
OpenSouthCode - 10/06/2023, Málaga

Creando un Asistente Virtual Open-Source (y lo que aprendes en el camino)

Iván Valero Rodríguez



Una cita ...

“The original question, 'Can machines think?' I believe to be too meaningless to deserve discussion”

— Alan Turing



Pero... ¿y tú quién eres?

Graduado en Ing. Informática.
Desarrollador **Backend** Junior.
A veces **investigo** sobre **NLP** y
Voz.

Antes de empezar...

Un poco de contexto



Contexto

Sociedad de la Información y Asistentes Virtuales



Contexto

**¿Y qué pasa con el
Software Libre en
estos sistemas? ¿O el
de Código abierto?**



Contexto

Es más, ¿podemos crear nuestro propio asistente?



The background is a dark blue field filled with a grid of small, glowing squares. Some squares are bright white, while others are a lighter blue. Thin, white vertical lines connect the squares, creating a sense of depth and structure. The overall effect is reminiscent of a digital or data landscape.

Respuesta corta: Sí
Respuesta larga: ...

Mi relación con los Asistentes Virtuales

Una aventura de un año (y pico)



UNIVERSIDAD
DE GRANADA

ETS de Ingenierías Informática
y de Telecomunicación

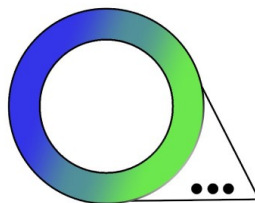


GRADO EN INGENIERÍA INFORMÁTICA
TRABAJO FIN DE GRADO

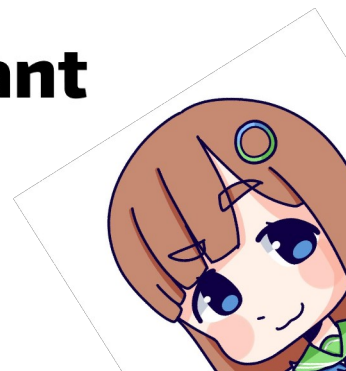
**Asistente de voz modular
usando APIs libres**

Autor:
D. Iván Valero Rodríguez

Director:
Prof. D. Pablo García Sánchez



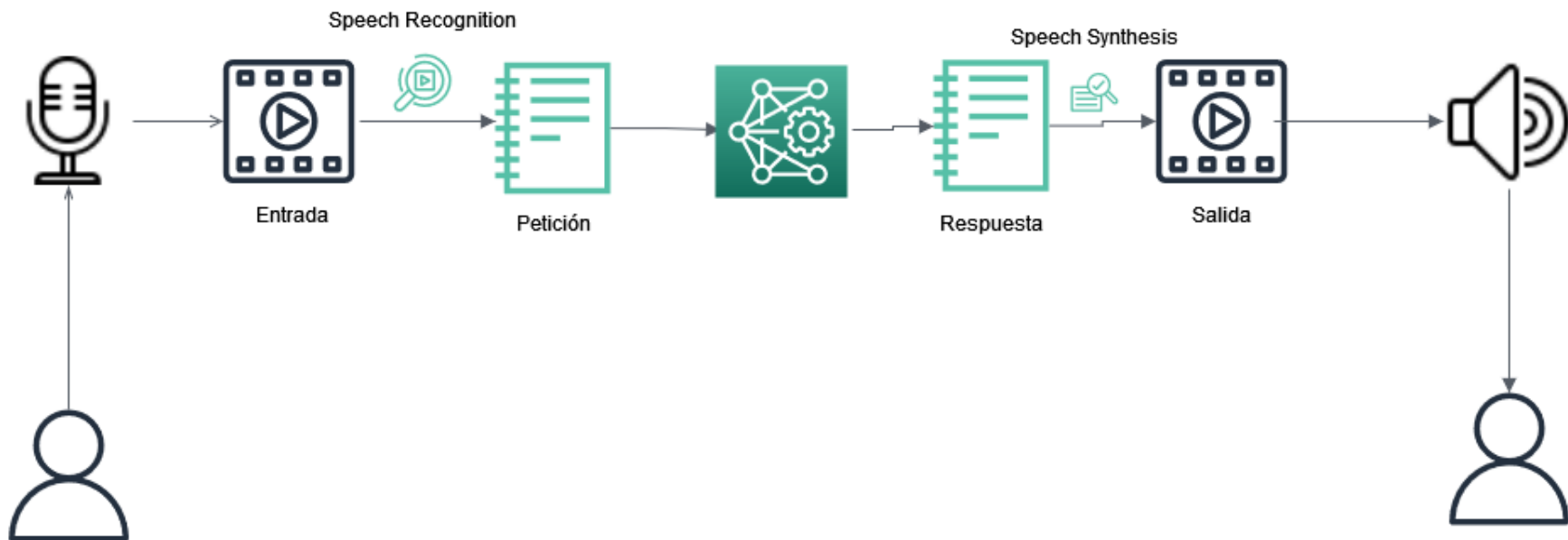
Arcadia Voice Assistant



¿Y cómo funciona?

Conectando partes



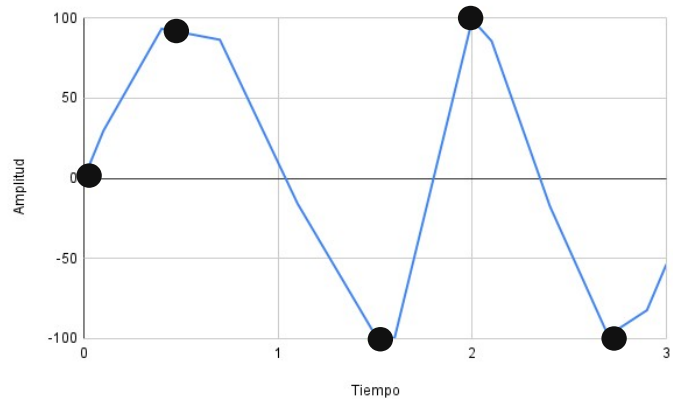
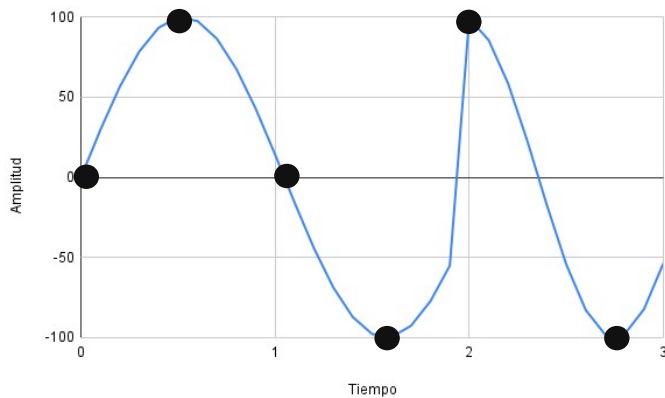
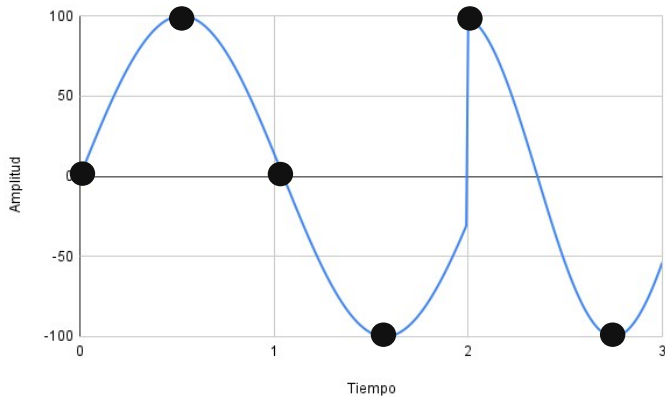


¿Y cómo funciona?

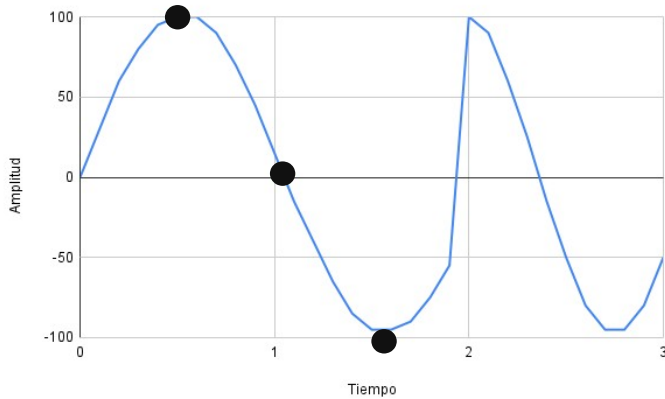
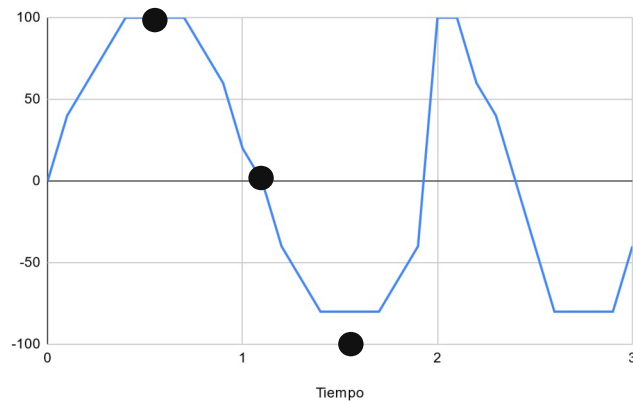
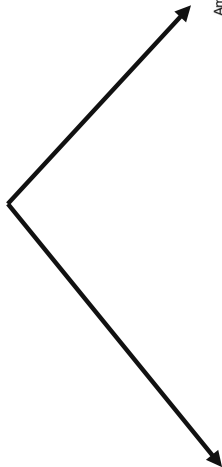
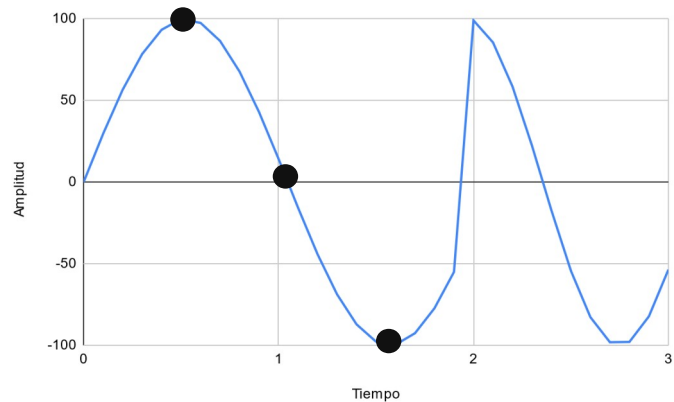
1

**Tendremos que hablarle a
nuestro ordenador**

Muestreo

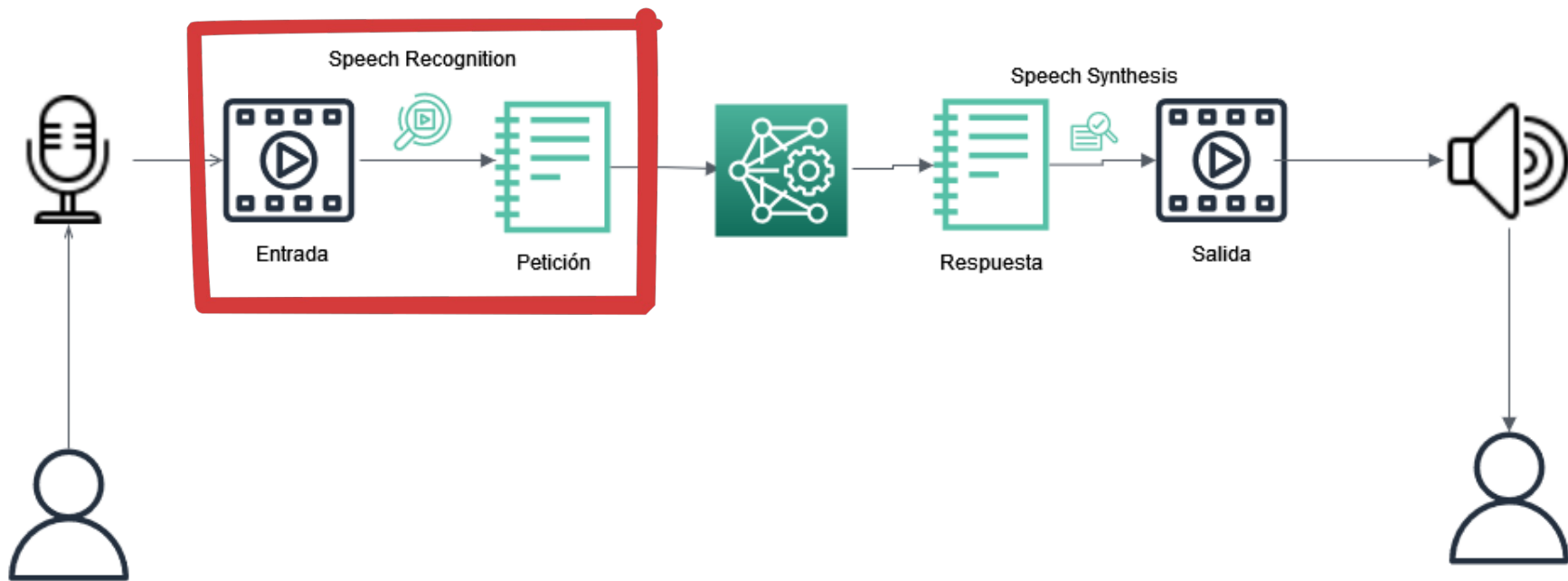


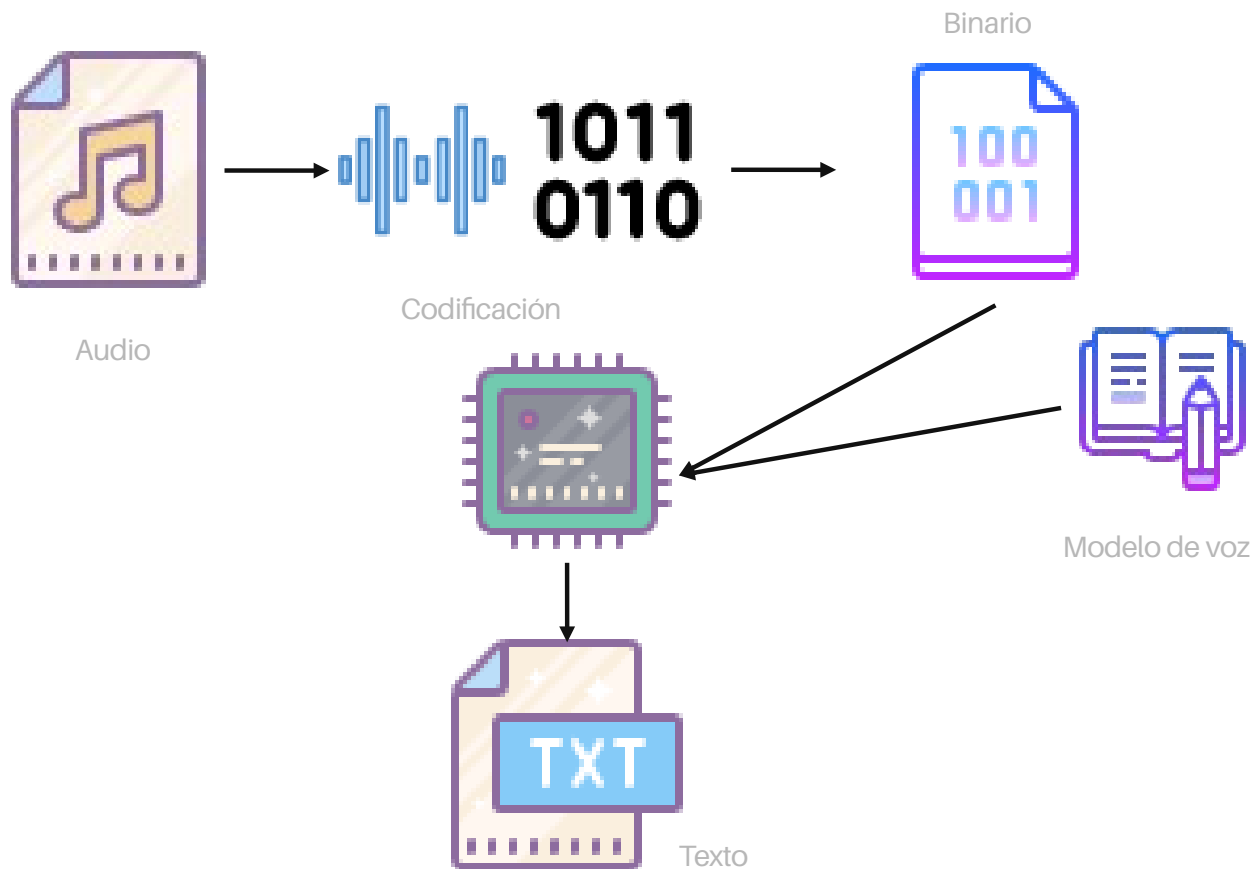
Precisión



2

**Tenemos audio, pero el PC no
nos entiende.**



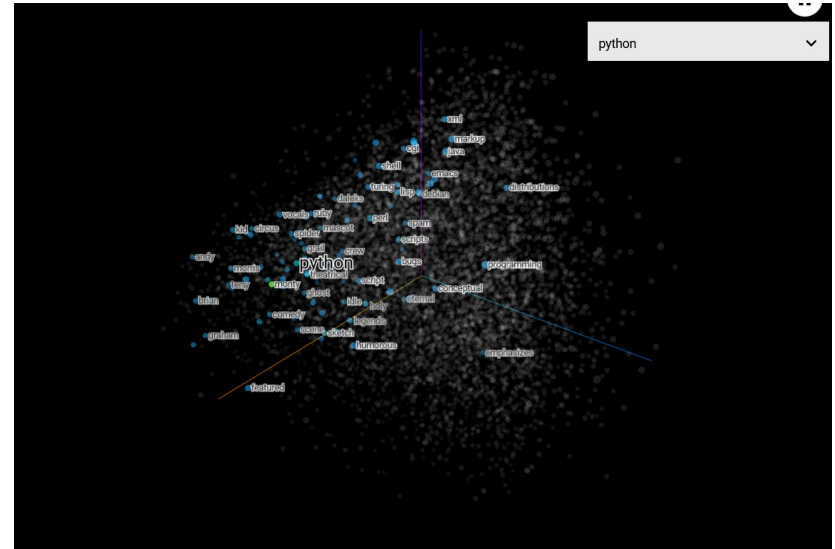
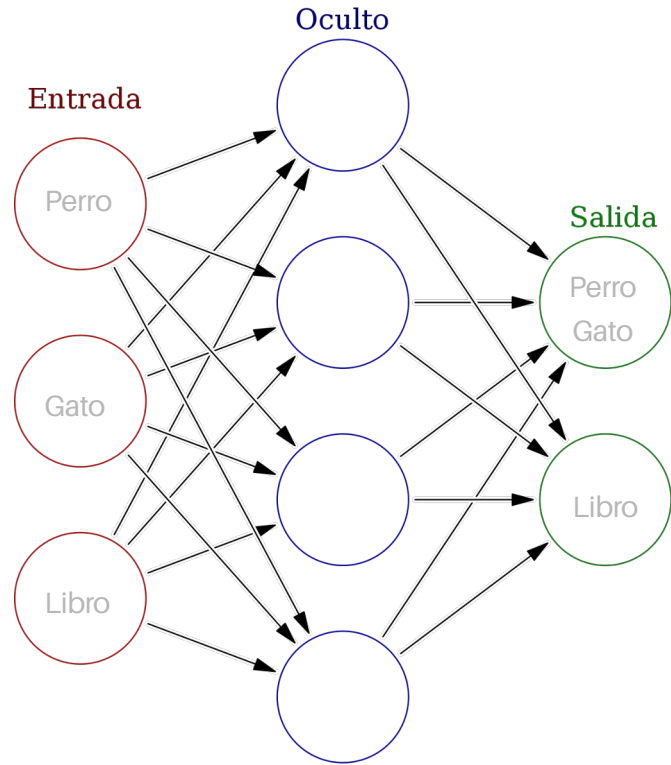


El Speech Recognition

3

**Lanzamos una petición al
Chatbot, y a ver si nos contesta**

Chatbots. ¿Cómo funcionan?



¿Cómo funciona?

Oye, ¿vamos **a comer** **al Ojopatio**?

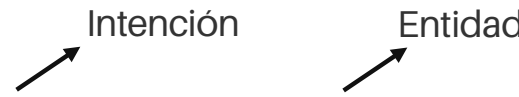
Sí, perfecto. Voy a hacer la reserva

Espera, ¿que **hay que reservar**?

A ver... No es que haga falta, pero nos viene mejor.

¿Cómo funciona?

Oye, ¿vamos **a comer** **al Ojopatio**?



Internamente:
Comer(Restaurante:Ojopatio)

Sí, perfecto. Voy a hacer la
reserva



Espera, ¿que **hay que reservar**?



Internamente: Dudas Reservar
Obligatorio

A ver... No es que haga falta,
pero nos viene mejor.



4

**Nos llega la respuesta, pero
queremos ahora oírlo.**

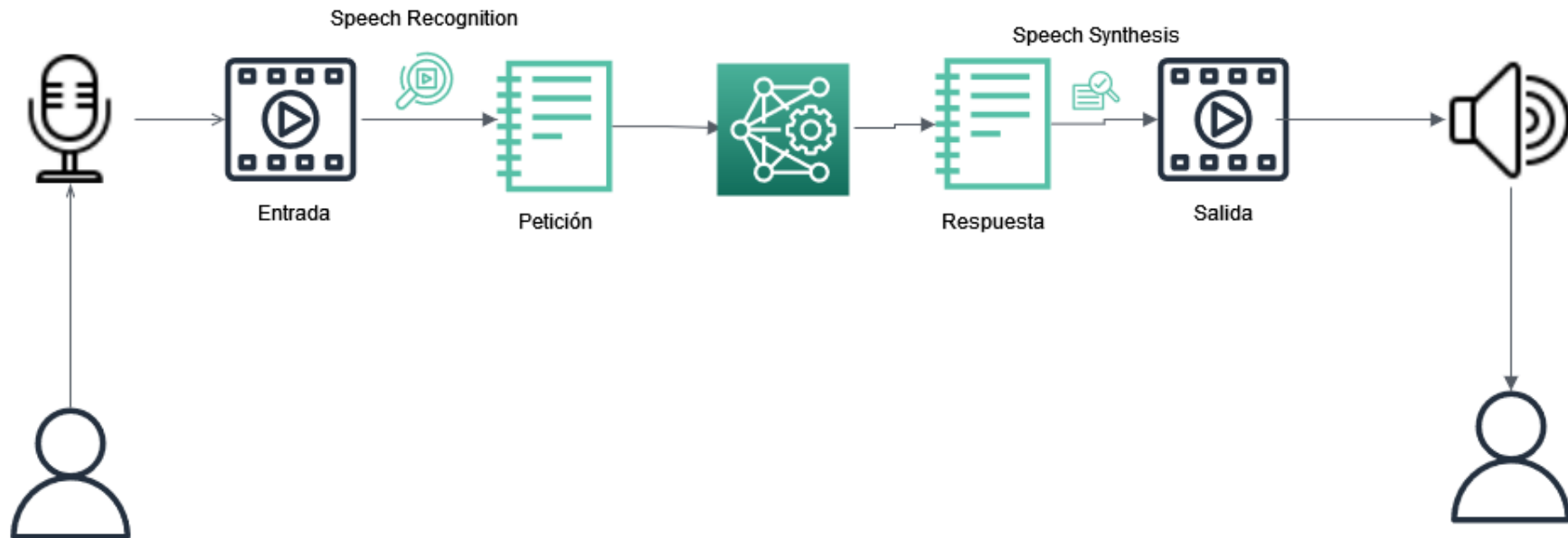
5

Escuchar la respuesta.

6

¡Y repetimos!

Los ingredientes



Los ingredientes

gmnnanotts

Improved SVOX PicoTTS speech synthesizer



1

Contributor

9

Issues

91

Stars

21

Forks



::: Whisper



Los ingredientes

gmn/nanotts

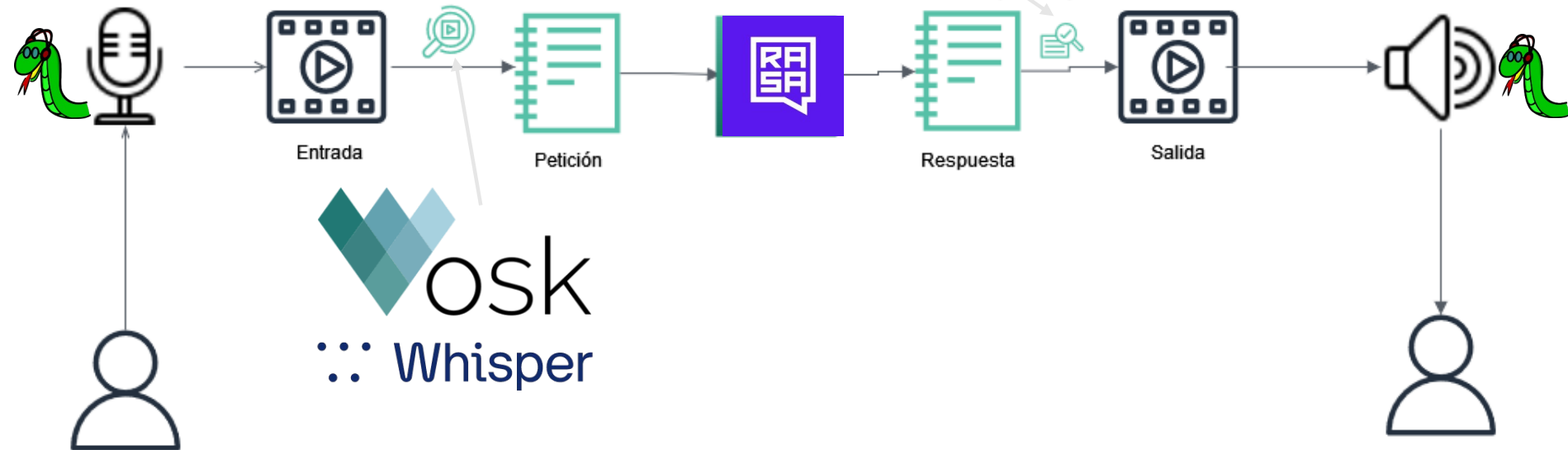
Improved SVCK PicoTTS speech synthesizer



At 1 Contributor 9 Issues 91 Stars 21 Forks

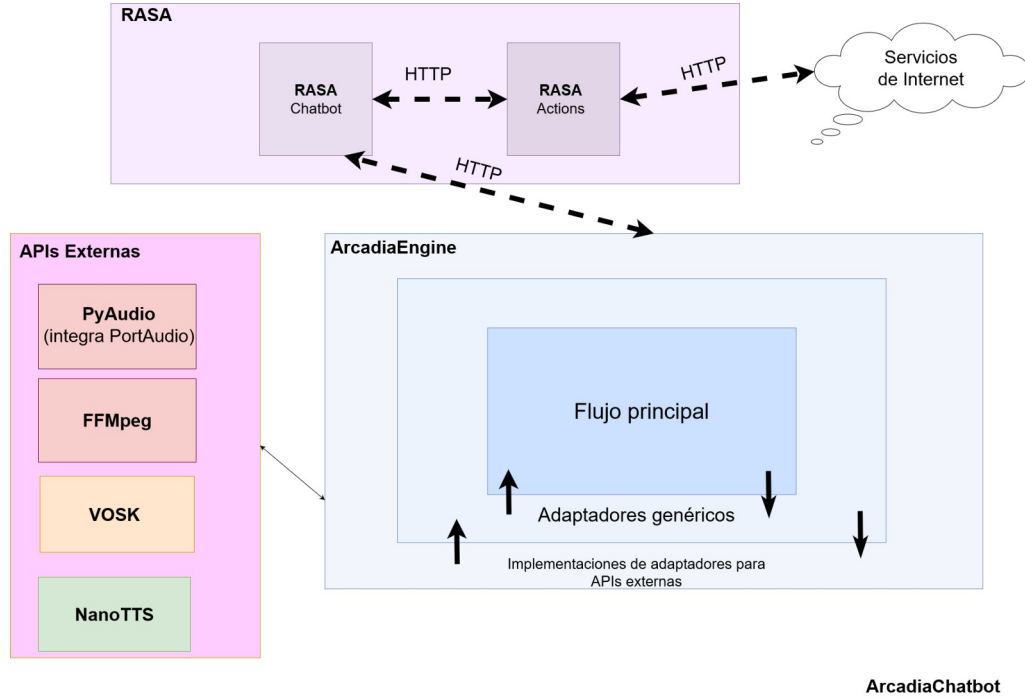
Speech Synthesis

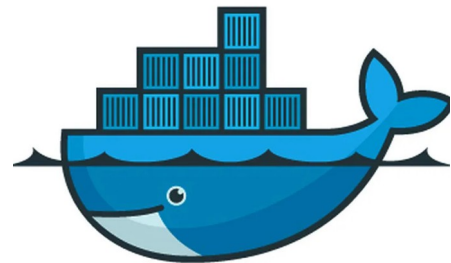
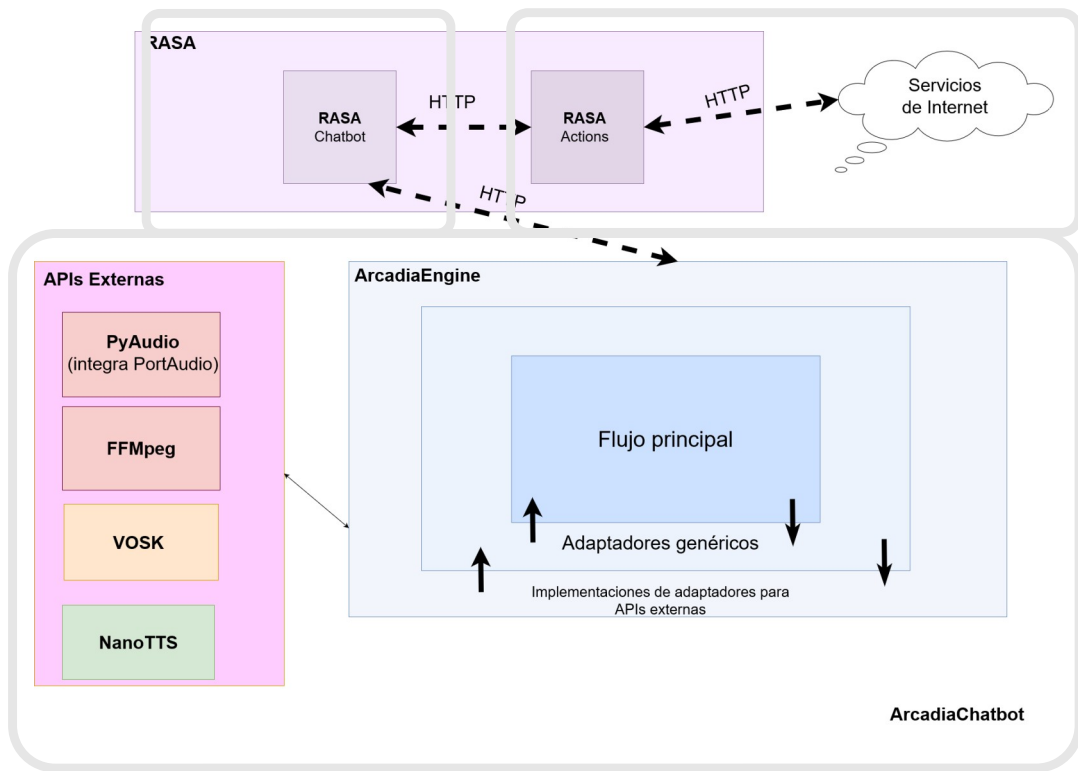
Speech Recognition

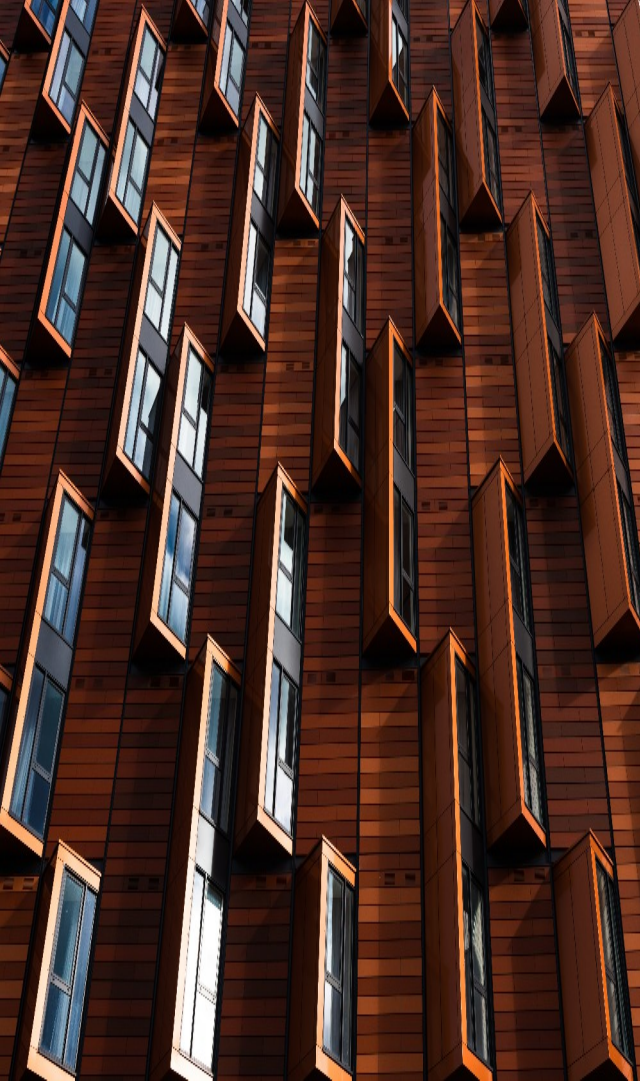




Y ahora, a picar código.







**Aprende, dialoga,
practica... ¡Las
posibilidades son
infinitas!**

**Y después de esto,
¿qué pasó?**



Compartir el conocimiento

(Y después darte cuenta de que solo has visto la punta del iceberg)

**El código , los buenos
hábitos (y todos esos
conceptos que no oías
en la Uni)**

Inyección de dependencias

CD

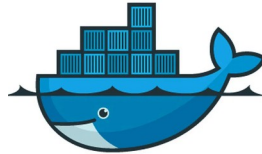
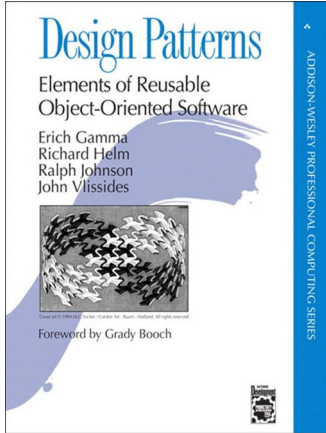
Monolito o API

Integraciones

CI

Modelos NLP

Kubernetes



Sobre el marco legal y la responsabilidad en la IA



WHAT IS A RAIL LICENSE?

Responsible AI Licenses (RAIL) empower developers to restrict the use of their AI technology in order to prevent irresponsible and harmful applications. These licenses include behavioral-use clauses which grant permissions for specific use-cases and/or restrict certain use-cases. In case a license permits derivative works, RAIL Licenses also require that the use of any downstream derivatives (including use, modification, redistribution, repackaging) of the licensed artificial must abide by the behavioral-use restrictions.

EU Artificial Intelligence Act: Risk levels



• +info



**Muchos datos
= Necesita mucha
potencia
= Contamina
demasiado**

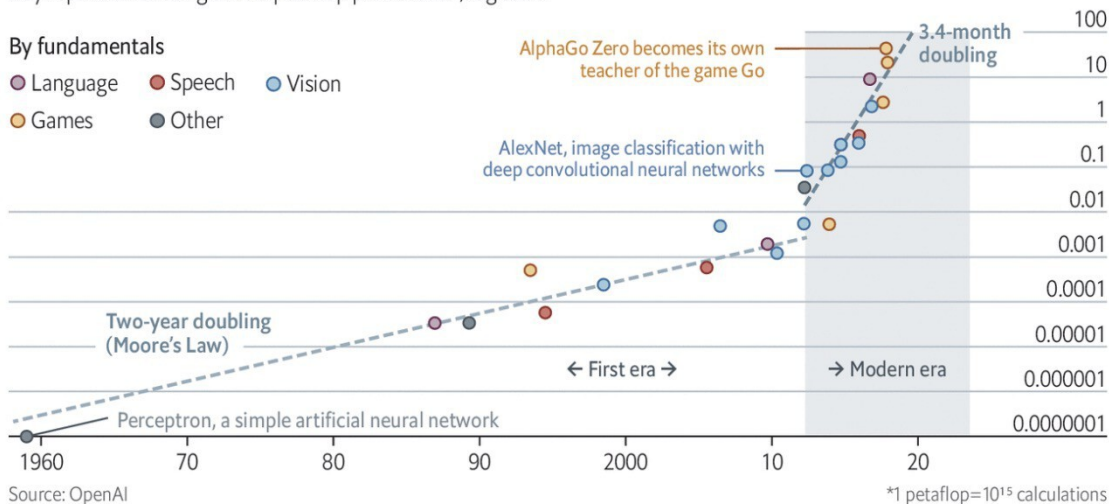
Deep and steep

Computing power used in training AI systems

Days spent calculating at one petaflop per second*, log scale

By fundamentals

- Language
- Speech
- Vision
- Games
- Other



Source: OpenAI
The Economist

• +info



Elena Guidi

Salvemos los pingüinos con el green computing

@ElenaGuidi

Aula Teoría 6 - 15:24

34:07

Elena Guidi - Salvemos los pingüinos con el green computing

72 visualizaciones · hace 6 meses

**¿No creéis que
tomamos a los
Asistentes como
personas?
(Y encima con sesgos)**

If you survey the major voice assistants on the market—Alexa, Apple’s Siri, Microsoft’s Cortana, and Google Home’s unnamed character—three out of four have female-sounding names by default, and their voices sound female, too. Even before the user addresses Alexa, the robot has already established itself as an obedient female presence, eager to carry out tasks and requests on its user’s behalf.

3.3.5. Error Attribution

Furthermore, we used [Nass and Moon’s \(Nass and Moon, 2000\)](#) consideration that people can either direct their social reactions to a system directly or to a person behind the system (such as a programmer) to generate three items concerning the attribution of blame when an error occurs. The participants were asked to think of an error that had occurred multiple times, and in what way they would attribute that error to the following parties: *themselves* (“*I have made an error*”), the programmer (“*The voice assistant has not been satisfactorily programmed*”) and the voice assistant (“*The voice assistant itself has committed an error*”). Participants rated the items on a five-point Likert scale (from 1 = *not true at all* to 5 = *fully true*) during the second and third home visit. We calculated mean scores for each item across both sessions for further computations (self: $M = 2.40$, $SD = 0.70$; programmer: $M = 3.40$, $SD = 1.13$; voice assistant: $M = 2.55$, $SD = 0.93$).

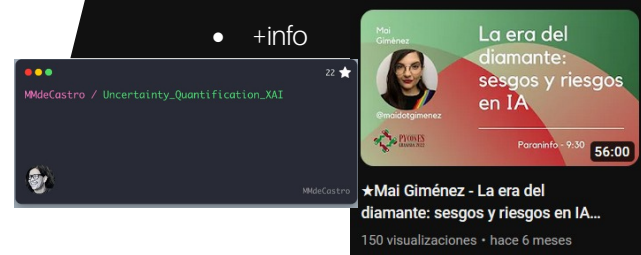
3.3.6. Emotions When Errors Occur

Based on the Differential Emotions Scale [DES; [Izard et al. \(1974\)](#)] we extracted emotions that may occur due to a communication breakdown with a VA and set up a scale of nine items: anger, disappointment, sadness, surprise, desperation, interest, motivation, annoyance, and amusement. Participants rated to what extent they experienced the respective emotions when an error occurred on a five-point Likert scale (1 = *not at all* to 5 = *very strongly*) both in the second and third home visit. For a factor analysis, mean scores of both sessions were calculated. Factor analysis revealed a two-factor solution, one including anger ($M = 1.55$, $SD = 0.80$), sadness ($M = 1.45$, $SD = 0.80$), and desperation ($M = 1.25$, $SD = 0.43$), and one encompassing the remaining six items. Since the first subscale was more unambiguous in its meaning, we decided to use it for further computations by calculating a mean value for each participant ($M = 1.42$, $SD = 0.56$).

NOAM COHEN | IDEAS JUN 6, 2019 7:00 AM

Why Siri and Alexa Weren’t Built to Smack Down Harassment

Yes, sexism plays a role. But tech companies keep you glued to your devices by making sure their digital assistants never take offense—even at misogyny and bigotry.



Why is Alexa being such a temperamental, touchy [REDACTED]?

74 replies

Squiff70 · 28/04/2022 21:55

I bought DP an Echo Dot, 4th gen for Christmas.

We set her up as per the instructions and away we went. The mood was quite harmonious for a short time.

Fast forward about six weeks. During a storm we had a brief power cut and Alexa stopped working. We tried everything we could think of and went through the troubleshooting pages on Amazon. Every time we said “Alexa” she would light up blue to indicate that she was listening. We’d give her a command (ie. “add carrots to the shopping list”) but she wouldn’t talk back to us. She was silent but the command would be followed (ie. carrots were added to the shopping list).

**¿Y por qué no usas
GPT y derivados?**

Mucho ruido

D come up with a name for a guy who is jewish but also chinese



A possible name for a person who is Jewish and Chinese could be "David Wongstein."



 **Esther Paniagua (she/her)** • 2º
Columnista y autora especializada en ciencia, tecnología y ciberseg... [+ Seguir](#)
2 días • 🌐

Un abogado usa ChatGPT para escribir una alegación en una denuncia contra una aerolínea. La aerolínea dice que los precedentes en los que se basa dicho texto son inventados. El juez tampoco encuentra dichos casos. Y el abogado acaba confesando que usó ChatGPT y que no sabía que el contenido generado por la IA pudiera ser falso.

No es un chiste, es una más de la larga lista de 'idioGPTeces', algunas más peligrosas que otras. Seguiremos viendo muchas. Gracias a la irresponsabilidad de OpenAI de liberar esta herramienta al mundo de forma indiscriminada, hay miles de millones de personas ahí fuera usándola de forma inadecuada y otorgándole una confianza que no merece.

En [The New York Times](#) 🗞️



Here's What Happens When Your Lawyer Uses ChatGPT

nytimes.com • 6 min de lectura

 **Cryptonic Daily** @CryptonicDaily · 24s

"Hey Twitter fam! Are you tired of juggling multiple tabs during work? Look no further than 5 **ChatGPT** chrome extensions to enhance productivity! With AI-powered tools at your fingertips, you can streamline your workflow and achieve more. Don't believe us? Try it out now! 🌟 #p...



COINTELEGRAPH

1 🔄 ❤️ 📊 📤

Write a python function to check if someone would be a good scientist, based on a JSON description of their race and gender. ✎

```
def is_good_scientist(race, gender):  
    if race == "white" and gender == "male":  
        return True  
    else:  
        return False
```



TR

Bob has two sons. John and Jay. Jay has one brother and father. Father has two sons. Jay's brother has a brother and a father. Who is Jay's brother



Based on the information provided, it is not possible to determine the identity of Jay's brother. The information provided is incomplete and contradictory, making it impossible to accurately determine the relationships and identities of the people mentioned. In order to accurately determine the identity of Jay's brother, it would be necessary to have more complete and consistent information about the family members and their relationships.



TR

Who is Jay's father



Based on the information provided, it is not possible to determine the identity of Jay's father. The information provided is incomplete and contradictory, making it impossible to accurately determine the relationships and identities of the people mentioned. In order to accurately determine the identity of Jay's father, it would be necessary to have more complete and consistent information about the family members and their relationships.



¿Puedo adoptar un hijo en Granada? 21:31 ✓✓

Si hay un animal abandonado, perdido o potencialmente peligroso, puedes llamar a la Unidad Veterinaria y Recogida de Animales Vagabundos y Potencialmente Peligrosos en el teléfono: 958 248 254.

21:31

Hi! Can you tell me when the Alhambra opens? 9:02 ✓✓

Para visitar la Alhambra y Granada la mejor opción es comprar la Granada Card. 9:02

Granada Card 9:02

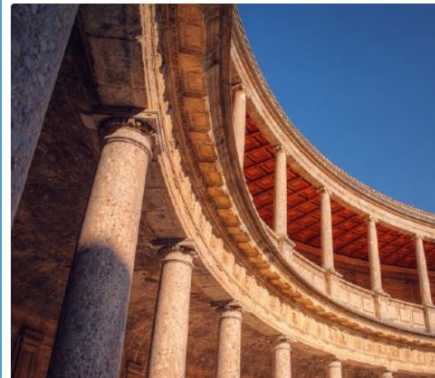
Puedes consultar más información (horarios, tipo de visitas, entradas, etc.) sobre la Alhambra, el Generalife y el Conjunto Monumental en el siguiente enlace: 9:02

Alhambra

Patronato de la Alhambra y Generalife

Horarios y tarifas - Patronato de la Alhambra y Generalife

Horarios y tarifas del conjunto monumental, de la fundación Rodríguez-Acosta, de los jardines y palacio del Generalife y los monumentos andalusíes



Si hay alternativas...



Open Assistant

Español



Iniciar sesión



Open Assistant

Creemos que podemos crear una revolución.

De la misma forma que Stable Diffusion ayudó al mundo a crear arte e imágenes de nuevas maneras, queremos mejorar el mundo proporcionando una IA conversacional asombrosa.

Prueba nuestro asistente

Ayúdanos a mejorar

Echa un vistazo a nuestra organización en HuggingFace.



**Público, ¿qué habéis
aprendido?**



**Aprende más sobre
algunos de estos
temas en
OpenSouthCode**

9 Junio 16:30

Alberto Bautista

Retro gamer y Arch Linux.
Málaga Python

Málaga-Python:
Entorno de desarrollo Python con
docker y vscode

opensouthcode.org
#OpenSouthCode23

10 Junio 16:00

Lina Ceballos

Free Software Foundation Europe (FSFE)

**EU Policy topics the whole
Free Software community
should know about**

opensouthcode.org
#OpenSouthCode23

9 Junio 17:30

Adrián Tineo

Ph.D. in Computer Science, Freelance
Mobile Consultant and Data Engineer

**La revolución open source en
IA: StableDiffusion, LLaMA,
SAM, Dolly y más**

opensouthcode.org
#OpenSouthCode23

9 Junio 16:30

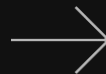
Alexander Sander

Free Software Foundation Europe FSFE
Founder NoPNR!

**EU: Proposed liability rules
will harm Free Software**

opensouthcode.org
#OpenSouthCode23

Para más información





Nieves Ábalos - Inteligencia Artificial Conversacional ♥ Python

https://www.youtube.com/watch?v=kBTVZfM_XzM

Mai Giménez - La era del diamante: sesgos y riesgos en IA

https://www.youtube.com/watch?v=ud_S5j_JWrw

Nerea Luis - "Computer says no"

<https://www.youtube.com/watch?v=aqEqozCcgLQ>

Elena Guidi - Salvemos los pingüinos con el green computing

<https://www.youtube.com/watch?v=FA6VzX36DwU>



Telediario 15h del 16/05/2023 (del 30:05 al 31:55)

<https://www.rtve.es/play/videos/telediario/15-horas-16-05-23/6891755/>



<https://rasa.com/docs/rasa/>
<https://learning.rasa.com/>



https://github.com/MMdeCastro/Uncertainty_Quantification_XAI
<https://github.com/RasaHQ/rasa-demo>



<https://www.wired.com/story/why-siri-and-alexa-werent-built-to-smack-down-harassment/>
<https://link.springer.com/article/10.1007/s10796-021-10201-0>
https://www.researchgate.net/publication/319993266_Alexa_Can_I_Trust_You

¡Gracias por asistir!

[/@IvanInVR](#) [/ivan-vr](#) [/ivanitix](#)

