

**POISON** PILLS AND ANTIDOTES,

---

**INOCULATING RELEVANCE  
FEEDBACK**

# FROM COMPLEX INFORMATION NEEDS TO THE RELEVANT RESULTS

---







Google Search

I'm Feeling Lucky

Google.nl offered in: [Nederlands](#)

## POISON PILLS

- ▶ Despite the fact that they are clicked as relevant pages, they **hurt** the performance of feedback!
  - ▶ Off-topic terms



## EXAMPLE

---



nobel prize winners

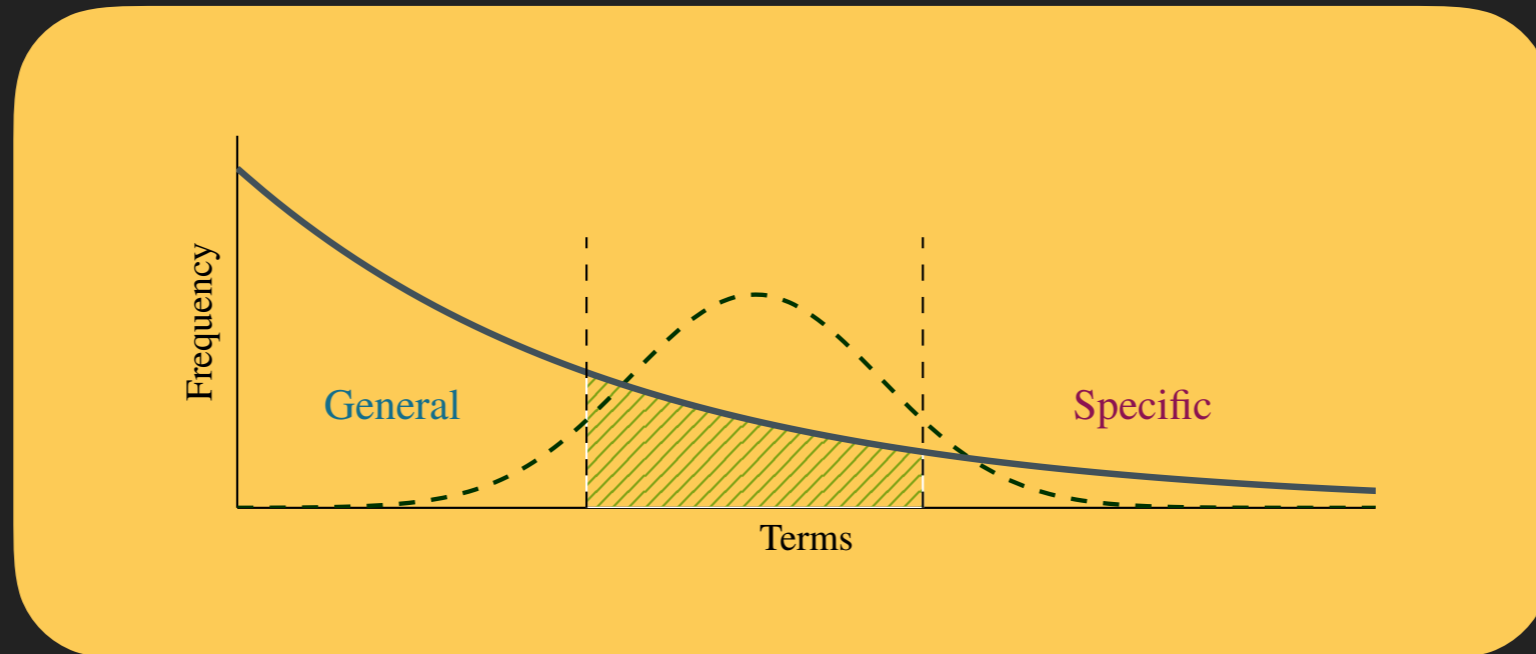


- ▶ One of the pages you clicked is about Yasser Arafat (1994)
  - ▶ it has a long discussion about "middle east issues"
    - ▶ Most frequent words: Palestine, Israel, attack, ...

**TOPIC DRIFT**

# SIGNIFICANT WORDS

---



# TOY EXAMPLE

QUERY: "Q" 🔍

**D1**  
Q  
A A A A  
B B  
D D D D

**D2**  
Q  
X X  
Y W  
D D D D D

**D3**  
Q  
B B  
C C C  
D D D D D

**D4**  
Q  
S T  
U V  
D D D D D

CLICKED

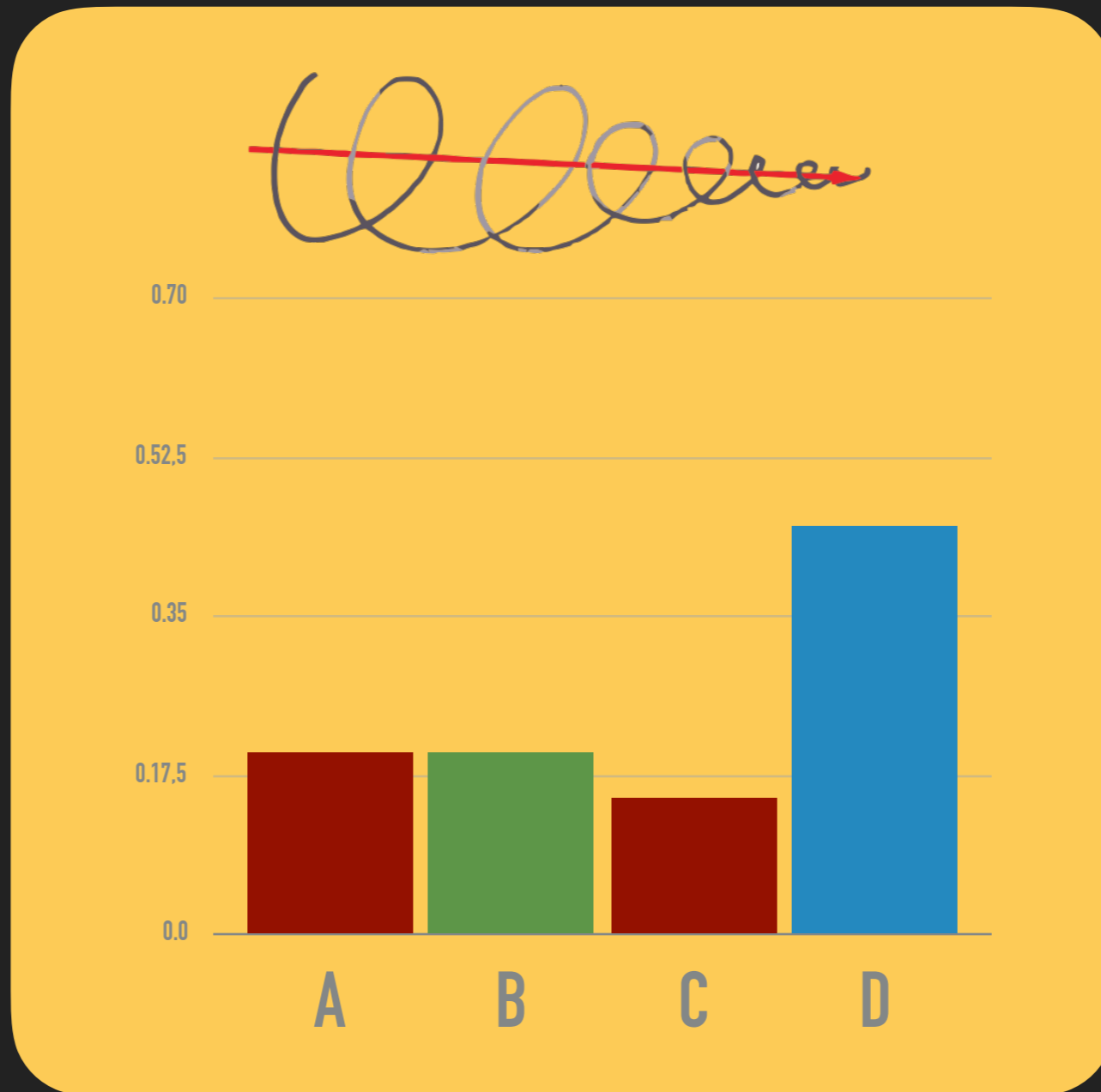
Q Q  
A A A A  
B B B B  
C C C  
D D D D D D D D

NEW EXPANDED QUERY:  
"Q B"



# SIGNIFICANT WORDS

---



**DON'T BUILD YOUR MODEL BASED ON  
PROPERTYLESS COMMON OBSERVATIONS  
OR POISONOUS RARE OBSERVATIONS,**

my model says