



Entendiendo los retos de las tecnologías del habla

Nieves Ábalos
Jaca, 13 de julio de 2022

 **monoceros**
LABS



Nieves Ábalos Serrano

Chief Product Officer & Cofundadora [Monoceros Labs](#)

Twitter: [@nieves_as](#)

LinkedIn: [/in/nievesabalosserrano](#)

- **Universidad de Granada:** Ingeniera Informática, Máster y PhD sin terminar en sistemas de diálogo.
- **Dpto Innovación BEEVA** (BBVA Next Technologies): I+D en conversacional y NLP.
- **Monoceros Labs:** *emprendedora*, estrategia y gestión de producto, diseño conversacional y desarrollo conversacional (lo que toque).
- **Comunidad:**
 - **Women in Voice Spain (Lead)**, 
 - **Alexa Champion**, 
 - **AWS Community Builder**. 

Voice-first y Multimodalidad

Creamos **interfaces conversacionales adaptadas** a los dispositivos del mercado, diseñadas para la **interacción por voz** como primer modo de interacción, y con apoyo en pantallas (multimodalidad).



Alexa Skill: Pasapalabra (2021)



Más videos en el
[Canal de Youtube](#) de
Monoceros Labs

<https://www.youtube.com/watch?v=HsJf5M-16EI>

Tecnologías del habla

Iniciandonos en los retos

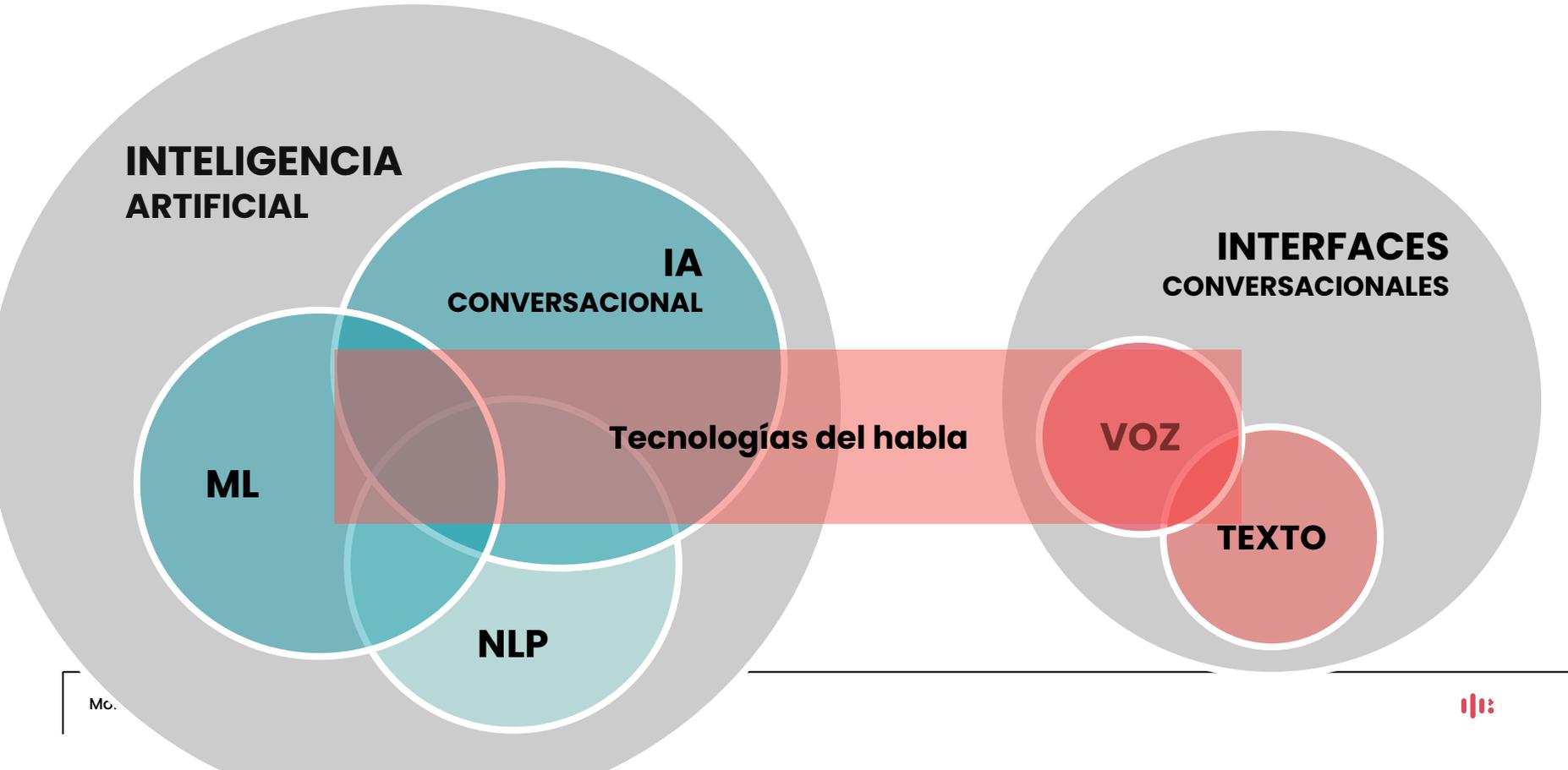


¿Qué son las tecnologías del habla?

Es el **conjunto de técnicas** que permiten la **interacción** entre personas y máquinas mediante el **uso de la voz**.

Además se engloban en las Tecnologías del Lenguaje.

Inteligencia Artificial + Canal / Modo



IA Conversacional



Conversational AI: Computers That Talk

Conversational AI systems are computers that people can interact with simply by having a conversation, our most natural form of interaction. In short, it is what allows us to talk to voice-driven technologies like Amazon Alexa and ask about the weather, order products online, and even call a cab, simply by using the language we already know.

With conversational AI, voice-enabled devices like Amazon Echo are finally starting to enable the sort of magical interactions we've dreamed of for decades (think: Star Trek computer). Through a [voice user interface \(VUI\)](#), voice services like Alexa can communicate with people in ways that feel effortless, solve problems, and get smarter over time.



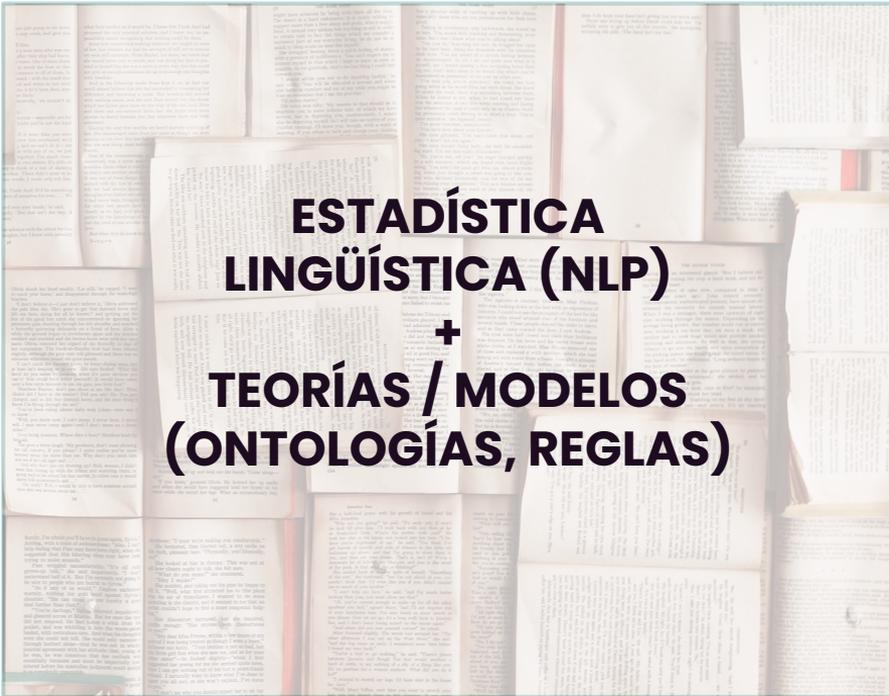
Teaching Computers to Converse

Conversations can be conceptually and emotionally complex; they entail much more than simple input and output. When we talk to each other, how we say something also matters as much as *what* we say.

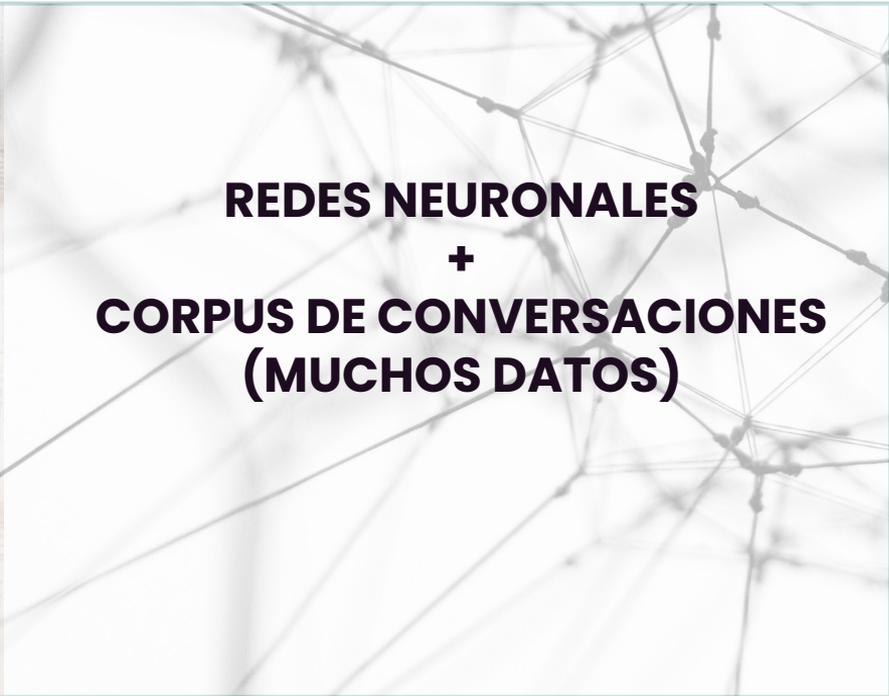
Computers can't grasp these nuances, and that's where voice design comes in. A well-designed VUI is flexible, and takes into account these unwritten rules of conversation. It enables computers to think and talk as we do, and not like a robot.

El objetivo es hacer que las personas interactuemos de **manera natural conversando con la tecnología**, teniendo en cuenta que las conversaciones pueden ser emocionalmente complejas, además de las reglas no escritas de la comunicación.

Tecnologías: **ayer** y **hoy**



**ESTADÍSTICA
LINGÜÍSTICA (NLP)
+
TEORÍAS / MODELOS
(ONTOLOGÍAS, REGLAS)**



**REDES NEURONALES
+
CORPUS DE CONVERSACIONES
(MUCHOS DATOS)**

El proceso comunicativo

Tecnologías del habla

**¿Habéis hablado
con Alexa
alguna vez?**

¿Cómo conversamos con las interfaces de voz?

A través de un intercambio de turnos, usando la voz o la escritura, y sobre todo, el **lenguaje natural**.

Aparecen retos en la comunicación:

- **contexto** de la conversación
- **semántica** de la frase
- canal: ruido
- ...

(¡que ya existen entre personas!).



¿Cómo conversamos con estas interfaces?



Objetivo o
intención

diálogo (o sesión)

hola, bienvenida
a **tu biblioteca musical**.
¿en qué te puedo ayudar?

TURNO DE ALEXA

prompt (con pregunta)

pon música

TURNO DE ELLA

frase o utterance

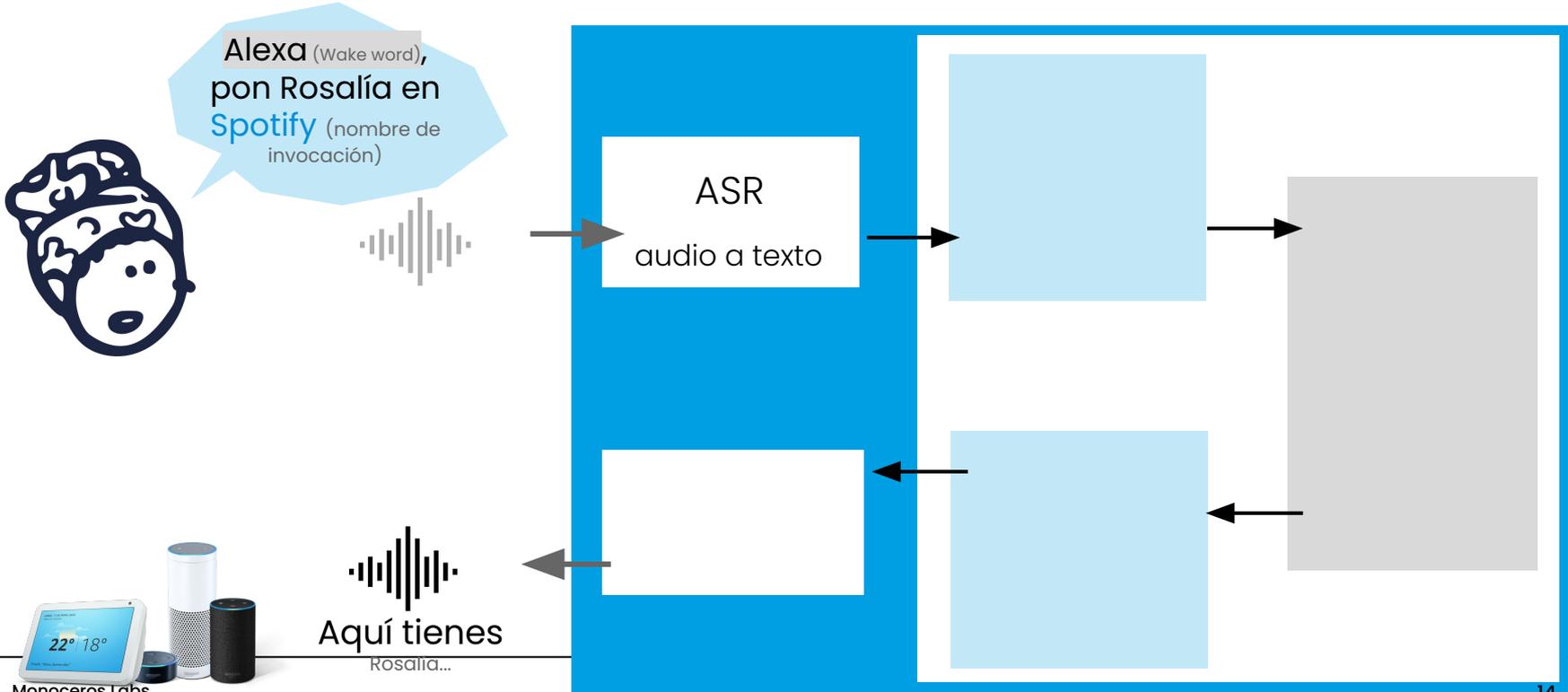
Claro, reproduciendo música
aleatoria desde *Spotify*...

TURNO DE ALEXA

prompt (y acción)



Las fases



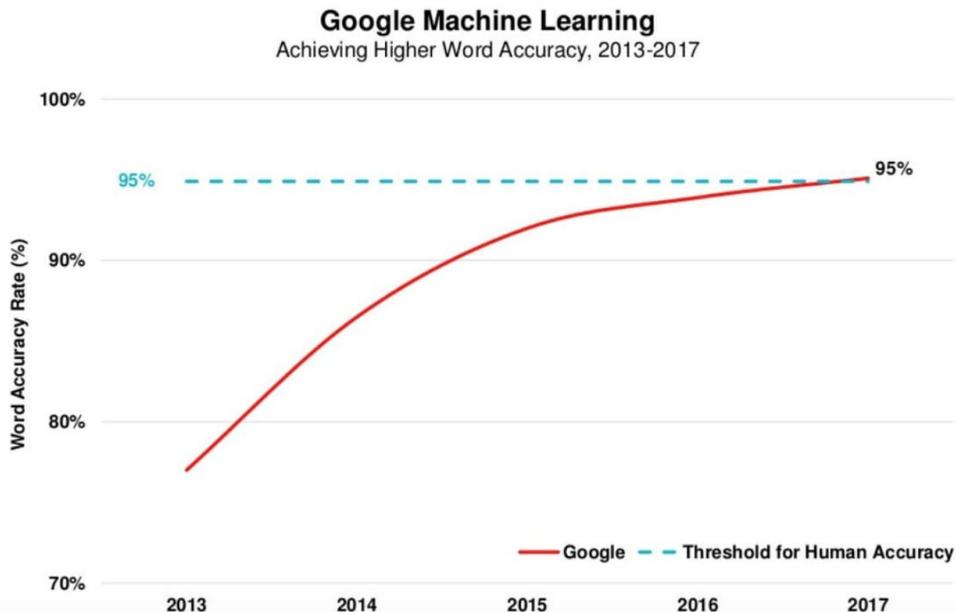
2013... ¡ASR en el umbral de la capacidad humana!



La mejora en la capacidad de cómputo (nube) **redujo el tiempo** de procesamiento de **enormes cantidades de datos** (speech) >> mejores modelos de ASR



Computing power and artificial intelligence are largely behind the advances in this space. With massive amounts of speech data combined with faster processing, speech recognition has hit an inflection point where its capabilities are roughly on par with humans. The graph below is from [Mary Meeker's 2017 Internet Trends report](#). It plots Google's word accuracy rate which recently broke the 95% threshold for human accuracy.



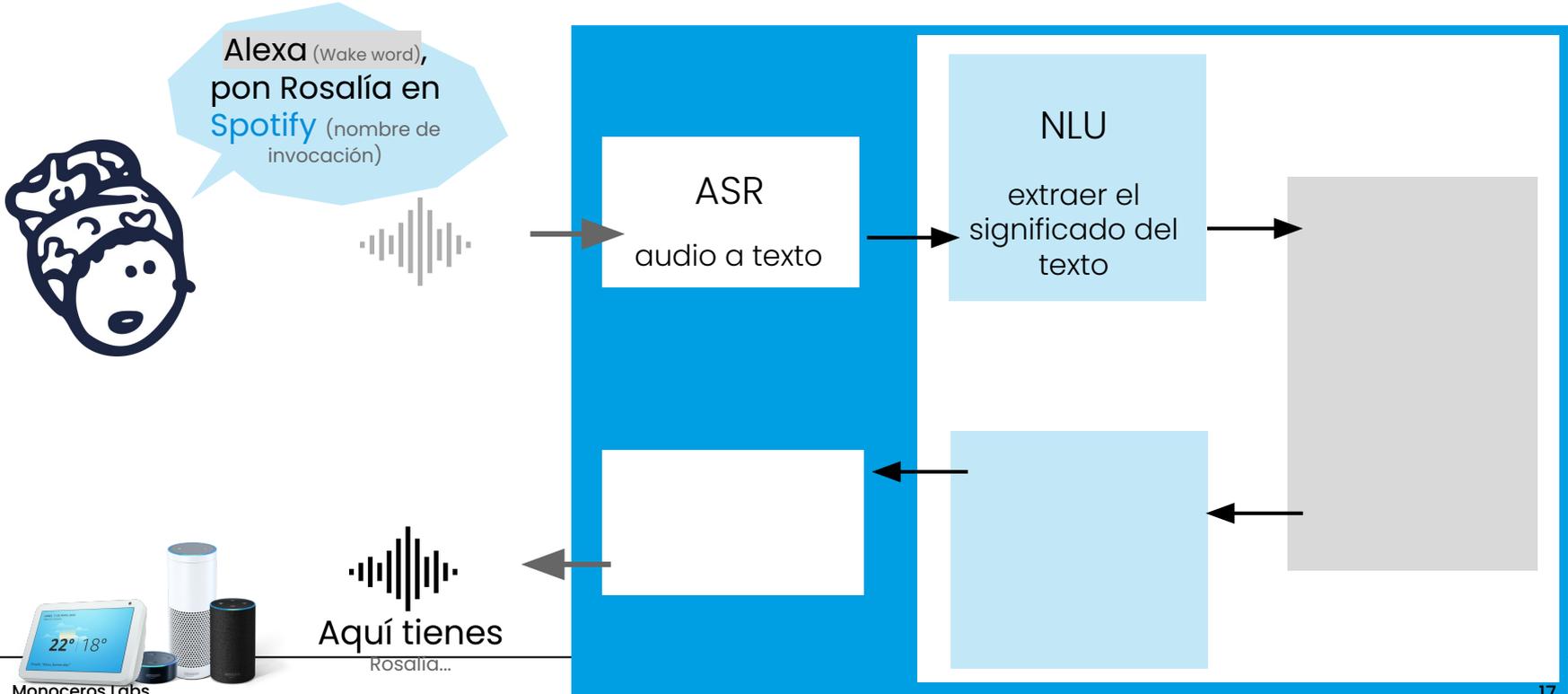
El ruido y

The Cocktail Party Problem

El reto del ASR: tratar de identificar al hablante y separar lo que dice, en entornos de mucho ruido (como una fiesta).



Las fases



El NLU es un **modelo** que entrenamos

nombre de invocación / invocation name

Alexa,

wake word

abre **Spotify**...

... y pon música.

frase o utterance

... y quiero escuchar a Rosalía.

frase o utterance

frases de entrenamiento para el NLU

El NLU se **entrena** así

frase o *utterance*:

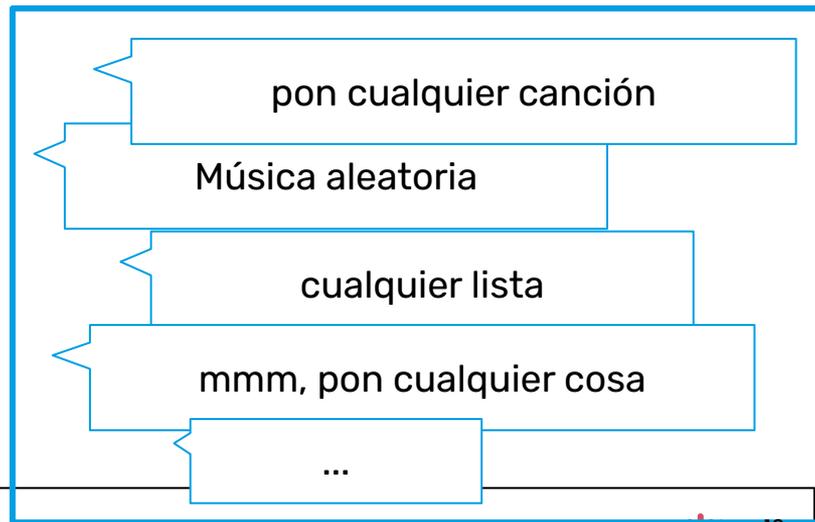
pon música

intención del usuario o *intent*

escucharRandom

Clasificamos diferentes frases según la intención del usuario:

- aportamos semántica,
- indicamos la información relevante al caso de uso (entidades).



El NLU se **entrena** así

frase o *utterance*:

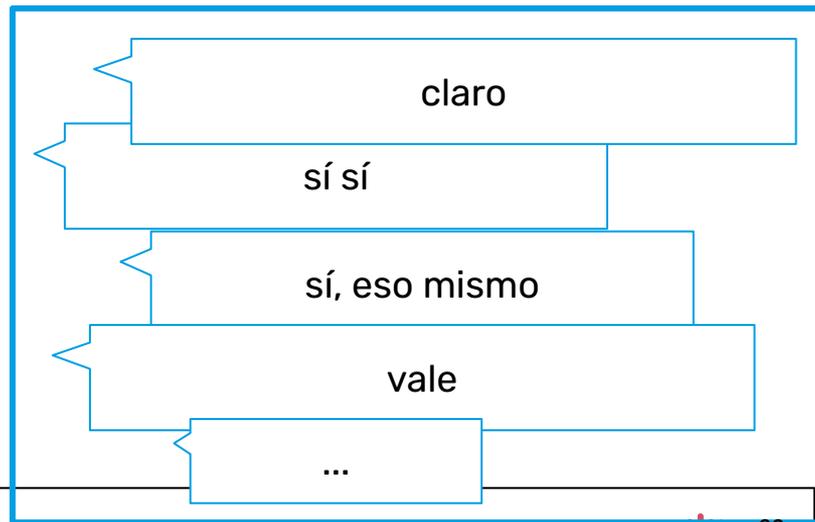
Sí

intención del usuario o *intent*

decirSi_____

Clasificamos diferentes frases según la intención del usuario:

- aportamos semántica,
- indicamos la información relevante al caso de uso (entidades).



El NLU se **entrena** así

frase o *utterance*:

quiero escuchar a

Rosalía

valor

intención del usuario o *intent*

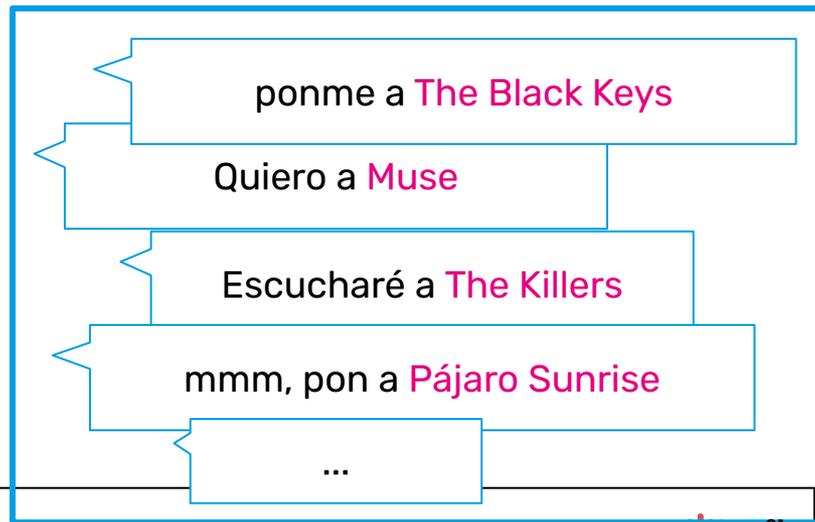
escucharArtista

entidad o *slot*

artista

Clasificamos diferentes frases según la intención del usuario:

- aportamos semántica,
- indicamos la información relevante al caso de uso (entidades).



El NLU es un **modelo** que contiene

intención: escucharArtista

quiero escuchar {artista}
pon a {artista}
...

frases de entrenamiento

intención: decirSi

sí
vale
...

frases de entrenamiento

entidad: artista

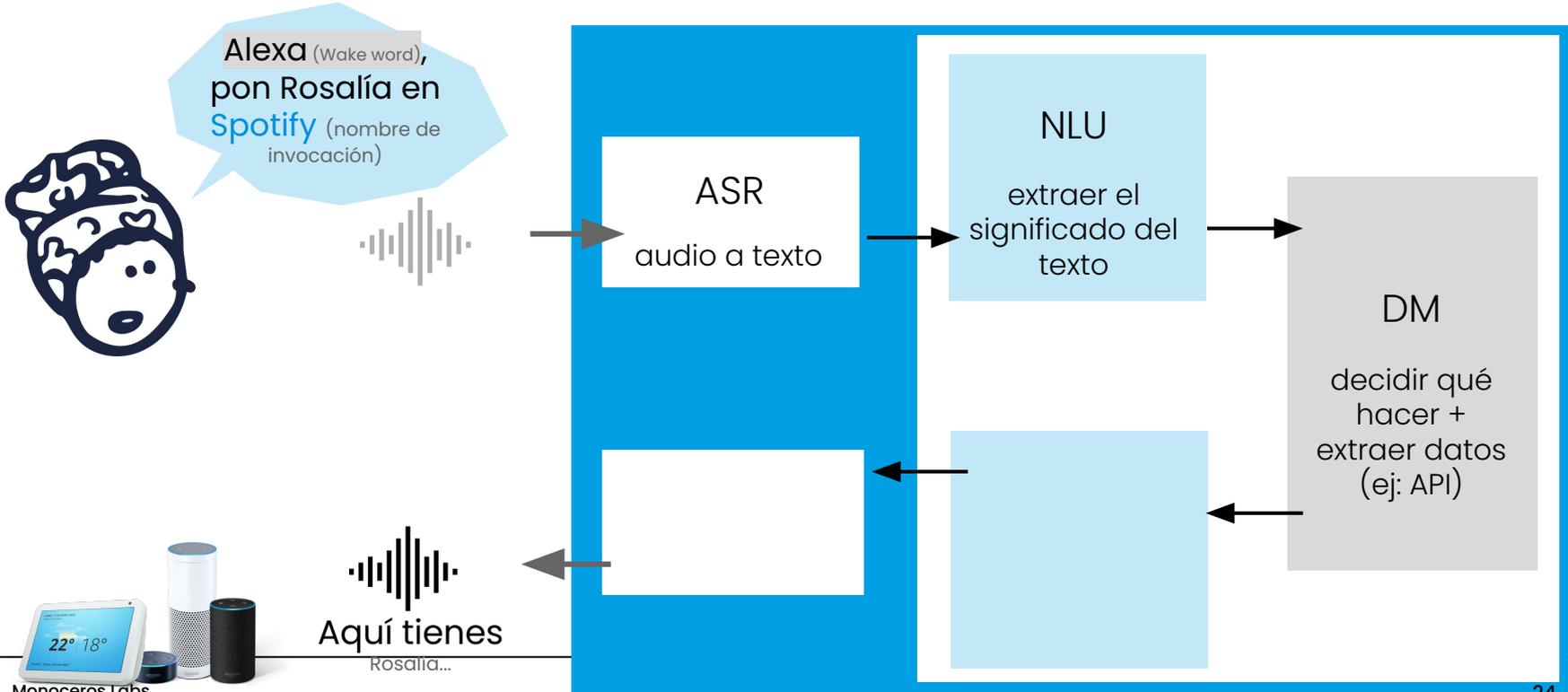
Bob Dylan	Bob, Dylan...
The Beatles	los escarabajos
...	...

valores y **sinónimos**

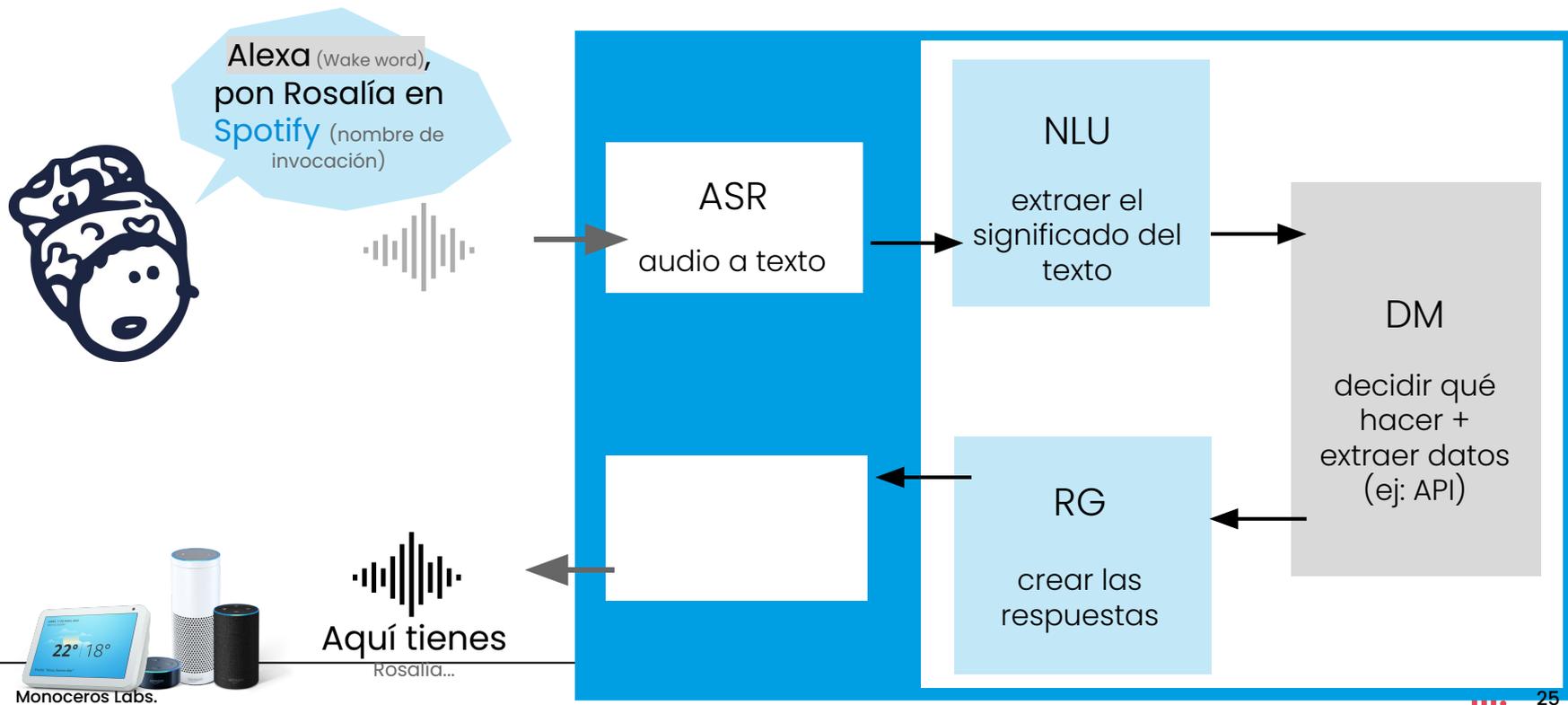
El NLU contiene: **todas** las *utterances* del caso de uso, clasificadas de la manera más óptima en los *intents* (o cajas) de nuestro caso de uso, y **todas** las *entidades* con los *valores* que podemos entender.

**Entrenar un modelo
NLU = enseñarle a
entender cosas
nuevas**

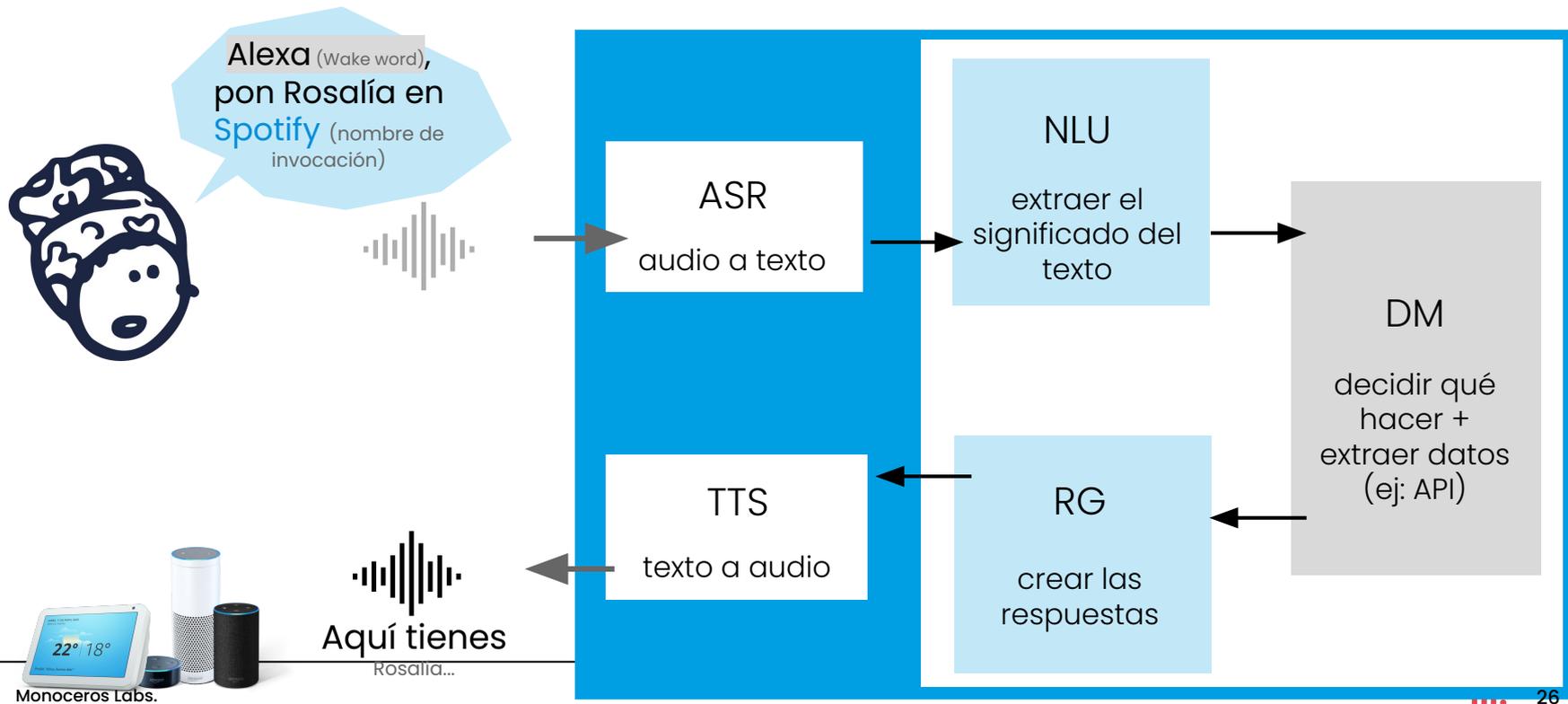
Las fases



Las fases



Las fases



Voiceflow 101

Cómo crear una interfaz de voz de manera sencilla.

Voiceflow: introducción

The screenshot shows the Voiceflow website homepage. At the top, there is a navigation bar with the Voiceflow logo, menu items for 'By Role', 'Resources', 'Enterprise', 'Pricing', and 'Company', and buttons for 'Login' and 'Sign Up'. A blue banner below the navigation bar reads 'NEW Explore Our New Templates Library – Get Started Today'. The main heading is 'Where teams create Voice Assistants', with 'Voice Assistants' in a large blue font and a small blue box containing the name 'Sam' next to it. Below the heading, a sub-headline states 'Voiceflow helps teams design, prototype and build conversational assistants, faster.' There are two primary call-to-action buttons: 'Get Started - It's free' and 'Watch video 2 mins'. At the bottom, three bullet points list features: 'Collaborative no-code design', 'Supports every NLU', and 'Product support'.

voiceflow.com

NEW Explore Our New Templates Library – Get Started Today

Voiceflow

By Role Resources Enterprise Pricing Company

Login Sign Up

Where teams create Voice Assistants

Sam

Voiceflow helps teams design, prototype and build conversational assistants, faster.

Get Started - It's free

Watch video
2 mins

- ✓ Collaborative no-code design
- ✓ Supports every NLU
- ✓ Product support

Voiceflow: introducción

- ¿Qué podemos hacer con Voiceflow?
 - Prototipos y desarrollos de aplicaciones de voz para Alexa, Google Assistant y otros asistentes personalizados.

Voiceflow: introducción

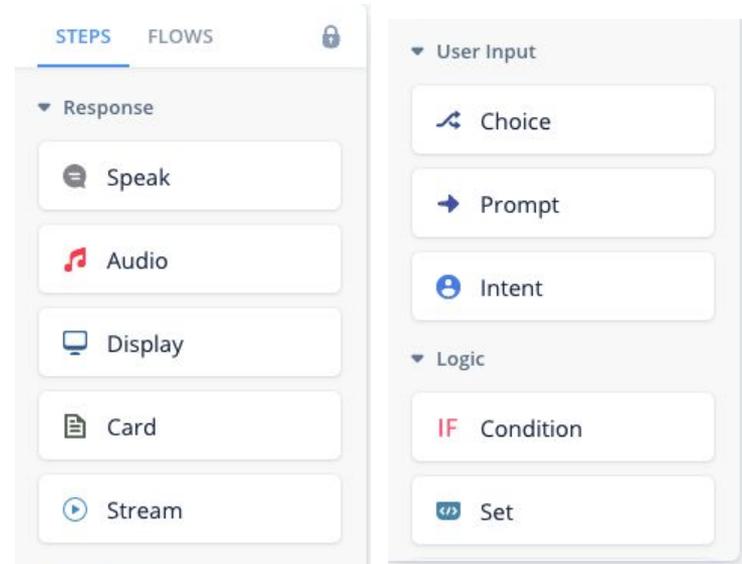
- ¿Qué podemos hacer con Voiceflow?
 - Prototipos y desarrollos de aplicaciones de voz para Alexa, Google Assistant y otros asistentes personalizados.
- ¿Cómo lo hacemos?
 - Con el modelo NLU, llamado **Model** (propio de Voiceflow, pero conectado con los intents por defecto de Alexa o Google Assistant).
 - Creamos **intents** con **utterances** y **slots**. Los slots extraídos de la conversación son **variables**.
 - Con el **comportamiento** de la app de voz, conectando componentes para crear la estructura de la conversación.

Voiceflow: componentes

Los componentes de Voiceflow son:

- **Steps** (o Bloques), divididos en:
 - Response
 - User Input
 - Logic
- **Flows** (agrupaciones de Steps)

Conectamos Steps mediante **flechas** (*) que representan **paths** o caminos en la conversación.



Voiceflow: las vistas de **diseño** y test

The screenshot displays the Voiceflow design interface for a chatbot. At the top, the header includes the Voiceflow logo, navigation icons, the text "Academia de Inglés / Home", and buttons for "Share", a checkmark, and a play button.

The main workspace shows a flowchart with the following components:

- Response Block:** Contains the text "¡Hola! Te doy la bienvenida al chatbot de la academia. ¿En qué puedo ayudarte?".
- ELEGIR CURSO Block:** A choice block with three options: "pedir_curso_adultos", "pedir_curso_seniors", and "pedir_curso_ninos".
- New Block 3:** A response block with the text "Curso adultos".
- New Block 3 copy copy:** A response block with the text "Curso seniors".
- New Block 3 copy:** A response block with the text "Curso niños".

Arrows indicate the flow from the initial response to the "ELEGIR CURSO" block, and then from each of its three options to the corresponding "New Block 3" response blocks.

On the left side, a sidebar menu is visible with the following categories and items:

- Steps:** Response (Speak, Audio, Display, Card, Stream), User Input (Choice, Prompt, Intent), Logic, Integration.
- Flows:** (locked)

At the bottom right, there is a chatbot icon with a notification badge and a red bar with the number "32".

Voiceflow: las vistas de diseño y test

< Back esc to return to designer 00:00 Share Prototype

TEST SETTINGS

- Debug Mode** Show the paths, variables and flows you're using while you test
- Intent Confidence Score** Show the natural language model's confidence out of 1.00 when matching trained intents.
- Guided Navigation** Stop and wait for path selection on IF and Custom Action steps.

ELEGIR CURSO

- pedir_curso_adultos
- pedir_curso_seniors
- pedir_curso_ninos

DIALOG

TRAINING

Run your project

Run your project to interact with it using text, voice or buttons. See more.

Run Test

Recuerda poner el "Debug mode" y el "Intent Confidence Score" al testear en la web de Voiceflow.



TTS

Voiceflow 101

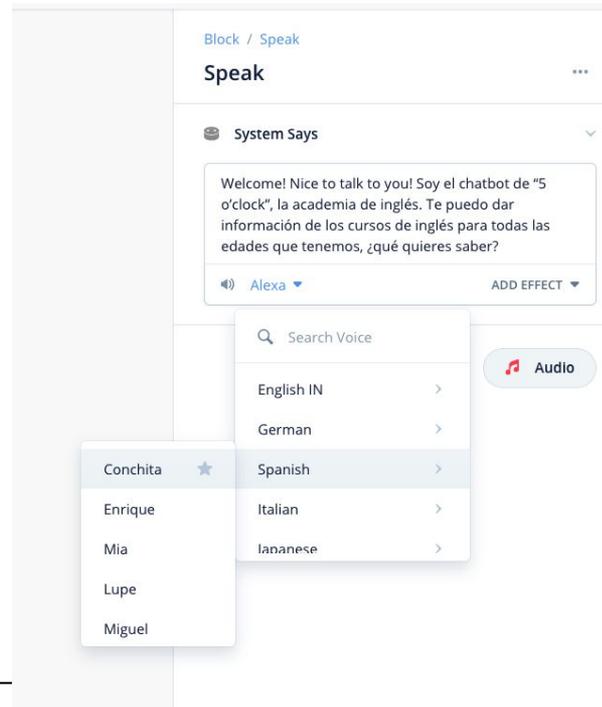
Voiceflow: step o bloque **Speak**

El bloque básico, en el que el asistente habla.

En él, se puede añadir audios, elegir la voz del TTS (en el idioma que sea)...

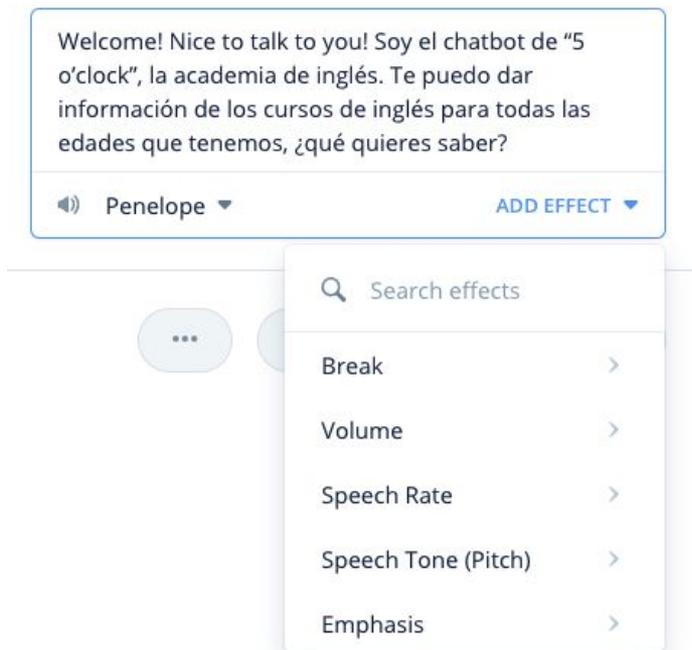
NOTA: Para español, si dejamos por defecto la voz de Alexa, en la consola de Voiceflow suena en spanglish.

Hay que cambiar a una voz española (ej: Conchita) y marcarlo como estrella (para no tener que estar cambiando todos los Speaks de aquí en adelante).



Voiceflow: step o bloque **Speak**

... y añadir efectos a la voz, que consiste en adaptarla con SSML (lenguaje de marcado para el habla).



SSML

Speech Synthesis Markup Language (W3C)



Speech Synthesis Markup Language (SSML) Version 1.1

W3C Recommendation 7 September 2010

This version:

<http://www.w3.org/TR/2010/REC-speech-synthesis11-20100907/>

Latest version:

<http://www.w3.org/TR/speech-synthesis11/>

Previous version:

<http://www.w3.org/TR/2010/PR-speech-synthesis11-20100223/>

Editors:

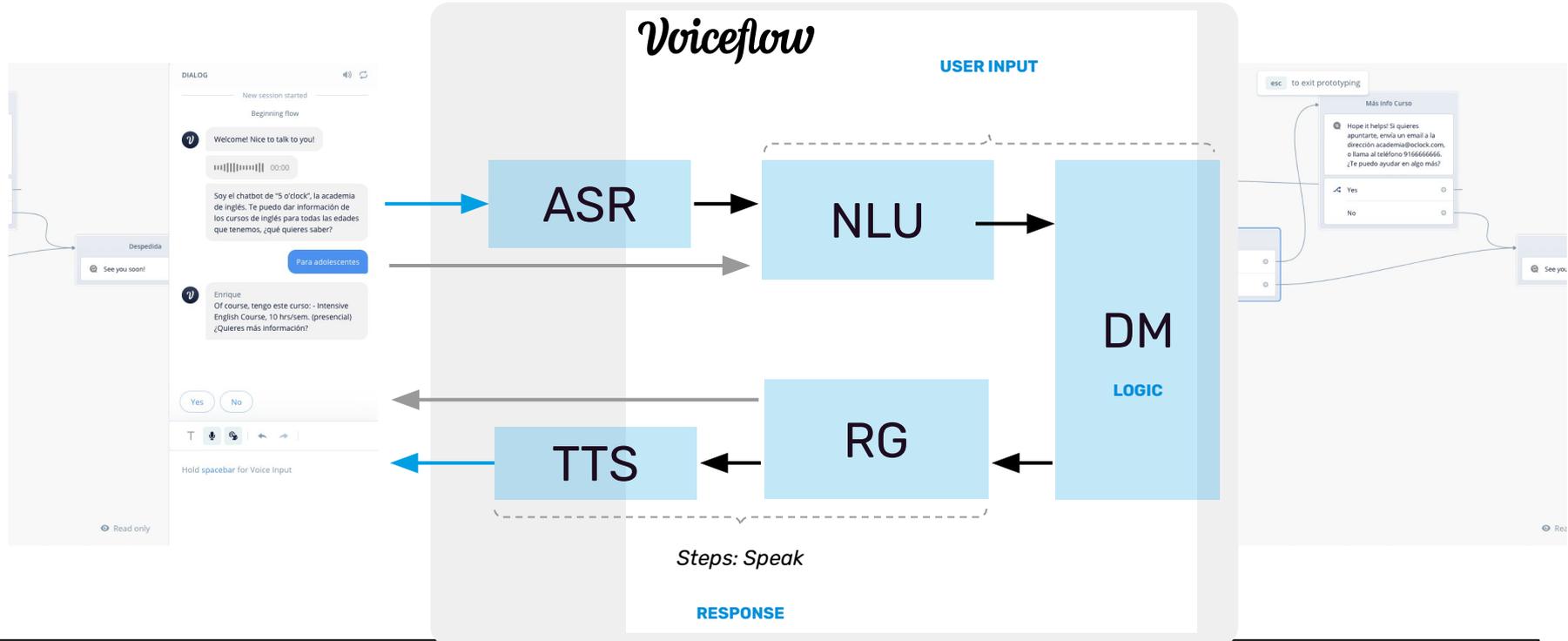
Daniel C. Burnett, Voxeo (formerly of Vocalocity and Nuance)
双志伟 (Zhi Wei Shuang), IBM

Authors:

Paolo Baggia, Loquendo
Paul Bagshaw, France Telecom
Michael Bodell, Microsoft
黄德智 (De Zhi Huang), France Telecom
楼晓雁 (Lou Xiaoyan), Toshiba

IPA: <https://tophoneics.com/>

Testeando en Web (voz y texto)



NLU

Voiceflow 101

Voiceflow: el modelo NLU (Model)

The screenshot displays the Voiceflow interface for a chatbot named "Academia de Inglés". The interface includes a top navigation bar with a logo, a search icon, a text input field, a microphone icon, a chat icon, and a grid icon (highlighted with a red circle). The main workspace shows a flowchart with several steps:

- Initial Step:** A text block containing the message: "¡Bienvenido! Te doy la bienvenida al chatbot de la academia 5 o'clock. ¿En qué puedo ayudarte?".
- ELEGIR CURSO:** A choice step with three options: "pedir_curso_adultos", "pedir_curso_seniors", and "pedir_curso_ninos".
- Flow Transitions:** Arrows connect the choice step to three separate "New Block 3" blocks, each containing a "Curso" option.
- Final Blocks:** Three "New Block 3" blocks, each containing a "Curso" option: "Curso adultos", "Curso seniors", and "Curso niños".

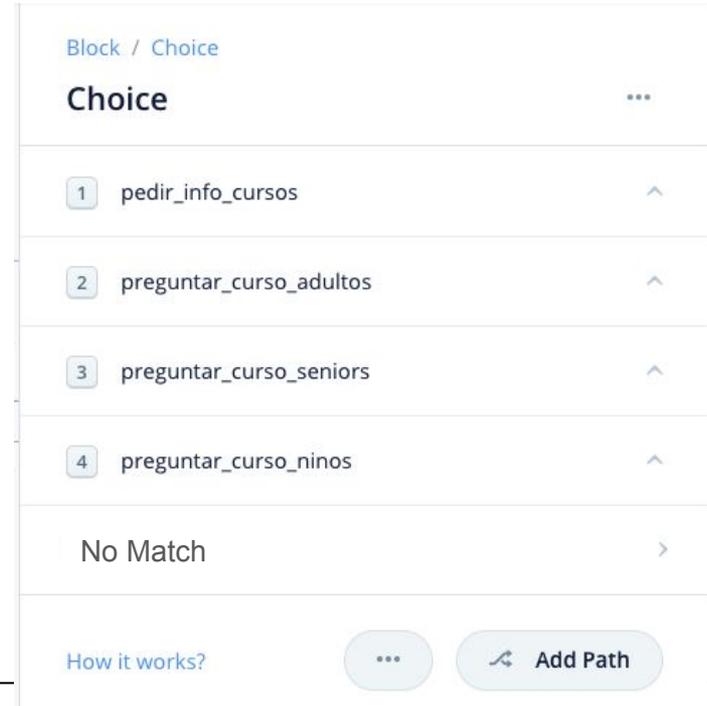
The left sidebar shows a "Steps" panel with various action types under "Response" (Speak, Audio, Display, Card, Stream) and "User Input" (Choice, Prompt, Intent). The bottom right corner features a chatbot icon with a notification badge and a red "1" indicator.

Voiceflow: step o bloque **Choices** (1)

Son las elecciones que presentamos al usuario.

Por defecto, con **path 1** (Paths de la conversación) y **no match**.

Permite que la conversación no termine, pues esperamos una respuesta del usuario.



The screenshot shows the configuration for a 'Choice' block in Voiceflow. At the top, it says 'Block / Choice' and 'Choice' with a three-dot menu icon. Below this, there are four numbered paths, each with an upward arrow icon:

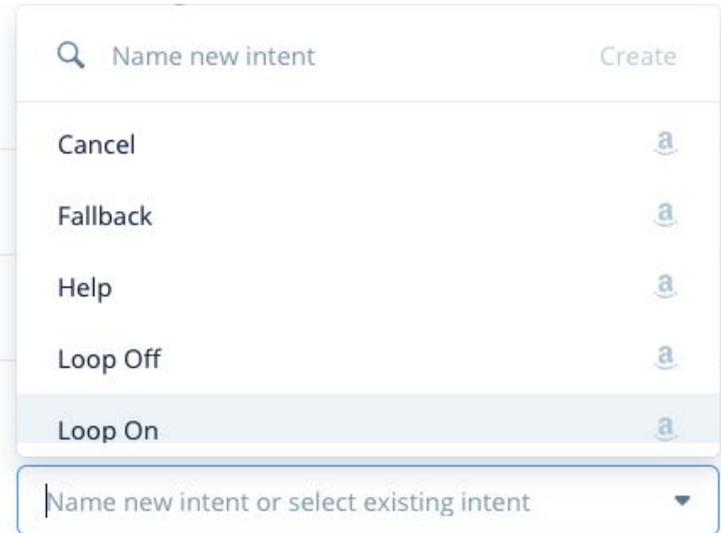
- 1 pedir_info_cursos
- 2 preguntar_curso_adultos
- 3 preguntar_curso_seniors
- 4 preguntar_curso_ninos

Below the numbered paths is a 'No Match' option with a rightward arrow icon. At the bottom of the interface, there is a 'How it works?' link, a three-dot menu icon, and an 'Add Path' button with a plus icon.

El NLU de Voiceflow: **Intents**

Podemos crear Intents propios de nuestro caso de uso, o utilizar los que vienen por defecto ya creados.

Si elegiste Alexa, podrás usar los intents por defecto de esta plataforma (ver el símbolo de Amazon a la derecha de cada nombre).



Un intent creado en un paso, es reutilizable en otros pasos (lo puedes ver en la vista [Model](#)).

El NLU de Voiceflow: **Intents**

The screenshot shows the configuration for an intent named "preguntar_curso_adultos". At the top, the intent name is displayed with a dropdown arrow. Below it, a search bar contains the text "preguntar_curso_adultos". The "Utterances" section is expanded, showing three user responses: "para adolescentes", "qué hay de inglés para adultos", and "Necesito un curso de inglés para adultos". Each utterance has a minus sign icon to its right. The "Slots" section is collapsed, indicated by an upward arrow.

¡Ojo con las eñes en los nombres de los intents!

Voiceflow: step o bloque Choices (2)

No Match: lo que sucede cuando el usuario dice cualquier otra cosa de las opciones que le damos.

Seleccionar el tipo de acción:

- PATH: Redirigimos al usuario al camino o paso correcto.
- REPROMPTS: ponemos las respuestas en orden / o aleatorio.

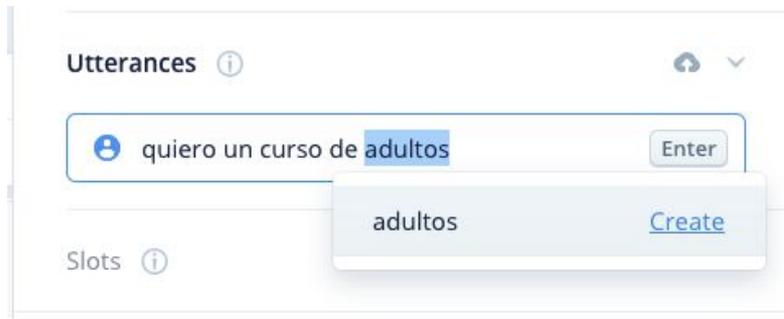
The screenshot shows the configuration for a 'No Match' action in Voiceflow. At the top, there are navigation links: 'Block / Choice / No Match' and a three-dot menu. Below this, the 'No Match Type' is set to 'Reprompts' (checked) and 'Path' (unchecked). A dropdown menu is open for 'Reprompt 1', showing a text box with the prompt: 'Perdona, no te entendí, dime adultos, seniors o niños. ¿Qué tipo de curso quieres?'. Below the text box, there is a speaker icon, the name 'Conchita', and an 'ADD EFFECT' button with a downward arrow. At the bottom of the interface, there are three buttons: a three-dot menu, 'System', and 'Audio'.

El NLU de Voiceflow: Slots

Desde Voiceflow, podemos crear **slots** para extraer información que nos digan los usuarios.

Subrayamos en la utterance lo que queremos que sea un slot.

Un Slot creado en un paso, es reutilizable en otros pasos (lo puedes ver en la vista **Model**).



El NLU de Voiceflow: Slots

The screenshot displays the Voiceflow development environment. On the left, the 'VARIABLES' panel lists: TIPO_CURSO niños, TIMESTAMP 0, SESSIONS 1, USER_ID TEST_USER, and PLATFORM alexa. The central workspace shows a workflow diagram with three main steps: 'Bienvenida' (Welcome! Nice to talk to you!), 'Más información General' (Ahora mismo tenemos tres cursos...), and 'Elección Curso' (preguntar_curso). A 'esc to exit prototyping' button is visible at the top. On the right, the 'DIALOG' panel shows a simulation where the user input 'Niños' is matched to the 'preguntar_curso' slot with 100.00% confidence. The system response lists three course options: 'cursos Adultos', 'cursos Seniors', and 'cursos Niños'. The 'Send' button at the bottom right is highlighted.

El NLU de Voiceflow: Slots

The image shows the Voiceflow interface with a 'NEW SLOT' dialog box open on the left and a 'Choice' block configuration on the right.

NEW SLOT Dialog:

- Slot Name:** {tipo_curso} (highlighted in red)
- Slot Name Input:** tipo_curso
- Slot Type:** Custom
- Values:**
 - adultos
 - adolescentes, adulto, adolescente
 - seniors
 - mayores de cincuenta y cinco, ser
 - niños
 - niño, jóvenes, jóven
- Add custom value:** (link)
- Slot Color:** A row of colored circles with a red checkmark on the pink one.
- Create Slot:** (button)

Choice Block Configuration:

- Block / Choice:** Choice
- Utterances:**
 - 1 preguntar_curso
 - preguntar_curso
 - quiero un curso de adultos
- Slots:** adultos (highlighted in a tooltip with a 'Create' button)
- Else:** Reprompt
- How it works?:** Add Path

El NLU de Voiceflow: Slots

Utilizamos `{ }` para el nombre del slot, añadimos valores y sinónimos.

Pueden ser obligatorios u opcionales (como en Dialogflow).

Consejo: En las utterances, añade variaciones en el orden de los slots, para que recoja las diferentes frases que el usuario

puede decir.
Monoceros Labs.

The screenshot shows the configuration for a slot named 'preguntar_curso'. It includes a search bar with the slot name, an 'Utterances' section with a button to 'Add synonyms of the user response', and a list of utterances. One utterance is 'quiero un curso de {tipo_curso}', where the slot name is highlighted in a red box. Below this, the 'Slots' section shows that '{tipo_curso}' is 'is not required'.

¿Qué es un **reprompt**?



En Alexa (y VUIS) tenemos una limitación técnica: **el tiempo que está abierto el micrófono** para escuchar la respuesta del usuario cuando le toca decir algo. Es de unos 7 u 8 segundos.

¿Le da tiempo al usuario a responder? Pues depende...

En este ejemplo, usamos el **reprompt** para dar más tiempo al usuario a responder.

¿Qué es un **reprompt**?

Reprompt: la acción de volver a preguntar o decir lo último que se dijo al usuario cuando hay un problema de comprensión (ASR/NLU).

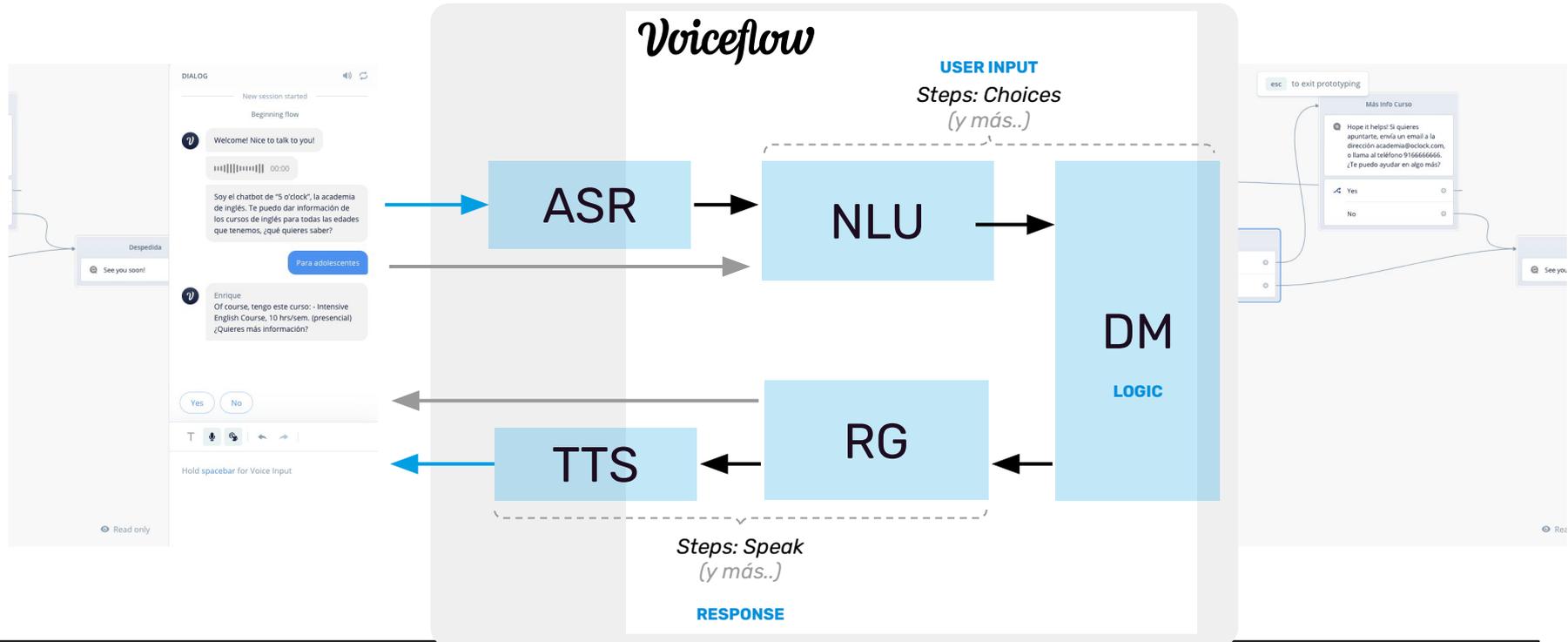
En **Alexa** (plataforma): es la **respuesta** que da cuando..

- el usuario no dice nada / Alexa no escucha nada
- el usuario dice algo que no está en nuestro NLU, y no tenemos Fallback.

¿En **Voiceflow**? Un poco lío...

El reprompt es sinónimo de Fallback para los Choices/Else. También se puede añadir un "No Response". También si elegiste Alexa, podrás añadir el intent Fallback.

Testeando en Web (**voz** y texto)



NLU + DM

Voiceflow 101

El NLU de Voiceflow: Slot Filling

Usuario: quiero información del curso

IC: ¿de qué curso? (falta {tipoCurso})

Usuario: del curso de niños

IC: ok, para niños tengo...

El NLU de Voiceflow: Slot Filling

Cuando trabajamos con **Slots** o **Entities**, quizás sea necesario recogerlas de manera obligatoria para poder continuar la conversación.

Marcaremos el Slot como **Required Entities**.

The screenshot displays the Voiceflow NLU configuration interface. On the left, a list of intents is shown, with 'dar_noticias_categoria' selected. The right panel provides a detailed view of this intent, including a search bar, a list of sample utterances with the 'categoria' entity highlighted in green, and a 'Required entities' section where 'categoria' is marked as required. An 'Entity reprompt' dialog box is also visible, showing the prompt '¿de qué categoría?' and the 'Alexa' voice profile.

El NLU de Voiceflow: Slot Filling

Pulsa en la entidad, y añade un prompt en **Entity Reprompt**. Esta es la pregunta que haremos para recoger el valor de esa entidad si no llega.

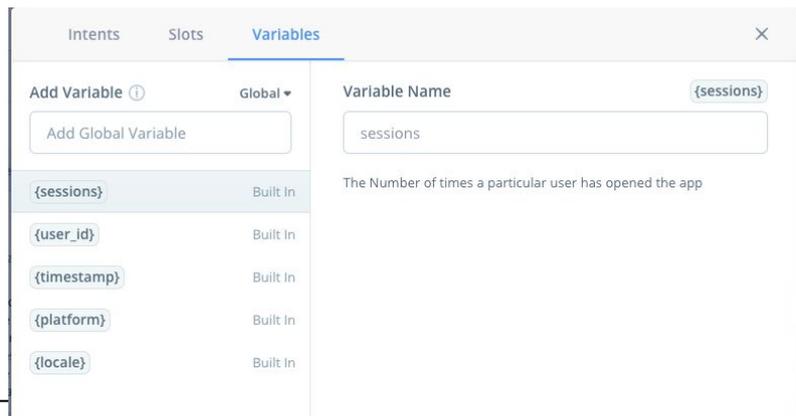
The screenshot displays the Voiceflow interface for configuring the 'dar_noticias_categoria' entity. On the left, a list of intents includes 'dar_noticias_categoria', which is selected. Below the list, the 'Entity reprompt' field is visible, containing the text '¿de qué categoría?'. On the right, the configuration for the 'dar_noticias_categoria' entity is shown, including a list of sample phrases and a 'Required entities' section with 'categoria' listed.

El NLU de Voiceflow: Variables

Desde Voiceflow, podemos utilizar lo recogido en las entidades a través de las [variables](#).

Hacemos uso de estas en las respuestas, utilizando `{}`.

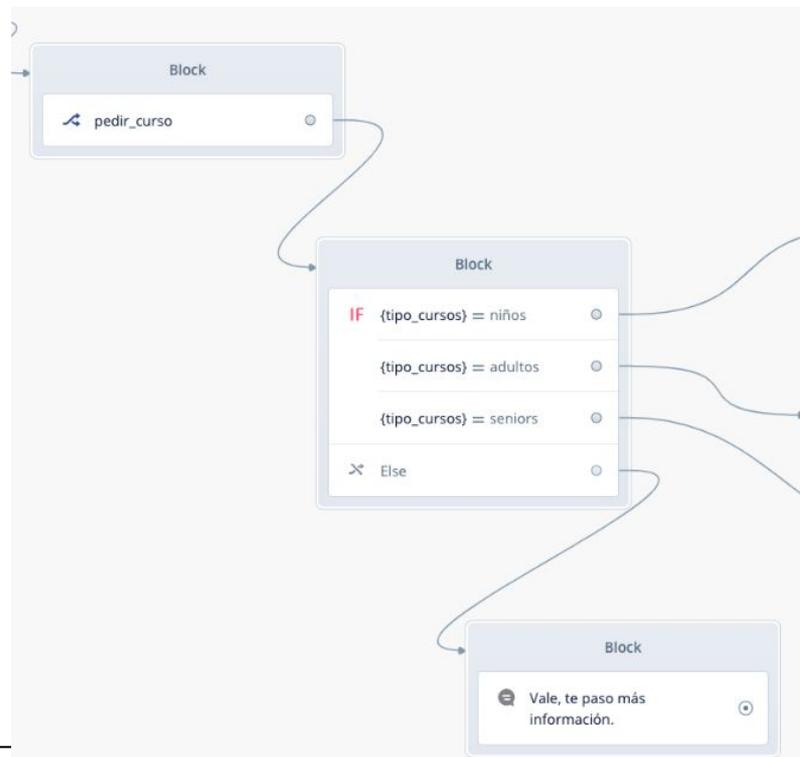
Además podemos crear otras nuevas, y utilizar variables globales (lo puedes ver en la vista [Model](#)).



Voiceflow: Bloque IF

Desde Voiceflow, utilizamos las variables para dirigir a caminos según los valores que toma la variable.

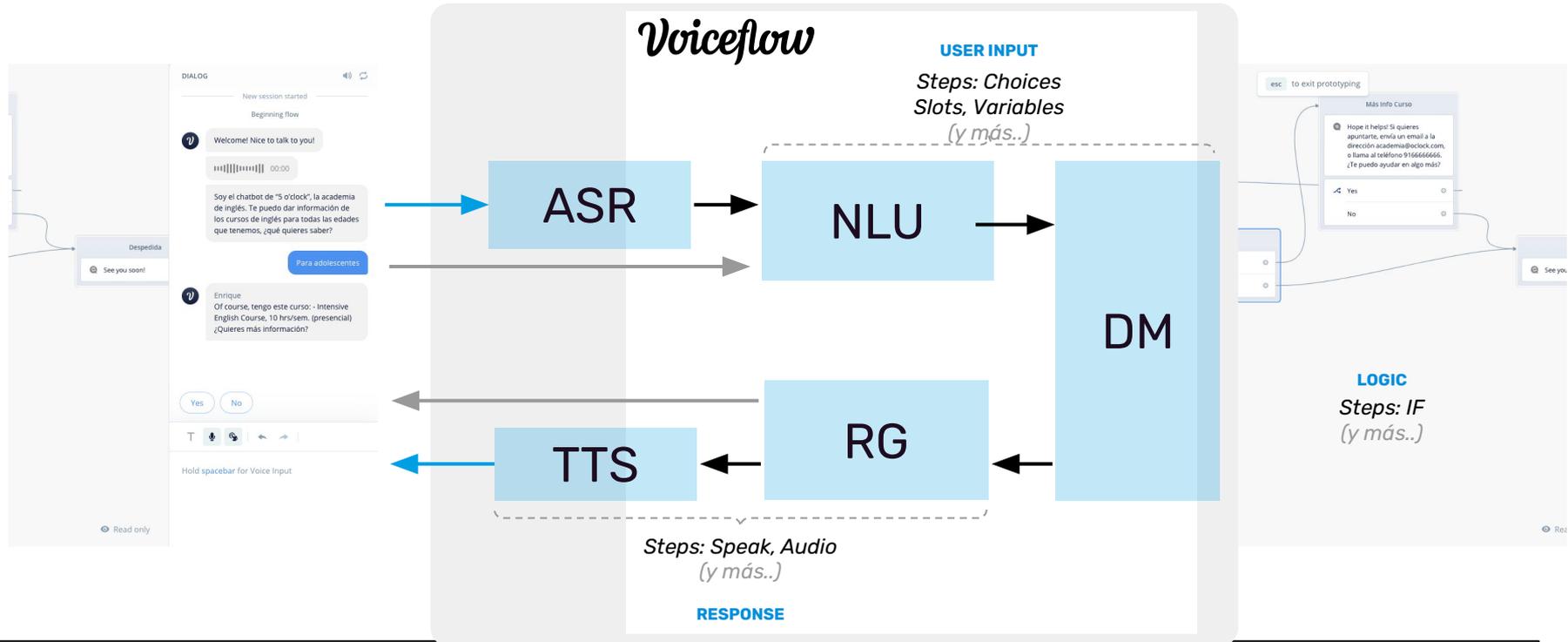
Utilizaremos el **bloque IF** para ello, evaluando las condiciones que creamos que utilizan estas variables.



Voiceflow: Integrations

The screenshot shows the Voiceflow interface for setting up an Amazon Alexa skill. On the left, a sidebar contains a 'v' logo, 'Integrations', and a list of options: 'Channel' (Amazon Alexa), 'Developer' (API), and 'Integration 3' (highlighted in a dark blue box). The main content area is titled 'Amazon Alexa' and features a light blue banner at the top stating 'This skill is linked on Amazon Developer Console' with a 'More Info' link. Below the banner, a progress indicator shows 'Basic Skill Info' as the current step. The 'Basic Skill Info' section includes a 'Display Name' field containing 'Academia de Inglés', two circular 'Large Icon' and 'Small Icon' placeholders (both with image icons), and a 'Next' button. To the right of the form, explanatory text reads: 'Display Name is what we display for your Skill on Voiceflow/Amazon' and 'Icons are what will be displayed for your Skill in the Amazon web store.'

Testeando en Web (**voz** y texto)



Probar en Alexa

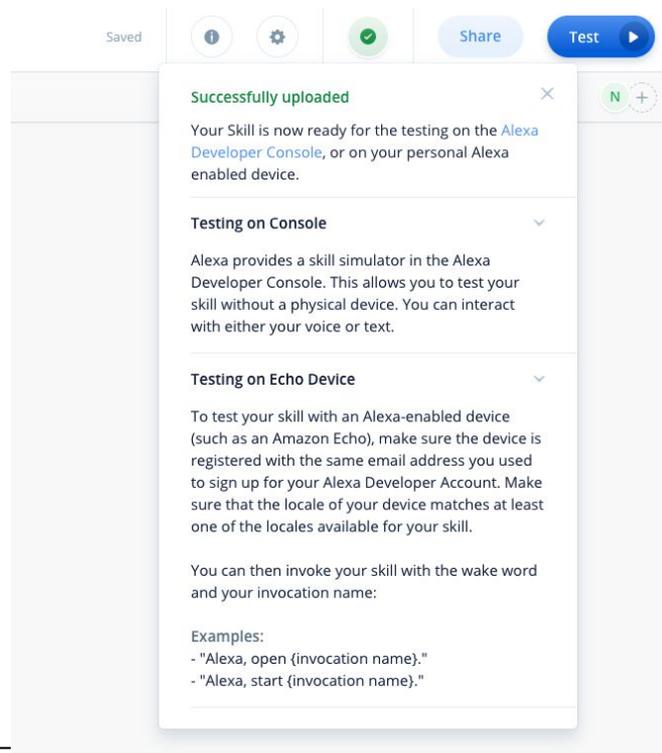
Voiceflow 101

Voiceflow: Conectar con Alexa

Necesitas tener una cuenta de Alexa Developer Console (ADC). Puedes crearla [aquí](#) e ir a la sección de [Alexa Developer](#).

Te permite sincronizar Voiceflow con ella, y podrás testear en:

- dispositivo Echo (o aplicación de móvil de Amazon Alexa).
- apartado Test en la ADC.



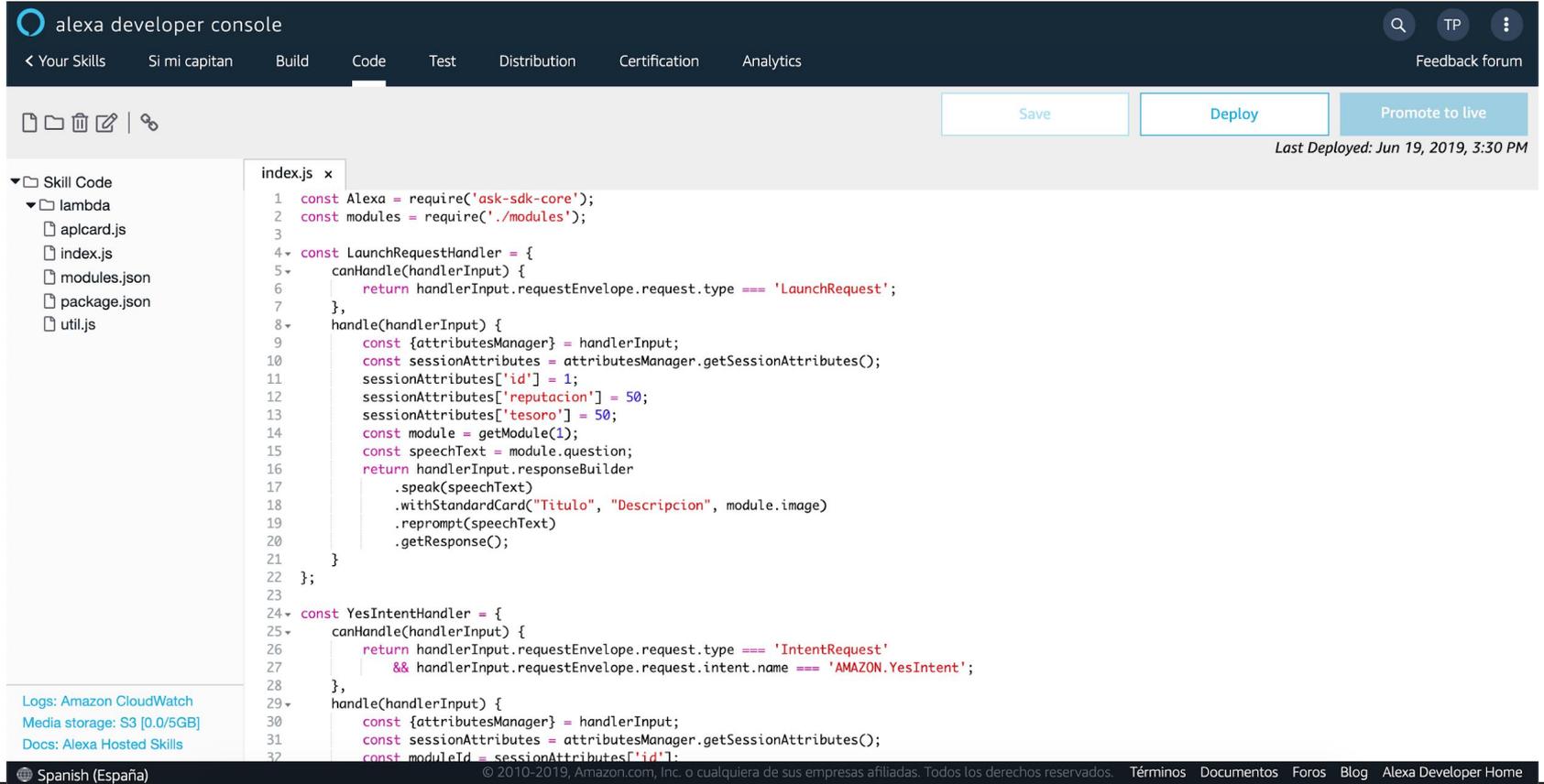
NLU de Alexa: **Alexa Developer Console**

The screenshot shows the Alexa Developer Console interface for a skill in Spanish (ES). The navigation bar includes 'Your Skills', 'Vevo', 'Build', 'Code', 'Test', 'Distribution', 'Certification', and 'Analytics'. The main content area is titled 'Intents / elegir' and displays 'Sample Utterances (20)'. A search box contains the text 'What might a user say to invoke this intent?'. Below the search box, five sample utterances are listed, each with the slot '{habitacion}' highlighted in orange:

- juguemos en el {habitacion}
- juguemos en la {habitacion}
- jugar en el {habitacion}
- jugar en la {habitacion}
- vamos a jugar en el {habitacion}

The left sidebar shows the 'Interaction Model' section with a tree view of intents. The 'elegir' intent is selected and highlighted in blue. Other intents visible include 'cantarVevoVevo', 'adivinar', 'cosa', 'darPista', and several 'Built-In Intents' like 'AMAZON.CancelIntent'.

DM de Alexa: **Alexa Developer Console (Node.js)**



The screenshot displays the Alexa Developer Console interface. At the top, the navigation bar includes "Your Skills", "Si mi capitan", "Build", "Code", "Test", "Distribution", "Certification", and "Analytics". The "Code" tab is active, showing a code editor for "index.js". The code defines two handlers: "LaunchRequestHandler" and "YesIntentHandler". The "LaunchRequestHandler" sets session attributes for 'id', 'reputacion', and 'tesoro', and returns a response with a standard card and a reprompt. The "YesIntentHandler" returns an intent name of 'AMAZON.YesIntent'. On the right side, there are buttons for "Save", "Deploy", and "Promote to live". Below the code editor, there are links for "Logs: Amazon CloudWatch", "Media storage: S3 [0.0/5GB]", and "Docs: Alexa Hosted Skills". At the bottom, there is a footer with "Spanish (Español)", copyright information, and links for "Términos", "Documentos", "Foros", "Blog", and "Alexa Developer Home".

```
index.js x
1  const Alexa = require('ask-sdk-core');
2  const modules = require('./modules');
3
4  const LaunchRequestHandler = {
5    canHandle(handlerInput) {
6      return handlerInput.requestEnvelope.request.type === 'LaunchRequest';
7    },
8    handle(handlerInput) {
9      const {attributesManager} = handlerInput;
10     const sessionAttributes = attributesManager.getSessionAttributes();
11     sessionAttributes['id'] = 1;
12     sessionAttributes['reputacion'] = 50;
13     sessionAttributes['tesoro'] = 50;
14     const module = getModule(1);
15     const speechText = module.question;
16     return handlerInput.responseBuilder
17       .speak(speechText)
18       .withStandardCard("Titulo", "Descripcion", module.image)
19       .reprompt(speechText)
20       .getResponse();
21   }
22 };
23
24 const YesIntentHandler = {
25   canHandle(handlerInput) {
26     return handlerInput.requestEnvelope.request.type === 'IntentRequest'
27       && handlerInput.requestEnvelope.request.intent.name === 'AMAZON.YesIntent';
28   },
29   handle(handlerInput) {
30     const {attributesManager} = handlerInput;
31     const sessionAttributes = attributesManager.getSessionAttributes();
32     const moduleId = sessionAttributes['id'];
```

Save Deploy Promote to live

Last Deployed: Jun 19, 2019, 3:30 PM

Skill Code
lambda
apcard.js
index.js
modules.json
package.json
util.js

Logs: Amazon CloudWatch
Media storage: S3 [0.0/5GB]
Docs: Alexa Hosted Skills

Spanish (Español) © 2010-2019, Amazon.com, Inc. o cualquiera de sus empresas afiliadas. Todos los derechos reservados. Términos Documentos Foros Blog Alexa Developer Home

TTS de Alexa: **Alexa Developer Console**

The screenshot shows the Alexa Developer Console interface. At the top, there's a navigation bar with "alex developer console" and a search icon. Below it, a secondary navigation bar contains "Your Skills", "Si mi capitan", "Build", "Code", "Test", "Distribution", "Certification", and "Analytics". The "Test" tab is active. Below this, there are settings for "Skill testing is enabled in:" (set to "Development"), "Skill I/O" (checked), "Device Display" (checked), and "Device Log" (unchecked). The main content area has three tabs: "Alexa Simulator", "Manual JSON", and "Voice & Tone" (highlighted). The "Voice & Tone" tab shows a simulator with a language dropdown set to "Spanish (ES)" and a microphone icon. A blue speech bubble contains the text: "buenas. Me alegro de volver a verte. El juego consiste en adivinar palabras de cosas que hay en cualquier habitación de mi casa virtual." Below the simulator, there's a "Skill I/O" section with "JSON Input" and "JSON Output" views. The JSON Input view shows a JSON object with fields like "version", "session", "application", "user", "context", and "device". The JSON Output view shows a JSON object with fields like "body", "version", "response", and "directives".

Alexa Developer Console: Distribution

The screenshot shows the Alexa Developer Console interface. At the top, there is a dark navigation bar with the 'alex developer console' logo and a breadcrumb trail: < Your Skills > Academia de Inglés > Build > Code > Test > Distribution > Certification > Analytics. The 'Distribution' tab is selected. On the left, a sidebar contains 'Skill Preview' (with a dropdown arrow), 'Spanish', 'Privacy & Compliance', and 'Availability' (highlighted in blue). The main content area is titled 'Availability' with a sub-header 'Put your skill into customers hands.' Below this, a section asks 'Who should have access to this skill? *' with two radio button options: 'Public' (selected) and 'Alexa for Business Organizations'. A 'Beta Test' section follows, with a sub-header 'Give access to skill beta testers by adding their email addresses. Once Beta Test is enabled, you can continue to edit your skill until you are ready to submit your skill for Amazon review.' A green success message states 'Your Beta Test is Active' and 'Academia de Inglés Beta Test is available to 1 tester. They will have access to this skill through Tuesday, May 11, 2021.' Below this message are 'Copy Invite URL' and 'End Test' buttons. A form for 'Beta Test Administrator Email Address' includes an input field and an 'Add' button. At the bottom, it says 'Manage Access to your Skill Beta Test (1 total tester, 0 active testers)'.

alex developer console

< Your Skills > Academia de Inglés > Build > Code > Test > Distribution > Certification > Analytics

Skill Preview ▼

Spanish

Privacy & Compliance

Availability

Availability

Put your skill into customers hands.

Who should have access to this skill? *

Public
Give access to anyone who enables your skill.

Alexa for Business Organizations
Give access to select businesses by adding organizations specific ARNs. This feature is only available in the United States.

Beta Test

Give access to skill beta testers by adding their email addresses. Once Beta Test is enabled, you can continue to edit your skill until you are ready to submit your skill for Amazon review.

✓ **Your Beta Test is Active**
Academia de Inglés Beta Test is available to 1 tester. They will have access to this skill through Tuesday, May 11, 2021. [Copy Invite URL](#) [End Test](#)

Beta Test Administrator Email Address

[Add](#)

Manage Access to your Skill Beta Test (1 total tester, 0 active testers)

Alexa Developer Console: Distribution

The screenshot displays the Alexa Developer Console interface for managing beta testers. The top navigation bar includes 'Your Skills', 'Academia de Inglés', 'Build', 'Code', 'Test', 'Distribution', 'Certification', and 'Analytics'. The left sidebar shows 'Skill Preview' (Spanish), 'Privacy & Compliance', and 'Availability' (highlighted). The main content area is titled 'Manage Access to your Skill Beta Test (1 total tester, 0 active testers)'. It features a table with columns for 'EMAIL ADDRESS', 'STATUS', and 'ACTIONS'. Below the table, there are two options for adding testers: 'Add beta testers' with a text input field for email addresses and an 'Add' button, and 'Or bulk upload a list of emails' with a dashed box for file upload and a 'Browse to CSV' button. A 'Download template' link is also present. At the bottom right, there is a 'Beta testing is enabled' status indicator.

alex developer console

< Your Skills Academia de Inglés Build Code Test Distribution Certification Analytics

Skill Preview

Spanish

Privacy & Compliance

Availability

Manage Access to your Skill Beta Test (1 total tester, 0 active testers)

<input type="checkbox"/>	EMAIL ADDRESS ▲	STATUS	ACTIONS
<input type="checkbox"/>			

▼ Add beta testers

Enter tester email addresses Invites remaining: 499

Delimit email addresses by semicolon or line by line

Add

Or bulk upload a list of emails Download template

Browse to CSV

<input type="checkbox"/>		Invite sent	Send reminder Remove tester
--------------------------	--	-------------	---

< 1 – 1 of 1 Testers > [View all](#)

Beta testing is enabled

NLU + DM

Voiceflow 102

El NLU de Voiceflow: Capture

Usuario: quiero información del curso para adultos

IC: vale, dime tu nombre

Usuario: Nieves ← CAPTURE: Slot type (AMAZON.Name)

IC: ok, Nieves, para niños tengo...

El NLU de Voiceflow: Slots

Usuario: quiero información del curso para adultos

IC: vale, dime tu nombre

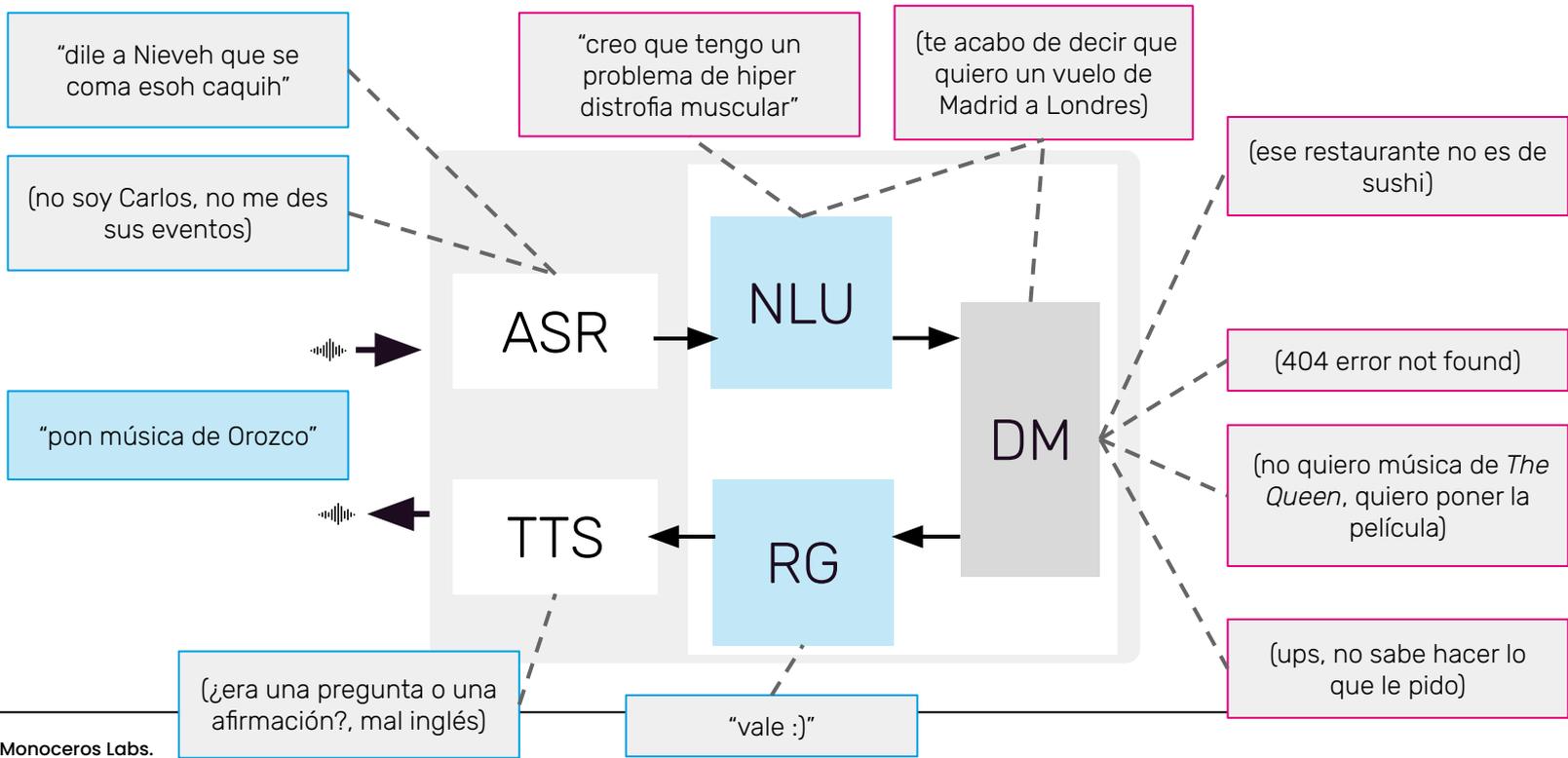
Usuario: me llamo Nieves ← CHOICE: Intent (utterance) + Slot type..

IC: ok, Nieves, para niños tengo...

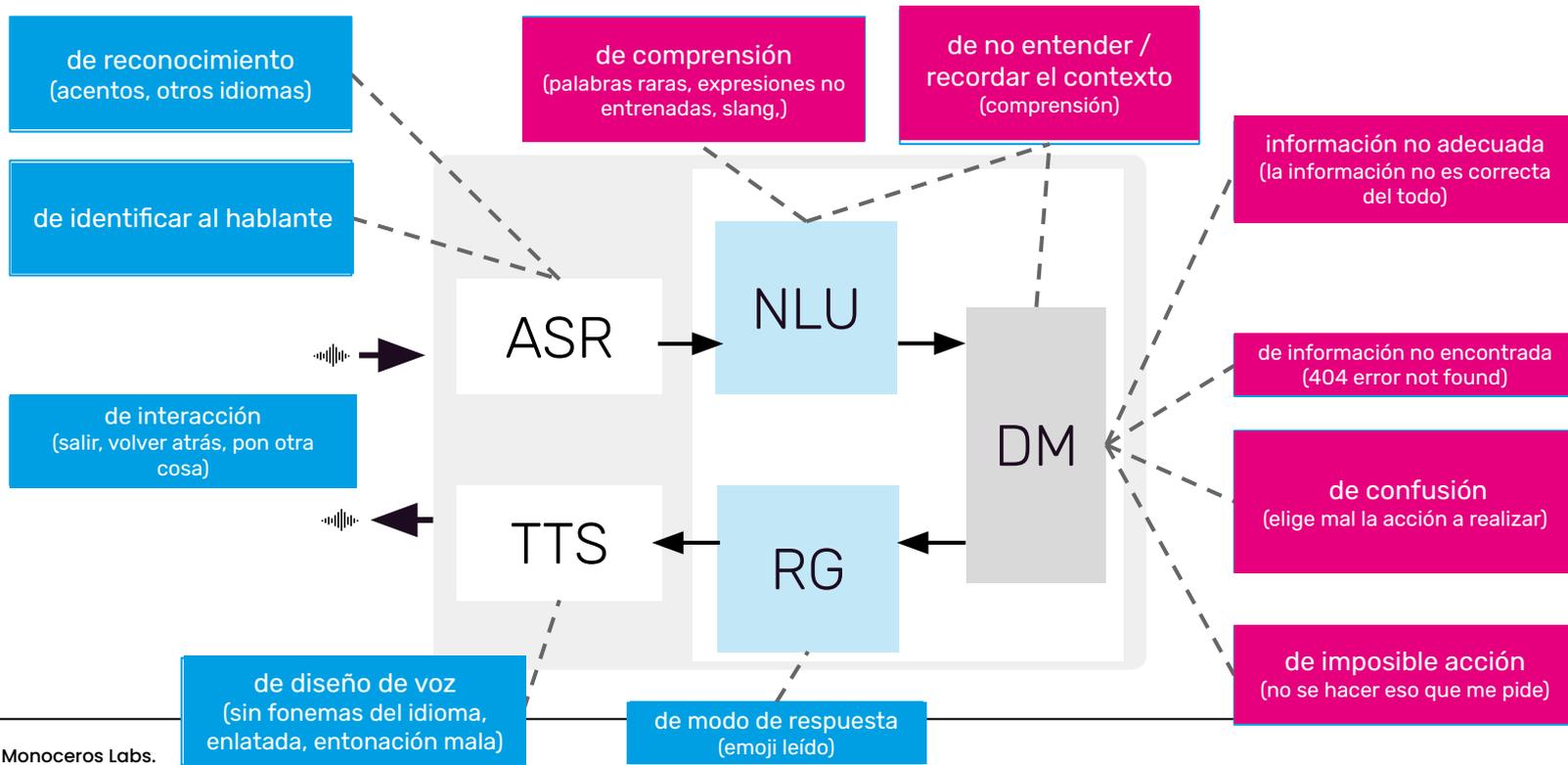
Problemas en la comunicación

Tecnologías del habla: limitaciones y desafíos

Errores en la conversación



Errores en la conversación



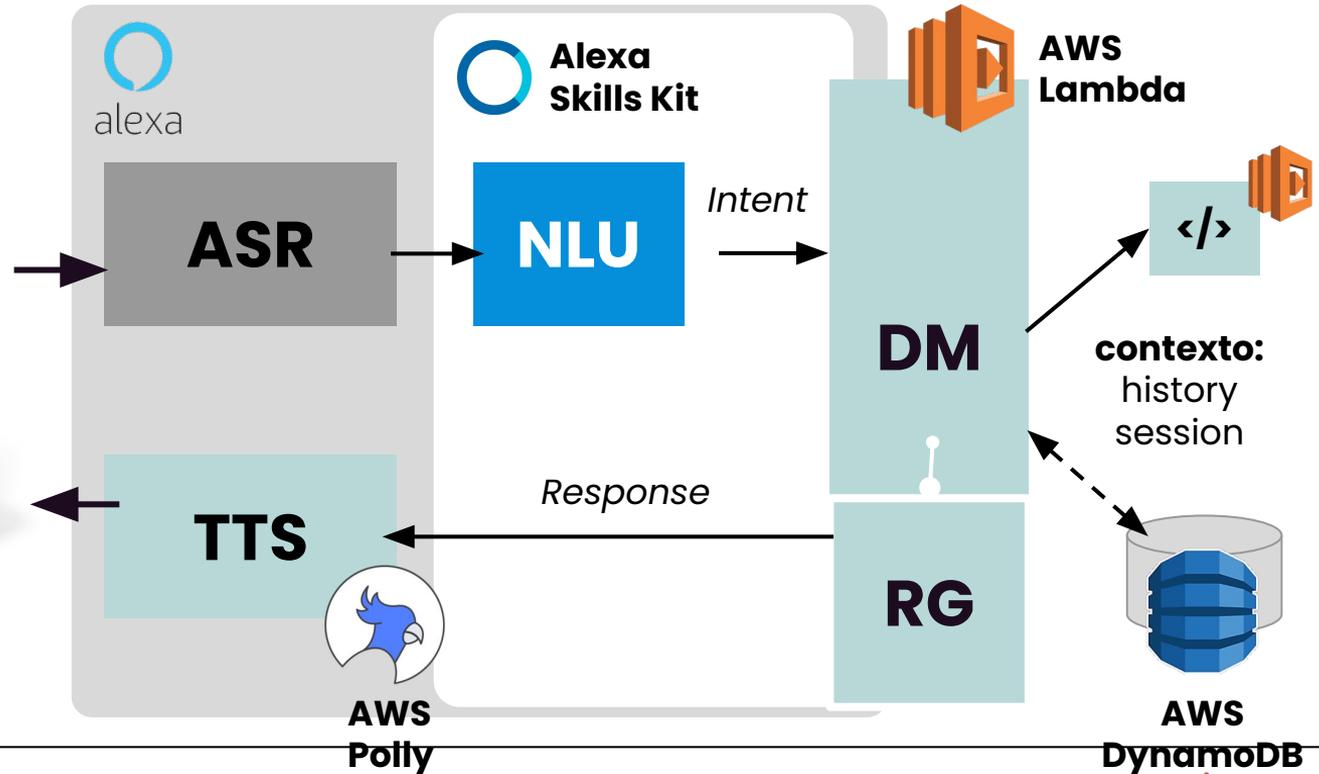
Tecnologías en la nube

Tecnologías del habla

Tecnologías de Amazon Web Services



Digo:
"Alexa,
abre
Pasapalabra"

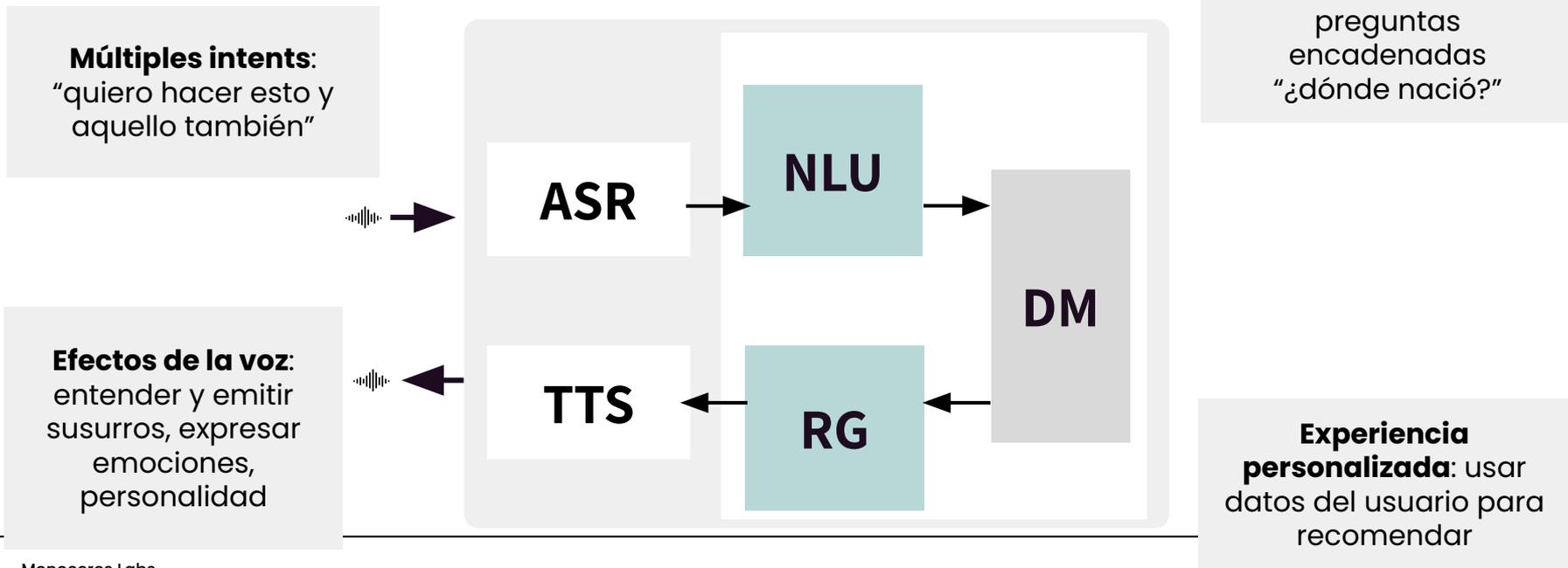


¿Y mañana?

Tecnologías del habla

¿Tecnología en fase temprana?

Advanced features (*state-of-the-art*, disponibles en algunos asistentes, o en inglés) que van orientadas a buscar la naturalidad en la conversación.



¿Tecnología en fase temprana?

Advanced features (*state-of-the-art*, disponibles en algunos asistentes, o en inglés) que van orientadas a buscar la naturalidad en la conversación.

The screenshot shows the Alexa developer documentation website. The header includes the Alexa logo and 'developer documentation'. Below the header, there are navigation links for 'Alexa Skills Kit', 'Alexa Voice Service', and 'Alexa Connect Kit'. A breadcrumb trail indicates the current page is 'About Alexa Conversations'. The left sidebar contains a table of contents for the 'Alexa Conversations Developer Guide', with 'About Alexa Conversations' selected. The main content area features a heading 'About Alexa Conversations' and a note box listing supported locales: (GA) en-US and (Beta) en-AU, en-CA, en-IN, en-GB, de-DE, ja-JP, es-ES, es-US. Below the note, the text describes Alexa Conversations as a deep learning-based approach to dialog management that enables natural, human-like voice experiences on Alexa. It explains that the service helps skills respond to a wide range of phrases and unexpected conversational flows, and provides conversational memory for long, two-way interactions. The text concludes by stating that developers can create a skill that uses Alexa Conversations to manage the entire skill experience, or extend an existing skill to use Alexa Conversations for more complex tasks.

alex developer documentation

Alexa Skills Kit Alexa Voice Service Alexa Connect Kit

Alexa > Documentation > Alexa Skills Kit > Alexa Conversations > About Alexa Conversations

Alexa Conversations Developer Guide

Collapse All | Expand All

- About Alexa Conversations
- How Alexa Conversations Works
- What's New in Alexa Conversations
- Guided Development Experience
- Command-Line Experience (Beta)

About Alexa Conversations

Note: Alexa Conversations currently supports the following locales:

- (GA) en-US
- (Beta) en-AU, en-CA, en-IN, en-GB, de-DE, ja-JP, es-ES, es-US

Alexa Conversations is a deep learning-based approach to dialog management that enables you to create natural, human-like voice experiences on Alexa. Alexa Conversations helps skills respond to a wide range of phrases and unexpected conversational flows, and gives skills the conversational memory to sustain long, two-way interactions between Alexa and the user.

You can create a skill that uses Alexa Conversations to manage the entire skill experience, or you can extend an existing skill with Alexa Conversations. For example, your skill can use your existing code to handle simple interactions. Then, your skill can delegate dialog management to Alexa Conversations for tasks that involve many two-way conversations with the user.

¿Tecnología en fase temprana?

Advanced features (*state-of-the-art*, disponibles en algunos asistentes, o en inglés) que van orientadas a buscar la naturalidad en la conversación.

Use New **Alexa Emotions** and Speaking Styles to Create a More Natural and Intuitive Voice Experience

Catherine Gao Nov 26, 2019

Share: [f](#) [in](#) [t](#)

Game Skills Content Skills Design News



We're excited to introduce two new Alexa capabilities that will help create a more natural and intuitive voice experience for your customers. Starting today, you can enable Alexa to respond with either a happy/excited or a disappointed/empathetic tone in the US. Emotional responses are particularly relevant to skills in the gaming and sports categories. Additionally, you can have Alexa respond in a speaking style that is more suited for a specific type of content, starting with news and music. Speaking styles are curated text-to-speech voices designed to create a more delightful customer experience for

¿Tecnología en fase temprana?

Advanced features (*state-of-the-art*, disponibles en algunos asistentes, o en inglés) que van orientadas a buscar la naturalidad en la conversación.

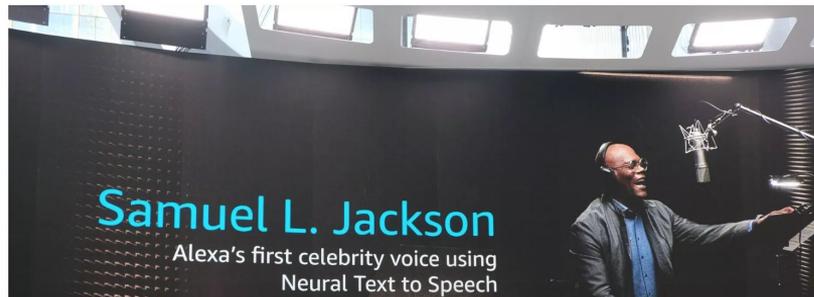
TECH \ AMAZON \ AMAZON ALEXA \

All the new features coming to Alexa, including a new voice, frustration mode, and Samuel L. Jackson

Also new: bilingual support, and more troubleshooting features like 'frustration detection'

By [Chaim Gartenberg](#) | [@cgartenberg](#) | Sep 25, 2019, 1:26pm EDT

f   SHARE



¿Tecnología en fase temprana?

Advanced features (*state-of-the-art*, disponibles en algunos asistentes, o en inglés) que van orientadas a buscar la naturalidad en la conversación.

Amazon's neural TTS can model speaking styles with only a few hours of recordings

KYLE WIGGERS @KYLE_L_WIGGERS NOVEMBER 19, 2018 6:55 AM



Above: Amazon Echo Sub

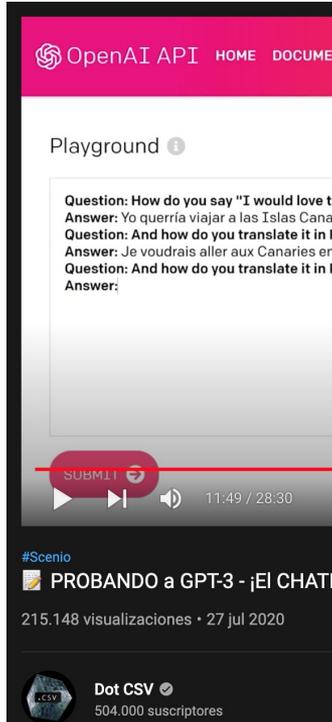


¿Tecnología en fase temprana?

Advanced features (*state-of-the-art*, disponibles en algunos asistentes, o en inglés) que van orientadas a buscar la naturalidad en la conversación.



Y lo que vendrá (... y sus retos)



The screenshot shows the OpenAI API Playground interface. At the top, there is a pink navigation bar with the OpenAI logo and the text "OpenAI API HOME DOCUME". Below this, the word "Playground" is displayed with a help icon. The main area contains a chat log with the following text:

Question: How do you say "I would love to
Answer: Yo querría viajar a las Islas Cana
Question: And how do you translate it in F
Answer: Je voudrais aller aux Canaries en
Question: And how do you translate it in F
Answer:

Below the chat log is a red "SUBMIT" button with a refresh icon. At the bottom of the chat area, there are playback controls (play, pause, volume) and a progress indicator showing "11:49 / 28:30".

Below the chat area, there is a social media-style post with the following details:

- Hashtag: #Scenio
- Title: PROBANDO a GPT-3 - ¡EL CHATE
- Views: 215.148 visualizaciones
- Date: 27 jul 2020
- Profile: Dot CSV (verified), 504.000 suscriptores

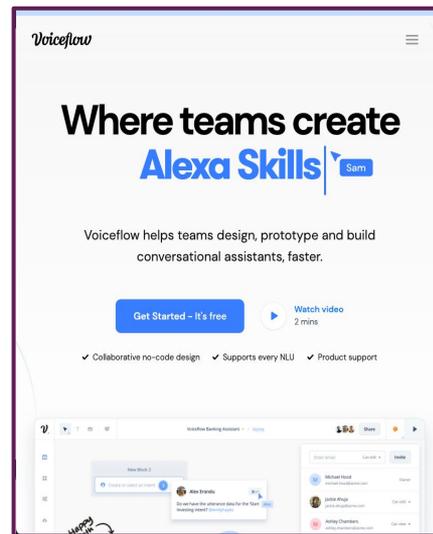
Cursos y libros



<https://plataforma.keepcoding.io/p/curso-desarrollo-alexa-skills>



<https://monoceros.xyz/blog/libros-interfaces-voz/>



<https://www.voiceflow.com/>



Es hora de crear conversaciones

Síguenos

[Twitter](#) / [LinkedIn](#)

Léenos

www.monoceros.xyz

[Blog](#)

¿Hablamos?

hello@monoceros.xyz