

Advances in IR Evaluation

Ben Carterette

Evangelos Kanoulas

Emine Yilmaz



The
University
Of
Sheffield.

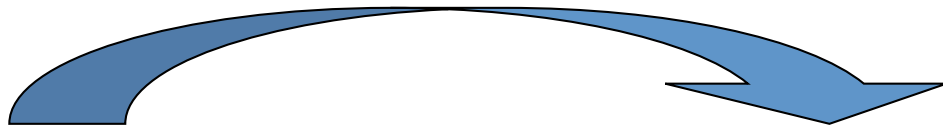


Course Outline

- Intro to evaluation
 - Evaluation methods, test collections, measures, comparable evaluation
- Low cost evaluation
- Advanced user models
 - Web search models, novelty & diversity, sessions
- Reliability
 - Significance tests, reusability
- Other evaluation setups

Traditional Experiment

- uses of alternative dispute resolution
- job search vancouver washington
- poem of arrival of columbus



Results

Web

[12 CIKM 2003, New Orleans, Louisiana, USA](#)
 Proceedings of the 2003 ACM CIKM International Conference on Information and Knowledge Management, New Orleans, Louisiana, USA, November 2-8, 2003. ...
[www.informatik.uns-trsr.de/~toydb/conf/cikm/cikm2003.html](#) - 52k - [Cached](#) - [Similar pages](#)

[CFP: CIKM 2003, msg#00007](#)
 Only the highest caliber papers submitted to CIKM 2003 will be accepted. We have a special interest in papers that bridge the areas of databases and ...
[csdlr.com/ml/information-retrieval_web/2003-03/msg00007.html](#) - 21k - [Cached](#) - [Similar pages](#)

[CIKM 2003 Technical Program, msg#00001](#)
 Twelfth International Conference on Information and Knowledge Management CIKM 2003
 November 3-8, 2003 Hotel Inter-Continental New Orleans, LA, ...
[csdlr.com/ml/information-retrieval_web/2003-03/msg00001.html](#) - 25k - [Cached](#) - [Similar pages](#)
[More results from csdlr.com »](#)

[ACM CIKM 2003: Preliminary call for papers](#)
 ACM CIKM 2003: PRELIMINARY CALL FOR PAPERS, Twelfth International Conference on Information and Knowledge Management (CIKM), November 2-8, 2003 ...
[do.wiley.com/10.1002/asi.10232](#) - [Similar pages](#)
 by I Dates - 2003

[cikm 2003](#)
 The CIKM 2003 web page, The CIKM 2005 web page, The CIKM 2004 web page, The CIKM 2003 Web Page, The CIKM 2002 Web Page, The CIKM 2001 Web page, The CIKM ...
[www.splyfu.com/Term.aspx?i=314867](#) - 76k - [Cached](#) - [Similar pages](#)

[Multi-Objective Optimisation for Information Access Tasks DRAFT](#) ...
 2003 ABSTRACT, Download: pdf, by Michelle J. Fisher, Jonathan E. Fieldsend ...
[citeseer.ist.psu.edu/061812.html](#) - 25k - [Cached](#) - [Similar pages](#)
 by MJ Fisher - All 2 versions

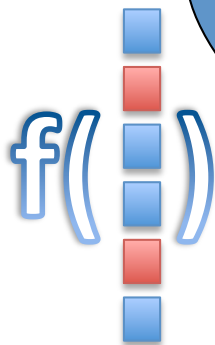
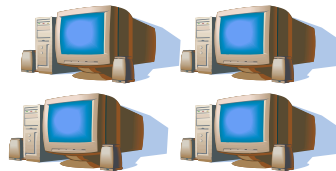
[12 CIKM 2003 New Orleans, Louisiana, USA](#)
 proceedings (DBLPconf/cikm2003) (sic) (Proceedings of the 2003 ACM CIKM International Conference on Information and Knowledge Management New Orleans Louisiana ...
[vidbarc.org/dblp/db/conf/cikm/cikm2003.html](#) - 52k - [Cached](#) - [Similar pages](#)

[ESRI Training and Education- Library, Conference Proceedings, ACM ...](#)
 Giordano Adam, 2003 ACM CIKM International Conference, 2003. Categorizing web queries according to geographical locality. Luis Gravano, 2003 ACM CIKM ...
[training.esri.com/campus/library/bibliography/Browse.cfm?referencetype=&libid=section=Conference%2CFProc...](#) - 32k - [Cached](#) - [Similar pages](#)

[CIKM '03 Proceedings of the twelfth international conference on ...](#)
 The Organizing Committee, as well as the sponsors of CIKM 2003, ... To the extent that we are successful in these goals, CIKM 2003 will have served its ...
[portal.acm.org/citation.cfm?id=956863&dl=portal&dl=ACM&type=proceeding&id=956863&par...](#) - [Similar pages](#)

[Conference on Information and Knowledge Management \(CIKM\)](#)
 CIKM has a strong tradition of workshops devoted to emerging areas of database ... The CIKM 2005 web page, The CIKM 2004 web page, The CIKM 2003 Web Page ...
[www.ce.umbc.edu/cikm/](#) - 7k - [Cached](#) - [Similar pages](#)

Search Engines



Judges



Today's Outline

- Evaluation based on user models
 - Gain/Discount Framework
 - Modeling browsing behavior
 - RBP, AP, DCG, ERR, EBU
 - Modeling utility
 - NDCG, GAP
- Novelty and diversity
- Query sessions

Measures of Effectiveness

- Determine how good the system is at finding and ranking relevant documents
- An effectiveness measure should be correlated to the user's experience
- Thus interest in effectiveness measures based on explicit models of user interaction

Discounted Gain Model

- Simple model of user interaction:
 - User steps down ranked results one-by-one
 - Gains something from relevant documents
 - Increasingly less likely to see documents deeper in the ranking

Discounted Gain Model

1. What is the chance a user will read a document?
2. What does the user gain by reading a document?

Model Browsing Behavior

black powder
ammunition

1

Position-based models

The chance of observing a document depends on the position of the document in the ranked list.

1

[Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able ammunition, hunting supplies, shooti unbeatable prices.
www.ableammo.com/catalog/default.js

2

[Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smokele largest on hand inventory of obsolete old, obsolete and ...
[www.buffaloarms.com/browse.cfm?2%](http://www.buffaloarms.com/browse.cfm?2%2)

3

[www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bjACTION=enter&thispage=index...>

4

[Black Powder Hunting, Shooting, and](#)
Count on Able Ammo to provide you shooting supplies, gun accessories, i
www.ableammo.com/catalog/default.js

5

[Black Dawge Cartridge - Black powder](#)
Black Dawge Cartridges and black powder June 14, 2011. Changes in the .45 caliber .45 cal, 450 ...
www.blackdawgecartridge.com

6

[Black Powder Ammunition - Comparison](#)
Black Powder Ammunition - 154 results Explosion, 20 Sticks, Avery Outdoor Series Black ...
www.nexttag.com/black-powder-amm

7

[Buffalo Arms - Ammunition - Black Powder](#)
WARNING: Discharging firearms in public handling ammunition may result in explosion
www.buffaloarms.com/browse.cfm?2,3

8

[Black powder ammunition | Shop by price](#)
Black powder ammunition - Find the ammunition on sale. Shop by price, coupons, and deals at ...
www.thefind.com/sports/info-black-powder

9

[Black Powder Ammunition - Sportsmans](#)
Black powder ammunition, bullets, and outdoor superstore your one stop hunting ammunition
www.sportsmansoutdoorsuperstore.com/ammunition

10

[Murphy Powder Ammunition - W.A. Murphy](#)
black, powder, loaded, ammunition . per box: 32-20 wcf 115 mfp; 26.50; 3 145 mfp
www.murhypowder.com/ammunition

...

Rank Biased Precision

[Moffat and Zobel, TOIS08]

black powder ammunition

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- ...

[Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able
ammunition, hunting supplies, shooti
unbeatable prices.
www.ableammo.com/catalog/default.js

[Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smokele
largest on hand inventory of obsolete
old, obsolete and ...
[www.buffaloarms.com/browse.cfm?2%](http://www.buffaloarms.com/browse.cfm?2%26)

[www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bj>
ACTION=enter&thispage=index...

[Black Powder Hunting, Shooting, and](#)
Count on Able Ammo to provide you
shooting supplies, gun accessories, i
www.ableammo.com/catalog/default.js

[Black Dawge Cartridge - Black powder](#)
Black Dawge Cartridges and black po
June 14, 2011. Changes in the .45 ca
.45 cal, 450 ...
www.blackdawgecartridge.com

[Black Powder Ammunition - Compar](#)
Black Powder Ammunition - 154 resi
Explosion, 20 Sticks, Avery Outdoor
Series Black ...
www.nextag.com/black-powder-amm

[Buffalo Arms - Ammunition - Black P](#)
WARNING: Discharging firearms in pc
handling ammunition may result in ex
www.buffaloarms.com/browse.cfm?2,3

[Black powder ammunition | Shop bl](#)
Black powder ammunition - Find the
ammunition on sale. Shop by price, c
coupons, and deals at ...
www.thefind.com/sports/info-black-po

[Black Powder Ammunition - Sports](#)
Black powder ammunition, bullets, t
outdoor superstore your one stop hur
www.sportsmansoutdoorsuperstore.co
ammunition

[Murphy Powder Ammunition - W.A. I](#)
black, powder, loaded, ammunition .
per box: 32-20 wcf 115 mfp; 26.50; 3
145 mfp
www.murhypowder.com/ammunition

Toss a biased coin (θ)

If HEADS, observe next
document

IF TAILS, stop

Rank Biased Precision

black powder
ammunition

Let $\theta=0.8$

- 1 [Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able
ammunition, hunting supplies, shooti
unbeatable prices.
- 2 [Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smokele
largest on hand inventory of obsolete
old, obsolete and ...
[www.buffaloarms.com/browse.cfm?2%](http://www.buffaloarms.com/browse.cfm?2%...)
- 3 [www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bj...>
ACTION=enter&thispage=index...
- 4 [Black Powder Hunting, Shooting, and](#)
Count on Able Ammo to provide you
shooting supplies, gun accessories, i
www.ableammo.com/catalog/default.p...
- 5 [Black Dawge Cartridge - Black powd](#)
Black Dawge Cartridges and black pc
June 14, 2011. Changes in the .45 ca
.45 cal, 450 ...
www.blackdawgecartridge.com
- 6 [Black Powder Ammunition - Compar](#)
Black Powder Ammunition - 154 res
Explosion, 20 Sticks, Avery Outdoor
Series Black ...
www.nextag.com/black-powder-amm...
- 7 [Buffalo Arms - Ammunition - Black P](#)
WARNING: Discharging firearms in pc
handling ammunition may result in e
www.buffaloarms.com/browse.cfm?2,3
- 8 [Black powder ammunition | Shop bl](#)
Black powder ammunition - Find the
ammunition on sale. Shop by price, c
coupons, and deals at ...
www.thefind.com/sports/info-black-pc...
- 9 [Black Powder Ammunition - Sports](#)
Black powder ammunition, bullets, b
outdoor superstore your one stop hur
www.sportsmansoutdoorsuperstore.co...
ammunition
- 10 [Murphy Powder Ammunition - W.A. |](#)
black, powder, loaded, ammunition .
per box: 32-20 wcf 115 mfp; 26.50: 3
145 mfp
www.murphy powder.com/ammunition
- ...



$$0.532 < \theta$$

$$0.933 \geq \theta$$

Rank Biased Precision

black powder
ammunition

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- ...

[Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able
ammunition, hunting supplies, shoot
 unbeatable prices.
www.ableammo.com/catalog/default.js

[Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smokele
 largest on hand inventory of obsolete
 old, obsolete and ...
[www.buffaloarms.com/browse.cfm?2%](http://www.buffaloarms.com/browse.cfm?2%2)

[www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bj>
 ACTION=enter&thispage=index...

[Black Powder Hunting, Shooting, and](#)
 Count on Able Ammo to provide you
 shooting supplies, gun accessories, i
www.ableammo.com/catalog/default.js

[Black Dawge Cartridge - Black powder](#)
Black Dawge Cartridges and **black po**
 June 14, 2011. Changes in the .45 ca
 .45 cal, 450 ...
www.blackdawgecartridge.com

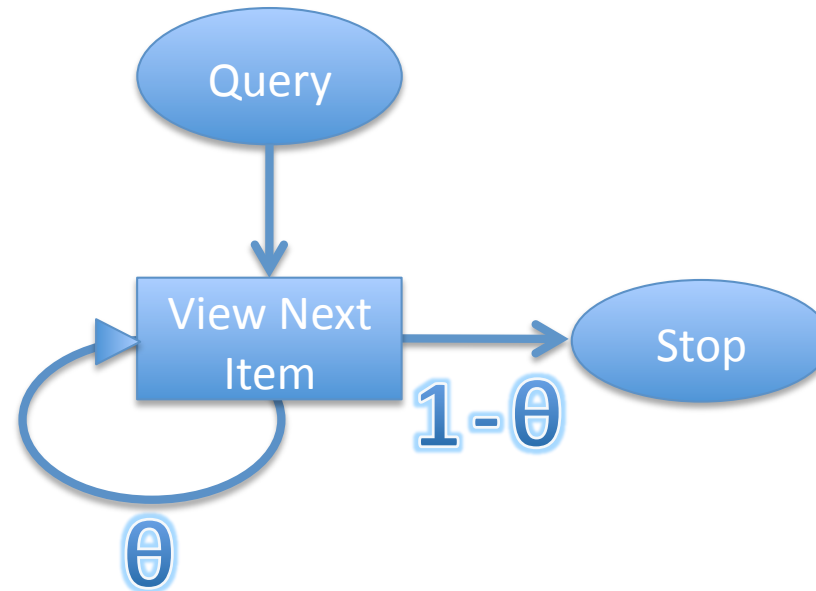
[Black Powder Ammunition - Compar](#)
Black Powder Ammunition - 154 res
 Explosion, 20 Sticks, Avery Outdoor
 Series **Black** ...
www.nextag.com/black-powder-amm

[Buffalo Arms - Ammunition - Black P](#)
 WARNING: Discharging firearms in pc
 handling **ammunition** may result in e
www.buffaloarms.com/browse.cfm?2,3

[Black powder ammunition | Shop bl](#)
Black powder ammunition - Find the
ammunition on sale. Shop by price, c
 coupons, and deals at ...
www.thefind.com/sports/info-black-pc

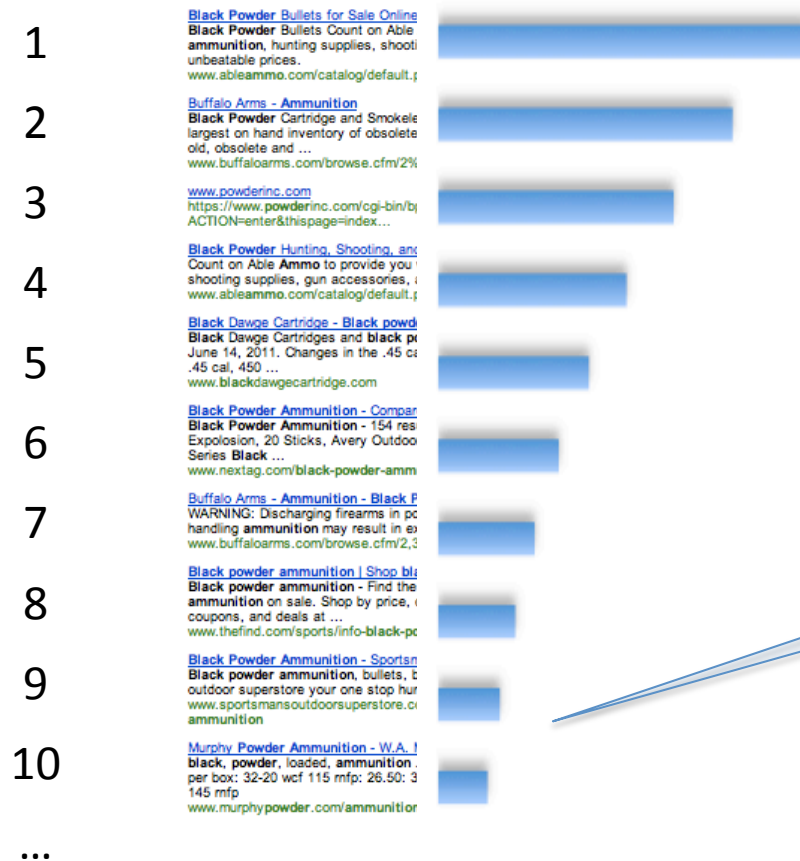
[Black Powder Ammunition - Sports](#)
Black powder ammunition, bullets, b
 outdoor superstore your one stop hur
www.sportsmansoutdoorsuperstore.co
ammunition

[Murphy Powder Ammunition - W.A. I](#)
black, powder, loaded, **ammunition** .
 per box: 32-20 wcf 115 mfp; 26.50: 3
 145 mfp
www.murphy powder.com/ammunition



Rank Biased Precision

black powder ammunition



$$RBP = (1 - \vartheta) \sum_{i=1}^{\infty} rel_i \vartheta^{k-1}$$

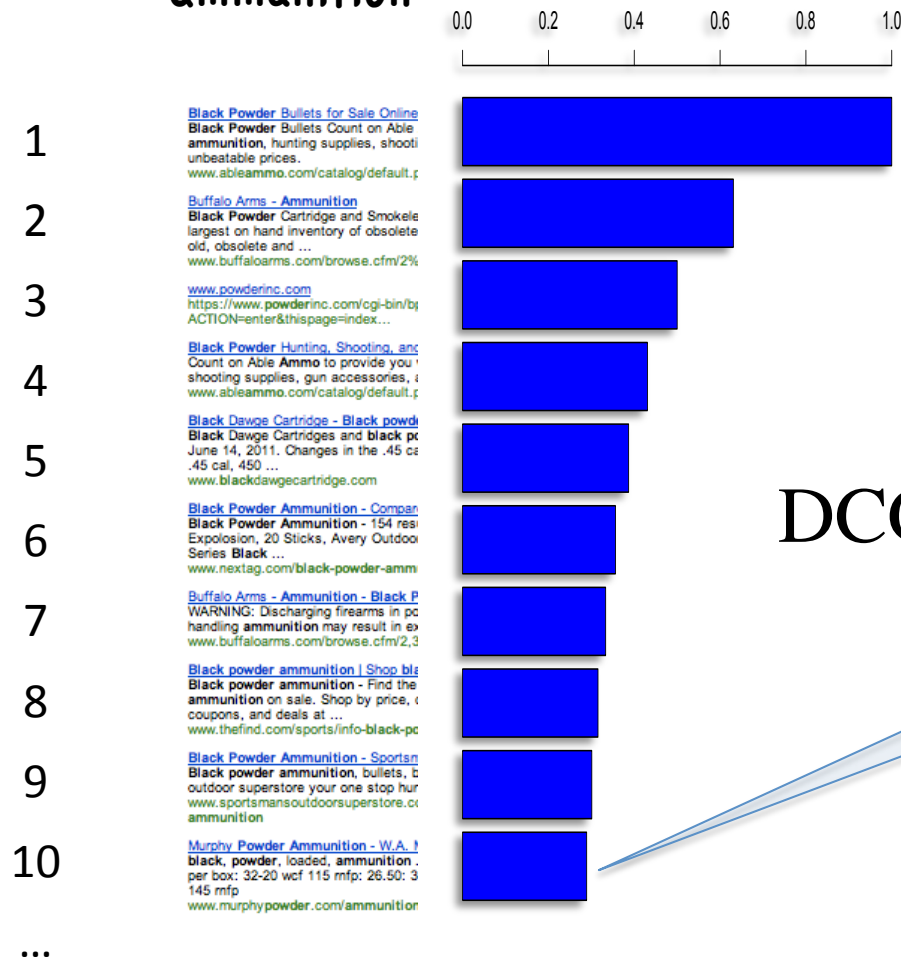
$$= \sum_{i=1}^{\infty} rel_i \vartheta^{k-1} (1 - \vartheta)$$

Relevance discounted by geometric distribution

Discounted Cumulative Gain

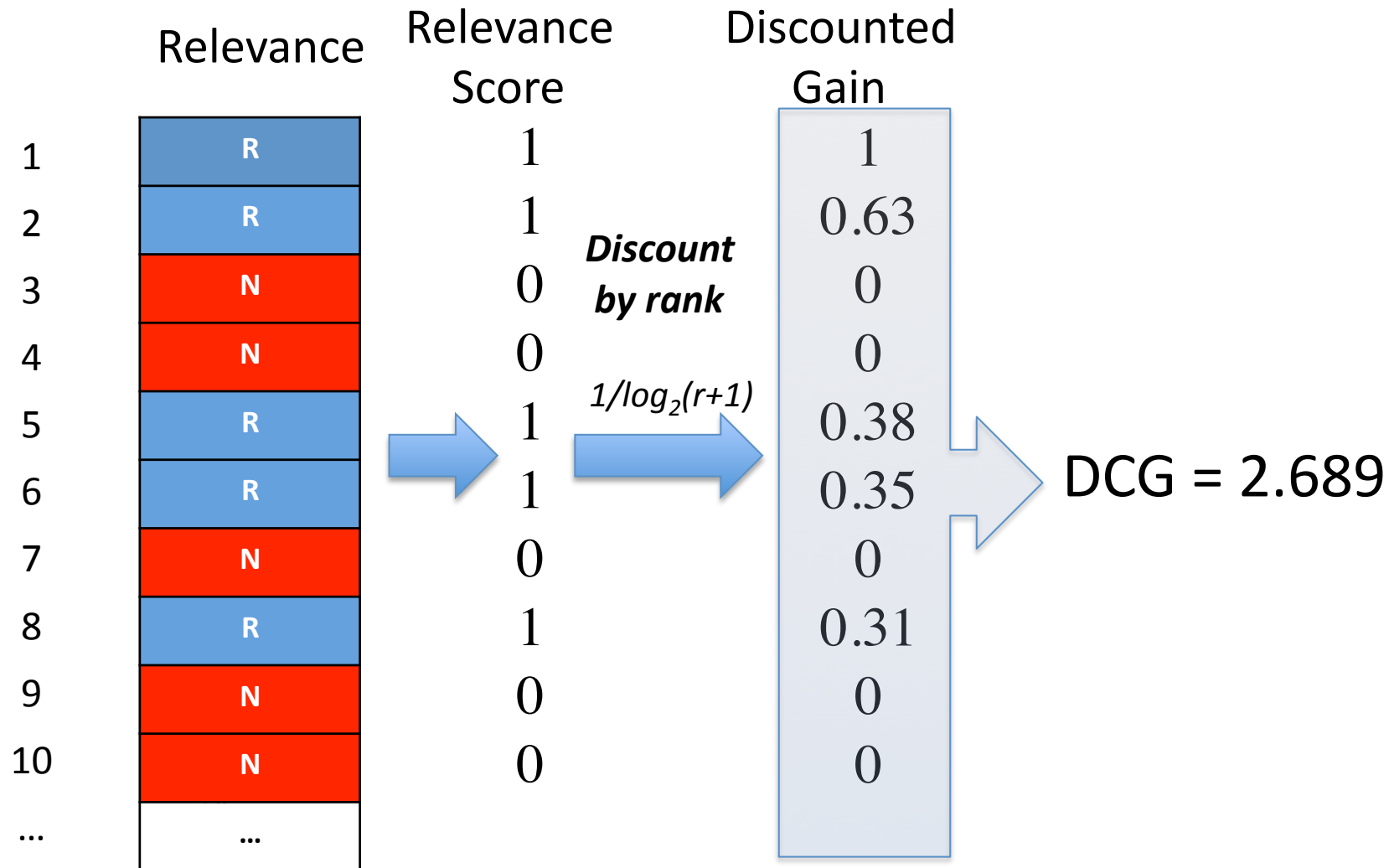
[Järvelin and Kekäläinen SIGIR00]

black powder ammunition

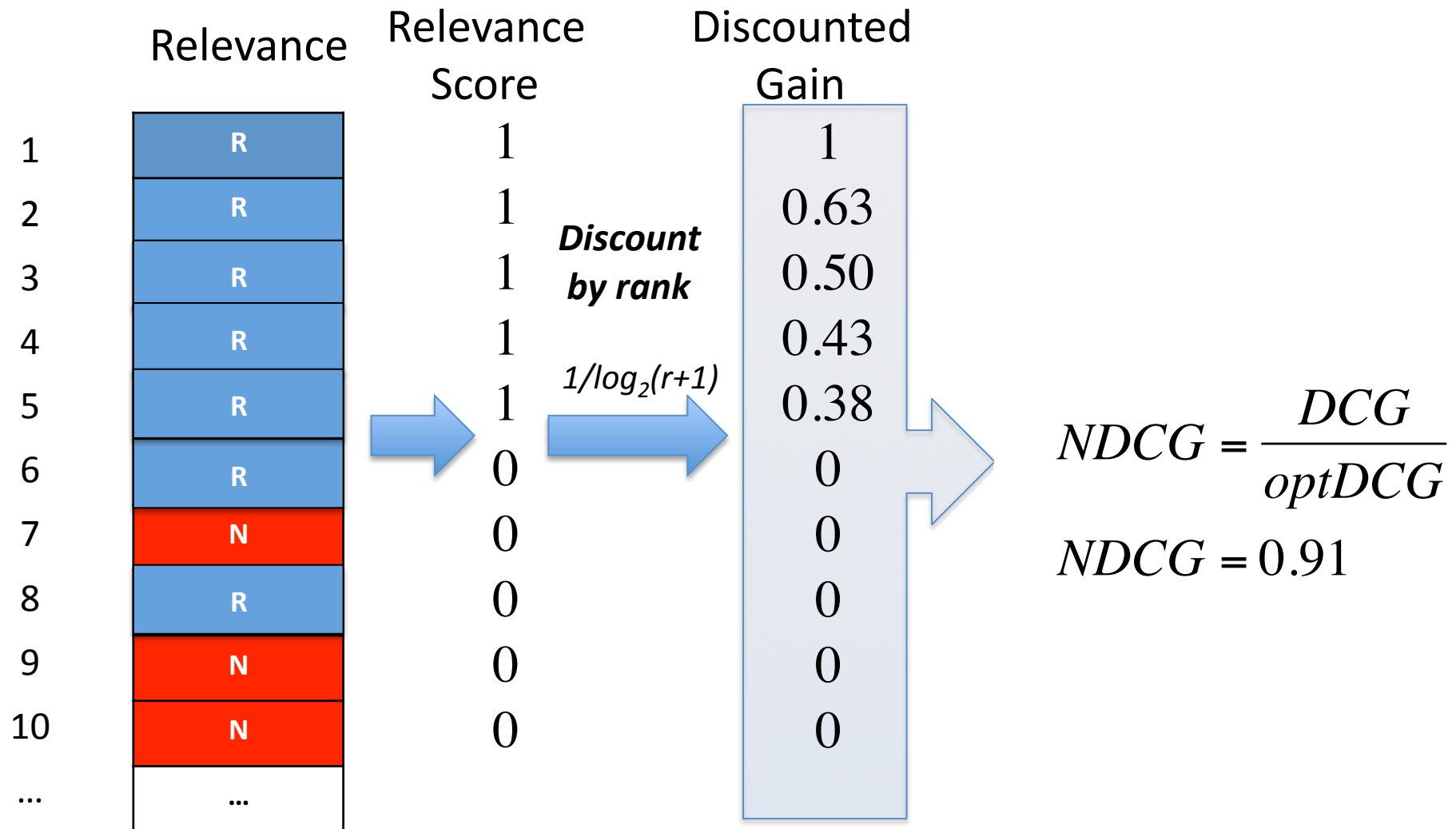


$$DCG = \sum_{i=1}^{\infty} rel_i \frac{1}{\log_2(1+i)}$$

Discounted Cumulative Gain



Discounted Cumulative Gain



Discounted Cumulative Gain

- DCG can be written as:

$$\sum_{r=1}^N P(\text{user visits doc } r) \cdot \text{Utility}_r$$

- Discount function models the probability that the user visits (clicks on) the document at rank r
 - Currently, $P(\text{user clicks on doc } r) = \log_2(r+1)$

Discounted Cumulative Gain

- Instead of stopping probability, think about viewing probability

$$P(\text{view doc at } k) = \sum_{i=k}^{\infty} P(i) = F(k)$$

- This fits in discounted gain model framework:

$$M = \sum_{k=1}^{\infty} rel_k F(k)$$

Model Browsing Behavior

black powder
ammunition

1

Position-based models

The chance of observing a document depends on the position of the document in the ranked list.

1

[Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able ammunition, hunting supplies, shooti unbeatable prices.
www.ableammo.com/catalog/default.js

2

[Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smokele largest on hand inventory of obsolete old, obsolete and ...
[www.buffaloarms.com/browse.cfm?2%](http://www.buffaloarms.com/browse.cfm?2%2)

3

[www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bjACTION=enter&thispage=index...>

4

[Black Powder Hunting, Shooting, and](#)
Count on Able Ammo to provide you shooting supplies, gun accessories, i
www.ableammo.com/catalog/default.js

5

[Black Dawge Cartridge - Black powder](#)
Black Dawge Cartridges and black powder June 14, 2011. Changes in the .45 caliber .45 cal, 450 ...
www.blackdawgecartridge.com

6

[Black Powder Ammunition - Comparison](#)
Black Powder Ammunition - 154 results Explosion, 20 Sticks, Avery Outdoor Series Black ...
www.nexttag.com/black-powder-amm

7

[Buffalo Arms - Ammunition - Black Powder](#)
WARNING: Discharging firearms in public handling ammunition may result in explosion
www.buffaloarms.com/browse.cfm?2,3

8

[Black powder ammunition | Shop by price](#)
Black powder ammunition - Find the ammunition on sale. Shop by price, coupons, and deals at ...
www.thefind.com/sports/info-black-powder

9

[Black Powder Ammunition - Sportsmans](#)
Black powder ammunition, bullets, and outdoor superstore your one stop hunting ammunition
www.sportsmansoutdoorsuperstore.com/ammunition

10

[Murphy Powder Ammunition - W.A. Murphy](#)
black, powder, loaded, ammunition . per box: 32-20 wcf 115 mfp; 26.50; 3 145 mfp
www.murphy powder.com/ammunition

...

Model Browsing Behavior

black powder ammunition

1
2
3
4
5
6
7
8
9
10
...

[Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able
ammunition, hunting supplies, shooti
unbeatable prices.
www.ableammo.com/catalog/default.p

[Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smokele
largest on hand inventory of obsolete
old, obsolete and ...
www.buffaloarms.com/browse.cfm?2%

[www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bj>
ACTION=enter&thispage=index...

[Black Powder Hunting, Shooting, and](#)
Count on Able Ammo to provide you
shooting supplies, gun accessories, i
www.ableammo.com/catalog/default.p

[Black Dawg Cartridge - Black powder](#)
Black Dawg Cartridges and black po
June 14, 2011. Changes in the .45 ca
.45 cal, 450 ...
www.blackdawgcartridge.com

[Black Powder Ammunition - Compar](#)
Black Powder Ammunition - 154 res
Explosion, 20 Sticks, Avery Outdoor
Series Black ...
www.nextag.com/black-powder-amm

[Buffalo Arms - Ammunition - Black P](#)
WARNING: Discharging firearms in pc
handling ammunition may result in ex
www.buffaloarms.com/browse.cfm?2,3

[Black powder ammunition | Shop bl](#)
Black powder ammunition - Find the
ammunition on sale. Shop by price, c
coupons, and deals at ...
www.thefind.com/sports/info-black-pc

[Black Powder Ammunition - Sports](#)
Black powder ammunition, bullets, b
outdoor superstore your one stop hur
www.sportsmansoutdoorsuperstore.co
ammunition

[Murphy Powder Ammunition - W.A. I](#)
black, powder, loaded, ammunition .
per box: 32-20 wcf 115 mfp: 26.50: 3
145 mfp
www.murphy powder.com/ammunition

2

Cascade-based models

Algorithm 1 The cascade user model

Require: R_1, \dots, R_{10} the *relevance* of the 10 URLs on the result page.

- 1: $i = 1$
 - 2: User examines position i .
 - 3: **if** $\text{random}(0,1) \leq R_i$ **then**
 - 4: User is satisfied with the document in position i and stops.
 - 5: **else**
 - 6: $i \leftarrow i + 1$; go to 2
 - 7: **end if**
-

NonRel : $R=0.2$

Rel : $R=0.8$

Rank Biased Precision

black powder
ammunition

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- ...

[Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able ammunition, hunting supplies, shooti unbeatable prices.
www.ableammo.com/catalog/default.aspx

[Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smokele largest on hand inventory of obsolete old, obsolete and ...
[www.buffaloarms.com/browse.cfm?2%](http://www.buffaloarms.com/browse.cfm?2%20)

[www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bjACTION=enter&thispage=index...>

[Black Powder Hunting, Shooting, and](#)
Count on Able Ammo to provide you shooting supplies, gun accessories, i
www.ableammo.com/catalog/default.aspx

[Black Dawg Cartridge - Black powder](#)
Black Dawg Cartridges and black powder June 14, 2011. Changes in the .45 caliber, 450 ...
www.blackdawgcartridge.com

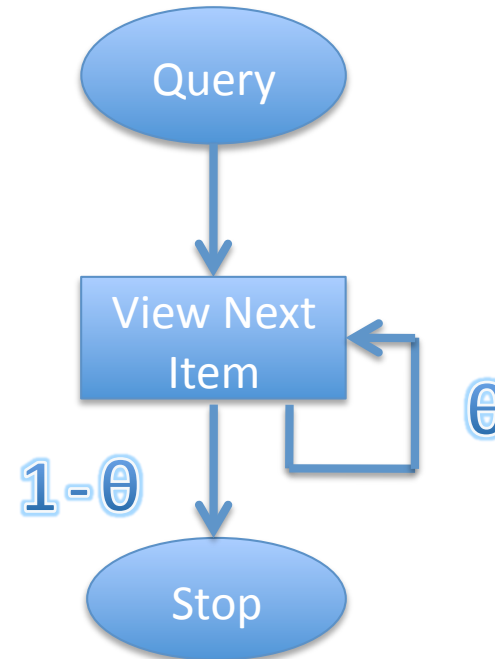
[Black Powder Ammunition - Comparison](#)
Black Powder Ammunition - 154 res Explosion, 20 Sticks, Avery Outdoor Series Black ...
www.nexttag.com/black-powder-amm

[Buffalo Arms - Ammunition - Black Powder](#)
WARNING: Discharging firearms in public handling ammunition may result in explosion
www.buffaloarms.com/browse.cfm?2,3

[Black powder ammunition | Shop by price](#)
Black powder ammunition - Find the ammunition on sale. Shop by price, coupons, and deals at ...
www.thefind.com/sports/info-black-powder

[Black Powder Ammunition - Sportsmans](#)
Black powder ammunition, bullets, outdoor superstore your one stop hunting
www.sportsmansoutdoorsuperstore.com/ammunition

[Murphy Powder Ammunition - W.A. Murphy](#)
black, powder, loaded, ammunition . per box: 32-20 wcf 115 mfp: 26.50: 3 145 mfp
www.murphy powder.com/ammunition



Expected Reciprocal Rank

[Chapelle et al CIKM09]

black powder
ammunition

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- ...

[Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able ammunition, hunting supplies, shoot unbeatable prices.
www.ableammo.com/catalog/default.aspx

[Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smoke largest on hand inventory of obsolete old, obsolete and ...
[www.buffaloarms.com/browse.cfm?2%](http://www.buffaloarms.com/browse.cfm?2%20)

[www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bjACTION=enter&thispage=index...>

[Black Powder Hunting, Shooting, and Count on Able Ammo to provide you shooting supplies, gun accessories, and](#)
www.ableammo.com/catalog/default.aspx

[Black Dawg Cartridge - Black powder Black Dawg Cartridges and black powder](#)
June 14, 2011. Changes in the .45 cal, 450 ...
www.blackdawgcartridge.com

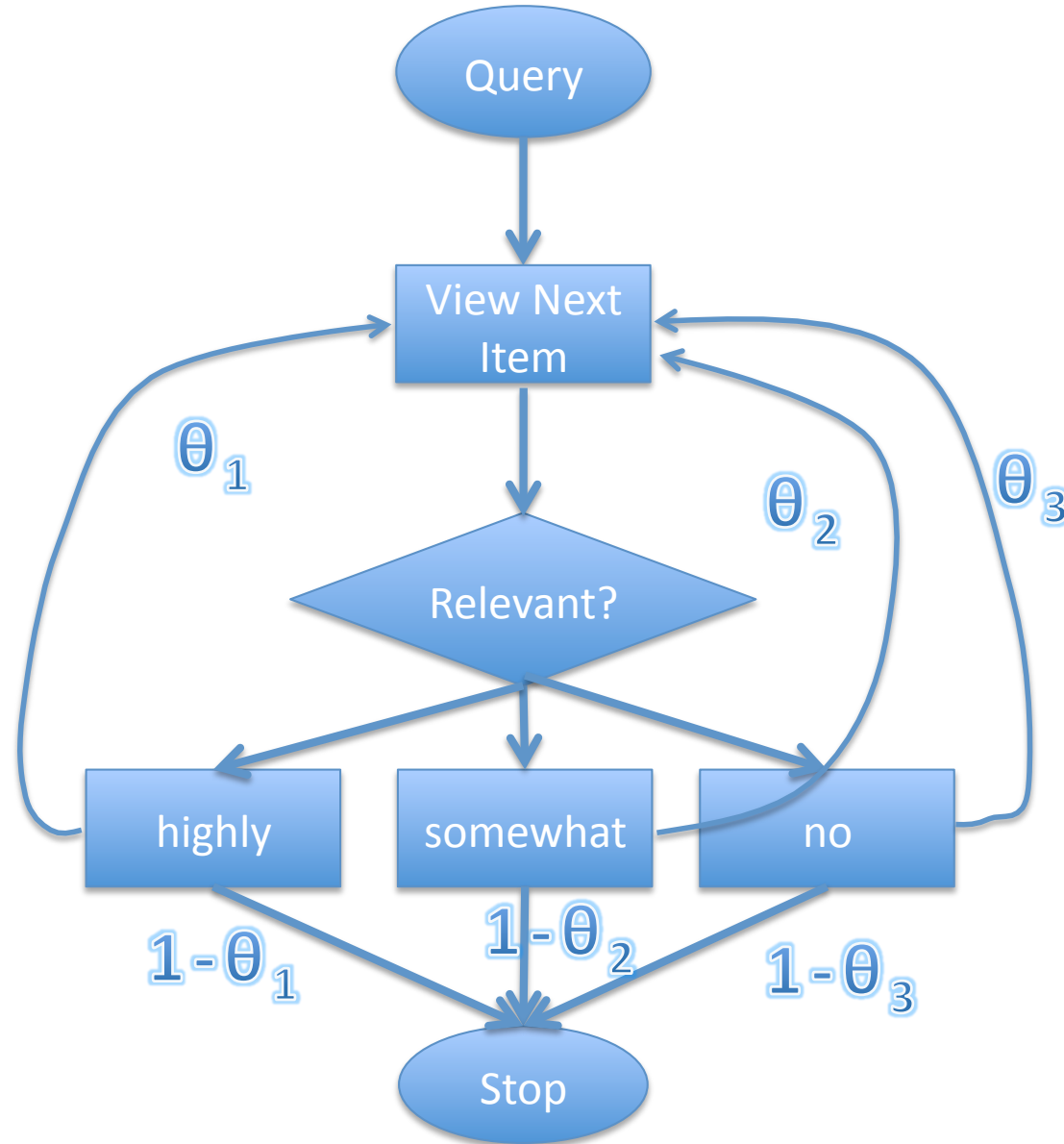
[Black Powder Ammunition - Comparison](#)
Black Powder Ammunition - 154 res Explosion, 20 Sticks, Avery Outdoor Series Black ...
www.nexttag.com/black-powder-amm

[Buffalo Arms - Ammunition - Black Powder](#)
WARNING: Discharging firearms in improper handling ammunition may result in explosion.
www.buffaloarms.com/browse.cfm?2,3

[Black powder ammunition | Shop by price](#)
Black powder ammunition - Find the ammunition on sale. Shop by price, coupons, and deals at ...
www.thefind.com/sports/info-black-powder

[Black Powder Ammunition - Sportsman's Outdoor Superstore](#)
Black powder ammunition, bullets, and outdoor superstore your one stop for ammunition.
www.sportsmansoutdoorsuperstore.com/ammunition

[Murphy Powder Ammunition - W.A. Murphy](#)
black, powder, loaded, ammunition . per box: 32-20 wcf 115 mfp: 26.50: 3 145 mfp
www.murphy powder.com/ammunition



In reality...

- The probability that a user visits a document is affected by
 - Relevance of the documents already visited
 - The snippet quality
 - Query type
 - ...

Query Log Reality

- Users tend to stop search if they are satisfied or frustrated

Relevance	P(Stop R)
Bad	0.49
Fair	0.41
Good	0.37
Excellent	0.53
Perfect	0.76

- P(click) highly affected by the snippet quality

Relevance	P(C R)
Bad	0.50
Fair	0.49
Good	0.45
Excellent	0.59
Perfect	0.79

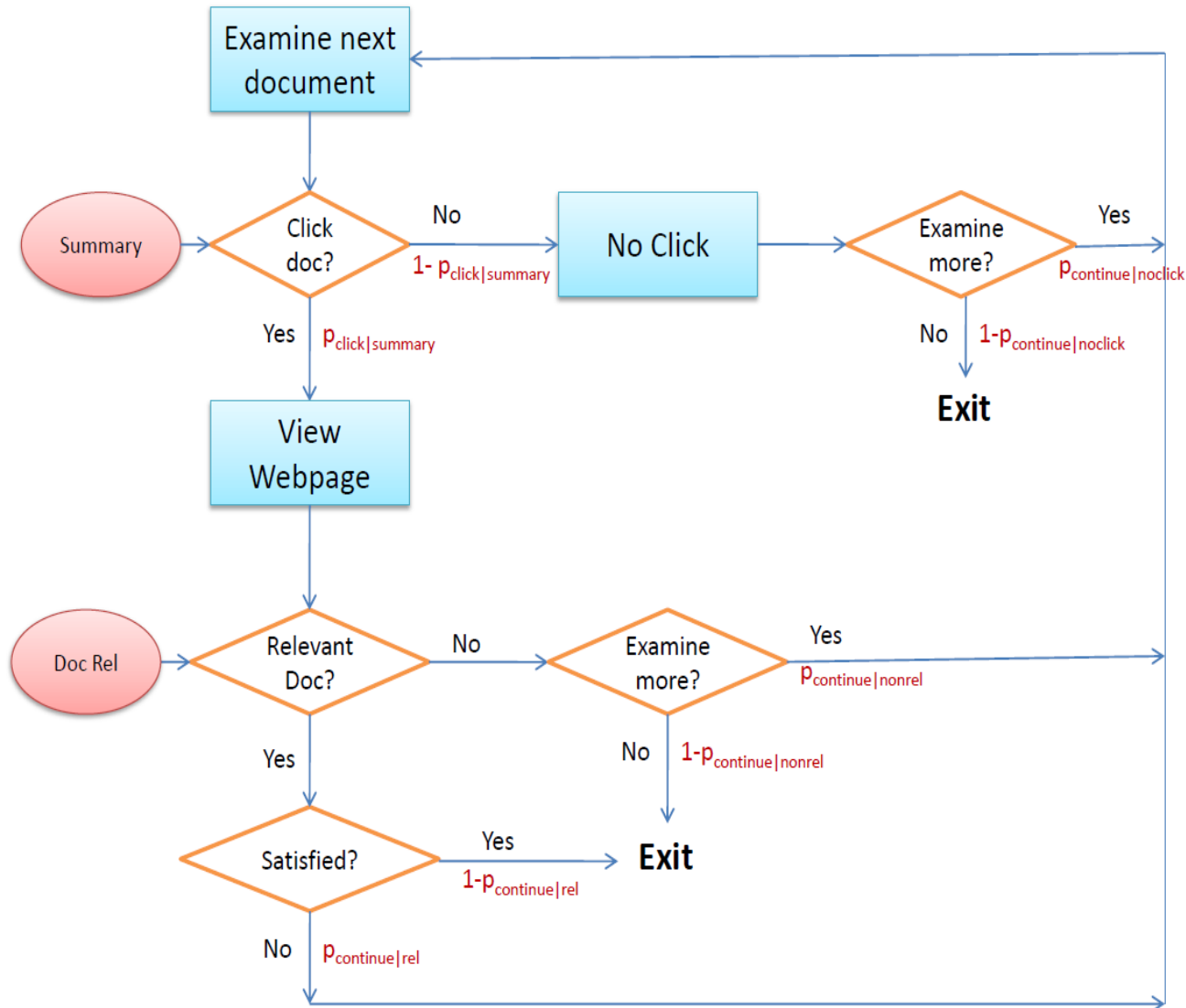
Query Log Reality

- Users behave differently for different queries
 - Informational queries
 - Navigational queries

	Navigational		Informational	
	P(C R)	P(Stop R)	P(C R)	P(Stop R)
Bad	0.632	0.587	0.516	0.431
Fair	0.569	0.523	0.455	0.357
Good	0.526	0.483	0.442	0.349
Excellent	0.700	0.669	0.533	0.458
Perfect	0.809	0.786	0.557	0.502

Expected Browsing Utility

[Yilmaz et al CIKM10]



$$D_{EBU}(r) = P(E_r) \cdot P(C | R_r)$$

$$EBU = \sum_{r=1}^n D_{EBU}(r) \cdot R_r$$

Comparison: NDCG vs EBU: Session Likelihood

- Discount function models $P(\text{click})$
- How well does the discount function predict actual clicks? : $\log(P(\text{Click Sequence} | \text{Metric}))$

$$P(\text{Session} | M) = \prod_{\forall r, doc_r \text{ clicked}} P(C_r | M) \cdot \prod_{\forall r, doc_r \text{ not clicked}} (1 - P(C_r | M))$$

	Session likelihood
NDCG, log discount	-2.204
NDCG, 1/r discount	-1.939
EBU	-1.582

Comparison: NDCG vs EBU: Session Likelihood

- Discount function models $P(\text{click})$
- How well does the discount function predict actual clicks? : $\log(P(\text{Click Sequence} | \text{Metric}))$

$$P(\text{Session} | M) = \prod_{\forall r, doc_r \text{ clicked}} P(C_r | M) \cdot \prod_{\forall r, doc_r \text{ not clicked}} (1 - P(C_r | M))$$


	Session likelihood	
	Navigational	Informational
NDCG, log discount	-2.204	-2.382
NDCG, 1/r discount	-1.939	-2.065
EBU	-1.582	-2.270

Discounted Gain Model

1. What is the chance a user will read a document?
2. What does the user gain by reading a document?


Discounted Cumulative Gain

	Relevance	Relevance Score
1	R	1
2	R	1
3	N	0
4	N	0
5	R	1
6	R	1
7	N	0
8	R	1
9	N	0
10	N	0
...	...	

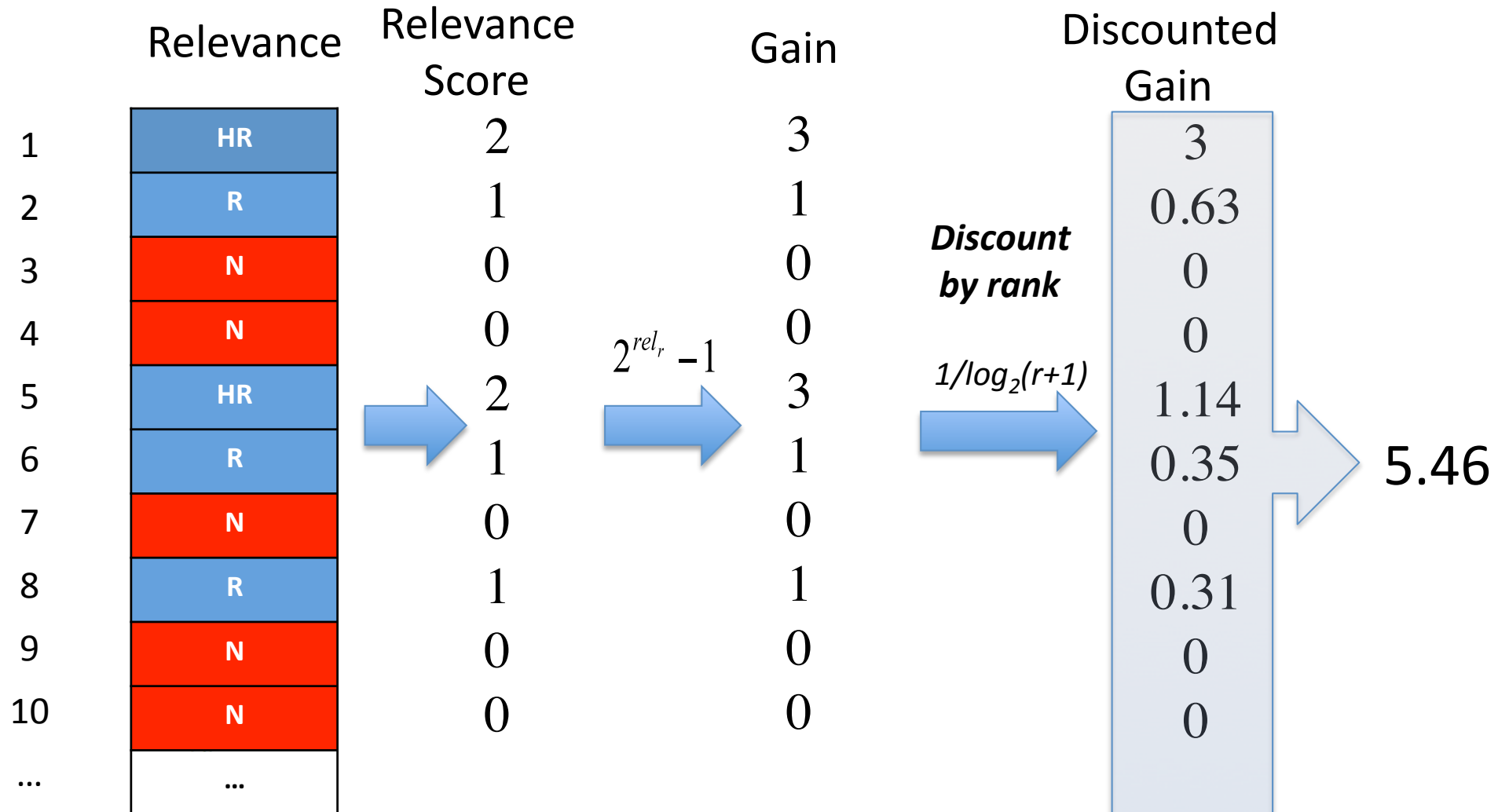


Discounted Cumulative Gain

	Relevance	Relevance Score
1	HR	2
2	R	1
3	N	0
4	N	0
5	HR	2
6	R	1
7	N	0
8	R	1
9	N	0
10	N	0
...	...	



Discounted Cumulative Gain



Graded Average Precision

[Robertson et al. SIGIR10]

One document is *more useful* than another

One possible meaning:

one document is *useful to more users* than another

Hence the following:

assume grades of relevance...

... but that user has a threshold relevance grade

which defines a binary view

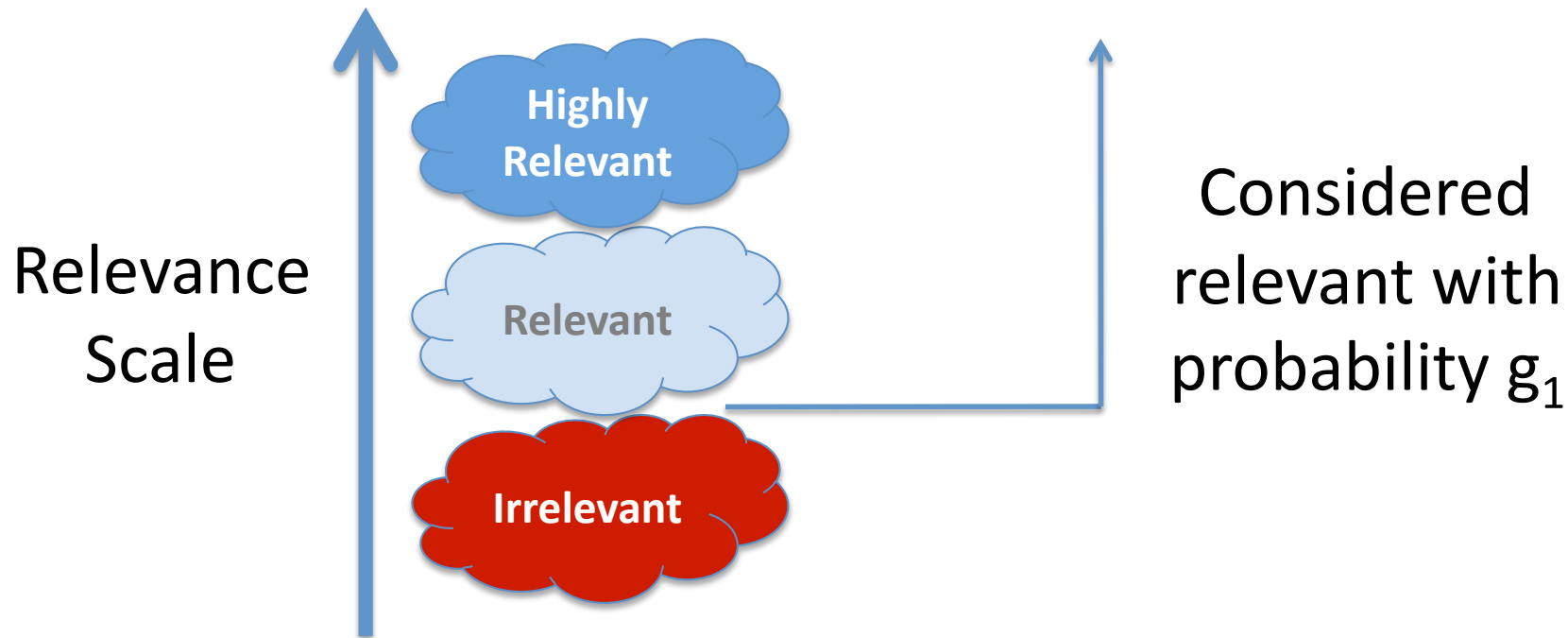
different users have different thresholds

described by a *probability distribution* over users

Graded Average Precision

[Robertson et al. SIGIR10]

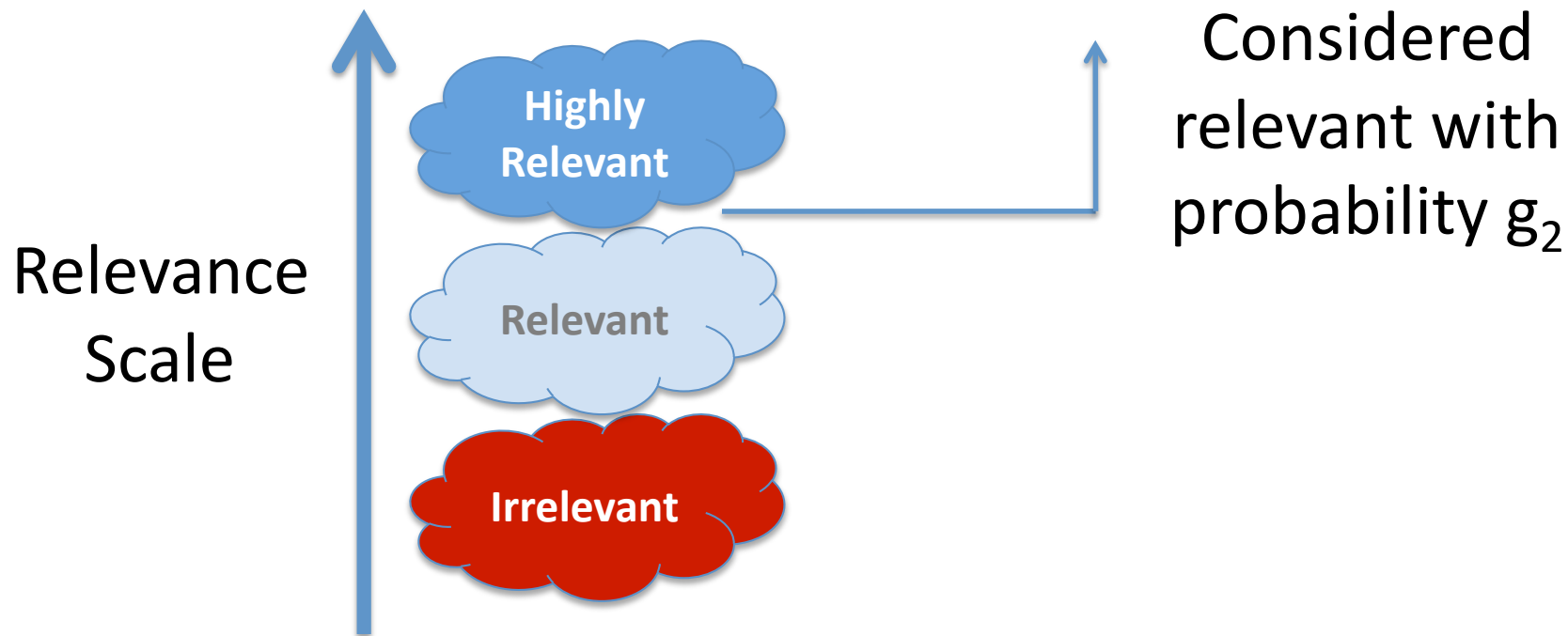
- User has binary view of relevance
 - by thresholding the relevance scale



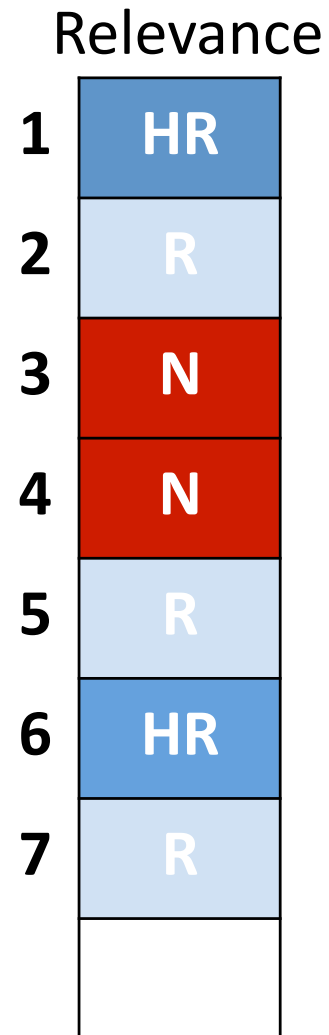
Graded Average Precision

[Robertson et al. SIGIR10]

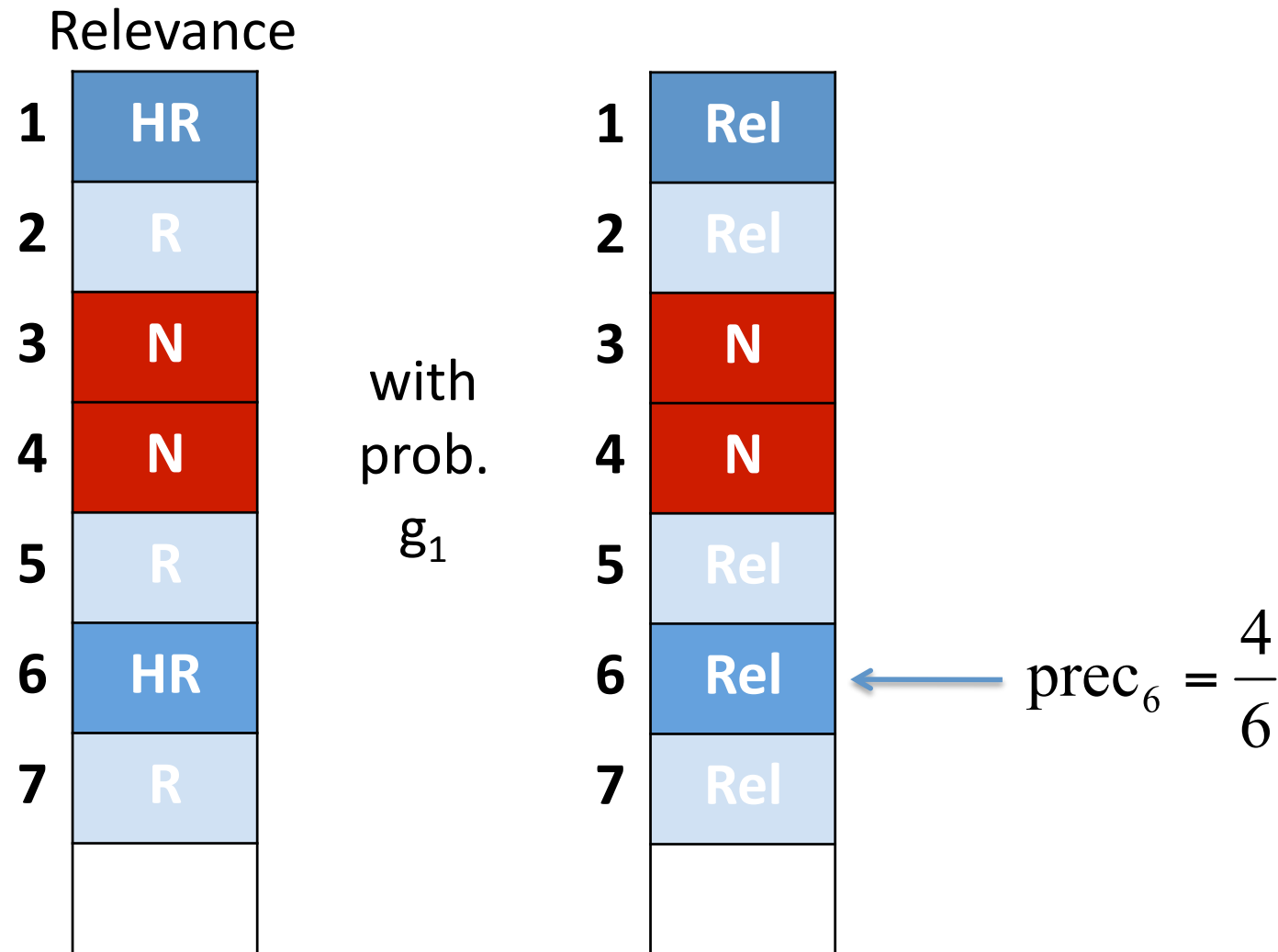
- User has binary view of relevance
 - by thresholding the relevance scale



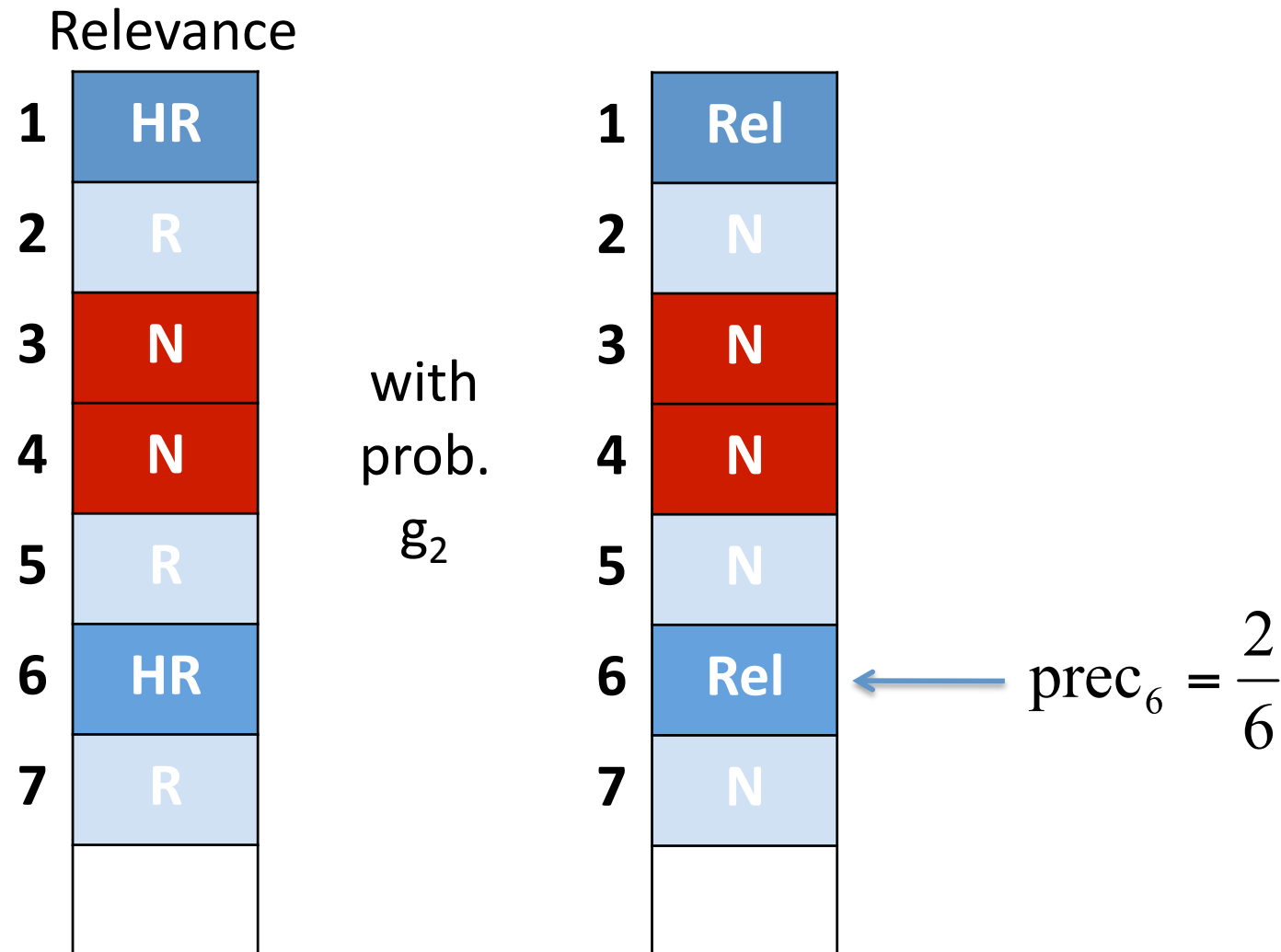
Graded Average Precision (GAP)



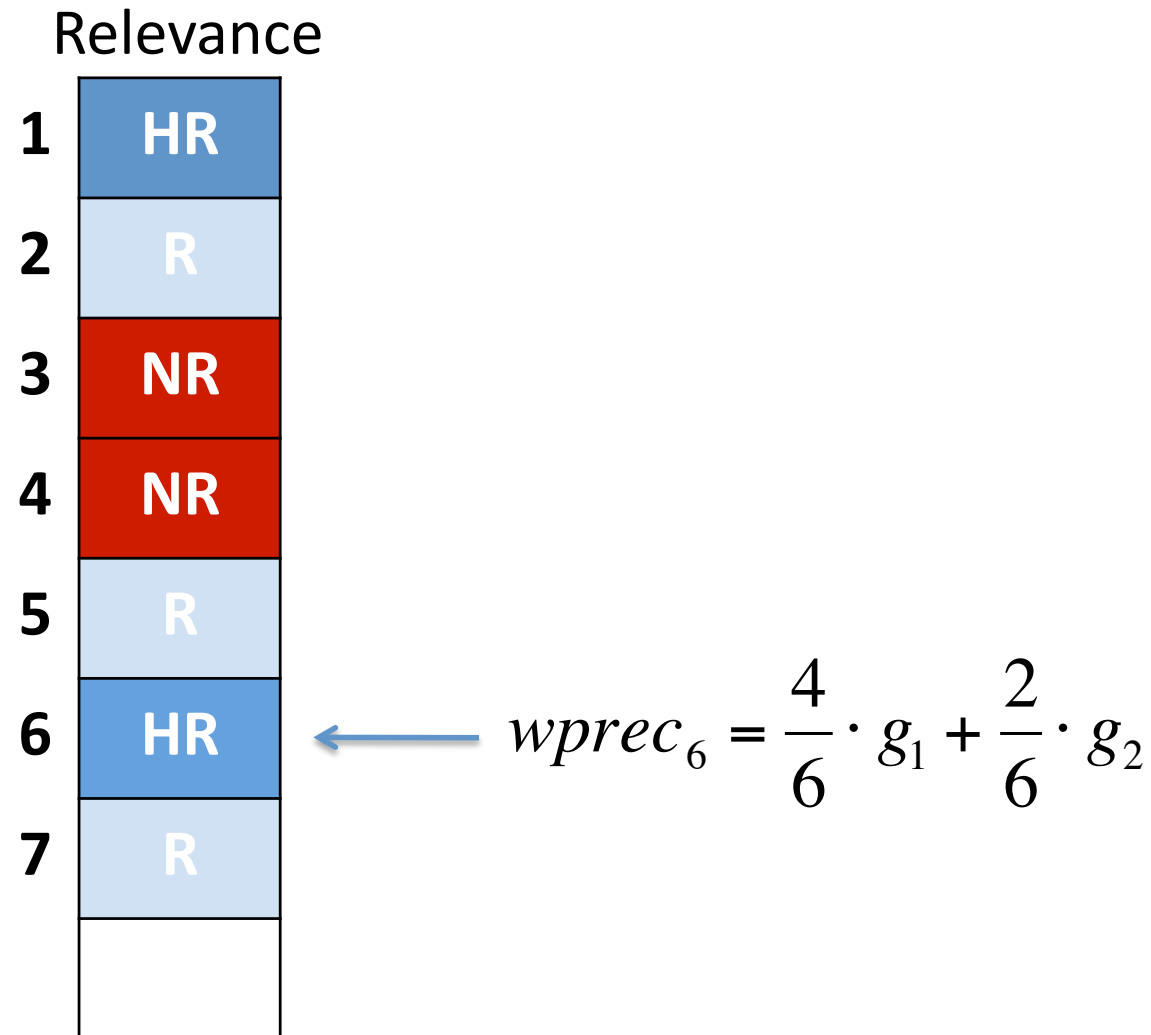
Graded Average Precision (GAP)



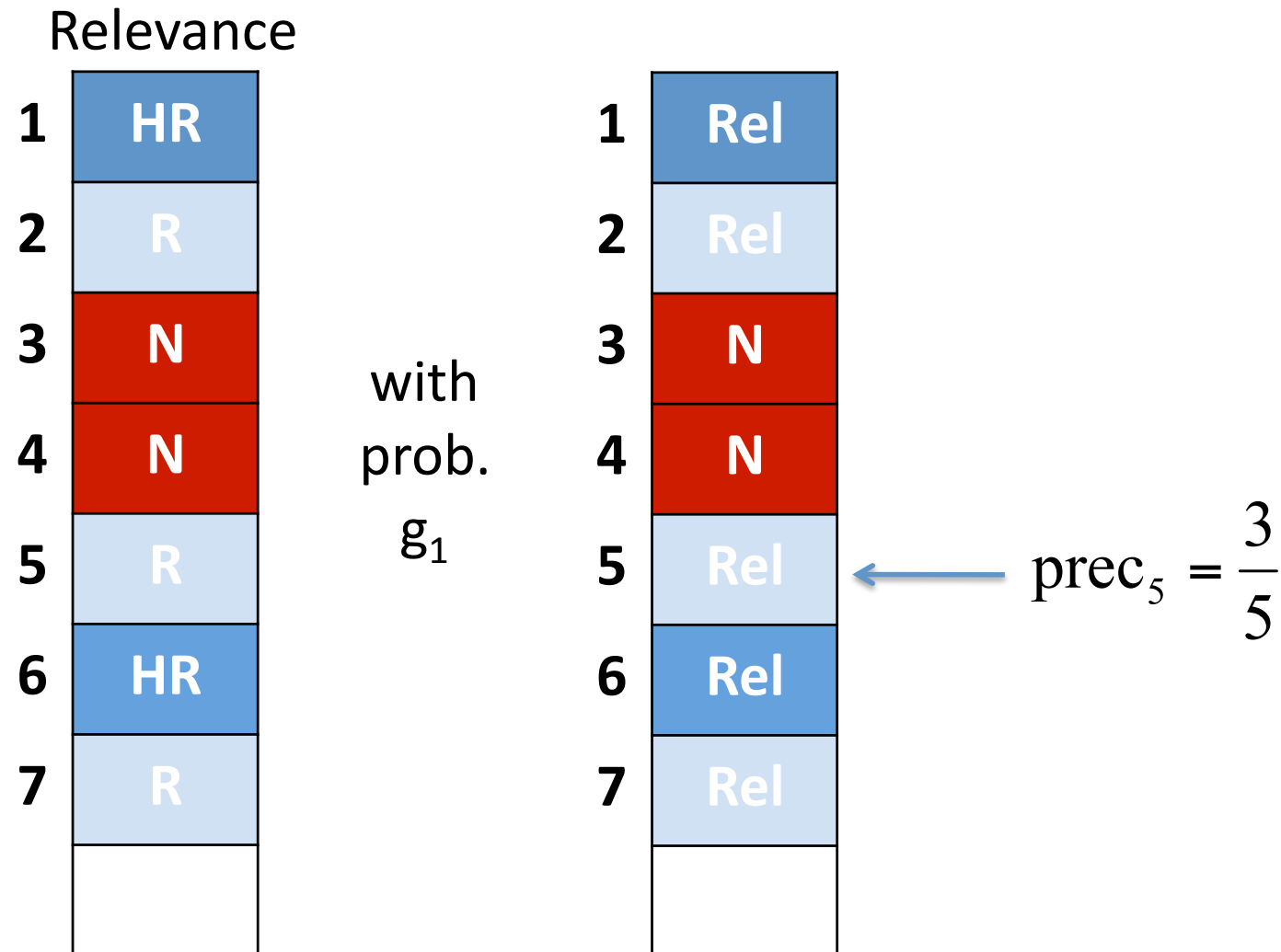
Graded Average Precision (GAP)



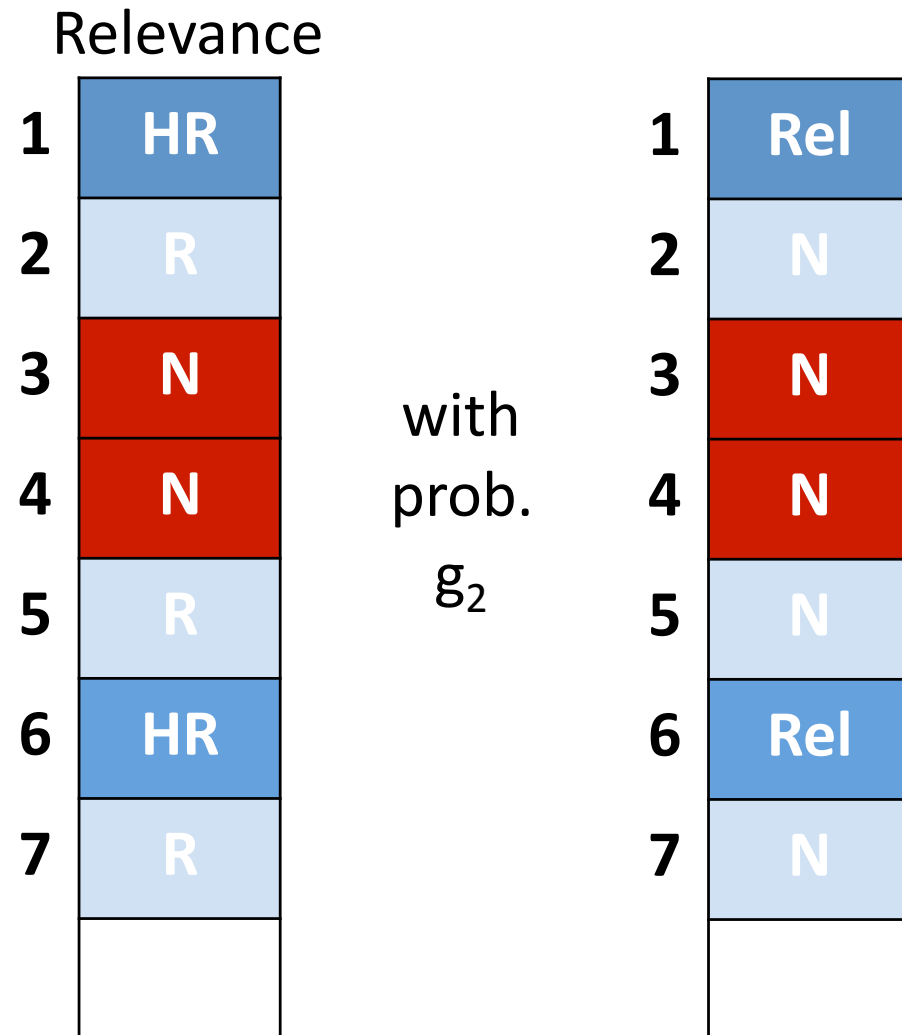
Graded Average Precision (GAP)



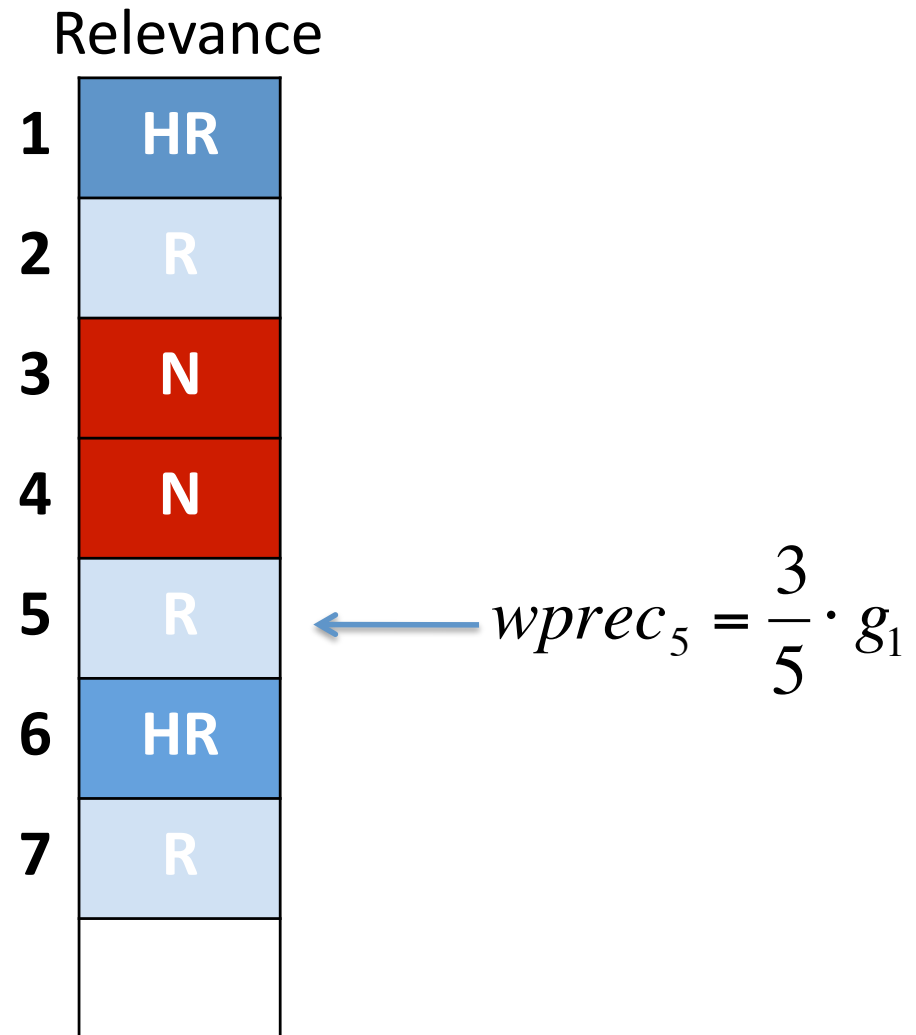
Graded Average Precision (GAP)



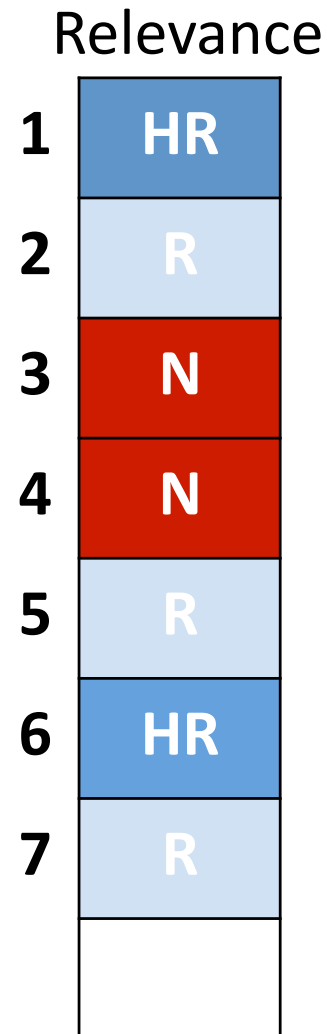
Graded Average Precision (GAP)



Graded Average Precision (GAP)



Graded Average Precision (GAP)



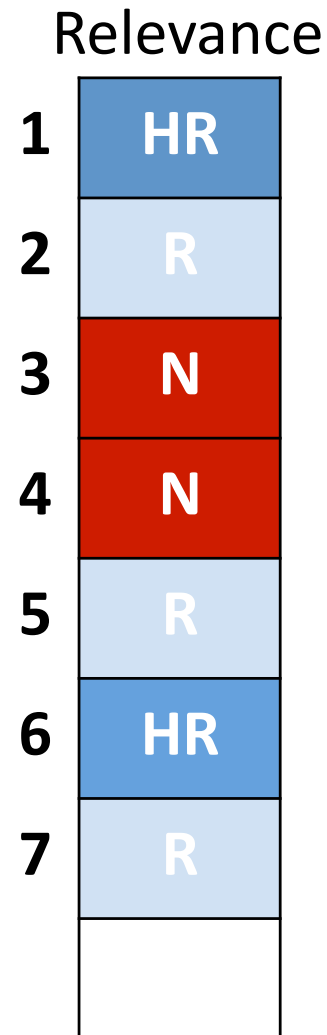
- Weighted Precision at rank n :

$$wprec_n = \sum_{j=1}^{i_n} \left(\frac{1}{n} |d_m : m \leq n, i_m \geq i_n| \right) \cdot g_j$$

i_n : relevance grade of doc at rank n

Binary Precision when grade j is considered the threshold

Graded Average Precision (GAP)



- Cumulative Weighted Precision of list of size N :

$$cwprec = \sum_{n=1}^N wprec_n$$

Graded Average Precision (GAP)

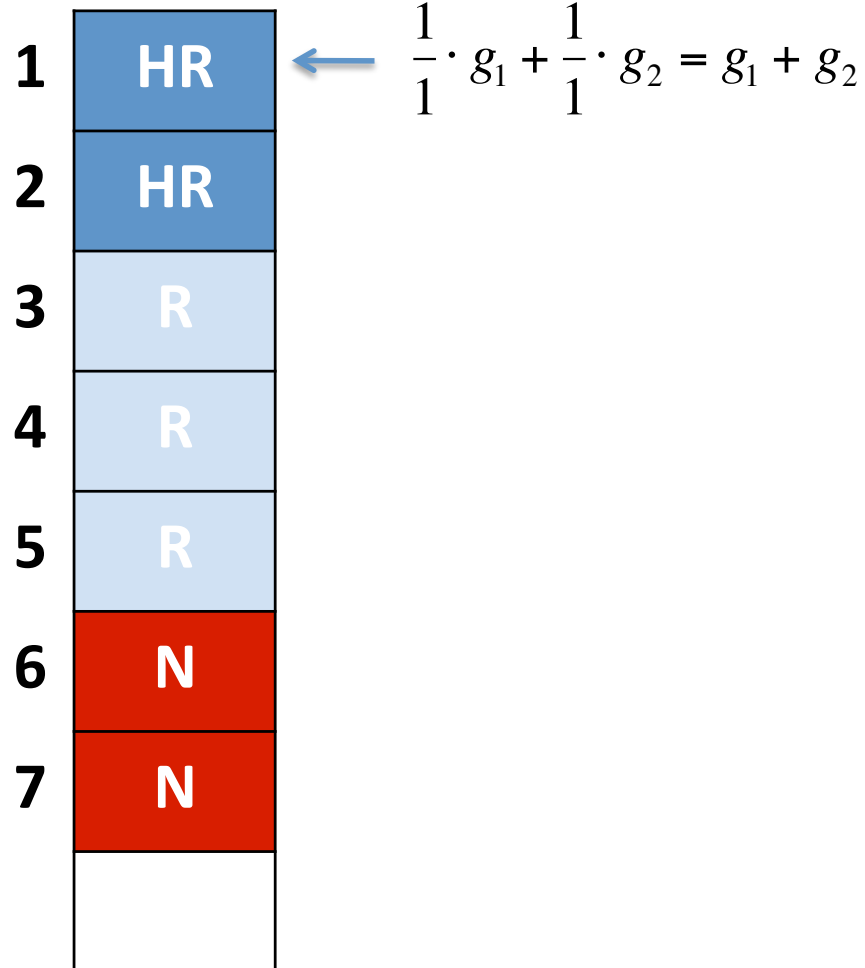
Optimal Ranking

1	HR
2	HR
3	R
4	R
5	R
6	N
7	N

- Cumulative Weighted Precision of the optimal ranking

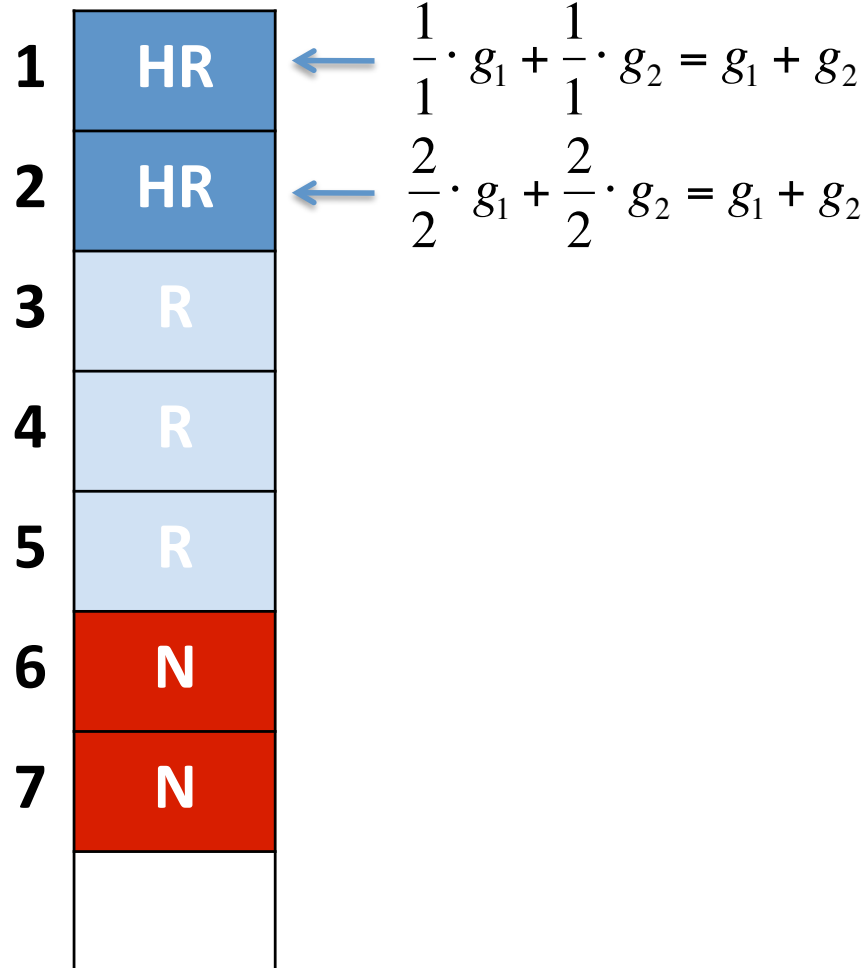
Graded Average Precision (GAP)

Optimal Ranking



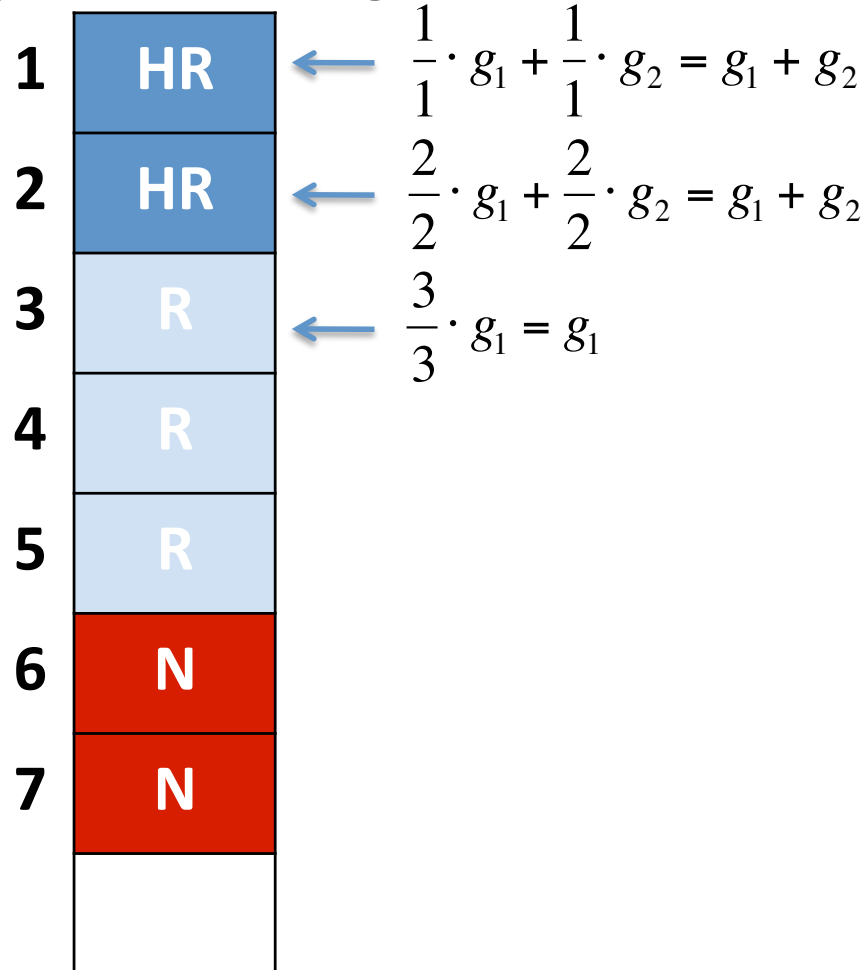
Graded Average Precision (GAP)

Optimal Ranking



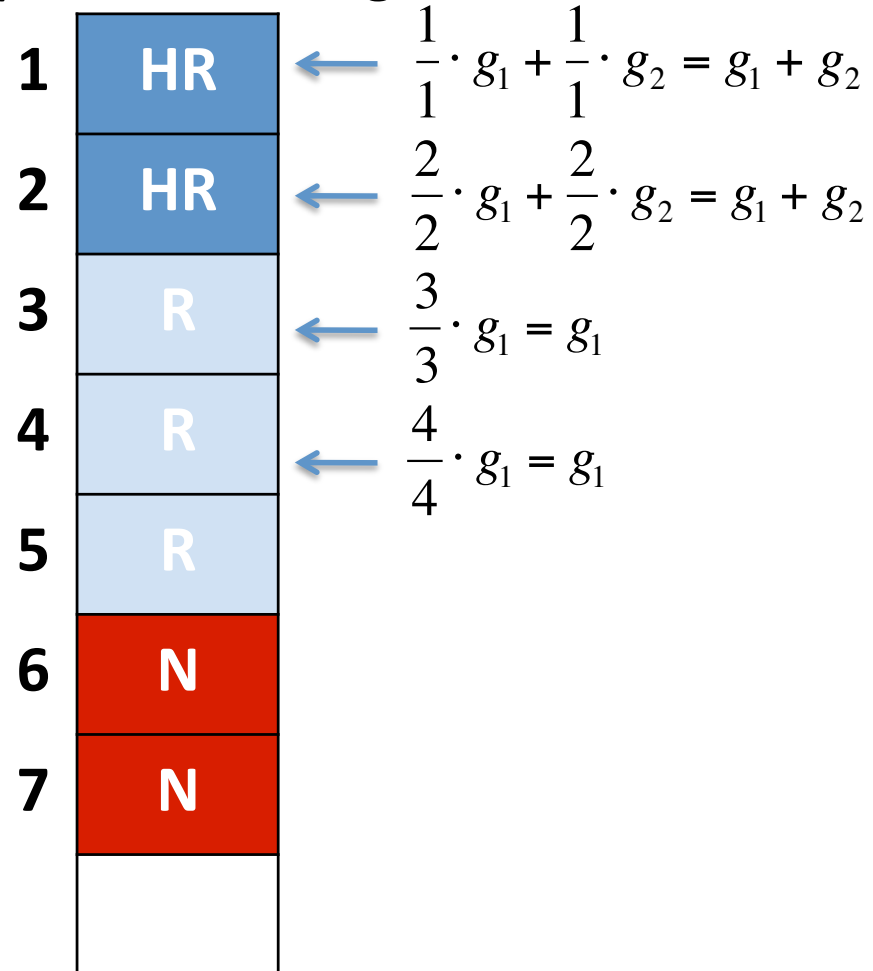
Graded Average Precision (GAP)

Optimal Ranking



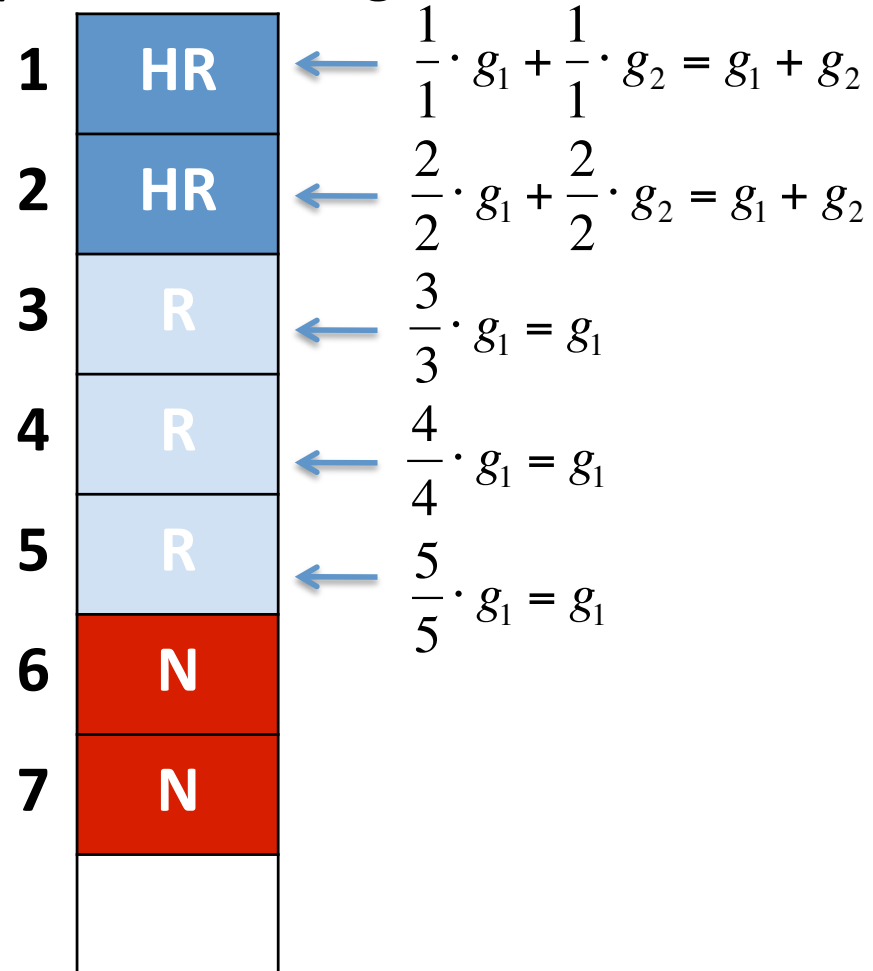
Graded Average Precision (GAP)

Optimal Ranking



Graded Average Precision (GAP)

Optimal Ranking



Graded Average Precision (GAP)

Optimal Ranking

1	HR
2	HR
3	R
4	R
5	R
6	N
7	N

- Cumulative Weighted Precision of the optimal ranking:

$$\sum_{j=1}^c R_j \sum_{i=1}^j g_i$$

where R_j is the total number of documents in grade j

Graded Average Precision (GAP)

Relevance

1	HR
2	R
3	N
4	N
5	R
6	HR
7	R

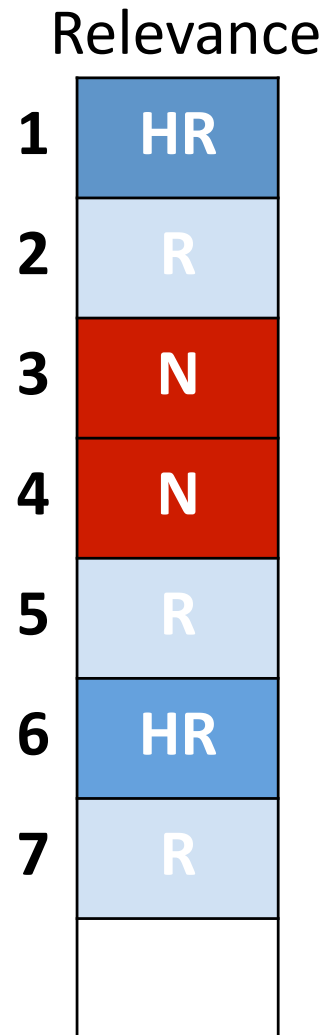
- Graded Average Precision

$$GAP = \sum_{n=1}^N \sum_{j=1}^{i_n} \left(\frac{1}{n} |d_m : m \leq n, i_m \geq i_n| \right) \cdot g_j$$

$$\sum_{j=1}^c R_j \sum_{i=1}^j g_i$$

Cumulative Weighted Precision

Graded Average Precision (GAP)



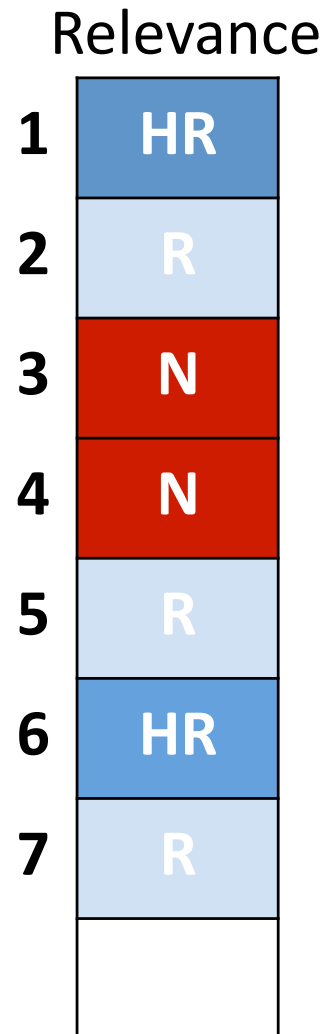
- Graded Average Precision

$$GAP = \frac{\sum_{n=1}^N \sum_{j=1}^{i_n} \left(\frac{1}{n} |d_m : m \leq n, i_m \geq i_n| \right) \cdot g_j}{\sum_{j=1}^c R_j \sum_{i=1}^j g_i}$$

$$\sum_{j=1}^c R_j \sum_{i=1}^j g_i$$

Cumulative Weighted Precision of Optimal List

Graded Average Precision (GAP)



- Graded Average Precision

$$GAP = \frac{\sum_{n=1}^N \sum_{j=1}^{i_n} \left(\frac{1}{n} |d_m : m \leq n, i_m \geq i_n| \right) \cdot g_j}{\sum_{j=1}^c R_j \sum_{i=1}^j g_i}$$

- In the binary case, GAP reduces to average precision

Probabilistic Interpretation of GAP

- Average Precision
 - $\Pr(d_m \text{ is relevant} \mid d_n \text{ is relevant})$
rank $n \leq$ rank m
- GAP
 - $\Pr(d_m \text{ is considered rel.} \mid d_n \text{ is considered rel.})$
rank $n \leq$ rank m

Today's Outline

- Evaluation based on user models
 - Gain/Discount Framework
 - Modeling browsing behavior
 - RBP, AP, DCG, ERR, EBU
 - Modeling utility
 - NDCG, GAP
- Novelty and diversity
- Query sessions

Novelty

- The redundancy problem:
 - the first relevant document contains some information
 - every document after that with the same information is worth less to the user
- Novelty retrieval attempts to ensure that ranked results do not have much redundancy

Example

- oil-producing nations
 - members of OPEC
 - North Atlantic nations
 - South American nations
- 10 relevant articles about OPEC probably not as useful as one relevant article about each group
 - And one relevant article about all oil-producing nations might be even better

How to Evaluate?

- One approach:
 - List subtopics or aspects of the topic
 - Judge each document relevant or not to each possible subtopic
- For oil-producing nations, subtopics could be names of nations
 - Saudi Arabia, Russia, Canada, ...

Subtopic Relevance Example

D1: West Africa's Gulf of Guinea is surrounded by oil-rich countries: Nigeria, Equatorial Guinea, Cameroon, Gabon and Angola. ...

D2: ... Iran ... said OPEC would make little change ... and [UAE said] current oil prices are reasonable ... oil giant Russia which is second only to Saudi Arabia in oil output ...

D3: ... Iraq's oil infrastructure ... Russia ... increased its output by 7.7 percent ...

D4: ... Libya, a member of the 11-nation Organisation of Petroleum Exporting Countries (OPEC) ... oil production in Iraq ... The GCC groups the gas- and oil-rich states of Bahrain, Kuwait, Oman, Qatar, Saudi Arabia and the United Arab Emirates (UAE)

Evaluation Measures

- Subtopic recall and precision (Zhai et al., 2003)
 - Subtopic recall at rank k:
 - Count unique subtopics in top k documents
 - Divide by total number of known unique subtopics
 - Subtopic precision at recall r:
 - Find least k at which subtopic recall r is achieved
 - Find least k at which subtopic recall r *could possibly* be achieved (by a perfect system)
 - Divide latter by former
 - Models a user that wants all subtopics
 - and doesn't care about redundancy as long as they are seeing new information

Subtopic Relevance Evaluation

D1: West Africa's Gulf of Guinea is surrounded by oil-rich countries: Nigeria, Equatorial Guinea, Cameroon, Gabon and Angola. ...

D2: Chad, a big, dirt-poor landlocked country in the heart of Africa, officially entered the ranks of oil-producing countries Friday ...

D3: Chad, a big, dirt-poor landlocked country in the heart of Africa, officially entered the ranks of oil-producing countries Friday ...

D4: ... will mark the entry of Chad into the elite club of world oil-producing nations ... Cameroon's Paul Biya ... Equatorial Guinea's Teodoro Obiang Nguema, Omar Bongo of Gabon ... [all of which] are oil producers.

D1: West Africa's Gulf of Guinea is surrounded by oil-rich countries: Nigeria, Equatorial Guinea, Cameroon, Gabon and Angola. ...

D2: ... Iran ... said OPEC would make little change ... and [UAE said] current oil prices are reasonable ... oil giant Russia which is second only to Saudi Arabia in oil output ...

D3: ... Iraq's oil infrastructure ... Russia ... increased its output by 7.7 percent ...

D4: ... Libya, a member of the 11-nation Organisation of Petroleum Exporting Countries (OPEC) ... oil production in Iraq ... The GCC groups the gas- and oil-rich states of Bahrain, Kuwait, Oman, Qatar, Saudi Arabia and the United Arab Emirates (UAE)

Diversity

- Short keyword queries are inherently ambiguous
 - An automatic system can never know the user's intent
 - (Another person can't even really know the original user's intent!)
- Diversification attempts to retrieve results that may be relevant to a space of possible intents

Evaluation Measures

- Subtopic recall and precision can be used
 - This time with judgments to intents rather than subtopics
- Measures that know about intents:
 - “Intent-aware” family of measures (Agrawal et al.)
 - α -nDCG (Clarke et al.)
 - ERR-IA (Chapelle et al.)

Intent-Aware Measures

- Assume there is a probability distribution $P(i | Q)$ over intents for a query Q
 - Probability that a randomly-sampled user means intent i when submitting query Q
- The intent-aware version of a measure is its weighted average over this distribution

[The Official Site of the Arizona Cardinals](#) - [翻译此页]
www.azcardinals.com/ - 网页快照
 Official team site with audio and video clips, team news, depth charts, transactions, statistics, and player profiles.

[The Official Site of The St. Louis Cardinals | cardinals.com: Homepage](#) - [翻译此页]
mlb.mlb.com/index.jsp?c_id=stl - 网页快照
 2 days ago - The Cardinals acquired starter Edwin Jackson and parted ways with Colby Rasmus on Wednesday in a deal with the Blue Jays that could ...

[Cardinal \(bird\) - Wikipedia, the free encyclopedia](#) - [翻译此页]
[en.wikipedia.org/wiki/Cardinal_\(bird\)](http://en.wikipedia.org/wiki/Cardinal_(bird)) - 网页快照
 The Cardinals or Cardinalidae are a family of passerine birds found in North and South America. The South American cardinals in the genus Paroaria are ...

[Cardinal - Wikipedia, the free encyclopedia](#) - [翻译此页]
en.wikipedia.org/wiki/Cardinal - 网页快照
 Jump to: navigation, search. Look up Cardinal, Cardinals, cardinal, or ...
 Cardinal (bird) - Cardinal (Catholicism) - Cardinal number - Cardinal direction
 显示更多来自 wikipedia.org 的搜索结果

[Northern Cardinal, Identification, All About Birds - Cornell Lab ...](#) - [翻译此页]
www.allaboutbirds.org/guide/northern_cardinal/id - 网页快照
 Learn how to identify Northern Cardinal, its life history, cool facts, sounds and calls, and watch videos. The male Northern Cardinal is perhaps responsible ...

[CATHOLIC ENCYCLOPEDIA: Cardinal](#) - [翻译此页]
www.newadvent.org > Catholic Encyclopedia > C - 网页快照
 A dignitary of the Roman Church and counsellor of the pope.

[Arizona Cardinals | NFL at CBSSports.com](#) - [翻译此页]
www.cbssports.com/nfl/teams/page/ARI/arizona-cardinals - 网页快照
 Complete Arizona Cardinals NFL Football Coverage at CBSSports.com.

cardinals 的图片搜索结果 - 举图图片



[St Louis Cardinals news, opinion, photos, and more | Bleacher Report](#) - [翻译此页]
bleacherreport.com/st-louis-cardinals - 网页快照
 Get the latest St Louis Cardinals news, photos, rankings, lists and more on Bleacher Report.

[Arizona Cardinals Football Clubhouse - ESPN](#) - [翻译此页]
espn.go.com/nfl/team/_name/ari/arizona-cardinals - 网页快照
 Arizona Cardinals scores, news, schedule, players, stats, photos, rumors, and highlights on ESPN.com.

[St. Louis Cardinals Schedule, Roster, Stats, Blog Posts - MLB ...](#) - [翻译此页]
www.sbnation.com/mlb/teams/st-louis-cardinals - 网页快照
 The St. Louis Cardinals, the Toronto Blue Jays, and the Chicago White Sox fulfilled a number of trade rumors at once by getting Edwin Jackson to the Cardinals ...

$$P(\text{NFL football} \mid Q) = 0.35$$

$$P(\text{MLB baseball} \mid Q) = 0.35$$

$$P(\text{Cardinalidae} \mid Q) = 0.2$$

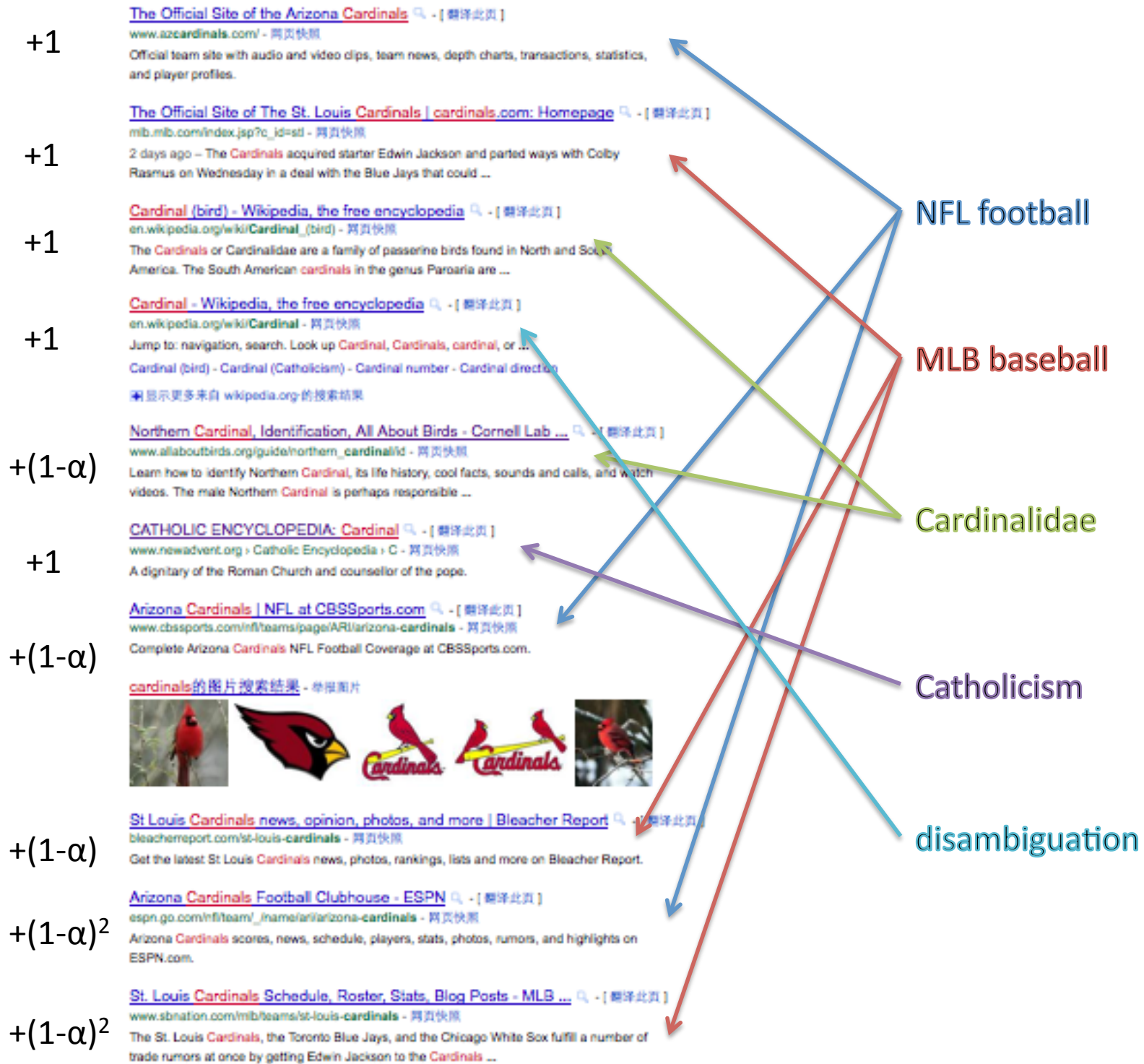
$$P(\text{Catholicism} \mid Q) = 0.08$$

$$P(\text{disambiguation} \mid Q) = 0.02$$

$$P@10\text{-IA} = 0.35*0.3 + 0.35*0.3 + 0.2*0.2 + 0.08*0.1 + 0.02*0.1 = 0.23$$

α -nDCG

- α -nDCG is a generalization of nDCG that accounts for both novelty and diversity
- α is a geometric penalization for redundancy
 - Redefine the gain of a document:
 - +1 for each subtopic it is relevant to
 - $\times(1-\alpha)$ for each document higher in the ranking that subtopic already appeared in
- Discount is the same as usual



ERR-IA

- Intent-aware version of ERR
- But it has appealing properties other IA measures do not have:
 - ranges between 0 and 1
 - submodularity: diminishing returns for relevance to a given subtopic -> built-in redundancy penalization
- Also has appealing properties over α -nDCG:
 - Easily handles graded subtopic judgments
 - Easily handles intent distributions

Today's Outline

- Evaluation based on user models
 - Gain/Discount Framework
 - Modeling browsing behavior
 - RBP, AP, DCG, ERR, EBU
 - Modeling utility
 - NDCG, GAP
- Novelty and diversity
- Query sessions

Why sessions?

- Current evaluation framework
 - Assesses the effectiveness of systems over one-shot queries
- Users reformulate their initial query
- Still fine if ...
 - optimizing system for one-shot queries led to optimal performance over an entire session

Why sessions?


When was the DuPont Science Essay Contest created?

Initial Query : **DuPont Science Essay Contest**

Reformulation : **When was the DSEC created?**

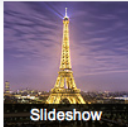
- e.g. retrieval systems should accumulate information along a session

Paris

Paris 

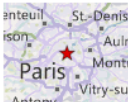
[Web](#) [Maps](#) [Weather](#) [Flights](#) [Images](#) [Videos](#) [News](#)

ALL RESULTS 1-11 of 15,800,000 results · [Advanced](#)




Paris, France
City Population: 2,138,551
Country Population: 65,447,374
Country World Rank: 20

[Maps](#)




[Weather](#)



45° F
52° / 37°
Clear

[Flight deals](#)



\$680
IAD > CDG
[See all deals](#)

[Attractions](#)

- Tour Eiffel
- Arc de Triomphe Paris
- Centre Pompidou
- Sainte-Chapelle

SHARE [Facebook](#) [Twitter](#) [Messenger](#) [Email](#)

[Paris - Wikipedia, the free encyclopedia](#)

[History](#) · [Geography](#) · [Cityscape](#) · [Culture](#)

Paris (in French) is the capital and largest city in France. It is situated on the river Seine, in northern France, at the heart of the Île-de-France region (or **Paris Region** ...

en.wikipedia.org/wiki/Paris · [Cached page](#)

[The Paris Pages / Les Pages de Paris / Paris.Org / The Web's First ...](#)

The **Paris Pages**™ / Les Pages de **Paris**™ / **Paris.Org**™ / Online Since 14 July 1994, It's The Internet's First **Paris** Web Site.

www.paris.org · [Cached page](#)

[Paris travel guide - Wikitravel](#)

[Districts](#) · [Understand](#) · [Get in](#) · [Get around](#)

Paris, the cosmopolitan capital of France, is - with 2.2 million people living in the dense central city and almost 12 million people living in the whole metropolitan area ...

wikitravel.org/en/Paris · [Cached page](#)

[Paris convention and visitors office - Official website - Paris ...](#)

Paris convention and visitors office diffuses all information to organise your stay or your trip in **Paris**: hotels and loadings, museums, monuments, going out, events, restaurants.

en.parisinfo.com · [Cached page](#)

[Paris Las Vegas Hotel](#)

Experience a first rate Las Vegas hotel at **Paris Las Vegas**. Find Las Vegas entertainment & enjoy one of the best Las Vegas casinos at **Paris Las Vegas** now!

Extend the evaluation framework

From one query evaluation

To multi-query sessions evaluation

Construct appropriate test collections

Rethink of evaluation measures

Basic test collection

- A set of information needs

What do we know about black powder ammunition?

– A static sequence of m queries

Initial Query :	black powder ammunition
1 st Reformulation :	black powder wiki
2 nd Reformulation :	gun powder wiki
...	...
($m-1$) th Reformulation :	history of gunpowder

Experiment

black powder ammunition

1

[Black Powder Bullets for Sale Online](#)
Black Powder Bullets Count on Able **ammunition**, hunting supplies, shooti unbeatable prices.
www.ableammo.com/catalog/default.js

2

[Buffalo Arms - Ammunition](#)
Black Powder Cartridge and Smokele largest on hand inventory of obsolete old, obsolete and
www.buffaloarms.com/browse.cfm?2%

3

[www.powderinc.com](#)
<https://www.powderinc.com/cgi-bin/bj ACTION=enter&thispage=index...>

4

[Black Powder Hunting, Shooting, and](#)
Count on Able **Ammo** to provide you shooting supplies, gun accessories, i
www.ableammo.com/catalog/default.js

5

[Black Dawg Cartridge - Black powder](#)
Black Dawg Cartridges and **black powder** June 14, 2011. Changes in the .45 cal .45 cal, 450 ...
www.blackdawgcartridge.com

6

[Black Powder Ammunition - Compar](#)
Black Powder Ammunition - 154 res Explosion, 20 Sticks, Avery Outdoor Series **Black** ...
www.nexttag.com/black-powder-amm

7

[Buffalo Arms - Ammunition - Black P](#)
WARNING: Discharging firearms in pc handling **ammunition** may result in ex
www.buffaloarms.com/browse.cfm?2,3

8

[Black powder ammunition | Shop big](#)
Black powder ammunition - Find the **ammunition** on sale. Shop by price, t coupons, and deals at ...
www.thefind.com/sports/info-black-pc

9

[Black Powder Ammunition - Sports](#)
Black powder ammunition, bullets, t outdoor superstore your one stop hur
www.sportsmansoutdoorsuperstore.co ammunition

10

[Murphy Powder Ammunition - W.A. J](#)
black powder, loaded, **ammunition** . per box: 32-20 wcf 115 mfp: 26.50: 3 145 mfp
www.murphy powder.com/ammunition

...

black powder wiki

[Gunpowder - Wikipedia, the free](#)
Characteristics · Coming · Combu
Gunpowder, also known since the of sulfur charcoal, and potassium potassium ...
en.wikipedia.org/wiki/Black_pow

[Black powder substitute - Wikipe](#)
Types of substitutes · Measurem
A **black powder** substitute is a re muzzleloading and cartridge firear advantages over **black** ...
en.wikipedia.org/wiki/Black_pow

[Black powder - Simple English W](#)
Black powder was the original gu explosive until the middle of the 1 more efficient ...
simple.wikipedia.org/wiki/Black_

[Blackpowder Courtesy - Dragon A](#)
Hawke is asked by the Viscount (Qunari, who has asked for him/he learns that the ...
dragonage.wikia.com/wiki/Black

[Blackpowder Promise - Dragon A](#)
Blackpowder Promise is a main q quest starts after Long Way Hom bodyguards at Dead ...
dragonage.wikia.com/wiki/Black

[Black Powder Mine - Guild Wars](#)
This page was last modified on 2 accessed 71,275 times. Content i wiki.guildwars.com/wiki/Black_P

[Black Powder - POTBS, Pirates](#)
Inexpensive hard cheeses are a circled milk, these che
potbs.wikia.com/wiki/Black_Pow

[Black Powder pistol - Wookieepe](#)
The **Black Powder** pistol was an slow to reload and pretty ineffecti collector ...
starwars.wikia.com/wiki/Black_F

[Black powder - HammerWiki, the](#)
Black powder or **gunpowder** is a l as a propellant in ranged weapons developed by ...
warhammeronline.wikia.com/wiki

[Black Powder Mine - GuildWars V](#)
Full: When **Black Powder** Mine is damage, become Blinded, and be Mine ends ...
guildwars.wikia.com/wiki/Black_

gun powder wiki

[Gunpowder - Wikipedia, the free encyclopedia](#)
Characteristics · Coming · Combustion ... · Sul
Gunpowder, also known since the late 19th ce of sulfur charcoal, and potassium nitrate. **Gunpowder** ...
en.wikipedia.org/wiki/Gunpowder

[History of gunpowder - Wikipedia, the free enc](#)
Original development · Medieval history · Early
Gunpowder was the first chemical explosive ar invention of nitrocellulose, nitroglycerin, smoke half of the ...
en.wikipedia.org/wiki/History_of_gunpowder

[Gun Powder - Resident Evil Wiki - The Reside](#)
The **Gun Powder** is a item found in Resident Ev Tool item to make ammo for your weapons, wtr combined ...
residentevil.wikia.com/Gun_Powder

[Gunpowder - Minecraft Wiki](#)
History · In Indev, mob drops were randomized drop **gunpowder** instead of Pork (seen here). C Beta 1.0, but ...
www.minecraftwiki.net/wiki/Gunpowder

Related Searches for **gun powder** wiki
Gun Powder Manufacturers Reloading Supp
Surplus Gunpowder How to Make G
Hercules Gun Powder Accurate Gun I

[Gunpowder - Simple English Wikipedia, the fre](#)
Gunpowder (or **gun powder**) is a mix of chemi salt(peter). It burns very rapidly, and creates gt space than the ...
simple.wikipedia.org/wiki/Gunpowder

[Gunpowder Plot - Simple English Wikipedia, th](#)
Origins · Planning · Discovery · Trial and execu
The **Gunpowder** Conspiracy of 1605 was a plan and VI of Scotland. Other names for the plot ar **Gunpowder** Plot A group ...
simple.wikipedia.org/wiki/Gunpowder_Plot

[Gunpowder - The TimeSplitters Wiki](#)
The **Gunpowder** is an obtainable piece of equip West, in the Story of TimeSplitters 2. Its primar from the ...
timesplitters.wikia.com/wiki/Gunpowder

[Gunpowder](#)
Gunpowder is an explosive mixture of sulfur, (also known as saltpeter/saltpeter) that burns r solids and ...
schools-wikipedia.org/wp/g/Gunpowder.htm

[Gunpowder - Vindictus Wiki](#)
A very unstable **powder** that will explode by he making bombs. It can be purchased from Aislin
www.vindictuswiki.com/wiki/Gunpowder

[Gunpowder - encyclopedia article - Citizendium](#)
For more information, see: Explosives. For mon **Gunpowder** (a generic term that includes black a propellant that has ...
en.citizendium.org/wiki/Gunpowder

Experiment

black powder
ammunition

black powder
wiki

gun powder
wiki

1
2
3
4
5
6
7
8
9
10
...



Construct appropriate test collections

Rethink of evaluation measures

What is a good system?

How can we measure “goodness”?

Measuring “goodness”

The user steps down a ranked list of documents and observes each one of them until a decision point and either

a) abandons the search, or

b) reformulates

While stepping down or sideways, the user accumulates utility

What are the challenges?

Evaluation over a single ranked list

black powder ammunition

1	Black Powder Bullets for Black Powder Bullets C ammunition, hunting su unbeatable prices. www.ableammo.com/ca
2	Buffalo Arms - Ammuni Black Powder Cartridge largest on hand invento old, obsolete and www.buffaloarms.com/b
3	www.powderinc.com https://www.powderinc.com ACTION=enter&thispag
4	Black Powder Hunting, Count on Able Ammo to shooting supplies, gun www.ableammo.com/ca
5	Black Dawg Cartridge Black Dawg Cartridges June 14, 2011. Change .45 cal, 450 ... www.blackdawgcartridge.com
6	Black Powder Ammuni Black Powder Ammuni Explosion, 20 Sticks, 7 Series Black ... www.nextag.com/black
7	Buffalo Arms - Ammuni WARNING: Discharging handling ammunition m www.buffaloarms.com/b
8	Black powder ammuni Black powder ammuni ammunition on sale. Si coupons, and deals at www.thefind.com/sports
9	Black Powder Ammuni Black powder ammuni outdoor superstore your www.sportsmansoutdoor.com ammunition
10	Murphy Powder Ammu black, powder, loaded, per box: 32-20 wcf 115 145 mfp www.murphy powder.com/ammunition
...	



Existing measures

- Session DCG [Järvelin et al ECIR 2008]
The user steps down the ranked list until rank k and reformulates [**Deterministic; no early abandonment**]
- Expected session utility [Yang and Lad ICTIR 2009]
The user steps down a ranked list of documents until a decision point and reformulates [**Stochastic; no early abandonment**]

Model-based measures

Probabilistic space of users following
different paths

- Ω is the space of all paths
- $P(\omega)$ is the prob of a user following a path ω in Ω
- M_ω is a measure over a path ω

$$esM = \sum_{\omega \in \Omega} P(\omega) M_\omega$$

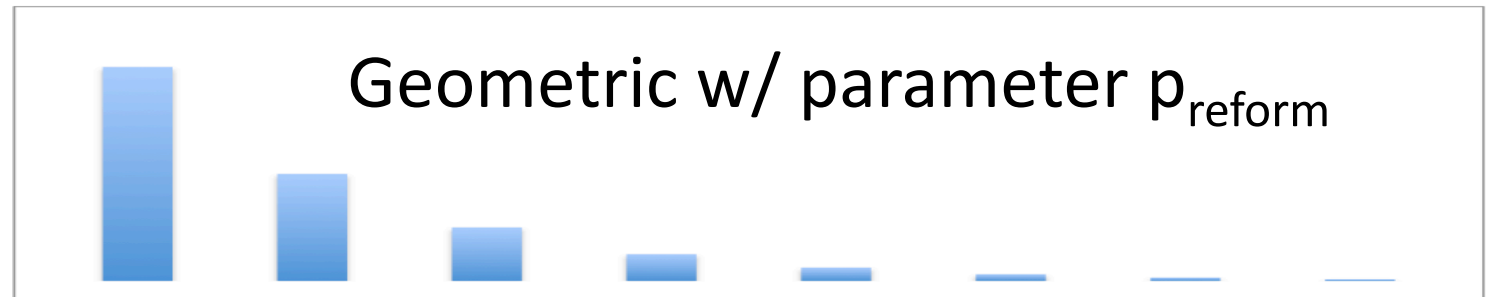
Probability of a path

Q1	Q2	Q3
N	R	R
N	R	R
N	R	R
N	R	R
N	R	R
N	N	R
N	N	R
N	N	R
N	N	R
N	N	R
...

(1) Probability of
abandoning at reform 2

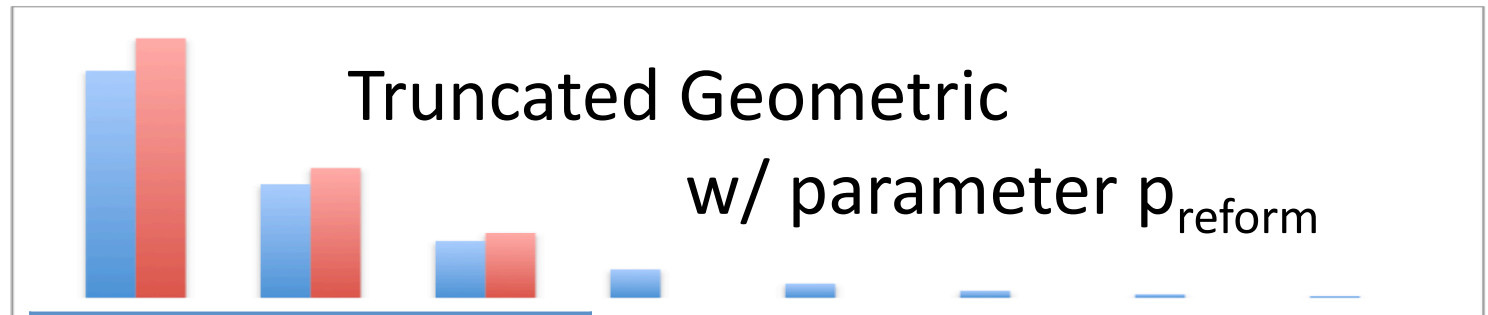
X

(2) Probability of
reformulating at rank 3



Q1	Q2	Q3
N	R	R
N	R	R
N	R	R
N	R	R
N	R	R
N	N	R
N	N	R
N	N	R
N	N	R
N	N	R
...

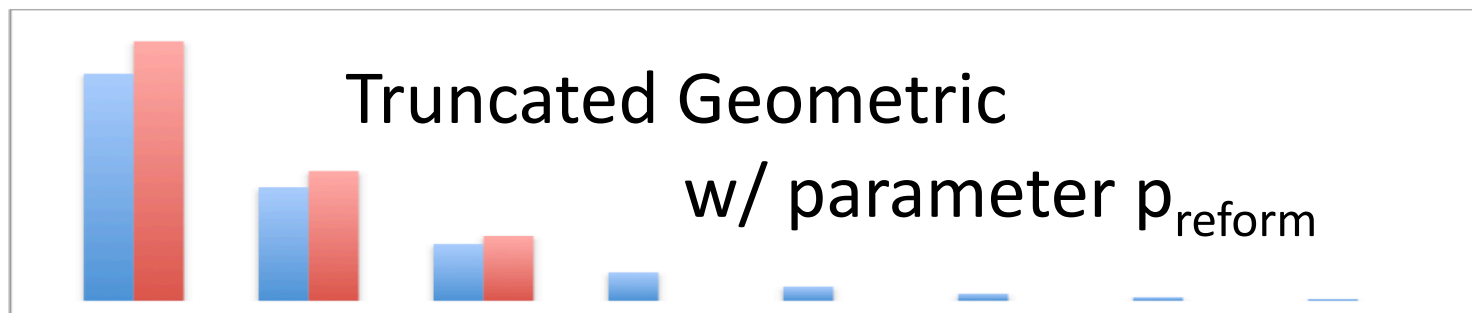
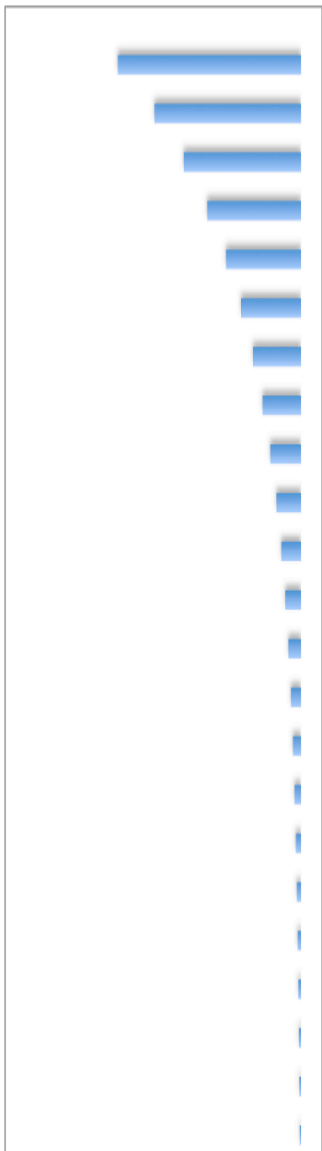
(1) Probability of abandoning the session at reformulation i



Q1	Q2	Q3
N	R	R
N	R	R
N	R	R
N	R	R
N	R	R
N	N	R
N	N	R
N	N	R
N	N	R
...

(1) Probability
of abandoning
the session at
reformulation i

Geometric w/ parameter p_{down}



Q1	Q2	Q3
N	R	R
N	R	R
N	R	R
N	R	R
N	R	R
N	N	R
N	N	R
N	N	R
N	N	R
...

Probability
(2) of reformulating
at rank j

Model-based measures

Probabilistic space of users following
different paths

- Ω is the space of all paths
- $P(\omega)$ is the prob of a user following a path ω in Ω
- M_ω is a measure over a path ω

$$esM = \sum_{\omega \in \Omega} P(\omega) M_\omega$$

Today's Outline

- Evaluation based on user models
 - Gain/Discount Framework
 - Modeling browsing behavior
 - RBP, AP, DCG, ERR, EBU
 - Modeling utility
 - NDCG, GAP
- Novelty and diversity
- Query sessions