

## Aufgabe 4 – Ranking

Gegeben ist die Teilmenge einer Term-Dokument-Matrix, welche die jeweilige Anzahl der Terme in den Dokumenten enthält. IDF ist ebenfalls gegeben. Die folgende Abfrage wird von einem Benutzer eingegeben: "beta dome". Folgenden Dokumente sind für den Nutzer relevant: Dok. 2, Dok. 3.

terms/docs	ant	beta	camp	dome	end	frank	golf
<b>Doc 1</b>	0	3	1	0	1	1	1
<b>Doc 2</b>	1	2	1	3	2	2	2
<b>Doc 3</b>	0	0	0	1	9	1	4
<b>Doc 4</b>	3	4	3	0	6	0	0
<b>Doc 5</b>	7	8	0	0	6	4	9

	<b>idf</b>	<b>log (idf)</b>	<b>normalized (log(idf))</b>
ant	240	2.4	0.4
beta	120	1.2	0.2
camp	1000	3.0	0.5
dome	260000	5.4	0.9
end	900000	6.0	1
frank	4000	3.6	0.6
golf	10	1.0	1/6

- Berechnen Sie das Ranking basierend auf der **tf (ohne Logarithmus)**
- Berechnen Sie ein neues Ranking, welches auf **tf\*normalized(log(idf))** basiert.
- Welches Ranking ist besser?