$ec{Q}_{m} = a\,ec{Q}_{o} + b\,rac{1}{|D_{r}|}\sum_{ec{D}_{i}\in D_{r}}ec{D}_{j} - c\,rac{1}{|D_{nr}|}\sum_{ec{D}_{k}\in D_{nr}}ec{D}_{k}$ 

# Rocchio Algoritmen

Av Adrian, Una and Øystein

## Hva blir det brukt til?

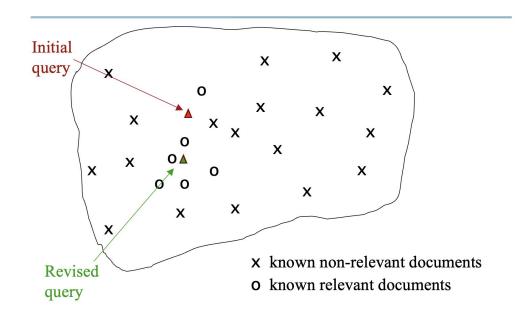
Standard algoritme for å inkorporere relevance feedback

Query ekspansjon, utvide søk til å gi mer relvant informasjon

Kan forbedre recall

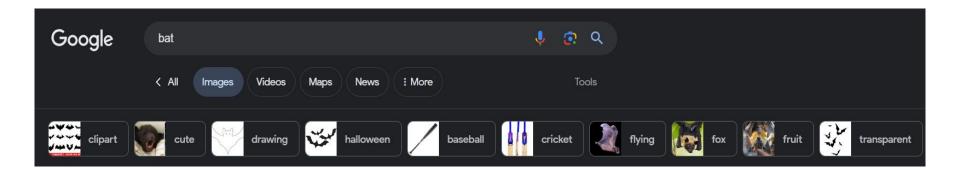
## Hvordan?

- Brukeren sender inn en enkel spørring.
- Systemet returnere resultater
- Brukeren markerer resultatene som relevante eller irrelevante
- Systemet beregner en bedre representasjon av informasjonen basert på brukerens tilbakemelding
- Systemet viser et revidert versjon av resultatene.



## Query expansion

Brukeren kan se og velge utvidelsesbegrepene(expansion terms).



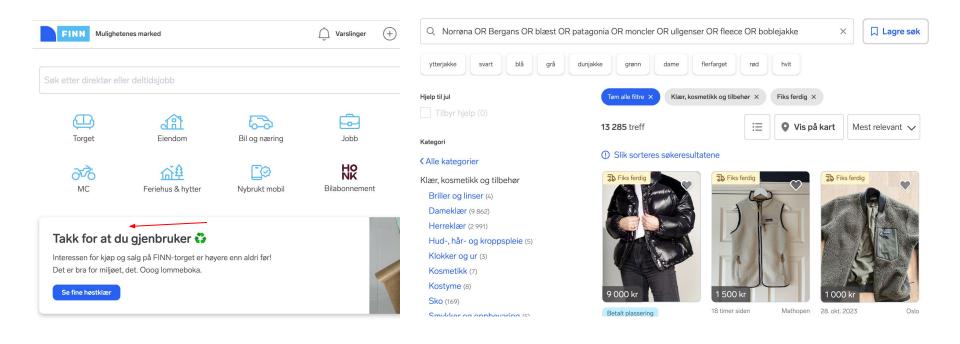
### Metoder

Selvlaget ordliste vedlikeholdt av menneskelige redaktører (f.eks. nøkkelord for publikasjoner som MedLine).

En håndlaget ordbok (f.eks. WordNet).

En automatisk generert ordbok.

Endring av søketeksten basert på søkelogganalyse (dvs. hva store søkemotorer gjør).



## Begrensninger

- Mismatch i brukerens og samlingens vokabular
- Eksempel: søker "laptop" og samlingen skriver "bærbar datamaskin"
- Ofte nødvendig med en domene-spesifikk ordbok