



# Instagram comment scraper

Moritz Eßer, Jonas Heinen



## (Potentiell) zu erfassende Daten

- Post:
  - Bild
  - Bildbeschreibung
  - Likes
  - Liste der Kommentare
  
- Kommentar:
  - Verfasser (anonymisiert durch ID)
  - Kommentar-Inhalt
  - Liste der Antworten (Kommentare)
  - Likes



# Manuelles Herunterladen

- Zeitaufwand (pro Post): < 1min ohne Aufbereitung der Daten
- Vorteile: unauffällig/nicht illegal
- Nachteile: Abbild zu einem Zeitpunkt; neue Kommentare werden nicht erfasst



## Erfasste Daten: Manuelles Herunterladen

- Post:
  - Bild
  - Bildbeschreibung ✓
  - Likes ✓
  - Liste der Kommentare ✓
  - (Accessibility Caption) ✓
  
- Kommentar:
  - Verfasser (anonymisiert durch ID) ✓
  - Kommentar-Inhalt ✓
  - Liste der Antworten (Kommentare) ✓
  - Likes ✓



# Automatisierung: Idee

- Umsetzung: python, selenium
- Vorteile: schnell, automatisiert, vollständig
- Nachteile: Anti-Bot-Maßnahmen, Bugs(?)



# Automatisierung: Umsetzung

- Zeitaufwand pro Post: 5.14s - 300.00s
- Vorteile : wiederverwendbar, automatisiert
- Nachteile:
  - Entwicklung: Zeitaufwendig
  - Unvorhergesehenes Verhalten : <https://www.instagram.com/p/CWbStR1jvE/>
  - HTTP 429 : Too many Requests



# Automatisierung: Demo



**Github**

<https://github.com/Zadest/insta-comment-extraction>