

# Esercitazione

## 14/05/2020

1. Caricamento dei dati sulle definizioni (file definizioni.xls o google doc presente su Moodle)
2. Preprocessing (su frequenza minima dei termini, stemming, etc. a vostra scelta)
3. Calcolo similarità tra definizioni (cardinalità dell'intersezione dei termini normalizzata su lunghezza minima tra le due, o varianti a scelta)
4. Aggregazione sulle due dimensioni (concretezza / specificità come da schema in basso)
5. Interpretazione dei risultati e scrittura di un piccolo report (da inserire nel vostro portfolio per l'esame)

	Astratto	Concreto
Generico	%	%
Specifico	%	%

Valori medi similarità

- Esperimento usando filtri e preprocessing a vostra scelta

(es. usando stemming e stopword removal come preprocessing)