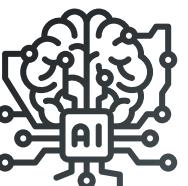


# Explainable Visual Question Answering (VQA)

## VQA Beispiel



Welche Nummer ist auf dem Bild zu sehen?



Die Nummer 32

### Motivation

- Erklärbarkeit von Black-Box-KI-Modellen mithilfe eines generierten „Wärmebildes“
- Barrierefreiheit durch eine sprachliche Ausgabe einer zusätzlich generierten textuellen Erklärung

### Technologie

Backend



PyTorch

Frontend



Streamlit

VQA-Modell



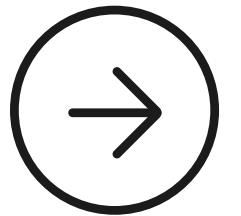
InstructBLIP

### Team

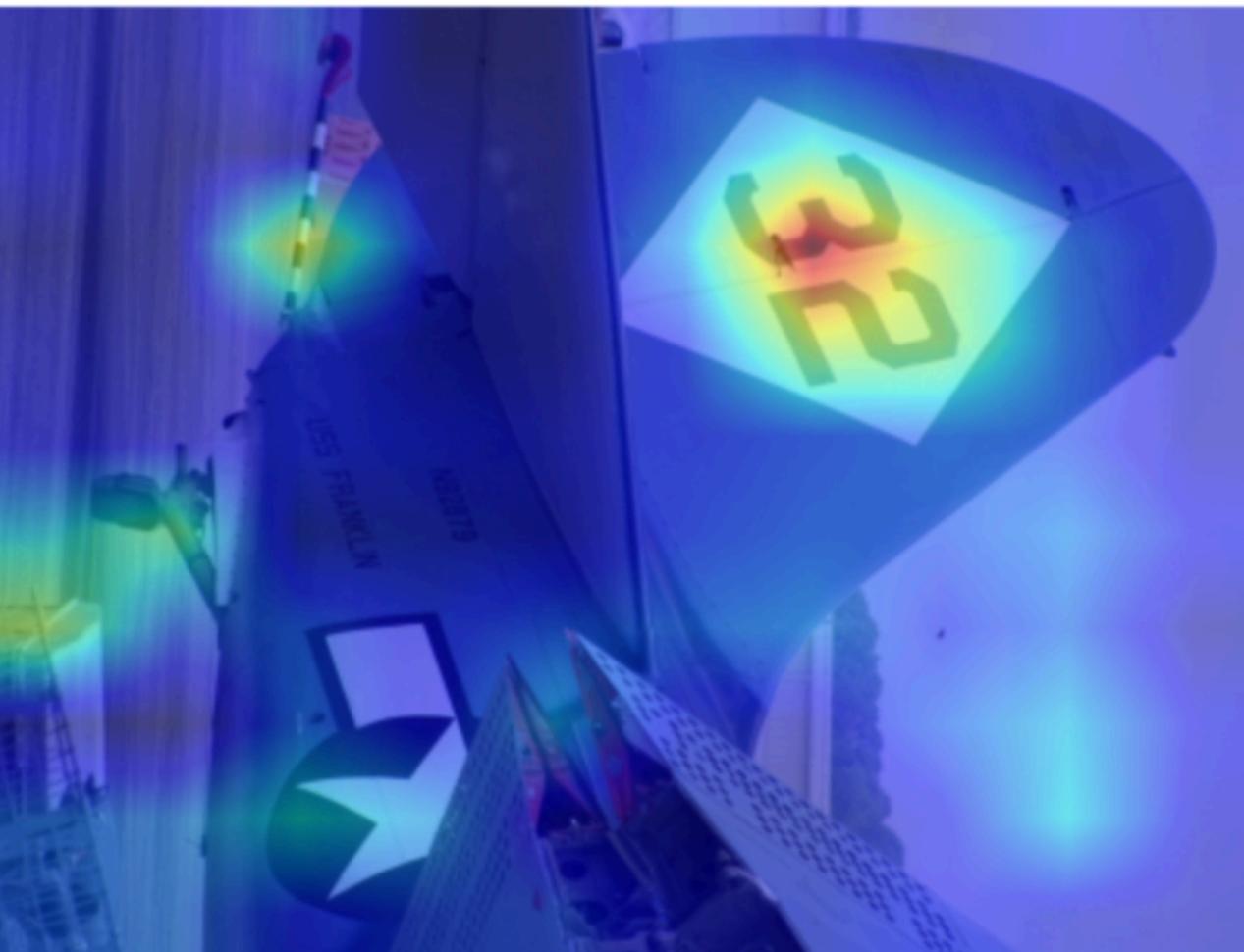
- Paula Schlötermann, CSM, 3. Semester
- Julia Lapp, CSM, 2. Semester

# Explainable Visual Question Answering (VQA)

## Visuelle Erklärung



 Welche Nummer ist auf dem Bild zu sehen?  
 Die Nummer 32

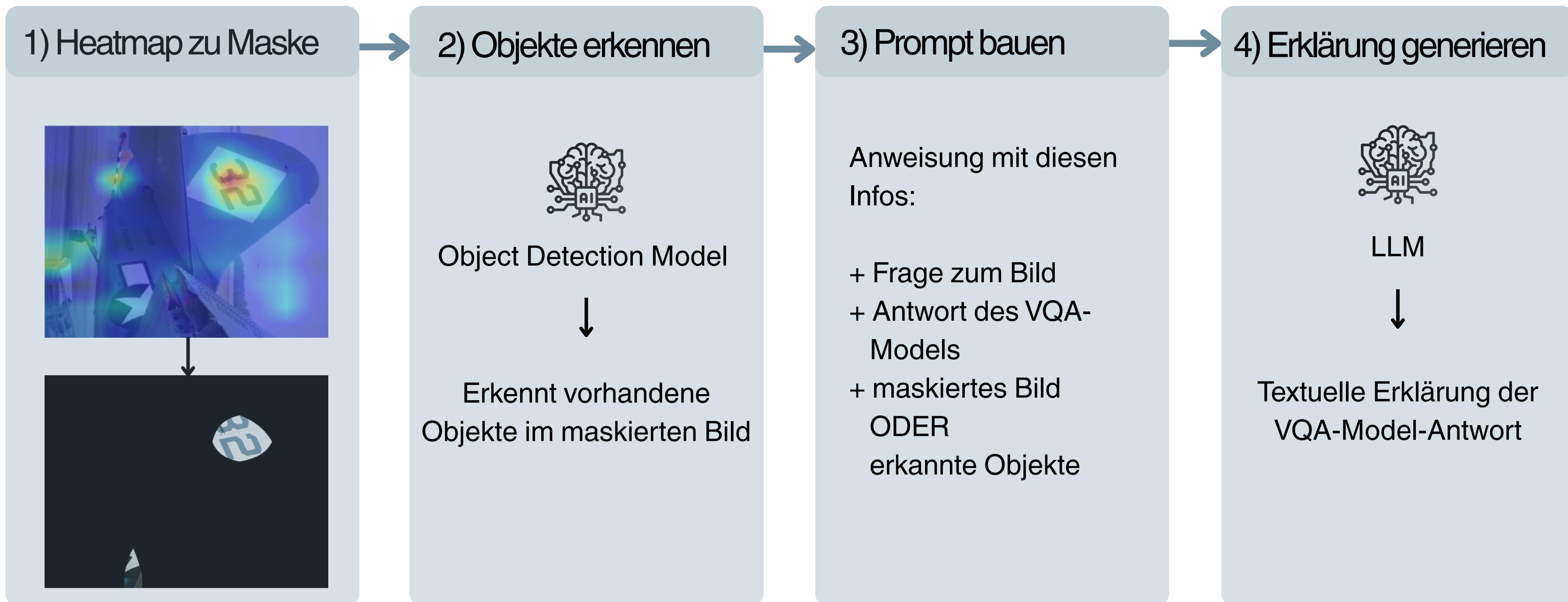


### SIDU-VQA Methode

- Systematisches Ausblenden einzelner Bildausschnitte
- Messung der Auswirkungen auf das KI-Verständnis
- Identifikation entscheidender Bildbereiche anhand von Antwortveränderungen

# Explainable Visual Question Answering (VQA)

## Textuelle Erklärung



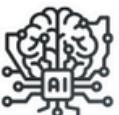
# Explainable Visual Question Answering (VQA)

## Overview

### Visual Question Answering



Welche Nummer ist auf dem Bild zu sehen?



Die Nummer 32



### 1) Visuelle Erklärung



### 2) Textuelle Erklärung

The visible part of the sign shows the digits “3” and “2,” forming the number 32.