

67-300 SEARCH ENGINES

---

# QUERY REFORMULATION AND RELEVANCE FEEDBACK

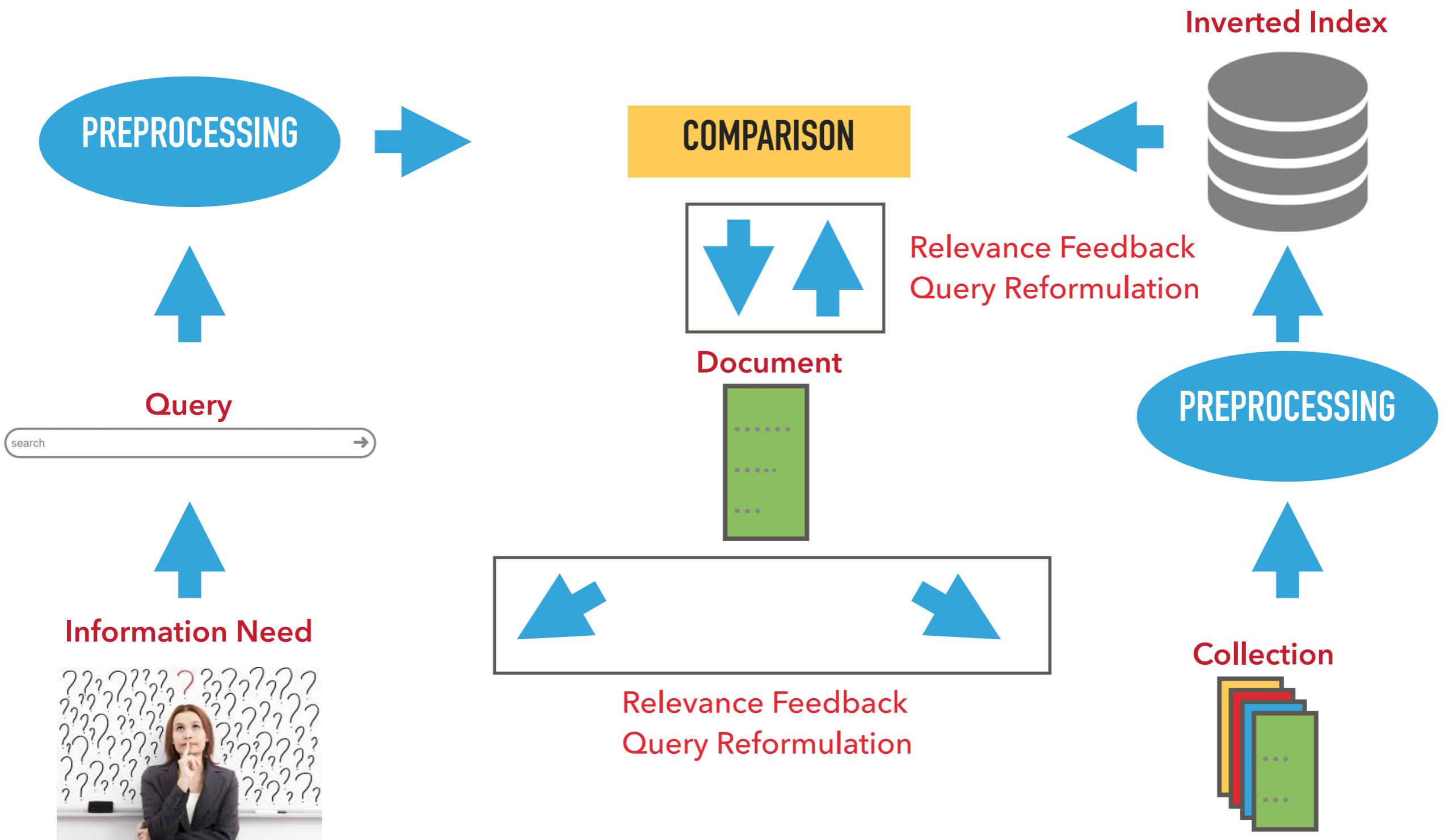
LECTURER: JOAO PALOTTI ([JPALOTTI@ANDREW.CMU.EDU](mailto:JPALOTTI@ANDREW.CMU.EDU))

29TH MARCH 2017

### LECTURE GOALS

- ▶ Query reformulation:
  - ▶ Global Methods
  - ▶ Local Methods
- ▶ Discussion
- ▶ Assignment 2

# LECTURE 6 - QUERY REFORMULATION AND RELEVANCE FEEDBACK



**Query**

 →

**Information Need**



---

# MOTIVATION



## ALL BEGAN WITH SYNONYMS



- ▶ Many concepts can be expressed in many different ways:
  - ▶ A search for **aircraft** could also find documents using **airplane**
  - ▶ A search for **cats** could also find documents using **kitten** or **feline**
  - ▶ A search for **heart attack** could also find documents using **myocardial infarction**



## ALL BEGAN WITH SYONYMS



- ▶ Many concepts can be expressed in many different ways:
  - ▶ A search for **aircraft** could also find documents using **airplane**
  - ▶ A search for **cats** could also find documents using **kitten** or **feline**
  - ▶ A search for **heart attack** could also find documents using **myocardial infarction**

**CAN WE TRANSFORM A QUERY Q  
INTO Q' THAT IS MORE EFFECTIVE?**



# THE TECHNIQUES FOR QUERY REFORMULATION

- ▶ There are two main families:
  - ▶ Global Methods (**do not look at result set**)
    - ▶ Expansion/reformulation with external thesaurus
    - ▶ Spelling correction
  - ▶ Local Methods (**look at the result set**)
    - ▶ Explicit Relevance feedback
    - ▶ (Implicit/Blind/Pseudo) Relevance feedback

---

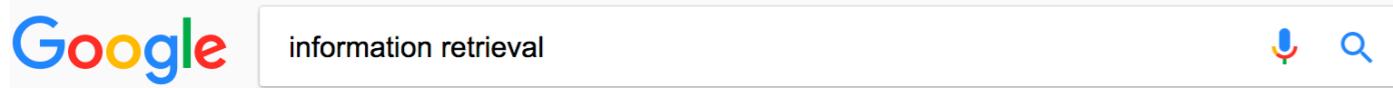
# GLOBAL METHODS

# GLOBAL METHODS

- ▶ We want to modify/reformulate a query:
  - ▶ Adding words not used in the query that are likely to increase user happiness...
  - ▶ Rephrasing some words to their synonyms...
- ▶ How can we do it?
  - ▶ Asking the user?
  - ▶ Automagically changing the queries?

## LECTURE 6 - QUERY REFORMULATION AND RELEVANCE FEEDBACK

# ASKING THE USER...



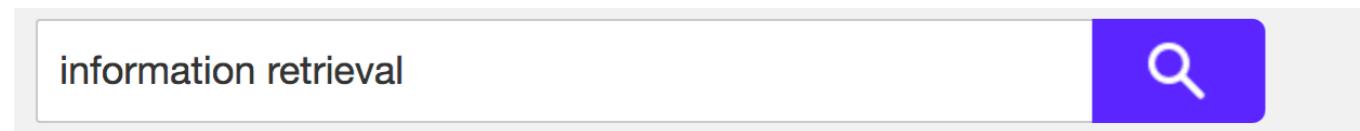
Searches related to information retrieval

information retrieval **pdf**      information retrieval **course**  
information retrieval **system**      information retrieval **techniques**  
information retrieval **journal**      information retrieval **ppt**  
information retrieval **book**      information retrieval **stanford**



Searches related to search engines

search **engine list**      **types of search engine**  
**internet** search engines      top 50 search engines  
search engines **movie**      alternative search engines  
**top** search engines      best search engines



Web    Images    Video    More ▾    Anytime ▾

Also try: information retrieval system, music information retrieval

### Also Try

information retrieval system  
music information retrieval  
introduction to information retrieval  
past performance information retrieval  
system

information retrieval brazil  
yahoo information retrieval  
information retrieval book  
information retrieval course

# LECTURE 6 - QUERY REFORMULATION AND RELEVANCE FEEDBACK

## NOT ASKING THE USER...



Web Images News

2,240,000 RESULTS Date Language Region

### Information retrieval - Wikipedia

[https://en.wikipedia.org/wiki/Information\\_retrieval](https://en.wikipedia.org/wiki/Information_retrieval) ▾

Information retrieval (IR) is the activity of obtaining information resources relevant to an information need from a collection of information resources.

[Overview](#) · [History](#) · [Model types](#) · [Performance and ...](#) · [Timeline](#)

### Introduction to Information Retrieval - Stanford NLP Group

[nlp.stanford.edu/IR-book](http://nlp.stanford.edu/IR-book) ▾

Introduction to Information Retrieval. This is the companion website for the following book. Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze ...

### Information Retrieval Journal - Springer

[link.springer.com/journal/10791](http://link.springer.com/journal/10791)

The journal provides an international forum for the publication of theory, algorithms, and experiments across the broad area of information retrieval.

### INFORMATION RETRIEVAL - Computing Science

[www.dcs.gla.ac.uk/Keith/Preface.html](http://www.dcs.gla.ac.uk/Keith/Preface.html) ▾

PREFACE TO THE FIRST EDITION (London: Butterworths, 1975) The material of this book is aimed at advanced undergraduate information (or computer) science students ...

### Information Retrieval – incl. option to publish open access

[www.springer.com/computer/database+management+&+information...](http://www.springer.com/computer/database+management+&+information...) ▾

The essential forum for theory, algorithms and experimentation in Information Retrieval and its Applications

### Information retrieval - Wikiquote

[https://en.wikiquote.org/wiki/Information\\_retrieval](https://en.wikiquote.org/wiki/Information_retrieval) ▾

Information retrieval is the activity of obtaining information resources relevant to an information need from a collection of information resources, and the part of ...

### Precision and recall - Wikipedia

[https://en.wikipedia.org/wiki/Recall\\_\(information\\_retrieval\)](https://en.wikipedia.org/wiki/Recall_(information_retrieval)) ▾

In pattern recognition and information retrieval binary classification, precision (also called positive predictive value) is the fraction of retrieved instances that ...

[Introduction](#) · [Definition ...](#) · [Definition ...](#) · [Probabilistic ...](#) · [F-measure](#)

### How Information Is Retrieved From Memory

<https://www.verywell.com/memory-retrieval-2795007> ▾

In order to use information, it must first be retrieved from memory. Learn more about memory retrieval as well as possible problems with this process.



Web Images News

48,800,000 RESULTS Date Language Region

### Dogpile - Official Site

[www.dogpile.com](http://www.dogpile.com) ▾

Dogpile.com makes searching the Web easy, because it has all the best search engines piled into one. Go Fetch!

### Google

[www.google.com](http://www.google.com) ▾

Search the world's information, including webpages, images, videos and more. Google has many special features to help you find exactly what you're looking for.

### Yahoo Search - Web Search

[search.yahoo.com](http://search.yahoo.com) ▾

The search engine that helps you find exactly what you're looking for. Find the most relevant information, video, images, and answers from all across the Web.

### Images of search engines

[bing.com/images](http://bing.com/images)



[See more images of search engines](#)

### Web search engine - Wikipedia

[https://en.wikipedia.org/wiki/Web\\_search\\_engine](https://en.wikipedia.org/wiki/Web_search_engine) ▾

A web search engine is a software system that is designed to search for information on the World Wide Web. The search results are generally presented in a ...

[History](#) · [How web search ...](#) · [Market share](#) · [Search engine bias](#)

### Ixquick - Official Site

[https://www.ixquick.com](http://https://www.ixquick.com) ▾

StartPage search engine provides search results from over ten best search engines in full privacy. Search anonymously with StartPage Search Engine!

### DuckDuckGo - Official Site

[duckduckgo.com](http://duckduckgo.com)

DuckDuckGo is the search engine that doesn't track you. We protect your search history from everyone – even us!

### Ask.com - Official Site

[www.ask.com](http://www.ask.com)

Offers search for web sites, images, news, blogs, video, maps and directions, local search and shopping.

### GLOBAL METHODS

- ▶ Query logs – There will be another lecture on query logs
- ▶ Thesaurus / Ontologies:
  - ▶ WordNet (general)
  - ▶ UMLS (medical)
  - ▶ Automatically? Manually?

### GLOBAL METHODS

- ▶ Query logs – There will be another lecture on query logs
- ▶ Thesaurus / Ontologies:
  - ▶ WordNet (general)
  - ▶ UMLS (medical)
  - ▶ Automatically? Manually?

**SUPPOSE WE SEARCH FOR “DOG”...**

## WORD NET

- ▶ Provides extensive word definition:

WordNet Search - 3.1  
- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations  
Display options for sense: (gloss) "an example sentence"

**Noun**

- S: (n) **dog**, [domestic dog](#), [Canis familiaris](#) (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) "*the dog barked all night*"
- S: (n) **frump**, **dog** (a dull unattractive unpleasant girl or woman) "*she got a reputation as a frump*"; "*she's a real dog*"
- S: (n) **dog** (informal term for a man) "*you lucky dog*"
- S: (n) **cad**, **bounder**, **blackguard**, **dog**, **hound**, **heel** (someone who is morally reprehensible) "*you dirty dog*"
- S: (n) **frank**, **frankfurter**, **hotdog**, **hot dog**, **dog**, **wiener**, **wienerwurst**, **weenie** (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- S: (n) **pawl**, **detent**, **click**, **dog** (a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward)
- S: (n) **andiron**, **firedog**, **dog**, **dog-iron** (metal supports for logs in a fireplace) "*the andirons were too hot to touch*"

**Verb**

- S: (v) **chase**, **chase after**, **trail**, **tail**, **tag**, **give chase**, **dog**, **go after**, **track** (go after with the intent to catch) "*The policeman chased the mugger down the alley*"; "*the dog chased the rabbit*"

## WORD NET

- ▶ Provides extensive word definition:

WordNet Search - 3.1  
- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations  
Display options for sense: (gloss) "an example sentence"

**Noun**

- S: (n) [dog](#), [domestic dog](#), [Canis familiaris](#) (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) "*the dog barked all night*"
- S: (n) [frump](#), [dog](#) (a dull unattractive unpleasant girl or woman) "*she got a reputation as a frump*"; "*she's a real dog*"
- S: (n) [dog](#) (informal term for a man) "*you lucky dog*"
- S: (n) [cad](#), [bounder](#), [blackguard](#), [dog](#), [hound](#), [heel](#) (someone who is morally reprehensible) "*you dirty dog*"
- S: (n) [frank](#), [frankfurter](#), [hotdog](#), [hot dog](#), [dog](#), [wiener](#), [wienerwurst](#), [weenie](#) (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- S: (n) [pawl](#), [detent](#), [click](#), [dog](#) (a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward)
- S: (n) [andiron](#), [firedog](#), [dog](#), [dog-iron](#) (metal supports for logs in a fireplace) "*the andirons were too hot to touch*"

**Verb**

- S: (v) [chase](#), [chase after](#), [trail](#), [tail](#), [tag](#), [give chase](#), [dog](#), [go after](#), [track](#) (go after with the intent to catch) "*The policeman chased the mugger down the alley*"; "*the dog chased the rabbit*"

Dog has 7 different meanings when used as a noun ([7 synsets](#))

Dog has only one meanings when used as a verb ([1 synset](#))

## WORD NET

- ▶ Provides extensive word definition:

The screenshot shows the WordNet Search interface version 3.1. At the top, there is a search bar with the placeholder "Word to search for: dog" and a "Search WordNet" button. Below the search bar are "Display Options" and a dropdown menu. A key is provided: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations. It also says "Display options for sense: (gloss) 'an example sentence'".

The results are categorized by part of speech. Under "Noun", there are seven entries, each starting with "S: (n)". The entries are:

- S: (n) [dog](#), [domestic dog](#), [Canis familiaris](#) (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) "the dog barked all night"
- S: (n) [frump](#), [dog](#) (a dull unattractive unpleasant girl or woman) "she got a reputation as a frump"; "she's a real dog"
- S: (n) [dog](#) (informal term for a man) "you lucky dog"
- S: (n) [cad](#), [bounder](#), [blackguard](#), [dog](#), [hound](#), [heel](#) (someone who is morally reprehensible) "you dirty dog"
- S: (n) [frank](#), [frankfurter](#), [hotdog](#), [hot dog](#), [dog](#), [wiener](#), [wienerwurst](#), [weenie](#) (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- S: (n) [pawl](#), [detent](#), [click](#), [dog](#) (a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward)
- S: (n) [andiron](#), [iredog](#), [dog](#), [dog-iron](#) (metal supports for logs in a fireplace) "the andirons were too hot to touch"

Under "Verb", there is one entry:

- S: (v) [chase](#), [chase after](#), [trail](#), [tail](#), [tag](#), [give chase](#), [dog](#), [go after](#), [track](#) (go after with the intent to catch) "The policeman chased the mugger down the alley"; "the dog chased the rabbit"

**SYNONYMS!**

Dog has 7 different meanings when used as a noun ([7 synsets](#))

Dog has only one meanings when used as a verb ([1 synset](#))

## WORD NET

- ▶ Provides extensive word definition:

The screenshot shows the WordNet Search interface version 3.1. At the top, there is a search bar labeled "Word to search for: dog" and a "Search WordNet" button. Below the search bar are "Display Options" and a "Key" explaining the abbreviations "S:" and "W:". The main content area is titled "Noun" and lists seven synsets for the noun "dog". Each synset includes a gloss and an example sentence. A red curved arrow points from the "SYNONYMS!" text below to the list of noun definitions. Another green curved arrow points from the "SYNONYMS!" text to the "Noun" title.

WordNet Search - 3.1  
- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:  [Search WordNet](#)

Display Options:  [Change](#)

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations  
Display options for sense: (gloss) 'an example sentence'

**Noun**

- S: (n) [dog](#), [domestic dog](#), [Canis familiaris](#) (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) "the dog barked all night"
- S: (n) [frump](#), [dog](#) (a dull unattractive unpleasant girl or woman) "she got a reputation as a frump"; "she's a real dog"
- S: (n) [dog](#) (informal term for a man) "you lucky dog"
- S: (n) [cad](#), [bounder](#), [blackguard](#), [dog](#), [hound](#), [heel](#) (someone who is morally reprehensible) "you dirty dog"
- S: (n) [frank](#), [frankfurter](#), [hotdog](#), [hot dog](#), [dog](#), [wiener](#), [wienerwurst](#), [weenie](#) (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- S: (n) [pawl](#), [detent](#), [click](#), [dog](#) (a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward)
- S: (n) [andiron](#), [firedog](#), [dog](#), [dog-iron](#) (metal supports for logs in a fireplace) "the andirons were too hot to touch"

**Verb**

- S: (v) [chase](#), [chase after](#), [trail](#), [tail](#), [tag](#), [give chase](#), [dog](#), [go after](#), [track](#) (go after with the intent to catch) "The policeman chased the mugger down the alley"; "the dog chased the rabbit"

**SYNONYMS!**

Dog has 7 different meanings when used as a noun (7 synsets)

Dog has only one meanings when used as a verb (1 synset)

## WORD NET

- ▶ Provides extensive word definition:

WordNet Search - 3.1  
- [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:  Search WordNet

Display Options: (Select option to change)

Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations  
Display options for sense: (gloss) 'an example sentence'

**Noun**

- S: (n) [dog](#), [domestic dog](#), [Canis familiaris](#) (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) "the dog barked all night"
- S: (n) [frump](#), [dog](#) (a dull unattractive unpleasant girl or woman) "she got a reputation as a frump"; "she's a real dog"
- S: (n) [dog](#) (informal term for a man) "you lucky dog"
- S: (n) [cad](#), [bounder](#), [blackguard](#), [dog](#), [hound](#), [heel](#) (someone who is morally reprehensible) "you dirty dog"
- S: (n) [frank](#), [frankfurter](#), [hotdog](#), [hot dog](#), [dog](#), [wiener](#), [wienerwurst](#), [weenie](#) (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- S: (n) [pawl](#), [detent](#), [click](#), [dog](#) (a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward)
- S: (n) [andiron](#), [firedog](#), [dog](#), [dog-iron](#) (metal supports for logs in a fireplace) "the andirons were too hot to touch"

**Verb**

- S: (v) [chase](#), [chase after](#), [trail](#), [tail](#), [tag](#), [give chase](#), [dog](#), [go after](#), [track](#) (go after with the intent to catch) "The policeman chased the mugger down the alley"; "the dog chased the rabbit"

**SYNONYMS!**

### RESEARCH QUESTION:

How can we find the correct synset to expand? Is the most popular synset the best one?

Dog has 7 different meanings when used as a noun ([7 synsets](#))

Dog has only one meanings when used as a verb ([1 synset](#))

## WORD NET

- ▶ Provides extensive word definition:

- ▶ Often we find meanings that are rarely used or not used anymore at all
- ▶ We can explore other relation between words:
  - ▶ **Hypernyms**: a word with a broad meaning that more specific words fall under; a superordinate. **For example, color is a hypernym of red.**
  - ▶ **Hyponyms**: a word of more specific meaning than a general or superordinate term applicable to it. **For example, spoon is a hyponym of cutlery.**
  - ▶ **Meronymy**: a word that denotes a constituent part or a member of something. **For example, tire is meronym of a car.**
  - ▶ **Holonymy**: the opposite of meronymy. **For example, a book is the holonym of a cover and pages.**

WordNet Search - 3.1  
 - [WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:  Select option to change  Change  
 Key: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations  
 Display options for sense: (gloss) "an example sentence"

**Noun**

- S: (n) [dog](#), [domestic dog](#), [Canis familiaris](#) (a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds) "the dog barked all night"
- S: (n) [frump](#), [dog](#) (a dull unattractive unpleasant girl or woman) "she got a reputation as a frump"; "she's a real dog"
- S: (n) [dog](#) (informal term for a man) "you lucky dog"
- S: (n) [cad](#), [bounder](#), [blackguard](#), [dog](#), [hound](#), [heel](#) (someone who is morally reprehensible) "you dirty dog"
- S: (n) [frank](#), [frankfurter](#), [hotdog](#), [hot dog](#), [dog](#), [wiener](#), [wienerwurst](#), [weenie](#) (a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll)
- S: (n) [pawl](#), [detent](#), [click](#), [dog](#) (a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward)
- S: (n) [andiron](#), [firedog](#), [dog](#), [dog-iron](#) (metal supports for logs in a fireplace) "the andirons were too hot to touch"

**Verb**

- S: (v) [chase](#), [chase after](#), [trail](#), [tail](#), [tag](#), [give chase](#), [dog](#), [go after](#), [track](#) (go after with the intent to catch) "The policeman chased the mugger down the alley"; "the dog chased the rabbit"

```
from nltk.corpus import wordnet as wn

for i,j in enumerate(wn.synsets('dog')):
    print "Meaning", i + 1
    print "Synset ID:", j.name()
    print "Definition:",j.definition()
    print "Synonyms:", ", ".join(j.lemma_names())

print

Meaning 1
Synset ID: dog.n.01
Definition: a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds
Synonyms: dog, domestic_dog, Canis_familiaris

Meaning 2
Synset ID: frump.n.01
Definition: a dull unattractive unpleasant girl or woman
Synonyms: frump, dog

Meaning 3
Synset ID: dog.n.03
Definition: informal term for a man
Synonyms: dog

Meaning 4
Synset ID: cad.n.01
Definition: someone who is morally reprehensible
Synonyms: cad, bounder, blackguard, dog, hound, heel

Meaning 5
Synset ID: frank.n.02
Definition: a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll
Synonyms: frank, frankfurter, hotdog, hot_dog, dog, wiener, wienerwurst, weenie

Meaning 6
Synset ID: pawl.n.01
Definition: a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward
Synonyms: pawl, detent, click, dog

Meaning 7
Synset ID: andiron.n.01
Definition: metal supports for logs in a fireplace
Synonyms: andiron, firedog, dog, dog-iron

Meaning 8
Synset ID: chase.v.01
Definition: go after with the intent to catch
Synonyms: chase, chase_after, trail, tail, tag, give_chase, dog, go_after, track
```

```

from nltk.corpus import wordnet as wn

for i,j in enumerate(wn.synsets('dog')):
    print "Meaning", i + 1
    print "Synset ID:", j.name()
    print "Definition:",j.definition()
    print "Synonyms:", ", ".join(j.lemma_names())
    print "Hypernym:",
    for hyper in j.hypernyms():
        print hyper.lemma_names(),
    print
    print "Hyponyms:",
    for hypo in j.hyponyms():
        print hypo.lemma_names(),
    print
    print

```

Meaning 1  
Synset ID: dog.n.01  
Definition: a member of the genus *Canis* (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds  
Synonyms: dog, domestic\_dog, *Canis familiaris*  
Hypernym: [u'canine', u'canid'] [u'domestic\_animal', u'domesticated\_animal']  
Hyponyms: [u'basenji'] [u'corgi', u'Welsh\_corgi'] [u'cur', u'mongrel', u'mutt'] [u'dalmatian', u'coach\_dog', u'carriage\_dog'] [u'Great\_Pyrenees'] [u'griffon', u'Brussels\_griffon', u'Belgian\_griffon'] [u'hunting\_dog'] [u'lapdog'] [u'Leonberg'] [u'Mexican\_hairless'] [u'Newfoundland', u'Newfoundland\_dog'] [u'pooch', u'doggie', u'doggy', u'barker', u'bow-wow'] [u'poodle', u'poodle\_dog'] [u'pug', u'pug-dog'] [u'puppy'] [u'spitz'] [u'toy\_dog', u'toy'] [u'working\_dog']

Meaning 2  
Synset ID: frump.n.01  
Definition: a dull unattractive unpleasant girl or woman  
Synonyms: frump, dog  
Hypernym: [u'unpleasant\_woman', u'disagreeable\_woman']  
Hyponyms:

Meaning 3  
Synset ID: dog.n.03  
Definition: informal term for a man  
Synonyms: dog  
Hypernym: [u'chap', u'fellow', u'feller', u'fella', u'lad', u'gent', u'blighter', u'cuss', u'bloke']  
Hyponyms:

Meaning 4  
Synset ID: cad.n.01  
Definition: someone who is morally reprehensible  
Synonyms: cad, bounder, blackguard, dog, hound, heel  
Hypernym: [u'villain', u'scoundrel']  
Hyponyms: [u'perisher']

Meaning 5  
Synset ID: frank.n.02  
Definition: a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll  
Synonyms: frank, frankfurter, hotdog, hot\_dog, dog, wiener, wienerwurst, weenie  
Hypernym: [u'sausage']  
Hyponyms: [u'Vienna\_sausage']

### WHAT ELSE?

- ▶ WordNet is useful, but sometimes it is just too generic...
- ▶ We could do a much better job if we could manually create the thesaurus for our specific needs:
  - ▶ UMLS (Unified Medical Language System)
  - ▶ MeSH (Medical Subject Headings)

### WHAT ELSE?

- ▶ WordNet is useful, but sometimes it is just too generic...
- ▶ We could do a much better job if we could manually create the thesaurus for our specific needs:
  - ▶ UMLS (Unified Medical Language System)
  - ▶ MeSH (Medical Subject Headings)

**Manually creating something? Trade-offs?**

# AUTOMATIC EXPANSION WITH WORD RELATION

- ▶ Can we explore word co-occurrences to improve query expansion?
- ▶ Sample approach:
  - ▶ She wore neither makeup nor jewellery, but her blouse and skirt
  - ▶ She had changed her clothes and washed the makeup from her face.
  - ▶ She wore no makeup on her skin but some alluring perfume made her smell like heaven.

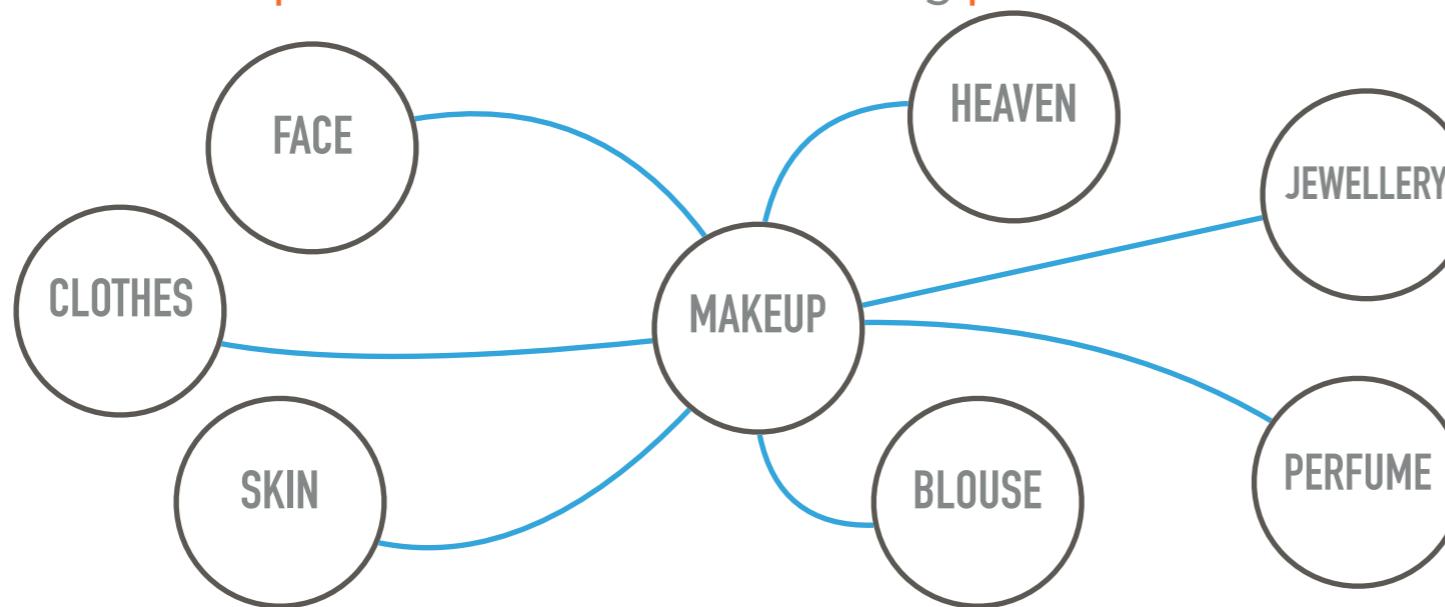
# AUTOMATIC EXPANSION WITH WORD RELATION

- ▶ Can we explore word co-occurrences to improve query expansion?
- ▶ Sample approach:
  - ▶ She wore neither **makeup** nor **jewellery**, but her **blouse** and **skirt**
  - ▶ She had changed her **clothes** and washed the **makeup** from her **face**.
  - ▶ She wore no **makeup** on her **skin** but some alluring **perfume** made her smell like **heaven**.



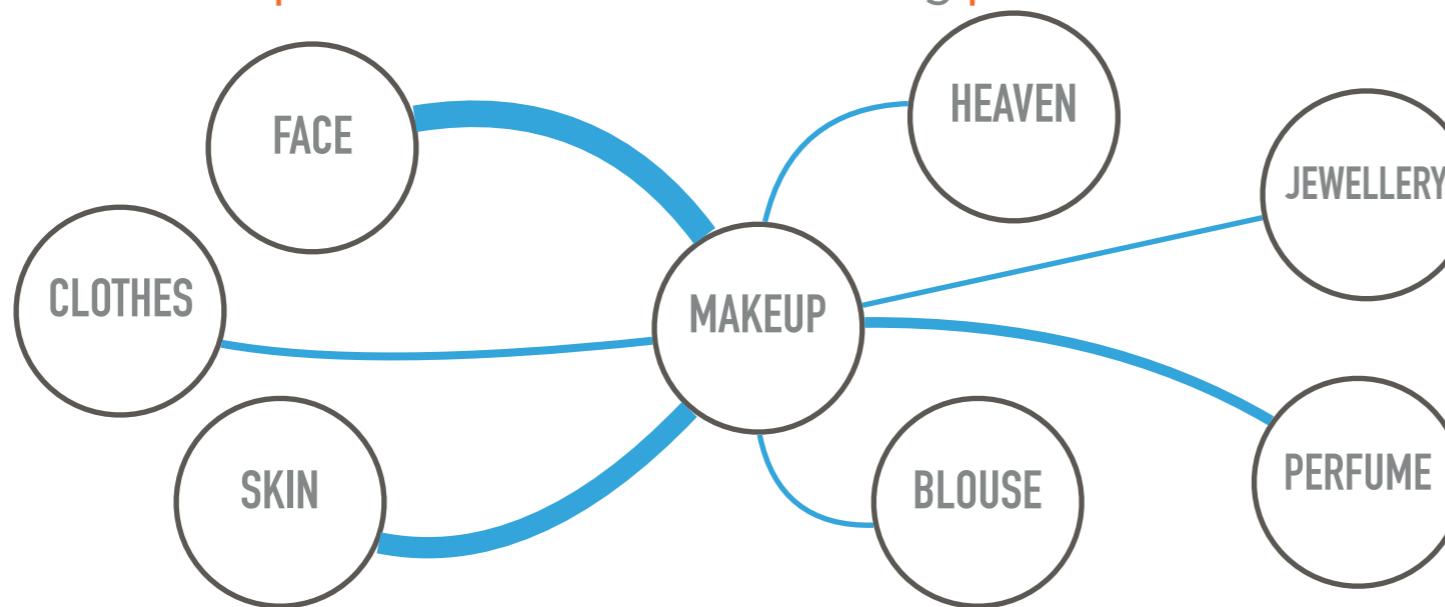
# AUTOMATIC EXPANSION WITH WORD RELATION

- ▶ Can we explore word co-occurrences to improve query expansion?
- ▶ Sample approach:
  - ▶ She wore neither **makeup** nor **jewellery**, but her **blouse** and **skirt**
  - ▶ She had changed her **clothes** and washed the **makeup** from her **face**.
  - ▶ She wore no **makeup** on her **skin** but some alluring **perfume** made her smell like **heaven**.



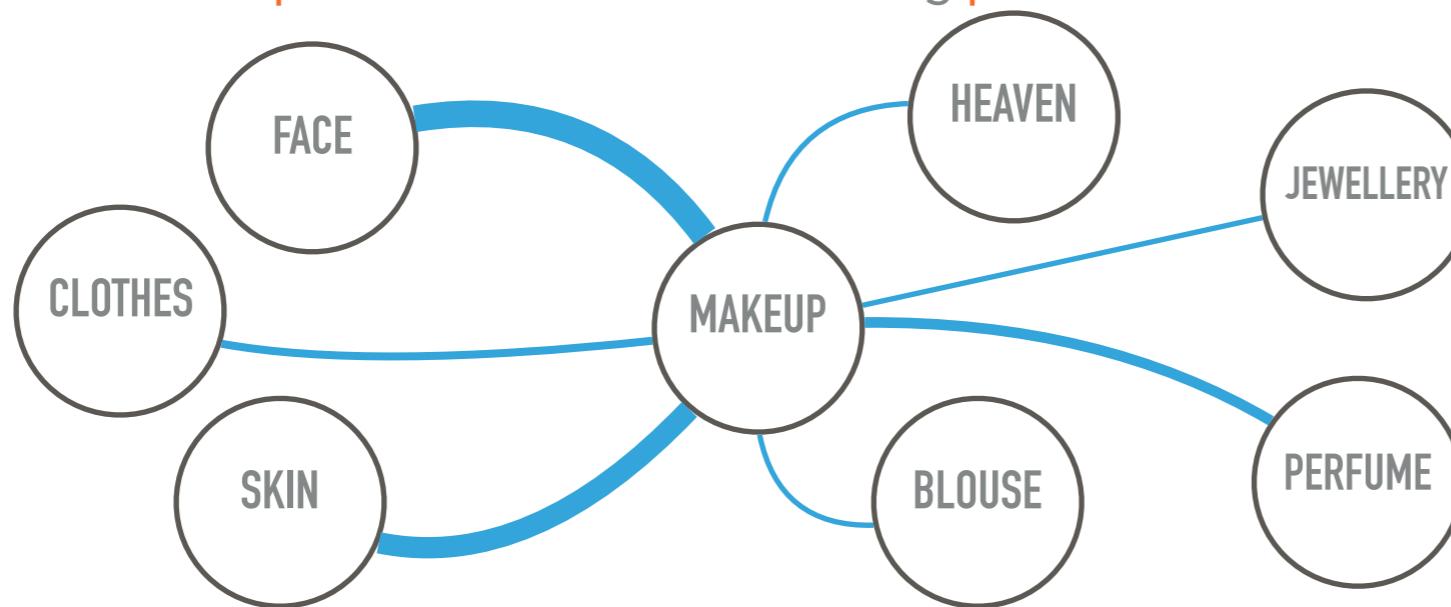
# AUTOMATIC EXPANSION WITH WORD RELATION

- ▶ Can we explore word co-occurrences to improve query expansion?
- ▶ Sample approach:
  - ▶ She wore neither **makeup** nor **jewellery**, but her **blouse** and **skirt**
  - ▶ She had changed her **clothes** and washed the **makeup** from her **face**.
  - ▶ She wore no **makeup** on her **skin** but some alluring **perfume** made her smell like **heaven**.



# AUTOMATIC EXPANSION WITH WORD RELATION

- ▶ Can we explore word co-occurrences to improve query expansion?
- ▶ Sample approach:
  - ▶ She wore neither **makeup** nor **jewellery**, but her **blouse** and **skirt**
  - ▶ She had changed her **clothes** and washed the **makeup** from her **face**.
  - ▶ She wore no **makeup** on her **skin** but some alluring **perfume** made her smell like **heaven**.



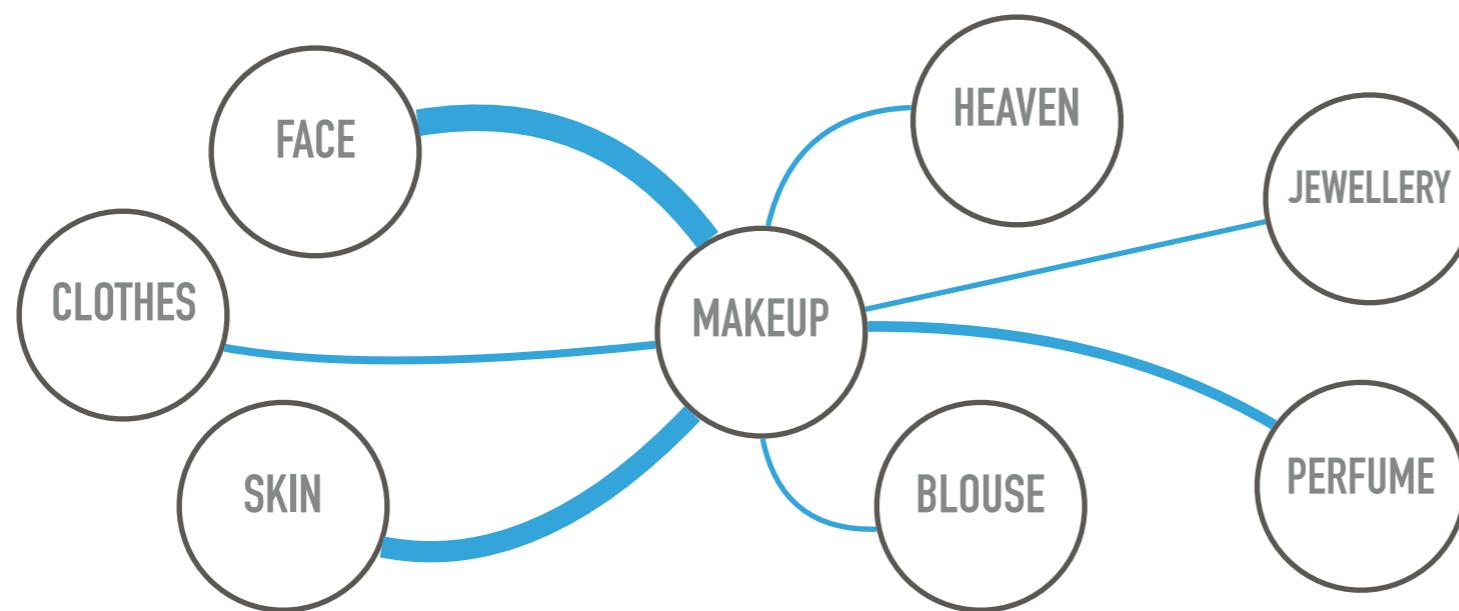
Pick top X (e.g. 3) nouns  
that co-occur with  
keyword (makeup):

1. Face
2. Skin
3. Perfume

## AUTOMATIC EXPANSION WITH WORD RELATION

| Word        | Nearest neighbors                                |
|-------------|--|
| absolutely  | absurd, whatsoever, totally, exactly, nothing    |
| bottomed    | dip, copper, drops, topped, slide, trimmed       |
| captivating | shimmer, stunningly, superbly, plucky, witty     |
| doghouse    | dog, porch, crawling, beside, downstairs         |
| makeup      | repellent, lotion, glossy, sunscreen, skin, gel  |
| mediating   | reconciliation, negotiate, case, conciliation    |
| keeping     | hoping, bring, wiping, could, some, would        |
| lithographs | drawings, Picasso, Dali, sculptures, Gauguin     |
| pathogens   | toxins, bacteria, organisms, bacterial, parasite |
| senses      | grasp, psyche, truly, clumsy, naive, innate      |

► **Figure 9.8** An example of an automatically generated thesaurus. This example is based on the work in Schütze (1998), which employs latent semantic indexing (see Chapter 18).



Pick top X (e.g. 3) nouns  
that co-occur with  
keyword (makeup):

1. Face
2. Skin
3. Perfume

---

# LOCAL METHODS

# RELEVANCE FEEDBACK

- ▶ The key idea of these methods is that search is (or at least should be) an iterative processes

# RELEVANCE FEEDBACK

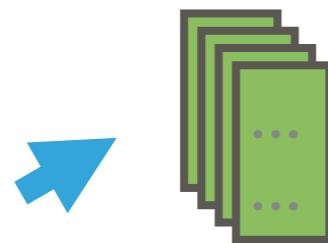
- ▶ The key idea of these methods is that search is (or at least should be) an iterative processes



Information Need

# RELEVANCE FEEDBACK

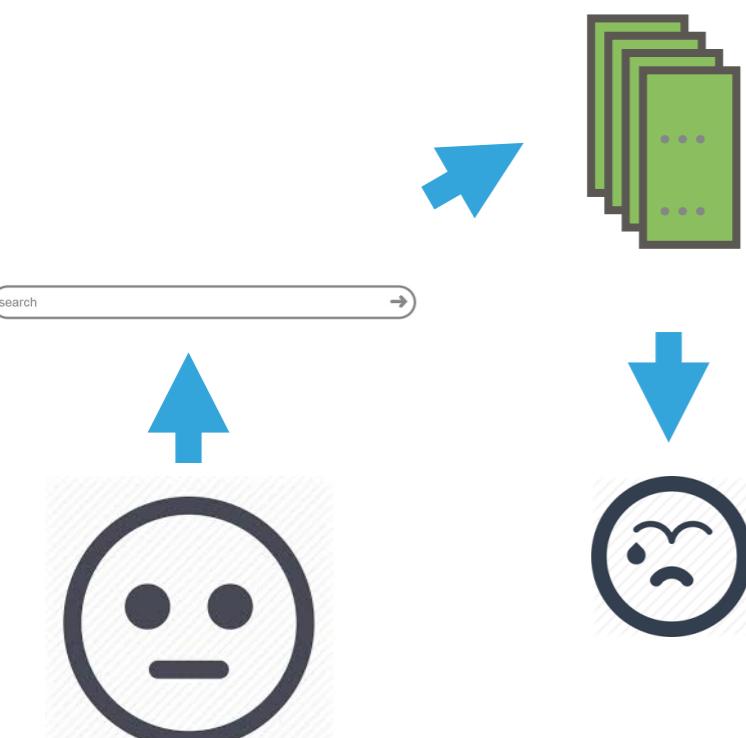
- ▶ The key idea of these methods is that search is (or at least should be) an iterative processes



Information Need

# RELEVANCE FEEDBACK

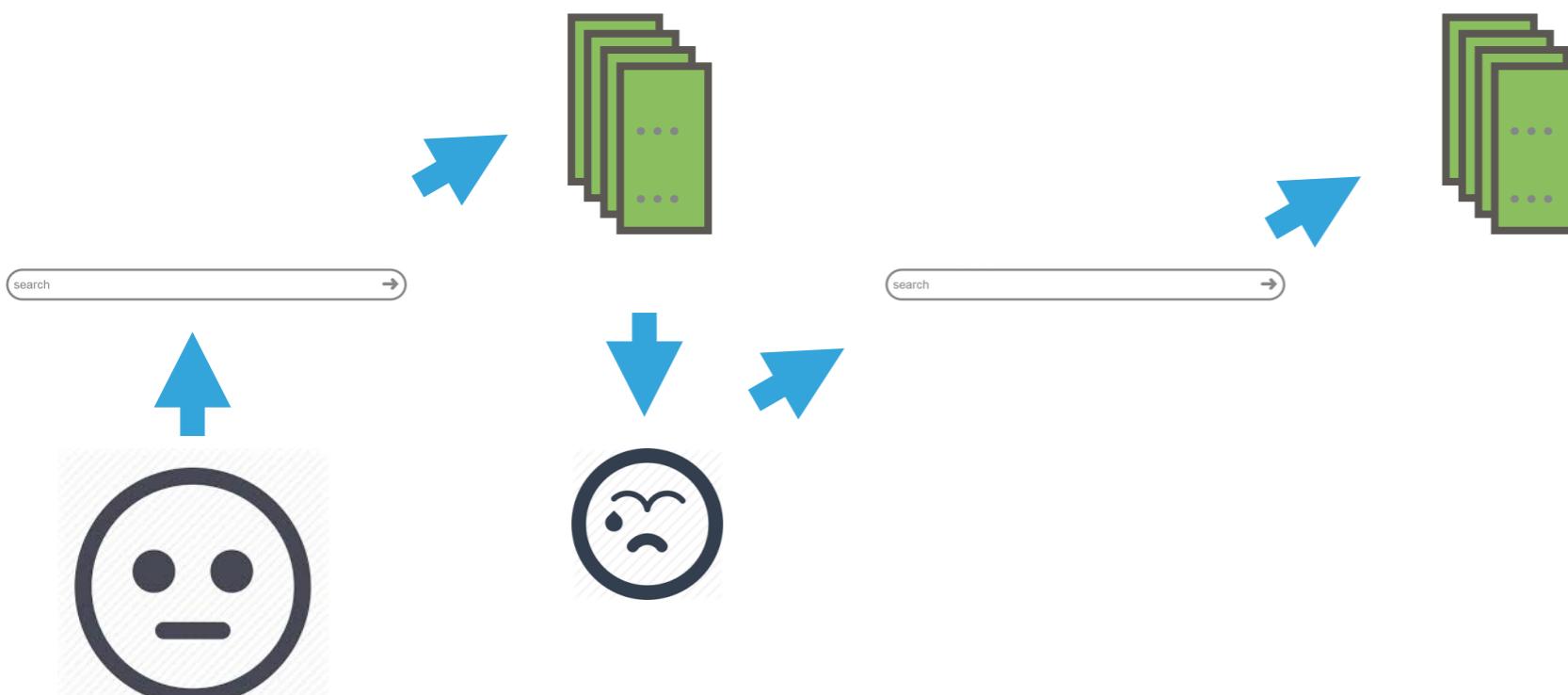
- ▶ The key idea of these methods is that search is (or at least should be) an iterative processes



Information Need

# RELEVANCE FEEDBACK

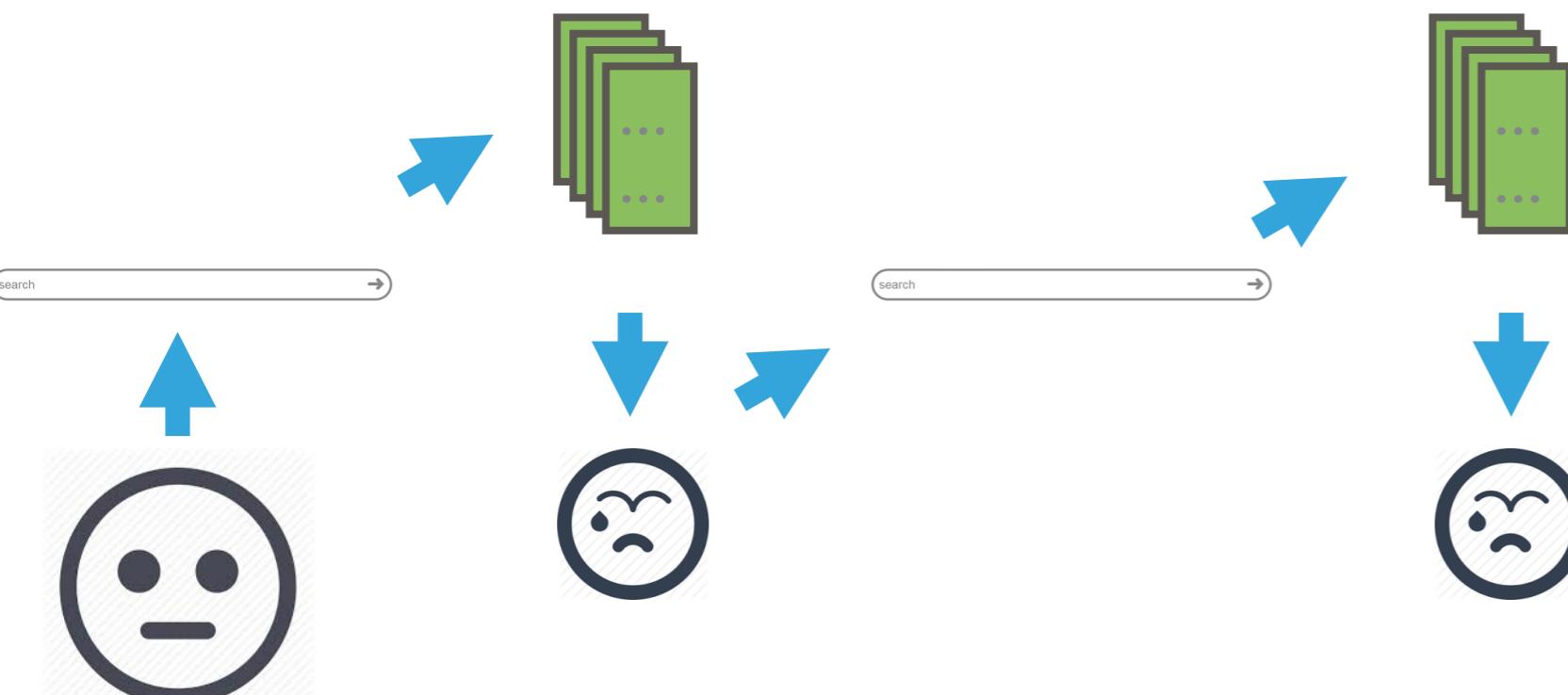
- ▶ The key idea of these methods is that search is (or at least should be) an iterative processes



Information Need

# RELEVANCE FEEDBACK

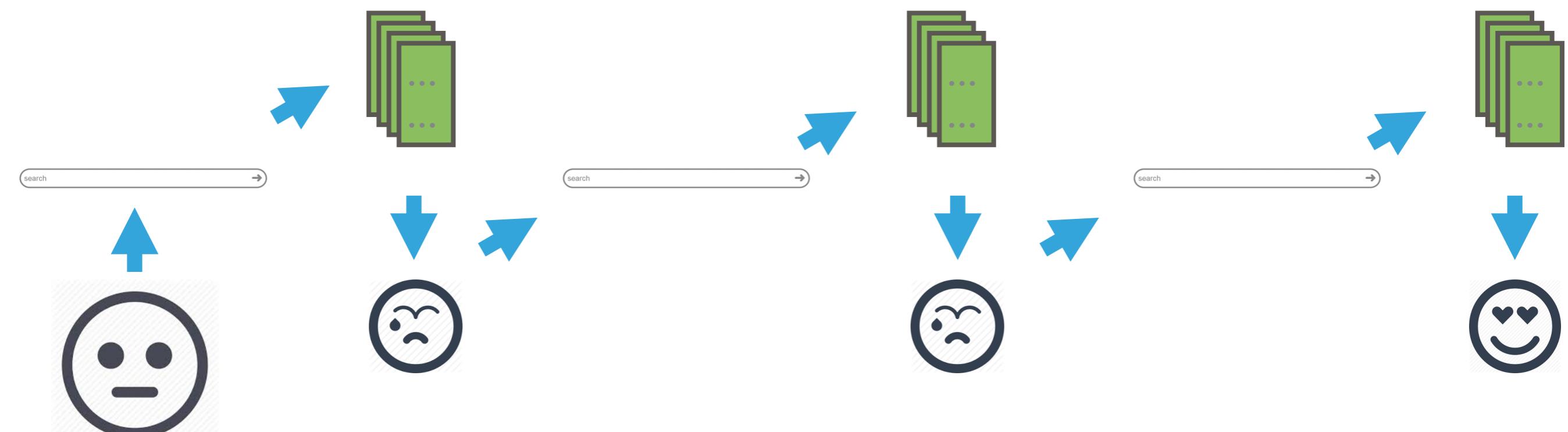
- ▶ The key idea of these methods is that search is (or at least should be) an iterative processes



Information Need

# RELEVANCE FEEDBACK

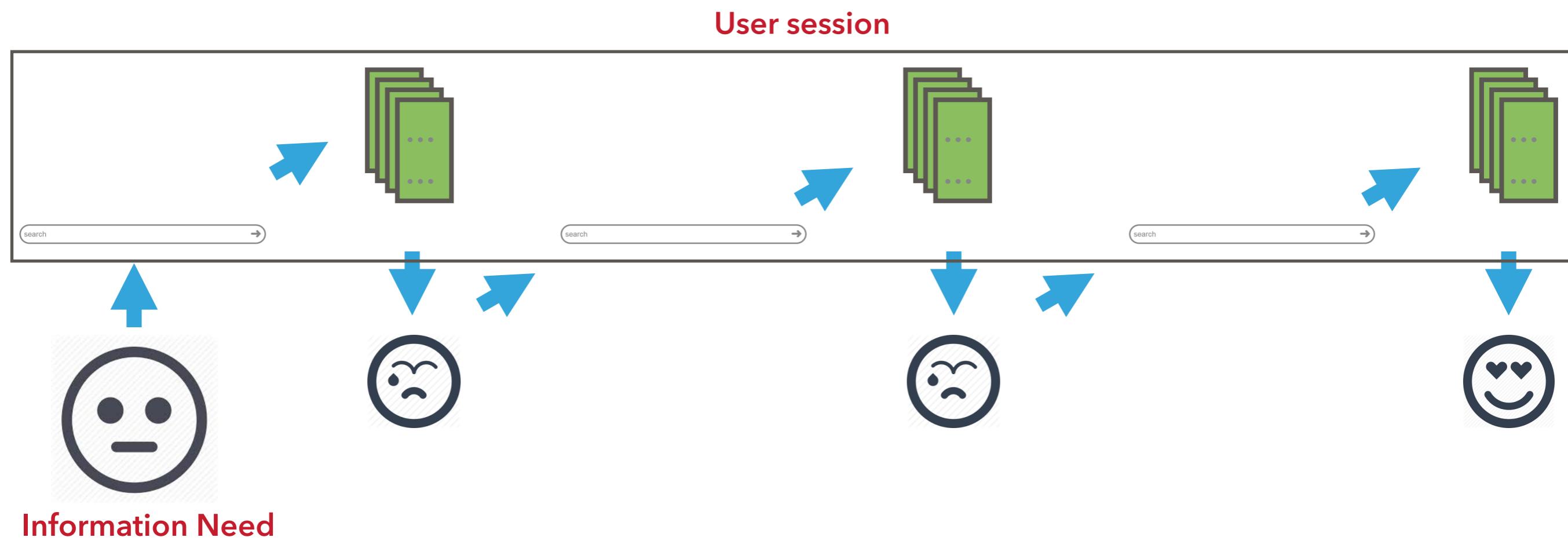
- ▶ The key idea of these methods is that search is (or at least should be) an iterative processes

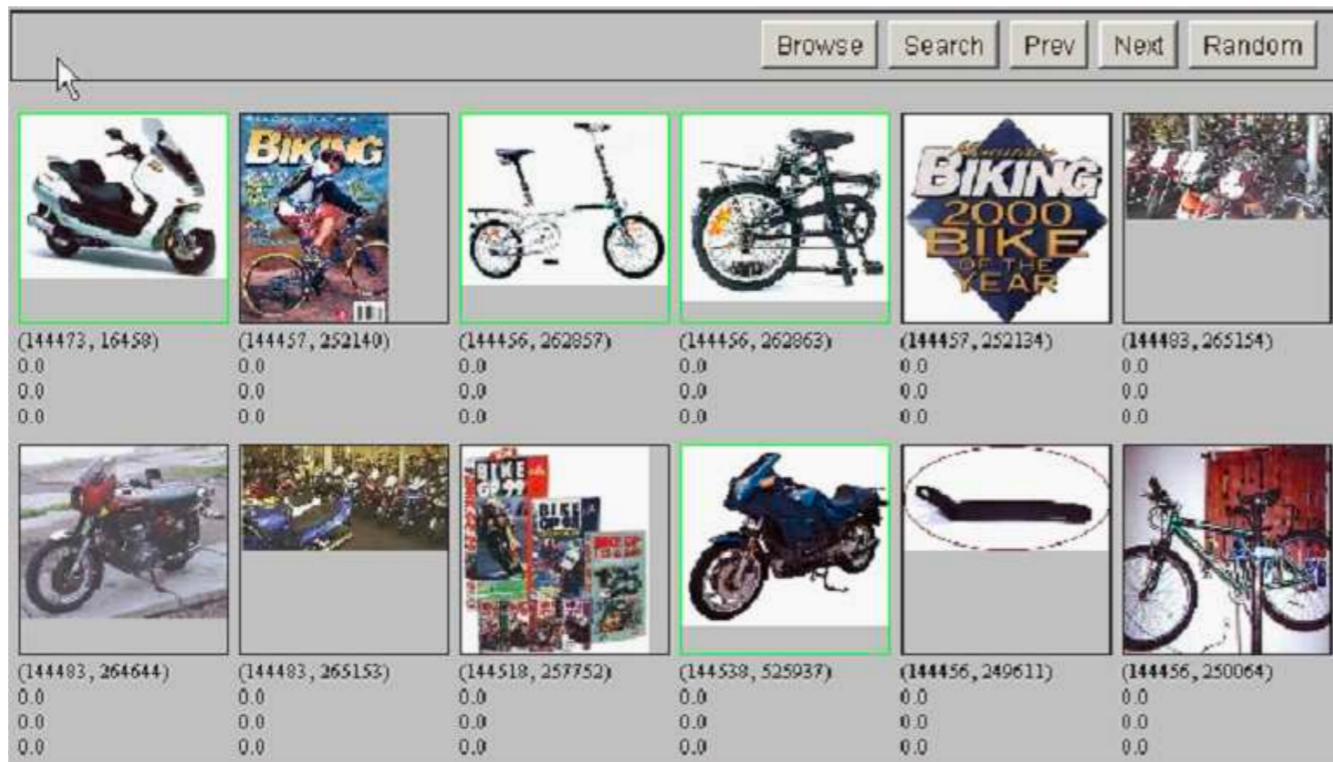


Information Need

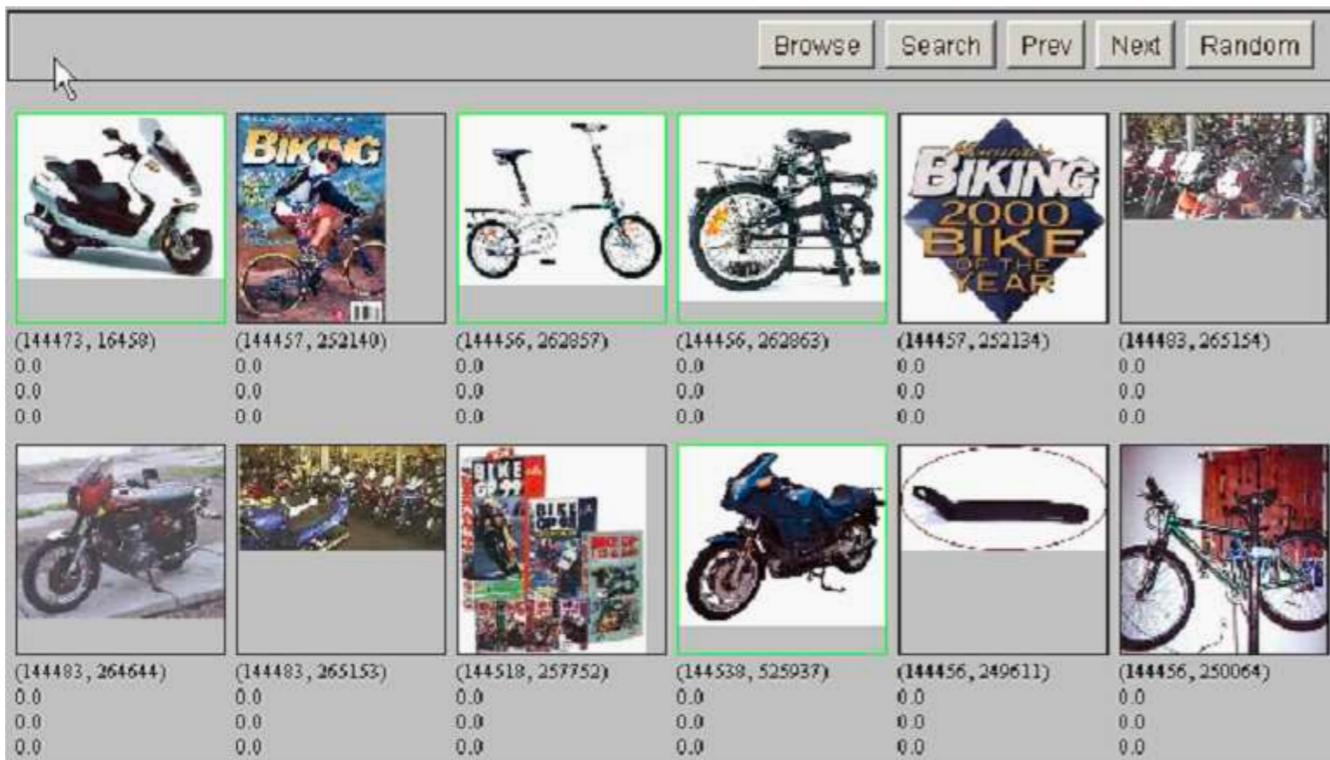
# RELEVANCE FEEDBACK

- ▶ The key idea of these methods is that search is (or at least should be) an iterative processes

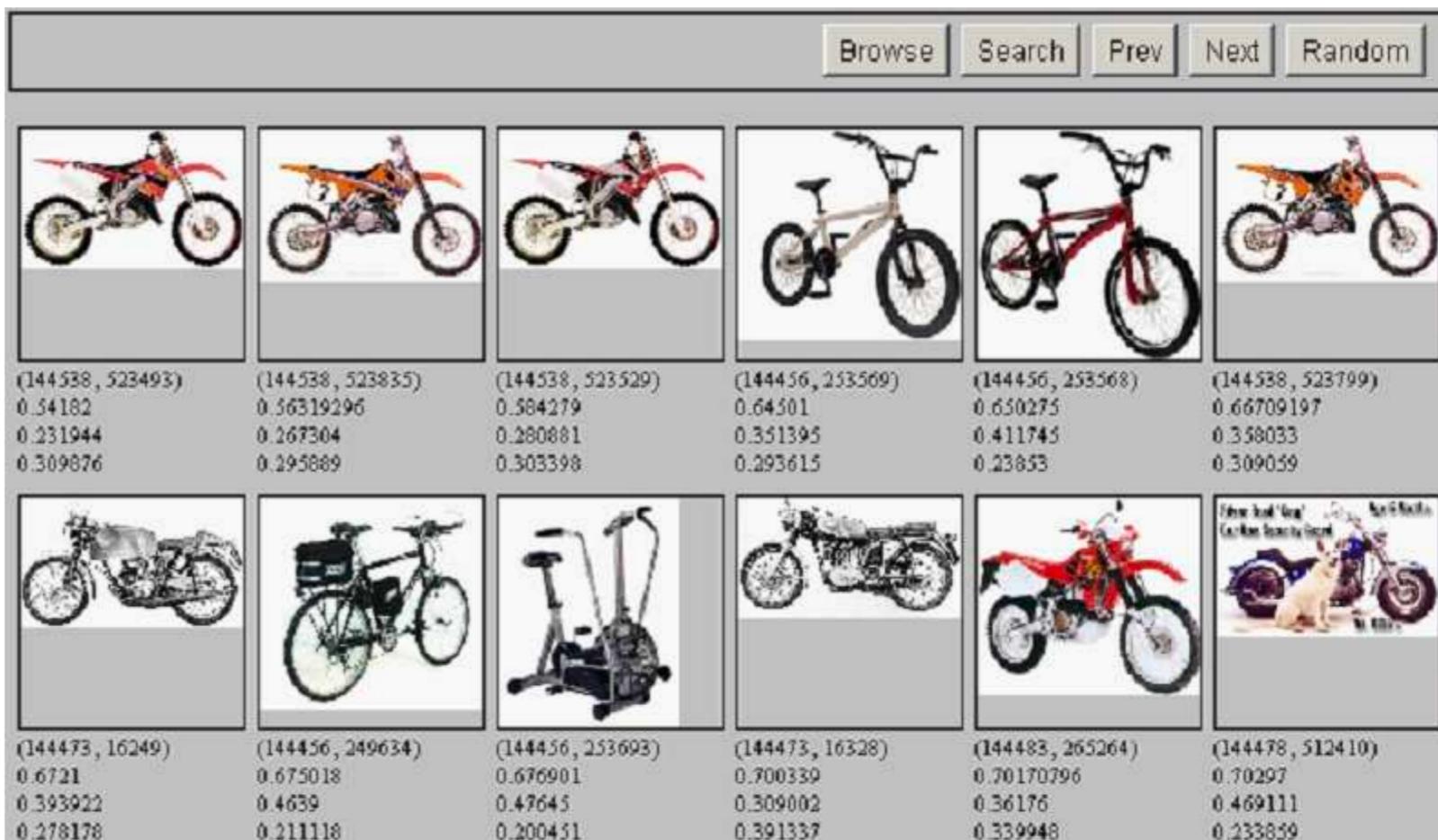




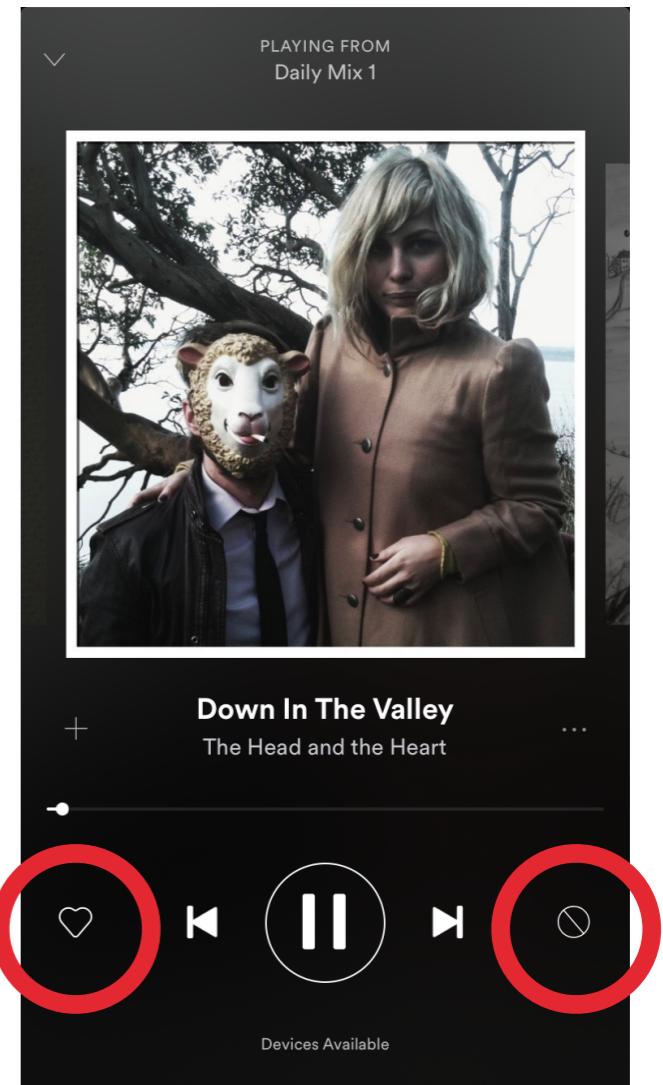
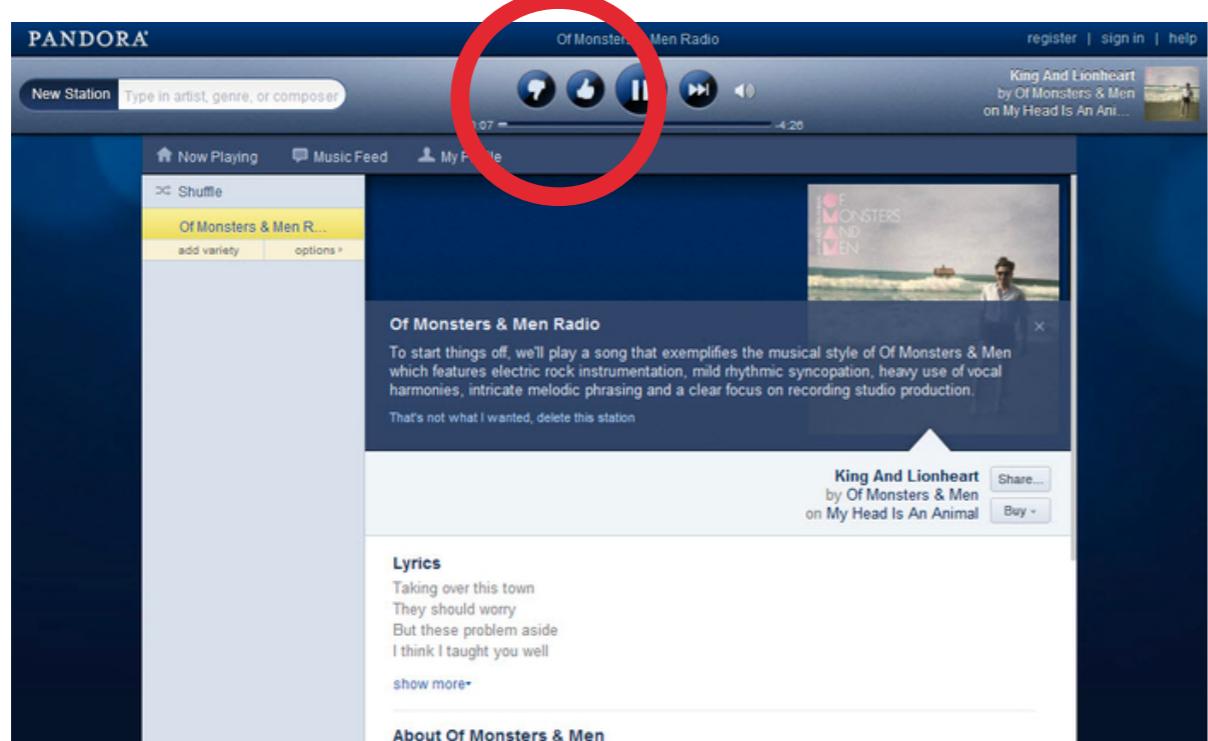
Local methods exploit the idea that it may be  
**difficult to formulate a good query**  
when you don't know the collection well



Local methods exploit the idea that it may be **difficult to formulate a good query** when you don't know the collection well



**BUT** it is **easy to judge** particular documents, and so it makes sense to engage in **iterative query refinement** of this sort



What are some negative feedbacks on Facebook?

- (a) Query: New space satellite applications
- (b)
  - + 1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
  - + 2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
  - 3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
  - 4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
  - 5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
  - 6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
  - 7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
  - + 8. 0.509, 12/02/87, Telecommunications Tale of Two Companies

- (a) Query: New space satellite applications
- (b)
- + 1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
  - + 2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
  - 3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
  - 4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
  - 5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
  - 6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
  - 7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
  - + 8. 0.509, 12/02/87, Telecommunications Tale of Two Companies
- 2.074 new 15.106 space  
30.816 satellite 5.660 application  
5.991 nasa 5.196 eos  
4.196 launch 3.972 aster  
3.516 instrument 3.446 arianespace  
3.004 bundespost 2.806 ss  
2.790 rocket 2.053 scientist  
2.003 broadcast 1.172 earth  
0.836 oil 0.646 measure

(a) Query: New space satellite applications

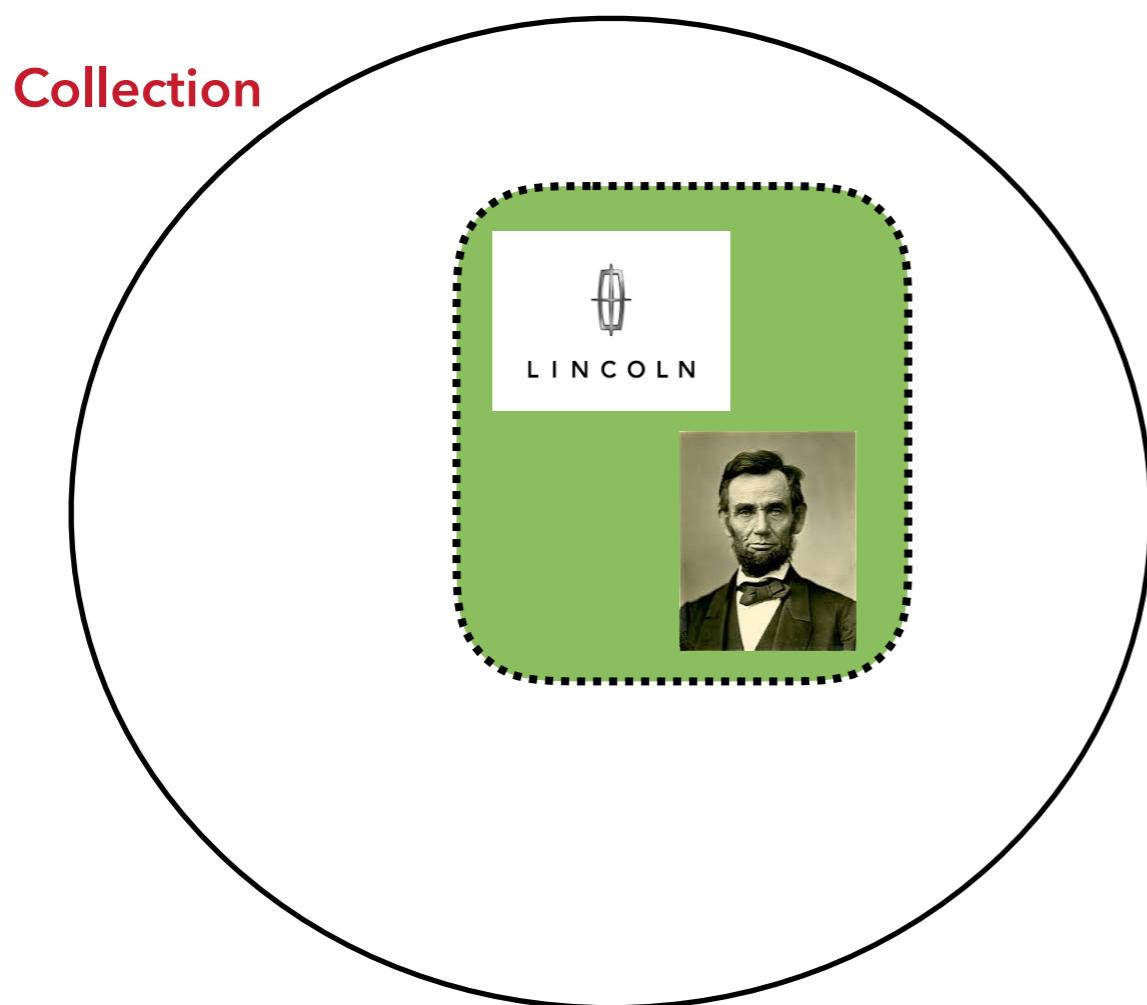
- (b)
- + 1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
  - + 2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
  - 3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
  - 4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
  - 5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
  - 6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
  - 7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
  - + 8. 0.509, 12/02/87, Telecommunications Tale of Two Companies

|                  |                   |
|------------------|-------------------|
| 2.074 new        | 15.106 space      |
| 30.816 satellite | 5.660 application |
| 5.991 nasa       | 5.196 eos         |
| 4.196 launch     | 3.972 aster       |
| 3.516 instrument | 3.446 arianespace |
| 3.004 bundespost | 2.806 ss          |
| 2.790 rocket     | 2.053 scientist   |
| 2.003 broadcast  | 1.172 earth       |
| 0.836 oil        | 0.646 measure     |

- (a) Query: New space satellite applications
- (b)
- + 1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
  - + 2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
  - 3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
  - 4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
  - 5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
  - 6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
  - 7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
  - + 8. 0.509, 12/02/87, Telecommunications Tale of Two Companies
- (d)
- \* 1. 0.513, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
  - \* 2. 0.500, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
  - 3. 0.493, 08/07/89, When the Pentagon Launches a Secret Satellite, Space Sleuths Do Some Spy Work of Their Own
  - 4. 0.493, 07/31/89, NASA Uses 'Warm' Superconductors For Fast Circuit
  - \* 5. 0.492, 12/02/87, Telecommunications Tale of Two Companies
  - 6. 0.491, 07/09/91, Soviets May Adapt Parts of SS-20 Missile For Commercial Use
  - 7. 0.490, 07/12/88, Gaping Gap: Pentagon Lags in Race To Match the Soviets In Rocket Launchers
  - 8. 0.490, 06/14/90, Rescue of Satellite By Space Agency To Cost \$90 Million
- |                  |                   |
|------------------|-------------------|
| 2.074 new        | 15.106 space      |
| 30.816 satellite | 5.660 application |
| 5.991 nasa       | 5.196 eos         |
| 4.196 launch     | 3.972 aster       |
| 3.516 instrument | 3.446 arianespace |
| 3.004 bundespost | 2.806 ss          |
| 2.790 rocket     | 2.053 scientist   |
| 2.003 broadcast  | 1.172 earth       |
| 0.836 oil        | 0.646 measure     |

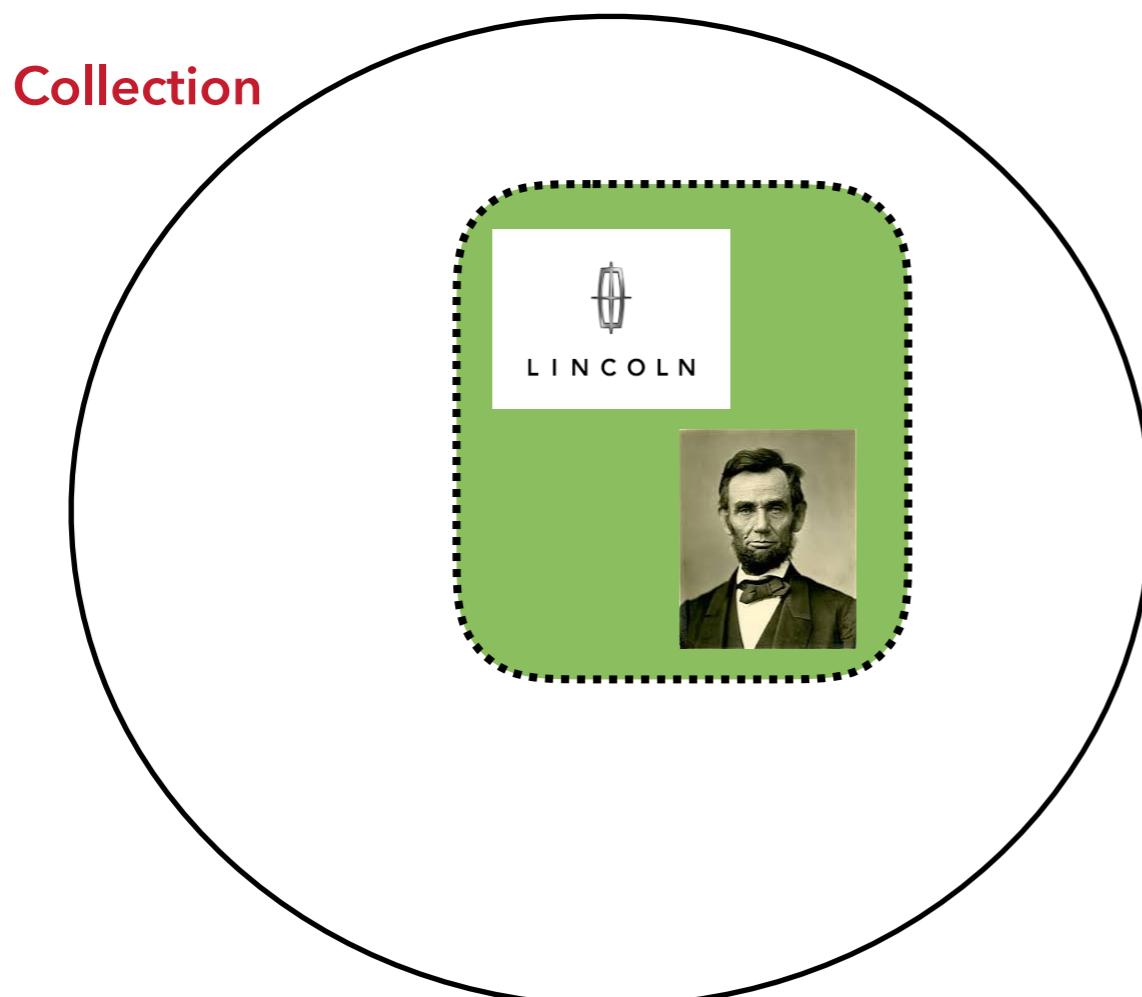
# PROBABILISTIC RELEVANCE FEEDBACK

- ▶ Query: “lincoln”



# PROBABILISTIC RELEVANCE FEEDBACK

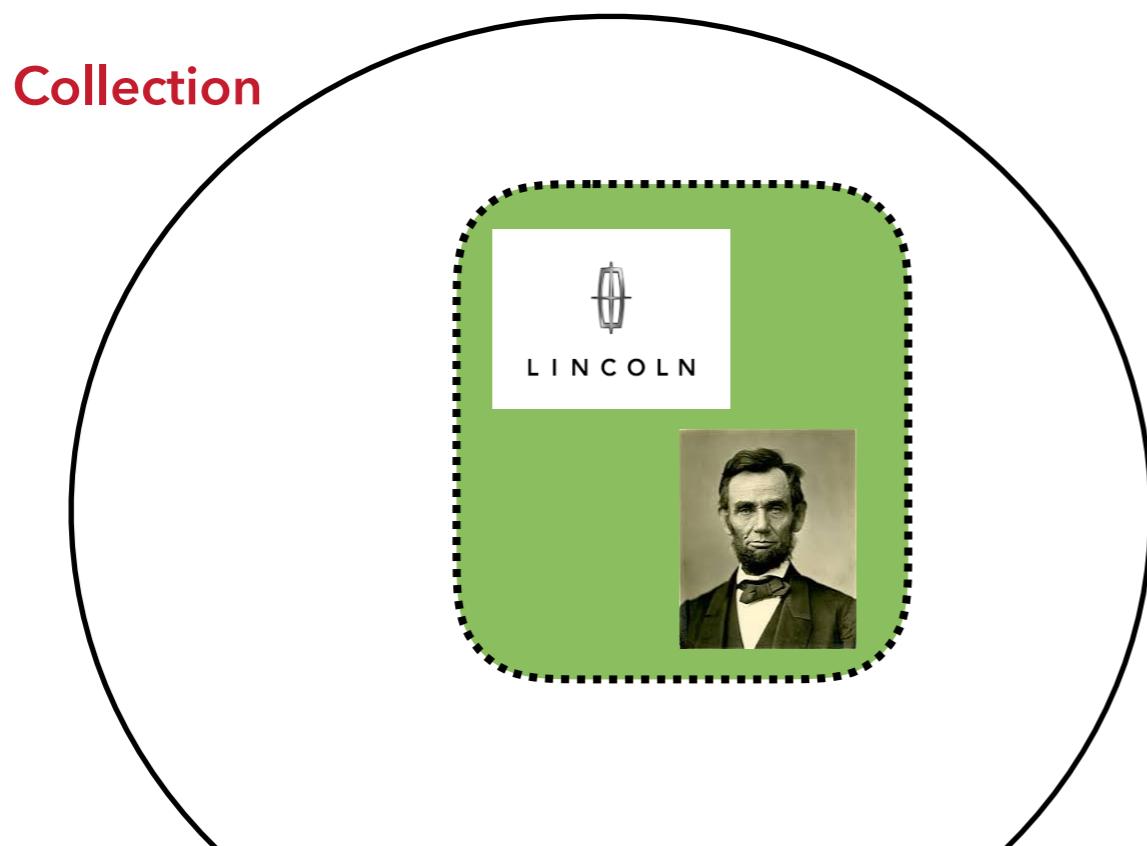
- ▶ Query: “lincoln”



- ▶ For every term  $t$ , what is the likelihood of  $t$  being present in relevant docs Vs. being present in non relevant docs.
- ▶ Term: “biography”
  - ▶ What is  $P(R = 1 | \text{“biography”})$  ?
  - ▶ What is  $P(R = 0 | \text{“biography”})$  ?
- ▶ Term: “industry”
  - ▶ What is  $P(R = 1 | \text{“industry”})$  ?
  - ▶ What is  $P(R = 0 | \text{“industry”})$  ?

# PROBABILISTIC RELEVANCE FEEDBACK

- ▶ Query: “lincoln”



$$\begin{aligned}\hat{P}(x_t = 1 | R = 1) &= |VR_t| / |VR| \\ \hat{P}(x_t = 1 | R = 0) &= (df_t - |VR_t|) / (N - |VR|)\end{aligned}$$

- ▶ For every term  $t$ , what is the likelihood of  $t$  being present in relevant docs Vs. being present in non relevant docs.
- ▶ Term: “biography”
  - ▶ What is  $P(R = 1 | \text{"biography"})$  ?
  - ▶ What is  $P(R = 0 | \text{"biography"})$  ?
- ▶ Term: “industry”
  - ▶ What is  $P(R = 1 | \text{"industry"})$  ?
  - ▶ What is  $P(R = 0 | \text{"industry"})$  ?

# PROBABILISTIC RELEVANCE FEEDBACK

- ▶ Suitable for probabilistic models (BIM, BM25, LM)

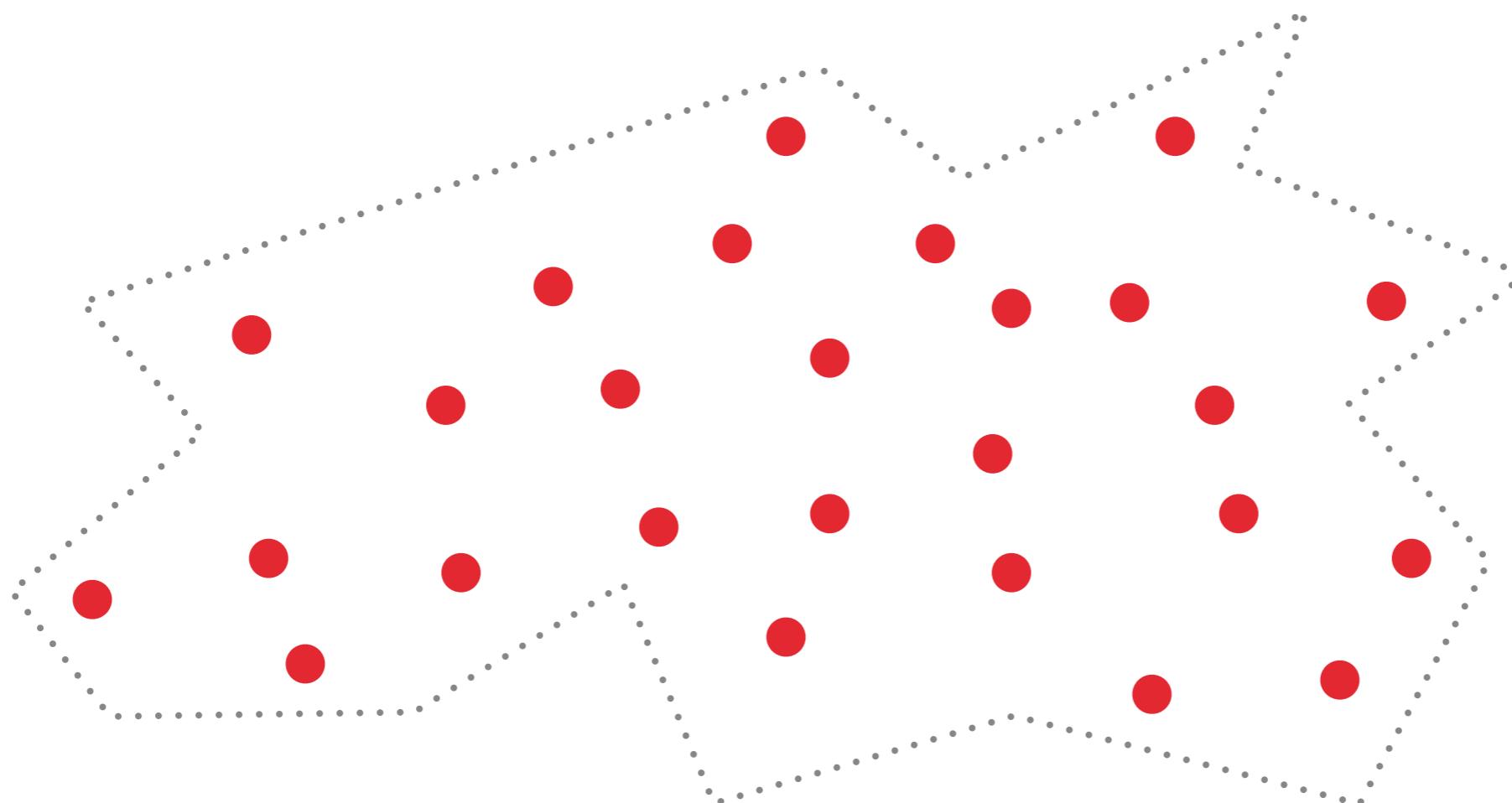
[LM? How? Suggestions?](#)

# PROBABILISTIC RELEVANCE FEEDBACK

- ▶ Suitable for probabilistic models (BIM, BM25, LM)  
  
[LM? How? Suggestions?](#)
  
- ▶ Can we use a similar approach for the vector space model?

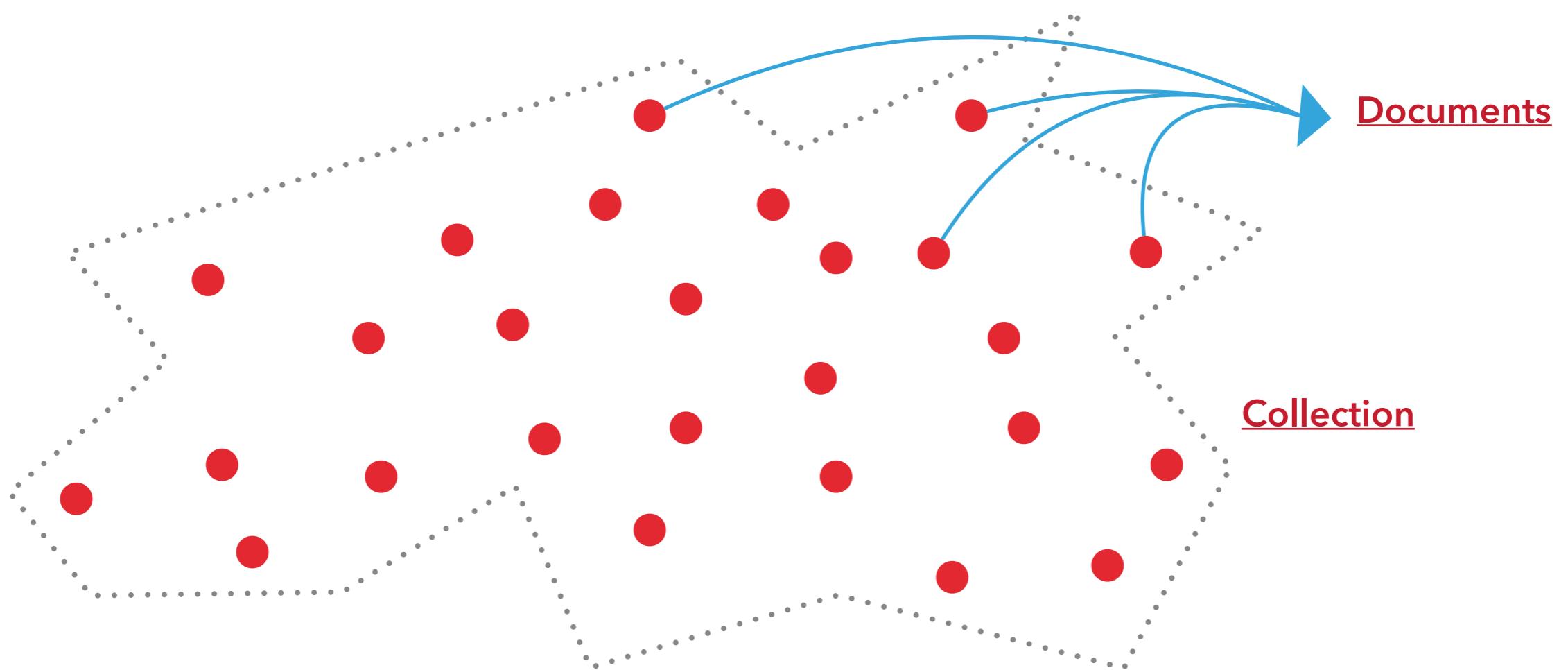
# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶ We can imagine our documents as points in a space



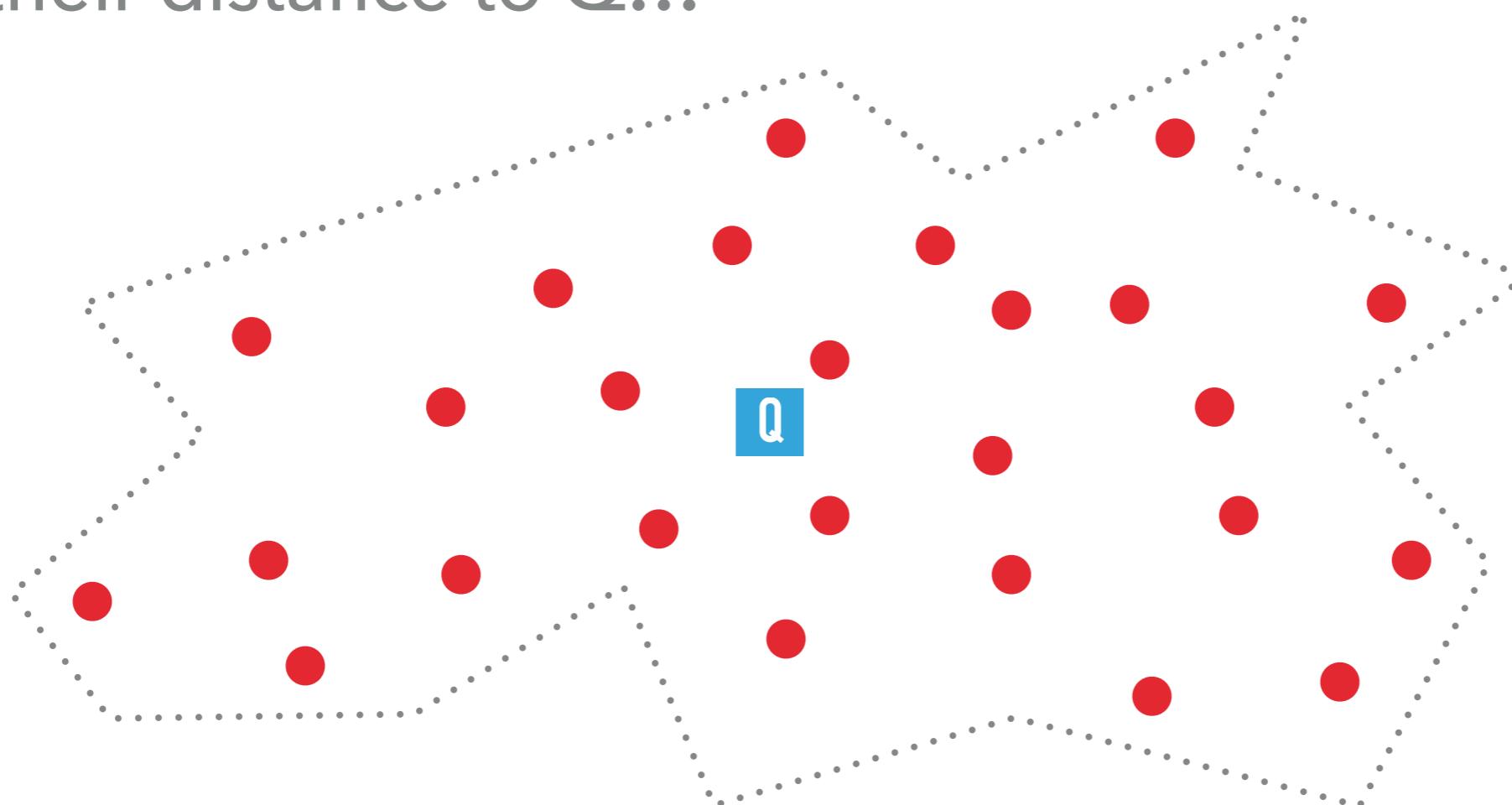
# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶ We can imagine our documents as points in a space



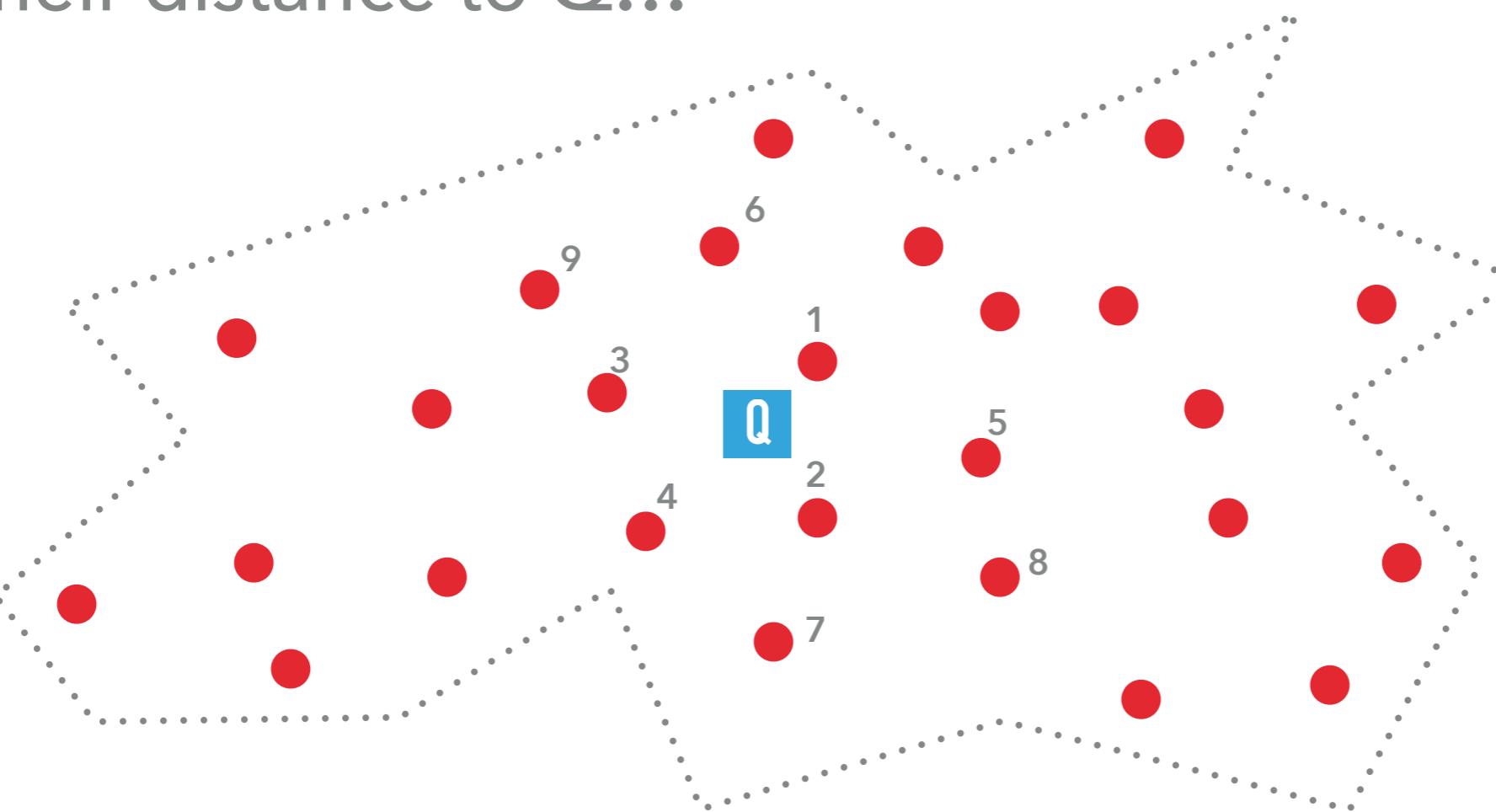
# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶ We can imagine our documents as points in a space
- ▶ Q is an initial query and we rank the documents according to their distance to Q...



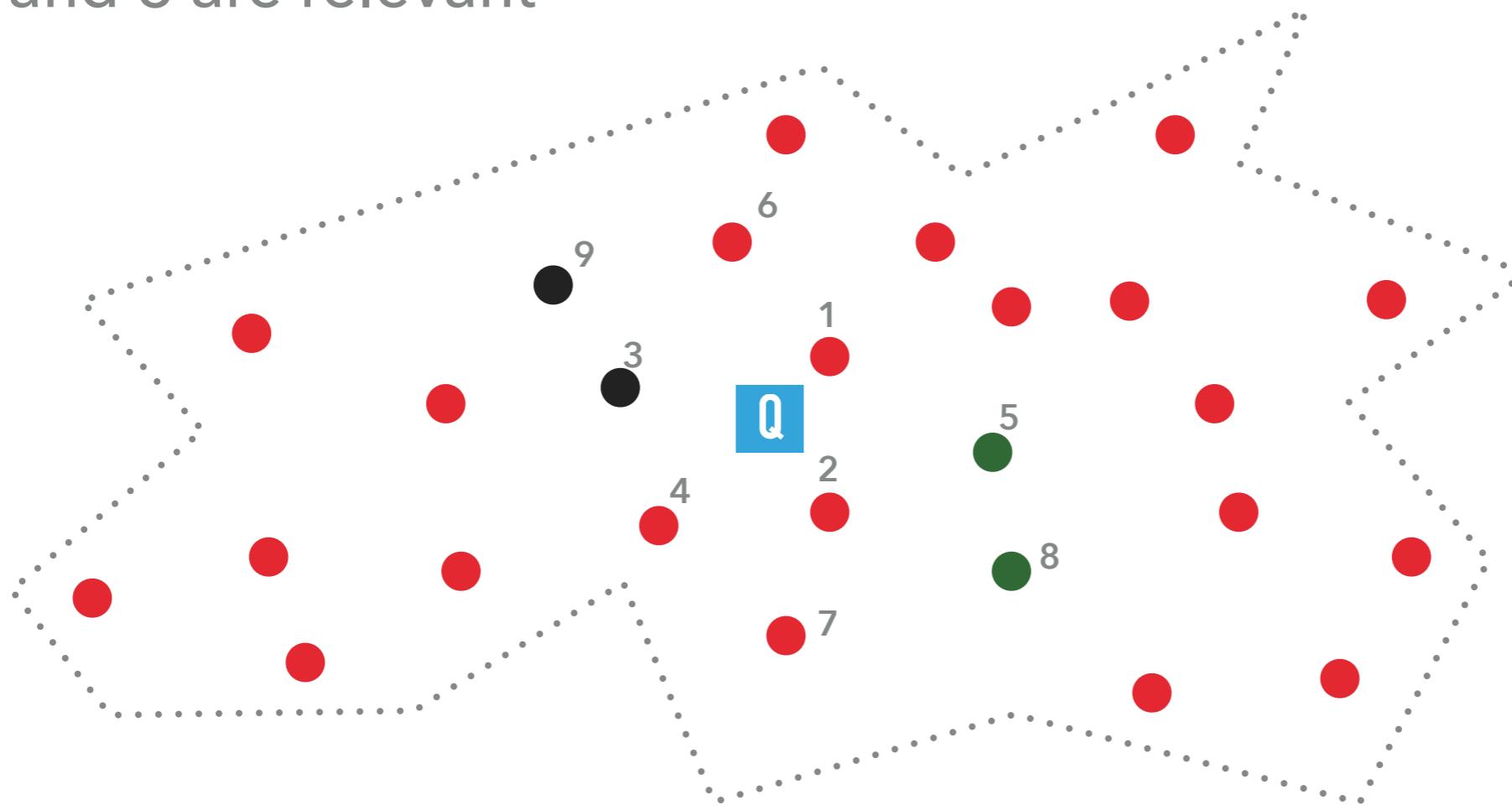
# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶ We can imagine our documents as points in a space
- ▶ Q is an initial query and we rank the documents according to their distance to Q...



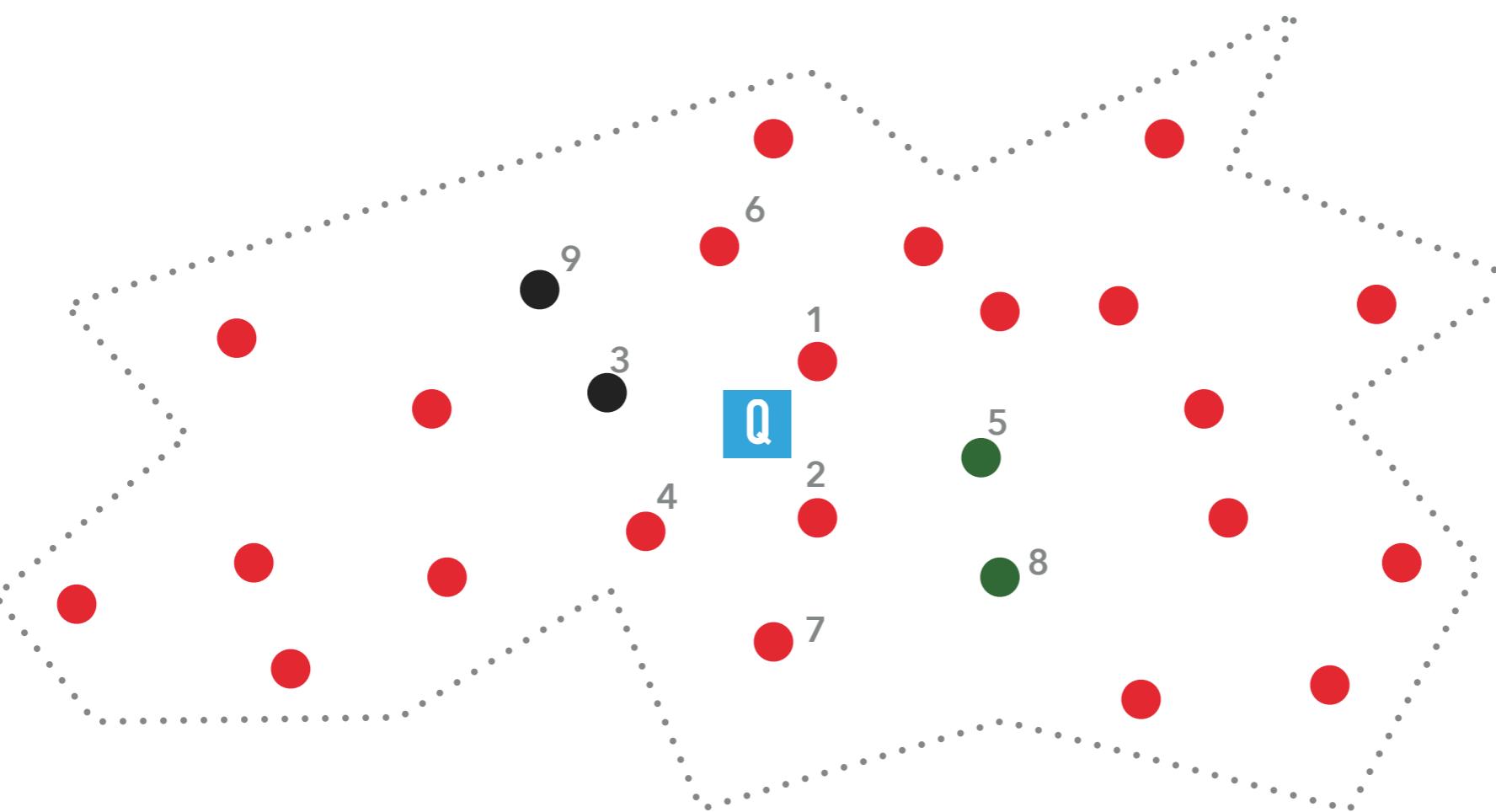
# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶ A user inspects some of the documents and judges them as following:
  - ▶ 3 and 9 are not relevant
  - ▶ 5 and 8 are relevant



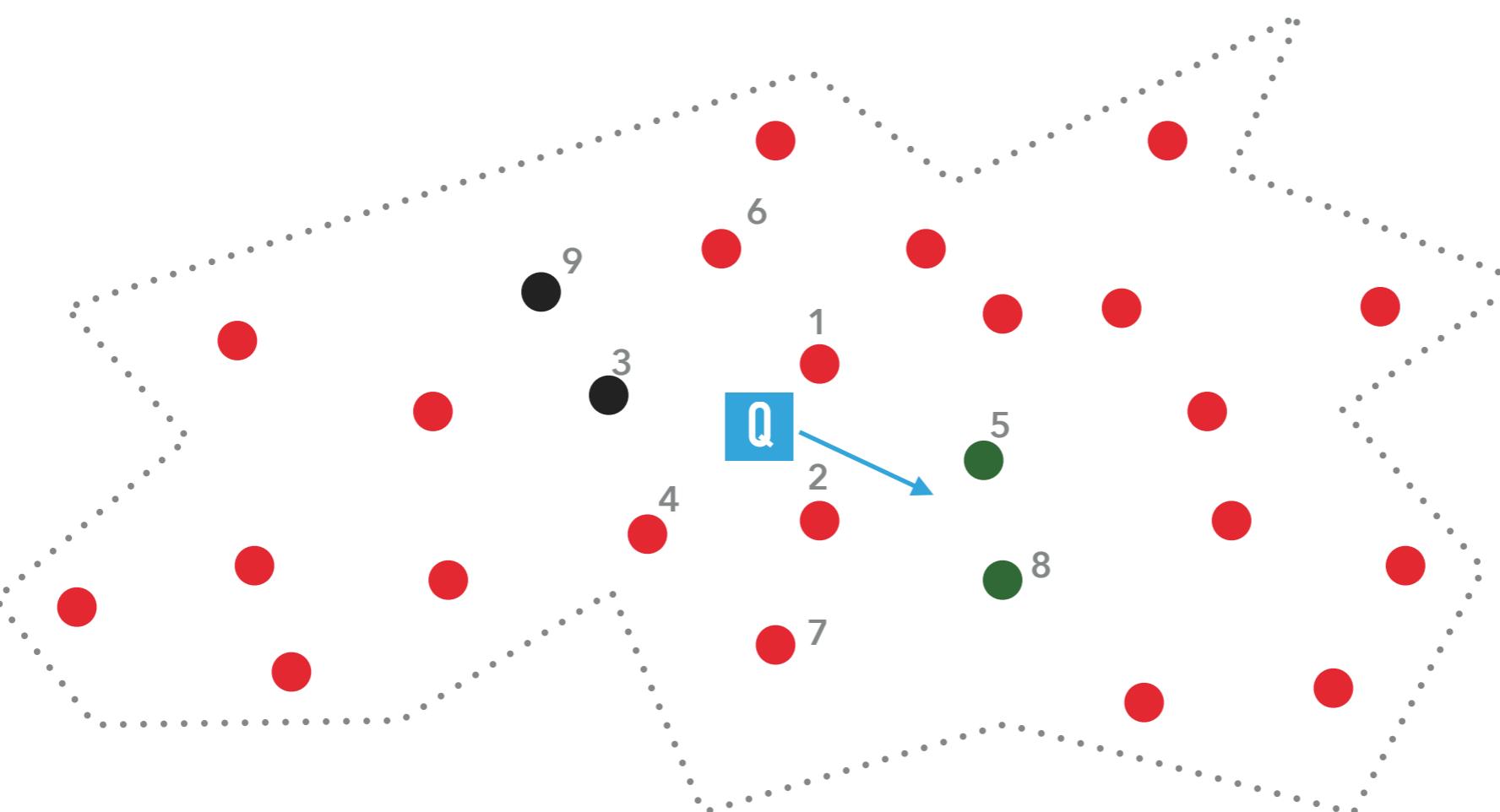
# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶ We use this information to retrieve better documents moving Q to another location in the space closer to the relevant documents and further to the non relevant ones



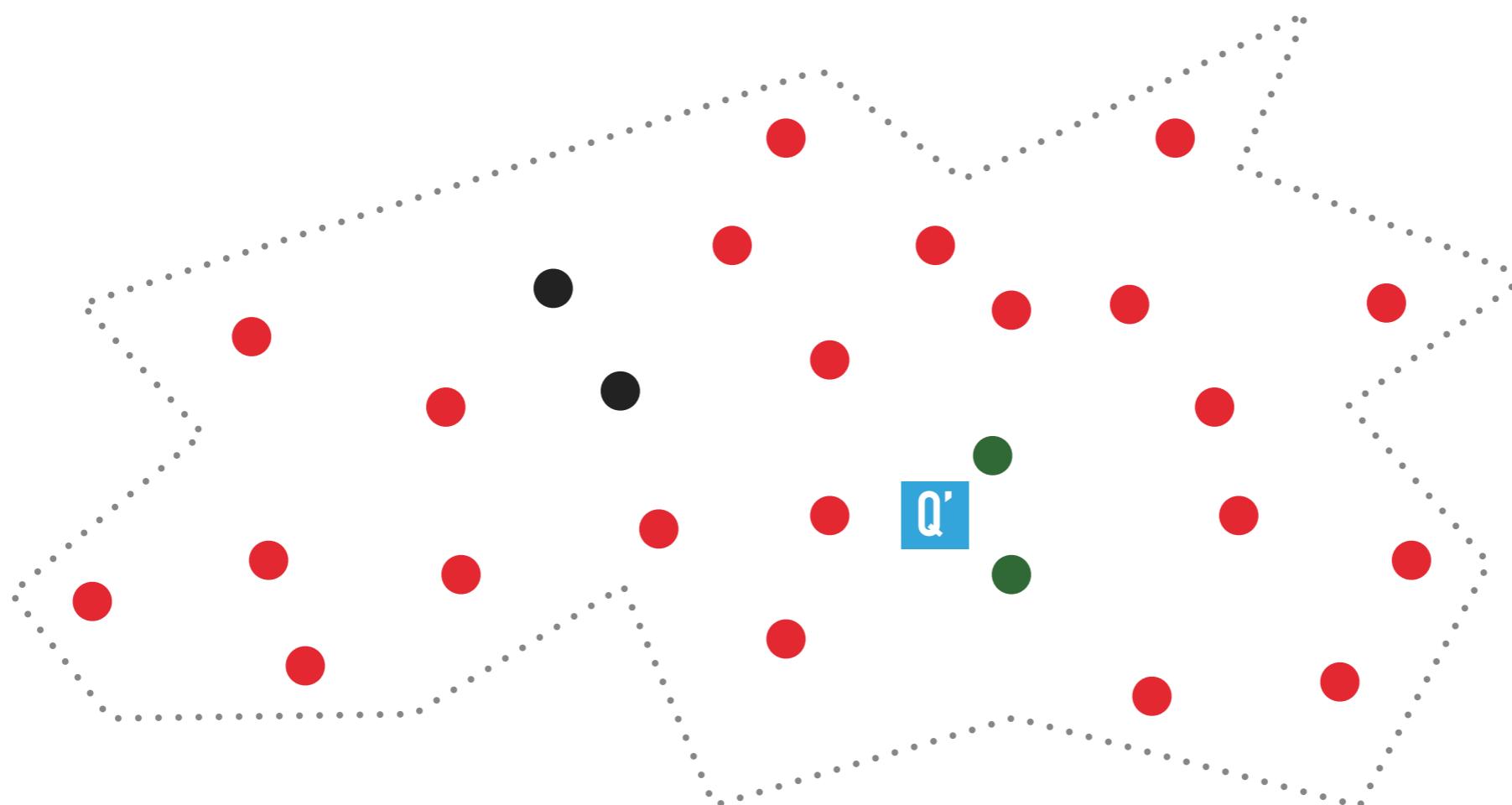
# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶ We use this information to retrieve better documents moving Q to another location in the space closer to the relevant documents and further to the non relevant ones



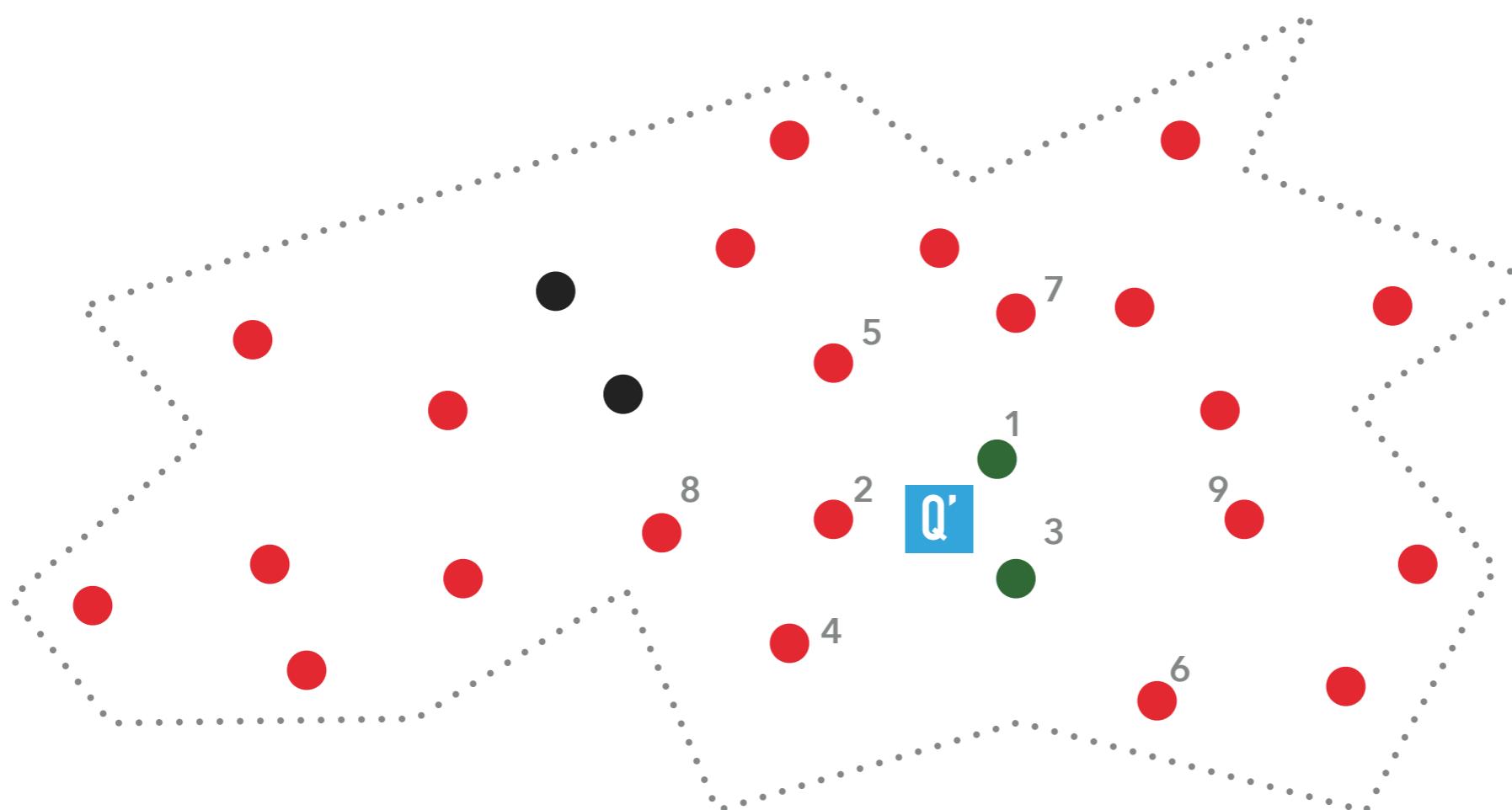
# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶  $Q'$  is the new modified query!



# ROCCCHIO ALGORITHM FOR RELEVANCE FEEDBACK

- ▶  $Q'$  is the new modified query!
- ▶ We rerank the documents according to their proximity to  $Q'$



### MATHEMATICALLY WE HAVE

$$\vec{q}_m = \alpha \vec{q}_0 + \beta \frac{1}{|D_r|} \sum_{\vec{d}_j \in D_r} \vec{d}_j - \gamma \frac{1}{|D_{nr}|} \sum_{\vec{d}_j \in D_{nr}} \vec{d}_j$$

### MATHEMATICALLY WE HAVE

$$\vec{q}_m = \alpha \vec{q}_0 + \beta \frac{1}{|D_r|} \sum_{\vec{d}_j \in D_r} \vec{d}_j - \gamma \frac{1}{|D_{nr}|} \sum_{\vec{d}_j \in D_{nr}} \vec{d}_j$$

### MATHEMATICALLY WE HAVE

$$\vec{q}_m = \boxed{\alpha \vec{q}_0} + \boxed{\beta \frac{1}{|D_r|} \sum_{\vec{d}_j \in D_r} \vec{d}_j} - \boxed{\gamma \frac{1}{|D_{nr}|} \sum_{\vec{d}_j \in D_{nr}} \vec{d}_j}$$

Original Query   Positive Feedback   Negative Feedback

### MATHEMATICALLY WE HAVE

$$\vec{q}_m = \boxed{\alpha \vec{q}_0} + \boxed{\beta \frac{1}{|D_r|} \sum_{\vec{d}_j \in D_r} \vec{d}_j} - \boxed{\gamma \frac{1}{|D_{nr}|} \sum_{\vec{d}_j \in D_{nr}} \vec{d}_j}$$


Original Query

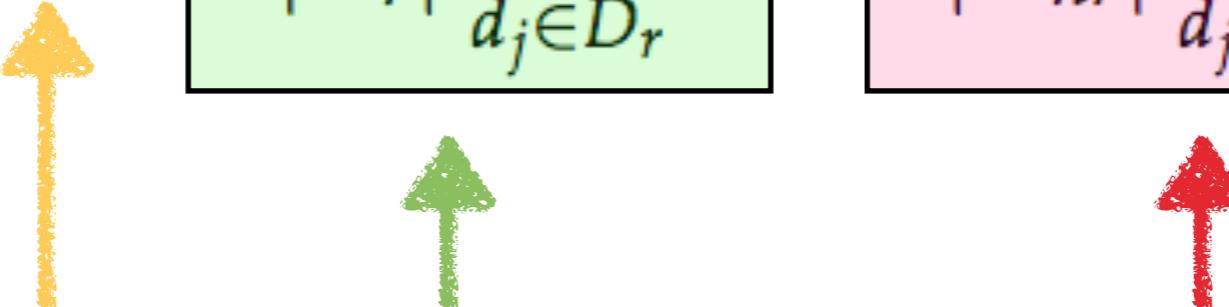
Positive Feedback

Negative Feedback

Moves **towards** the centroid of relevant documents  
 Moves **away** the centroid of relevant documents

Reasonable values for IR: alpha = 1, beta = 0.75, gamma = 0.15

### MATHEMATICALLY WE HAVE

$$\vec{q}_m = \alpha \vec{q}_0 + \beta \frac{1}{|D_r|} \sum_{\vec{d}_j \in D_r} \vec{d}_j - \gamma \frac{1}{|D_{nr}|} \sum_{\vec{d}_j \in D_{nr}} \vec{d}_j$$


Original Query

Positive Feedback

Negative Feedback

Moves **towards** the centroid of relevant documents

Moves **away** the centroid of relevant documents

Reasonable values for IR: alpha = 1, beta = 0.75, gamma = 0.15

Why? Would these values apply for other tasks such as music recommendation?

# LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings

# LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings: your query representation is wrong

# LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings: your query representation is wrong
- ▶ Cross-Language

### LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings: your query representation is wrong
- ▶ Cross-Language: your query representation is wrong

# LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings: your query representation is wrong
- ▶ Cross-Language: your query representation is wrong
- ▶ Vocabulary mismatch

# LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings: your query representation is wrong
- ▶ Cross-Language: your query representation is wrong
- ▶ Vocabulary mismatch: your query representation is wrong

# LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings: your query representation is wrong
- ▶ Cross-Language: your query representation is wrong
- ▶ Vocabulary mismatch: your query representation is wrong
- ▶ Relevant documents should be clustered together...

# LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings: your query representation is wrong
- ▶ Cross-Language: your query representation is wrong
- ▶ Vocabulary mismatch: your query representation is wrong
- ▶ Relevant documents should be clustered together...
- ▶ Examples of problems:
  - ▶ Soviet Union Vs Russia, Persia Vs Iran, Siam Vs Thailand
  - ▶ Disjunctive queries: Nobel prizes that worked at CMU
  - ▶ Disjunctive concepts: Felines, Computer Brands

# LIMITATIONS OF LOCAL METHODS

- ▶ Misspellings: your query representation is wrong
- ▶ Cross-Language: your query representation is wrong
- ▶ Vocabulary mismatch: your query representation is wrong
- ▶ Relevant documents should be clustered together...
- ▶ Examples of problems:
  - ▶ Soviet Union Vs Russia, Persia Vs Iran, Siam Vs Thailand
  - ▶ Disjunctive queries: Nobel prizes that worked at CMU
  - ▶ Disjunctive concepts: Felines, Computer Brands

Extra Limitation: Are users willing to provide feedback???

# PSEUDO RELEVANCE FEEDBACK

Query: New space satellite applications

1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
8. 0.509, 12/02/87, Telecommunications Tale of Two Companies

# PSEUDO RELEVANCE FEEDBACK

Query: New space satellite applications

1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
8. 0.509, 12/02/87, Telecommunications Tale of Two Companies

# PSEUDO RELEVANCE FEEDBACK

Query: New space satellite applications

1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
8. 0.509, 12/02/87, Telecommunications Tale of Two Companies



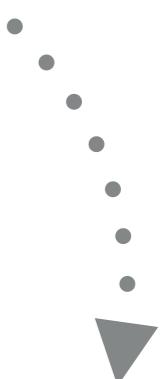
Assume that all top X (e.g. 5) documents are relevant!

# PSEUDO RELEVANCE FEEDBACK

Query: New space satellite applications

1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
8. 0.509, 12/02/87, Telecommunications Tale of Two Companies

Assume that all top X (e.g. 5) documents are relevant!



## VSM MODEL

Re-adjust our query vector representation moving it towards the relevant set

# PSEUDO RELEVANCE FEEDBACK

Query: New space satellite applications

1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
8. 0.509, 12/02/87, Telecommunications Tale of Two Companies

Assume that all top X (e.g. 5) documents are relevant!

### PROBABILISTIC MODEL

Add back to the initial query top Z (e.g. 10) terms in the relevant set

### VSM MODEL

Re-adjust our query vector representation moving it towards the relevant set

# PSEUDO RELEVANCE FEEDBACK

Query: New space satellite applications

1. 0.539, 08/13/91, NASA Hasn't Scrapped Imaging Spectrometer
2. 0.533, 07/09/91, NASA Scratches Environment Gear From Satellite Plan
3. 0.528, 04/04/90, Science Panel Backs NASA Satellite Plan, But Urges Launches of Smaller Probes
4. 0.526, 09/09/91, A NASA Satellite Project Accomplishes Incredible Feat: Staying Within Budget
5. 0.525, 07/24/90, Scientist Who Exposed Global Warming Proposes Satellites for Climate Research
6. 0.524, 08/22/90, Report Provides Support for the Critics Of Using Big Satellites to Study Climate
7. 0.516, 04/13/87, Arianespace Receives Satellite Launch Pact From Telesat Canada
8. 0.509, 12/02/87, Telecommunications Tale of Two Companies



Assume that all top X (e.g. 5) documents are relevant!

### PROBABILISTIC MODEL

Add back to the initial query top Z (e.g. 10) terms in the relevant set

### VSM MODEL

Re-adjust our query vector representation moving it towards the relevant set

Does it work?

Can we evaluate?



heart attack



Web Images News

31,400,000 RESULTS

Date ▾

Language ▾

Region ▾

## [Heart Attacks and Heart Disease – WebMD](#)

[www.webmd.com/heart-disease/guide/heart-disease-heart-attacks](http://www.webmd.com/heart-disease/guide/heart-disease-heart-attacks) ▾

What happens before, during and after a **heart attack**? WebMD explains the causes of **heart attacks**, which affect 1 million Americans every year, along with symptoms ...

## [Myocardial infarction - Wikipedia](#)

[https://en.wikipedia.org/wiki/Myocardial\\_infarction](https://en.wikipedia.org/wiki/Myocardial_infarction) ▾

**Myocardial infarction** (MI) or **acute myocardial infarction** (AMI), commonly known as a **heart attack**, occurs when blood flow stops to a part of the **heart** ...

[Signs and symptoms](#) · [Causes](#) · [Pathophysiology](#) · [Diagnosis](#) · [Prevention](#)

## [What Is a Heart Attack? - NHLBI, NIH](#)

<https://www.nhlbi.nih.gov/health/health-topics/topics/heartattack> ▾

A **heart attack** happens when the flow of oxygen-rich blood to a section of **heart** muscle suddenly becomes blocked and the **heart** can't get oxygen.

## [Videos of heart attack](#)

[bing.com/videos](http://bing.com/videos)



Demi Lovato - **Heart Attack**

[YouTube](#)

Enrique Iglesias - **Heart Attack**

[YouTube](#)

Trey Songz - **Heart Attack [Official]**

[YouTube](#)

Olivia Newton-John - **Heart Attack**

[YouTube](#)

[See more videos of heart attack](#)

## [Demi Lovato - Heart Attack - YouTube](#)

[www.youtube.com/watch?v=AByfaYcOm4A](https://www.youtube.com/watch?v=AByfaYcOm4A) ▾

By DemiLovatoVEVO · 4 min · 437M views

09/04/2013 · Video embedded · Demi's album CONFIDENT available now!

<http://smarturl.it/dls2> Amazon <http://smarturl.it/dlams2> Google Play ...

## [Heart Attack Information - American Heart Association](#)

[www.heart.org/HEARTORG/Conditions/HeartAttack/Heart-Attack](http://www.heart.org/HEARTORG/Conditions/HeartAttack/Heart-Attack_UCM...) UCM... ▾

A **heart attack** is a frightening, confusing and stressful experience. Life after a **heart attack** isn't easy, but we're with you all the way on your journey toward ...

## [Heart attack - Mayo Clinic](#)

[www.mayoclinic.org/diseases-conditions/heart-attack/basics/](http://www.mayoclinic.org/diseases-conditions/heart-attack/basics/) ... ▾

**Heart attack** — Comprehensive overview covers symptoms, causes, treatment of a **heart attack**.



heart attack



Web Images News

31,400,000 RESULTS

Date ▾

Language ▾

Region ▾

## [Heart Attacks and Heart Disease – WebMD](#)

[www.webmd.com/heart-disease/guide/heart-disease-heart-attacks](http://www.webmd.com/heart-disease/guide/heart-disease-heart-attacks) ▾

What happens before, during and after a **heart attack**? WebMD explains the causes of heart attacks, which affect 1 million Americans every year, along with symptoms ...



## [Myocardial infarction - Wikipedia](#)

[https://en.wikipedia.org/wiki/Myocardial\\_infarction](https://en.wikipedia.org/wiki/Myocardial_infarction) ▾

**Myocardial infarction** (MI) or acute **myocardial infarction** (AMI), commonly known as a **heart attack**, occurs when blood flow stops to a part of the **heart** ...

[Signs and symptoms](#) · [Causes](#) · [Pathophysiology](#) · [Diagnosis](#) · [Prevention](#)

## [What Is a Heart Attack? - NHLBI, NIH](#)

<https://www.nhlbi.nih.gov/health/health-topics/topics/heartattack> ▾

A **heart attack** happens when the flow of oxygen-rich blood to a section of **heart** muscle suddenly becomes blocked and the **heart** can't get oxygen.

## [Videos of heart attack](#)

[bing.com/videos](http://bing.com/videos)



Demi Lovato - Heart  
Attack

YouTube

Enrique Iglesias -  
Heart Attack

YouTube

Trey Songz - Heart  
Attack [Official]

YouTube

Olivia Newton-John  
- Heart Attack

YouTube

[See more videos of heart attack](#)

## [Demi Lovato - Heart Attack - YouTube](#)

[www.youtube.com/watch?v=AByfaYcOm4A](https://www.youtube.com/watch?v=AByfaYcOm4A) ▾

By DemiLovatoVEVO · 4 min · 437M views

09/04/2013 · Video embedded · Demi's album CONFIDENT available now!

<http://smarturl.it/dls2> Amazon <http://smarturl.it/dlams2> Google Play ...

## [Heart Attack Information - American Heart Association](#)

[www.heart.org/HEARTORG/Conditions/HeartAttack/Heart-Attack](http://www.heart.org/HEARTORG/Conditions/HeartAttack/Heart-Attack_UCM...) UCM... ▾

A **heart attack** is a frightening, confusing and stressful experience. Life after a **heart attack** isn't easy, but we're with you all the way on your journey toward ...

## [Heart attack - Mayo Clinic](#)

[www.mayoclinic.org/diseases-conditions/heart-attack/basics/](http://www.mayoclinic.org/diseases-conditions/heart-attack/basics/) ... ▾

Heart attack — Comprehensive overview covers symptoms, causes, treatment of a **heart attack**.



heart attack



Web Images News

31,400,000 RESULTS

Date ▾

Language ▾

Region ▾

## [Heart Attacks and Heart Disease – WebMD](#)

[www.webmd.com/heart-disease/guide/heart-disease-heart-attacks](http://www.webmd.com/heart-disease/guide/heart-disease-heart-attacks) ▾

What happens before, during and after a **heart attack**? WebMD explains the causes of heart attacks, which affect 1 million Americans every year, along with symptoms ...

## [Myocardial infarction - Wikipedia](#)

[https://en.wikipedia.org/wiki/Myocardial\\_infarction](https://en.wikipedia.org/wiki/Myocardial_infarction) ▾

**Myocardial infarction** (MI) or acute **myocardial infarction** (AMI), commonly known as a **heart attack**, occurs when blood flow stops to a part of the **heart** ...

[Signs and symptoms](#) · [Causes](#) · [Pathophysiology](#) · [Diagnosis](#) · [Prevention](#)



## [What Is a Heart Attack? - NHLBI, NIH](#)

<https://www.nhlbi.nih.gov/health/health-topics/topics/heartattack> ▾

A **heart attack** happens when the flow of oxygen-rich blood to a section of **heart** muscle suddenly becomes blocked and the **heart** can't get oxygen.

## [Videos of heart attack](#)

[bing.com/videos](http://bing.com/videos)



[Demi Lovato - Heart Attack](#)  
YouTube

[Enrique Iglesias - Heart Attack](#)  
YouTube

[Trey Songz - Heart Attack \[Official\]](#)  
YouTube

[Olivia Newton-John - Heart Attack](#)  
YouTube

[See more videos of heart attack](#)

## [Demi Lovato - Heart Attack - YouTube](#)

[www.youtube.com/watch?v=AByfaYcOm4A](https://www.youtube.com/watch?v=AByfaYcOm4A) ▾

By DemiLovatoVEVO · 4 min · 437M views

09/04/2013 · Video embedded · Demi's album CONFIDENT available now!

<http://smarturl.it/dls2> Amazon <http://smarturl.it/dlams2> Google Play ...

## [Heart Attack Information - American Heart Association](#)

[www.heart.org/HEARTORG/Conditions/HeartAttack/Heart-Attack](http://www.heart.org/HEARTORG/Conditions/HeartAttack/Heart-Attack_UCM...) UCM... ▾

A **heart attack** is a frightening, confusing and stressful experience. Life after a **heart attack** isn't easy, but we're with you all the way on your journey toward ...

## [Heart attack - Mayo Clinic](#)

[www.mayoclinic.org/diseases-conditions/heart-attack/basics/](http://www.mayoclinic.org/diseases-conditions/heart-attack/basics/) ... ▾

Heart attack — Comprehensive overview covers symptoms, causes, treatment of a **heart attack**.



heart attack



Web Images News

31,400,000 RESULTS

Date ▾

Language ▾

Region ▾

## [Heart Attacks and Heart Disease – WebMD](#)

[www.webmd.com/heart-disease/guide/heart-disease-heart-attacks](http://www.webmd.com/heart-disease/guide/heart-disease-heart-attacks) ▾

What happens before, during and after a **heart attack**? WebMD explains the causes of heart attacks, which affect 1 million Americans every year, along with symptoms ...

## [Myocardial infarction - Wikipedia](#)

[https://en.wikipedia.org/wiki/Myocardial\\_infarction](https://en.wikipedia.org/wiki/Myocardial_infarction) ▾

**Myocardial infarction** (MI) or acute **myocardial infarction** (AMI), commonly known as a **heart attack**, occurs when blood flow stops to a part of the **heart** ...

[Signs and symptoms](#) · [Causes](#) · [Pathophysiology](#) · [Diagnosis](#) · [Prevention](#)

## [What Is a Heart Attack? - NHLBI, NIH](#)

<https://www.nhlbi.nih.gov/health/health-topics/topics/heartattack> ▾

A **heart attack** happens when the flow of oxygen-rich blood to a section of **heart** muscle suddenly becomes blocked and the **heart** can't get oxygen.

## [Videos of heart attack](#)

[bing.com/videos](http://bing.com/videos)



Demi Lovato - Heart  
Attack

YouTube

Enrique Iglesias -  
Heart Attack

YouTube

Trey Songz - Heart  
Attack [Official]

YouTube

Olivia Newton-John  
- Heart Attack

YouTube

[See more videos of heart attack](#)

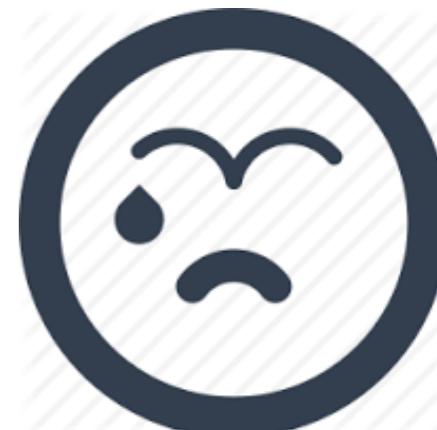
## [Demi Lovato - Heart Attack - YouTube](#)

[www.youtube.com/watch?v=AByfaYcOm4A](https://www.youtube.com/watch?v=AByfaYcOm4A) ▾

By DemiLovatoVEVO · 4 min · 437M views

09/04/2013 · Video embedded · Demi's album CONFIDENT available now!

<http://smarturl.it/dls2> Amazon <http://smarturl.it/dlams2> Google Play ...



## [Heart Attack Information - American Heart Association](#)

[www.heart.org/HEARTORG/Conditions/HeartAttack/Heart-Attack](http://www.heart.org/HEARTORG/Conditions/HeartAttack/Heart-Attack_UCM...) UCM... ▾

A **heart attack** is a frightening, confusing and stressful experience. Life after a **heart attack** isn't easy, but we're with you all the way on your journey toward ...

## [Heart attack - Mayo Clinic](#)

[www.mayoclinic.org/diseases-conditions/heart-attack/basics/](http://www.mayoclinic.org/diseases-conditions/heart-attack/basics/) ... ▾

Heart attack — Comprehensive overview covers symptoms, causes, treatment of a **heart attack**.

Web Images News

48,800,000 RESULTS

Date ▾

Language ▾

Region ▾



### Dogpile - Official Site

[www.dogpile.com](http://www.dogpile.com) ▾

Dogpile.com makes searching the Web easy, because it has all the best search engines piled into one. Go Fetch!



### Google

[www.google.com](http://www.google.com) ▾

Search the world's information, including webpages, images, videos and more. Google has many special features to help you find exactly what you're looking for.



### Yahoo Search - Web Search

[search.yahoo.com](http://search.yahoo.com) ▾

The search engine that helps you find exactly what you're looking for. Find the most relevant information, video, images, and answers from all across the Web.

### Images of search engines

[bing.com/images](http://bing.com/images)



[See more images of search engines](#)



### Web search engine - Wikipedia

[https://en.wikipedia.org/wiki/Web\\_search\\_engine](https://en.wikipedia.org/wiki/Web_search_engine) ▾

A web search engine is a software system that is designed to search for information on the World Wide Web. The search results are generally presented in a ...

[History](#) · [How web search ...](#) · [Market share](#) · [Search engine bias](#)



### Ixquick - Official Site

<https://www.ixquick.com> ▾

StartPage search engine provides search results from over ten best search engines in full privacy. Search anonymously with StartPage Search Engine!



### DuckDuckGo - Official Site

[duckduckgo.com](http://duckduckgo.com)

DuckDuckGo is the search engine that doesn't track you. We protect your search history from everyone – even us!



### Ask.com - Official Site

[www.ask.com](http://www.ask.com)

Offers search for web sites, images, news, blogs, video, maps and directions, local search and shopping.



### Metasearch Search Engine - Search.com

<https://www.search.com> ▾

Search the Web by searching the best engines from one place.



### Google

<https://www.google.co.uk>

Search; Images; Maps; Play; YouTube; News; Gmail; Drive; More. Calendar; Translate; Books; Shopping; Blogger; ... Advertising Programmes Business Solutions +Google ...

[Web](#) [Images](#) [News](#)

48,800,000 RESULTS

Date ▾

Language ▾

Region ▾

**Dogpile - Official Site**[www.dogpile.com](http://www.dogpile.com) ▾

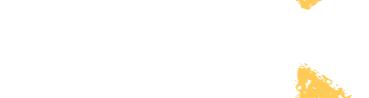
Dogpile.com makes searching the Web easy, because it has all the best search engines piled into one. Go Fetch!

**Google**[www.google.com](http://www.google.com) ▾

Search the world's information, including webpages, images, videos and more. Google has many special features to help you find exactly what you're looking for.

**Yahoo Search - Web Search**[search.yahoo.com](http://search.yahoo.com) ▾

The search engine that helps you find exactly what you're looking for. Find the most relevant information, video, images, and answers from all across the Web.

**Images of search engines**[bing.com/images](http://bing.com/images)[See more images of search engines](#)[Web](#) [Images](#) [News](#)

11-20 OF 48,800,000 RESULTS Date ▾ Language ▾ Region ▾

**Top 15 Most Popular Search Engines | March 2017**[www.ebizmba.com/articles/search-engines](http://www.ebizmba.com/articles/search-engines) ▾

Here are the top 15 Most Popular Search Engines ranked by a combination of continually updated traffic statistics.

**Bing**[https://www.bing.com](http://https://www.bing.com)

Bing helps you turn information into action, making it faster and easier to go from searching to doing.

**Second page?****Web search engine - Wikipedia**[https://en.wikipedia.org/wiki/Web\\_search\\_engine](https://en.wikipedia.org/wiki/Web_search_engine) ▾

A web search engine is a software system that is designed to search for information on the World Wide Web. The search results are generally presented in a ...

[History](#) · [How web search ...](#) · [Market share](#) · [Search engine bias](#)**Ixquick - Official Site**<https://www.ixquick.com> ▾

StartPage search engine provides search results from over ten best search engines in full privacy. Search anonymously with StartPage Search Engine!

**DuckDuckGo - Official Site**[duckduckgo.com](https://duckduckgo.com)

DuckDuckGo is the search engine that doesn't track you. We protect your search history from everyone – even us!

**Ask.com - Official Site**[www.ask.com](http://www.ask.com)

Offers search for web sites, images, news, blogs, video, maps and directions, local search and shopping.

**Metasearch Search Engine - Search.com**<https://www.search.com> ▾

Search the Web by searching the best engines from one place.

**Google**<https://www.google.co.uk>

Search; Images; Maps; Play; YouTube; News; Gmail; Drive; More. Calendar; Translate; Books; Shopping; Blogger; ... Advertising Programmes Business Solutions +Google ...

**Top 15 Most Popular Search Engines | March 2017**[www.ebizmba.com/articles/search-engines](http://www.ebizmba.com/articles/search-engines) ▾

---

# DISCUSSION

# WHAT DID WE SEE? WHAT SHOULD YOU KNOW?

## ► Query Modification (mostly expansion):

### 1. Using the retrieved documents:

- Relevance Feedback (Explicitly/Implicitly)

### 2. Not using the retrieved documents:

- Explores word relationships: synonyms, co-occurrence, so on

### DISCUSSION

- ▶ Query expansion is a great way to increase the number of retrieved documents that might be useful to the user (increase recall)
- ▶ However, we are risking to include garbage:
  - ▶ Query drift: hurt precision
  - ▶ Demi Lovato when search for heart attack? Seriously?

# DISCUSSION

- ▶ Query expansion with explicit relevance feedback
  - ▶ by far the best one
  - ▶ by far the hardest one (users...)
- ▶ Query expansion with word relationships (synonyms/hypernyms...) Vs Query expansion with pseudo relevance feedback

### DISCUSSION

- ▶ Query expansion with explicit relevance feedback
    - ▶ by far the best one
    - ▶ by far the hardest one (users...)
  - ▶ Query expansion with word relationships (synonyms/hypernyms...) Vs Query expansion with pseudo relevance feedback
- 
- The diagram consists of two curved red arrows. One arrow originates from the text 'Easier to explain' and points to the statement 'Query expansion with explicit relevance feedback'. The other arrow originates from the text 'Better overall empirical results' and points to the statement 'Query expansion with word relationships (synonyms/hypernyms...) Vs Query expansion with pseudo relevance feedback'.

### TODAY'S LECTURE IN THE STANFORD IR BOOK

- ▶ Chapter 9 - Relevance feedback and query expansion
- ▶ Chapter 9.1: Relevance feedback and pseudo relevance feedback
- ▶ Chapter 9.2: Global methods for query reformulation