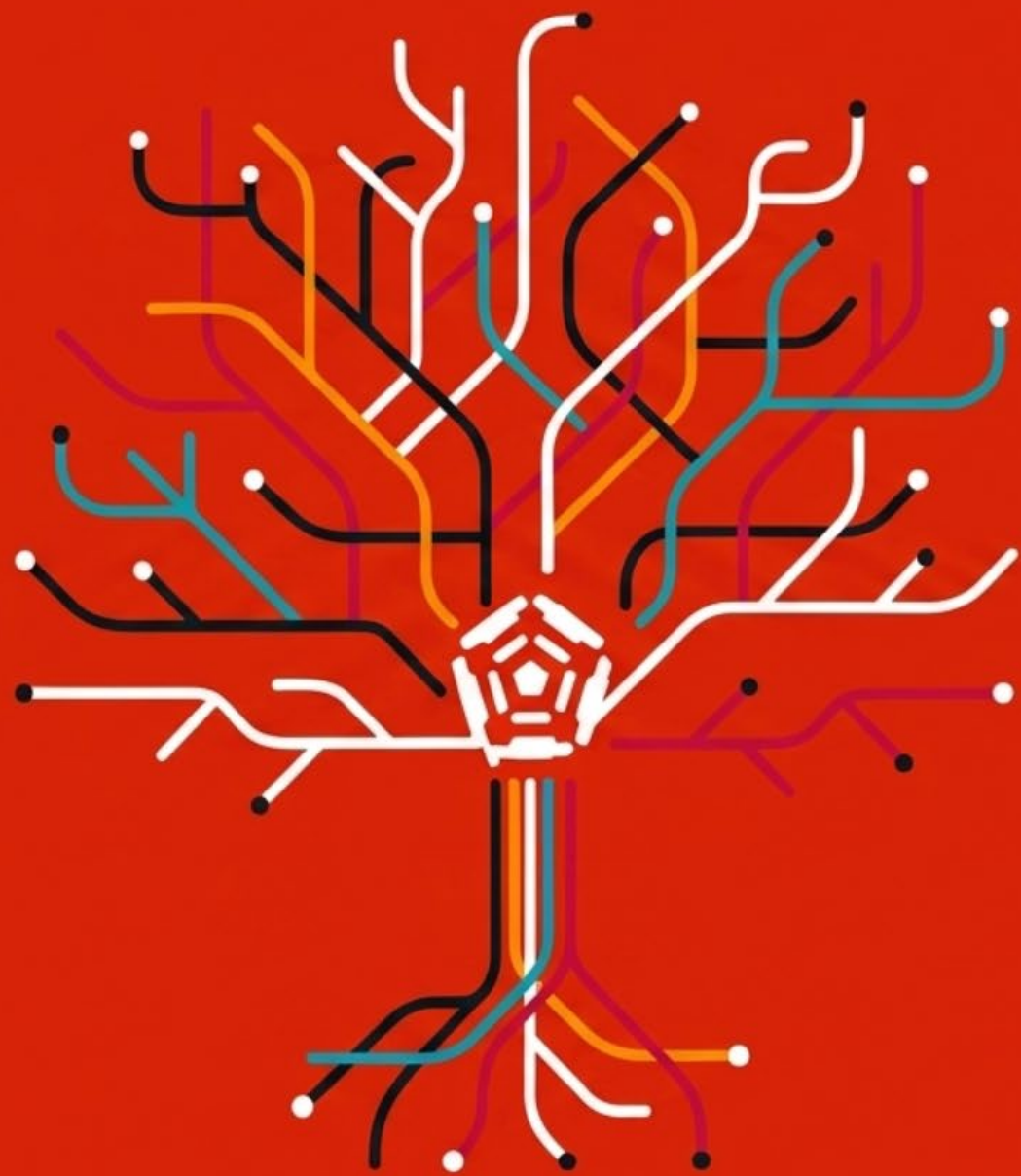
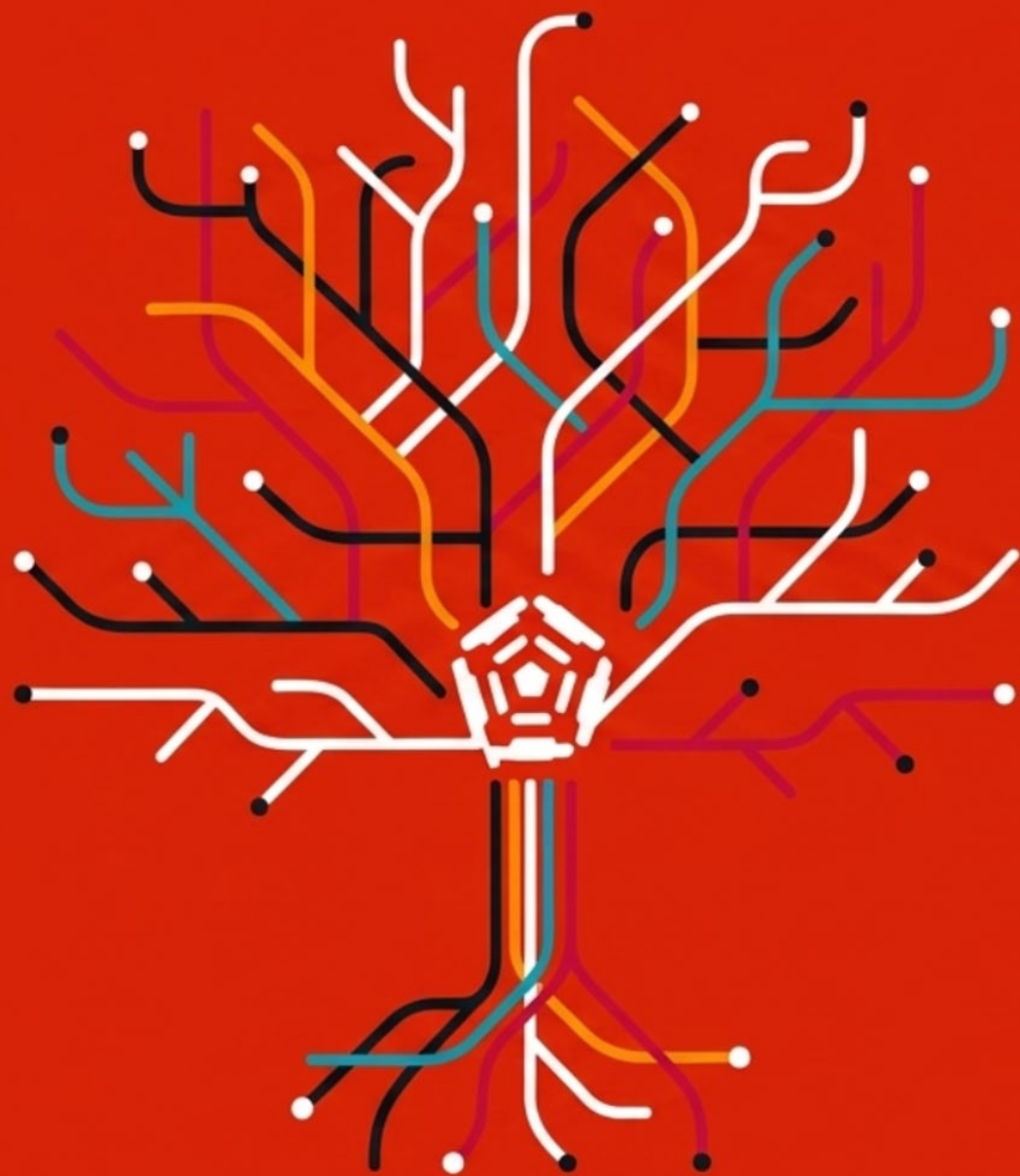




Universidad de Zaragoza | Curso 2026

¿QUÉ TIENE LA INTELIGENCIA ARTIFICIAL PARA MÍ ARCHIVO?

De la teoría a la práctica.



Cursos Extraordinarios
Universidad de Zaragoza 2026

¿Qué tiene la IA para mi archivo?
De la teoría a la práctica

Taller 3: Organización no
supervisada de
colecciones fotográficas y
detección de estructuras
de documentos.

Parte 1

Representación de imágenes y Organización No Supervisada

Descubriendo patrones y estructuras en colecciones de imágenes sin
etiquetas previas

Extracción de Características

Para organizar fotos, primero debemos "traducirlas" a números que la IA entienda.

CLIP: Embeddings Multimodales

Indexación Semántica

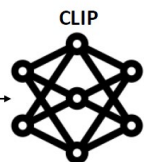
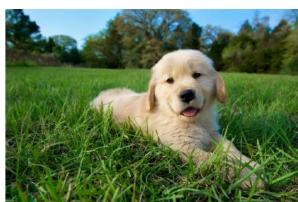
En lugar de descriptores visuales puros (CNNs estándar), empleamos CLIP para alinear el espacio visual y el textual.

- ⚡ **Comprensión Abstracta:** Agrupa por contexto, no solo por similitud de píxeles.
- 🔍 **Búsqueda Cruzada:** Permite realizar búsquedas en lenguaje natural inmediatamente sobre imágenes sin etiquetar.

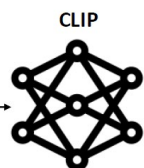
Como representamos las imágenes

Espacios semánticos y multimodalidad: CLIP (Contrastive Language-Image Pre-Training)

Combina un modelo de lenguaje con un modelo semántico de conocimiento de imágenes
Entrenado con más de 400M de pares imagen+texto

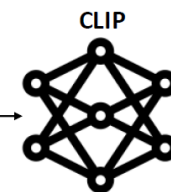


Semantic Vector
 1×512



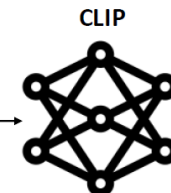
Semantic Vector
 1×512

"A photo of a cute dog on the grass"

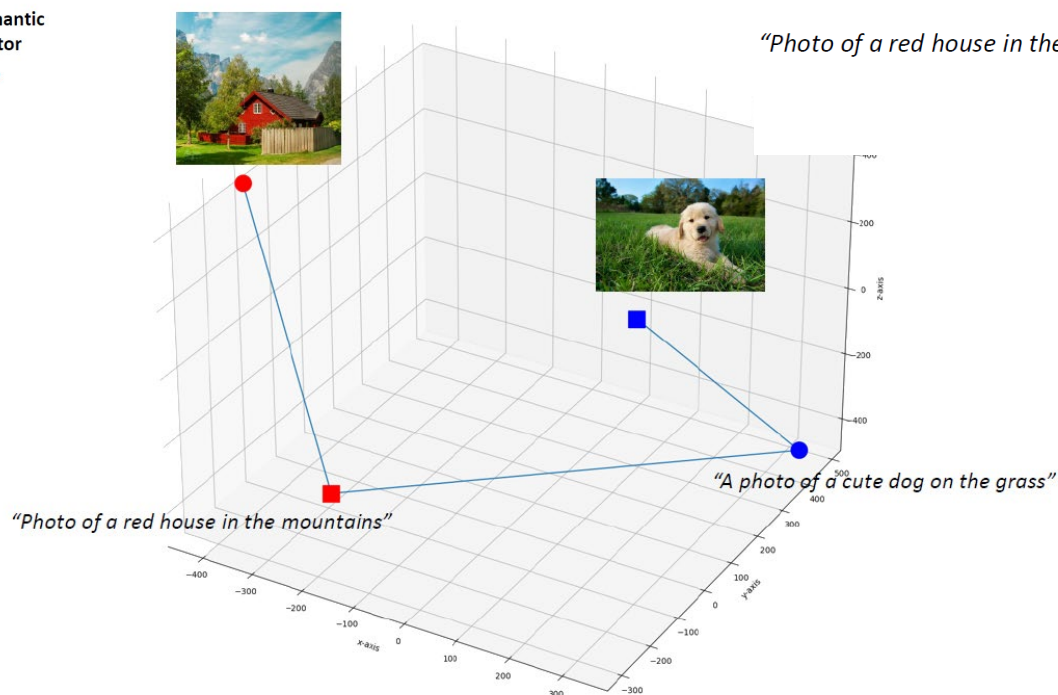


Semantic Vector
 1×512

"Photo of a red house in the mountains"



Semantic Vector
 1×512

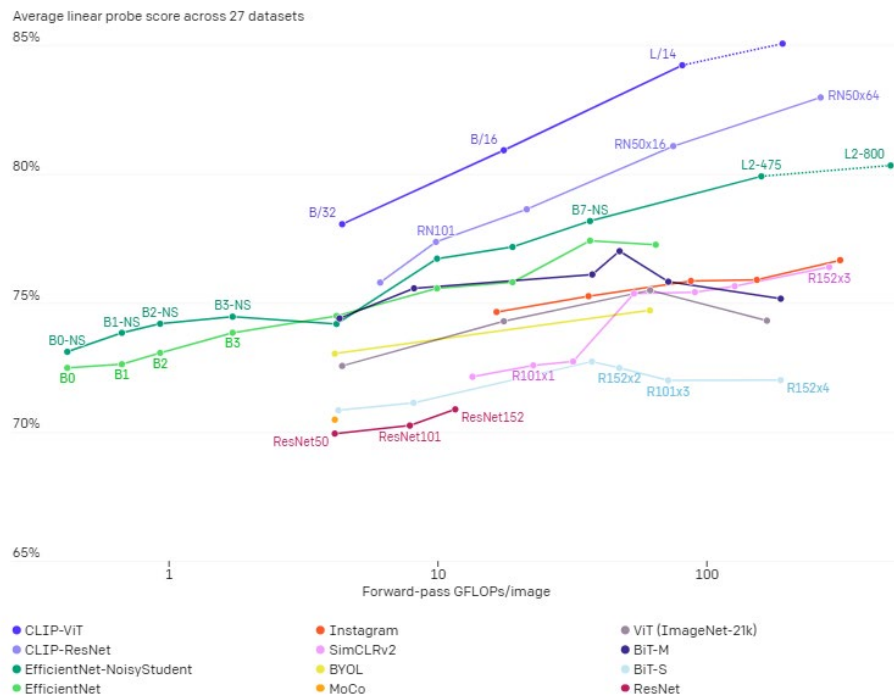


Casos de uso con CLIP

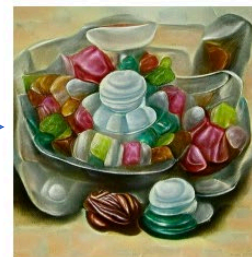
Generación de imágenes:

DALL.E de OpenAI y su sucesor DALL.E VQGAN-CLIP (código abierto) →

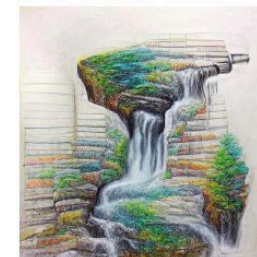
Clasificación de imágenes



Across a suite of 27 datasets measuring tasks such as fine-grained object classification, OCR, activity recognition in videos, and geo-localization, we find that CLIP models learn more widely useful image representations. CLIP models are also more compute efficient than the models from 10 prior approaches that we compare with.



(a) Oil painting of a candy dish of glass candies, mints, and other assorted sweets



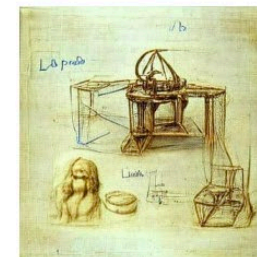
(b) A colored pencil drawing of a waterfall



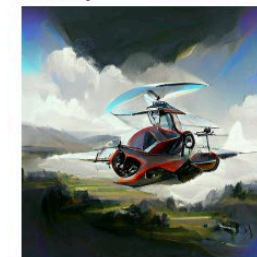
(c) A fantasy painting of a city in a deep valley by Ivan Aivazovsky



(d) A beautiful painting of a building in a serene landscape



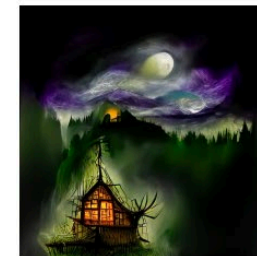
(e) sketch of a 3D printer by Leonardo da Vinci



(f) an autogyro flying car, trending on art-station



(g) an astronaut in the style of van Gogh



(h) Baba Yaga's house + fantasy art

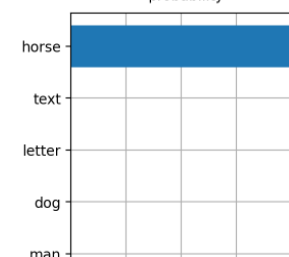
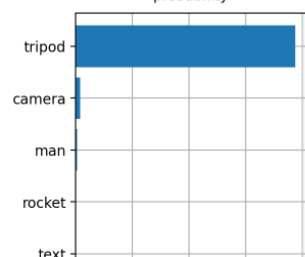
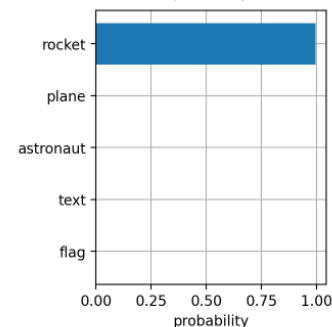
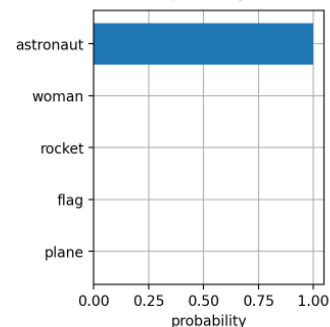
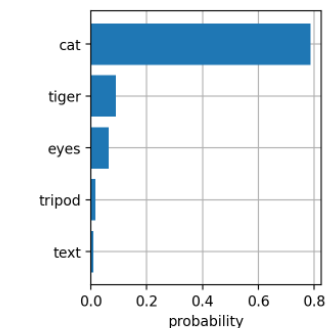
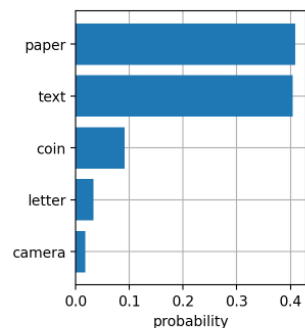
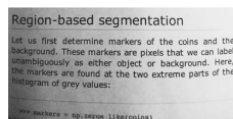


(i) pickled eggs, tempera on wood

CLIP en acción

<https://colab.research.google.com/drive/1O9AkGnR6dHzX-stQRIWDdmEnJv9Fbc4S>

Clasificación “zero-shot”



Geolocalización con CLIP

<https://huggingface.co/geolocal/StreetCLIP>

¿Dónde estoy?

Este es un modelo de clasificación de países basado en CLIP.

Choose an image...



Drag and drop file here

Limit 200MB per file • JPG, PNG, JPEG, AVIF, WEBP

Browse files



images.jpg 13.6KB



Top 10 countries:

	Country	Score
0	Argentina	56.3566
1	Bolivia	33.6038
2	Brazil	3.4332
3	Madagascar	2.5838
4	Uruguay	1.4587
5	Chile	0.6000
6	Peru	0.3102
7	Senegal	0.3075
8	India	0.2060
9	Mexico	0.1846

Top 10 cities from: Argentina

	Admin	Score
Puerto Iguazú	Misiones	68.773
Termas de Río Hondo	Santiago del Estero	2.0647
Gualectuaychú	Entre Ríos	2.0204
Perito Moreno	Santa Cruz	1.7624
Río Colorado	La Pampa	1.4637
Catamarca	Catamarca	1.437
Alto Río Senguer	Chubut	1.2853
Gualectuay	Entre Ríos	1.2193
Río Segundo	Córdoba	0.9734
Tandil	Buenos Aires	0.9054

Entrenado con un dataset de 1,1 millones de imágenes geo etiquetadas urbanas y rurales a nivel de calle a partir del modelo openai/clip-vit-large-patch14-336

Hardware: 4 NVIDIA A100

GPUs

Horas: 12

StreetCLIP fue entrenado con el objetivo de hacer coincidir las imágenes con el título correspondiente a la ciudad, región y país de origen de las imágenes.

<http://signal4.cps.unizar.es:8540/>

Aprendizaje no supervisado

No tenemos clases predeterminadas

- Aprendizaje por observación vs aprendizaje con ejemplos (supervisado)
- Vamos a explorar los datos para buscar en ellos estructuras intrínsecas

Aplicaciones del aprendizaje no supervisado

- Segmentación de clientes: Identificar diferentes grupos de clientes con comportamientos similares.
- Recomendaciones personalizadas: Sugerir productos o servicios basados en las similitudes de los clientes con otros.
- Detección de anomalías: Identificar datos inusuales que no se ajustan a patrones esperados.
- Análisis de textos: Agrupar documentos similares y descubrir temas subyacentes.

Aprendizaje supervisado vs No supervisado

Supervisado:

Datos: (x, y)

x es un dato, y es una etiqueta

Objetivo:

Aprender una función que mapea el dato x a la etiqueta y

Ejemplos:

Clasificación, regresión, detección de objetos, segmentación semántica, descripción de imágenes, descripción de imágenes, etc.

¿Cuánto cuesta etiquetar 1M de imágenes?

(Base de datos pequeña a mediana)
1.000.000 (imágenes)

× (10 segundos/imagen) (anotación rápida)

× (1/3600 horas/segundo)

× (15€ / hora) (salario bajo)

= **41.667 €**

No supervisado:

Datos: x

x es un dato, no hay etiquetas

Objetivo:

Aprender algunas estructuras ocultas subyacentes de los datos.

Ejemplos:

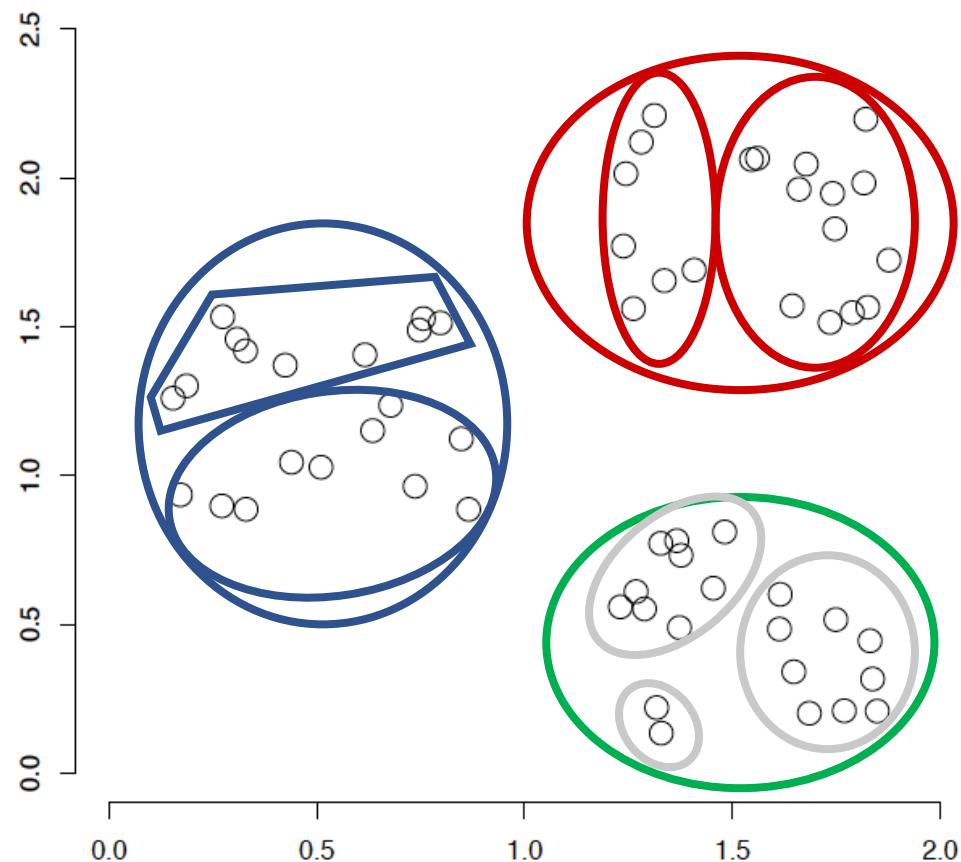
Agrupamientos, Reducción de la dimensionalidad, aprender características, estimar densidades, etc.

Aprendizaje no supervisado

Métodos básicos de aprendizaje no supervisado

Métodos de agrupamiento (Clustering)

Para los datos de la figura, ¿Cuántas agrupaciones/clusters podemos definir?



Técnicas de Clustering

Una vez extraídos los embeddings, aplicamos algoritmos de agrupación:

K-Means: Partición eficiente para grupos definidos.

HDBSCAN: Detección de grupos basados en densidad y ruido.

UMAP / t-SNE: Reducción de dimensionalidad para visualización de la colección.

Aprendizaje no supervisado

Métodos básicos de aprendizaje no supervisado

Métodos de agrupamiento (Clustering)

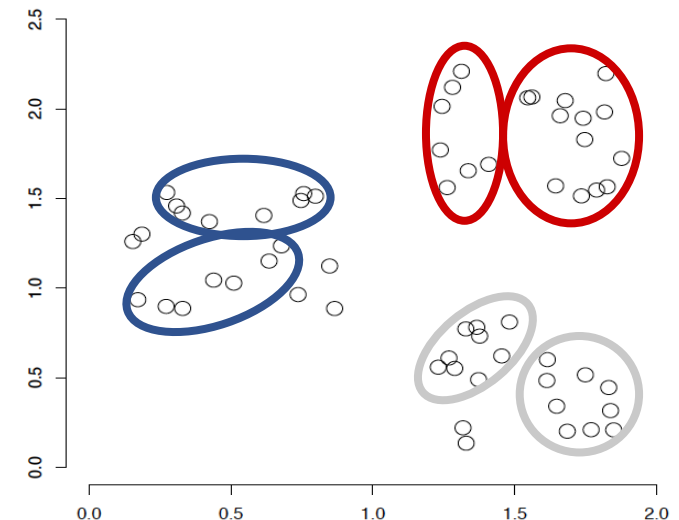
- **Basados en la densidad:**

- ✓ **DBSCAN (Density-Based Spatial Clustering of Applications with Noise):**

- Agrupa puntos densamente conectados.
- Identifican clusters como áreas densamente pobladas de puntos de datos, separadas por áreas de menor densidad.
- Son especialmente efectivos para detectar clusters de forma arbitraria.
- Adecuados para manejar ruido y outliers.
- Puede definirse un enfoque jerárquico:

HDBSCAN

clusters con diferentes densidades.



Organización no supervisada de colecciones fotográficas



Multimodal Search, Clustering & Visualization

Search Configuration | Collection: Periodicos

Search Query

Search by Text:

image of Madonna

OR

Search by Image:

Drag & Drop or Select Image

CLEAR
QUERY

START
SEARCH

SEARCH FOR
DUPLICATES

CLUSTER &
VISUALIZE

Detailed Results

Duplicate Results

UMAP Visualization

Network Visualization

Cluster Details

Cluster Representatives

Found 100 items



#0 Score: 0.308



#1 Score: 0.288



#2 Score: 0.284



#3 Score: 0.281



#4 Score: 0.279



#5 Score: 0.277



#6 Score: 0.277



#7 Score: 0.276



#8 Score: 0.276



#9 Score: 0.275



#10 Score: 0.275



#11 Score: 0.274



#12 Score: 0.274



#13 Score: 0.274



#14 Score: 0.274



#15 Score: 0.273



#16 Score: 0.272



#17 Score: 0.272

Búsqueda de segmentos inéditos en vídeos

Analisis de Archivo Lorca / RTVE

Objetivo

- Detectar que tramos de vídeo son **similares** (repetidos o ya emitidos)
- Detectar que tramos son **inéditos** dentro del conjunto analizado

Contexto de trabajo

- Colección de vídeos históricos sobre Federico Garcia Lorca y La Barraca
- Analisis automático para acelerar catalogación y revisado editorial

Como funciona el pipeline

1. Extracción de frames

- Se extraen frames con ffmpeg a 3 fps

2. Embeddings visuales

- Cada frame se codifica con modelo (qwen3VL2B, clip, etc.)

3. Similitud entre videos

- Se calcula matriz de similitud coseno $N \times M$ entre cada par de videos

4. Segmentacion

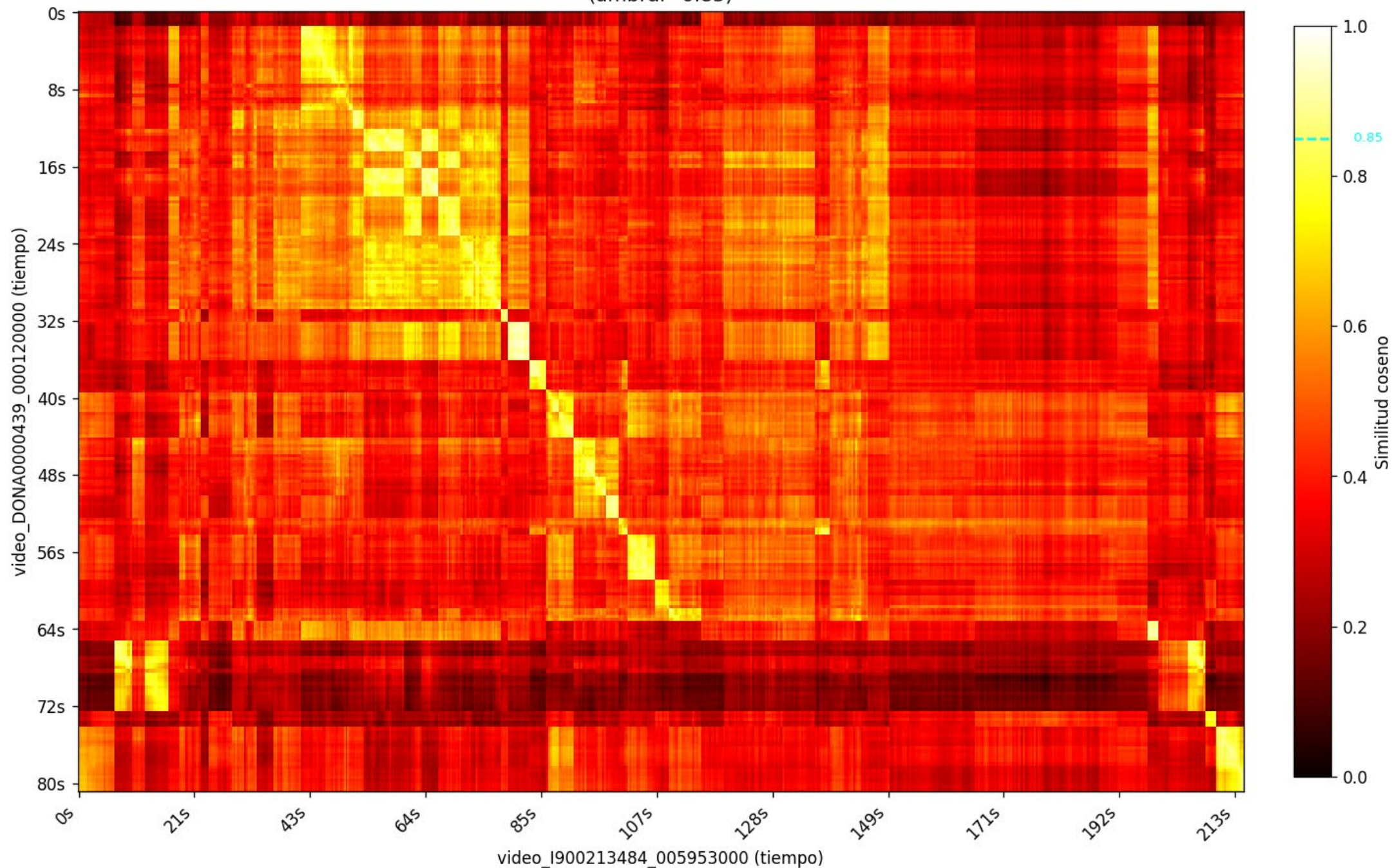
- Se clasifican intervalos como `similar` o `único`

Lógica de detección

Patron principal

- Si dos videos comparten contenido, en la matriz de similitud aparece una **diagonal** de valores altos

Similitud CLIP: video_DONA0000439_000120000 vs video_I900213484_005953000
(umbral=0.85)



¿Qué tiene la IA para mi archivo?

 Lorca — Análisis de Escenas

5 vídeos · 113 segmentos

[Segmento Similar](#)

VÍDEOS

...70486+Imagenes+inéditas.mp4

5 similar 5 inédito

...ORCA_PREVIEW_NITRATO F...

13 similar 14 inédito

...ORCA_PREVIEW_COPIA RTV...

14 similar 15 inédito

...o_DONA0000439_000120000....

7 similar 8 inédito

video_I900213484_005953000....

16 similar 16 inédito



▶ 1:14 / 3:35



video_I900213484_005953000

97s similar 117s inédito 215s total

■ Similar (≥2 vídeos) ■ Inédito (solo este vídeo)



Distribución de contenido

Similar: 45.3%

Inédito: 54.7%

● Este segmento también aparece en:

- LA BARRACA_LORCA_PREVIEW_COPIA RTVE
- LA BARRACA_LORCA_PREVIEW_NITRATO FE
- video_DONA0000439_000120000

SEGMENTOS (32)

0:00 → 0:06 **Inédito**
6.7s

0:06 → 0:09 **Similar**
2.7s ↔ LA BARRACA_LO...

0:09 → 0:12 **Inédito**
3.0s

0:12 → 0:16 **Similar**
4.0s ↔ LA BARRACA_LO...

RECUPERACION DE INFORMACIÓN MULTIMODAL BASADA EN ESPACIOS SEMÁNTICOS

Los sistemas tradicionales de IR (Recuperación de Información) se basan en el indexación de texto completo de los documentos en sí mismos o en metadatos que describen los datos en el caso de audio, imágenes o video.

Nuestro problema: solo contenido de video, **sin metadatos**

¿Cómo representamos el contenido solo de video?

- Consumir miles de horas de documentalistas creando metadatos.
- Utilizar nuevos sistemas de descripción automática de vídeos para crear metadatos. Grandes avances...
- Usar una representación conjunta en el espacio semántico.



Describir la información visual de manera que una computadora pueda entenderla.

Video shot description

GEMINI 2.5 Pro



Prompt: Give a formal description of the video to store as metadata. Try to guess characters, places, dates and any information useful for information retrieval.

- **Title:** Soviet Premier Nikita Khrushchev at a Beach with Officials, circa 1960.
- **Description:** This is short, silent, black-and-white archival footage showing a group of people on a sandy beach. The central figure, identified as Soviet Premier Nikita Khrushchev, is seen shirtless and wearing dark swim trunks. He is surrounded by and interacting with several men dressed in formal business suits. One woman in a coat and headscarf is also present. In the background, another man is visible in swimming briefs, and a large crane can be seen in the distance, suggesting proximity to a port or construction area. The footage captures an informal moment during what appears to be an official visit or inspection at a seaside location.
- **Subject (Probable):** The main subject is **Nikita Khrushchev** (First Secretary of the Communist Party of the Soviet Union, 1953-1964). Other individuals appear to be Soviet officials, aides, or members of his entourage.
- **Location (Probable):** The setting is likely a resort on the **Black Sea coast** of the former USSR, possibly near Khrushchev's dacha in **Pitsunda** (in the Abkhaz ASSR, Georgian SSR at the time) or another official government resort like those in Crimea.
- **Date (Estimated):** The film stock, attire, and Khrushchev's period in power suggest the footage was filmed between the **late 1950s and early 1960s**.
- **Keywords for Information Retrieval:** Nikita Khrushchev, Soviet Union, USSR, Cold War, beach, seaside, politics, leader, premier, officials, suit, swimwear, Black Sea, Pitsunda, Crimea, 1960s, 1950s, archival, vintage footage, historical.

Video shot description

GEMINI 2.5 Pro



Prompt: Give a formal description of the video to store as metadata **knowing that it was recorded in Spain**. Try to guest characters, places, dates and any information useful for information retrieval.

Title: Publicity Swim of Minister Manuel Fraga in Palomares, 1966

Date: March 8, 1966

Location: Palomares, Cuevas del Almanzora, Almería, Andalusia, Spain.

Synopsis: Short video clip depicting the Spanish Minister of Information and Tourism, Manuel Fraga Iribarne, on the beach at Palomares. Fraga, dressed in swimming trunks, is the central figure, surrounded by several other men dressed in formal business suits, as well as the U.S. Ambassador Angier Biddle Duke in swimming briefs. The group is engaged in conversation on the sand before entering the water. The background shows an undeveloped coastline with a crane visible in the distance.

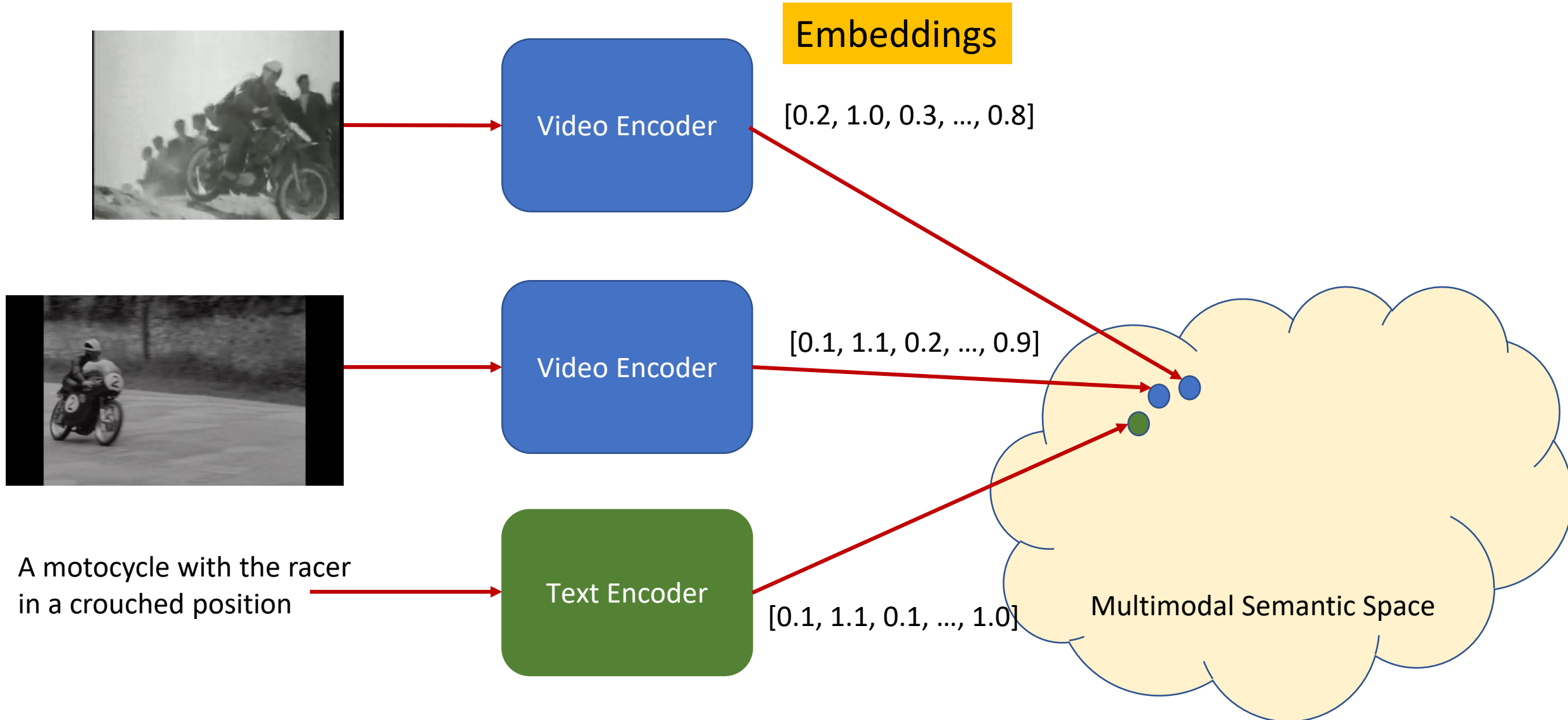
Historical Context and Significance: This footage documents a significant propaganda event organized by the Spanish government following the **Palomares Incident** of January 17, 1966. In that incident, a U.S. B-52 bomber and a KC-135 tanker collided, dropping four hydrogen bombs near the village, resulting in radioactive plutonium contamination on land. To counteract widespread fear and negative international press that threatened Spain's tourism industry and agricultural exports, Minister Fraga orchestrated this highly publicized swim. He, along with the U.S. Ambassador, swam in the sea to publicly "prove" that the area was safe from contamination. The event is a key example of state-managed public relations during the Franco regime and a notable episode of the Cold War.

Key Persons Identified:

- **Manuel Fraga Iribarne:** Spanish Minister of Information and Tourism.
- **Angier Biddle Duke:** United States Ambassador to Spain.

Keywords/Tags: Palomares Incident; Manuel Fraga Iribarne; Franco regime; Cold War; 1966; Almería; Nuclear Accident; Broken Arrow; Propaganda; Public Relations; U.S. Air Force; B-52; Hydrogen Bomb; Radioactive Contamination.

Video shot semantic representation: multimodal semantic space



Video shot semantic representation: multimodal semantic space

Problem building the multimodal semantic space

Needs to align textual and image/video semantic spaces

“A picture is worth a thousand words”



An image contains vast amounts of information, thousands of sentences can be related with an image.

- Men in business suits on the beach.
- People in formal suits and bathing suits.
- Two men in bathing suits chat with a group of men in suits.
- A woman wearing a headscarf laughs on the beach next to a group of men.
- An outdoor gathering with tents in the background.
- A group of people gathered on the sand with vehicles in the distance.
- Fraga Iribarne in the Palomares beach

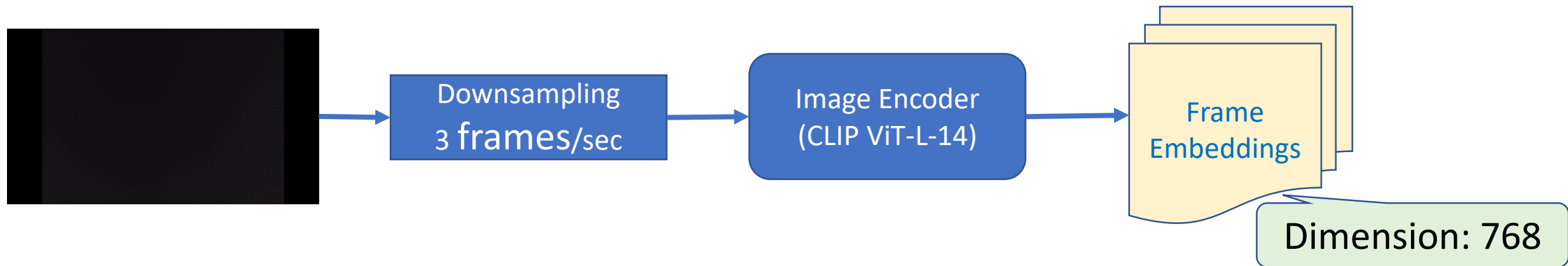
How is a video represented in the multimodal semantic space?

VIDEO ANALYSIS: from video to embeddings

INPUT: VIDEO (mp4 format)

1. STEP: DOWNSAMPLING VIDEO FRAME RATE TO 3 FRAMES/SEC
2. STEP: COMPUTE FRAME EMBEDDINGS USING AN IMAGE ENCODER (CLIP ViT-L-14)

OUTPUT: FRAME EMBEDDINGS (768 DIMENSIONAL VECTORS)



How is a video represented in the multimodal semantic space?

SEMANTIC ANALYSIS: from embeddings to semantic shots

INPUT: FRAME EMBEDDINGS (768D x (time video length in seconds x 3))

1. COARSE GROUPING: Find Clusters in the Semantic Space → SEMANTIC CLASSES
2. CLASS DIARIZATION: Distribute Semantic Classes in the Time-Line → SEMANTIC SCENES
3. FINE GROUPING: Find Clusters in the Semantic Scenes → SEMANTIC SHOTS
4. SHOT EMBEDDINGS: Use Embedding Mean Average over Semantic Shots

OUTPUT: SHOT EMBEDDINGS (768D x (#semantic shots))



How is a video represented in the multimodal semantic space?

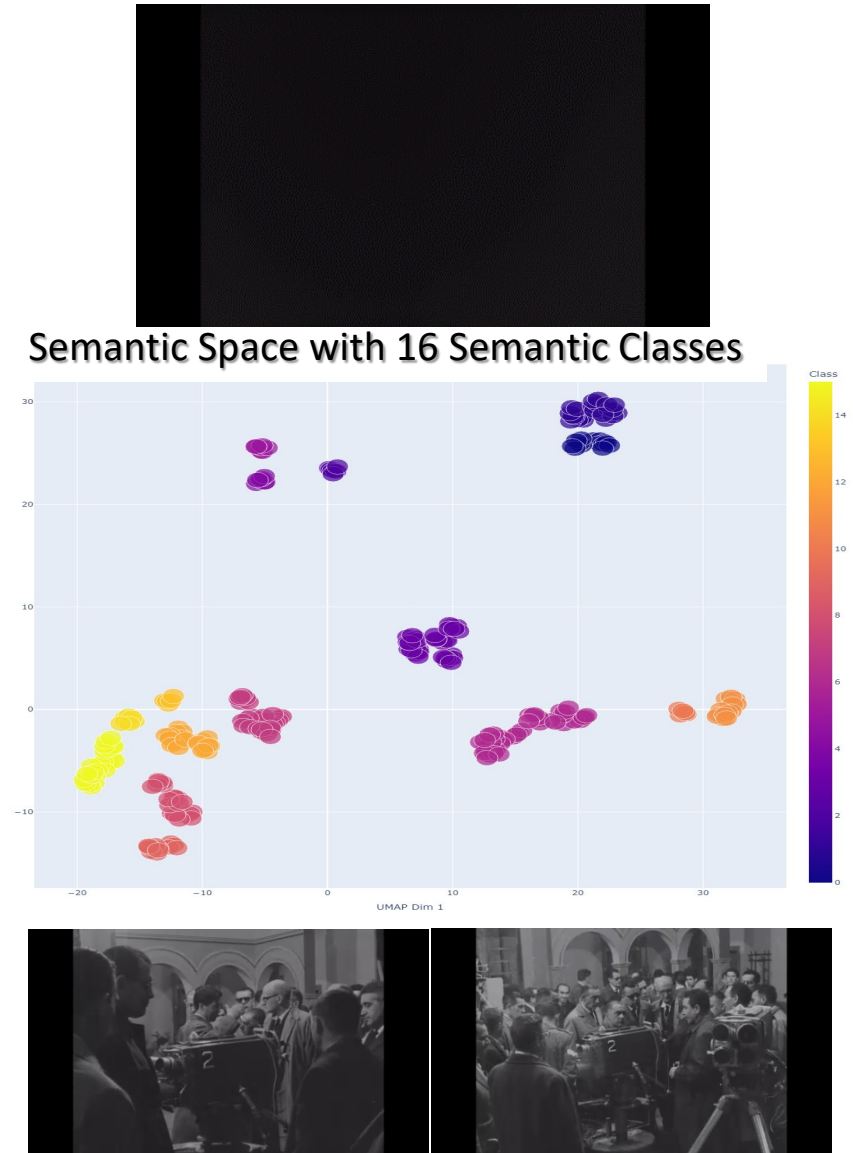
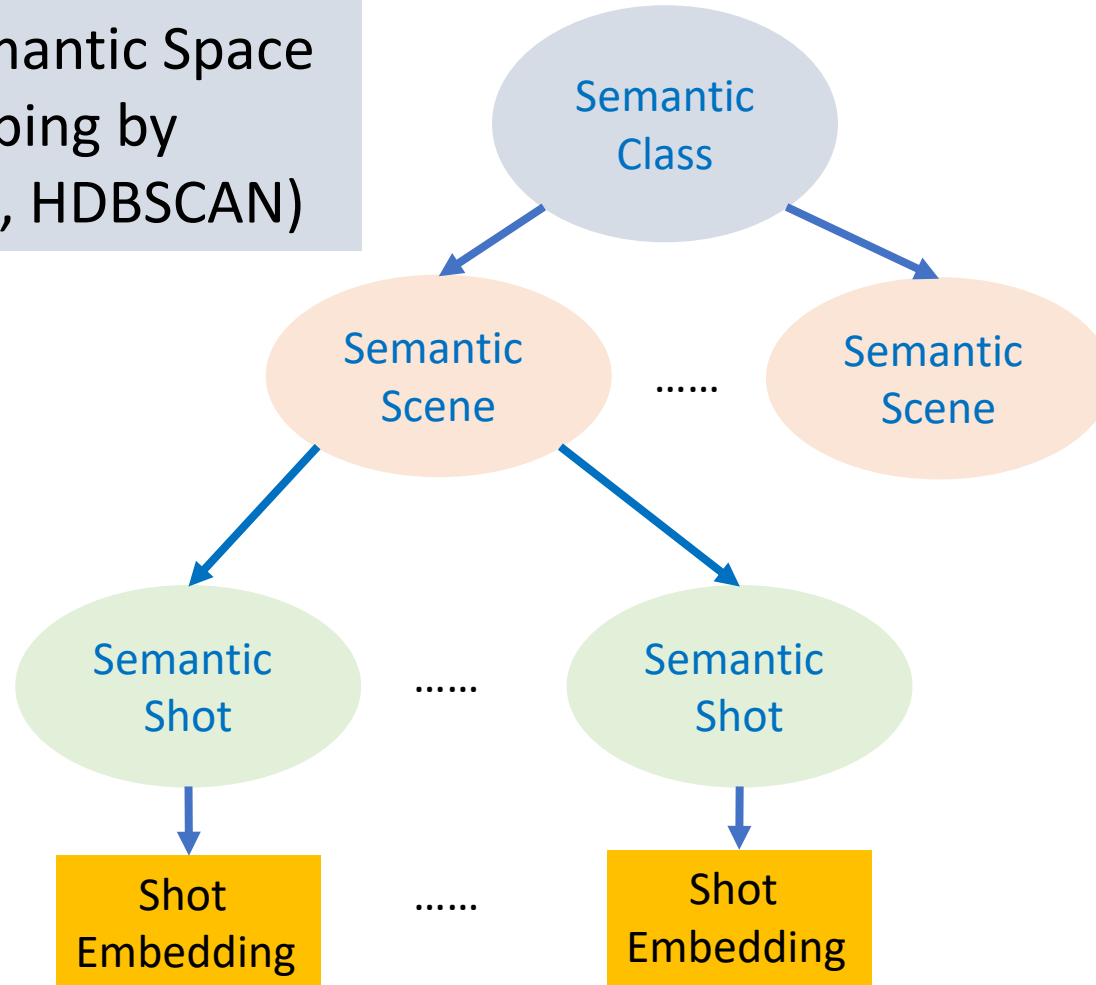
Semantic hierarchy

Clustering over Semantic Space
(coarse grouping by
density clustering, HDBSCAN)

Diarization over
Semantic Classes

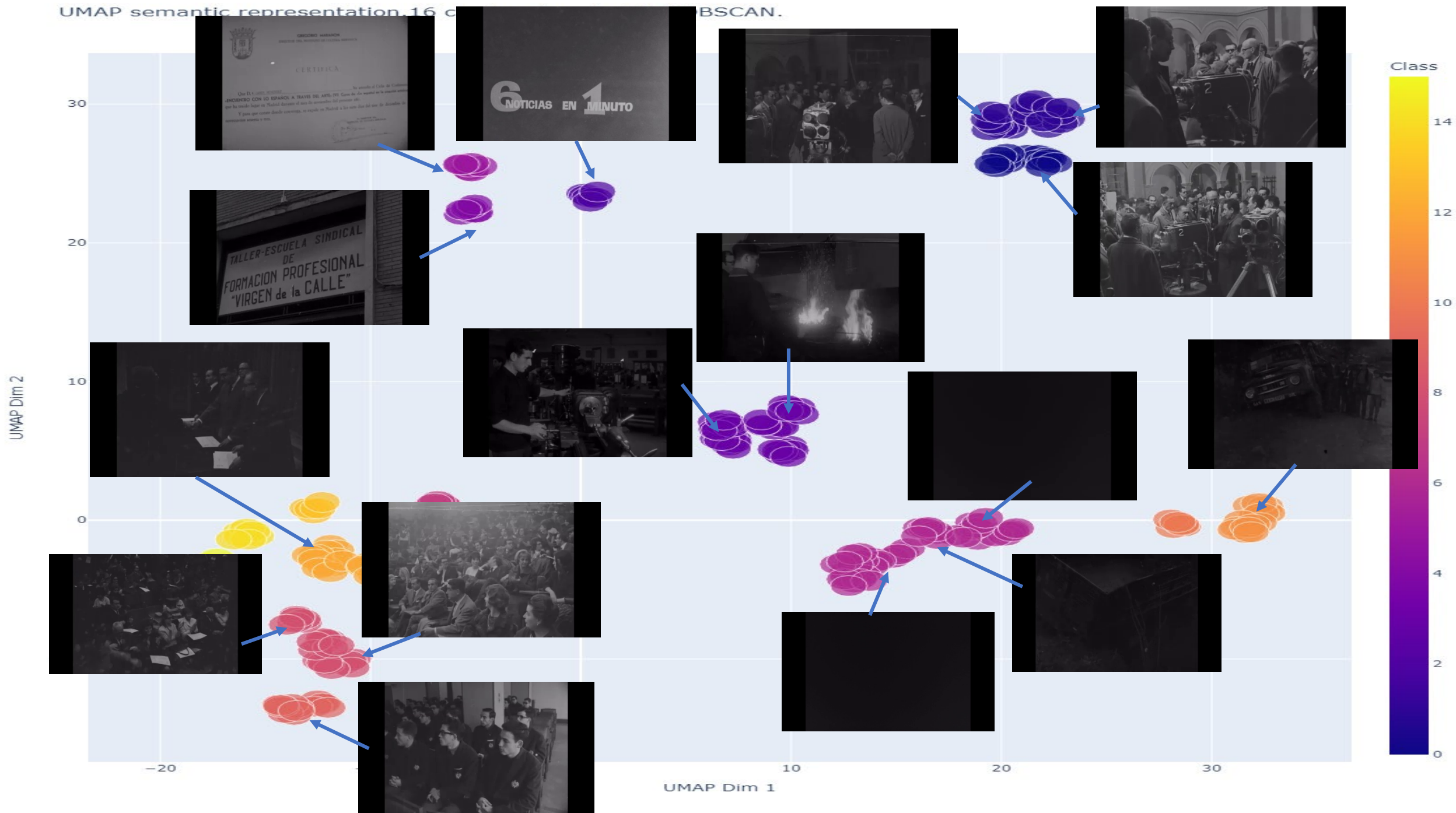
Clustering over
Semantic Scenes
(fine grouping,
HDBSCAN)

Mean Pooling



UMAP semantic representation, 16 c

BSCAN.



How is a video represented in the multimodal semantic space?

Video shots Timeline (23 scenes y 28 shots)



How is a video represented in the multimodal semantic space?

Database PoCArchiveRTVEDB

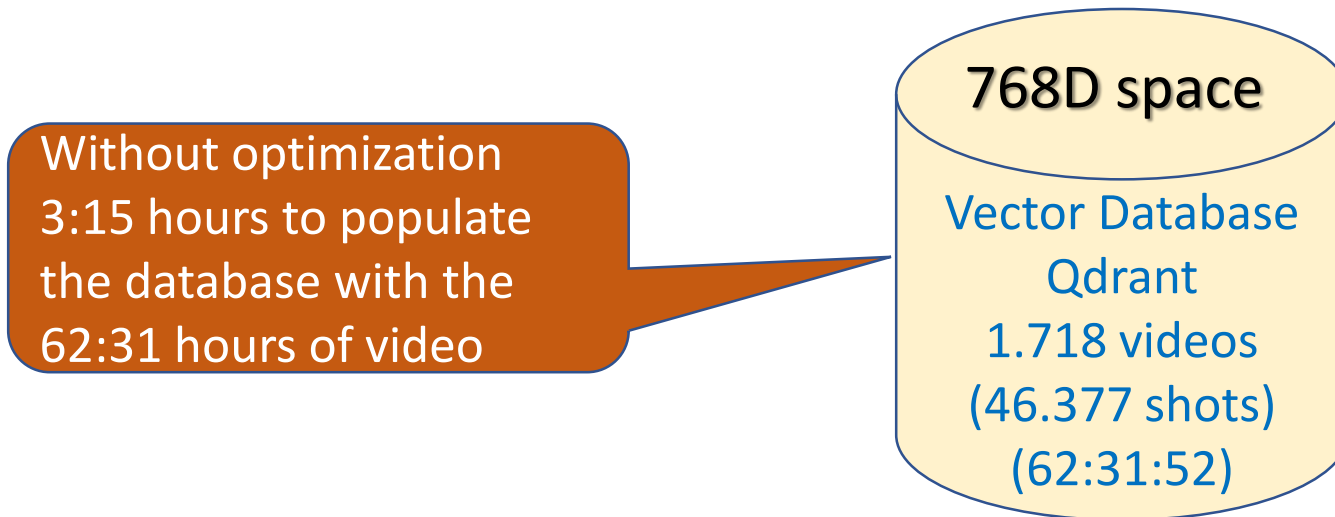
Phase I: 364 videos with a total duration of 27:35:02 (18.404 shots)

Phase II: 1.354 videos with a total duration of 34:56:50 (27.973 shots)

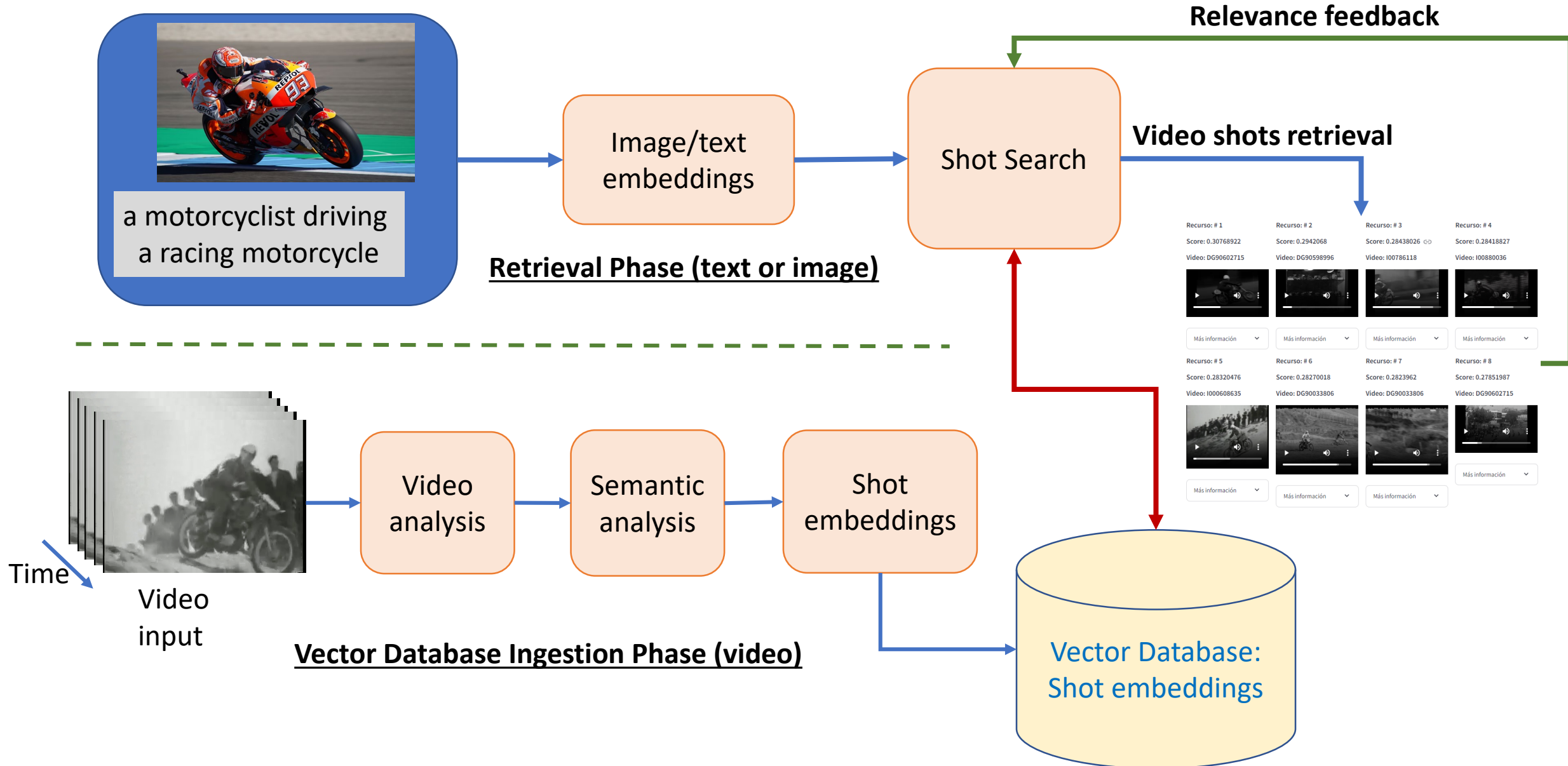
Total: 1.718 videos with a total duration of 62:31:52 (46.377 shots)

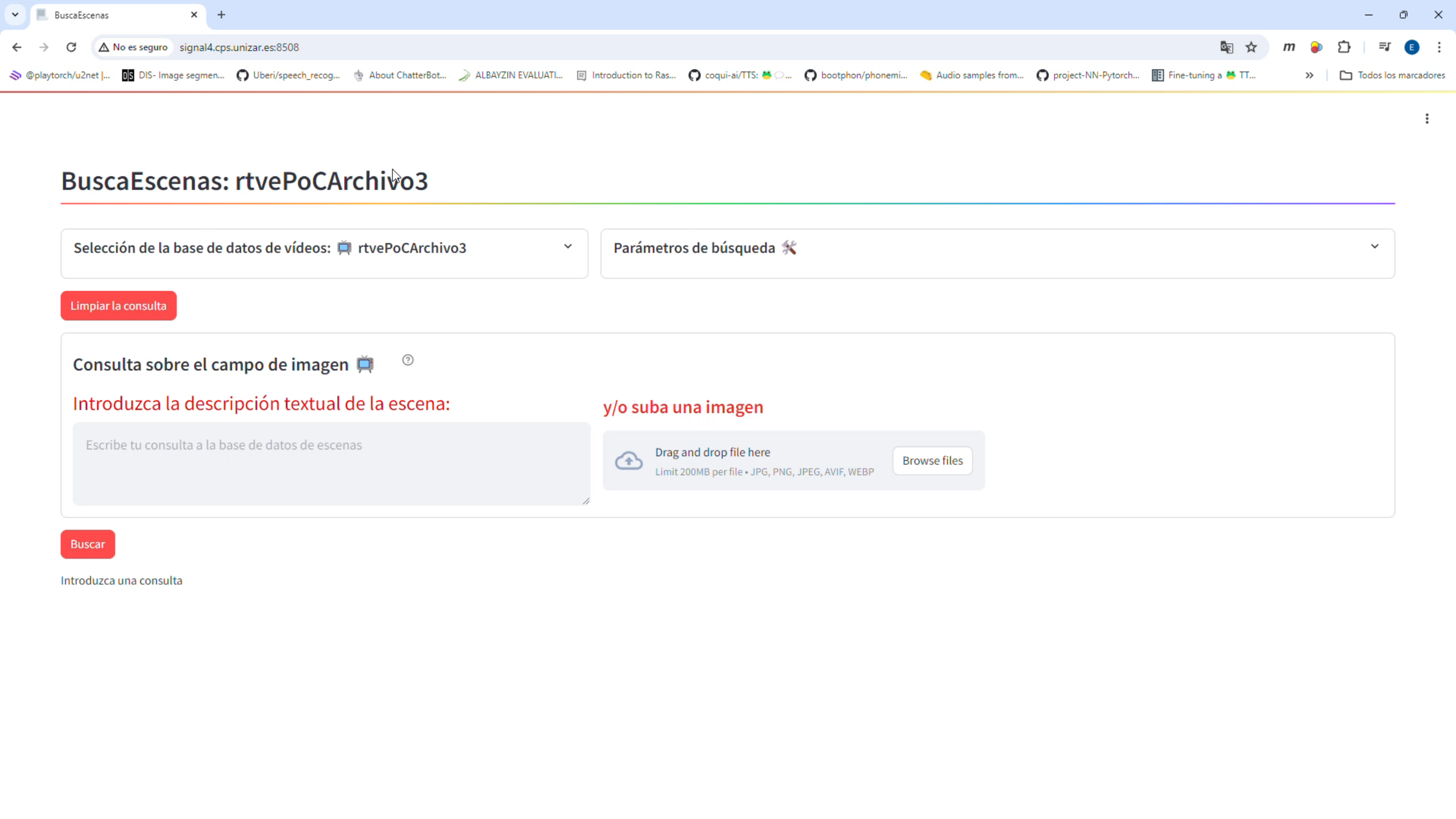
The videos are mainly from the 60's news archive, filmed in 16mm black and white and silent with no descriptive metadata in ARCA (the RTVE Archive).

They were filmed outdoors and feature a wide variety of shots and actions.



The multimodal retrieval system





Women working in a telephone switchboard

Recurso: # 1

Score: 0.3703646

Video: DG90337148



> Más información

Recurso: # 2 ↵

Score: 0.3452981

Video: I00837226

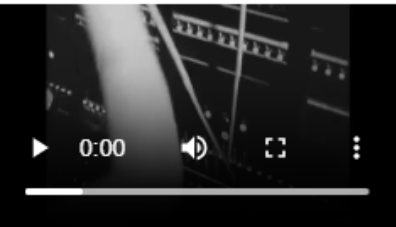


> Más información

Recurso: # 3

Score: 0.34363273

Video: I00864167

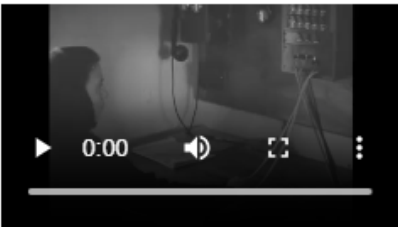


> Más información

Recurso: # 4

Score: 0.33399796

Video: I00863913



> Más información

Recurso: # 5

Score: 0.3258024

Video: DG90033223



> Más información

Recurso: # 6

Score: 0.32123908

Video: DG90033223

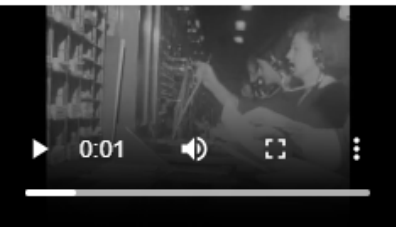


> Más información

Recurso: # 7

Score: 0.32046682

Video: I00837015



> Más información

Recurso: # 8

Score: 0.31972277

Video: I00864167



> Más información

Fraga Iribarne in a swimsuit on Palomares beach

Recurso: # 1

Score: 0.3223287

Video: I000169027



> Más información

Recurso: # 2 ⇄

Score: 0.3045665

Video: DG90021789

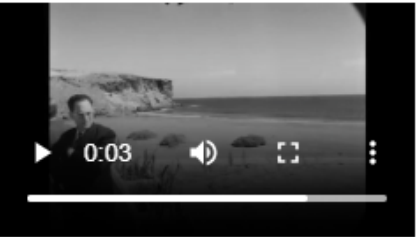


> Más información

Recurso: # 3

Score: 0.29866287

Video: I00869775



> Más información

Recurso: # 4

Score: 0.29417047

Video: I000169027



> Más información

Recurso: # 5

Score: 0.29329908

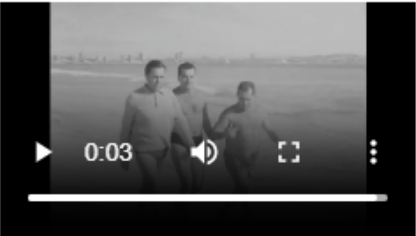
Video: DG90033888



Recurso: # 6

Score: 0.29264608

Video: I00909557



> Más información

Recurso: # 7

Score: 0.29040432

Video: I00903952

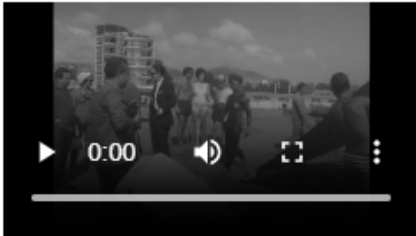


> Más información

Recurso: # 8

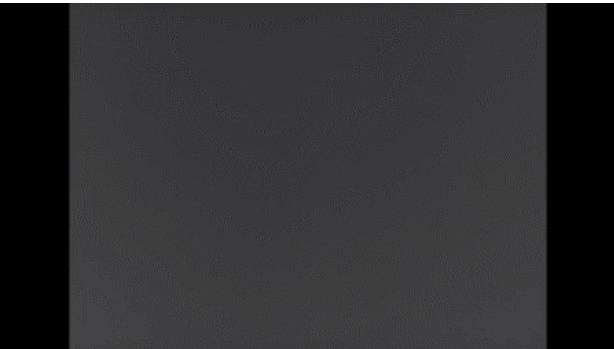
Score: 0.28269655

Video: I00903952



> Más información

FACTORY. People working in a Factory. (I00786311.mp4)



Recurso: # 1

Score: 0.26436812

Video: DG90347831



Más información



Recurso: # 2

Score: 0.2605429

Video: DG90347831



Más información



Recurso: # 3

Score: 0.25045383

Video: DG90347831



Más información



Recurso: # 4

Score: 0.2488241

Video: DG90033229



Más información



Recurso: # 5

Score: 0.24868599

Video: I00786096



Más información



Recurso: # 6

Score: 0.24504258

Video: DG90021789



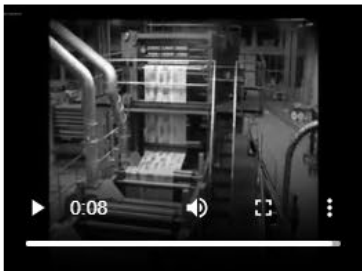
Más información



Recurso: # 7

Score: 0.24170998

Video: I900528772



Más información



Recurso: # 8

Score: 0.24167752

Video: DG90033601



Más información



Recurso: # 9

Score: 0.24079113

Video: DG90033692



Más información



Recurso: # 10

Score: 0.23922843

Video: DG90347831



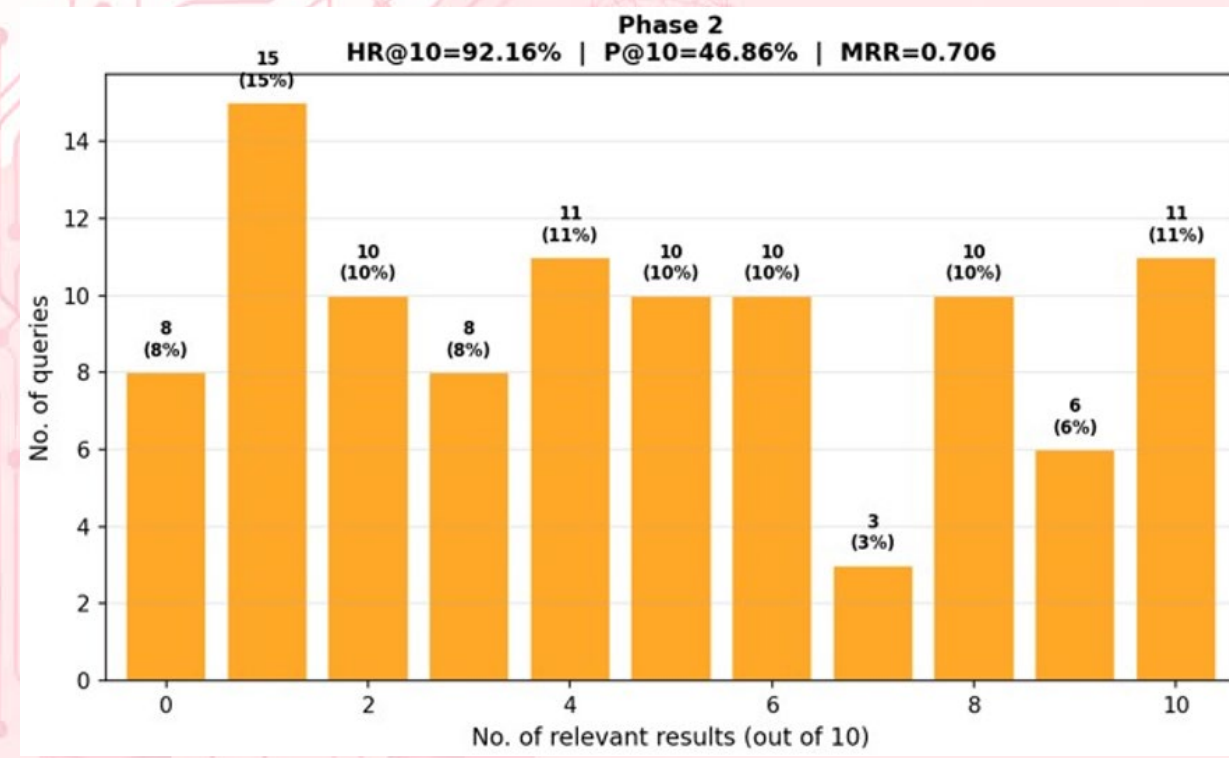
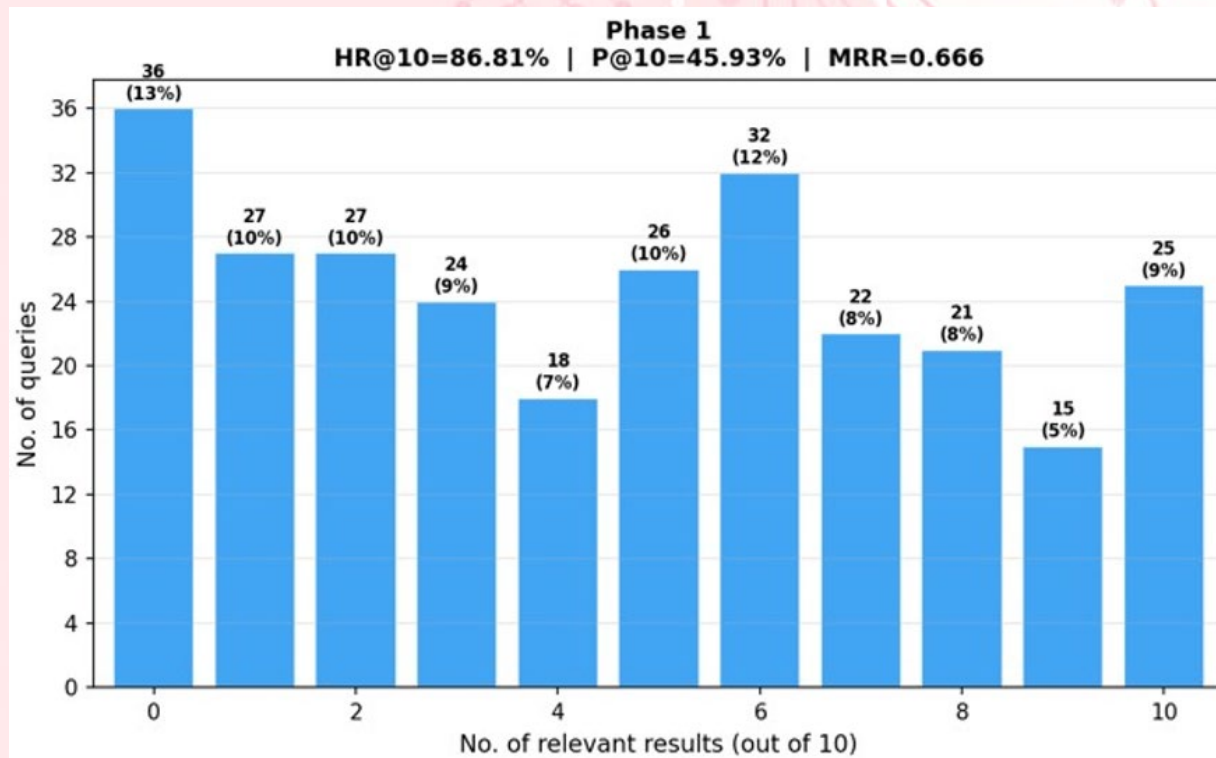
Más información



Hit Rate @K (HR@K): ¿Hay al menos un resultado relevante entre los K mejores?

Precision @K (P@K): ¿Qué proporción de los K mejores recuperados son relevantes?

Mean Reciprocal Rank (MRR): ¿En qué posición aparece el primer resultado relevante en la lista ordenada?



Parte 2

Detección de Estructura

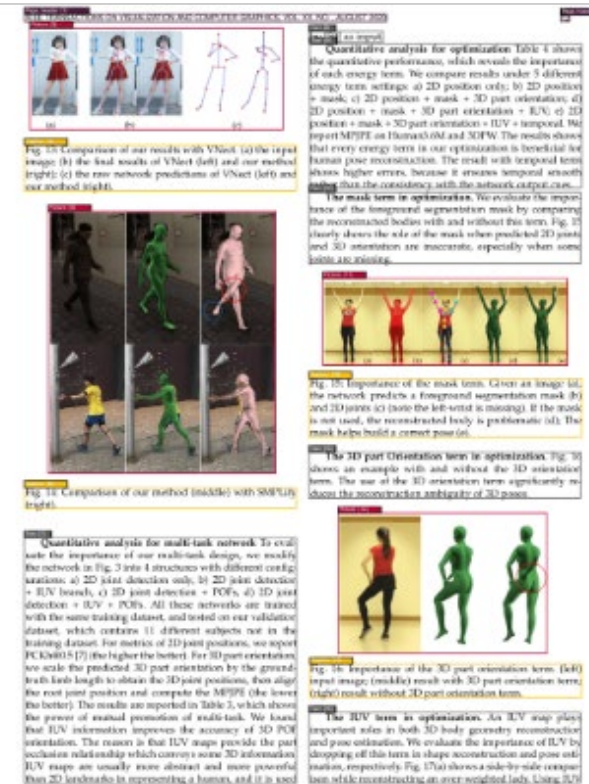
De la segmentación geométrica a la jerarquía lógica del documento

Análisis de la estructura del document (Layout)

Segmentación de Regiones

Identificación precisa de componentes estructurales en documentos digitalizados.

- **Roles Geométricos:** Texto, tablas, figuras y marcas de selección.
- **Roles Lógicos:** Títulos, autores, notas al pie y secciones.
- **Nivel de Píxel:** Segmentación semántica para bordes irregulares.



Telerín, análisis documentos

Detección de la estructura de los documentos

Primeros pasos:

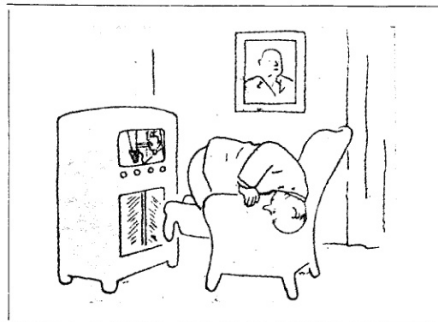
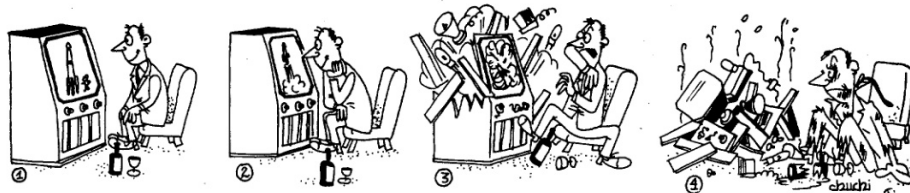
- Utilizar modelos específicos para el tratamiento de documentos:
 - PaddleOCR
 - Docling
- Utilizar modelos de lenguaje multimodales:
 - Qwen3-vl:8b
 - Gemini / GPT

Análisis de Layout Tradicional vs vLLMs

El Cambio de Paradigma

Los modelos clásicos (YOLO, Detectron, Docling, PaddleOCR) localizan cajas aisladas. Los **vLLMs** (**Vision LLMs**) procesan el documento de forma integral.

- 👁️ **Comprensión Contextual:** Relaciona visualmente las cabeceras con sus tablas correspondientes.
- 🔗 **Conversión Directa:** Capacidad de transformar una imagen médica o factura en un JSON limpio con un solo prompt.
- ✅ **Robustez ante el Ruido:** Los LLMs corrigen de manera implícita las anomalías del OCR utilizando el sentido del contexto lingüístico.
- ✅ **Flexibilidad Multilingüe y Temporal:** Capacidad nativa para procesar textos en varios idiomas y tipografías históricas.



AVISO A LOS SEÑORES SUSCRIPTORES

Rogamos a nuestros suscriptores que hasta ahora no hayan hecho efectivo el importe de su suscripción, tengan la bondad de hacerlo por medio de giro postal o telegráfico, cheque o talón, dirigido a nombre del ADMINISTRADOR GENERAL DEL MINISTERIO DE INFORMACIÓN Y TURISMO, Montequinza, 4, Madrid. Suplicamos perdonen las molestias que esta forma de pago les pueda ocasionar, pero la buena marcha de nuestra administración así lo exige, y les recordamos que el importe de la suscripción es de 150 pesetas para el año 1955.

RCA ESPAÑOLA

ASOCIADA DE LA
RADIO CORPORATION OF AMERICA

PALADINES DE LA TELEVISION EN EL MUNDO



Adquiera un televisor RCA
y gozará de la pureza máxima
en imagen y sonido



RCA ESPAÑOLA, S. A.
ASOCIADA DE LA RADIO CORPORATION OF AMERICA
Avenida de América y Hernández de Tejada (Ciudad Lineal) - Madrid (España)

NOTICIAS PINTORESCAS

TRAGAPERRAS

Si alguna vez visita usted Suttugar y compra un paquete de cigarrillos de cierta marca en una de las numerosas máquinas automáticas que abundan por las calles, no se extrañe de que el aparato, aunque no le haga una reverencia ni se quite un sombrero que no tiene, le diga: "Muchas gracias. Hasta la próxima."

Esta nueva maravilla ha llamado tanto la atención, que una emisora de televisión—sospechamos que como pronganda de los cigarrillos—ha considerado dedicarle un reportaje, sometiéndola a una entrevista. Naturalmente, la máquina sólo contestó: "Muchas gracias. Hasta la próxima", a cada pregunta. Pero no hay que desanimarse. Estamos seguros de que muy pronto se inventará un aparato de este tipo que será muy capaz de ganar en locuacidad a esos concursantes que suelen, al ser invitados por el locutor, decir: "Saludos a mi abuelita, que me estará escuchando."



CHQUITAJA



Por el Gobierno norteamericano han sido dados a conocer los detalles completos de un radio-receptor de pulsera, que puede captar emisiones y mensajes a 72 kilómetros de distancia. El receptor está formado por tres elementos: el receptor propiamente dicho, que se lleva en la muñeca como si se tratara de un reloj de pulsera, con una pequeñísima batería de 12,5 milímetros de ancho y 1,5 milímetros de largo; un auricular, que se introduce en el oído como los de algunos tipos de ayuda contra la sordera, y de la antena, que puede colocarse dentro de la correa que sujeta el receptor a la muñeca, y que no es necesaria cuando el usuario se encuentra en zonas urbanas.

BAÑO

En el Olympia de Londres se ha celebrado una exposición de edificaciones. Entre muchas curiosas novedades figuraba un cuarto de baño, donde podía verse una bañera diseñada por Denis Potts que lleva un panel de plástico que contiene un dispositivo de televisión. Mientras la señora de la casa tranquilamente toma su baño, puede observar por este dispositivo quién llama a la puerta del piso.... y según la categoría del visitante, el ama de casa inglesa podrá o no continuar tomando su baño.



RECOMENDADA

El director de una emisora de televisión norteamericana no ha tenido más remedio, para complacer a uno de los más importantes consejeros de la entidad, que darle un papel bastante importante en uno de los seriales a una actriz principiante. Pasados unos días, el consejero le pregunta:

—¿Cómo va la actriz que le recomendé?
—Muy bien—contesta—. Estoy haciendo dos seriales con ella.
—¿Dos?
—Sí; el primero y el último.

VIGILANCIA

La televisión declara la guerra a los cleptomáticos y los vigila. Unos grandes almacenes de París, ante el aumento de robos en sus locales, ha instalado un sistema de televisión interior con el que pueden vigilarse con tranquilidad todos los lugares del extraordinario establecimiento. De esta forma, y desde un despacho, varios vigilantes no pierden de vista a los sospechosos. Esta pequeña cámara de televisión va unida por un cable a los lugares que se desea vigilar, y así, cómodamente, sin moverse, se domina la situación y al cleptomático sospechoso.



Docling

Picture: 0.47



Picture: 0.40



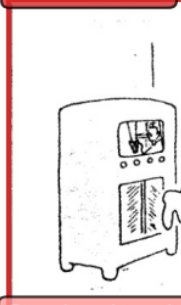
Picture: 0.70



Picture: 0.86



Picture: 0.97



Section-header: 0.92

Text: 0.98

... suscriptores que hasta ahora no hayan hecho efectivo el importe de su suscripción, tengan la bondad de hacerlo por medio de giro postal o telegráfico, cheque o talón, dirigido a nombre del ADMINISTRADOR GENERAL DEL MINISTERIO DE INFORMACIÓN Y TURISMO, Montequinza, 4, Madrid. Suplicamos perdonen las molestias que esta forma de pago les pueda ocasionar, pero la buena marcha de nuestra administración así lo exige, y les recordamos que el importe de la suscripción es de 150 pesetas para el año 1955.

Text: 0.32

ader: 0.88

Section-header: 0.31

Text: 0.33

Section-header: 0.32

Text: 0.36

Picture: 0.90



Caption: 0.32

er: 0.42

Adquiera un televisor RCA

y gozará de la pureza máxima

en imagen y sonido

Picture: 0.87

Section-header: 0.52

Text: 0.31

Text: 0.30

Text: 0.41

Section-header: 0.90

Section-header: 0.84

Text: 0.42

ader: 0.84

Text: 0.81

Text: 0.81

Text: 0.89

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Text: 0.81

Picture: 0.95



Section-header: 0.92

Text: 0.40

ader: 0.85

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Text: 0.32

Picture: 0.92



List-item: 0.32

Picture: 0.30

ader: 0.88

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.30

Text: 0.31

ader: 0.87

Text: 0.31

Text: 0.31

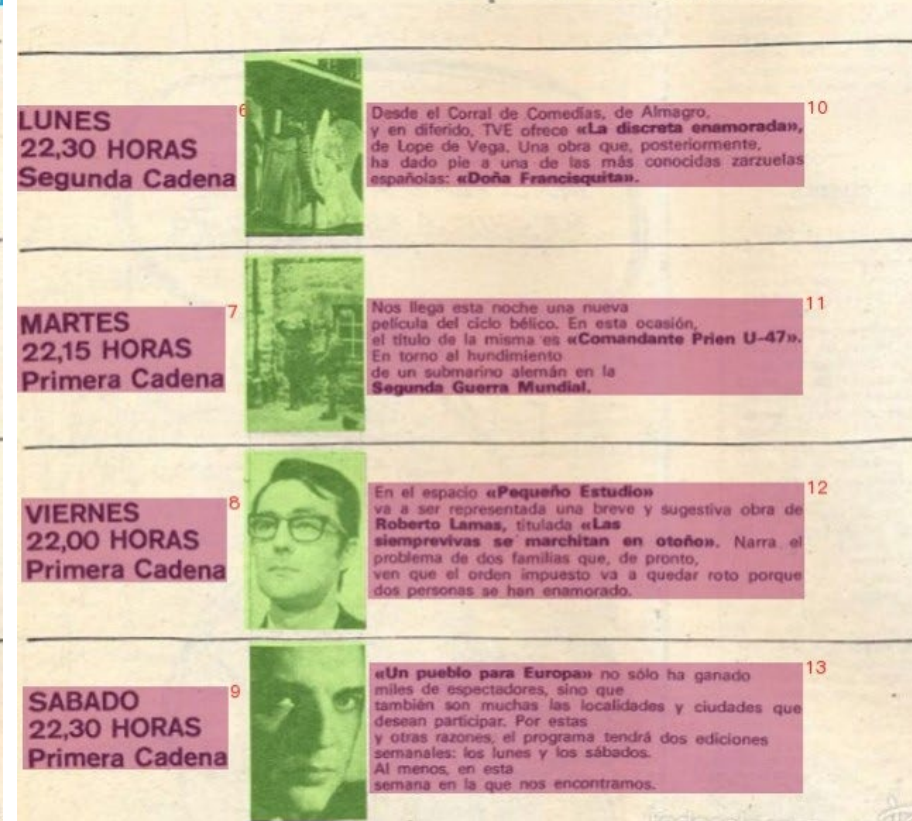
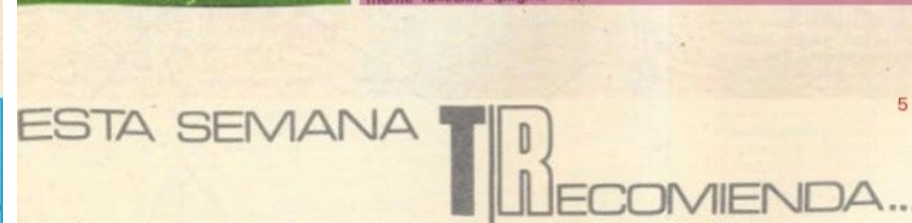
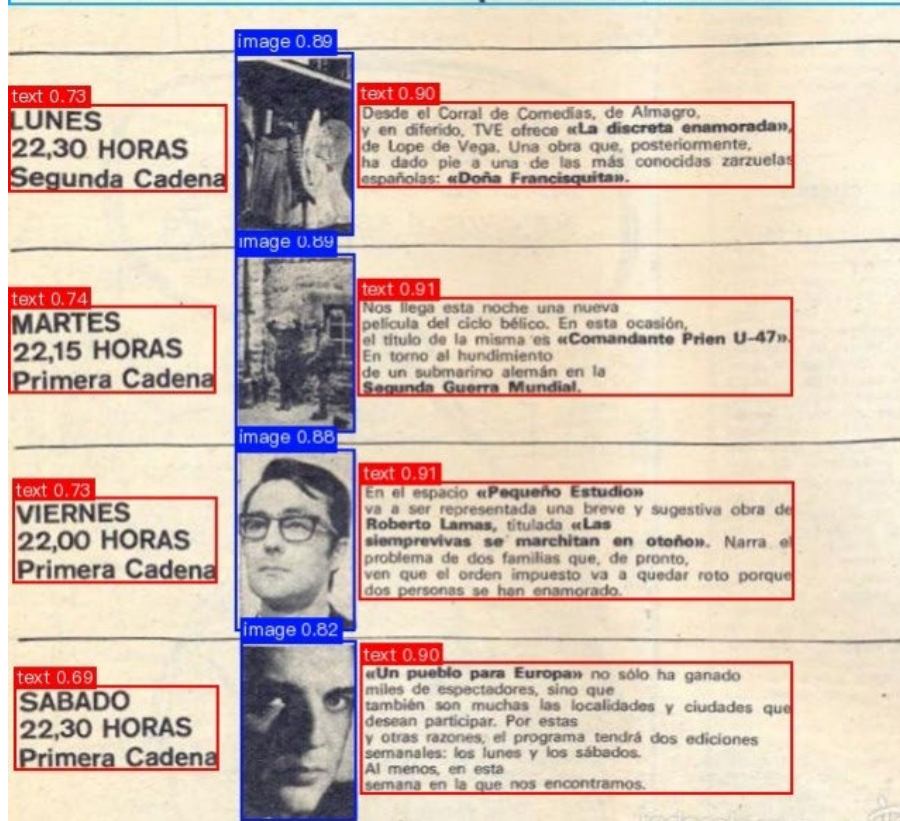
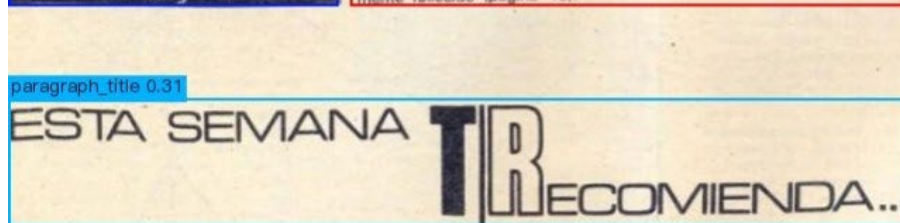
Text: 0.31

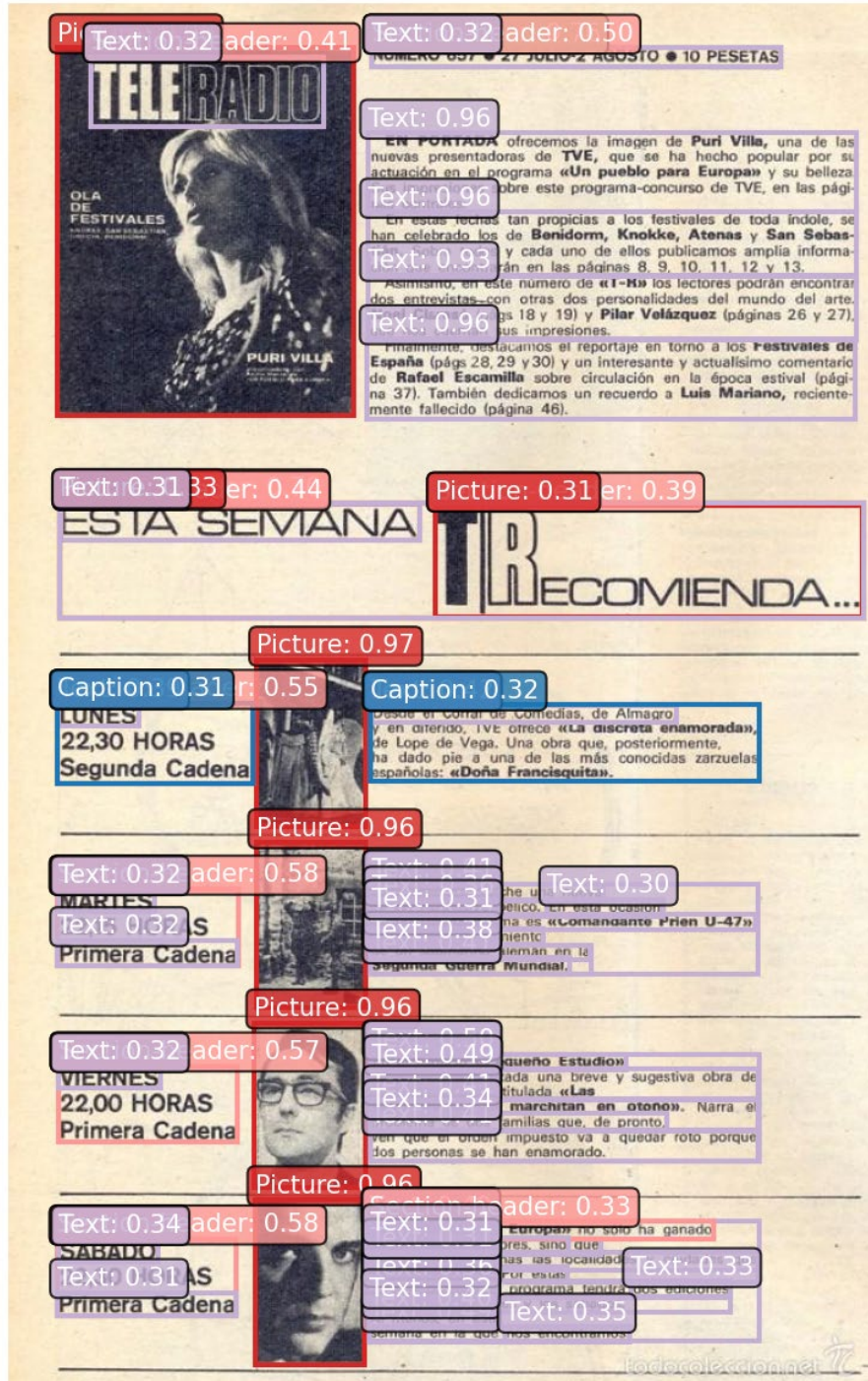
Text: 0.31

Text: 0.31

Text: 0.31

Text: 0.31





Prompt Gemini3: analiza la estructura de esta imagen y dame el texto escrito ordenado con la estructura original

NUMERO 657 • 27 JULIO-2 AGOSTO • 10 PESETAS

TELE RADIO

OLA DE FESTIVALES *(Imagen: PURI VILLA)*

[EDITORIAL / EN PORTADA]

EN PORTADA ofrecemos la imagen de **Puri Villa**, una de las nuevas presentadoras de **TVE**, que se ha hecho popular por su actuación en el programa «**Un pueblo para Europa**» y su belleza. Sus impresiones sobre este programa-concurso de TVE, en las páginas centrales.

En estas fechas tan propicias a los festivales de toda índole, se han celebrado los de **Benidorm, Knokke, Atenas y San Sebastián**. Sobre todos y cada uno de ellos publicamos amplia información que encontrarán en las páginas 8, 9, 10, 11, 12 y 13.

Asimismo, en este número de «T-R» los lectores podrán encontrar dos entrevistas con otras dos personalidades del mundo del arte. **Noel Clarasó** (págs 18 y 19) y **Pilar Velázquez** (páginas 26 y 27), que nos cuentan sus impresiones.

Finalmente, destacamos el reportaje en torno a los **Festivales de España** (págs 28, 29 y 30) y un interesante y actualísimo comentario de **Rafael Escamilla** sobre circulación en la época estival (página 37). También dedicamos un recuerdo a **Luis Mariano**, recientemente fallecido (página 46).

ESTA SEMANA TRECOMIENDA...

DÍA Y HORA	CANAL	DESCRIPCIÓN DEL PROGRAMA
LUNES 22,30 HORAS	Segunda Cadena	Desde el Corral de Comedias, de Almagro, y en diferido, TVE ofrece « La discreta enamorada », de Lope de Vega. Una obra que, posteriormente, ha dado pie a una de las más conocidas zarzuelas españolas: « Doña Francisquita ».
MARTES 22,15 HORAS	Primera Cadena	Nos llega esta noche una nueva película del ciclo bélico. En esta ocasión, el título de la misma es « Comandante Prien U-47 ». En torno al hundimiento de un submarino alemán en la Segunda Guerra Mundial.
VIERNES 22,00 HORAS	Primera Cadena	En el espacio « Pequeño Estudio » va a ser representada una breve y sugestiva obra de Roberto Lamas , titulada « Las siemprevivas se marchitan en otoño ». Narra el problema de dos familias que, de pronto, ven que el orden impuesto va a quedar roto porque dos personas se han enamorado.
SABADO 22,30 HORAS	Primera Cadena	« Un pueblo para Europa » no sólo ha ganado miles de espectadores, sino que también son muchas las localidades y ciudades que desean participar. Por estas y otras razones, el programa tendrá dos ediciones semanales: los lunes y los sábados. Al menos, en esta semana en la que nos encontramos.

Misión: Actúa como un sistema experto en extracción de datos estructurados (Document AI). Tu objetivo es transcribir y organizar la información de la imagen adjunta siguiendo estrictamente el esquema JSON proporcionado.

Directrices de Formato:

Nombres de Campos: Utiliza exclusivamente los nombres de campos definidos en la estructura de abajo.

Campos Vacíos: Si un dato no existe, usa null.

Coordenadas: Los bounding_box deben estimarse en una escala de 0 a 1000 (normalizada).

Instrucciones de Extracción:

Metadatos de la Publicación:

Extrae el nombre de la revista, fecha exacta y día de la semana.

Módulo de Artículos (Contenido Editorial):

article_id: Crea un identificador único (ej: art_titulo_programa).

full_text: Realiza una transcripción íntegra palabra por palabra. Corrige errores de OCR por guiones de fin de línea o manchas, pero mantén el estilo original. No resumas aquí.

brief_summary: Crea una síntesis propia de 1 o 2 frases.

media: Para cada foto asociada, describe visualmente el contenido y transcribe el pie de foto literal. Incluye el bounding box con las coordenadas de la posición de la imagen.

Módulo de Programación (Parrilla Horaria):

Crea bloques por cada programa con la siguiente información:

channel (ej: Primer Programa, Segundo Programa) y por Bloque Horario (Sobremesa, Tarde, Noche).

Para cada entrada, captura: hora y título.

Créditos Técnicos y Artísticos: Busca y extrae nombres de directores, realizadores, guionistas, presentadores y el reparto completo que aparezca bajo el título del programa en la parrilla. Presta especial atención a las listas de actores en letra pequeña bajo los títulos de la programación

linked_article_id: Si el programa tiene un artículo expandido en la misma página, vincúlalo usando el id creado en el punto 2.

Marcadores: Identifica si el programa se emite en "Color" o tiene patrocinios específicos (ej: Philips).

Módulo de Programación de Radio:

Crea bloques por emisora (station) incluyendo información sobre la misma en station_information (p.e. frecuencia o longitud de onda de emisión, localidad, etc.)

Identifica y extrae información sobre emisoras de radio, incluyendo el nombre de la emisora, horarios de los programas y descripción del contenido o secciones.

Estructura JSON Obligatoria:

```
{
  "metadata": {
    "source_magazine": "string",
    "publication_date": "YYYY-MM-DD",
    "day_of_week": "string"
  },
  "editorial_content": [
    {
      "article_id": "string (slug_del_titulo)",
      "section_title": "string",
      "main_title": "string",
      "full_text": "string (transcripción íntegra)",
      "brief_summary": "string (1-2 frases)",
      "media": [
        {
          "type": "image",
          "description": "string",
          "caption_literal": "string",
          "bbox": { "top": 0, "left": 0, "bottom": 0, "right": 0 }
        }
      ]
    }
  ]
},
```

```
"tv_schedule": {
  "blocks": [
    {
      "channel": "string (channel_1|channel_2)",
      "block_name": "string (SOBREMESA|TARDE|NOCHE|NONE)",
      "entries": [
        {
          "time": "HH:mm",
          "title": "string",
          "is_color": "boolean",
          "sponsor": "string|null",
          "description": "string|null",
          "credits": {
            "directors": ["string"],
            "cast": ["string"],
            "crew": ["string (guionistas, realizadores, etc)"]
          }
        }
      ]
    }
  ]
},
```



```
"radio_schedule": [
  {
    "station": "string",
    "station_information": "string",
    "entries": [
      {
        "time": "HH:mm",
        "title": "string",
        "content_description": "string|null"
        "linked_article_id": "string|null"
      }
    ]
  }
],
```

```
"advertising": [
  {
    "advertiser": "string (marca o empresa)",
    "ad_copy": "string (texto completo del anuncio)",
    "visual_elements": [
      {
        "description": "string (qué se ve en la imagen)",
        "bbox": { "top": 0, "left": 0, "bottom": 0, "right": 0 }
      }
    ]
  }
]
```

Pasos a seguir:

Identifica logotipos o bloques de texto que no pertenezcan a la programación ni a artículos de opinión

Analiza los encabezados "Primer programa" y "Segundo programa" para separar los canales.

Identifica los bloques tipográficos (negritas para títulos, cursivas para créditos).

Mapea cada programa de televisión o radico con su artículo extendido si existe, usando el article_id.

Usa listas (arrays) para campos con múltiples nombres (actores, invitados).

Si una información no está presente, usa null o no incluyas el campo.

Asegura que el JSON sea válido y esté bien anidado.

Idioma: Español.

Devuelve únicamente el objeto JSON.

Estructuras complejas: Gemini3 con prompt mejorado

```
{
  "metadatos": {
    "publicacion": "TeleRadio",
    "fecha": "1975-11-12",
    "dia_semana": "miércoles",
    "idioma": "español"
  },

  "programacion": {
    "primer_programa": [
      {
        "bloque": "NOCHE",
        "emisiones": [
          {
            "hora": "21:00", "titulo": "TELEDIARIO", "formato": "COLOR PHILIPS", "nota": "Segunda edición",
            "hora": "21:30", "titulo": "LOS RIOS", "formato": "COLOR PHILIPS", "id_referencia_articulo": "art_rios_tajo"
          },
          "creditos": { "textos": ["Pedro de Lorenzo"], "guion": ["Eduardo Delgado"], "direccion_realizacion": ["Ramón Comas"] }
        ],

        {
          "hora": "22:00", "titulo": "ESTE SEÑOR DE NEGRO", "formato": "COLOR PHILIPS", "id_referencia_articulo":
          "art_senor_negro_aventura", "creditos": { "autor": ["A. Mingote"], "direccion": ["Antonio Mercero"], "reparto": [ "José
          Luis López Vázquez", "Florinda Chico", "Mary Carmen Prendes", "Angeles Garralón", "Valeriano Andrés", "Angela
          Capilla", "Pepa Terrón", "Javier de Campos", "Adriana O'Neil" ] } },
          ...]
        }....}]
      }
    ]
  }
}
```



```
"articulos": [  
  {  
    "id": "art_ríos_tajo",  
    "seccion": "LOS RÍOS",  
    "titulo": "EL TAJO (1)",  
    "texto_literal": "Se trata del primer episodio de los dedicados al más largo y uno de los más importantes  
ríos de España; el guión está escrito por Eduardo Delgado sobre una idea de Pedro de Lorenzo. La realización  
en esta ocasión corresponde a Ramón Comas. El Tajo nace en la cordillera Ibérica, cerca de Casas de Fuente  
García (Teruel). Encajonado entre abruptos parajes, corre en dirección SE-NE hasta el lugar conocido por  
Buena Fuente donde se produce un brusco cambio de rumbo hacia la dirección SO. Ese tramo está regulado  
por grandes embalses que permiten el aprovechamiento integral de sus aguas, habiéndose instalado en  
Entrepeñas, uno de los complejos hidroeléctricos más importantes de España. Al descender a la meseta va en  
dirección Oeste, y flanqueado por el sistema Central al Norte y los Montes de Toledo al Sur llega al final de su  
primera etapa, Aranjuez, antigua residencia real, y centro de una fértil comarca agrícola.",  
    "resumen": "Documental sobre el curso alto del río Tajo, desde su nacimiento en Teruel hasta su llegada a  
Aranjuez.",  
    "imagen": {  
      "fichero": "/home/eduardo/PoC/imágenes/19751112-art_ríos_tajo-1.jpg"  
      "descripcion_visual": "Vista panorámica del río Tajo bordeando la ciudad de Toledo.",  
      "pie_de_foto": "El Tajo a su paso por Toledo."  
    }  
  }  
]
```



TELERÍN RTVE: El Asistente Cognitivo para Archivos Históricos

Consulta documental
mediante agentes LLM

Cátedra RTVE – Universidad de Zaragoza

Archivo de RTVE

Grupo investigación ViVolab -

INIA7AR

¡Hola! Soy **TELERÍN**, tu asistente de búsqueda en el archivo histórico de **RTVE**: Teleradio (1958-1969), Fotografías y vídeos. ¿Qué te gustaría buscar?

¿Qué puedo preguntarte?



Contexto

Las primeras emisiones de TVE no se conservan, ya que la televisión se realizaba en directo y solo se registraban algunos fragmentos en cine.



El 31 de diciembre de 1957 aparece la revista TELEDIARIO (más tarde TELERADIO), siendo una fuente de conocimiento de la oferta televisiva inicial.

En paralelo, RTVE dispone de una **colección fotográfica de la revista Teleradio**, con imágenes de personas, platós y espacios de los primeros años de televisión.

En el Archivo General de la Administración están disponibles partes de emisión de los años 60, complemento perfecto a la información de programación disponible en la revista TELERADIO



TELEDIARIO

PROGRAMA SEMANAL DE LA TVE

NUM. 1 • MADRID 31 DICIEMBRE 1957

PROPOSITO

Es muy concreto y claro: Aspiramos a una comunicación con el público. Queremos hacer llegar al telespectador el programa semanal puntualmente y esperamos que el público nos haga llegar sus impresiones sobre la orientación y ejecución de aquí. Esta comunicación ha de ser real y efectiva, y para ello lanzamos este TELEDIARIO. Los abonados en su calidad de tal y haciendo uso exclusivamente de su número, tendrán personalidad para dirigirse a T. V. E. El semanario actuará cumpliendo términos de televisión—de canal de comunicación.

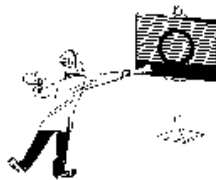
T. V. es como un pozo ancho y profundo, de gran capacidad, en el que caben muchas ideas. La programación no puede ser obra de una sola cabeza. Les invitamos a programar con nosotros. Nos sentiremos observados a diario por un público inteligente y sensible que sabe que T. V. es un espectáculo de hogar, y que el ideal es llegar a lograr, al propio tiempo que un medio informativo de primera calidad, un entretenimiento digno y honesto, que enriquezca la cultura y eleve el nivel espiritual de los españoles.

Televisión Española cuenta un año de vida. Es, pues, como un niño que por fuerza ha de crecer. El TELEDIARIO nace hoy; también crecerá. De ello se encargarán ustedes con nosotros.

Un cordial saludo y nuestros mejores deseos para el nuevo año.



El pasado día 22 se celebró en el Estadio Santiago Bernabéu, de Madrid, el último campeonato futbolístico de la capital, por primera vez dos equipos rivales de Madrid. Televisión Española no podía estar ajena al citado acontecimiento, y así a su disposición de servicio, instaló sus cámaras en el campo para ofrecer a los telespectadores una visión directa e instantánea de las vicisitudes del partido.



El fundamento técnico es sencillo. Impulsando un haz, se crea una imagen que viaja a la velocidad de la luz, es decir, siguiendo letra a letra cada carácter de la imagen. Y pasando desde el final de un rayo al comienzo del siguiente. En la cámara transmisora se forma una imagen óptica sobre una placa, como se forma la imagen fotográfica en el cristal sensible de una cámara.

Una placa de electrones recibe esta imagen de luz, la convierte en una imagen eléctrica.

Al pasar por cada punto de la imagen, el píxel pierde más o menos electrones, según que el píxel o punto sea más o menos brillante.

Esta pérdida de electrones se convierte en un impulso eléctrico que transmite por cada una emisión repetida.

En el receptor hay un haz de rayos catódicos, en el que también un píxel de electrones recibe una señal eléctrica. Transmite, con la misma velocidad exactísima que el píxel emisor, la imagen transmitida.

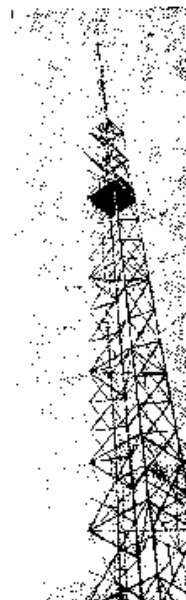
Cambia la señal recibida por el receptor es muy sencilla. La corriente de electrones es también más lenta, y el píxel marcado sobre la pantalla es muy brillante, y poca brillante cuando la señal recibida es débil, porque el píxel de electrones de la cámara para con un punto débilmente iluminado.

Queda un problema difícil: un haz de electrones parece ser tan sencillo. Lo complicado, por tanto, es la posición, de los dos píxeles de electrones, el emisor de la cámara y el del tubo receptor, que se mueven sobre la pantalla a una velocidad elevada, en que recorren 15.625 líneas por segundo.

Para conseguir esta coincidencia, al final de cada línea horizontal, se transmite un impulso eléctrico de sincronización horizontal, que corrige la posición del píxel en el receptor de manera que la coincidencia de posición de cada línea es perfecta.

Al llegar a la línea horizontal inferior, se ha acabado la imagen, y hay que pasar a la siguiente, es decir, los dos píxeles de electrones.

La cámara y el receptor deben estar a la primera línea horizontal siguiente, y de eso se encarga otro impulso de sincronismo, que es la misma sincronización vertical.



Reproducción en una línea de la línea de Televisión Española donde se transmiten los programas de televisión. Esta línea principal, de un tipo que se usa a menudo en la...

TELE radio

revista de actualidad

27 diciembre 1965-2 enero 1966



10 pesetas

Presencia de doce humoristas



Los Humoristas.



Los Humoristas.

Ricardo Arias y realizado por Ramón Díez y José María Quero, no ha de ser de música de baile, sino de música y humor, un espectáculo de variedades creado para compartir con los espectadores la alegría de las primeras horas del año y permitirles bailar al son de las músicas interpretadas, lo que es bien distinto de un programa concreto de música de baile. La presencia de doce humoristas explica claramente esta intención.

La elección de los actores ha respondido, en primer lugar, al deseo de llamar a los más representativos intérpretes españoles y, en segundo, a posibilitar un espectáculo moderno, como corresponde a la fecha; pero lo suficientemente valioso. Importa destacar la ilusión con que todos los cantantes y humoristas españoles se han brindado para actuar en este programa, la ilusión que todos han puesto en aparecer, cualquiera sea por unos minutos, en la pequeña pantalla para compartir con su público la alegría del año nuevo.

Las ausencias que pudieran darse no se deben sino a anteriores contratos que les impiden aparecer ante las cámaras en esa ocasión.

En líneas generales, el programa, por lo que su director nos ha contado, se ordenará en el sentido de introducir la actuación de un humorista con la de un cantante solista y la de un conjunto musical. Por eso andan aproximadamente equilibradas las listas de los nombres. Los humoristas, en breves actuaciones de cada uno, serán presentadores de quienes han de seguirles, al tiempo que desempeñarán, a su aire, la misión que les ha sido encomendada de facilitar el nuevo año a los telespectadores. Solistas y conjuntos interpretarán, cada uno, un par de obras de las que mayor popularidad han alcanzado durante el año en sus propias versiones.

Y hacia las tres de la madrugada, con el aire del nuevo año cruzado ya de risas y canciones, la despedida de un programa de bienvenida, el primero de 1968.

★ Al cierre de nuestra Revista han sido contratados para actuar en el programa de Fin de Año los siguientes intérpretes y humoristas, citados sin otro orden que el de esa misma contratación; es decir, puramente accidental.

Humoristas

TONY LEBLANC
ÁNGEL DE ANDRÉS
MANOLO GÓMEZ BUR
ANTONIO OZORES
CASEN
EMILIO LAGUNA
LALY SULDEVILLA
ZORI
SANTOS
LINA MORGAN
GUA
JUANITO NAVARRO

Solistas

LAURA
PSYCHEDELIC MISTRAL
ADRIANGELA
RAYMOND
MICHEL
JAIME MOREY
FEDERICO CABO
GELU
TATA TORRELLA
MINICHO
ROSALBA

Conjuntos

TITO MORA Y SU CONJUNTO
JOSÉ SIREX
LOS MUSTANG
RUDY VENTURA Y SU CONJUNTO
LOS TIFONES
LOS BOTNIS
MICKY Y LOS TONY
LOS BEATLES DE CADIZ
JUAN VICENTE TORREALBA Y SUS TONKALABEROS
LOS LILANEROS
LOS JOLLYS
LA TUNA DE FARMACIA
LOS TRES DE CASTILLA
LOS CUATRO DE LA TORRE

Dirección

RICARDO ARIAS

Realización

RAMÓN DÍEZ
JOSÉ MARÍA QUERO

Productores

RAMÓN MORENO
ALFREDO MONIS
JOSÉ MARÍA MORALES



Casen.



Laura.



Adriangela.

S.O.S. AVERIAS TV

La evolución de la televisión se ajusta

El establecimiento de la televisión

Galeón

TALLERES GALEÓN, S. A.

DELEGACIÓN MADRID: **LUIS MONTOLIU**
EDUARDO CL. CATEDRAL, M. 28014

TELEFONOS:

245 94 76

257 39 82

POVEDILLA, 5



AVERIAS TELEVISION

SERVICIO PERMANENTE:

246 25 24

ISABEL LA CATOLICA, 8

REPARACION TELEVISORES

TODAS MARCAS - RADIO - TOCADISCOS

ANTENAS INSTALADAS 1.ª CALIDAD: 750 PTAS.

¡RAPIDÍSIMO! - TEL. 2763557

SERVICIO PERMANENTE
A DOMICILIO Y FESTIVOS

Garantía A. Ingeniero

Duque de Sesto, 8 - MADRID



TELEVISION

LABORATORIO TECNICO



SERVICIO
PERMANENTE
DIARIO Y
FESTIVOS



MARCA REGISTRADA

DESCUENTOS ESPECIALES A LOS LECTORES DE TELE - RADIO

FARMACIA, 8 (entre Fuencarral y Hortaleza) - MADRID

REPARACIONES

RADIO - TRANSISTORES - TOCADISCOS

MAQUINAS DE ESCRIBIR
SUMAR Y CALCULAR
REPARACIONES
CONSERVACION
VENTAS



AVISOS: TELEFS. 231-62-72 y 231-74-99

TELE-URGENCIA

AVERIAS TELEVISION

SERVICIO TECNICO PERMANENTE A DOMICILIO

CLAUDIO COELLO, 48

Teléf. 225 85 00
246 34 64

PARTES DE EMISIÓN (AGA)

HOJA NUMERO, 5.- VIERNES 4 DE MARZO DE 1.966.-PR. SOIREMESA.

NUMERO, 6 .- TELEDIARIO (1ª EDICION)

6.- 15h -- ROTULOS.....
FILMADO..... 9'15"
DIRECTO..... 11'40"
FOTOGRAFÍAS..... 2'05"
T.- 15h 23' -- D.- 23' --

EMISORA..... MADRID.....
CANALES..... 2 y 4
CONTENIDO..... INFORMATIVO.....
REDACTOR JEFE.....
JEFE DE SERVICIO..... MARRERO.....
CASAS.....
REDACTORES..... BOFILL.....
M. ORS.....
BLANCO.....
ERQUICIA.....
GOMEZ TELLO.....
M. MEDINA.....
J.M. MARIN.....
MONT. CINE..... SATURIO.....
BRAVO.....
FOC. EN IMAGEN..... SANCHEZ.....
ZAMORA.....
LOC. EN OFF..... URIBARRI.....
OPACIO.....
REALIZADOR..... ALONSO.....
AYTE. REALIZADOR..... ARACIL.....
REGIDOR..... ERQUICIA.....

FILMACIONES NACIONALES.

INAUGURACIONES CUBILLO Y AREVALILLO.....
ALBERGUE DEL TORCAL.....
ENTRE COLAS: PRIMER VIERNES DEL MES EN
MEDINACELI.....
SINDICATO COMBUSTIBLE EN CORDOBA.....
ASAMBLEA PLENARIA SIND. GERONA.....

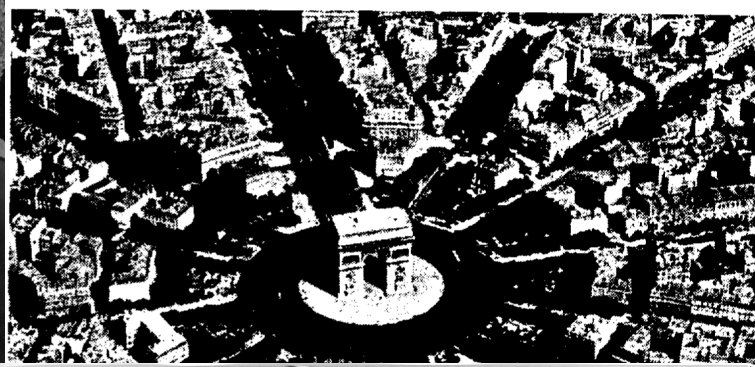
OPERADORES.

RAPADO.....
ROIG...
DE GREGOIO Y
TAPIADOR.....
POYATO.....
SANZ.....

FILMACIONES DEL EXTRANJERO.

COMENTARIO: GHANA Y GUINEA.....
(.) INUNDACIONES ARGENTINA.....
COLAS: VIETNAM.....
AVIADOR: "EMBAJADOR DE ISRAEL".....
PELEA EN MIAMI.....
DESFILE PIEL ESPAÑOLA EN ALEMANIA.....

(.) SE APRECIAN OBJETOS EXTRAÑOS EN EL FILM.



ds that will not go bad quickly. Many of his foods
ry, and some of them do not need further cooking.
ried fruit, jam, oatmeal for porridge, corn flakes,
ruits, and many other foods. He also sells soap,
that the grocer wears a special long coat over his
ad they protect the shopkeeper's clothes.



Algunos números:
Revista TeleRadio del 1 al 627 + 987 (especial 20 años emisiones)
+ 4 ejemplares cronología, teleclub y 10 y 50 aniversario TVE

Revistas analizadas

Concepto / Componente	Total
Ejemplares	628
Páginas completas	36.941
Imágenes recortadas	96.648
Total PNGs	133.589

Datos extraídos

Nombre de la Tabla	Registros	Tamaño (MB)
foto_clusters	57	0,04
personajes_historicos_TVE	576	7,20
content_advertising	14.693	1.600,00
content_editorial	36.496	2.600,00
content_faces	4.309	36,90
content_fotografias	1.383	17,70
content_fotografias_caras	4.865	22,20
content_image	96.727	2.600,00
content_others	12.691	714,00
content_radio_schedule	95.413	1.500,00
content_tv_schedule	127.519	2.100,00
video_faces	6.810	59,80
video_segments	7.405	140,90
videos	289	5,70
telerin_documents	763	10,30
telerin_schema_registry	5	0,07
Total Tablas CrateDB	410.001	11.414,81



TELERÍN

[Chat](#)[Imágenes](#)[Vídeo](#)

 **TELERÍN** Memoria de RTVE "Reviviendo el papel de 1958 a 1969"

[Conversación](#)[Nueva conversación](#)

🌈 ¡Hola! Soy **TELERÍN** 📺, tu asistente de búsqueda en el archivo histórico de **RTVE**: Teleradio (1958-1969), Fotografías y vídeos. ¿Qué te gustaría buscar?

[Ocultar capacidades](#)

¿Qué puedo hacer por ti?



Buscar programas de televisión o radio



Encontrar contenido por fecha, canal o año



Buscar artículos por tema



Hacer preguntas sobre el contenido de los programas

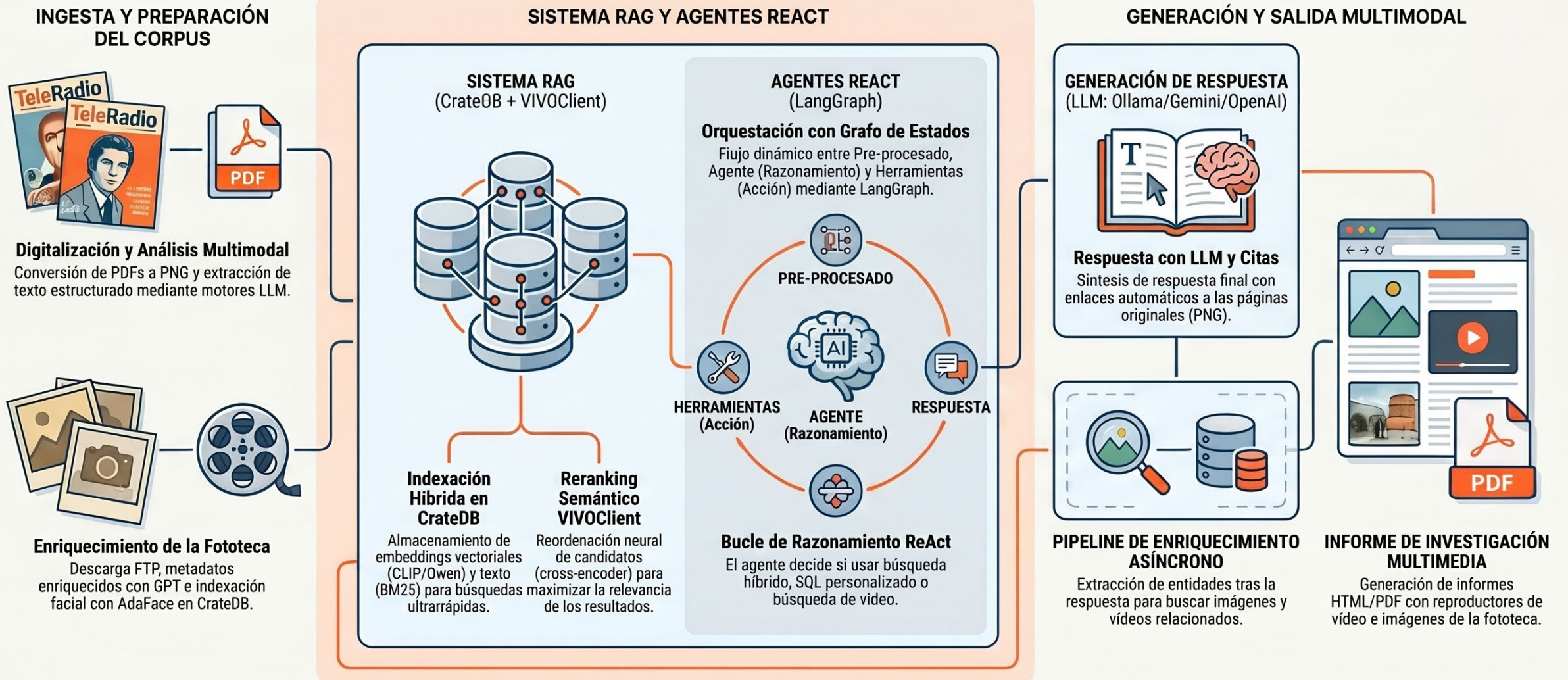


Buscar imágenes, fotografías y vídeos de personajes y programas



Proporcionarte información sobre la historia de la TV española

Arquitectura Inteligente de **TELERÍN RTVE**: Del Archivo Histórico al Asistente RAG



El Viaje del Dato en Telerín RTVE



Ingesta: del papel al vector

Convierte PDFs en texto estructurado y representaciones matemáticas (embeddings) almacenadas en CrateDB.



Guardrail y clasificación de intención

Filtra consultas fuera de dominio y decide instantáneamente si el usuario necesita datos, ayuda o un saludo.



Búsqueda híbrida y reordenación neuronal

Combina precisión léxica (BM25) con significado semántico (KNN) y un reranker que garantiza la máxima relevancia.



Respuesta chat con citas reales

Genera respuestas en lenguaje natural vinculadas directamente a las páginas originales de la revista mediante WebSockets.

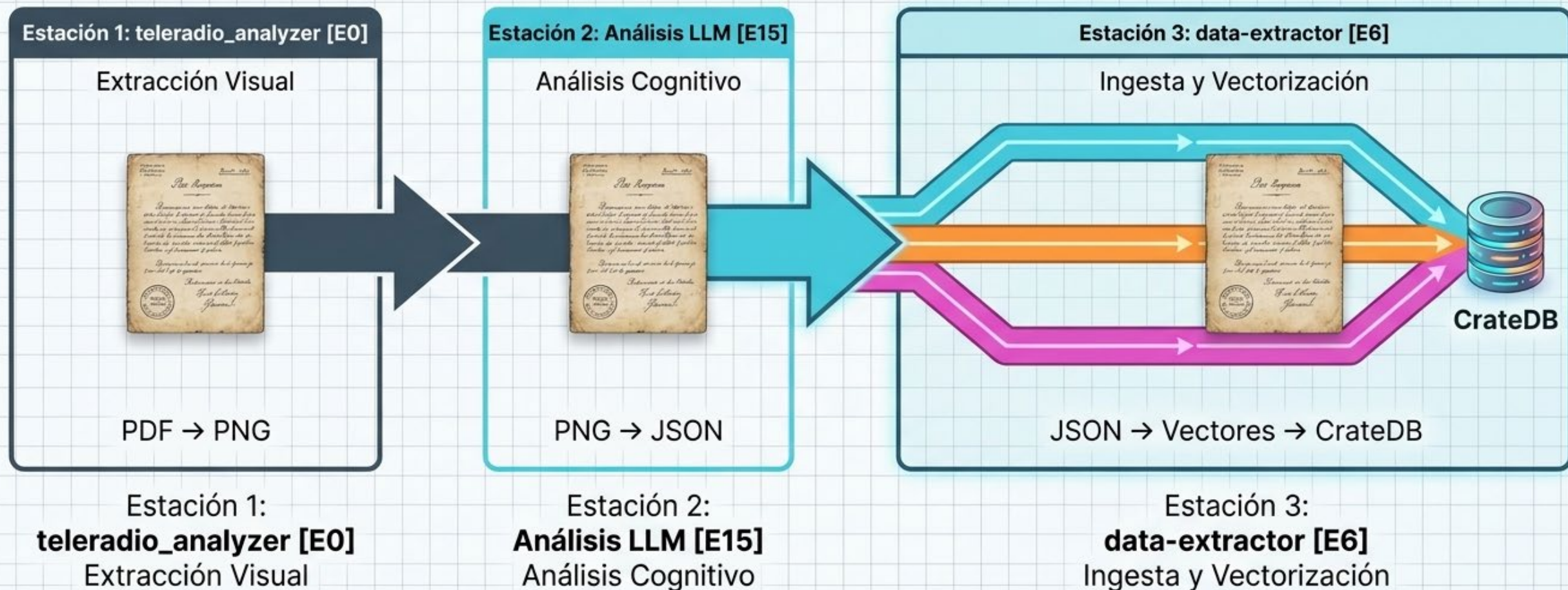


Enriquecimiento y reporte de investigación

Extrae entidades para buscar automáticamente fotos y vídeos relacionados, generando un informe final descargable.

Grafo de conocimiento: estructura datos conectando entidades con relaciones y propiedades. Búsquedas rápidas

Arquitectura Macro del Pipeline de Ingesta



El Puente JSON: Estructurando el Caos

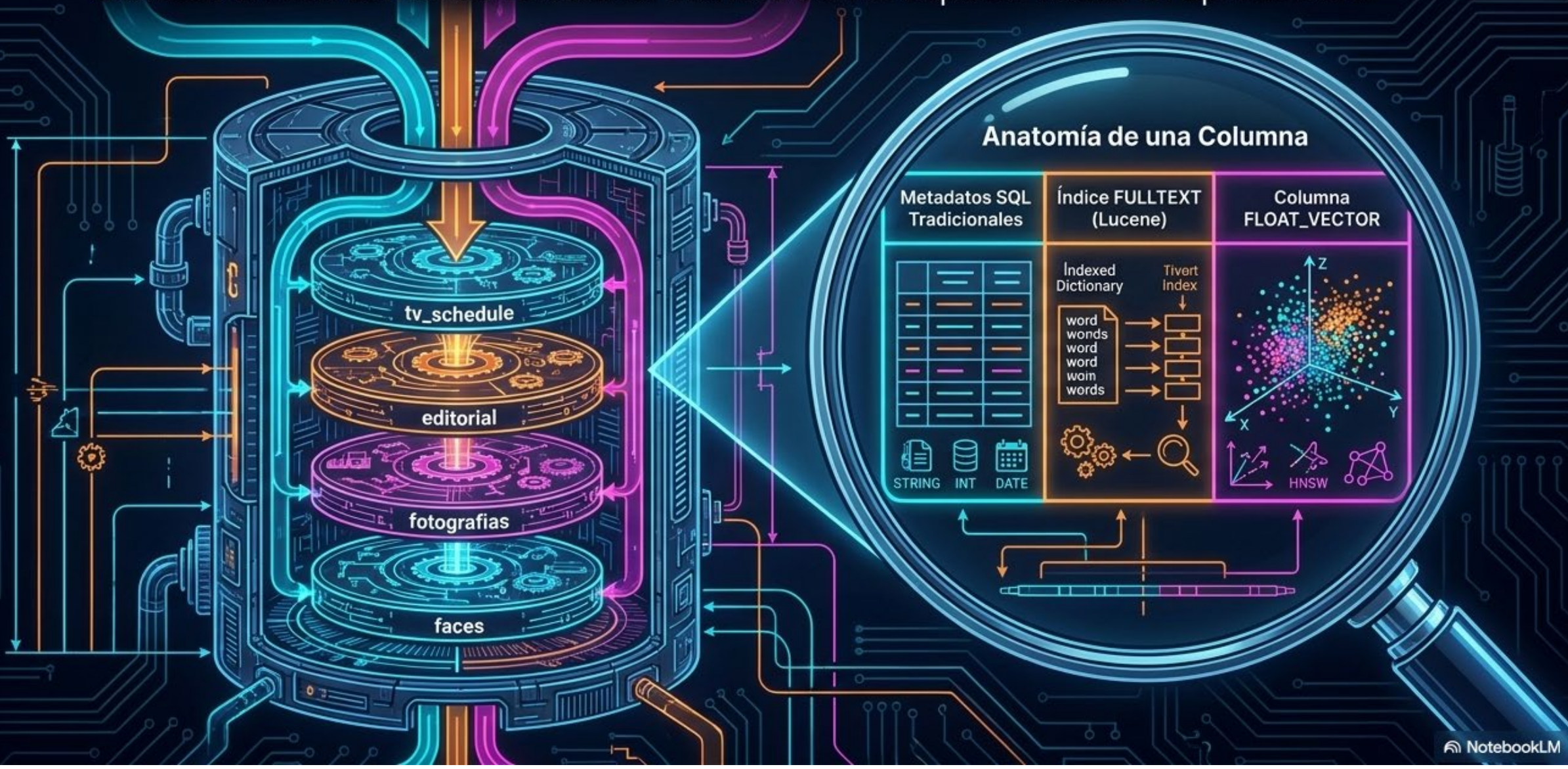
El formato intermedio universal. El LLM transforma el píxel en datos relacionales listos para la extracción vectorial.



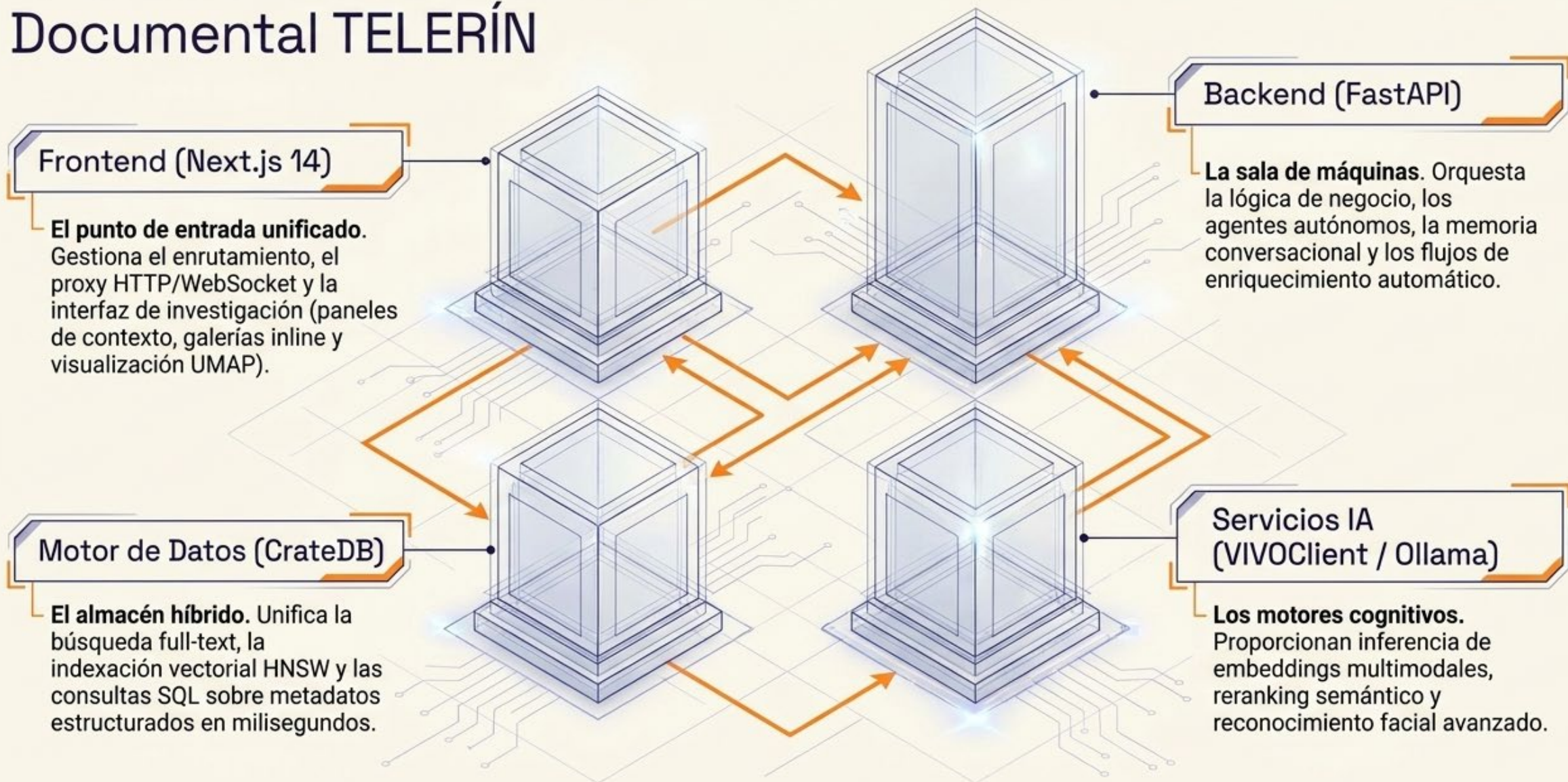
```
{
  "data": {
    "source_document": "TeleRadio_1960s.pdf",
    "processed_at": "2024-07-27T10:00:00Z",
    "extracted_content": {
      "tv_schedule": {
        "description": "Programación TV",
        "items": [
          { "time": "19:00", "program": "Noticias" },
          { "time": "20:00", "program": "Cine: Casablanca" }
        ]
      },
      "radio_schedule": {
        "description": "Radio",
        "items": [
          { "time": "08:00", "program": "Despertar Musical" }
        ]
      },
      "advertising": {
        "description": "Publicidad",
        "brand": "Telefunken",
        "product": "Televisor 21 Pulgadas"
      },
      "editorial": {
        "description": "Artículos",
        "headline": "El Futuro de la Televisión en Color",
        "author": "Carlos Martínez"
      }
    }
  }
}
```


Convergencia en CrateDB: El Almacén Multimodal

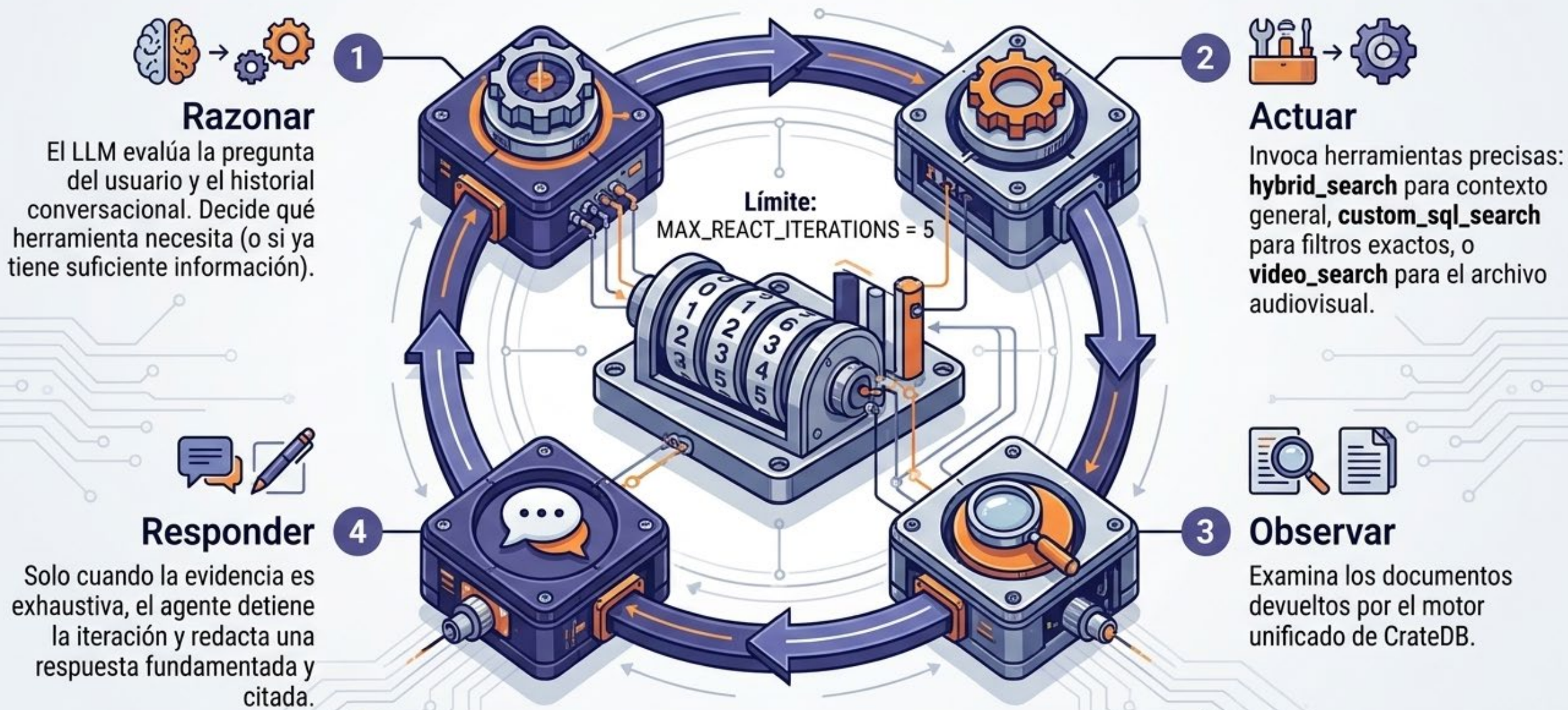
CrateDB unifica el mundo relacional estricto con el espacio vectorial aproximado.



Anatomía del Ecosistema Documental TELERÍN



Agentes Autónomos: El Paradigma ReAct impulsado por LangGraph



INTERFACE CHAT DE USUARIO

 TELERÍN ChatBot

 lleida@unizar.es  contraseña (salir)

Nueva conversación

☒ Seleccionar

HOY

Dame referencias de TeleRadio s...

1 turno

ESTA SEMANA

dame la programación de televisi...

1 turno

busca la programación de televisi...

1 turno

¿quién era Laurita y en qué pro...

1 turno

redacta un reportaje sobre las ser...

1 turno

redacta un reportaje sobre las ser...

1 turno

LangGraph · CrateDB · FastAPI

 OpenAI: gpt-5.3-chat-latest

 Reranker: qwen3rerank

 Memoria de RTVE "Reviviendo el papel de 1958 a 1969"

Chat

Imágenes

Video

?

Conversación

Generar informe

Nueva conversación

¿quién era Laurita y en qué programas participó en los años 60?

Laurita (nombre artístico de **Laura Valenzuela**) fue una de las figuras más populares de la televisión española en sus primeros años: **locutora y presentadora de TVE desde comienzos de los años 50**, además de actriz de cine y antigua modelo de alta costura (Documento 4  Revista#002, 1958-01-06, Página 3; Documento 2  Revista#154, 1960-12-05, Página 17). La propia actriz reconocía que la televisión fue el "trampolín" que impulsó su carrera y le dio gran popularidad (Documento 2  Revista#154, 1960-12-05, Página 17).

En los años 60, ya consolidada como una de las caras más conocidas de TVE, participó en varios programas destacados, especialmente de variedades y musicales:

"Cantamos contigo" (1968): programa musical con participación del público mediante votaciones telefónicas, presentado por Laurita Valenzuela (Documento 5  Revista#551, 1968-07-15, Página 29; Documento 8  Revista#560, 1968-09-18, Página 47).

"Galas del sábado" (finales de los 60): programa de variedades muy popular que presentaba junto a Joaquín Prat (Documento 1  Revista#593, 1969-05-10, Página 59; Documento 13  Revista#578, 1969-01-25, Página 57).

Además, en esta etapa seguía siendo una figura central de la televisión española, elegida incluso para eventos de gran relevancia como la presentación del Festival de Eurovisión celebrado en Madrid en 1969 (Documento 9  Revista#576, 1969-01-06, Página 11).

En resumen, Laurita Valenzuela fue una **presentadora clave en la historia de TVE**, que en los años 60 destacó sobre todo en programas musicales y de entretenimiento de gran audiencia, consolidando su enorme popularidad en España.

Si quieres, puedo detallarte todos sus programas desde los años 50 o profundizar en alguno en concreto.

— continúa aquí —

P.ej: ¿quién presentaba 'Caras Nuevas'? / busca fotos de Raphael / muéstrame vídeos de deportes





TELERÍN — Archivo histórico TeleRadio RTVE · 1958–1969



LOS FAMOSOS "TELEOPINAN"

- "La TV ha sido el trampolín que me ha empujado hacia arriba."
- "En los estudios de TV he aprendido a moverme con soltura ante las cámaras."

UN TEMOR:

"ME DA MIEDO QUE PASEN LOS AÑOS Y NO TENGA TIEMPO DE LLEGAR HASTA DONDE QUIERO."

UN DESEO:

"ME GUSTARIA COMPRARME CIENTOS DE PARES DE ZAPATOS."

UNA DECISION:

"SE ELIGE UN HORIZONTE Y HAY QUE SEGUIR HASTA EL FINAL."

Y UNA CONFIDENCIA:

"RECONOZCO QUE, A VECES, TENGO MALA IDEA."

LAURITA VALENZUELA

UNA MODELO DE ALTA COSTURA QUE HA TRIUNFADO EN EL CINE

A Laurita Valenzuela le hubiera gustado ser cirujano, porque le parece estupendo eso de poder cortar pedacitos a un persona y, después, con una simple aguja y un hilo, dejarlo todo como estaba.

Pero Laurita Valenzuela, a la hora de elegir su camino, pensó que el hacerse cirujano resultaba una cosa muy pesada, ya que había que pasarse más de ocho años en una Facultad, aprendiendo por dónde había que cortar, para no organizar en cada operación una catástrofe de ayuda.

Así que Laurita Valenzuela, que es una mujer con los pies bien pegados a la tierra, se miró al espejo a ver qué pasaba, se encontró francamente buena y comprendió que si se dedicaba a actriz de cine tenía media carrera aprobada por delante.

De modo que Laurita Valenzuela se puso a andar sin vacilaciones por la senda de la fama, aun a sabiendas de que, a pesar de todo, el asunto no iba a ser tan fácil, porque la media carrera que le quedaba aún por estudiar tenía lo suyo.

Sin embargo, Laurita Valenzuela ha conseguido aprenderse casi toda la asignatura y ahí la tienen ustedes dándole los primeros mordiscos a ese pan del triunfo, un poquito amargo y un poquito dulce.

Porque a Laurita Valenzuela, que tiene ideas simples y fundamentales de las cosas, lo que de verdad le hubiera gustado, en el fondo de sí misma, es haberse dejado de zarandajas y ser una sencilla y elemental madre de familia.

—Con muchos niños—me dice, y a Laurita Valenzuela, al decir esto, se le va la mirada muy lejos, quizá hasta ese rincón del recuerdo donde se le quedaron paradas sus primeras ilusiones.

—Muchos niños—me repite—. Pero no puedo separar a la mujer de la actriz, porque yo soy "yo y todas mis circunstancias". Y una actriz, cuando de verdad tiene vocación, se ve obligada a sacrificarlo todo a la carrera.

Yo le pregunto si merece la pena, y Laurita Valenzuela, afilando los labios y mirándose de golpe a los ojos, me responde:

—Se elige un horizonte y hay que seguir hasta el final.

—Usted ha luchado duramente.

—Sí. Pero lo mejor que he hecho en mi vida ha sido luchar con paciencia. Estoy orgullosa de no haber flaqueado en los momentos malos.

Laurita Valenzuela ha sido modelo de alta costura y presentadora de televisión.

—La televisión—me explica—ha sido el trampolín que me ha lanzado hacia arriba. Me ha dado popularidad y, sobre todo, he aprendido a moverme con soltura delante de las cámaras.

—¿Qué le queda por aprender?

—Tanto, que me da miedo que pasen los años y no tenga tiempo de llegar hasta donde quiero.

—¿Hasta dónde?

—Creo que ni yo misma lo sé. Todo lo que se sueña es, siempre, algo inconcreto. Sólo resulta posible concretar en las cosas intrascendentes. Por ejemplo, sé que me gustaría comprarme cien pares de zapatos.

—¿Y ¿por qué no se los compra, Laurita?

—Porque (todavía no he ganado bastante dinero para ello).

—¿Está en eso el tope de su felicidad?

—El dinero no me interesa demasiado. Lo importante es sentirme de acuerdo conmigo. Aunque no suelo conseguirlo, porque no soy muy inteligente. Pero, en cambio, tengo una gran fuerza de voluntad y he aprendido a encajar las decepciones.

Laurita Valenzuela sabe ponerse frente a sí misma y mirarse a los ojos:

—También reconozco que, a veces, tengo mala idea. Pero creo que, en el fondo, soy buena chica, sin ninguna complicación. Me gusta fregar el suelo, porque me encanta ver cómo va quedando limpio. Leo novelas policíacas y me da mucha rabia no adivinar nunca quién es el asesino. Y tengo mis virtudes y mis defectos, como todo el mundo.

Me dice una virtud:

—En mí, el corazón vence constantemente a la cabeza.

Y, luego, me dice un defecto:

—No soy una mujer muy equilibrada.

Además, Laurita Valenzuela se come las uñas.

Lo he visto yo.

Fernando ALBERT

(Fotos Montoya.)



Búsqueda Híbrida y Multimodal: Recuperación sin Puntos Ciegos

Converging Dual Pillar

Pilar Léxico (BM25)

Búsqueda de coincidencia exacta sobre índices invertidos.

Vital para nombres propios, fechas y títulos específicos.

Supera al TF-IDF clásico evitando la saturación de términos.



Pilar Semántico (HNSW)

Búsqueda vectorial mediante embeddings (1024-dim).

Resuelve el desajuste de vocabulario capturando conceptos afines (ej. 'tele' vs 'televisión').

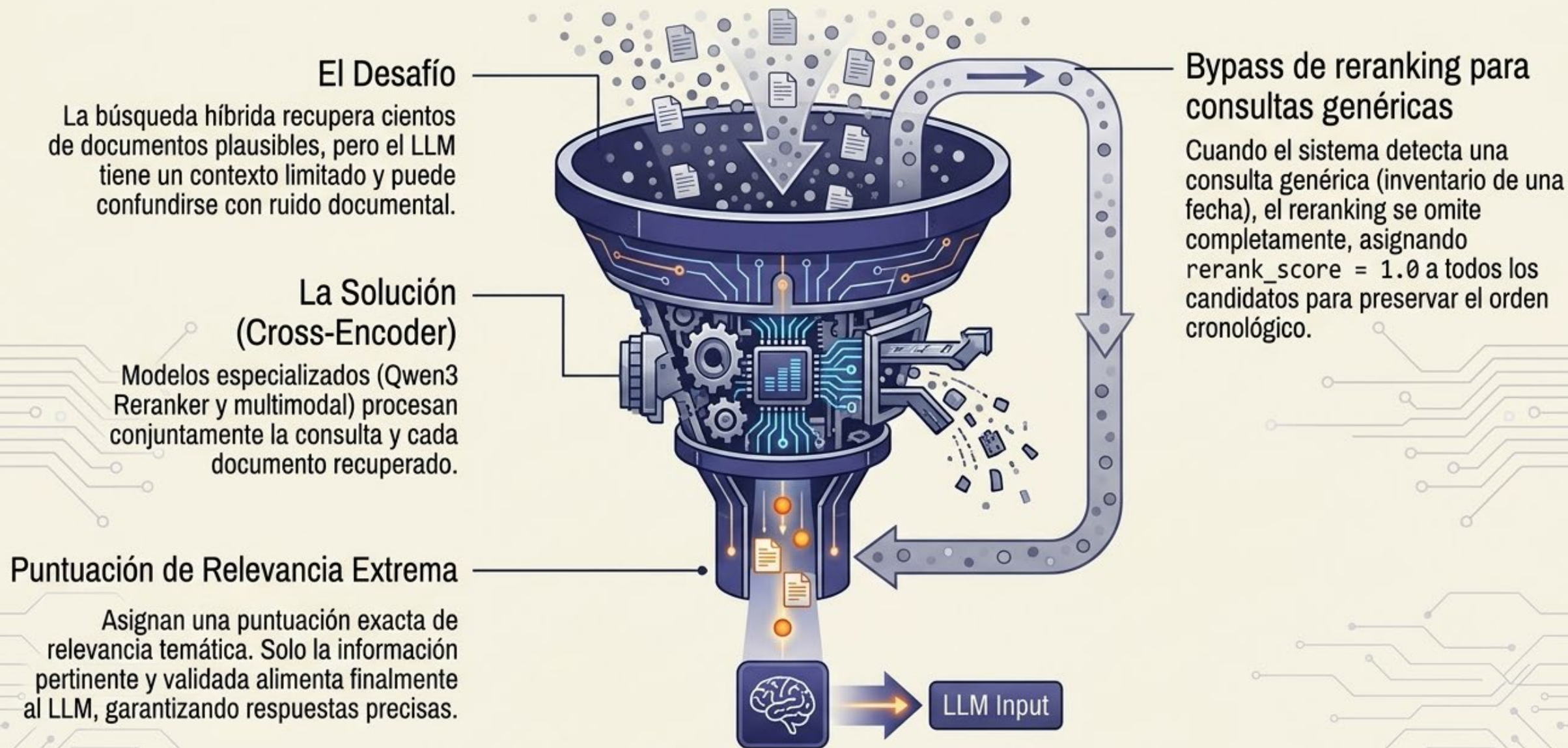


Fusión en CrateDB

Ambos motores se unifican en milisegundos. A esto se suma la búsqueda multimodal mediante embeddings CLIP para imágenes y AdaFace (IR-50) para reconocimiento de rostros históricos, mitigando la brecha de dominio (domain gap) del material en blanco y negro.

CrateDB integra motor Lucene nativo para el pilar léxico y utiliza HNSW con su estructura jerárquica multicapa para el pilar semántico, todo sobre una arquitectura share-nothing distribuida.

Reranking Neuronal: El Antídoto contra la Alucinación Documental



BLINDAJE DE TELERÍN RTVE: LAS 3 CAPAS DE SEGURIDAD (GUARDRAILS)

CAPA 1: PRE-CHECK TÉCNICO CON REGEX

Filtro instantáneo en guardrail.py que bloquea ataques DoS, inyecciones de prompt y secuestro de roles sin coste de LLM.

CAPA 2: FILTRO SEMÁNTICO CON LLM (GEMINI)

Análisis profundo de intención para evitar contenido político, daños reputacionales o extrapolar datos al presente.

CAPA 3: RESTRICCIÓN POR PROMPT DE GENERACIÓN

Instrucciones rígidas en tools.py que prohíben el uso de conocimiento externo al corpus documental de 1958-1969.



UMAP: Exploración Espacial de la Fototeca y Archivo Audiovisual

Más allá de las Palabras

Proyección algorítmica de reduccionalidad (UMAP) que mapea miles de fotografías y vídeos en un plano 2D interactivo.

Descubrimiento Visual

Permite a los documentalistas identificar clústeres temáticos ocultos (ej. retratos, exteriores, estudios de grabación) basándose puramente en la similitud visual capturada por embeddings CLIP, eliminando la dependencia de metadatos manuales incompletos.



Del Diálogo al Dossier: Informes de Investigación Automatizados



Un proceso asíncrono y no bloqueante analiza la respuesta final del agente y extrae personas, programas, canales y conceptos clave.

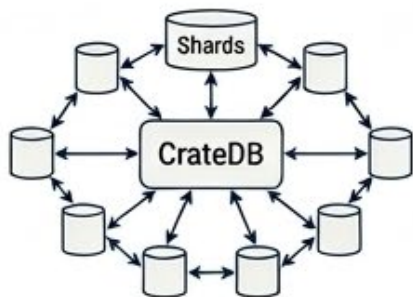
El sistema peina autónomamente la fototeca y el archivo de vídeo buscando coincidencias visuales para esas entidades, aplicando un filtrado estricto (score ≥ 0.5).

El Artefacto Final
Generación de un informe descargable (HTML/PDF) con tarjetas documentales, badges de puntuación de relevancia, y vídeo inline. Trabajo de días consolidado en segundos.

La Arquitectura: Cuatro Pilares del Motor Telerín

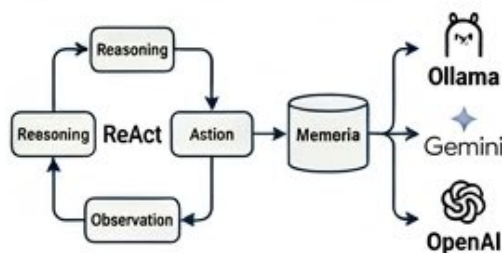
1. Base de Datos Híbrida (CrateDB)

- Motor Apache Lucene unificado.
- Fusión nativa de búsqueda léxica (FULLTEXT) y vectorial (HNSW).
- Arquitectura distribuida share-nothing.



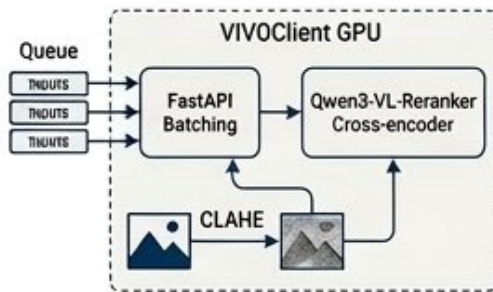
2. Orquestación Agéntica (LangGraph)

- Ciclo ReAct con grafo de estado finito.
- Memoria conversacional persistente con detección de follow-up.
- Agnóstico al backend LLM (Ollama, Gemini, OpenAI).



3. Reranking Neural (VIVOCient GPU)

- Batching dinámico en FastAPI.
- Cross-encoder multimodal (Qwen3-VL-Reranker) para scoring cruzado.
- Procesamiento facial adaptativo (CLAHE) para reducir el domain gap.

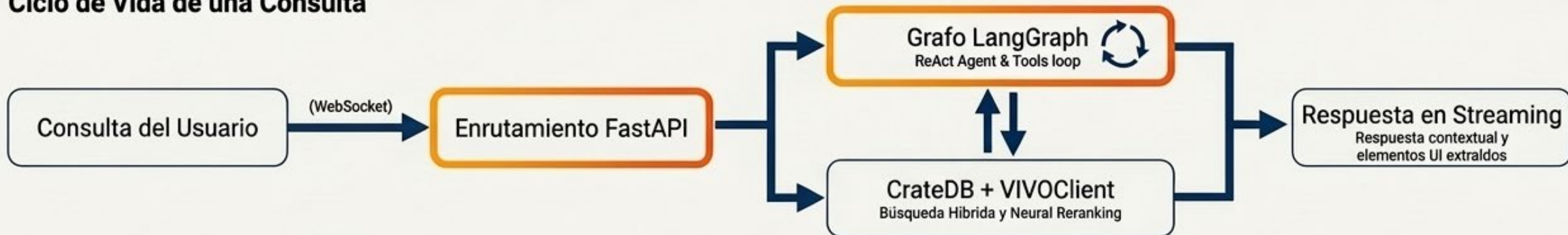


4. Guardarail y Generación de Respuesta

- Filtrado de seguridad y validación de respuestas.
- Post-procesamiento para consistencia y formato.
- Generación de respuestas seguras y contextualmente relevantes.



Ciclo de Vida de una Consulta



DEMO TELERÍN

