

# Evaluación de Resultados



UCR – ECCI

CI-2414 Recuperación de Información

Prof. Kryscia Daviana Ramírez Benavides

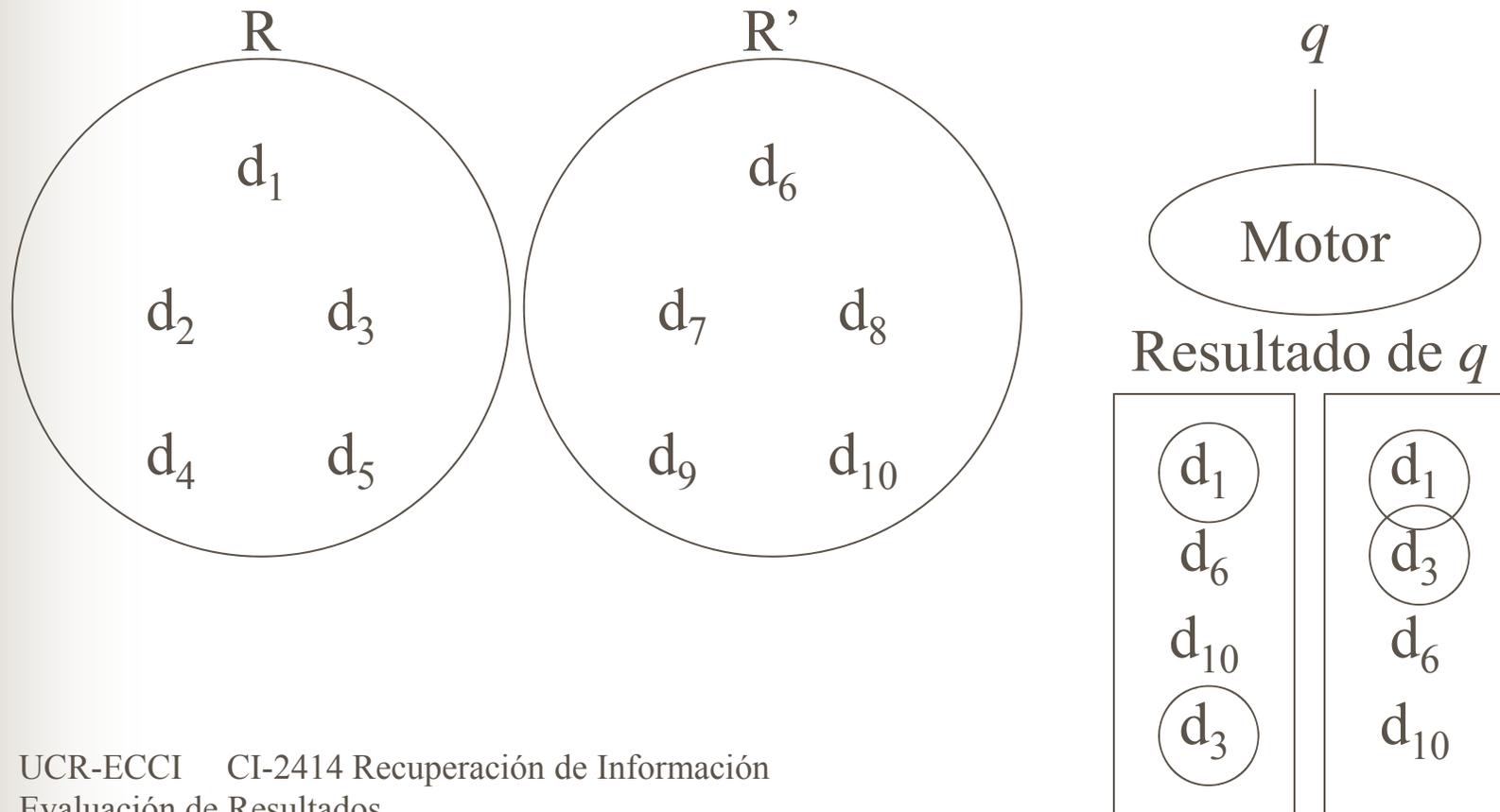


## Introducción

- En un modelo de RI la calidad de la respuesta es lo importante, aunque no por eso el tiempo y velocidad dejan de ser importantes.
- Para medir la bondad de un modelo de RI, aunque existen muchas medidas, la más popular es un diagrama de *precisión – recall* (*precisión – recuperación*).

## Introducción (cont.)

- Se agarra una colección y se clasifica:





## Conceptos Generales

- *Precisión*: El número de documentos recuperados que son relevantes.
- *Recall (índice de recuperación)*: El número de documentos relevantes que se recuperaron.
- *Generality (índice de relevancia)*: Proporción entre documentos relevantes y el tamaño de la colección.
- *Fallout (índice de fallo)*: Proporción entre documentos no relevantes recuperados y documentos no relevantes en la colección.
- Según la cantidad que se elige como relevantes se puede hacer aumentar una a costa de reducir la otra.

## Nomenclatura

- $Ra \rightarrow$  número de documentos relevantes que son recuperados
- $A \rightarrow$  número de documentos recuperados
- $R \rightarrow$  número de documentos relevantes para la consulta
- $P_i \rightarrow$  precisión en  $i$  de documentos recuperados
- $\overline{P} \rightarrow$  precisión promedio en  $n$  documentos recuperados
- $P_i(r) \rightarrow$  precisión en el nivel  $r$  de *recall*
- $\overline{P}(r) \rightarrow$  precisión promedio en el nivel  $r$  de *recall*

## Nomenclatura (cont.)

- $N_a \rightarrow$  número de documentos recuperados
- $N_q \rightarrow$  número de consultas realizadas
- $N \rightarrow$  número de documentos de la colección

# Fórmulas

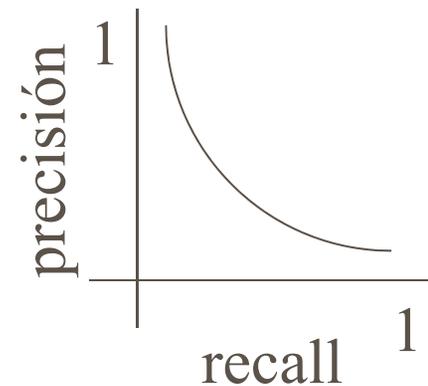
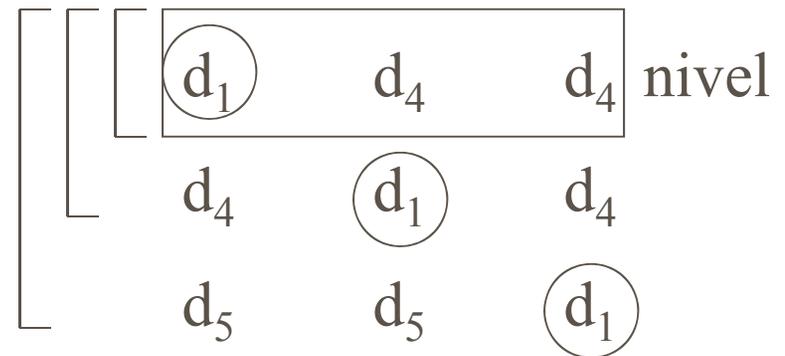
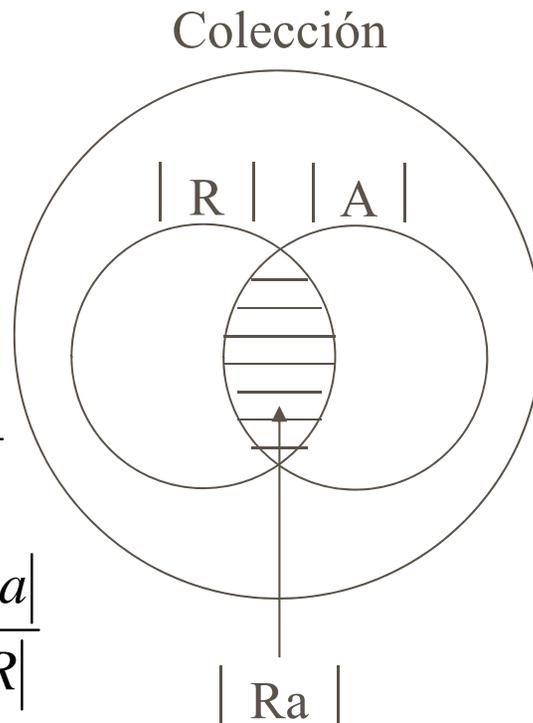
$$Precisión = \frac{|Ra|}{|A|} \quad Recall = \frac{|Ra|}{|R|}$$

$$\bar{P}(r) = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q}$$

$$\bar{P} = \sum_{i=1}^{N_a} \frac{P_i}{N_a}$$

$$Generality = \frac{|R|}{|N|}$$

$$Fallout = \frac{|A - Ra|}{|N - R|}$$





## Generalidades

- Es fácil tener alta precisión, basta con retornar el documento más relevante, pero entonces el *recall* es cercano a 0.
- Es fácil tener alto *recall*, basta retornar todos los documentos relevantes, pero ahora la precisión es cercana a 0.
- En ciertos casos, como la Web, es posible calcular la precisión, pero puede ser imposible calcular el *recall*.
  - Se puede hacer mediante muestreo.
- El objetivo es tratar de aumentar ambos al mismo tiempo.
- Entre valores más altos se obtenga en la precisión promedio significa que los documentos relevantes están más arriba.



## Generalidades (cont.)

- Para evaluar un modelo de RI se debe tener en cuenta todo el gráfico.
  - Gráfico *Precisión-Recall*: Si se quiere conocer cuál es la *precisión* para un valor arbitrario de *recall* se realiza una interpolación de orden 0 hacia a la izquierda, es decir, se toma el valor de *precisión* del primer *punto de observación* a la derecha.
  - Gráfico *Precisión en 11 niveles de Recall*: Esta gráfica es la más utilizada, se calcula la *precisión* en 11 puntos de *recall* correspondientes a 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 y 1.
    - La forma de construir esta gráfica es calcular para la consulta los valores de *precisión* en los 11 puntos de *cobertura*.



## Generalidades (cont.)

- Para evaluar un modelo de RI se debe tener en cuenta todo el gráfico.
  - Gráfico *Precisión Promedio en 11 niveles de Recall*: Esta gráfica es la más utilizada, se calcula la *precisión promedio* en 11 puntos de *recall* correspondientes a 0, 0.1, 0.2, ..., 0.9 y 1.
    - La forma de construir esta gráfica es calcular para cada consulta los valores de *precisión* en los 11 puntos de *cobertura* y hacer la media de todas las consultas en cada punto.
- **Interpolación:** Si se quiere conocer cuál es la precisión para un valor arbitrario de *recall* se toma el valor de precisión del primer punto de observación *precisión-recall* a la derecha.
- Por convenio se toma que la precisión se hace cero cuando ya no quedan documentos relevantes en los recuperados.

## Ejemplo

- $Colección = \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9, d_{10}, d_{11}, d_{12}, d_{13}, d_{14}, d_{15}\}$
- $Rq = \{d_1, d_2, d_5, d_7, d_{10}\}$
- Documentos recuperados del Motor de Búsqueda #1:  
 $Aq = \{d_1, d_2, d_4, d_5, d_7, d_8, d_{10}, d_{12}, d_{13}, d_{14}\}$
- Documentos recuperados del Motor de Búsqueda #2:  
 $Aq = \{d_1, d_4, d_3, d_2, d_8, d_9, d_{10}, d_5, d_7, d_{14}\}$

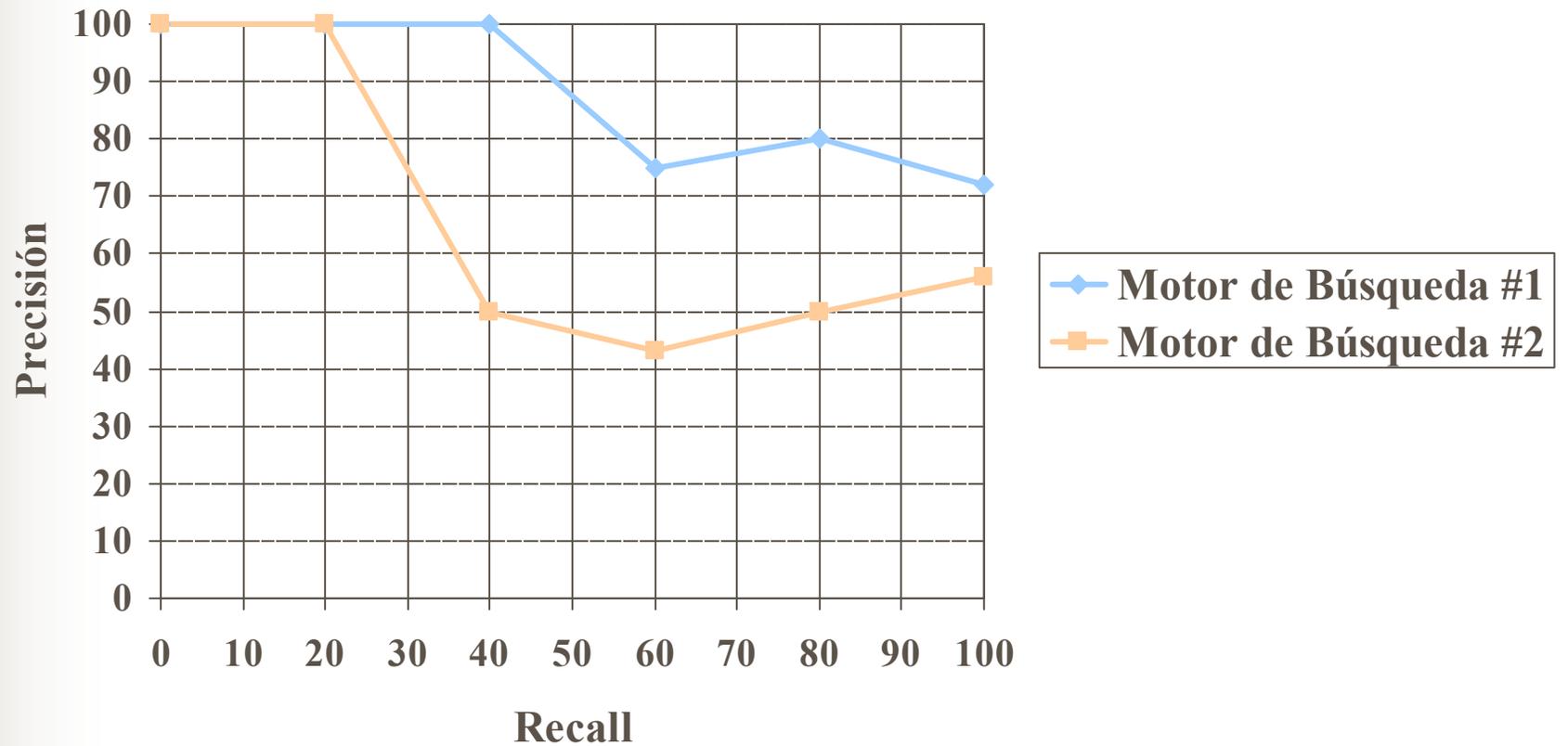
## Ejemplo (cont.): Cálculos Precisión-Recall Motor de Búsqueda #1

<b>Relevante</b>	<b>Documento</b>	<b>Precisión</b>	<b>Recall</b>
<b>Sí</b>	<b>1</b>	<b>1/1</b>	<b>1/5</b>
<b>Sí</b>	<b>2</b>	<b>2/2</b>	<b>2/5</b>
No	4	2/3	2/5
<b>Sí</b>	<b>5</b>	<b>3/4</b>	<b>3/5</b>
<b>Sí</b>	<b>7</b>	<b>4/5</b>	<b>4/5</b>
No	8	4/6	4/5
<b>Sí</b>	<b>10</b>	<b>5/7</b>	<b>5/5</b>
No	12	5/8	5/5
No	13	5/9	5/5
No	14	5/10	5/5

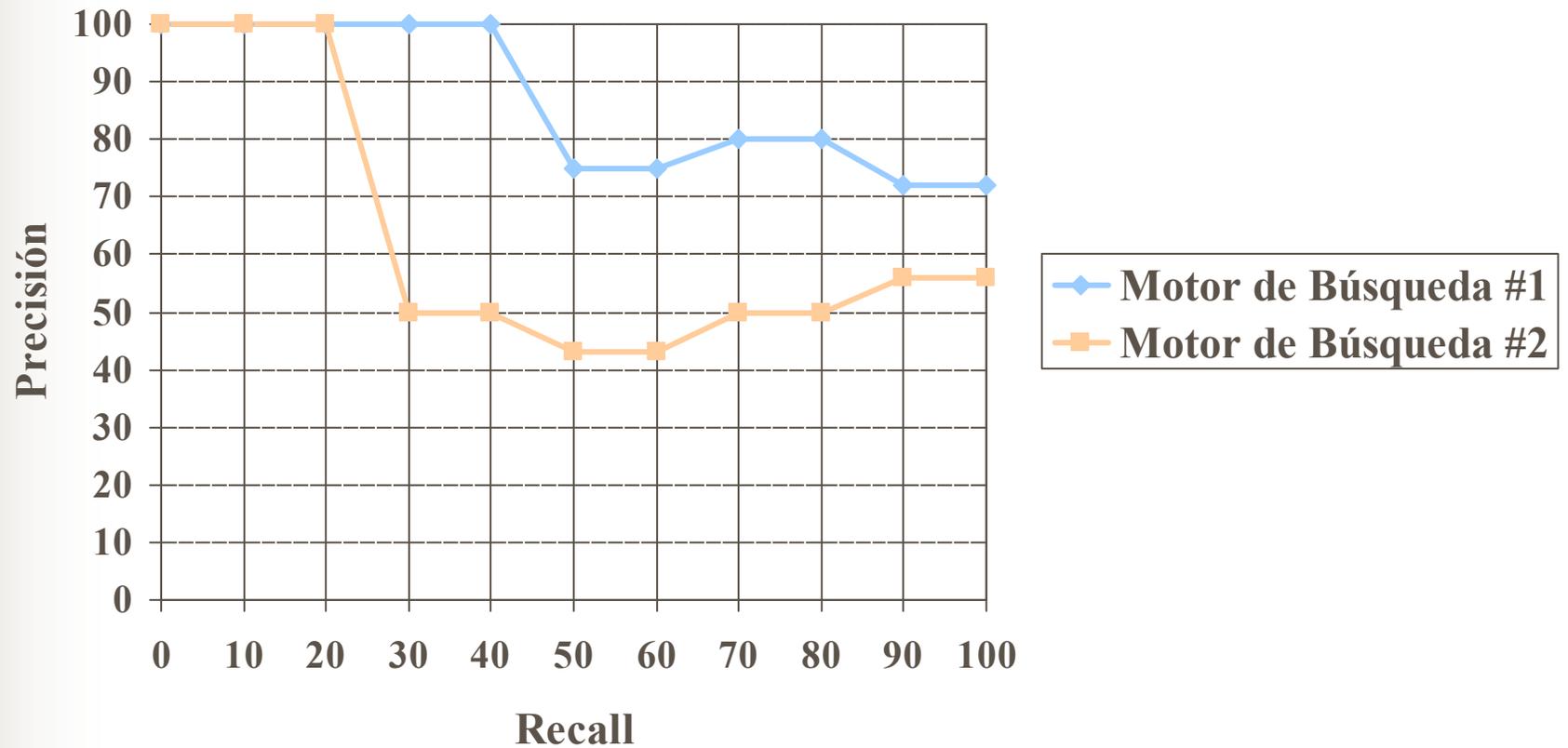
## Ejemplo (cont.): Cálculos Precisión-Recall Motor de Búsqueda #2

<b>Relevante</b>	<b>Documento</b>	<b>Precisión</b>	<b>Recall</b>
<b>Sí</b>	<b>1</b>	<b>1/1</b>	<b>1/5</b>
No	4	1/2	1/5
No	3	1/3	1/5
<b>Sí</b>	<b>2</b>	<b>2/4</b>	<b>2/5</b>
No	8	2/5	2/5
No	9	2/6	2/5
<b>Sí</b>	<b>10</b>	<b>3/7</b>	<b>3/5</b>
<b>Sí</b>	<b>5</b>	<b>4/8</b>	<b>4/5</b>
<b>Sí</b>	<b>7</b>	<b>5/9</b>	<b>5/5</b>
No	14	5/10	5/5

## Ejemplo (cont.): Gráfico Precisión-Recall



## Ejemplo (cont.): Gráfico Precisión en 11 niveles de Recall



## Ejemplo (cont.): Precisión Promedio en $n$ Documentos Recuperados

$$\bar{P}_{qMB1} = \frac{(1/1 + 2/2 + 2/3 + 3/4 + 4/5 + 4/6 + 5/7 + 5/8 + 5/9 + 5/10)}{10}$$

$$\bar{P}_{qMB1} = \frac{7.28}{10} = 0.728$$

$$\bar{P}_{qMB2} = \frac{(1/1 + 1/2 + 1/3 + 2/4 + 2/5 + 2/6 + 3/7 + 4/8 + 5/9 + 5/10)}{10}$$

$$\bar{P}_{qMB2} = \frac{5.05}{10} = 0.505$$

## Ejemplo (cont.): Cálculos Precisión Promedio en el Nivel $r$ de Recall

- $Colección = \{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9, d_{10}, d_{11}, d_{12}, d_{13}, d_{14}, d_{15}\}$
- $Rq_1 = \{d_1, d_2, d_5, d_7, d_{10}\}$
- $Rq_2 = \{d_4, d_8, d_{10}, d_{11}, d_{15}\}$
- $Rq_3 = \{d_3, d_6, d_9, d_{13}, d_{14}\}$
- Documentos recuperados del Motor de Búsqueda #1:
  - $Aq_1 = \{d_1, d_2, d_4, d_5, d_7, d_8, d_{14}\}$
  - $Aq_2 = \{d_5, d_4, d_7, d_8, d_{10}, d_{11}, d_{13}\}$
  - $Aq_3 = \{d_6, d_3, d_4, d_5, d_7, d_9, d_{14}\}$
- Documentos recuperados del Motor de Búsqueda #2:
  - $Aq_1 = \{d_1, d_2, d_3, d_4, d_7, d_9, d_{10}\}$
  - $Aq_2 = \{d_4, d_8, d_3, d_2, d_6, d_{15}, d_{11}\}$
  - $Aq_3 = \{d_3, d_2, d_4, d_6, d_9, d_8, d_{14}\}$

## Ejemplo: Cálculos Precisión Promedio en el Nivel $r$ de Recall - Motor de Búsqueda #1

<b>Consulta #1</b>	<b>Precisión Consulta #1</b>	<b>Consulta #2</b>	<b>Precisión Consulta #2</b>	<b>Consulta #3</b>	<b>Precisión Consulta #3</b>
<b>1</b>	<b>1/1</b>	<b>5</b>	<b>0/1</b>	<b>6</b>	<b>1/1</b>
<b>2</b>	<b>2/2</b>	<b>4</b>	<b>1/2</b>	<b>3</b>	<b>2/2</b>
<b>4</b>	<b>2/3</b>	<b>7</b>	<b>1/3</b>	<b>4</b>	<b>2/3</b>
<b>5</b>	<b>3/4</b>	<b>8</b>	<b>2/4</b>	<b>5</b>	<b>2/4</b>
<b>7</b>	<b>4/5</b>	<b>10</b>	<b>3/5</b>	<b>7</b>	<b>2/5</b>
<b>8</b>	<b>4/6</b>	<b>11</b>	<b>4/6</b>	<b>9</b>	<b>3/6</b>
<b>10</b>	<b>5/7</b>	<b>13</b>	<b>4/7</b>	<b>14</b>	<b>4/7</b>

## Ejemplo: Cálculos Precisión Promedio en el Nivel $r$ de Recall - Motor de Búsqueda #1

<b>Consulta #1</b>	<b>Recall Consulta #1</b>	<b>Consulta #2</b>	<b>Recall Consulta #2</b>	<b>Consulta #3</b>	<b>Recall Consulta #3</b>
1	1/5	5	0/5	6	1/5
2	2/5	4	1/5	3	2/5
4	2/5	7	1/5	4	2/5
5	3/5	8	2/5	5	2/5
7	4/5	10	3/5	7	2/5
8	4/5	11	4/5	9	3/5
10	5/5	13	4/5	14	4/5

## Ejemplo: Cálculos Precisión Promedio en el Nivel $r$ de Recall - Motor de Búsqueda #1

Niveles de Recall	Precisión Consulta #1	Precisión Consulta #2	Precisión Consulta #3	Precisión Promedio
0	1	0.5	1	$(1 + 1/2 + 1) / 3 = 0.833$
0.1	1	0.5	1	$(1 + 1/2 + 1) / 3 = 0.833$
0.2	1	0.5	1	$(1 + 1/2 + 1) / 3 = 0.833$
0.3	1	0.5	1	$(1 + 2/4 + 1) / 3 = 0.833$
0.4	1	0.5	1	$(1 + 2/4 + 1) / 3 = 0.833$
0.5	0.75	0.6	0.5	$(3/4 + 3/5 + 3/6) / 3 = 0.617$
0.6	0.75	0.6	0.5	$(3/4 + 3/5 + 3/6) / 3 = 0.617$
0.7	0.8	0.667	0.571	$(4/5 + 4/6 + 4/7) / 3 = 0.679$
0.8	0.8	0.667	0.571	$(4/5 + 4/6 + 4/7) / 3 = 0.679$
0.9	0.714	0	0	$(5/7 + 0 + 0) / 3 = 0.238$
1	0.714	0	0	$(5/7 + 0 + 0) / 3 = 0.238$

## Ejemplo: Cálculos Precisión Promedio en el Nivel $r$ de Recall - Motor de Búsqueda #1

Niveles de Recall	Precisión Promedio
0	0.833
0.1	0.833
0.2	0.833
0.3	0.833
0.4	0.833
0.5	0.617
0.6	0.617
0.7	0.679
0.8	0.679
0.9	0.238
1	0.238

## Ejemplo: Cálculos Precisión Promedio en el Nivel $r$ de Recall - Motor de Búsqueda #2

<b>Consulta #1</b>	<b>Precisión Consulta #1</b>	<b>Consulta #2</b>	<b>Precisión Consulta #2</b>	<b>Consulta #3</b>	<b>Precisión Consulta #3</b>
<b>1</b>	<b>1/1</b>	<b>4</b>	<b>1/1</b>	<b>3</b>	<b>1/1</b>
<b>2</b>	<b>2/2</b>	<b>8</b>	<b>2/2</b>	<b>2</b>	<b>1/2</b>
<b>3</b>	<b>2/3</b>	<b>3</b>	<b>1/3</b>	<b>4</b>	<b>1/3</b>
<b>4</b>	<b>2/4</b>	<b>2</b>	<b>1/4</b>	<b>6</b>	<b>2/4</b>
<b>7</b>	<b>3/5</b>	<b>6</b>	<b>2/5</b>	<b>9</b>	<b>3/5</b>
<b>9</b>	<b>4/6</b>	<b>15</b>	<b>3/6</b>	<b>8</b>	<b>3/6</b>
<b>10</b>	<b>4/7</b>	<b>11</b>	<b>4/7</b>	<b>14</b>	<b>4/7</b>

## Ejemplo: Cálculos Precisión Promedio en el Nivel $r$ de Recall - Motor de Búsqueda #2

<b>Consulta #1</b>	<b>Recall Consulta #1</b>	<b>Consulta #2</b>	<b>Recall Consulta #2</b>	<b>Consulta #3</b>	<b>Recall Consulta #3</b>
<b>1</b>	<b>1/5</b>	<b>4</b>	<b>1/5</b>	<b>3</b>	<b>1/5</b>
<b>2</b>	<b>2/5</b>	<b>8</b>	<b>2/5</b>	<b>2</b>	<b>1/5</b>
<b>3</b>	<b>2/5</b>	<b>3</b>	<b>2/5</b>	<b>4</b>	<b>1/5</b>
<b>4</b>	<b>2/5</b>	<b>2</b>	<b>2/5</b>	<b>6</b>	<b>2/5</b>
<b>7</b>	<b>3/5</b>	<b>6</b>	<b>2/5</b>	<b>9</b>	<b>3/5</b>
<b>9</b>	<b>3/5</b>	<b>15</b>	<b>3/5</b>	<b>8</b>	<b>3/5</b>
<b>10</b>	<b>4/5</b>	<b>11</b>	<b>4/5</b>	<b>14</b>	<b>4/5</b>

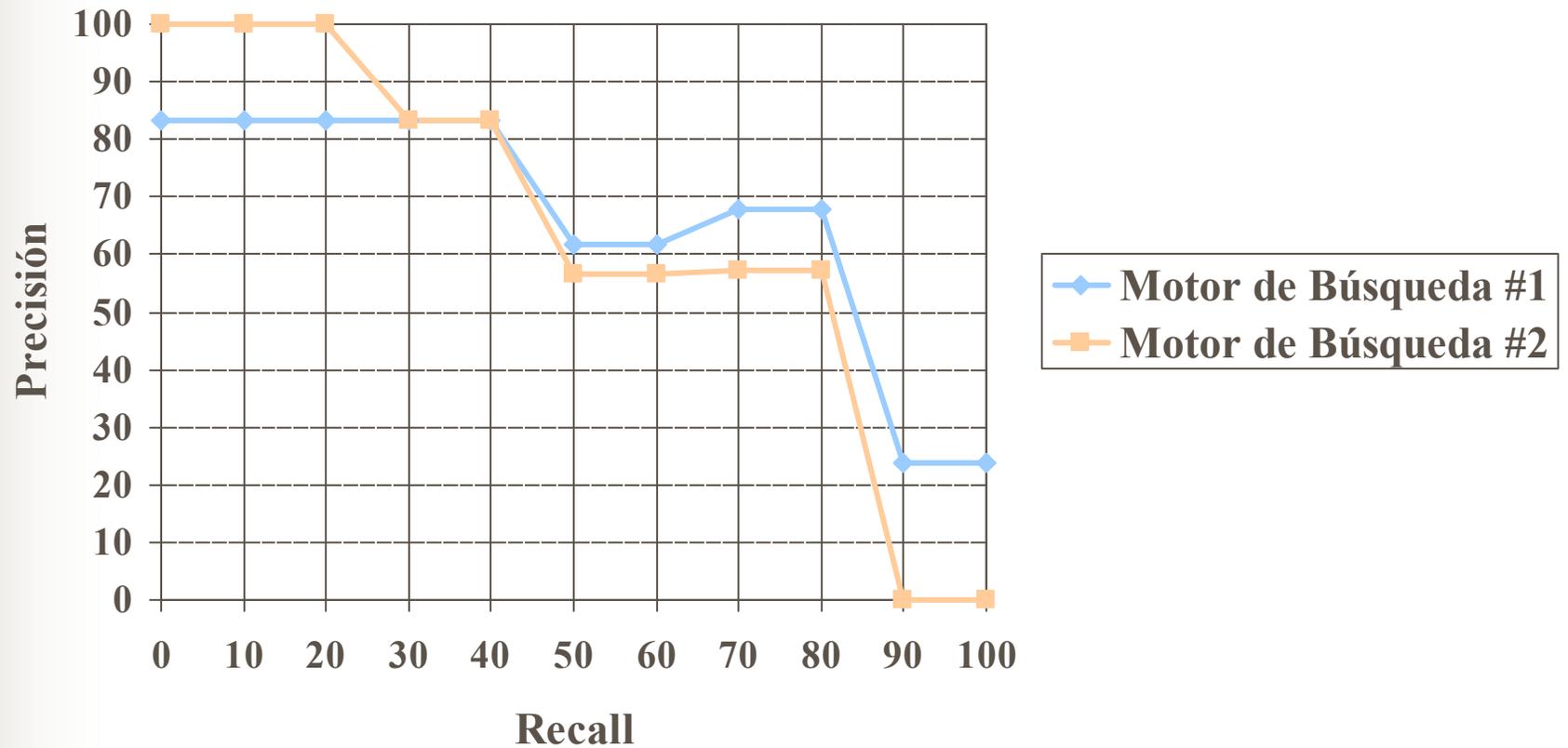
## Ejemplo: Cálculos Precisión Promedio en el Nivel $r$ de Recall - Motor de Búsqueda #2

Niveles de Recall	Precisión Consulta #1	Precisión Consulta #2	Precisión Consulta #3	Precisión Promedio
0	1	1	1	$(1 + 1 + 1) / 3 = 1$
0.1	1	1	1	$(1 + 1 + 1) / 3 = 1$
0.2	1	1	1	$(1 + 1 + 1) / 3 = 1$
0.3	1	1	0.5	$(1 + 1 + 2/4) / 3 = 0.833$
0.4	1	1	0.5	$(1 + 1 + 2/4) / 3 = 0.833$
0.5	0.6	3/6	3/5	$(3/5 + 3/6 + 3/5) / 3 = 0.567$
0.6	0.6	3/6	3/5	$(3/5 + 3/6 + 3/5) / 3 = 0.567$
0.7	0.571	0.571	0.571	$(4/7 + 4/7 + 4/7) / 3 = 0.571$
0.8	0.571	0.571	0.571	$(4/7 + 4/7 + 4/7) / 3 = 0.571$
0.9	0	0	0	$(0 + 0 + 0) / 3 = 0$
1	0	0	0	$(0 + 0 + 0) / 3 = 0$

## Ejemplo: Cálculos Precisión Promedio en el Nivel $r$ de Recall - Motor de Búsqueda #2

Niveles de Recall	Precisión Promedio
0	1
0.1	1
0.2	1
0.3	0.833
0.4	0.833
0.5	0.567
0.6	0.567
0.7	0.571
0.8	0.571
0.9	0
1	0

## Ejemplo (cont.): Gráfico Precisión Promedio en 11 puntos de Recall



## Otras Medidas – Precisión Promedio de Documentos Relevantes Recuperados $\bar{P}_{Ra} = \frac{\sum_{i=1}^{Ra} P_{Ra_i}}{Ra}$

- Obtiene un valor medio según se van considerando los documentos relevantes, para el total de documentos vistos.
  - $P_{Ra_i}$  = Precisión de cada documento relevante recuperado.
  - $Ra$  = Número de documentos relevantes recuperados.
- Esta medida favorece aquellos sistemas que obtienen antes los documentos relevantes.
- Un valor alto puede enmascarar valores muy bajos de *recall*, si se recuperan pocos documentos relevantes en buenas posiciones, el valor de esta medida sería alto, pero el *recall* sería pequeño.

# Ejemplo: Precisión Promedio de Documentos Relevantes Recuperados

## ■ Ejemplo:

- $Rq = \{d_1, d_2, d_5, d_7, d_{10}\}$
- Documentos recuperados del Motor de Búsqueda #1:  
 $\{d_1, d_2, d_4, d_5, d_7, d_8, d_{10}, d_{12}, d_{13}, d_{14}\}$
- Documentos recuperados del Motor de Búsqueda #2:  
 $\{d_1, d_4, d_3, d_2, d_8, d_9, d_{10}, d_5, d_7, d_{14}\}$

$$\bar{P}_{Ra_{qMB1}} = \frac{\sum_{i=1}^{Ra} P_{Ra_i}}{Ra} = \frac{(1/1 + 2/2 + 3/4 + 4/5 + 5/7)}{5} = 0.853$$

$$\bar{P}_{Ra_{qMB2}} = \frac{\sum_{i=1}^{Ra} P_{Ra_i}}{Ra} = \frac{(1/1 + 2/4 + 3/7 + 4/8 + 5/9)}{5} = 0.597$$

## Otras Medidas – Precisión-R

$$\text{Precisión} - R = \frac{Ra_R}{A_R}$$

- Obtiene la precisión para una consulta una vez vistos una cantidad de documentos igual al número de documentos relevantes para esa consulta.
  - $Ra_R$  = Número de documentos relevantes entre los  $A_R$  recuperados.
  - $A_R$  = Número de los primeros  $R$  documentos recuperados.
  - $R$  = Número de documentos relevantes para la consulta.
- Esta medida individual permite observar el comportamiento del algoritmo de recuperación para todas las consultas de la colección de prueba.
- También se puede obtener una Precisión-R promedio para toda la colección de consultas, pero puede ser una media poco caracterizadora del sistema.

## Ejemplo: Precisión-R

- $Rq = \{d_1, d_2, d_5, d_7, d_{10}\}$
- Documentos recuperados del Motor de Búsqueda #1:  
 $\{d_1, d_2, d_4, d_5, d_7, d_8, d_{10}, d_{12}, d_{13}, d_{14}\}$
- Documentos recuperados del Motor de Búsqueda #2:  
 $\{d_1, d_4, d_3, d_2, d_8, d_9, d_{10}, d_5, d_7, d_{14}\}$

$$\text{Precisión} - R_{qMB1} = \frac{Ra_R}{A_R} = \frac{4}{5} = 0.8$$

$$\text{Precisión} - R_{qMB2} = \frac{Ra_R}{A_R} = \frac{2}{5} = 0.4$$

## Otras Medidas – Histograma de Precisión

- Obtiene una comparación entre dos algoritmos de recuperación utilizando la diferencia entre los valores correspondientes de la Precisión-R para cada una de las consultas en ambos algoritmos.
- Para cada consulta,  $i$ , se tiene un valor:

$$RP_{A/B}(i) = RP_A(i) - RP_B(i)$$

- $A$  y  $B$  son algoritmos de recuperación.
- $RP_A(i)$  = Precisión-R del algoritmo de recuperación  $A$  para la consulta  $i$ .
- $RP_B(i)$  = Precisión-R del algoritmo de recuperación  $B$  para la consulta  $i$ .



## Otras Medidas – Histograma de Precisión (cont.)

- Un valor 0 de esta medida indica que ambos algoritmos tienen igual calidad para la consulta  $i$ , en términos de la Precisión-R.
- Un valor positivo de esta medida indica una mejor calidad del algoritmo  $A$  para la consulta  $i$ , en términos de la Precisión-R.
- Un valor negativo de esta medida indica una mejor calidad del algoritmo  $B$  para la consulta  $i$ , en términos de la Precisión-R.

## Ejemplo: Histograma de Precisión

- $Rq_1 = \{d_1, d_2, d_5, d_7, d_{10}\}$

- $Rq_2 = \{d_4, d_8, d_{10}, d_{11}, d_{15}\}$

- $Rq_3 = \{d_3, d_6, d_9, d_{13}, d_{14}\}$

- Documentos recuperados del Motor de Búsqueda #1:

$$Aq_1 = \{d_1, d_2, d_4, d_5, d_7, d_8, d_{14}\} \quad RP_{MB1}(q_1) = 4/5 = 0.8$$

$$Aq_2 = \{d_5, d_4, d_7, d_8, d_{10}, d_{11}, d_{13}\} \quad RP_{MB1}(q_2) = 3/5 = 0.6$$

$$Aq_3 = \{d_6, d_3, d_4, d_5, d_7, d_9, d_{14}\} \quad RP_{MB1}(q_3) = 2/5 = 0.4$$

- Documentos recuperados del Motor de Búsqueda #2:

$$Aq_1 = \{d_1, d_2, d_3, d_4, d_7, d_9, d_{10}\} \quad RP_{MB2}(q_1) = 3/5 = 0.6$$

$$Aq_2 = \{d_4, d_8, d_3, d_2, d_6, d_{15}, d_{11}\} \quad RP_{MB2}(q_2) = 2/5 = 0.4$$

$$Aq_3 = \{d_3, d_2, d_4, d_6, d_9, d_8, d_{14}\} \quad RP_{MB2}(q_3) = 3/5 = 0.6$$

## Ejemplo (cont.): Histograma de Precisión

- Motor de Búsqueda #1:

$$RP_{MB1}(q_1) = 4/5 = 0.8$$

$$RP_{MB1}(q_2) = 3/5 = 0.6$$

$$RP_{MB1}(q_3) = 2/5 = 0.4$$

- Motor de Búsqueda #2:

$$RP_{MB2}(q_1) = 3/5 = 0.6$$

$$RP_{MB2}(q_2) = 2/5 = 0.4$$

$$RP_{MB2}(q_3) = 3/5 = 0.6$$

$$RP_{MB1/MB2}(q_1) = RP_{MB1}(q_1) - RP_{MB2}(q_1)$$

$$RP_{MB1/MB2}(q_1) = 0.8 - 0.6 = 0.2$$

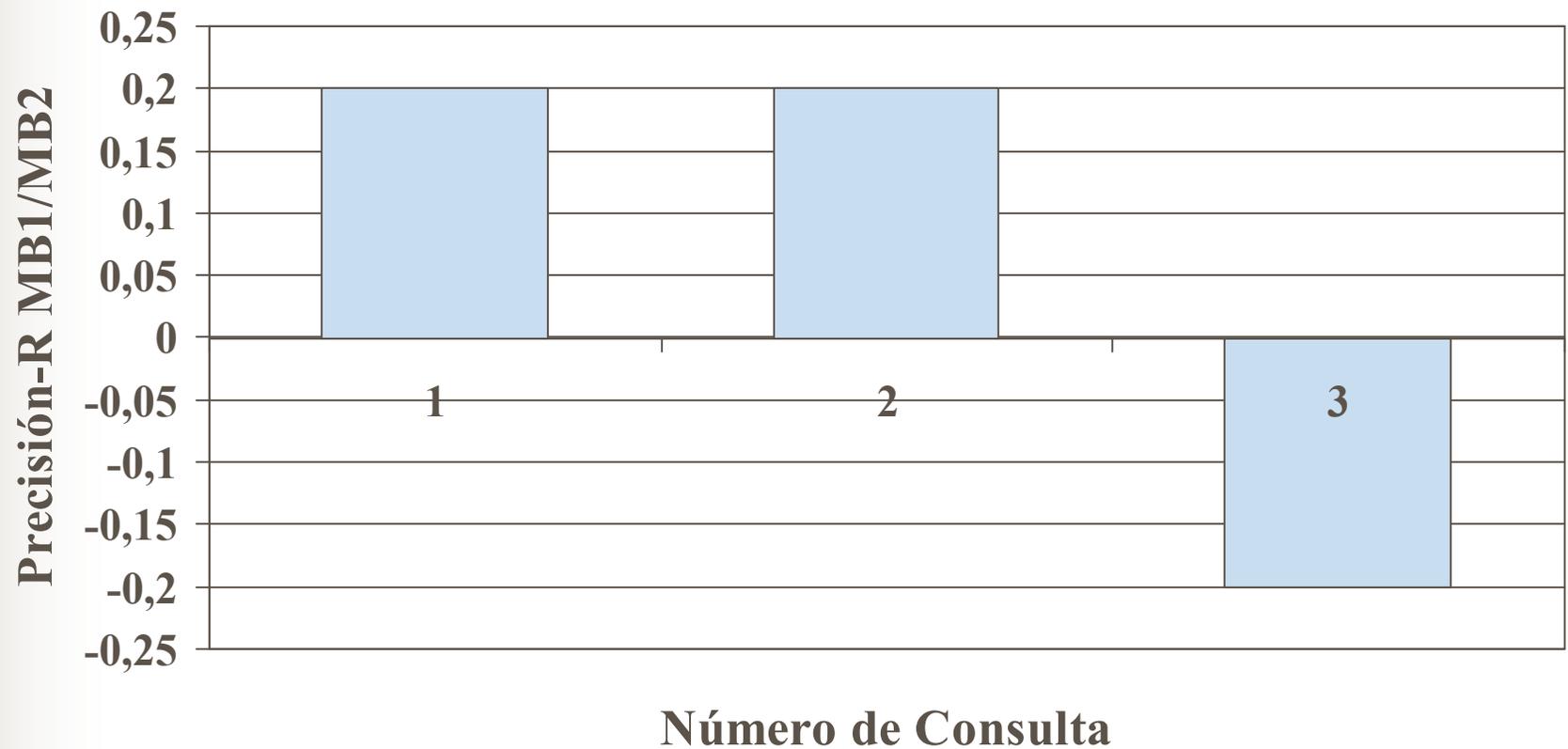
$$RP_{MB1/MB2}(q_2) = RP_{MB1}(q_2) - RP_{MB2}(q_2)$$

$$RP_{MB1/MB2}(q_2) = 0.6 - 0.4 = 0.2$$

$$RP_{MB1/MB2}(q_3) = RP_{MB1}(q_3) - RP_{MB2}(q_3)$$

$$RP_{MB1/MB2}(q_3) = 0.4 - 0.6 = -0.2$$

## Ejemplo (cont.): Histograma de Precisión





## Otras Medidas – Tablas Estadísticas Resumen

- Son tablas en las que aparecen los resultados, primero individuales y al final la media, para todas las consultas del experimento.
- Es habitual encontrarse datos como el número de consultas usadas, el número de documentos relevantes para cada consulta, el número de documentos relevantes recuperados para cada consulta, el número de documentos recuperados para cada consulta, los valores interpolados de precisión, Precisión-R, etc.



## TREC (*Text REtrieval Conference*)

- Las colecciones del TREC están compuestas por 3 partes:
  - Los documentos
  - Los ejemplos de preguntas de información (llamados *tópicos* (*topics*))
  - Conjunto de documentos relevantes para cada ejemplo de preguntas de información



## TREC (*Text REtrieval Conference*) (cont.)

- La estructura de los documentos está SGML, mide alrededor de 1 GB, hay 6 discos CD-ROM.
- Los tópicos son un conjunto de consultas que son usadas para testear un nuevo algoritmo de ranking. Cada consulta es una descripción de una necesidad de información en LN.
- Cada conjunto de documentos relevantes es un conjunto de documentos que se obtuvo por el tópico, es obtenido de una piscina de posibles documentos relevantes.



## TREC (*Text REtrieval Conference*) (cont.)

- La piscina de documentos es evaluada por asesores humanos que deciden la relevancia de cada documento de acuerdo al tópico.
- La técnica que utilizan para evaluar relevancia se llama *método pooling* y está basada en dos suposiciones:
  - La inmensa mayoría de los documentos relevantes es coleccionada en la piscina ensamblada.
  - Los documentos que no están en la piscina pueden ser considerados no relevantes.



## TREC (*Text REtrieval Conference*) (cont.)

- Tareas principales de RI:
  - *Tarea ad hoc*: Conjunto de nuevas preguntas que son corridas cada vez más en una BD de documentos corregidos.
  - *Tarea routing*: Conjunto de preguntas corregidas que son corridas en una BD donde los documentos son continuamente cambiados.
- Colecciones:
  - Colección que entrena al algoritmo.
  - Colección que testea el algoritmo.

## TREC (cont.)

### ■ Tipos de evaluación:

- *Summary table statistics*: Tabla que resume estadísticas relativas a la tarea (# tópicos, # documentos recuperados, # documentos relevantes recuperados, # documentos relevantes que pudieron ser recuperados).
- *Recall-precision averages*: Tabla o gráfico con precisión promedio en 11 niveles de *recall* estándar.
- *Document level averages*: Precisión promedio que especifica los documentos.
- *Average precision histogram*: Gráfico (cada tópico separado) que incluye una medida para cada tópico.



## Referencias Bibliográficas

- La información fue tomada de:
  - Libro de texto del curso.
- The Text Retrieval Conference (TREC). URL:  
<http://trec.nist.gov/>.