

## Aufgabe 1: Kardinalitätsschätzung (10P)

Abgabe-Deadline: 2023-11-14 8:00 im ELO

Gegeben sind die Tabellen eines Videoportals:

Videos(vnr, titel, genre, sprache)

Schlagworte(vnr, schlagwort)

Schlagworte.vnr ist ein Fremdschlüssel Videos.vnr

Die Kardinalitäten der Tabellen bzw. Spalten sind wie folgt:

$|Videos| = 100.000.000$ ,  $|Schlagworte| = 400.000.000$

$|Videos.vnr| = 100.000.000$ ,  $|Videos.titel| = 50.000.000$

$|Videos.genre| = 20$ ,  $|Videos.sprache| = 25$

$|Schlagworte.vnr| = 50.000.000$ ,  $|Schlagworte.schlagwort| = 400$

Führen Sie eine Schätzung durch, wie viele Zeilen die folgenden Ausdrücke liefern (je 2P).  
Gehen Sie von Gleichverteilung und Werteunabhängigkeit zwischen den Attributen aus.

1.  $\pi_{vnr}(Schlagworte)$
2.  $\sigma_{genre='Musik' \vee genre='Musikvideo'}(Videos)$
3.  $\sigma_{genre='Musik' \wedge sprache='Deutsch'}(Videos)$
4.  $Videos \bowtie_{Videos.vnr = Schlagworte.vnr} Schlagworte$
5.  $\pi_{vnr, titel} \sigma_{Schlagworte.schlagwort='lustig'}(Videos \bowtie_{Videos.vnr = Schlagworte.vnr} Schlagworte)$

## Aufgabe 2: Anfrageoptimierung

Überführen Sie diese Anfrage in eine äquivalente Anfrage gemäß den Heuristiken zur Anfrageoptimierung aus der Vorlesung:

$\pi_{schlagwort} \sigma_{genre='Musik'} \sigma_{Videos.vnr = schlagworte.vnr} (Videos \times Schlagworte)$

$\cup \pi_{schlagwort} \sigma_{genre='Musikvideo'} \sigma_{Videos.vnr = schlagworte.vnr} (Videos \times Schlagworte)$

## Aufgabe 3: Relationale Algebra

1. Formulieren Sie in Ihren eigenen Worten, was die Anfrage 5 aus Aufgabe 1 liefert.
2. Formulieren Sie in Ihren eigenen Worten, was die Anfrage aus Aufgabe 2 liefert.
3. Formulieren Sie auf dem Relationenschema von Aufgabe 1 einen Ausdruck der relationalen Algebra, der Ihnen die Titel von derjenigen Videos ausgibt, die mindestens ein Schlagwort haben, welches auch das Video mit der VNR 3517 besitzt.