



# Recuperación de Información Multimedia

**Juan Manuel Barrios**

[juan.barrios@impresee.com](mailto:juan.barrios@impresee.com)

<http://juan.cl/mir>

## Introducción

Septiembre 2018

# Curso CC5213

- **Recuperación de Información Multimedia**
  - **Multimedia Information Retrieval**
- 6 Créditos
- Martes y Jueves 16:15 – 17:45
- Requisito (basta uno):
  - CC5206 – Introducción a la Minería de Datos
  - EL4106 – Inteligencia Computacional
  - CC4102 – Diseño y Análisis de algoritmos
- Equivalente con CC5204

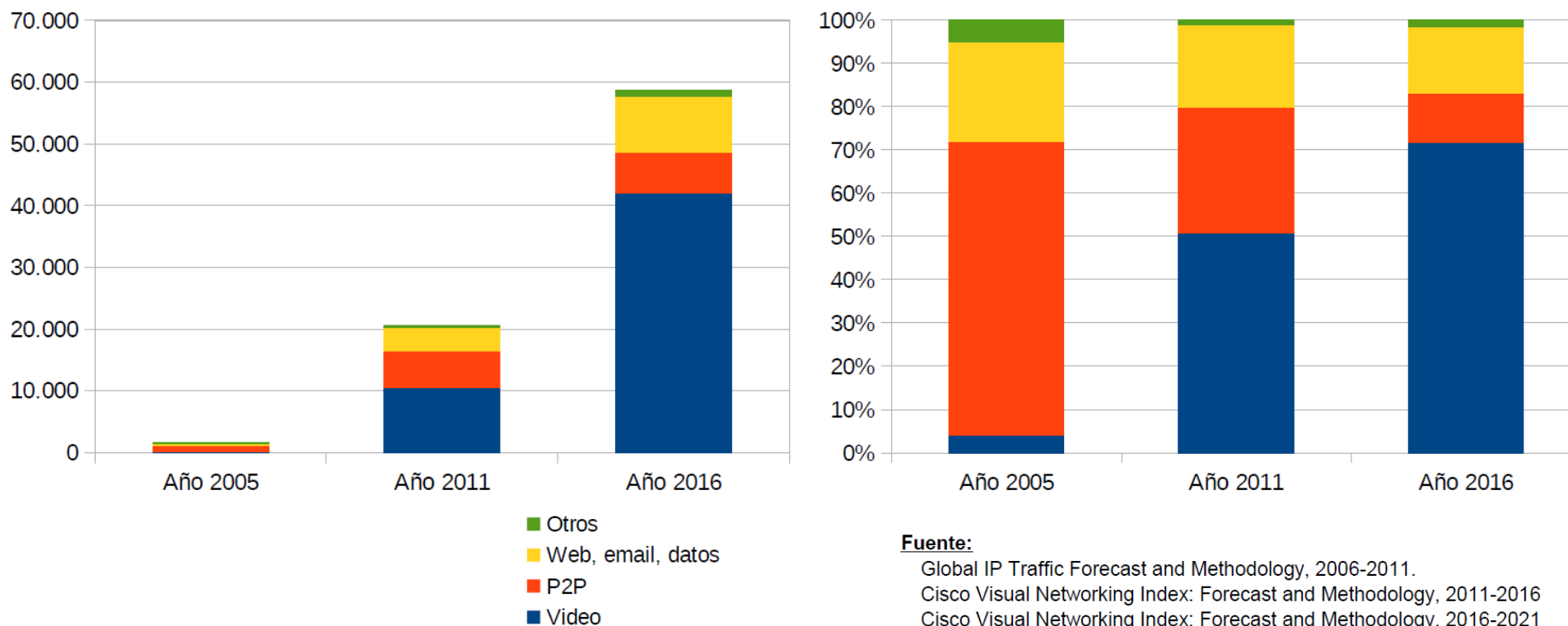
# Recuperación de Información

- **Information Retrieval (IR)** estudia cómo representar, organizar, almacenar y acceder a **información** existente en documentos
  - Objetivo: Recuperar (*retrieve*) documentos **relevantes** a la necesidad de **información** del usuario
  - Documentos: página web, texto de emails, conjunto de tags, fichas de libros, etc.
  - Consulta: frase, keywords, preguntas

# Evolución del uso de Internet

- 2016: más del 70% del tráfico corresponde a videos

## Tráfico de datos de usuarios de Internet [PB mensual]



# Recuperación de Información Multimedia

## ■ Multimedia Information Retrieval (MIR)

- Buscar documentos **multimedia relevantes** a la necesidad de **información** del usuario
- Documentos: Audio, Imagen, Video, Objetos 3D, etc. Ej: fotos personales, películas, grabación de cámaras de seguridad, música sin ID3, etc.
- La búsqueda **NO requiere metadatos** ni etiquetas, debe analizar el **contenido multimedia** (píxeles, samples de audio, frames de video)
- Consulta: keywords, “by-example” (buscar algo parecido a un documento modelo), “by-sketch” (buscar usando bosquejos)

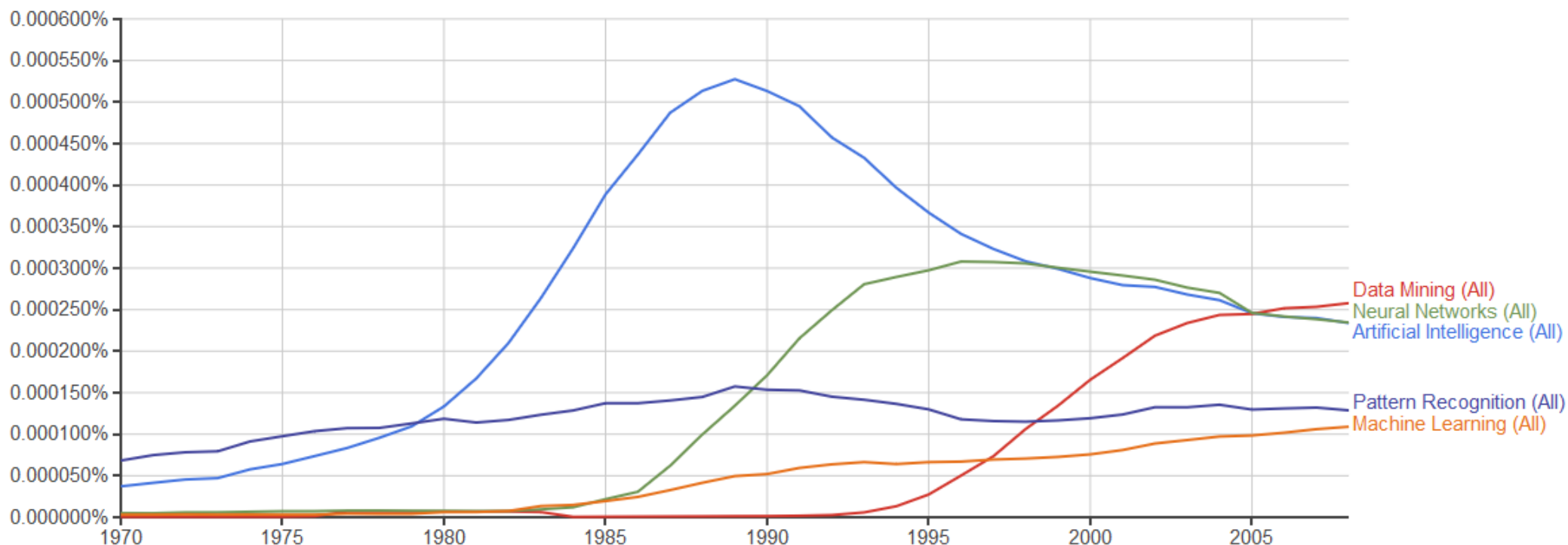


# Áreas Involucradas en MIR

- Information Retrieval
- Artificial Intelligence
- Pattern Recognition
- Machine Learning
- Deep Learning
- Computer Vision
- Data Visualization
- Data Mining
- Data Science

# Uso de Términos (1970-2008)

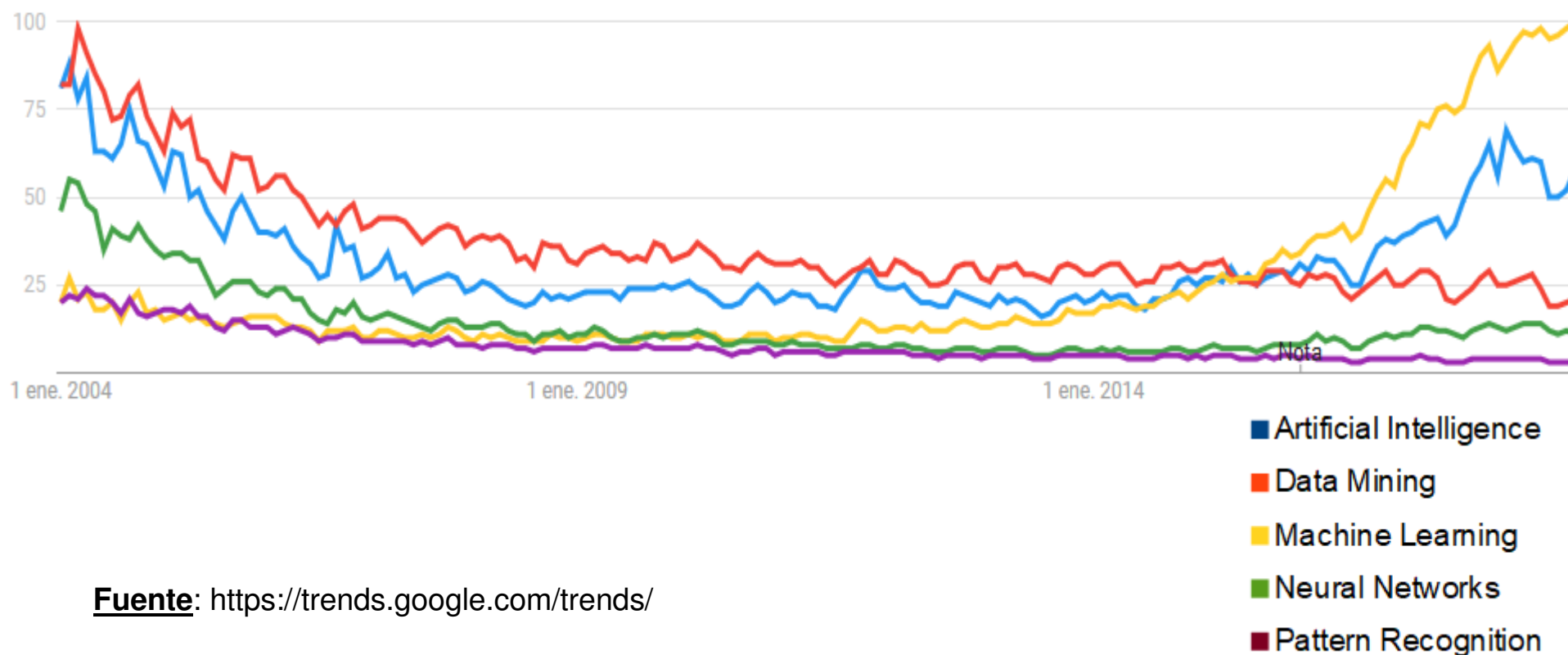
## ■ Google Books (apariciones en libros)



**Fuente:** <https://books.google.com/ngrams>

# Uso de Términos (desde 2004)

## ■ Google Trends (términos en el buscador)



**Fuente:** <https://trends.google.com/trends/>

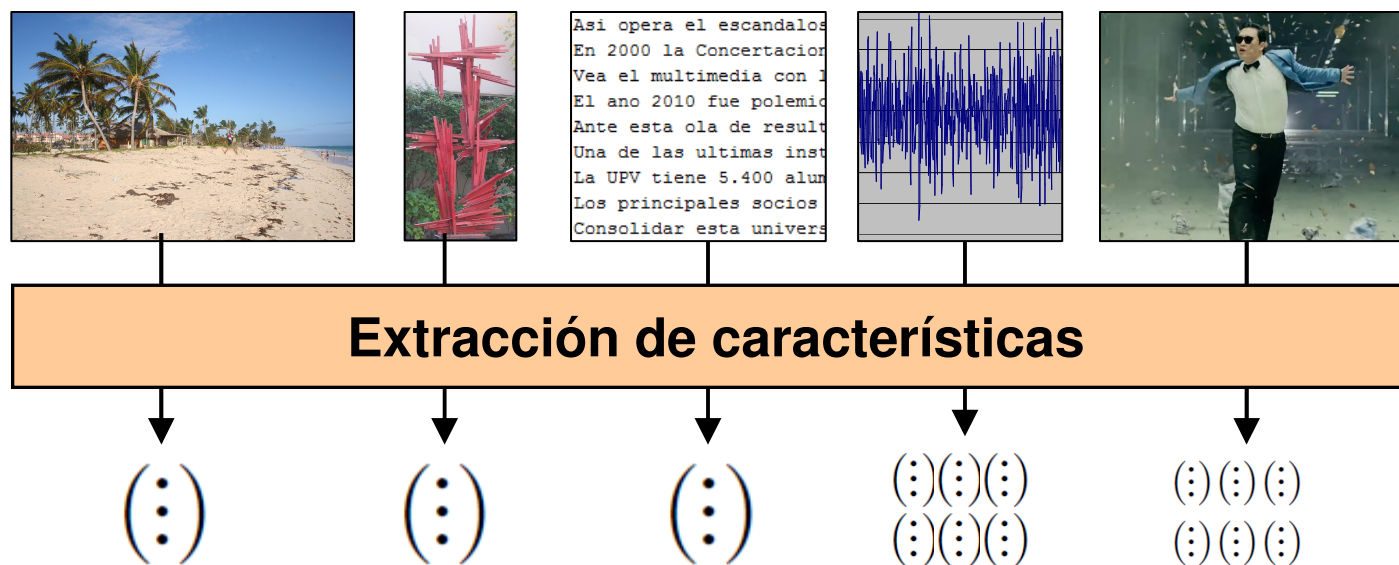


# Temas a estudiar en el semestre

- Parte 1: Descripción de contenido multimedia
  - Métodos para analizar contenido de imágenes, audio, videos y texto
- Parte 2: Búsqueda por similitud e Indexamiento
  - Métodos para resolver eficientemente búsquedas en espacios vectoriales y métricos
- Parte 3: Aplicaciones e Investigación
  - Técnicas actuales de descripción y búsqueda usando Codebooks y Deep Learning

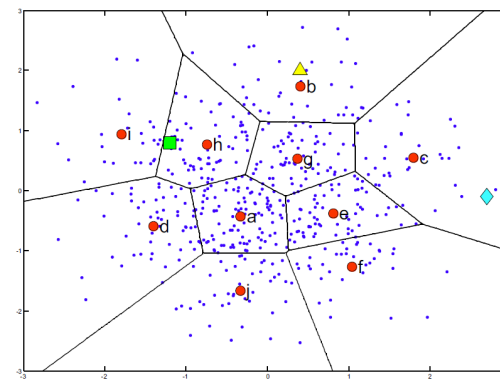
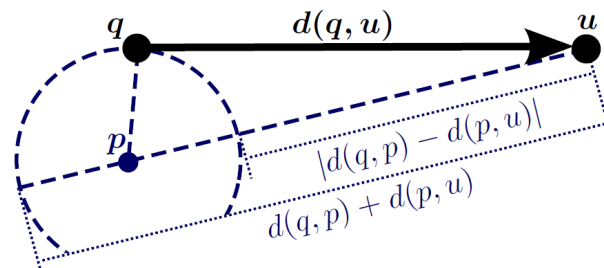
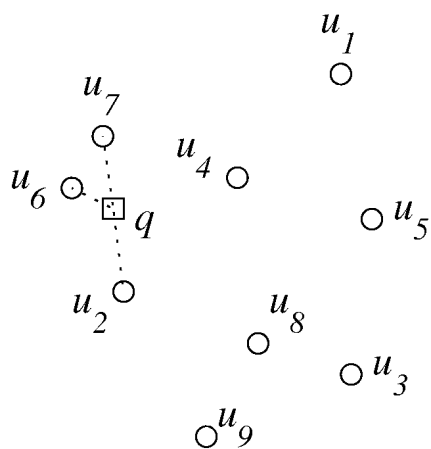
# Descripción del Contenido

- El contenido de cada documento multimedia (foto, canción, video, email, etc.) se representa por uno o más vectores



# Búsqueda por Similitud

- Resolver búsquedas eficientemente entre millones de vectores
- Índices para espacios vectoriales y métricos



# Casos de Estudio

- Durante el semestre se estudiarán métodos para resolver diferentes problemas tipo
- Problema 1: Dada una imagen de consulta, buscar otras imágenes parecidas

Imagen de  
consulta



# Casos de Estudio

- Buscar imágenes parecidas según algún criterio (color, forma, etc.)

Resultados



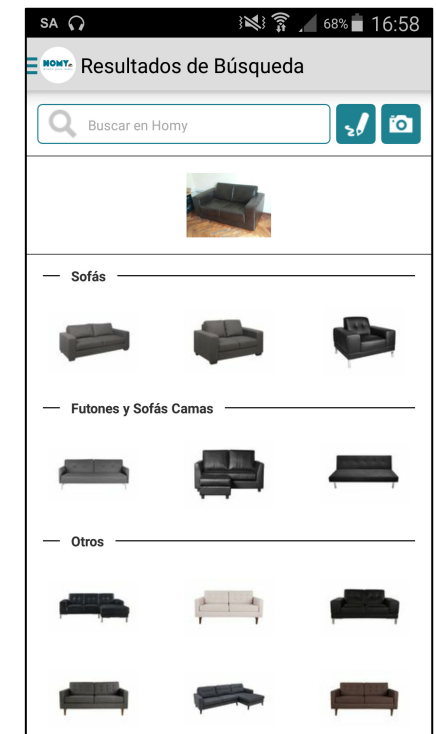
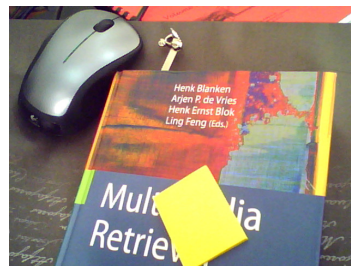
Ej: Dada la imagen de una playa encontrar otras imágenes parecidas.

# Casos de Estudio

- Problema 2: Dado un catálogo de objetos, reconocer el objeto fotografiado



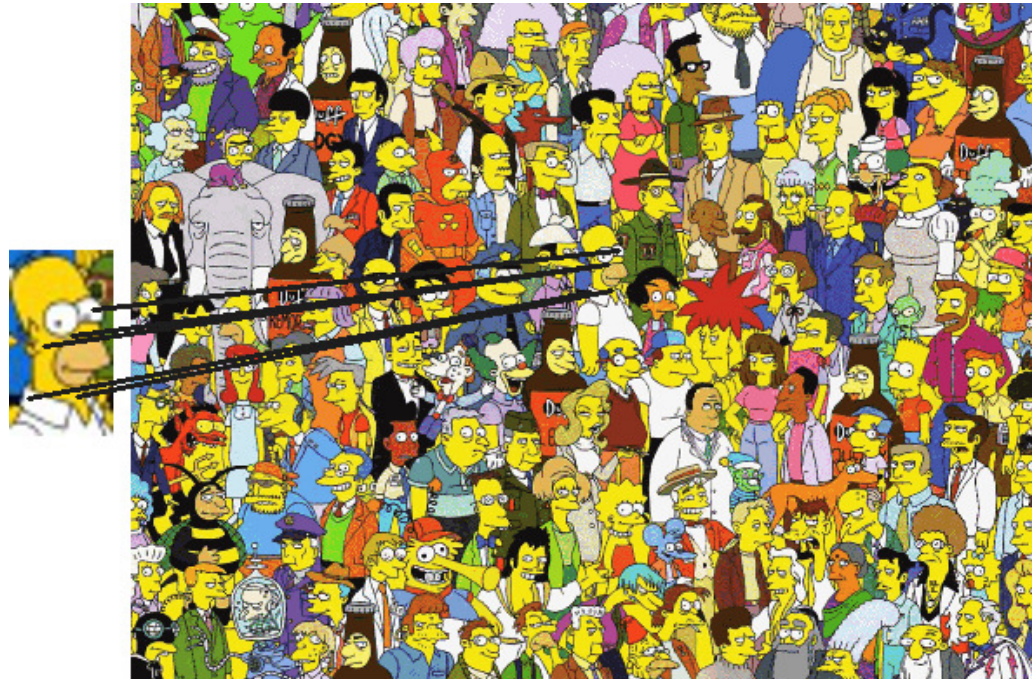
Imagen de consulta



Ver: <http://www.impresee.com/>

# Casos de Estudio

- Además de la ocurrencia, determinar la ubicación y pose del objeto encontrado



# Casos de Estudio

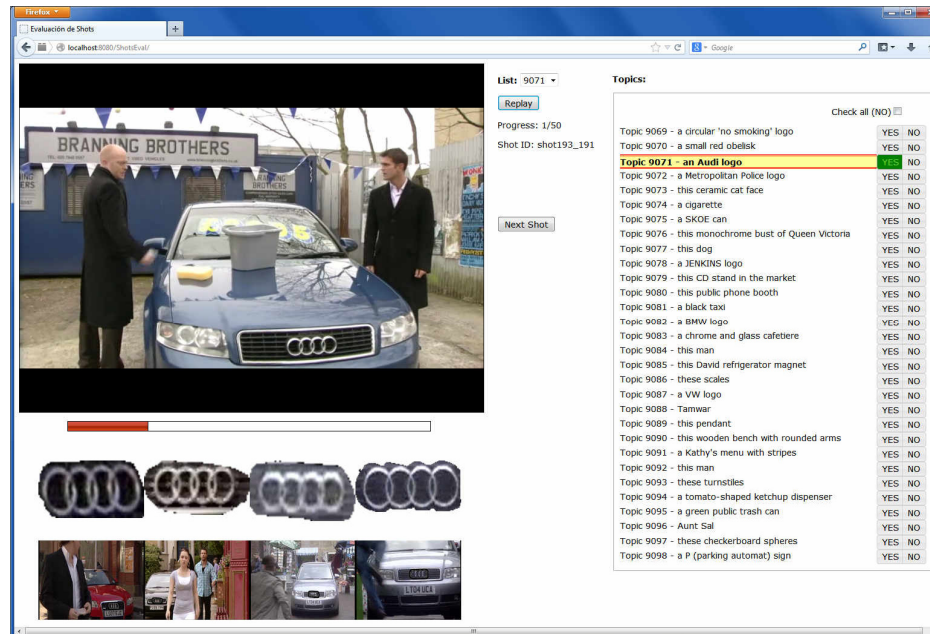
- Problema 3: Dado un video de Internet determinar la escena original de la que proviene.
- Problema 3: Reconocer una canción según un trozo de audio (“Shazam”)





# Casos de Estudio

- Problema 4: Buscar la aparición de algún producto, logo, etc. en televisión



Ej.: Buscar las apariciones del logo Audi

© BBC EastEnders

Ver: <http://www.robots.ox.ac.uk/~vgg/research/vgoogle/>

# Casos de Estudio

- Problema 5: Buscar eventos en videos, ingresando texto o imágenes de ejemplo, sin requerir de metadatos ni etiquetado previo de los datos

Ej: Buscar videos donde aparezca un vehículo estacionando

The screenshot shows a web browser window with the URL 'localhost:8080/EventEval/'. The main content area displays a video player showing a street scene with several cars. To the right of the video player, there is a control panel with 'Replay' and 'Next Video' buttons, and a progress indicator showing 'Progress: 2/30' and 'Video ID: 776704' with a score of '0.98949'. Below this, there is an 'Events:' section with a 'Check All (NO)' checkbox and a list of event categories. The 'Event 37 - Parking a vehicle' is highlighted in yellow and has a green checkmark in the 'NO' column. Below the event list, there is a 'Hide Samples' button and two small thumbnail images: one showing a car parked in a lot and another showing an airplane on a tarmac.

Event	YES	NO
Event 31 - Beekeeping	YES	NO
Event 32 - Wedding shower	YES	NO
Event 33 - Non-motorized vehicle repair	YES	NO
Event 34 - Fixing musical instrument	YES	NO
Event 35 - Horse riding competition	YES	NO
Event 36 - Felling a tree	YES	NO
<b>Event 37 - Parking a vehicle</b>	YES	<b>NO</b>
Event 38 - Plowing fetch	YES	NO
One or more people navigate a vehicle into a designated space.	YES	NO

# Muchas Aplicaciones Novedosas y Problemas Abiertos

Google: Flujo de videos en YouTube hace imposible eliminar todo el contenido terrorista

La compañía justificó que la gran cantidad de videos no se puede detectar toda la propaganda

29 de Enero de 2015 | 09:26 | AP

Google sigue los pasos de Facebook e implementará inteligencia artificial contra contenidos terroristas

La empresa enfocará sus esfuerzos a través de YouTube, un

19 de Junio de 2017 | 16:39 | DPA

Liga inglesa demanda a YouTube por violar derechos de autor

La asociación que maneja el campeonato inglés inició la acción legal en Nueva York, ante los numerosos videos con imágenes de la Premier League que están disponibles en el sitio web.

La aplicación móvil Shazam permitirá a sus usuarios identificar objetos

La herramienta que reconoce la música permite a los usuarios identificar productos en un supermercado

05 de Marzo de 2015 | 16:42 | Reuters

Apple confirma la adquisición de Shazam y aventura "interesantes planes" para el futuro

Las compañías deberán esperar a comenzar a integrar ambos servicios

11 de Diciembre de 2017 | 15:12 | Emol

Facebook reconocerá automáticamente canciones que suenan de fondo

La red social también podrá identificar los programas de televisión que ven sus usuarios mientras escriben mensajes.

Aplicación chilena encuentra productos con solo dibujarlos o sacarles una foto

• Creada por investigadores universitarios, la herramienta analiza imágenes y busca los objetos en



De la idea al prototipo

109:17 | DPA

Spotify amplía su alcance más allá de la música, sumando videos y podcasts a su servicio

La compañía de Daniel Ek estrenó hoy un nuevo modelo que busca ayudar al descubrimiento de contenido en la app. También sumó nuevas funciones para escuchar música al hacer deporte.

20 de Mayo de 2015 | 13:33 | Emol

# Actividades de Evaluación

## 1. **Mini-Controles** (individual)

- Una pregunta escrita al inicio de clases, 15 minutos, sin apuntes

## 2. **Tareas de Programación** (individual)

- Resolver problemas prácticos con Python 3 o C++ 11

## 3. **Presentación de Lecturas** (grupal)

- Exposición de publicaciones científicas (a elección dentro de opciones dadas)

## 4. **Proyecto de Síntesis** (grupal)

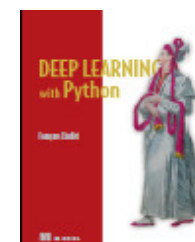
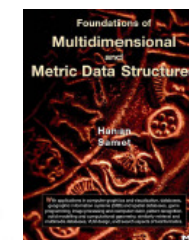
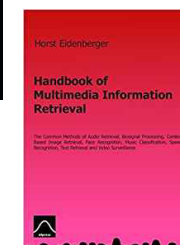
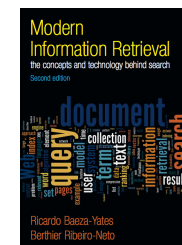
- Resolver algún problema a elección que requiera aplicar los contenidos del curso
- Presentación de resultados al finalizar el semestre

# Evaluación

- Siete mini-controles, cada uno se puede rendir dos veces
  - **NC** =  $(C1 + \dots + C7) / 7$
- Tres tareas de programación
  - **NT** =  $(T1 + T2 + T3) / 3$
- Dos presentaciones de lecturas
  - **NL** =  $(L1 + L2) / 2$
- Proyecto final (T=trabajo realizado, P=presentación final)
  - **NP** =  $(T + P) / 2$
- Para aprobar:  
**NC**  $\geq$  4.0, **NL**  $\geq$  4.0, **NT**  $\geq$  4.0, **NP**  $\geq$  4.0
- Nota Final:  
**(NC + NL + NT + NP) / 4**

# Bibliografía

- **Modern Information Retrieval.** Baeza-Yates, Ribeiro-Neto, 2011.
- **Handbook of Multimedia Information Retrieval.** Eidenberger. 2012.
- **Multimedia Retrieval.** Blanken, de Vries, Blok, Feng. 2007.
- **Foundations of Multidimensional and Metric Data Structures.** Samet. 2006.
- **Deep Learning with Python.** Chollet. 2018.



Sitio web del curso: <http://juan.cl/CC5213-2018b/>  
user=**CC5213** pass=**uchile**

# Bibliografía Extra

- **The Essential Guide to Image/Video Processing.** Bovik. 2009.
- **H.264 and MPEG-4 Video Compression.** Richardson. 2003.
- **Digital Image Processing.** Gonzalez, Woods. 2008.
- **Data Mining: The Textbook.** Aggarwal. 2015.
- **Similarity Search The Metric Space Approach.** Zezula, Amato, Dohnal, Batko. 2006.
- **Computer Vision. Algorithms and Applications.** Szeliski. 2011.
- **Deep Learning: A Practitioner's Approach.** Patterson, Gibson. 2017.

