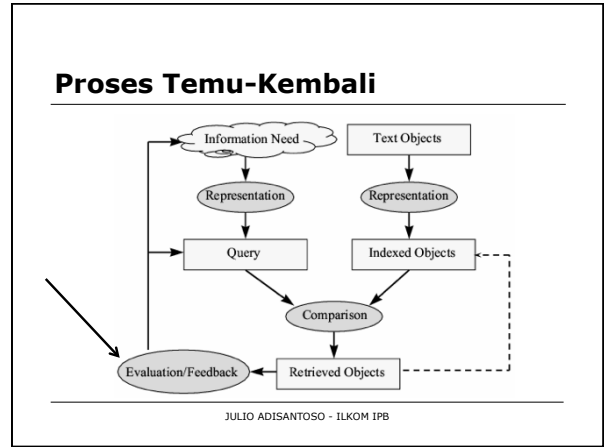


# KOM341 Temu Kembali Informasi

---

KULIAH #5  
• Evaluasi IR



### Contoh query : sby

Presiden Republik Indonesia - Dr. H. **Susilo Bambang Yudhoyono**  
 Presiden SBY menyerahkan Pita Cita Pahlawan Peran 2008 kepada Kasu-SDM Polda ...  
 Presiden SBY mengesahkan hal itu dalam sambutannya saat menyerahkan Pita ...  
[www.presidenby.info - 236 - Tembakola - Halaman Resmi](#)

**Foto** **Pusat** **Pemerintah**  
**Kolom** **Pada**  
**Catatan** **Batas** **Dana**  
**Kabinet** **Indonesia** **Reruntu** **Buang** **Dara**

*Hasil temuan lainnya dari presidenby.info*

Presiden Republik Indonesia - Dr. H. **Susilo Bambang Yudhoyono**  
 Presiden SBY eseperti banyak rakyat menanggapinya saat pada 9 September 1949 ... **Susilo Bambang Yudhoyono** meraih julukan terbak **AKABRI** Darat tahun 1973 ...  
[www.presidenby.info/index.php?option=com\\_content&view=article&layout=edit - 126 - Tembakola - Halaman Resmi](#)

**YouTube - SBY** **ngamuk**  
 SBY masih bisa ketawa ketawa, asap seperti itu?sequal: [http://www.youtube.com/watch?v=DualH0KQZYA](#)

**Susilo Bambang Yudhoyono** - Wikipedia bahasa Indonesia  
 Yudhoyono yang dipanggil Sus oleh orang tuaanya dan populer dengan panggilan **SBY** lahir di Pacitan, Jawa Timur pada 9 September 1949 ...  
[id.wikipedia.org/wiki/Susilo\\_Bambang\\_Yudhoyono - 626 - Tembakola - Halaman Resmi](#)

**delikinet** situs warta era digital | **SBY** Geli Tanggapi Kasus P-00 ...  
 Presiden **Susilo Bambang Yudhoyono** menyampaikan jawaban di sidang pers di Gedung Sate ...  
[www.delikinet.com/read/2009/12/16/1602251920227359389-981-tanggapi-kasus-p00-sby-bambang - 339 - Tembakola - Halaman Resmi](#)

**Antenasas 17 Agustus 1945 Surabaya**  
**Wabakar** Resmi Universitas 17 Agustus 1945 Surabaya  
[www.antenasas17ag.45 - 238 - Tembakola - Halaman Resmi](#)

JULIO ADISANTOSO - ILKOM IPB

### Contoh query: flu burung

**Flu burung** - Wikipedia bahasa Indonesia, ensiklopedia bebas  
 Ada dua tipe demam yang bagian dan arteri seperti flu burung tergolong ke arteri atau bagian ke (Purba) ...  
[id.wikipedia.org/wiki/Flu\\_burung - 406 - Tembakola - Halaman Resmi](#)

**Demam Berdarah** **Indonesia**  
 Flu burung adalah suatu penyakit menular yang disebabkan oleh virus ... Saat ini, penyebab flu burung adalah Highly Pathogenic Avian Influenza Virus ...  
[www.medika.go.id/index.php?option=com\\_content&view=article&layout=edit - 306 - Tembakola - Halaman Resmi](#)

**perti FLU BURUNG**  
 Jema Berhari - POF/Adobe Acrobat - [Virus ISM](#)  
 Penyakit flu burung atau flu avian influenza (AI) merupakan penyakit yang disebabkan oleh virus influenza tipe A. Umumnya tipe ini disebabkan pada burung dan unggas ...  
[www.komradpgo.go.id/flu\\_burung.html - 298 - Tembakola - Halaman Resmi](#)

**Serba Serbi Flu Burung**  
 Flu burung atau Avian influenza adalah penyakit menular yang disebabkan oleh virus influenza tipe A. Umumnya tipe ini disebabkan pada burung dan unggas ...  
[www.komradpgo.go.id/flu\\_burung.html - 298 - Tembakola - Halaman Resmi](#)

**Flu Burung (Avian Influenza)** - **Libana Pertanian**  
 Penyakit flu burung atau Avian influenza (AI) merupakan penyakit yang disebabkan oleh virus influenza yang menyerang burung/unggas/air ...  
[www.libana.pertanian.go.id/flu\\_burung.html - 154 - Tembakola - Halaman Resmi](#)

**TempoInteraktif.Com - Flu Burung**  
 Flu burung merupakan infeksi virus influenza A subtype H5N1 (Pernamagukan, Di Indonesia) yang pada umumnya menyerang unggas, burung dan ayam lain ...  
[www.tempointeraktif.com/koran/2004/02/28/flu\\_20040228-04\\_24.html - 204 - Tembakola - Halaman Resmi](#)

**Mengetahui dan Menakuti Penyakit Flu Burung** - [www.dinikawalel.com](#)  
 28 Sep 2007 - Penyakit flu burung dapat menula dari hewan ke manusia ... Sementara di Indonesia, penyakit flu burung menjadi penyakit zoonosis (dat ...

JULIO ADISANTOSO - ILKOM IPB

### Evaluasi IR

Indikator yang dapat diukur:

- Seberapa cepat dia meng-indeks
  - Banyaknya dokumen/jam
  - Terkait dengan ukuran dokumen
- Seberapa cepat dia mencari
  - Fungsi dari ukuran indeks

Indikator yang sulit diukur:

- Seberapa "bahagia" user ??
  - Kesesuaian jawaban hasil query
  - Ketepatan
- Seberapa "mudah" user ??
  - Menyusun query

JULIO ADISANTOSO - ILKOM IPB

### Versi user

- Seberapa tergantung setting.
- Web mesin pencari:
  - User mendapatkan apa yg diinginkan dan kembali melanjutkan pencarian
- Situs eCommerce:
  - User mendapatkan apa yang diinginkan dan membuat pemesanan
- Umumnya adalah mengukur kepuasan user.

JULIO ADISANTOSO - ILKOM IPB

## Kepuasan user

- Mengukur relevansi hasil query.
- Masalah: bagaimana mengukur relevansi?
- Tiga elemen:
  - Koleksi dokumen
  - Kumpulan query
  - Pasangan query-dokumen relevan dan tidak relevan (relevance judgments)

JULIO ADISANTOSO - ILKOM IPB

## Evaluasi IR

- Kebutuhan informasi diterjemahkan ke dalam bentuk query.
- Relevansi lebih kepada kesesuaian dengan kebutuhan informasi, bukan pada query.
- Contoh kebutuhan informasi: Saya mencari informasi tentang jenis kacang yang dapat meningkatkan kolesterol dan menimbulkan resiko pada jantung.
- Query: ***kacang kolesterol resiko jantung***

JULIO ADISANTOSO - ILKOM IPB

## Standard test collections

- The Cranfield collection. Collected in the United Kingdom starting in the late 1950s, it contains 1398 abstracts of aerodynamics journal articles, a set of 225 queries, and exhaustive relevance judgments of all (query, document) pairs.
- Text Retrieval Conference (TREC) by The U.S. National Institute of Standards and Technology (NIST). these test collections comprise 6 CDs containing 1.89 million documents and relevance judgments for 450 information needs.

JULIO ADISANTOSO - ILKOM IPB

## Standard test collections

- NII Test Collections for IR Systems (NTCIR). Similar sizes to the TREC collections, focusing on East Asian language and cross-language information retrieval, where queries are made in one language over a document collection containing documents in one or more other languages. <http://research.nii.ac.jp/ntcir/data/data-en.html>
- Cross Language Evaluation Forum (CLEF). This evaluation series has concentrated on European languages and cross-language information retrieval. <http://www.clef-campaign.org/>

JULIO ADISANTOSO - ILKOM IPB

## Standard test collections

- Reuters-21578 and Reuters-RCV1. Reuters-21578 collection: 21578 newswire articles. RCV1 : Reuters Corpus Volume 1 (RCV1), consisting of 806,791 documents.
- 20 Newsgroups. This is another widely used text classification collection, collected by Ken Lang. It consists of 1000 articles from each of 20 Usenet newsgroups, it contains 18941 articles.

JULIO ADISANTOSO - ILKOM IPB

## Evaluation of retrieval sets

- Precision:
  - rasio dokumen yang di-retrieve adalah relevan
  - $P(\text{relevant}|\text{retrieved})$

$$\text{Precision} = \frac{\#(\text{relevant items retrieved})}{\#(\text{retrieved items})}$$

- Recall:
  - rasio dokumen relevan yang di-retrieve
  - $P(\text{retrieved}|\text{relevant})$

$$\text{Recall} = \frac{\#(\text{relevant items retrieved})}{\#(\text{relevant items})}$$

JULIO ADISANTOSO - ILKOM IPB

### Evaluation of retrieval sets

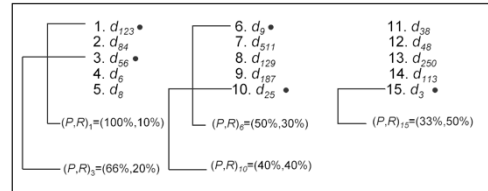
	Relevant	Not Relevant
Retrieved	tp	fp
Not Retrieved	fn	tn

- Precision =  $P = tp / (tp + fp)$
- Recall =  $R = tp / (tp + fn)$
- Accuracy =  $(tp + tn) / (tp + fp + fn + tn)$

JULIO ADISANTOSO - ILKOM IPB

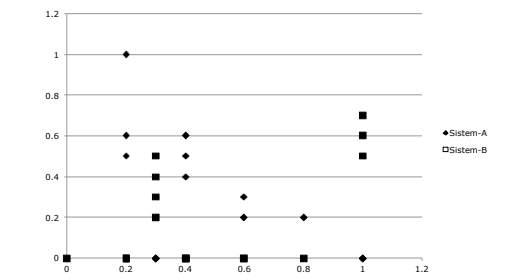
### Contoh

- $R_q = \{d_3, d_5, d_9, d_{25}, d_{39}, d_{44}, d_{56}, d_{71}, d_{89}, d_{123}\}$
- Ten relevant documents
- A ranking of the documents for the given query  $q$



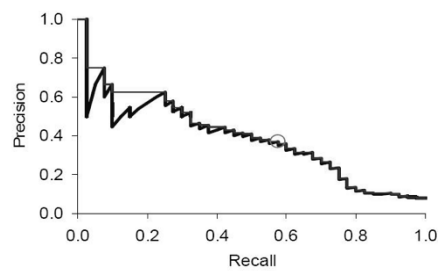
JULIO ADISANTOSO - ILKOM IPB

### Membandingkan 2 sistem IR



JULIO ADISANTOSO - ILKOM IPB

### Kurva R-P



JULIO ADISANTOSO - ILKOM IPB

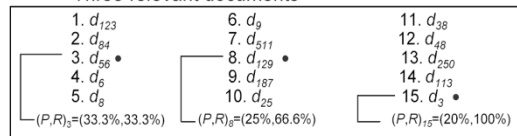
### Interpolasi

- Nilai R-P tiap query berbeda sehingga sulit membandingkan antar metode.
- Perlu dilakukan interpolasi.
- Cara interpolasi:
  - Menghubungkan titik
  - Menghubungkan titik maksimum
  - Menghubungkan titik minimum
  - Menghubungkan titik rata-rata

JULIO ADISANTOSO - ILKOM IPB

### Contoh

- $R_q = \{d_3, d_{56}, d_{129}\}$
- Three relevant documents



JULIO ADISANTOSO - ILKOM IPB

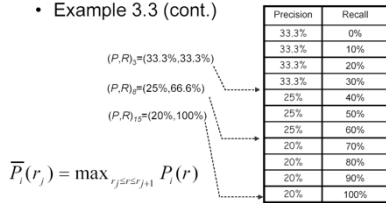
### Interpolasi

- Interpolated Precisions at standard recall levels

$$\bar{P}(r_j) = \max_{r_j \leq r \leq r_{j+1}} P(r)$$

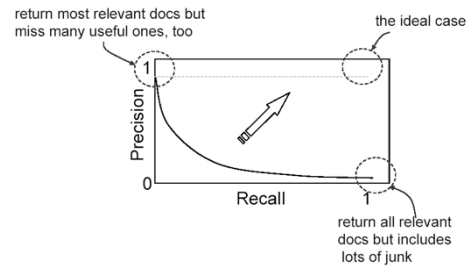
- the  $j$ -th standard recall level (e.g.,  $r_5$  is recall level 50%)

- Example 3.3 (cont.)



JULIO ADISANTOSO - ILKOM IPB

### Trade-off



JULIO ADISANTOSO - ILKOM IPB

### Precision Rata-rata

- Ada  $k$  buah query  $\{q_1, q_2, \dots, q_k\}$
- Untuk query tertentu, hitung titik P/R untuk setiap dokumen yang relevan, pada titik recall standar.
- Hitung rata-rata Precision setiap query pada setiap titik recall yang standar.

$$P(r) = \frac{1}{k} \sum_{j=1}^k P_j(r)$$

JULIO ADISANTOSO - ILKOM IPB

### Contoh

Query	q1	q2	q3
Dokumen relevan	d1, d8, d10, d120, d15	d8, d9, d25, d40, d78, d85, d88, d100	d7, d10, d12, d20
Hasil query (sesuai ranking)	d10, d50, d8, d19, d100, d30, d15, d80, d92, d65	d100, d90, d32, d65, d78, d25, d88, d95, d62, d120	d10, d15, d90, d7, d95, d12, d120, d30, d20, d100

JULIO ADISANTOSO - ILKOM IPB

### MAP: Mean average precision

- Nilai Precision rata-rata diperoleh dari top  $k$  dokumen, setiap kali suatu dokumen yang relevan diperoleh
- Menghindari interpolasi, tidak menggunakan titik recall yang tetap.
- MAP untuk koleksi query adalah rata-ratanya.

$$MAP(Q) = \frac{1}{|Q|} \sum_{j=1}^{|Q|} \frac{1}{m_j} \sum_{k=1}^{m_j} Precision(R_{jk})$$

JULIO ADISANTOSO - ILKOM IPB

### MAP: Contoh

1. $d_{123}$ • ( $P=1.0$ )	6. $d_9$ • ( $P=0.5$ )	11. $d_{38}$
2. $d_{84}$	7. $d_{511}$	12. $d_{48}$
3. $d_{56}$ • ( $P=0.66$ )	8. $d_{129}$	13. $d_{250}$
4. $d_6$	9. $d_{187}$	14. $d_{113}$
5. $d_8$	10. $d_{28}$ • ( $P=0.4$ )	15. $d_3$ • ( $P=0.3$ )

$(1.0+0.66+0.5+0.4+0.3)/5=0.57$

JULIO ADISANTOSO - ILKOM IPB

### R-Precision

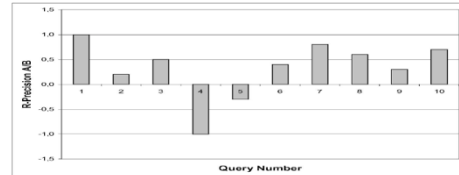
- Generate a single value summary of ranking by computing the precision at the  $R$ -th position in the ranking
  - Where  $R$  is the total number of relevant docs for the current query

1. $d_{123}$ ●	6. $d_9$ ●	11. $d_{38}$
2. $d_{84}$ ●	7. $d_{511}$	12. $d_{48}$
3. $d_{56}$ ● ■	8. $d_{129}$ ■	13. $d_{250}$
4. $d_6$	9. $d_{87}$	14. $d_{113}$
5. $d_8$	10. $d_{25}$ ●	15. $d_3$ ● ■

$R_A = \{d_5, d_9, d_{25}, d_{38}, d_{48}, d_{56}, d_{71}, d_{89}, d_{123}\}$       $R_B = \{d_3, d_{56}, d_{123}\}$   
 •10 relevant documents (●)     •3 relevant document (■)  
 =>  $R$ -precision =  $4/10=0.4$      =>  $R$ -precision =  $1/3=0.33$

JULIO ADISANTOSO - ILKOM IPB

### Precision histograms



- A positive  $RP_{A/B}(i)$  indicates that the algorithm  $A$  is better than  $B$  for the  $i$ -th query and vice versa

JULIO ADISANTOSO - ILKOM IPB

### F-measure

- The weighted harmonic mean of precision and recall:

$$F = \frac{1}{\alpha \frac{1}{P} + (1-\alpha) \frac{1}{R}} = \frac{(\beta^2 + 1)PR}{\beta^2 P + R} \quad \text{where} \quad \beta^2 = \frac{1-\alpha}{\alpha}$$

- Balanced F-measure : bobot  $P$  sama dengan bobot  $R$ , artinya  $\alpha=1/2$  atau  $\beta = 1$ , sehingga

$$F_{\beta=1} = \frac{2PR}{P+R}$$

JULIO ADISANTOSO - ILKOM IPB

### Assessing relevance

- Menentukan relevansi dokumen terhadap suatu query menggunakan pooling dari beberapa ahli.
- Menilai hasil menggunakan statistik Kappa:

$$kappa = \frac{P(A) - P(E)}{1 - P(E)}$$

dimana  $P(A)$ : proporsi banyaknya penilai yang setuju,  $P(E)$ : persetujuan yang merupakan kebetulan

JULIO ADISANTOSO - ILKOM IPB

### Contoh

		Judge 2 Relevance		
		Yes	No	Total
Judge 1 Relevance	Yes	300	20	320
	No	10	70	80
Total		310	90	400

- $P(A) = (300 + 70)/400 = 370/400 = 0.925$
- Pooled marginals
  - $P(nr) = (80 + 90)/(400 + 400) = 170/800 = 0.2125$
  - $P(r) = (320 + 310)/(400 + 400) = 630/800 = 0.7878$
- $P(E) = P(nr)^2 + P(r)^2 = 0.2125^2 + 0.7878^2 = 0.665$
- Statistik Kappa:  
 $(0.925 - 0.665)/(1 - 0.665) = 0.776$
- Kesimpulan nilai Kappa:
  - $Kappa \geq 0.8$  = persetujuannya baik
  - $0.67 \leq Kappa < 0.8$  = persetujuan yang fair
  - $Kappa < 0.67$  = ditolak

JULIO ADISANTOSO - ILKOM IPB