Google Developer Student Clubs
University of Granada

Asistentes Virtuales Open Source (y lo que no te quieren contar)

Iván V.R
Backend Junior @
Nazaríes-Intelligenia
Arcadia Voice Assistant Dev

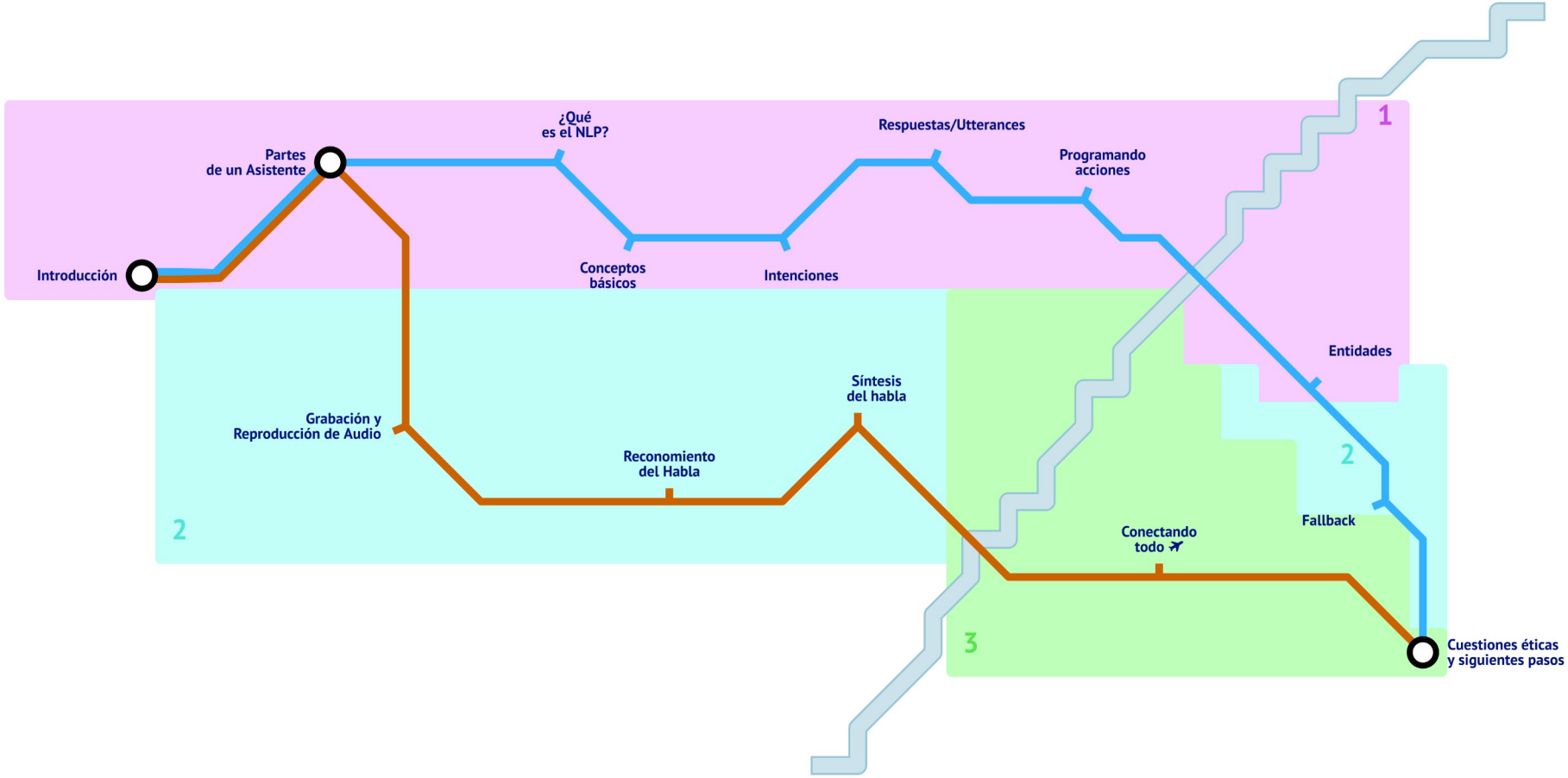
Una cita ...

“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.
En general, **hago cosas**.


Introducción
Partes de un Asistente
¿Qué es el NLP?
Conceptos básicos
Intenciones
Respuestas/Utterances
Programando acciones
Entidades
Fallback
Conectando todo
Cuestiones éticas y siguientes pasos
Grabación y Reproducción de Audio
Reconomiento del Habla
Síntesis del habla
1
2
2
3

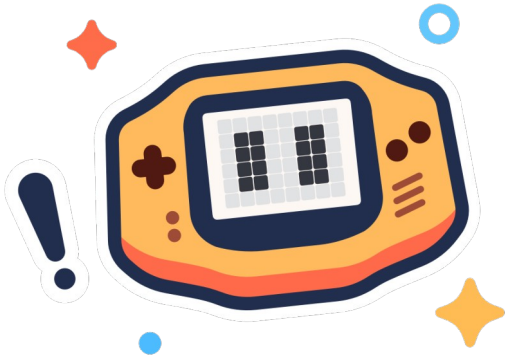
Antes de empezar...

Un poco de contexto

Contexto

Sociedad de la Información y Asistentes Virtuales

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



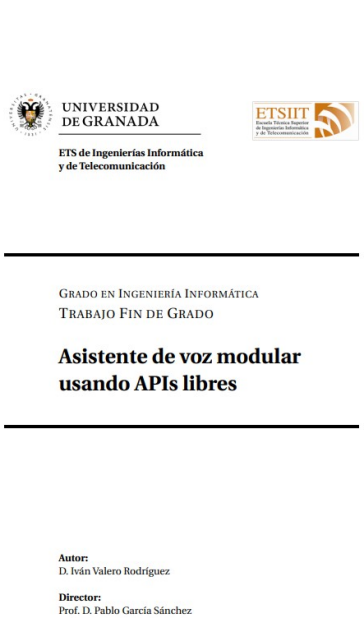
Contexto

Es más, ¿podemos crear nuestro propio asistente?


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

ETSIIT

GRADO EN INGENIERÍA INFORMÁTICA
TRABAJO FIN DE GRADO

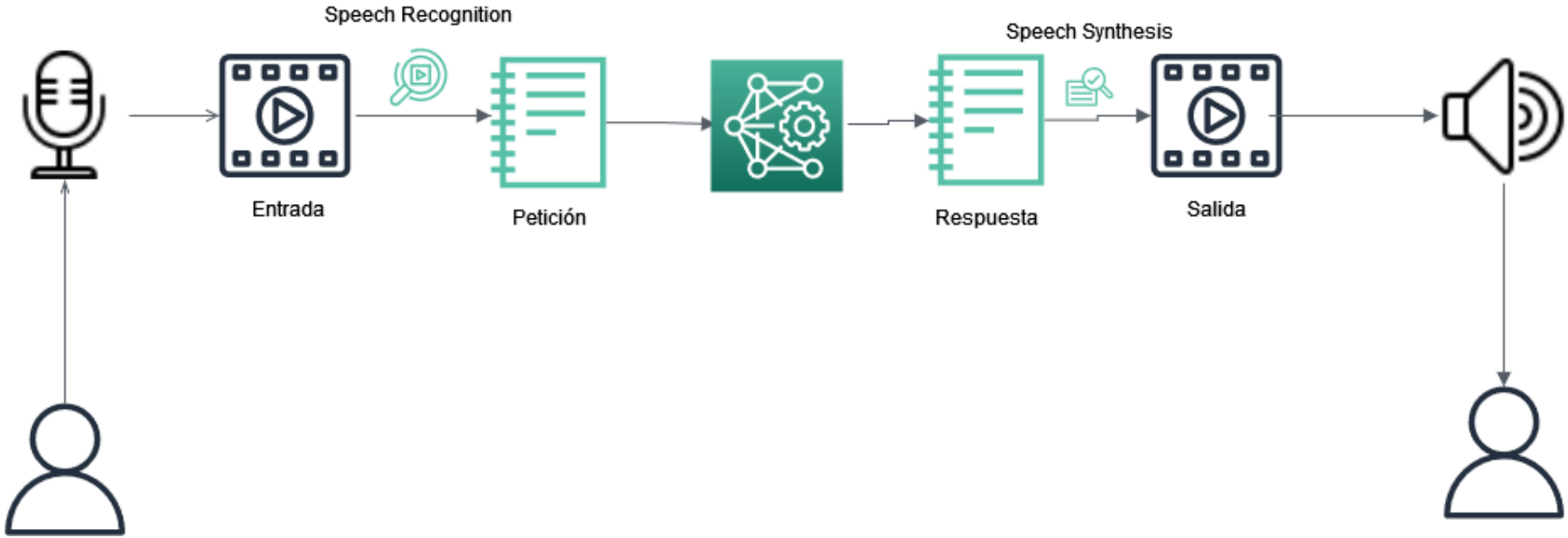
Asistente de voz modular usando APIs libres

Autor:
D. Iván Valero Rodríguez

Director:
Prof. D. Pablo García Sánchez

Arcadia
Voice Assistant

¿Y cómo funciona?
Conectando partes


Speech Recognition
Speech Synthesis
Entrada
Petición
Respuesta
Salida

1

Tendremos que hablarle a nuestro ordenador

2

Tenemos audio, pero el PC no nos entiende.

3

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

4

Nos llega la respuesta, pero queremos ahora oírlo.

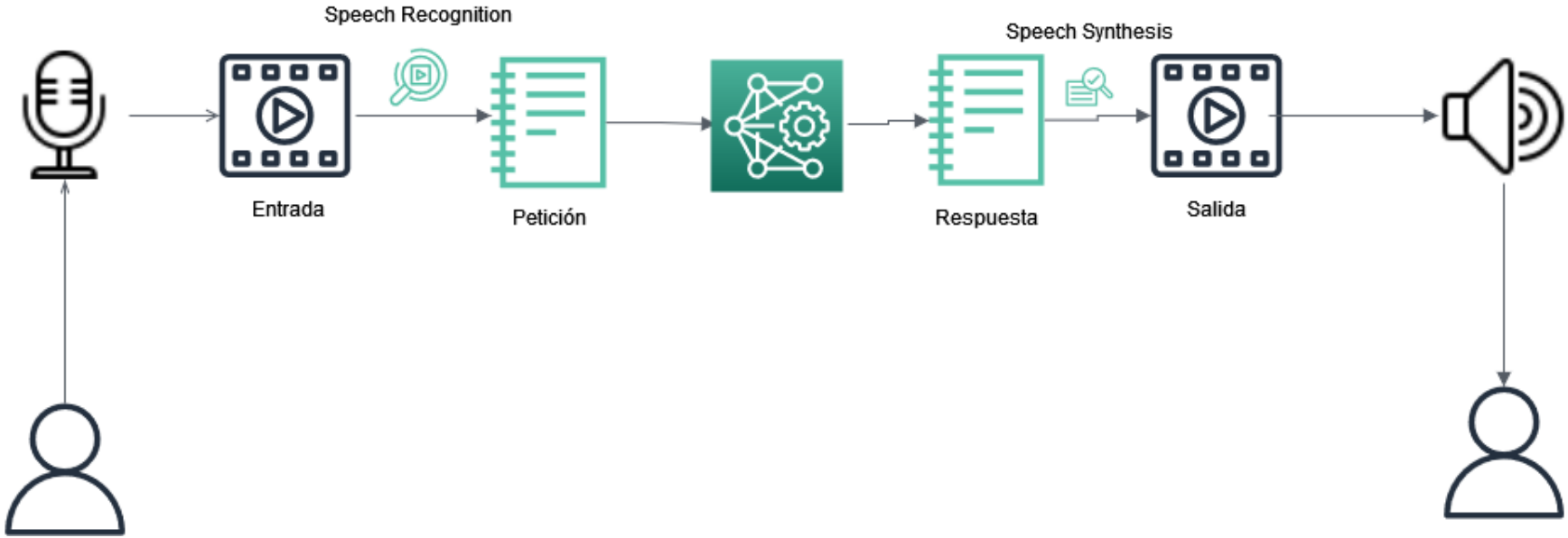
5

Escuchar la respuesta.

6

¡Y repetimos!

Los ingredientes



Los ingredientes

gmn/**nanotts**

Improved SVOX PicoTTS speech synthesizer

1 Contributor · 9 Issues · 91 Stars · 21 Forks

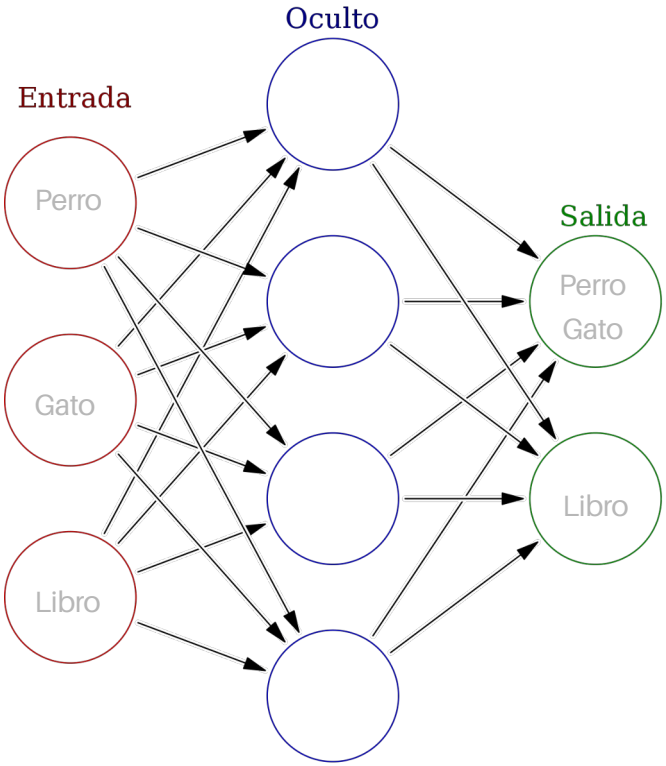
Whisper

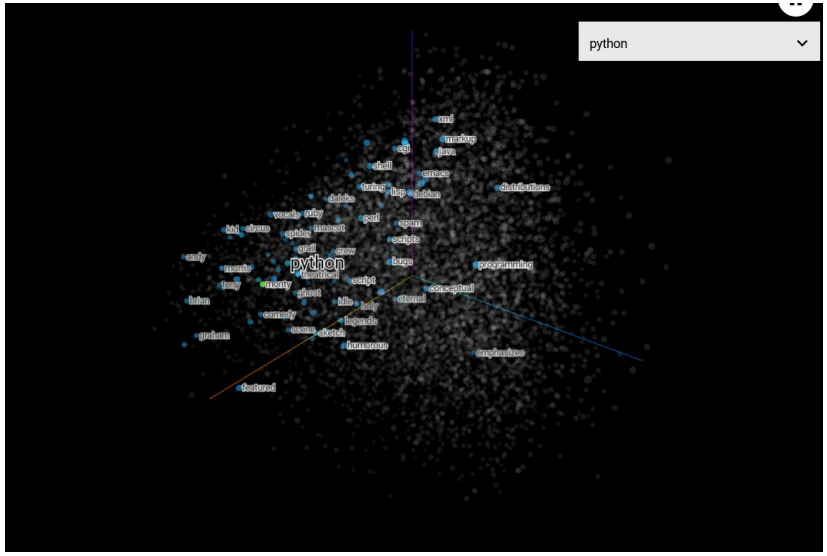
Los ingredientes

P1

Chatbots y NLU: ¿Cómo chatean las máquinas con nosotr@s?

Chatbots. ¿Cómo funcionan?





Oye, ¿vamos a tapear a la Posada?

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?

Intención Entidad

Oye, ¿vamos a tapear a la Posada?

Respuesta

Sí, perfecto. Voy a hacer la reserva

Internamente:
Comer(Restaurante:Posada)

Intención

Espera, ¿que hay que reservar?

Respuesta

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

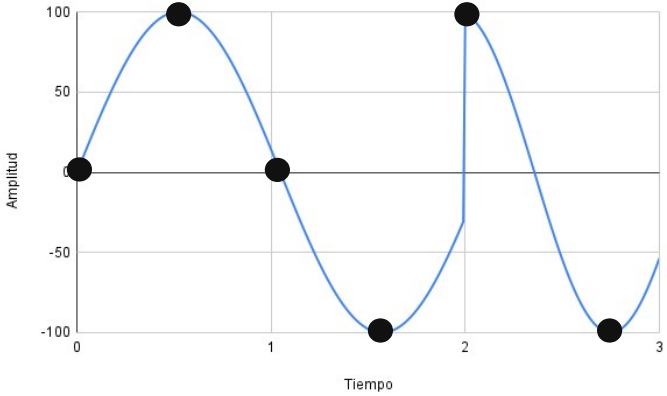
Internamente: Dudas Reservar Obligatorio

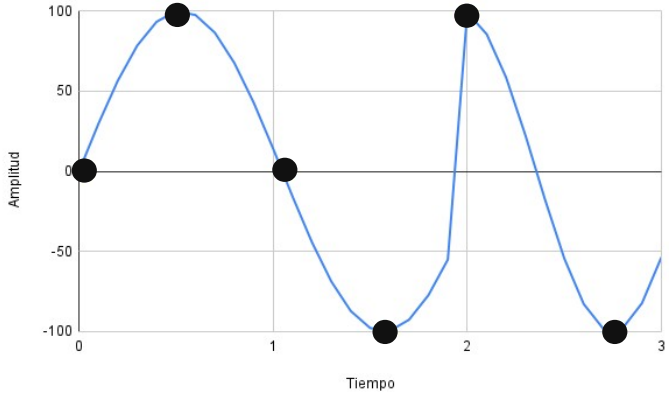

Y ahora, a picar código.

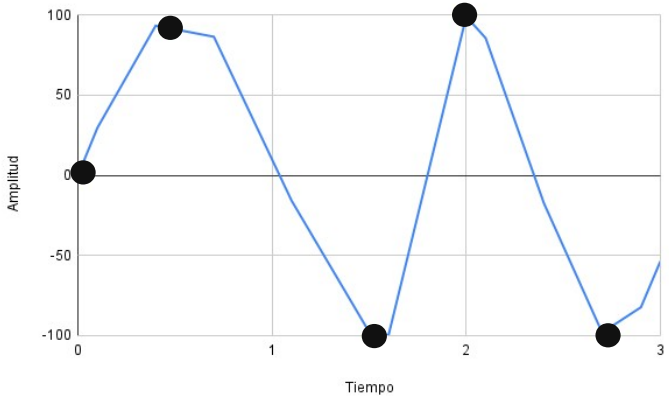
P2

Trabajando con la voz: ¿Ese ordenador me está hablando?

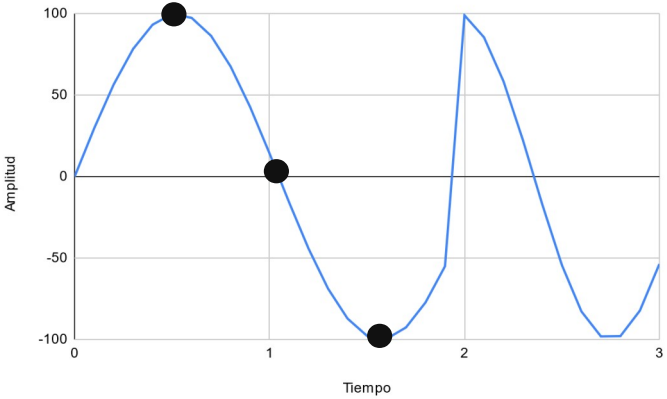
Muestreo

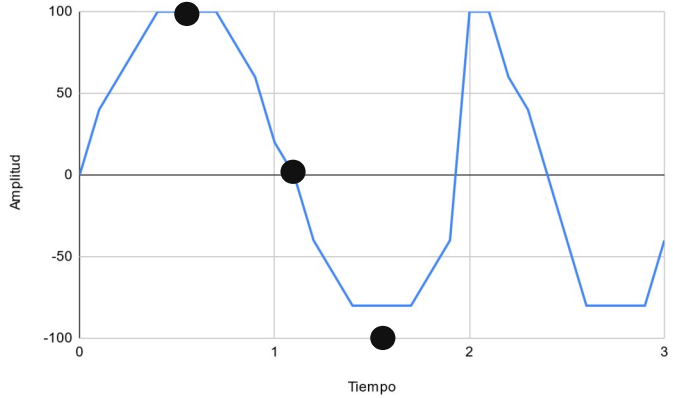

100
50
0
-50
-100
Amplitud
0
1
2
3
Tiempo

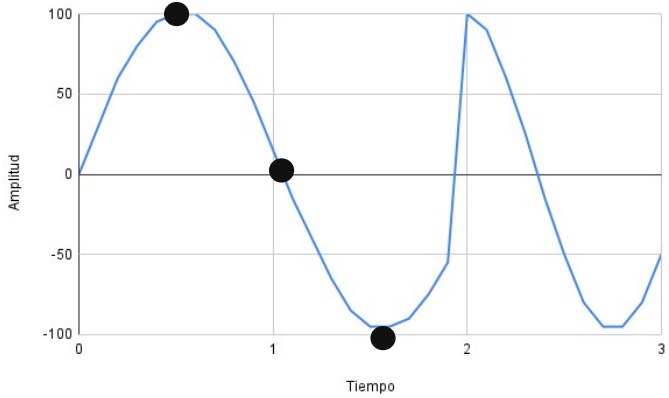

100
50
0
-50
-100
Amplitud
0
1
2
3
Tiempo

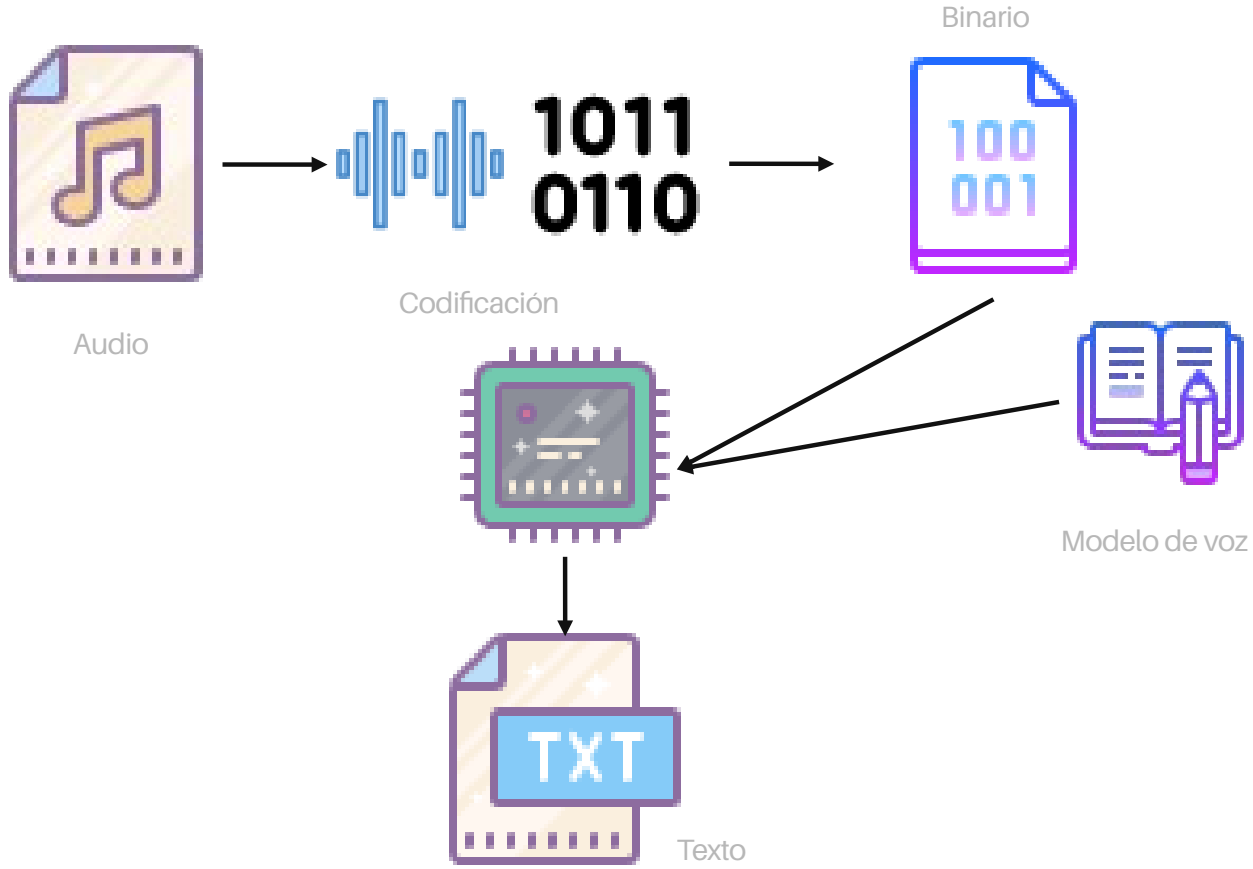

100
50
0
-50
-100
Amplitud
0
1
2
3
Tiempo

Precisión


100
50
0
-50
-100
Amplitud
0
1
2
3
Tiempo

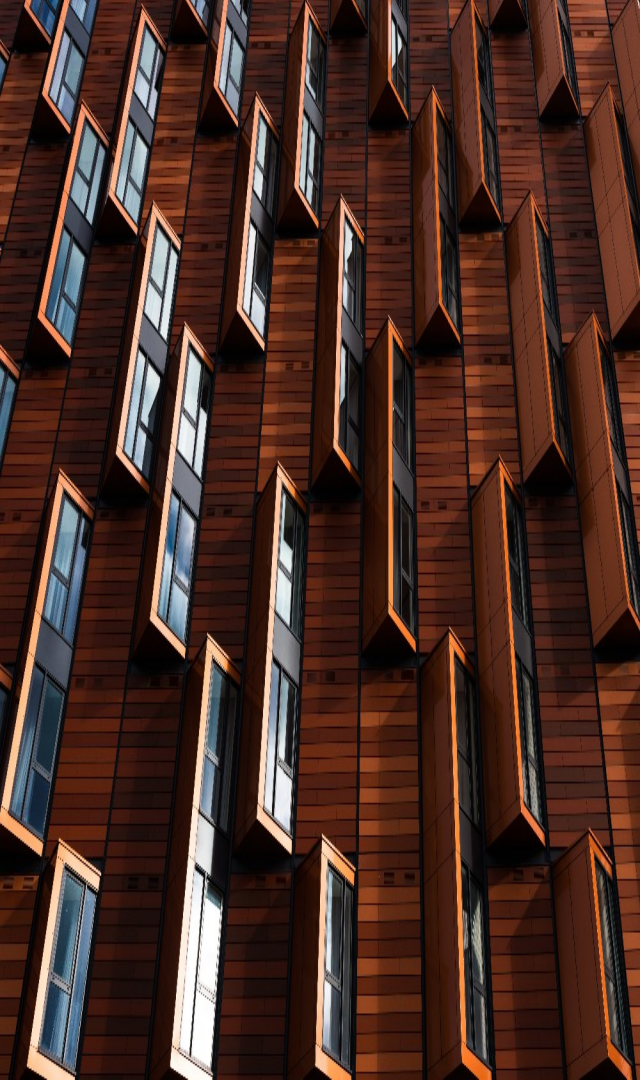

100
50
0
-50
-100
Amplitud
0
1
2
3
Tiempo


100
50
0
-50
-100
Amplitud
0
1
2
3
Tiempo



El Speech Recognition


¿Más código? Más 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)

Sobre el marco legal y la responsabilidad en la IA

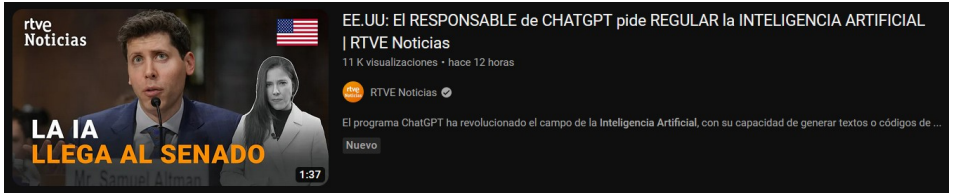
WHAT IS A RAIL LICENSE?

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.

- +info

Nerea Luis - "Computer says no"

95 visualizaciones • hace 6 meses



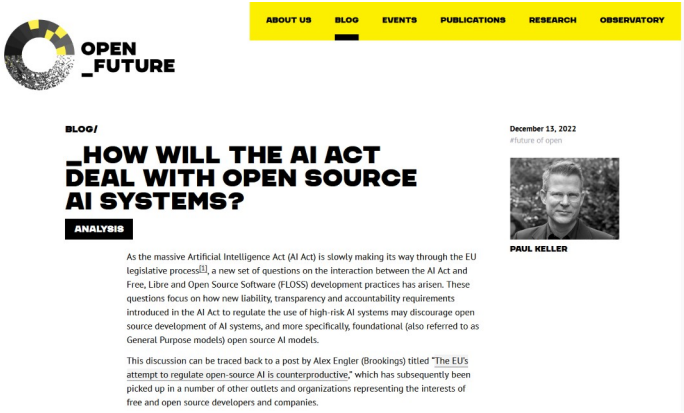
EE.UU: El RESPONSABLE de CHATGPT pide REGULAR la INTELIGENCIA ARTIFICIAL | RTVE Noticias

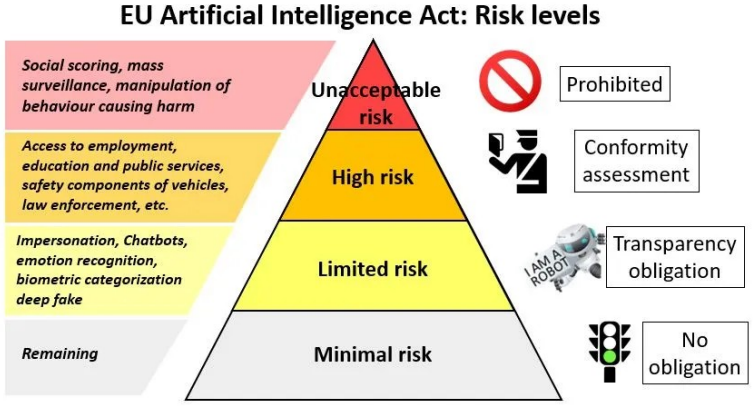
11 K visualizaciones • hace 12 horas

RTVE Noticias

El programa ChatGPT ha revolucionado el campo de la **Inteligencia Artificial**, con su capacidad de generar textos o códigos de ...

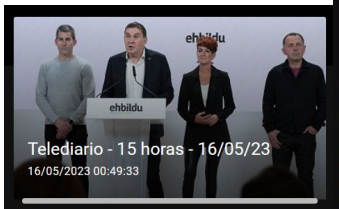
Nuevo





Muchos datos
= Necesita mucha potencia
= Contamina demasiado

• +info



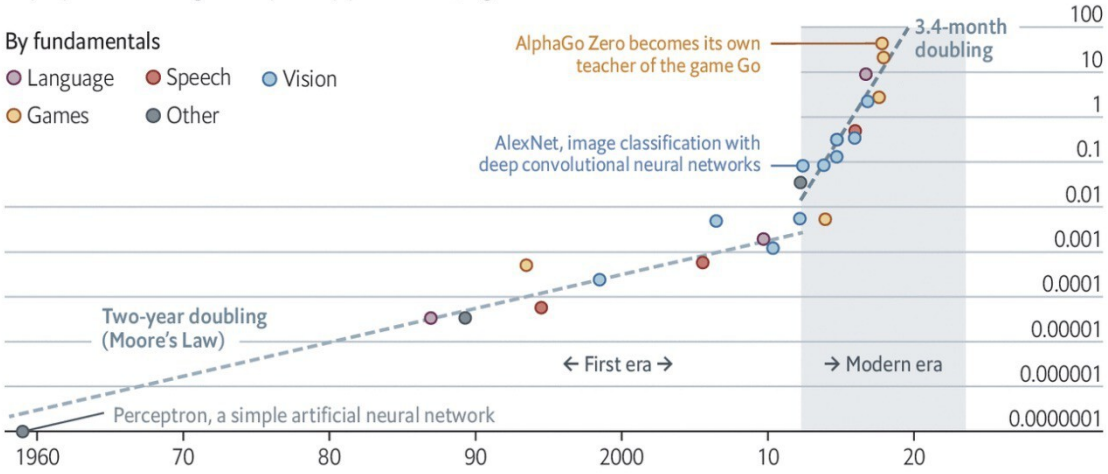
Elena Guidi - Salvemos los pingüinos con el green computing

72 visualizaciones • hace 6 meses

Deep and steep

Computing power used in training AI systems

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



Source: OpenAI

*1 petaflop=10^{15} calculations

The Economist

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

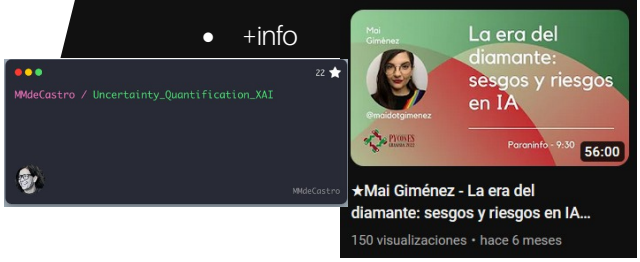
If you survey the major voice assistants on the market—Alexa, Apple's Siri, Microsoft's Cortana, and Google Home's unnamed character—three out of four have female-sounding names by default, and their voices sound female, too. Even before the user addresses Alexa, the robot has already established itself 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: themself ("*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).



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).

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.

Público, ¿qué habéis aprendido?

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!