



Universidade Federal do Rio de Janeiro (UFRJ)
Departamento de Ciência da Computação (DCC)



Recuperação da Informação (MAB605)

Avaliação da Recuperação

Profa. Giseli Rabello Lopes

Roteiro

- Coleções de Teste
- Implementação de avaliação da recuperação
- Referências

Coleções de Teste

- Composta por:
 - Coleção de Documentos
 - Consultas (chamadas de tópicos)
 - Julgamentos de relevância – quais documentos são relevantes para cada consulta
- Exemplos:
 - Cranfield; Time Magazine; CACM; LA Times; Folha de São Paulo; Wall Street Journal

Cranfield Collection

- Primeira coleção de teste desenvolvida no *Cranfield College of Aeronautics* (Inglaterra) no final dos anos 60
 - 1.400 abstracts de documentos sobre aeronáutica
 - 225 consultas
 - Cada documentos julgado por mais de um avaliador – mais de meio milhão de julgamentos de relevância
 - $225 \times 1400 \times 2 = 630.000$ julgamentos de relevância!

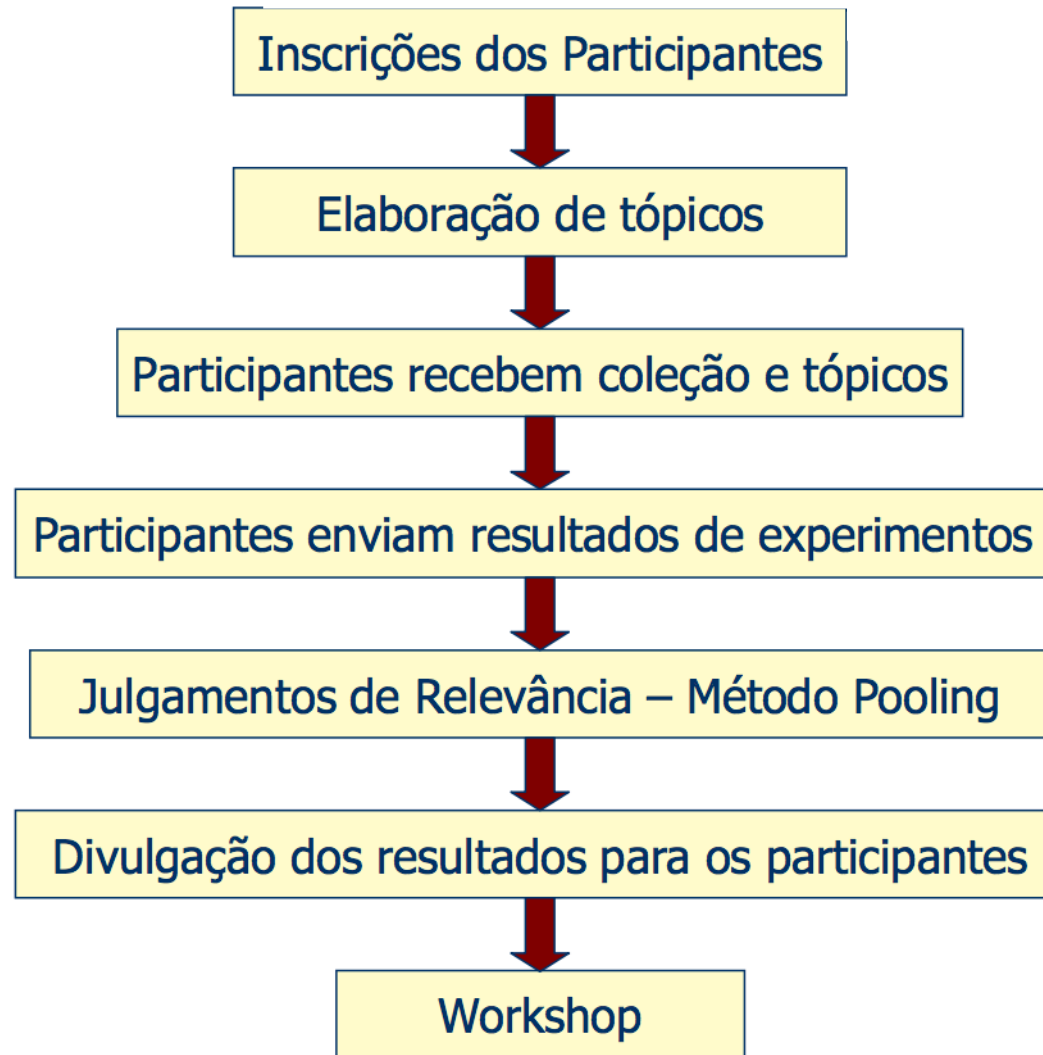
Construindo uma coleção de teste

- Problema: julgar todos os documentos para todas as consultas é impraticável
 - Ex.: LA Times = 50 consultas x 113 mil docs x 2 avaliadores = 11.300.000 julgamentos

Construindo uma coleção de teste

- Solução: **Método *Pooling***
 - Repositórios de documentos são criados para cada consulta a partir dos resultados de experimentos enviados por diferentes grupos
 - Apenas os top n documentos de cada grupo são adicionados ao repositório
 - Os avaliadores julgam somente documentos que estão no repositório
 - Os documentos não julgados são considerados irrelevantes
 - Como o repositório é construído com documentos retornados por vários sistemas, existe a probabilidade de que todos os documentos relevantes sejam encontrados

Campanhas de avaliação de sistemas de RI



Exercício - Avaliação da recuperação

- Sabendo quais são os documentos relevantes para uma consulta (**R**), calcule as métricas de avaliação a seguir para os modelos vetorial e BM25, implementados nas aulas anteriores (considere que seriam recuperados apenas os documentos com similaridades positivas):
 - Revocação e precisão (considerando todos documentos recuperados)
 - Medida F1 (considerando todos documentos recuperados)
 - Revocação e precisão a cada relevante recuperado
 - Precisão interpolada para cada um dos 11 níveis padrão de revocação
 - Precisão média para a consulta (MAP_i)
-

Exercício - Relembrando

- Exemplo de entradas:

```
M=['O peã e o caval são pec de xadrez. O caval é o  
melhor do jog.'];  
'A jog envolv a torr, o peã e o rei.';  
'O peã lac o boi';  
'Caval de rodei!';  
'Polic o jog no xadrez.']; //conjunto de documentos  
stopwords=['a', 'o', 'e', 'é', 'de', 'do', 'no',  
'são']; //lista de stopwords  
q='xadrez peã caval torr'; //consulta  
separadores=[' ',',','.', '!','?']; //separadores  
para tokenizacao  
R=[1; 2]; //identificador dos documentos relevantes  
para a consulta q
```



Universidade Federal do Rio de Janeiro (UFRJ)
Departamento de Ciência da Computação (DCC)



Recuperação da Informação (MAB605)

Dúvidas?

Profa. Giseli Rabello Lopes
giseli@dcc.ufrj.br
CCMN - DCC - Sala E-2012

