

## Die Datenscouts

### **Fallschilderung:**

Der Video-Streaming-Dienst *MyTube* ist ein großes Online-Unternehmen, das in seinem Online-Portal Videoclips zu vielen Themenbereichen anbietet. Ein wichtiges Geschäftsziel des Unternehmens ist es, mehr über die Interessen und Vorlieben seiner Nutzer:innen zu erfahren, um ihnen passende und nützliche Videos zu empfehlen.

Dazu greift *MyTube* auf die Daten seiner Nutzer:innen zurück, die diese bei der Nutzung hinterlassen, wenn sie z.B. ein Video anklicken und sich dieses ansehen. Diese Daten werden in einem Datensatz gesammelt, der auch als Nutzer-Historie bezeichnet wird. Über die Nutzer-Historie erfährt das Unternehmen zum Beispiel, wie lange oder zu welcher Tageszeit sich ein Nutzer ein Video ansieht oder welches Thema ihn beim Videokonsum interessiert.

Auf Grundlage solcher Datensätze führt *MyTube* Datenanalysen durch und erhält dadurch individuelle Steckbriefe über seine Nutzer:innen. Aus diesen Steckbriefen lässt sich z.B. ablesen, welche Interessen, Hobbys, Vorlieben oder alltäglichen Nutzungsgewohnheiten eine Person hat.

Bei einem unglücklichen Vorfall sind dem Unternehmen jedoch einige Steckbriefe durcheinandergeraten, sodass unklar ist, welcher Datensatz zu welchem Steckbrief gehört.

### **Aufgabe:**

**Nun seid Ihr als Datenscouts gefordert. Wie könnt Ihr dem Unternehmen helfen, die Datensätze den richtigen Steckbriefen zuzuordnen?**

### **Fragen:**

Falls notwendig, kann die anschließende Diskussion mit diesen Fragen angeregt werden:

- Welche Plattformen nutzen die Schüler:innen selbst?
- Welche Daten werden bei der Nutzung erzeugt?
- Was verraten die Nutzungsdaten über das Leben und den Alltag einer Person?
- Was wissen Online-Unternehmen über ihre Nutzer:innen?
- Welche Nutzungsdaten sind aussagekräftig, welche weniger?
- Wie könnten Lösungsansätze für die Bearbeitung der Aufgabe aussehen?
- Was sind geeignete Herangehensweisen, um größere Datensätze zu analysieren?

