

# Examen CC52D 2004/01

Ricardo Baeza-Yates - Carlos Castillo

**Tiempo: 2.00 horas, con apuntes.**

## Pregunta 1

- i. Indique cuáles son los problemas con la noción de distancia entre consultas usando modelo vectorial, y en qué aspectos es mejor la noción de distancia entre consultas presentada (charla de C. Hurtado).
- ii. Indique cuáles son los beneficios de disminuir la dimensionalidad del espacio de documentos (charla de G. Dupret).

## Pregunta 2

Explique tres formas distintas de usar las consultas a un buscador para mejorar el buscador mismo.

## Pregunta 3

Indique cuáles de las técnicas de visualización vistas en clase le parecen más relevantes para el caso de un sistema de búsqueda de e-mails personales. Dibuje/diagrame y explique ejemplos de usos concretos para éste caso.

## Pregunta 4

Muestre una arquitectura de software para un sistema que permite buscar noticias publicadas en sitios Web de diarios chilenos, explicando el objetivo y cuáles son los problemas de diseño más importantes para cada componente.