#### Instituto de Investigación en Ingeniería de Aragón (I3A), Universidad de Zaragoza



HOME RESEARCH V PROIECTS V TEAM PUBLICATIONS DEMOS NEWS INTRANET WIKI



Reseachers





Alfonso Ortega











**Antonio Miguel** 

Research lines





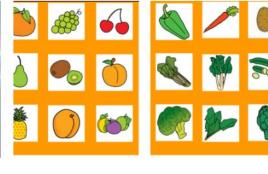
Speaker/Face Verification and identification Language identification Speaker/Face Diarization Acoustic event detection & classification Speech enhancement and audio quality assessment

Leer más →



Multimedia Content Retrieval & Indexing

Classification and segmentation of audiovisual documents Analysis and retrieval of audiovisual content Multimodal person and event recognition



#### Aumentative and Alternative Communication & Paralinguistics

Automatic assessment of pathological speech Pictograms Training assistant

Leer más →



Technologies for Human-Machine Interaction

Robust voice modeling and processing Automatic speech recognition Natural language processing

https://vivolab.i3a.es/

https://i3a.unizar.es/

https://catedrartve.unizar.es/

Más de 25 años investigado en tecnologías y sistemas de IA aplicadas al audio, el habla y el lenguaje. Continua transferencia al sector productivo.



# Cátedra RTVE - UNIZAR

Fecha de creación: 10 de Julio de 2017

# CONTRA ALBA NATVE

#### Objetivo:

realización de actividades de formación, investigación, estudio y divulgación en el área de las Tecnologías de la Información y de las Comunicaciones relacionadas con el Big Data aplicado al análisis de los contenidos audiovisuales y sonoros.

#### **Actividades:**

#### Formación y divulgación:

Seminarios, charlas, conferencias y cursos extraordinarios para acercar la IA a áreas no tecnológicas.

#### Investigación y estudio:

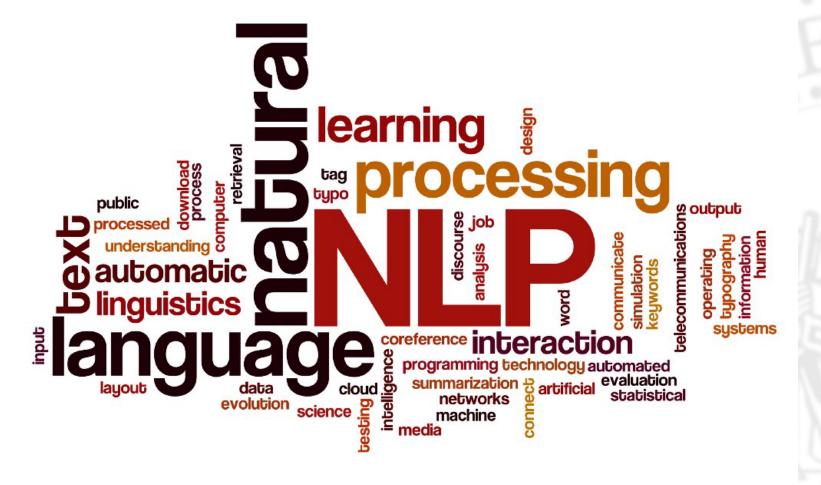
Seguimiento del estado del arte en tecnologías de IA asociadas al análisis y metadatado de contenidos audiovisuales.

Investigación en espacios semánticos multimodales para la análisis, búsqueda y recuperación de contenido audiovisual y sonoro.

#### Retos tecnológicos:

Congreso IberSpeech, Red Temática en Tecnologías del Habla: Retos 2018, 2020, 2022, 2024 Voz a Texto, Diarización y asignación de identidad, Diarización multimodal, Sincronización subtítulos

# La Inteligencia Artificial al rescate: buscando una aguja en el pajar de los archivos audiovisuales



**Speech and Language Processing (3rd ed. draft)** 

**Dan Jurafsky and James H. Martin** 

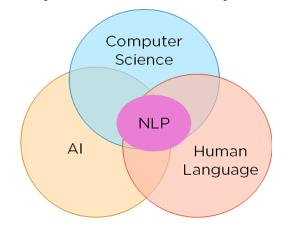
https://web.stanford.edu/~jurafsky/slp3/

# ¿Qué es el procesamiento del lenguaje natural?

✓ Término genérico que abarca todo aquello que permite a las máquinas procesar el lenguaje humano tanto en forma escrita, verbal, o visual.

# ¿Porqué es importante el procesamiento del lenguaje natural?

✓ Componente/Capacidad fundamental de los sistemas de IA.



#### Capacidades de un sistema de IA

- Percepción
- Aprendizaje
- Representación del conocimiento
- Razonamiento



# Comprensión de la información

El objetivo final es comprender el mensaje codificado en el lenguaje.

### Comprender

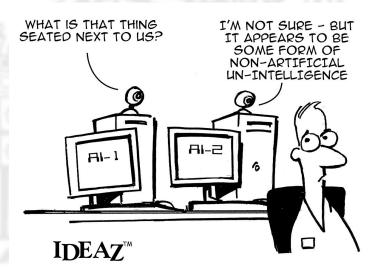
Percibir y tener una idea clara de lo que se dice, se hace o sucede o descubrir el sentido profundo de algo.

Implica entender conceptos y procesos para poder explicarlos y describirlos

de forma adecuada.



Nos proporciona herramientas para representar el *conocimiento* 



#### Representación del conocimiento

formalismos de representación del conocimiento:

Redes Semánticas (relaciones semánticas entre objetos en una red)

Oso

Tiene

Es un

Pelo

Mamífero

Frames(marcos), colección de datos estructurados slots (propiedades) & fillers (valores) & Métodos

#### Reglas

Si < condición>

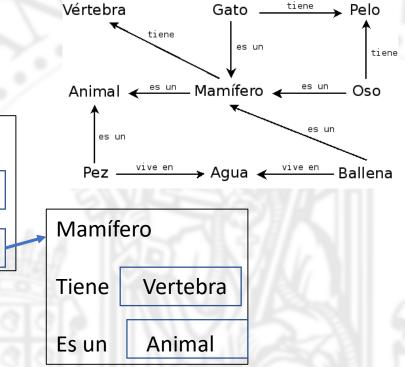
entonces < conclusión>

#### Ontologías

esquemas de representación del conocimiento basados en redes semánticas o marcos.

#### Espacios semánticos

Representación vectorial capaz de capturar el significado

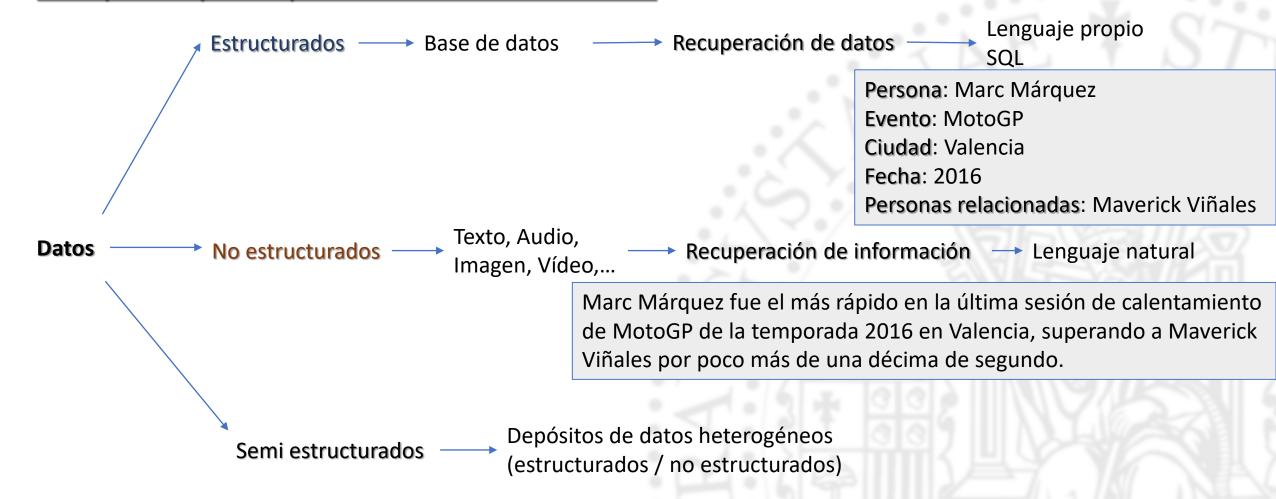


Procesamiento del Lenguaje Natural Image-to-text Text-to-Image ¿Qué tareas podemos hacer? Data-to-Text Part Of Speech Text-to-Speech Stance Detection Text-to-Data and Text Preprocessing Fake News and Hate Speech Detection Feature Extraction **Text Reasoning** SMILE! OUR AI WILL TAKE A PICTURE OF Classification YOUR FACE AND WILL KNOW EXACTLY WHAT **NLP Tasks** YOU WOULD LIKE FOR DINNER TONIGHT. Chatbots formation Retrieva Text-to-Text

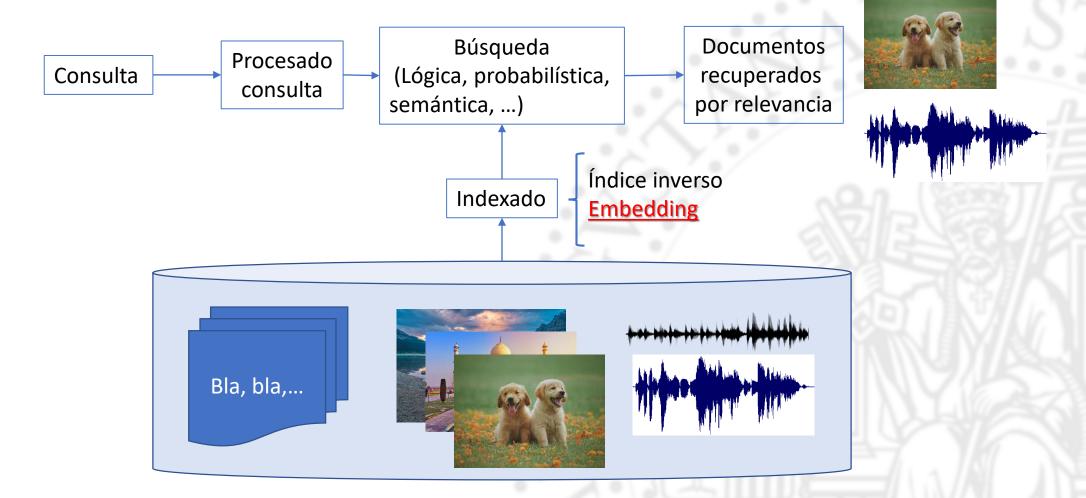
Image by NLPlanet

KDNUGGETS.COM

#### Búsqueda y recuperación de información



# Búsqueda y recuperación de información

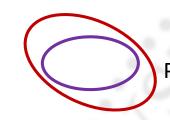


Medida de prestaciones en búsquedas y recuperación de información:

Precisión y Exhaustividad/Sensibilidad (Precision/Recall), Valor-F (F-score)

Precisión: Probabilidad de que un documento recuperado sea relevante.

$$P = \frac{|\{documentos\ relevantes\ y\ recuperados\}|}{|\{documentos\ recuperados\}|}$$



|{Documentos relevantes}|

|{Documentos recuperados}|



Exhaustividad: Probabilidad de que un documento relevante sea recuperado en una búsqueda.

Colección documentos

$$R = \frac{|\{documentos\ relevantes\ y\ recuperados\}|}{|\{documentos\ relevantes\}|}$$



Valor-F: Valor único ponderado de la Precisión y la Exhaustividad.

$$F = 2 \frac{Precisión \times Exhaustividad}{Precisión + Exhaustividad}$$

$$F_{\beta} = (1 + \beta^2) \frac{Precisión \times Exhaustividad}{\beta^2 \times Precisión + Exhaustividad}$$

Búsqueda semántica: ¿cómo representamos las palabras/documentos en una máquina?

La máquina trabaja con números, .....

.... luego debemos trasformar las palabras/documentos a números

#### Opción 1.

Utilizamos un código numérico, p.e. las numeramos de forma correlativa

netflix	1	
encontrar	2	
juego	3	
calamar	4	
nuevo	5	
fenómeno	6	
mundial	7	
ni	8	
siquiera	9	
propio	10	
plataforma	11	
contar	12	

¿tiene algún significado el valor numérico?

¿Podemos calcular la proximidad semántica?, ¿tiene sentido?

#### Opción 2.

Reflexionemos, ¿qué buscamos?

- ✓ Queremos representar el significado de unidades lingüísticas (palabras/documentos)
- ✓ Queremos definir una medida de similitud semántica entre unidades
- ✓ Queremos que sea una representación numérica densa: "embeddings"

#### En definitiva:

Un espacio matemático de representación compacto donde cada palabra/documentos es un vector numérico de valores reales. Estos vectores capturan las relaciones semánticas entre las palabras/documentos, de manera que palabras/documentos similares en significado tienen representaciones vectoriales cercanas en el espacio. Estos vectores los denominamos "Embeddings" y constituyen lo que denominaremos <u>espacio semántico</u>

¿cómo lo construimos?

Semántica distribucional / Modelos de lenguaje



¿Cómo conocemos el significado de una palabra?

John Rupert Firth, "You shall know a word by the Company it keeps"

Ludwig Wittgenstein, "The meaning of a word is its use in language"

Hay una botella de *Belikin* sobre la mesa A todo el mundo le gusta la *Belikin* No bebas *Belikin* si tienes que conducir La *Belikin* se fabrica con granos de cebada germinada

¿qué podemos deducir sobre la palabra *Belikin*?

Miramos las palabras que acompañan

Buscamos la similitud semántica con otras palabras ya conocidas

... y deducimos que la *Belikin* debe ser una bebida similar a...





# Semántica distribucional

Matriz términos-documentos: buscar documentos similares

Cada fila representa una palabra del vocabulario

Cada columna representa un documento

Cada celda: tf-idf:

 $\mathsf{tf}_{t,d}$ :Frecuencia de aparición de un término t en el documento d

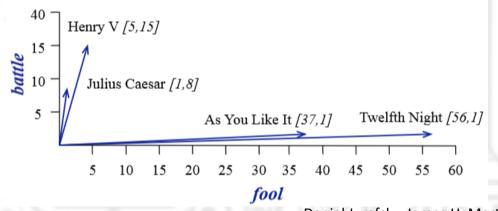
 $\mathrm{idf}_{t,D}$ : inversa frecuencia de aparición término t en el conjunto de documentos D

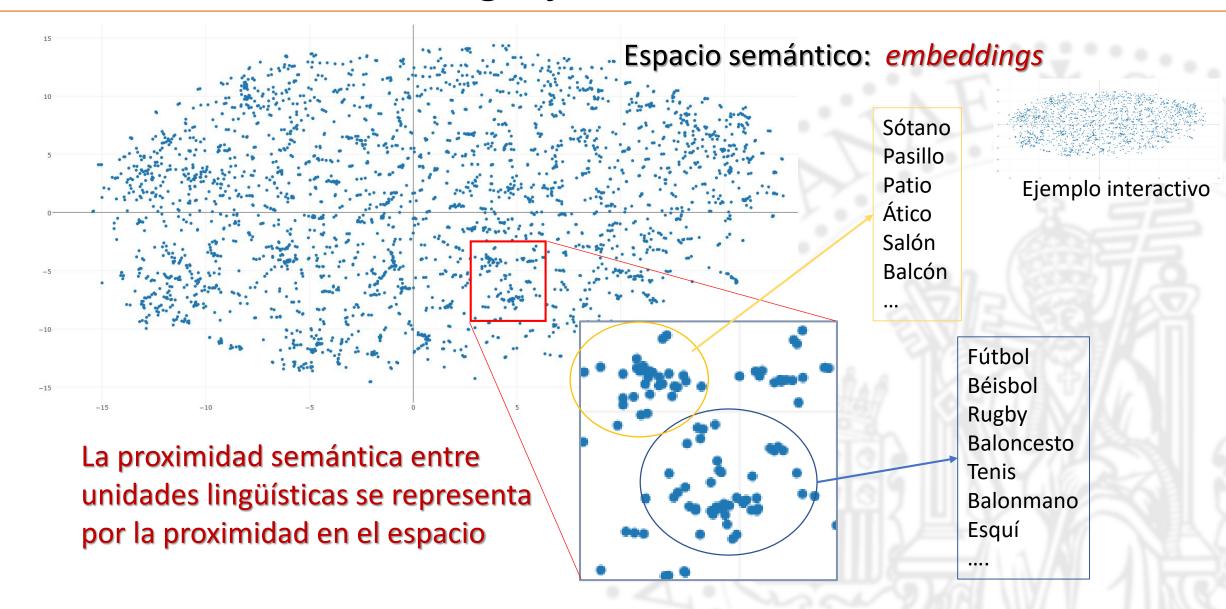
Cada documento: un vector de cuentas de dimension V (tamaño vocabulario)

#### Dos documentos son similares si sus vectores son similares

As You Like It	Twelfth Night	Julius Caesar	Henry V
$\Box$	$\Box$	8	[15]
2	2	12	36
37	58	1	5
5	117	0	0
	П		$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\begin{bmatrix} 8 \\ 12 \end{bmatrix}$

Figure 15.2 The term-document matrix for four words in four Shakespeare plays. The red boxes show that each document is represented as a column vector of length four.





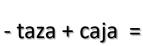
#### Relaciones semánticas

vector[Reina] = vector[Rey] - vector[Hombre] + vector[Mujer]

# Rey Reina Mujer Hombre



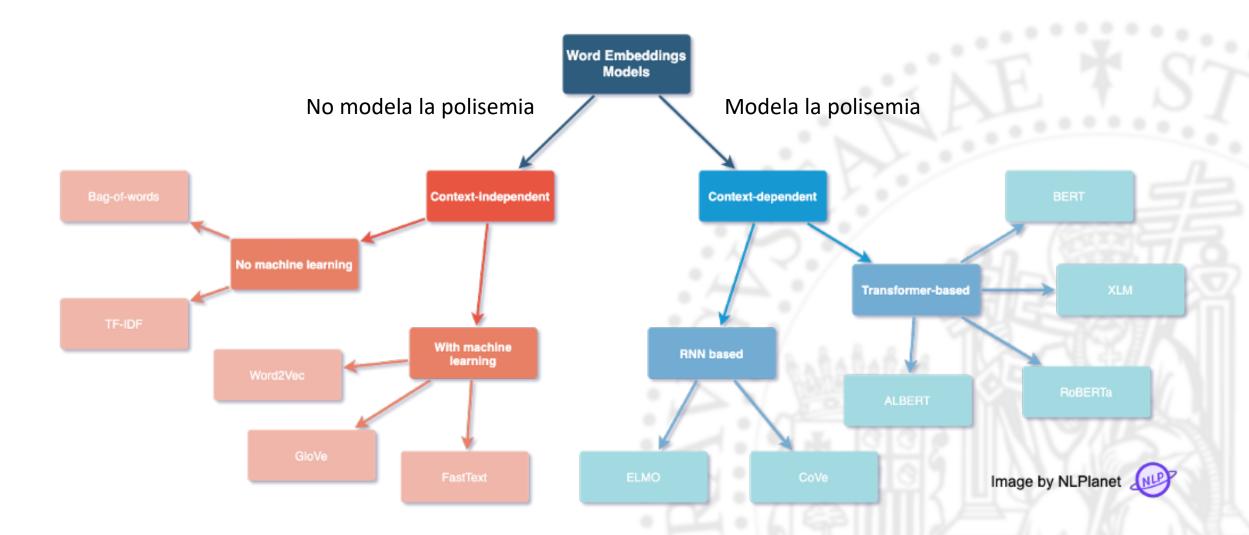




#### Imágenes próximas



(Kiros, Salakhutdinov, Zemel, TACL 2015)





# Modelos de lenguaje

- Representación matemática de la utilización de la lengua
- Comprender, generar y predecir el texto en función del contexto Ejemplos de modelos:
  - Bolsa de palabras (Bag-of-Words): contar frecuencia de aparición de las palabras en un texto, no se tiene en cuenta el orden o la estructura gramatical. Clasificación de textos y análisis de sentimiento.
  - Modelos basados en reglas: utilizar reglas gramaticales de la lengua para analizar y comprender el texto. Reglas específicas según la tarea. Uso en análisis morfológico
  - Modelos estadísticos: modelar la probabilidad de ocurrencia de palabras o secuencias de palabras en un texto. Uso en reconocimiento automático del habla, traducción automática y generación de texto.
  - <u>Modelos neuronales</u>: aprender patrones y representaciones del lenguaje.
     Estado del arte actual. Útil para casi todo.

#### Modelos de Lenguaje de N-gramas

La probabilidad de la palabra actual solo depende de las N-1 anteriores

```
P(unidas | la,alta,comisionada,de,las,naciones) ≈ P(unidas) (Unigrama)
o
P(unidas | la,alta,comisionada,de,las,naciones) ≈ P(unidas | naciones) (Bigrama)
o
P(unidas | la,alta,comisionada,de,las,naciones) ≈ P(unidas | las,naciones) (Trigrama)
```

#### Frases generadas con un UNIGRAMA

norte justamente se cualquier líneas por determinadas la llevará venía junto víctima creará hacen humos capacidad puede económica carnaval para enterrado fundamento de acepto argentinos mareante publicado nova diamantes son horteras hamalainen añadir bin sistemáticamente misteriosas presupuestario conocido además preside mejor intermón esta jessica gravamen dramas que así grecia imputada ocurrió veremos delito personajes orografía escandalizados o e espero establece establecieron pero cuatro continúan pp aplazada panamá quedaban hubiera tendremos proyectos por dato dando hansch santa cuánto intactos aportan del pertenece

#### Frases generadas de forma automática con el modelo de bigramas

recientemente en toda la champions en el chelsea y padre fueron de mediados del gobierno de hollywood con consecuencias si gana las listas para mi mujer y aplaudido antes de uno nuevos indicios de practicantes pero siguen evacuados ya lo que vamos hasta la posibilidad de la feroz en resaltar la causa ya lo desmienten con las filtraciones grandes se podría alargarse tiro en la ampliación de los mossos en el más remedio porque en especial esfuerzo y deslizamientos de la ley navarra de murcia en este poeta escritor que la desaprueba la guardia entra dentro de información el sesenta por último día en calidad extraordinaria atento y qué ser humano

#### Frases generadas de forma automática con el modelo de trigramas

señor fischler presentará ahora es una de ellas estaba ingresado por una posible sanción de nueve grados más esperados como la afirmación de que dice debe cesar económica que el tribunal así lo haya clarificado todavía no han logrado calar una cuenta en matar uno de los reclusos etarras pero por encima de todo lo que sucede hoy



Generación de texto a partir de un vocabulario de caracteres (2016)

Las temperaturas bajarán en el norte y en el sur, lloviznas en el este y en Galicia

y en Asturias, con algunas nubes más por la tarde, mientras que en las Islas Baleares se mantienen los cielos básicamente despejados. Aquí tenemos las nubes que se extienden por muy poco espacios de las próximas horas. De hecho, esta es la situación de cara a la jornada del domingo, continuaremos hablando de una situación marcada por la estabilidad. Por tanto también con intervalos de nubes altas y medias en la costa mediterránea. En general tiempo estable en el centro y en el oeste del país. Esto se va a notar más porque el viento es de componente norte que va a traer un tiempo muy parecido al de hoy. En el resto del país con cielos despejados, pero también con nieblas en el norte de las islas de mayor relieve.



#### ChatGPT (2023)

Las temperaturas bajarán en el norte y en el sur, anticipando un cambio en el clima que se avecina en las próximas semanas. Se espera que las regiones septentrionales experimenten un descenso gradual en las temperaturas, con mínimas que podrían llegar a alcanzar valores cercanos a los 10 grados Celsius. En el sur, las temperaturas también mostrarán una tendencia a la baja, con valores máximos que podrían oscilar alrededor de los 20 grados Celsius.

Por otro lado, se pronostican lloviznas dispersas en el este y en Galicia. Estas precipitaciones serán de intensidad moderada y se extenderán durante varios días, proporcionando alivio a las zonas que han experimentado periodos de sequía. Se recomienda a los residentes tomar precauciones y estar preparados para condiciones climáticas húmedas, así como a los conductores adaptar su conducción a las carreteras mojadas.

En el resto del país, se espera un clima mayormente estable, con cielos despejados y temperaturas suaves. Las zonas centrales y occidentales disfrutarán de días soleados y cálidos, con máximas que podrían alcanzar los 25 grados Celsius.

En resumen, se anticipa un cambio en las temperaturas con descensos en el norte y sur, lloviznas en el este y en Galicia, y condiciones climáticas agradables en el resto del país. Se recomienda estar atentos a los pronósticos actualizados y tomar las precauciones necesarias para adaptarse a las condiciones climáticas variables.

#### <u>Estructurar la información</u> <u>GPT-3: engine="text-davinci-001"</u>

Hazme una lista en formato json con la persona, oficio, nacionalidad y año de nacimiento.

Bryan Adams, el fotógrafo encargado de realizar el calendario, es un cantante, guitarrista, compositor, fotógrafo y filántropo canadiense.

Anne Erin Annie Clark, conocida artísticamente como St. Vincent, es una cantautora y multiinstrumentista estadounidense. Es ganadora de tres Premios Grammy por Mejor Canción de Rock.

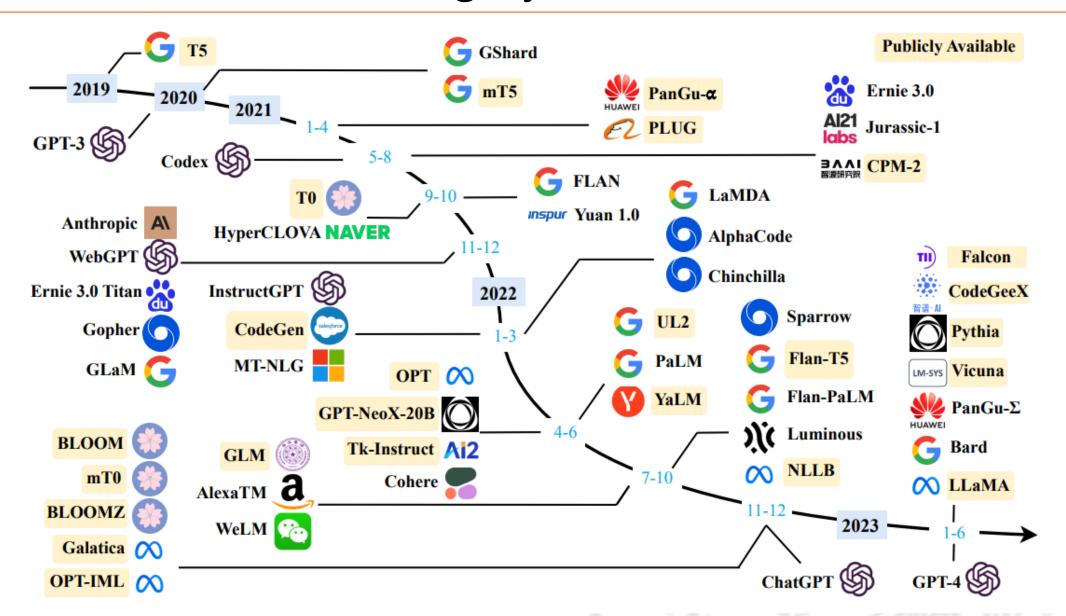
Kali Uchis, es una cantante, compositora, actriz, directora y diseñadora colomboestadounidense, saltó a la fama internacional en dos mil veintiuno con el gran éxito de su canción Telepatía.

Diamonte Quiava Valentin Harper, es una cocnocida rapera estadounidense.

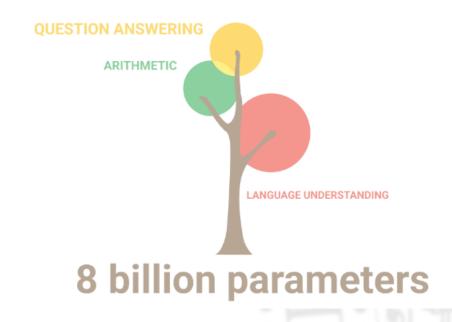
Cher es una de las figuras más destacas del este calendario.

```
"person": "Bryan Adams",
"oficio": "cantante, guitarrista, compositor, fotógrafo y filántropo",
"nacionalidad": "canadiense",
"año de nacimiento": "1958"
"person": "Anne Erin Annie Clark",
"oficio": "cantautora y multiinstrumentista",
"nacionalidad": "estadounidense",
"año de nacimiento": "1982"
"person": "Kali Uchis",
"oficio": "cantante, compositora, actriz, directora y diseñadora",
"nacionalidad": "colomboestadounidense",
"año de nacimiento": "1994"
"person": "Diamonte Quiava Valentin Harper",
"oficio": "rapera",
"nacionalidad": "estadounidense",
"año de nacimiento": "1996"
"person": "Cher",
"oficio": "cantante",
"nacionalidad": "estadounidense",
"año de nacimiento": "1946"
```

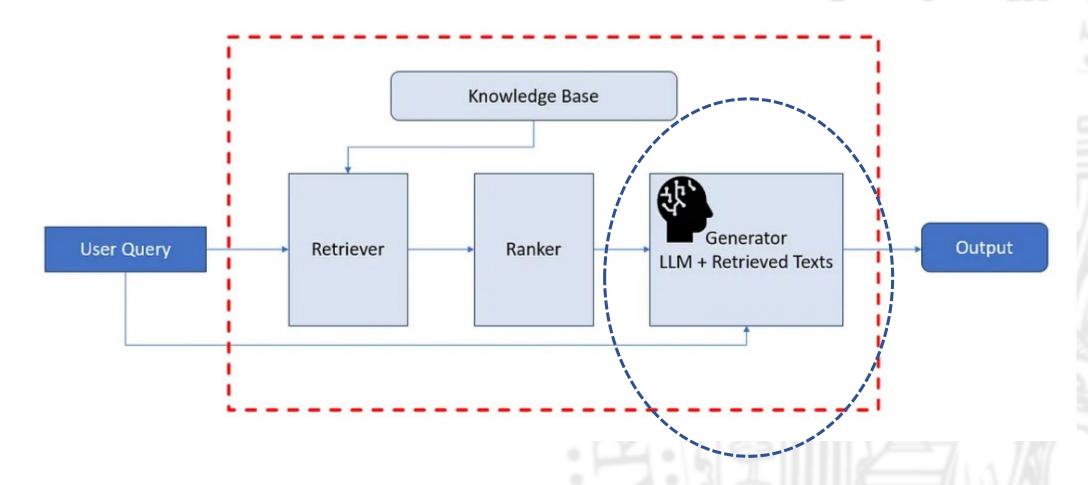




# Habilidades emergentes



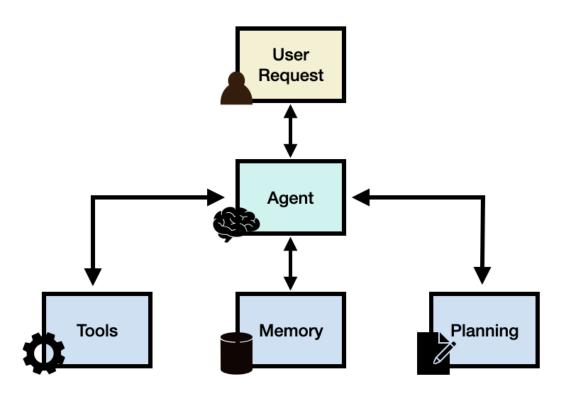
Con la llegada de los modelos de lenguaje de gran tamaño (LLM)



Un paso más: Sistemas de recuperación avanzados basados en agentes

#### Agente:

Usar un modelo de lenguaje para elegir la secuencia de acciones a realizar para resolver una consulta



#### Herramientas:

- Calculadora
- Búsqueda Wikipedia
- Búsqueda abierta en google
- Búsqueda de imágenes en google
- Búsqueda de vídeos en youtube
- Búsqueda en bases de datos privadas
- Interaccionar con el usuario ante dudas



Ejemplo:

Pregunta: busca la relación que hay entre el pueblo de Calasanz y el conde de la Viña

Invoking: `Wikipedia` with `Pueblo de Calasanz`

Page: Aragon

Summary: Aragon (ARR-ə-gən, US also -gon, -gohn; Spanish and Aragonese: Aragón [aɾaˈɣon]; Catalan: Aragó [əɾəˈɣo]) is

an autonomous community in Spain, ......

Invoking: 'Wikipedia' with 'Conde de la Viña'

Page: Concha y Toro

Summary: Concha y Toro is the largest producer and exporter of wines from Latin America.....

Invoking: `user\_input` with `{'query': 'No se encontró información sobre la relación entre el pueblo de Calasanz y el conde de la Viña. ¿Me puedes ayudar?'}`

Enter your answer: El conde de la Viña tenía una casa en Calasanz

Según la información disponible, no se encontró una relación específica entre el pueblo de Calasanz y el conde de la Viña. Sin embargo, mencionas que el conde de la Viña tenía una casa en Calasanz. Lamentablemente, no tengo acceso a información detallada sobre propiedades o residencias específicas de personas. Te recomendaría buscar información adicional en fuentes locales o históricas para obtener más detalles sobre la relación entre el conde de la Viña y su casa en Calasanz.

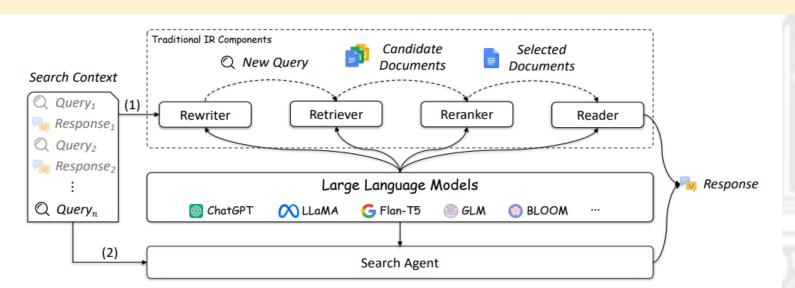


#### Respuesta directa del modelo de lenguaje (gpt-3.5-turbo-0125):

La relación entre el pueblo de Calasanz y el conde de la Viña se remonta a la época medieval. El conde de la Viña era un noble que poseía tierras y propiedades en la zona de Calasanz. Como señor feudal, tenía autoridad sobre el pueblo y sus habitantes.

El conde de la Viña ejercía su poder y controlaba los recursos del pueblo, como la tierra y los impuestos. Además, tenía la responsabilidad de proteger y defender a los habitantes de Calasanz en caso de conflictos o amenazas externas.

A lo largo de los años, la relación entre el conde de la Viña y el pueblo de Calasanz pudo haber variado dependiendo de las circunstancias históricas y las dinámicas sociales. Sin embargo, en general, se puede decir que el conde de la Viña era una figura de autoridad y poder en el pueblo, y su influencia se extendía a diversos aspectos de la vida de los habitantes de Calasanz.



#### Agentes autónomos:

Utilizar LLMs e internet

Plantear un objetivo

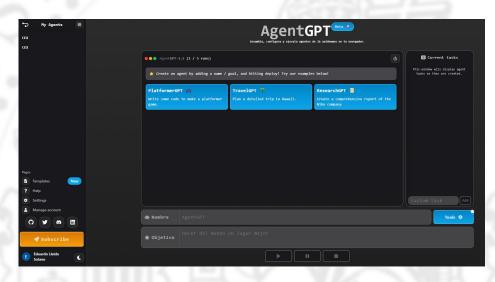
Crear tareas a partir del objetivo



#### **AutoGPT**



AgentGPT



https://agentgpt.reworkd.ai/es

https://babyagi.org/ https://github.com/Significant-Gravitas/Auto-GPT

https://babyagi-ui.vercel.app/es

#### CLIP: modelo multimodal de OpenAl

Combina un modelo de lenguaje de inglés con un modelo semántico de conocimiento de imágenes

Entrenado con mas de 400M de pares imagen+texto

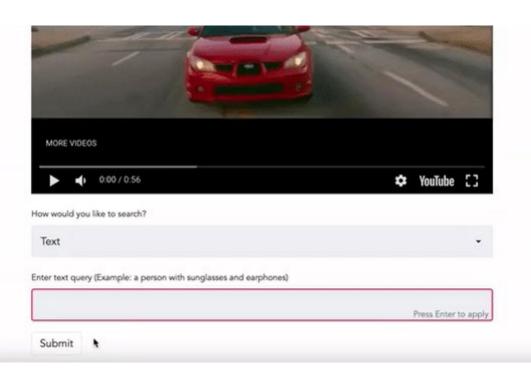


Semantic

#### Algunas aplicaciones

CLIP + LLM: Descripción de imágenes

#### Búsqueda en vídeos con lenguaje natural





A couple of people standing next to an elephant.



A wooden table sitting in front of a window.



A bunch of bananas sitting on top of a table.



A woman holding a plate with a piece of cake in front of her face.



A wooden table topped with lots of wooden utensils.



A red motorcycle parked on top of a dirt field.

#### Casos de uso con CLIP

<u>Detección de imágenes con contenido inadecuado (NSFW-Not safe/suitable for work)</u>: Similitud entre la interpretación de CLIP del texto y la interpretación de CLIP de la imagen.



```
"matches": [ {"text": "the blue car is on the left, the red car is on the right"},
{"text": "the blue car is on the right, the red car is on the left"},
{"text": "the blue car is on top of the red car"},
{"text": "the blue car is below the red car"}]}],

"the blue car is on the left, the red car is on the right" 0.5232442617416382
"the blue car is on the right, the red car is on the left" 0.32878655195236206
"the blue car is below the red car" 0.11064132302999496
"the blue car is on top of the red car" 0.03732786327600479
```



#### Descripción de segmentos de vídeo



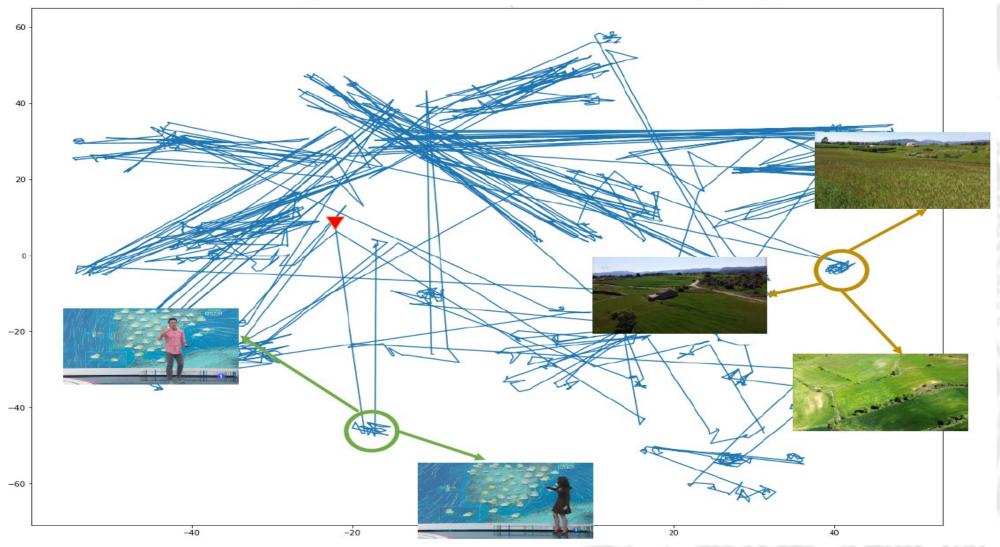
"Chef and his partner were seen chatting to one another as they prepared the meal . Chef was joined by his wife, who was also in the show . The video shows the chef slicing the meat with a knife"

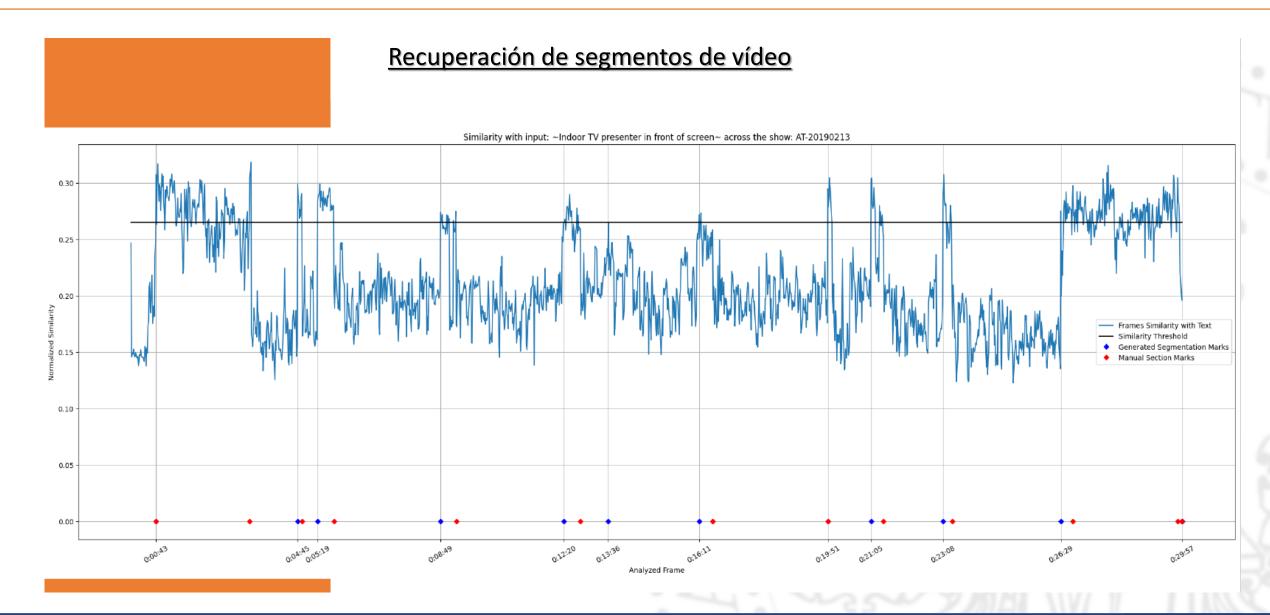
#### Descripción de segmentos de vídeo



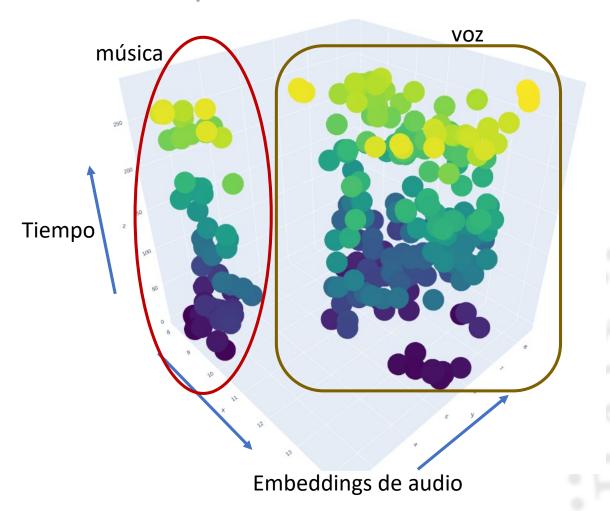
"The presenter was seen walking along the beach with his arms outstretched . The weather is expected to be mild and sunny with showers and windy"

#### Semantic concept evolution across TV program





#### AudioCLIP: espacio semántico Audio-Texto-Vídeo



- ✓ Programas de televisión: Subtítulos / Audio / Vídeo
- ✓ AudioCLIP permite consultas-respuestas multimodales



#### Nuevos entornos de desarrollo



https://github.com/salesforce/LAVIS

Benchmark, Technical Report, Documentation, Jupyter Notebook Examples, Blog

LAVIS - A Library for Language-Vision Intelligence

LLaVA: Large Language-and-Vision Assistant

https://llava-vl.github.io/



[Input Image]



[Input Instruction] What is unusual about this image?

[Model Response]

InstructBLIP 🦝



The unusual aspect of this image is that a man is ironing clothes on the back of a yellow taxi cab while it is driving down a bus city street. He is standing on top of the car's trunk, holding an iron and using it to press his clothing items. This situation is not only unconventional but also potentially hazardous, as the man could be at risk of falling off the vehicle or causing an accident due to his actions.

GPT-4



The unusual thing about this image is that a man is ironing clothes on an ironing board attached to the roof of a moving taxi.

LLaVA

The unusual aspect of this image is a man ironing clothes on the back of a minivan or van. This is not a typical place to perform this activity, as one would usually iron clothes in a more stationary and safe location, such as a home, using a regular ironing board. The scene depicted in the image is peculiar as it involves a makeshift ironing setup on a vehicle, which can be both unsafe and unconventional. Additionally, it is not clear how the man is able to maintain balance and stability while ironing clothes in such an unstable environment.

MiniGPT-4

The image shows a man standing on the back of a yellow taxi, with a yellow shirt and black pants, and a blue backpack on his back. The taxi is driving on a city street with othr cars and taxis in the background.



"These are frames of a video. Create a short voiceover script in the style of a super excited brazilian sports narrator who is narrating his favorite match. He is a big fan of Messi, the player who scores in this clip. Use caps and exclamation marks where needed to communicate excitement. Only include the narration, your output must be in english. When the ball goes into the net, you must scream GOL either once or multiple times.",



AND THERE HE GOES, MESSI WITH THE BALL, LIKE A MAGICIAN ON THE FIELD, DODGING ONE, TWO, THREE... UNSTOPPABLE! LOOK AT HIM GO, THE CROWD IS ROARING!!! CAN YOU BELIEVE THIS?? HE'S TAKEN ON THE WHOLE DEFENSE, HE'S A ONE-MAN SHOW LADIES AND GENTLEMEN... HE SHOOTS... GOOOOOOOOOOO!!!! MESSI, MESSI, MESSIIIII!!! UNBELIEVABLE, WHAT A GOAL, WHAT A GOAL! GLORIOUS, ABSOLUTELY GLORIOUS! THE STADIUM EXPLODES IN JOY! THIS IS FOOTBALL MAGIC AT ITS FINEST! ONLY MESSI, ONLY MESSI! GOL, GOL, GOOOOOOL!!!

```
PROMPT = ("The uploaded series of images is from a single video."

"The frames were sampled every {FRAME_EXTRACTION_FREQUENCY_SECONDS} seconds."

"Make sure it takes about {FRAME_EXTRACTION_FREQUENCY_SECONDS // 2} seconds to voice the description of each frame."

"Use exclamation points and capital letters to express excitement if necessary."

"Describe the video using Marv Albert style.")
```



"Alright, HERE WE GO! The game is TIED at 115! The crowd is on their FEET! The clock's ticking down! Dribbling at the top... He's waiting for the FINAL SHOT! JUST UNDER 6 SECONDS LEFT! He steps back... TAKES THE THREE... \n\nYES! HE NAILS THE THREE-POINTER AT THE BUZZER! OH, BABY, WHAT A SHOT! The crowd is going NUTS! Teammates are ECSTATIC! Celebration on the court as CONFETTI flies! \n\nThe arena is ELECTRIC, FANS are hugging, the team is swarmed with joy! What a CLUTCH PERFORMANCE to win the GAME and the SERIES! Unbelievable FINISH! Absolute MADNESS in the arena! And the hugs keep coming, WHAT A NIGHT!"

#### Ejemplo 1:

Representacion semántica de la noticias

Base de datos vectorial: Qdrant (<a href="https://qdrant.tech/">https://qdrant.tech/</a>)

Embeddings texto: multi-qa-MiniLM-L6-cos-v1 (Huggingface)

Total de noticias: 1.136.531 con 4.223.682 vectores indexados

Campos indexados: Texto / Título / Imagen / Resumen

Noticias: 40 periódicos españoles de mayo de 2023 a noviembre de 2023

http://signal4.cps.unizar.es:8515



#### Ejemplo 2:

Buscador de noticias de periódicos multimodal: texto e imagen

Búsqueda híbrida: índice invertido + embeddings

Base de datos vectorial: Qdrant (<a href="https://qdrant.tech/">https://qdrant.tech/</a>)

Embeddings texto: multi-qa-MiniLM-L6-cos-v1 (Huggingface)

Embeddings multimodales: clip-ViT-L-14 (OpenAI)

Total de noticias: 1.136.531 con 4.223.682 vectores indexados

Campos indexados: Texto / Título / Imagen / Resumen

Noticias: 40 periódicos españoles de mayo de 2023 a noviembre de 2023

http://signal4.cps.unizar.es:8514



#### Ejemplo 3:

Agente: Busca información en la base de datos de noticias, wikipedia y youtube.

Agente: gpt-3.5-turbo-0125

Herramientas: WikipediaQueryRun, YouTubeSearchTool,

QdrantSearchNewsTool, GoogleSearchTool

Base de datos privada: Beword Qdrant (https://qdrant.tech/)

Embeddings texto: multi-qa-MiniLM-L6-cos-v1 (Huggingface)

Embeddings multimodales: clip-ViT-L-14 (OpenAI)

Total de noticias: 1.136.531 con 4.223.682 vectores indexados

Campos indexados: Texto / Título / Imagen / Resumen

Noticias: 40 periódicos españoles de mayo de 2023 a noviembre de 2023

http://signal4.cps.unizar.es:8513/

