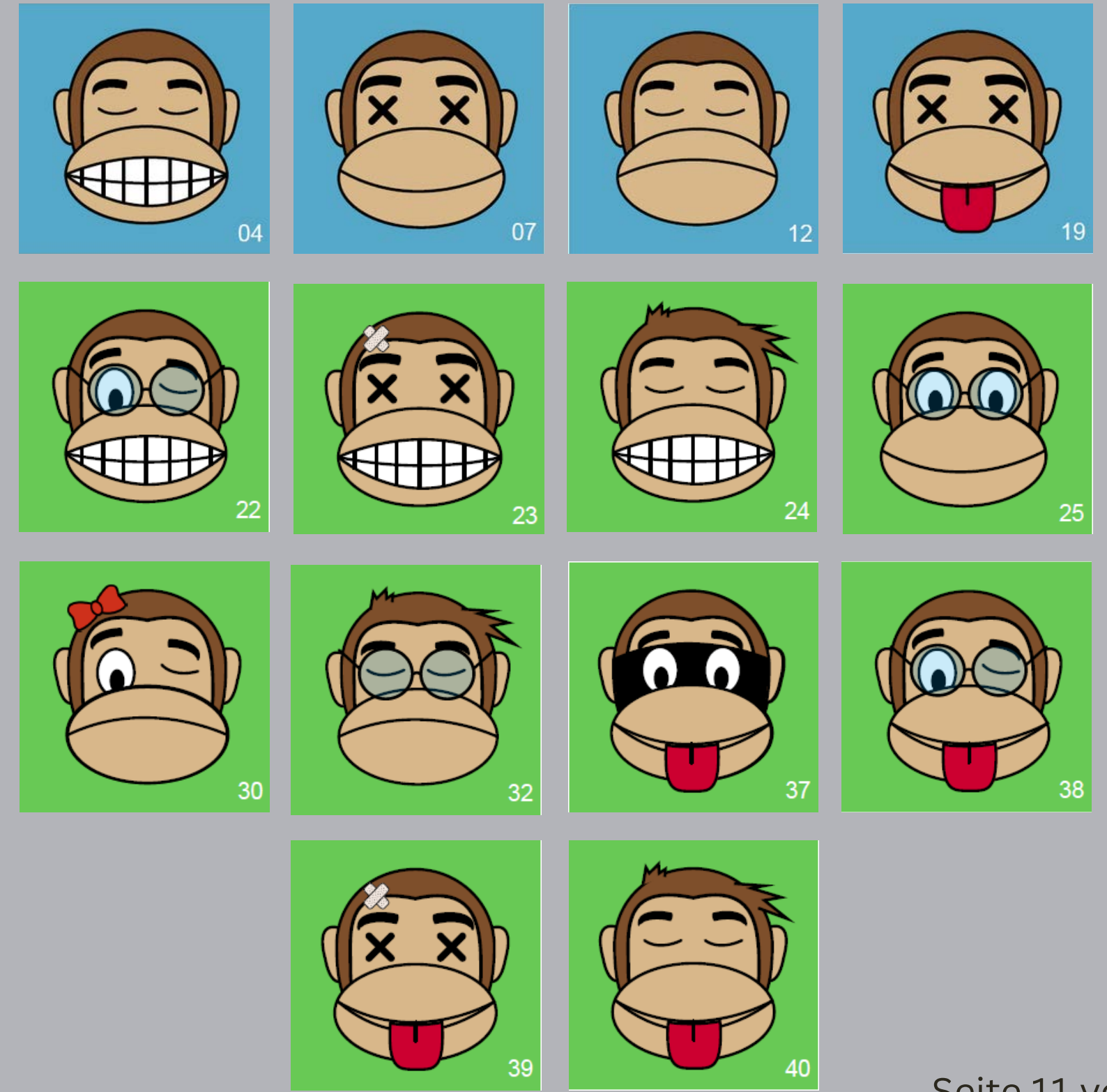
- Tauscht Eure Kriterien mit einer anderen Gruppe.
- Entscheidet anhand der Kriterien der anderen Gruppe, welche der Äffchen aus den Testdaten beißen.
- Notiert eure Entscheidungen.

Trainingsdaten 2:

Beißt



Beißt nicht



Testdaten 2: Welcher Affe beißt?



03

beißt nicht



06

beißt



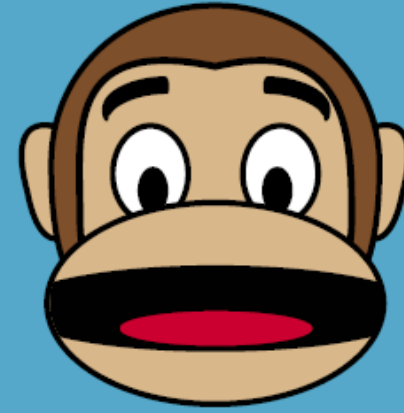
08

beißt nicht



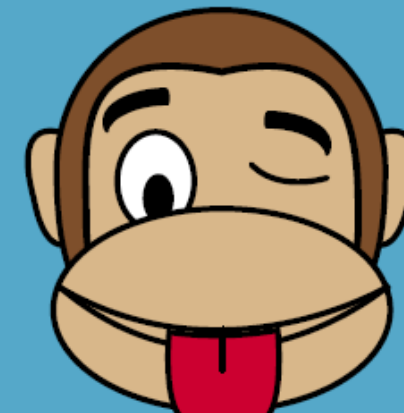
11

beißt nicht



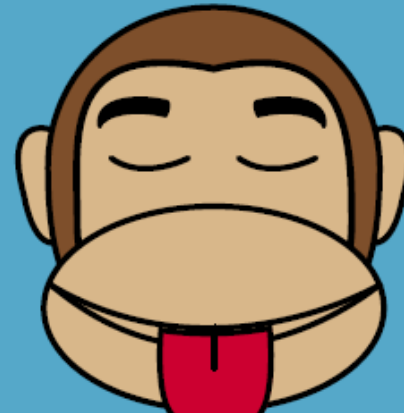
13

beißt



18

beißt



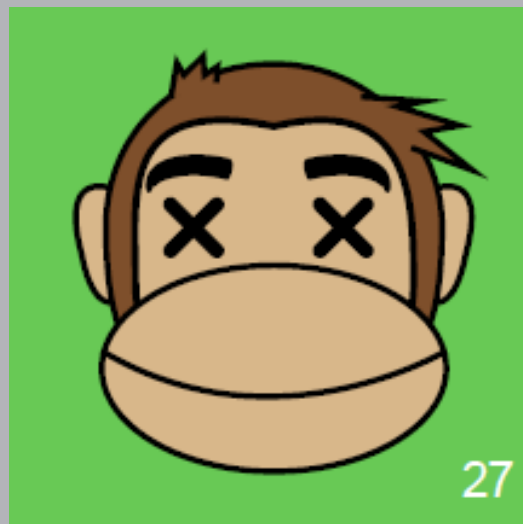
20

beißt nicht



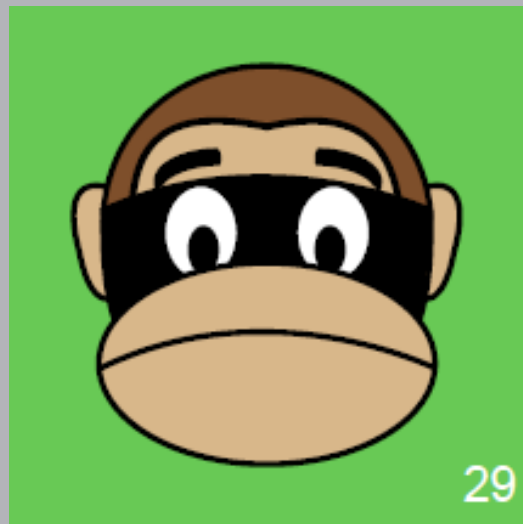
26

beißt nicht



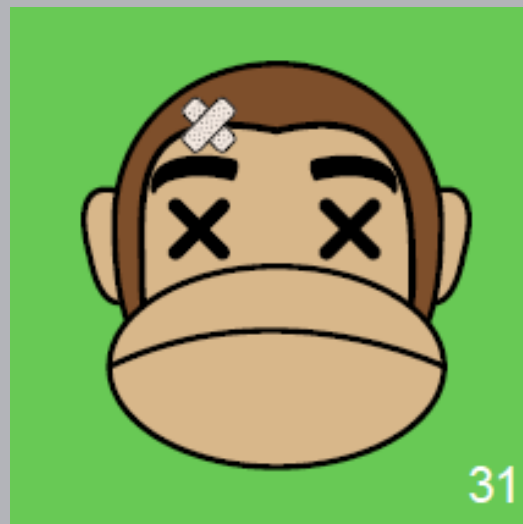
27

beißt nicht



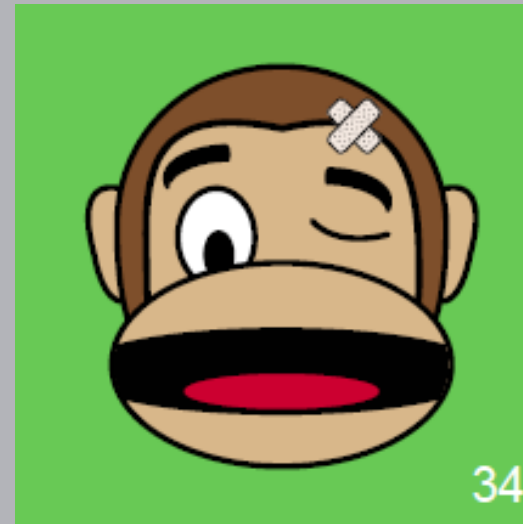
29

beißt nicht



31

beißt nicht



34

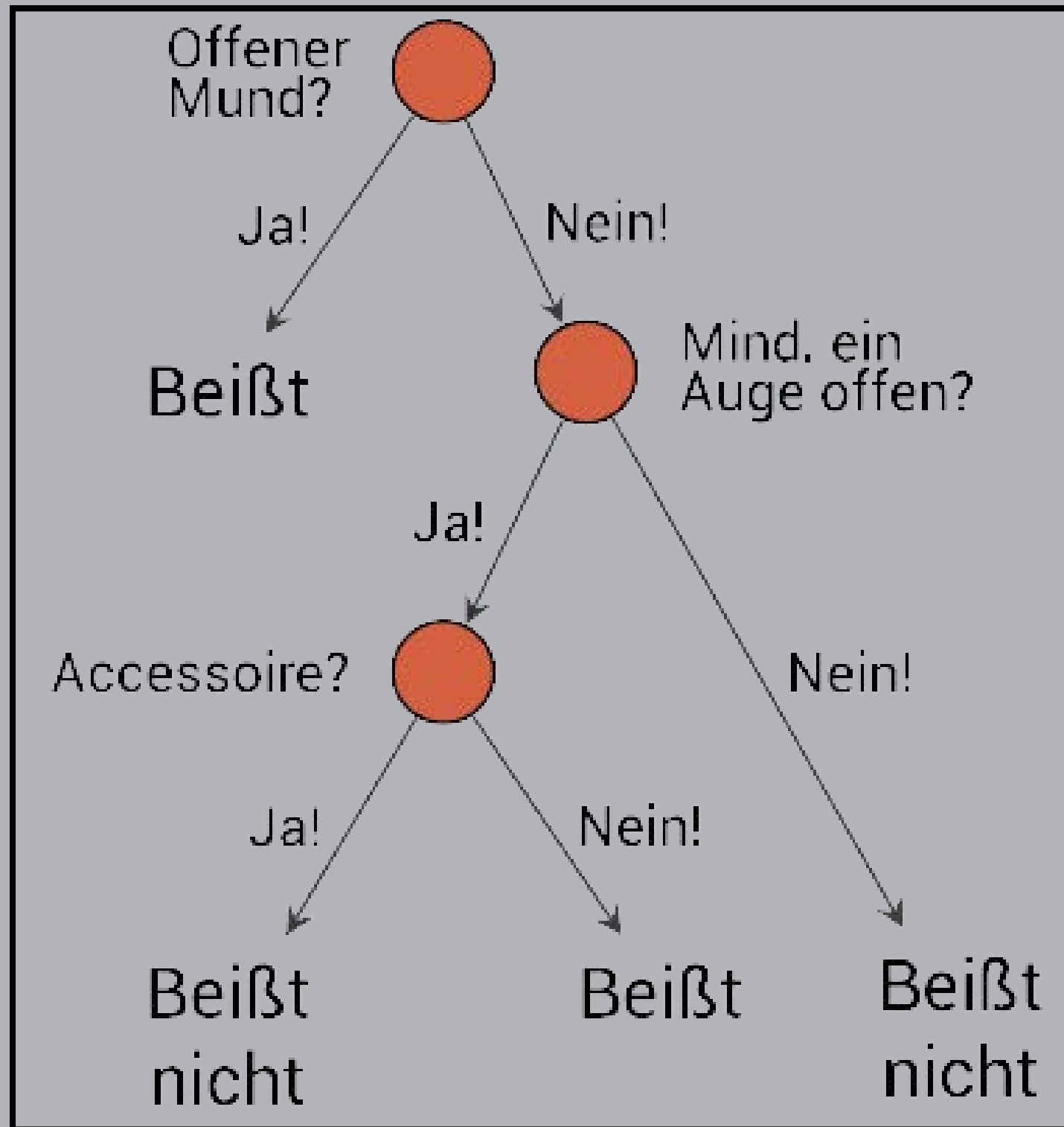
beißt



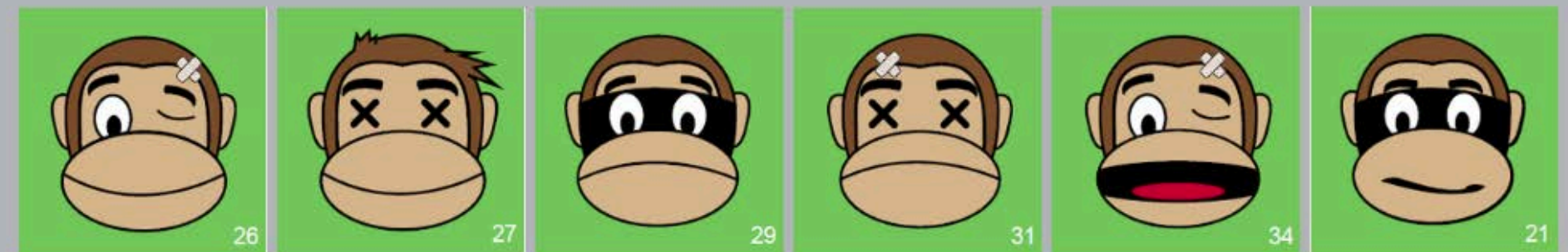
21

beißt nicht

Auflösung: Kriterien für die Testdaten 2



beißt nicht beißt beißt nicht beißt nicht beißt beißt beißt nicht

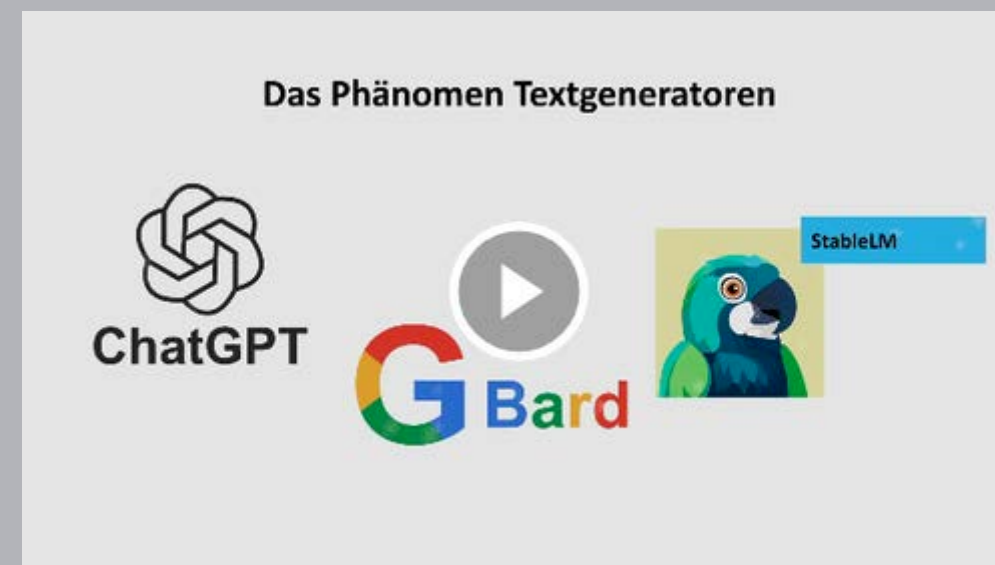


beißt nicht beißt nicht beißt nicht beißt nicht beißt beißt nicht

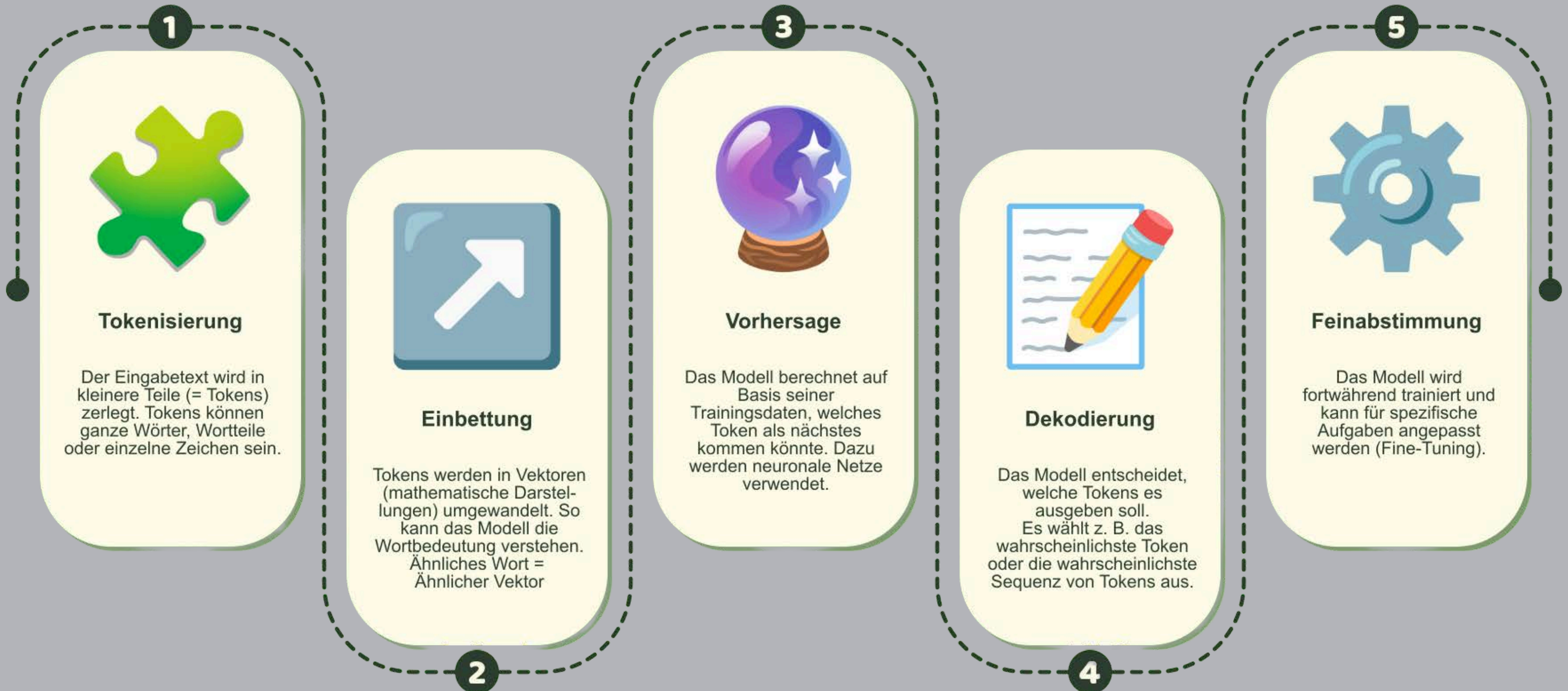
Wie funktioniert ein LLM?

Textgenerator, mit dem man hinter die Kulissen schauen kann:

Soekid_{GPT}



Wie funktioniert ein LLM?



Arbeitsaufträge zum Prompten



1 Prompten ausprobieren (10 Min.)

- Überlegt euch ein Thema, das sich mit Schule, Wirtschaft oder aktuellen Themen beschäftigt, wie etwa die Erstellung von Lernplänen, Nachhaltigkeit, Marketingstrategien oder Unternehmensgründungen.
- Formuliert **mindestens drei unterschiedliche Prompts** zu diesem Thema. Die Prompts sollen Fragen oder Anweisungen an eine KI enthalten, zum Beispiel: „Erläutere die Vorteile von E-Commerce.“
- Nutzt das KI-Tool **KI Chat** über den angegebenen Link, um eure Prompts einzugeben.
- **Wichtiger Hinweis: Gebt niemals persönliche Daten von euch oder anderen in die KI ein!**
- Kopiert die Antworten der KI in ein Word-Dokument und speichert es in eurem Ordner.

Arbeitsaufträge zum Prompten

2 Prompts überarbeiten (15 Min.)

- Überarbeitet eure Prompts basierend auf folgender Anleitung und testet sie erneut.
- Speichert und vergleicht die Antworten mit den vorherigen bezüglich ihrer Qualität.



Probleme & Gefahren bei der Nutzung von LLMs

Halluzinationen:

- Ein überzeugend klingendes Ergebnis ist falsch bzw. weitestgehend frei erfunden.
- Überprüfe die Ergebnisse daher immer auf Plausibilität und fordere die KI auf, die verwendeten Quellen anzuzeigen.



Datenschutz:



- Viele LLMs verwenden deine Eingaben, um die KI weiter zu trainieren oder um Daten anderweitig weiterzuverarbeiten.
- Berechnungen erfolgen oft auf ausländischen Servern, wodurch personenbezogene Daten möglicherweise unzureichend geschützt sind.
- Gib deshalb niemals personenbezogene Daten von dir oder anderen in die KI ein!

Training beeinflusst Ergebnisse

- Jedes LLM wird mit unterschiedlichen Trainingsdaten gefüttert und erhält unter Umständen Vorgaben für seine Antworten.
- Daher sind die Antworten der KI nicht immer neutral; so können z. B. geschichtliche Darstellungen je nach Herkunftsland der KI variieren.

