

# Digital Libraries WS 2018/2019

## Übungsblatt 4

Aaron Winziers - 1176638; Michael Wolz - 1195270

26. November 2018

### Aufgabe 1

a)

Term	$D_1$	$D_2$	$D_3$	$D_4$
Weihnachten	1	1	1	1
Schnee	1	0	1	0
Skifahren	0	1	0	0
Apfelstrudel	0	0	0	1
Eislaufen	1	0	0	0
Seilbahn	0	1	0	0
Langlauf	0	0	1	0
Zimt	0	0	0	1
Glühwein	1	1	0	0
Eis	0	0	1	1
Punsch	0	1	0	0

**b)**

Term		Dokumente
Weihnachten	$\Rightarrow$	1,2,3,4
Schnee	$\Rightarrow$	1,3
Skifahren	$\Rightarrow$	2
Apfelstrudel	$\Rightarrow$	4
Eislaufen	$\Rightarrow$	1
Seilbahn	$\Rightarrow$	2
Langlauf	$\Rightarrow$	3
Zimt	$\Rightarrow$	4
Glühwein	$\Rightarrow$	1,2
Eis	$\Rightarrow$	3,4
Punsch	$\Rightarrow$	2

**c)**

$$\textit{Weihnachten} = \{1,2,3,4\}$$

$$\textit{Eis} = \{3,4\}$$

$$\textit{Punsch} = \{2\}$$

$$\textit{Eis} \textbf{ OR } \textit{Punsch} = \{3,4\} \cup \{2\} = \{2,3,4\}$$

$$\textit{Weihnachten} \textbf{ AND } (\textit{Eis} \textbf{ OR } \textit{Punsch}) = \{1,2,3,4\} \cap \{2,3,4\} = \{2,3,4\}$$

**d)**

Es werden Weihnachtliche Eis- oder Punsch-Rezepte gesucht.  $D_2, D_3, D_4$  sind die Dokumente die zurückgegeben werden, jedoch ist  $D_4$  das einzige relevante Dokument.  $D_2$  ist nicht relevant da es vermutlich Informationen über z.B. ein Skigebiet beinhaltet wo es Stände oder Hütten gibt wo Glühwein angeboten wird und  $D_3$  weil das Eis-Term nicht auf Speiseeis bezogen ist. Daher:

$$\text{Precision} = \frac{|\{D_4\}|}{|\{D_4\} \cup \{D_2, D_3\}|} = \frac{1}{3}$$

$$\text{Recall} = \frac{|\{D_4\}|}{|\{D_4\}|} = 1$$

Da  $D_2$  und  $D_3$  aus den oben genannten Gründen nicht relevant sind zum Informationsbedürfniss aber trotzdem als relevant geliefert werden, ist Recall nicht gleich 1

e)

Angenommen die angegebene Dokumente umfassen alle die zu durchsuchen sind, ist der Term "Weihnachten in allen suchen nicht relevant. In jeder möglichen Suche wird die Inklusion des Terms dazu führen dass entweder alle oder durch das Einsetzen von einem **NICHT** keine Dokumente zurückgegeben. Sollten es andere Dokumente geben, wird der Term sicher stellen dass Dokumente 1 bis 4 immer oder nie mitgeliefert werden in einer Suche.