## 

## UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Facultad de ciencias

# Reporte de Skill: Can Clues

Diseño e Implementación de Interfaces de Usuario Basadas en Voz

Semestre 2023-1

Introducción

La tecnología nos facilita la vida y nos da soluciones a nuestros problemas, por ello para tener una mejor interacción computadora-usuario es necesario el uso de una interfaz que permita su uso de manera eficiente y fácil, sin embargo, hay que notar que hay diferentes tipos de interfaces, desde aquella mostrada en una consola sólo con texto, hasta una interactiva que se muestra de manera gráfica, éstas son de las más utilizadas, por ejemplo, en dispositivos móviles, aunque en estos últimos años se ha popularizado la interacción mediante la interfaz de voz con el uso de comandos, esto debido a la existencia de dispositivos inteligentes que se conectan por WiFi, como focos, tiras de luz, cámaras, enchufes, cerraduras, etc. Antes de que existieran estas interfaces por voz, había que controlarlos con un dispositivo móvil que tuviera la aplicación instalada, ante la necesidad de hacer este proceso más fácil y rápido se crearon los dispositivos a los que sólo les indicas por voz lo que tiene que hacer, sin necesidad de tomar el dispositivo y acceder a la aplicación, que nos permite tener las manos libres para hacer otras cosas, también su alta de popularidad se debe a lo fácil y concreto que vuelve al acceso a información, por otro lado, también pueden proporcionar entretenimiento, música, noticias, e incluso que nos ayudan a organizar nuestra vida con tan sólo pedírselo.

Son importantes las interfaces por voz dado que nos trae el beneficio de facilidad, eficiencia y hacer otras actividades sin tener que suspenderlas para enfocarnos en otra, pero sobre todo nos ayuda a una interacción más natural, casi como si hablaras con otra persona, por lo que el usuario se siente en confianza y se desenvuelve mejor sin importar la edad.

Para entender mejor de lo que hablamos, se tocarán algunos términos así como la explicación de la skill desarrollada durante el curso.

* ¿Qué es una interfaz de usuario inteligente?

Las interfaces de usuario inteligentes son interfaces cuyo propósito es mejorar el rendimiento, efectividad, naturalidad y, en general, la usabilidad de las interacciones entre el humano y la computadora, ya sea representando, entendiendo o ejecutando cierta colección de modelos.

Para poder cumplir su principal objetivo las interfaces de usuario inteligentes deben analizar la información que tienen sobre el usuario, saber las actividades que están permitidas, tener la capacidad de comunicarse con el usuario entendiendo correctamente los comandos que éste pida y realizando la acción que corresponda basada en la información recolectada.

* Diseño centrado en el usuario

Es definido por la [Usability Professionals Association](https://www.upassoc.org/) (UPA) como un enfoque de diseño cuyo proceso está dirigido por información sobre las personas que van a hacer uso del producto.

Más puntualmente debemos tener claro a qué público/usuario va dirigido nuestra skill, una vez decidido esto entonces recabaremos la información del cliente como sus necesidades, intereses, metas y limitaciones con ello comenzaremos en el desarrollo de la skill en la cual haremos pruebas y mejoraremos los errores que se vayan obteniendo.

Para el desarrollo es necesario tomar en cuenta lo aprendido en teoría y laboratorio, los términos se verán a continuación.

* Skills de Alexa

Alexa es el asistente virtual controlado por voz creado por Amazon, lanzado [en noviembre de 2014](https://www.xatakahome.com/altavoces/amazon-echo-el-altavoz-portatil-que-controlaras-con-tu-voz) junto a su línea de altavoces inteligentes Echo. Todos los dispositivos Alexa vienen con un conjunto de habilidades básicas preestablecidas, pero tenemos la opción de añadir otras nuevas o aumentar la funcionalidades de las existentes, mejorando así el dispositivo. A estas aplicaciones se les llama Skills y han ayudado a cubrir las necesidades de miles de usuarios, las cuales pueden ir desde ayuda en la cocina hasta simple entretenimiento.

El propósito de este reporte es presentar el desarrollo que tuvo una nueva skill creada por nosotras usando las herramientas que Amazon nos ofrece.

Definición del Problema de la Skill

Nuestra skill, Can Clues, busca proporcionar entretenimiento al usuario mediante una historia interactiva donde se tienen que tomar ciertas decisiones para llegar a diferentes finales.

En nuestro caso, Alexa toma el papel de storyteller pues ella narra lo que está pasando en la historia en tiempo real. Además, funciona como intermediaria pues a través de ella es que se reciben las decisiones que toma el usuario.

Debido a que en cada punto se dan dos opciones que bifurcan la historia, tuvimos que hacer uso de un árbol para determinar a dónde te llevaba cada ruta. Aunque no hay un camino predeterminado que siga el usuario, todas sus decisiones seguirán perteneciendo al mismo conjunto de opciones.

Es importante brindar este tipo de entretenimiento (una historia interactiva) en una skill de Alexa porque el usuario tiende a usar la vista constantemente por lo que puede terminar cansándose de mirar una pantalla, por ejemplo. Por lo que es conveniente tener una forma de divertirse que sólo sea auditiva, además de que puede ser más inmersiva pues cuenta con diferentes efectos de sonido.

Justificación

Nosotras tratamos de crear una historia más larga y elaborada a comparación de las que normalmente se ven en aplicaciones similares. Buscamos cubrir esa parte del público que gusta más de escuchar y realmente meterse en la historia, tipo podcast. También permitimos que el usuario pueda parar en cualquier momento para acomodarnos a los horarios complicados y así reanudar el juego en el momento que quiera dando el contexto de donde se quedó para que pueda recordar y no sólo elegir sin saber cuál es el camino por donde quiere ir.

Se puede jugar varias veces sin llegar a aburrirse por dos principales razones. La primera es que al principio se puede elegir tres objetos de un conjunto, esto nos da la posibilidad de acceder a ciertas pistas si traes el objeto necesario, cabe destacar que no hay ningún objeto malo, todos te permiten descubrir algo, así que por cada partida puedes elegir diferentes objetos y explorar otros caminos. La segunda razón es que se puede llegar a diferentes finales, por diferentes caminos, ya que la historia está construida como si fuera un árbol donde hay más de una hoja final.

Por otro lado, Can Clues ayuda al usuario a concentrarse al tener que poner mucha atención a la historia, así como retención de información para acceder a ciertos candados. Creemos que el poner atención y retener información es algo que nos ayuda a la vida diaria, ya sea escolar o de trabajo e incluso personal.

Diseño de la Solución

* Análisis de Usuario

Para tener un análisis más puntual tomaremos un perfil de usuario con sus necesidades.

**Nombre:** Bryan González

**Descripción:** Chico de 23 años que disfruta de resolviendo problemas y acertijos

**Comportamiento:**

* Se levanta temprano para el home office y atender clases virtuales
* Pasa mucho tiempo en casa
* No le gusta mucho socializar
* No disfruta de la lectura, pero sí de la música y podcasts.

**Motivaciones:**

* Ser un buen empleado y estudiante
* Invertir más tiempo para sí mismo
* Buscar entretenimiento lo desafíe mentalmente
* Entretenimiento que pueda retomar en sus tiempos libres

**Necesidades:**

* Distraerse de estar siempre frente a una pantalla
* Se aburre de las rutinas por lo que busca algo que active su mente
* Aprender cómo sus acciones afectan el desenlace de una situación
* Buscar diferentes formas de entretenimiento

**Frustraciones:** Los juegos predecibles.

Así que generalizando podríamos decir que va dirigido a un usuario que le guste escuchar historias por un rato largo, descubrir pistas analizando la mejor decisión y que esté cansado de estar usando un dispositivo con interfaz gráfica.

* ¿Cómo ayuda la skill al usuario a lograr sus objetivos?

Proporciona al usuario un tiempo de distracción y relajación cuando tenga un tiempo libre, si el usuario busca tener un tiempo para sí mismo puede invertir este descanso en una historia entretenida tipo podcast con la diferencia de que él decide el rumbo que desea tomar.

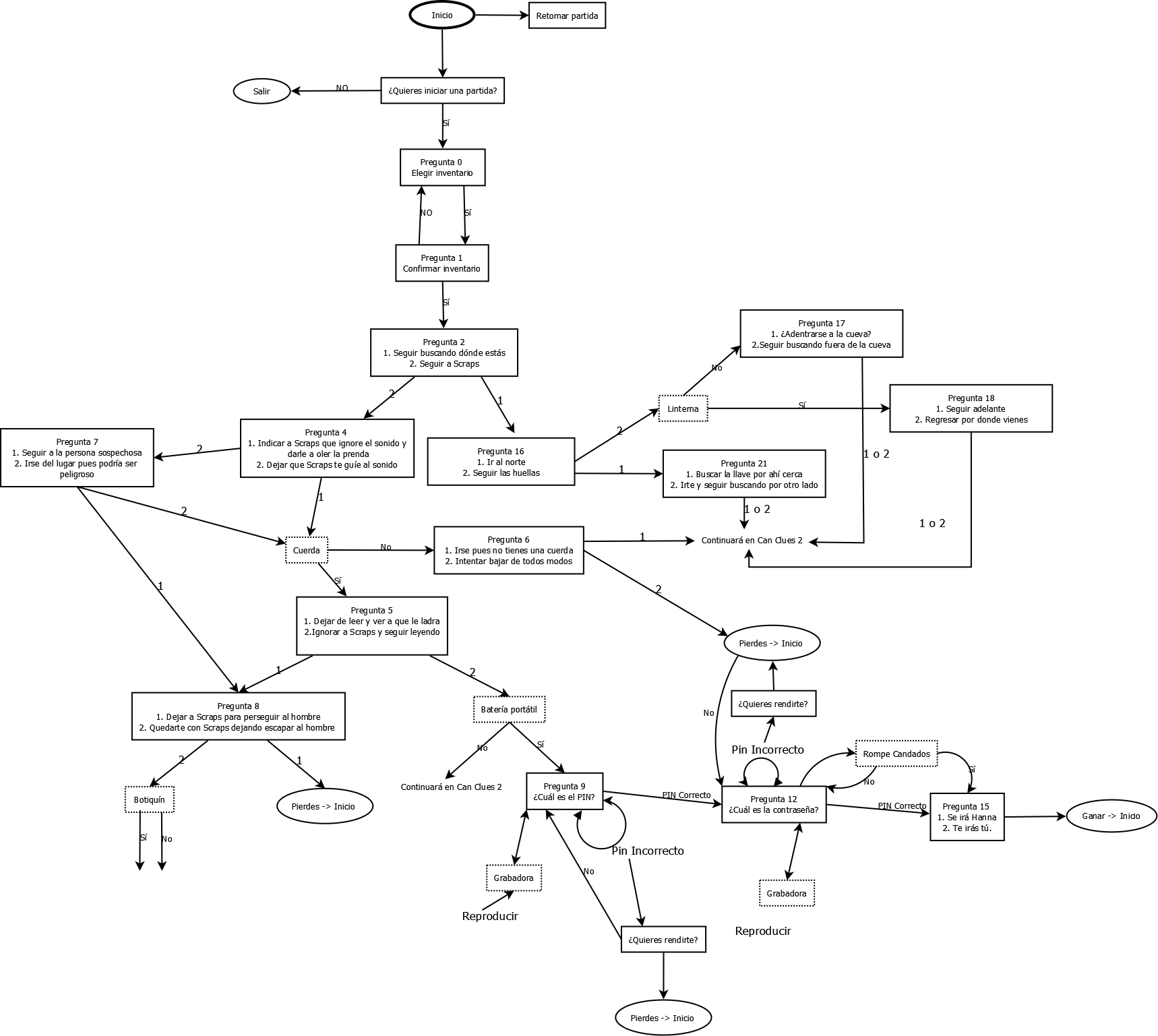
La skill, como ya hemos mencionado, ayuda a la práctica de retención de información así que mantiene su mente activa y lo ayuda a aprender acerca de las decisiones que toma ya que pueden traer consecuencias.

* Descripción del funcionamiento de la skill

Primero explicaremos cada intent que tiene la aplicación pues son el núcleo de su funcionamiento:

* AMAZON.YesIntent: Tomamos este intent predeterminado de Amazon que sirve para cachar todas las veces que el usuario diga una forma de afirmación, por ejemplo: sí, continuar, aceptar, etc. Para saber a qué nos está respondiendo le dimos un id a cada pregunta, entonces antes de llevar a cabo una acción, primero revisa qué número de pregunta es para poder actuar en respuesta.
* ObjetoIntent: Se encarga de ver cuáles objetos ha elegido el usuario para llevar en su inventario. Tienen que ser tres y en caso de no recibirlos le pedirá al usuario que le dé los faltantes.
* [UnoDosIntent](https://developer.amazon.com/alexa/console/ask/build/custom/amzn1.ask.skill.264d8a2a-cfad-4fc4-bb05-9fbc4216ef1b/development/es_MX/intents/UnoDosIntent): Este intent sirve para obtener la opción que elige el usuario a cada pregunta que bifurca la historia. Usa el mismo sistema de ID de preguntas que el [AMAZON.YesIntent](https://developer.amazon.com/alexa/console/ask/build/custom/amzn1.ask.skill.264d8a2a-cfad-4fc4-bb05-9fbc4216ef1b/development/es_MX/intents/AMAZON.YesIntent).
* AMAZON.NoIntent: Tomamos este intent predeterminado de Amazon que sirve para cachar todas las veces que el usuario diga una forma de negación, por ejemplo: no, no quiero, etc. Usa el mismo sistema de ID de preguntas que el [AMAZON.YesIntent](https://developer.amazon.com/alexa/console/ask/build/custom/amzn1.ask.skill.264d8a2a-cfad-4fc4-bb05-9fbc4216ef1b/development/es_MX/intents/AMAZON.YesIntent).
* CandadoIntent: Este intent se activa cuando el usuario se encuentra con uno de los candados del juego. Se asegura de que se le hayan dado 4 números, si no es así, pide al usuario que se los dé. Dependiendo del número de intentos que se han hecho para tratar de abrir el candado, dará diferentes pistas o la opción de rendirse.
* ReproducirIntent: Cuando el usuario diga reproduce, reproducir, etc., se encargará de reproducir lo que se haya guardado en la grabadora (lo cuál se hace automáticamente dependiendo de qué pasos haya tomado). Sólo se activa si tiene la grabadora en el inventario.
* PuertaIntent: Igual a CandadoIntent, pero recibe 6 letras en vez de 4 números.
* RepiteIntent: Como dice su nombre, repite lo último que haya dicho Alexa.
* AMAZON.HelpIntent: Le repite al usuario las reglas del juego y que puede usar el comando “repite” como sea necesario.
* AMAZON.CancelIntent o AMAZON.StopIntent: Cuando el usuario pide a Alexa que pare o salga de la aplicación, en automático se guarda el estado de la partida para poder continuar con ella después si se desea.

Para mostrar la interacción del usuario con la skill, nos basamos en un diagrama de flujo que tiene preguntas que hará Alexa y mediante la respuesta del usuario se seguirá por una de las ramas hasta finalizar la conversación.

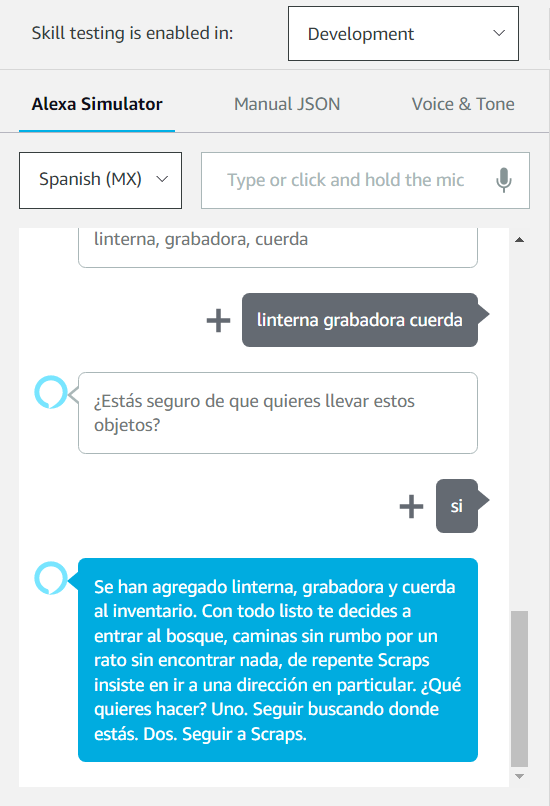


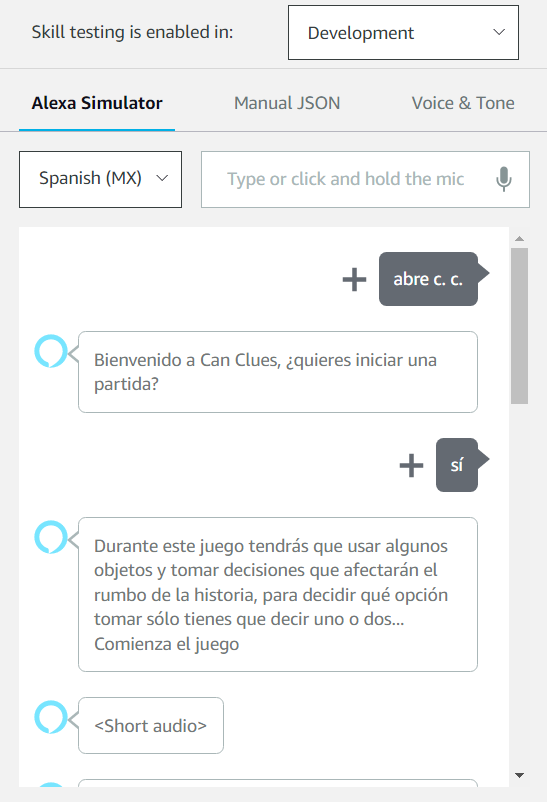
Daremos a conocer los distintos casos de usos dependiendo de las peticiones del usuario.

| Solicitud | Respuesta |
| --- | --- |
| Abre c.c. | Bienvenido a Can Clues ¿Quieres iniciar una partida? |
| Abre c.c | Bienvenido a Can Clues. Hay una partida guardada, ¿quieres continuar? |
| 1 | Depende de la parte de la historia en la que estés, comenzará a decirle el contexto. |
| 2 | Depende de la parte de la historia en la que estés, comenzará a decirle el contexto. |
| {objeto1} {objeto2} {objeto3} | ¿Estás seguro de que quieres llevar estos objetos? |
| Ayuda | Para elegir tu respuesta sólo basta con que digas 1 o 2. Si quieres volver a oír lo último que se dijo di repite. |
| Repite | Te dice la última parte de la historia que te contó. |
| Reproducir | Reproduce lo que se haya guardado en la grabadora de tener una. |
| Para | Se ha guardado la partida en este momento. Adiós. |
| Cierra la skill | Se ha guardado la partida en este momento. Adiós. |

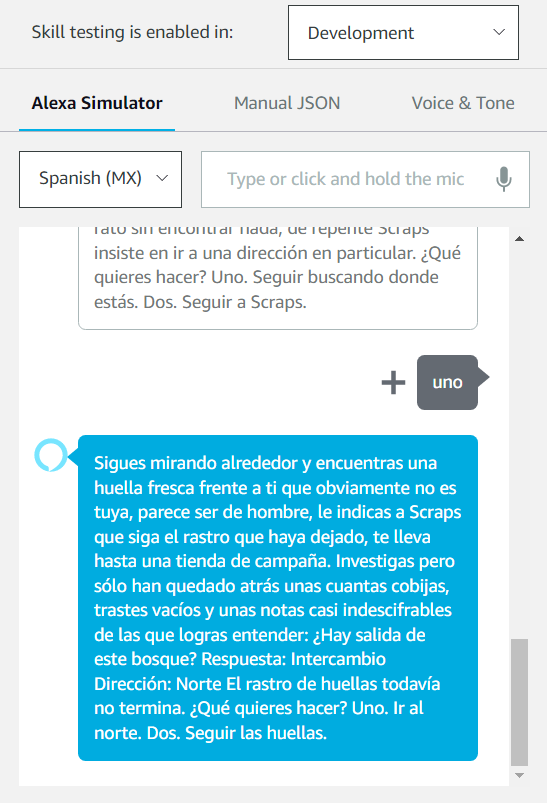
Se pondrá un ejemplo de ejecución con las peticiones del usuario.

El usuario inicia la petición “*abre c.c”.* Inicia la skill dando la bienvenida, si hay una partida guardada lo notificará y dará el contexto de donde se quedó el usuario; si no hay una partida guardada entonces explicará las reglas y comenzará el juego.

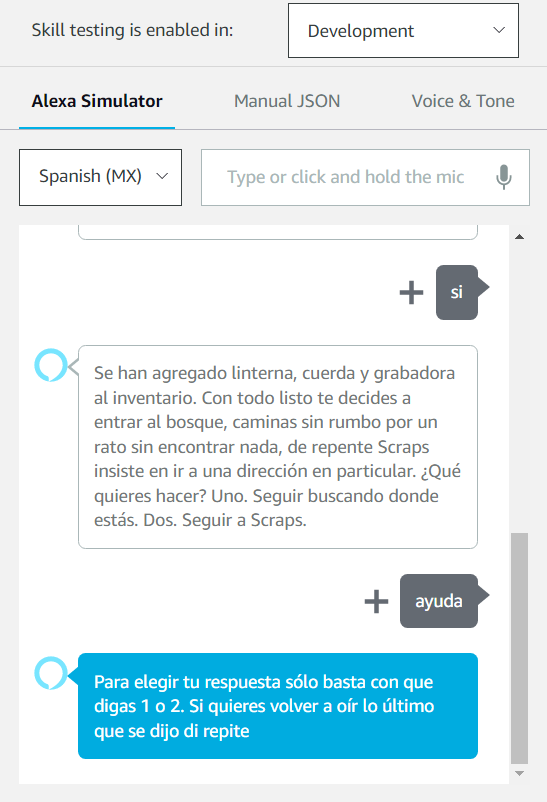




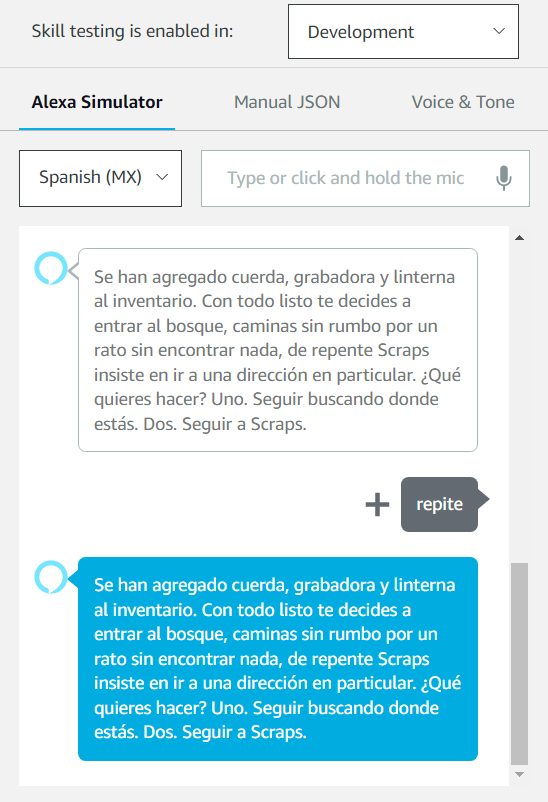
Alexa dice un conjunto de objetos de los cuales permite elegir 3, el usuario indica la petición diciendo “*{objeto1} {objeto2} {objeto 3}”* por ejemplo “*linterna grabadora cuerda”* y una vez que el usuario está seguro le notifica que los guardará en el inventario.



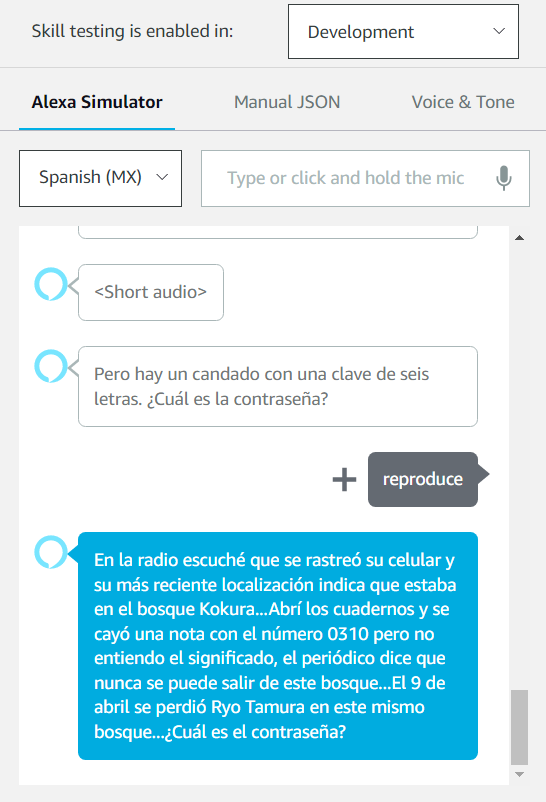
Cuando el usuario elige la opción 1 como petición diciendo “*uno”*, Alexa narrará lo que sigue en la rama del camino y al finalizar nuevamente hará una pregunta. Análogamente funciona para la opción 2.



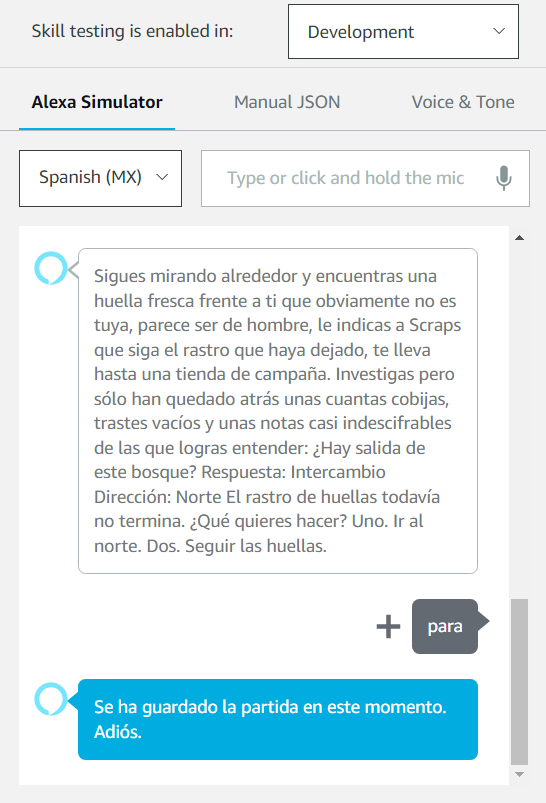
Cuando el usuario hace la petición de *“ayuda”* , la skill le recordará las instrucciones y como pedir que se repita lo último.



Para la petición de *“repite” ,* la skill te recuerda lo último que te mencionó ya que nuestros diálogos están separados por bloques o nodos.



Cuando el usuario tiene en su inventario la grabadora entonces podrá hacer la petición de *“reproduce”* tal que la skill le dirá todos los datos recabados hasta el momento, de esta manera podría descubrir más fácil las claves.



Ante la petición de *“para”*  la skill notificará que se guardó la partida y dirá una despedida.

* Manejo de confirmaciones

El manejo de confirmaciones es importante porque de esta manera le notificamos al usuario que la petición que requiere fue correctamente recibida.

Para el desarrollo de nuestra skill hicimos uso de las confirmaciones implícitas, que son aquellas que se le da a entender al usuario que se le ha entendido dentro de la respuesta. Por ejemplo, cuando el usuario selecciona sus tres objetos y está seguro de que son los objetos que quiere entonces Alexa procede a hacer una confirmación diciendo *“Se han agregado {objeto1}, {objeto2} y {objeto3} al inventario, Con todo listo te decides a entrar al bosque…”* como se puede notar toma lo que se le acaba de decir y lo hace parte de la conversación.

Por otro lado manejamos la confirmación sin diálogo que se encarga de completar una acción que no requiere una respuesta hablada. Es decir, si hay una demora por parte de Alexa para completar una acción entonces regresa un tipo de confirmación como *“Ok”* o *“Entendido”* para que el usuario sepa que está en proceso su petición. Por ejemplo, si el usuario dicta una clave incorrecta para abrir un candado, Alexa reproduce un audio donde se escucha que está girando el candado y después prosigue a decirte que la clave es incorrecta para luego dar una pista, entonces podemos decir que es este tipo de confirmación pues no repite la clave que se le dió.

* Marcadores conversacionales.

Conforme pasa el tiempo buscamos que la inteligencia artificial pueda ser más apegado a un lenguaje humano, es decir, una conversación más natural, esto lo logramos con el uso de marcadores ya que dirá *“frases”*  para mejorar la experiencia, por ejemplo: *“primero, finalmente, gracias, lo siento, buen trabajo, qué mal, etc”*

Una de las partes donde usamos estos marcadores es al momento de revisar el inventario del usuario. Si trae, por ejemplo, una cuerda y ésta le ayuda a descubrir ciertas pistas entonces la skill dice: *“pero… ¡Genial!, traes una cuerda en tu inventario y la usas para bajar…”*.O en otro caso donde se revisa si trae la batería y si no es así, dice: *“¡Lástima! No traes la batería portátil, deberás seguir buscando”*.

Otro ejemplo de marcador es cuando pide que le digas un objeto ya que no te entendió anteriormente o no le diste 3 objetos diferentes, es decir, *“Finalmente, dime el tercer objeto”*.

Por último, cuando pierdes la partida hace una conversación más natural donde no sólo te dice *“Pierdes”*, sino que hace énfasis con un tono diferente diciendo *“Cuando ya no puedes más te dejas ir y… ¡Has perdido!”*.

* Manejo de errores

El manejo de errores tiene gran importancia en cualquier tipo de proyecto y este no es la excepción debido a que mejora la experiencia de usuario y lo vuelve funcional. Un error recurrente al que le dimos un manejo adecuado es Recognized but Not Handled donde el usuario dice una palabra que entiende Alexa, pero no hay un intent que lo cache, en nuestra skill pasa algo similar, cuando el usuario no le da 3 objetos diferentes, Alexa toma los que se hayan dicho y pide los que falten uno por uno; lo mismo pasa para cuando se le tiene que dar un PIN o contraseña.

Cuando se responde algo incorrecto a una pregunta, por ejemplo, se pide que elija *“1”* o *“2”*, pero si el usuario nos dice *“3”*, entonces Alexa le responde que no es una respuesta válida para esa pregunta.

Conclusiones

Primero analizaremos las dificultades que tuvimos en el desarrollo de la skill y cómo fueron solucionadas:

1. Dado que es una historia consecutiva pensamos en dividir por bloques o párrafos la historia de tal manera que después de narrarlo le hiciéramos la pregunta de decisión, es decir, en el diagrama de flujo se puede ver como nodos.
2. Tuvimos problemas al implementar el intent para ingresar un PIN pues tenía problemas para distinguirlo del que sólo recibía 1 y 2, pero se solucionó al darle un ID a las preguntas.
3. Al principio tuvimos la dificultad de colocar sonidos externos para ambientar la historia, pero después incluimos los sonidos que están disponibles para la skill.
4. Un problema constante al cual no le pudimos dar explicación ni tampoco los ayudantes, fue que por un momento dejaba de responder la skill en media ejecución y había que ir al código a modificar cualquier cosa para que reaccionara (incluso algo que estuviera comentado). Eventualmente se quitó este error por sí solo.
5. Por último, lo más complicado fue pensar en una historia donde hubiera diferentes caminos y darle los cierres adecuados.

* Ventajas y desventajas

| Ventajas | Desventajas |
| --- | --- |
| Cuenta con comandos fáciles que el usuario puede recordar que son:  abre c.c., 1,2, ayuda, reproducir, repite. | No cuenta con una exploración libre tipo “escape room”, por lo que aún le falta para llegar a ser más natural la interacción. |
| Tiene una narración más natural con uso de marcadores y confirmaciones además de corrección en la mayoría de los errores. Por todo esto se vuelve una skill un tanto más profesional. | Hay un error el cual no pudimos resolver y los ayudantes nos estaban ayudando pero se nos acabó el tiempo.  El error consiste en que si le das una palabra que no reconoce el intent, te marca un error y te pide de nuevo la palabra, eso está bien, pero si pasa una segunda vez, la skill deja de responder. |
| Cuenta con varios caminos, un conjunto de objetos y diferentes finales, todo esto para que se pueda jugar más de una vez. | Dado que es una historia un tanto compleja ya que no contiene ciclos y debemos apegarnos a los objetos que llevan, se vuelve larga la historia y por lo tanto aún hay ramas sin cerrar. |
| No sólo es un juego de selección 1 o 2, también es un juego de selección de objetos y de claves. | Puede llegar a cansar al público que quiere una historia corta y concisa. |
| Cuenta con sonidos y distintas voces que nos ayuda a distinguir entre la narración de Alexa y los personajes. |  |
| Permite una experiencia más detallada en la cual se pueden descubrir pistas. |  |

* ¿Qué se puede mejorar?

1. Debemos corregir el error planteado en las desventajas
2. Terminar las ramas que faltan por cerrar
3. Incluir más sonidos que ambientan la historia
4. Volverlo más conversacional con la libertad de que el usuario explore sin que Alexa lo lleve de la mano.
5. La manera de decir el inventario para que el usuario no pida que se le repita más veces.

Gracias si leíste hasta aquí 🙂 por favor disfruta nuestra skill en el link de abajo…

[¡Pica aquí!](https://github.com/Kalato-Pari/Can-Clues-Version-Final.git)

Bibliografía:

Jalil, N. IntechOpen. (2021, 15 junio). *Introduction to Intelligent User Interfaces (IUIs)* Recuperado 4 de diciembre de 2022, de <https://www.intechopen.com/chapters/77196>

Embou. (2021, 25 marzo). *Qué son las Skills de Alexa y cómo instalarlas*. Recuperado 4 de diciembre de 2022, de <https://www.embou.com/blog/que-son-las-skills-de-alexa-y-como-instalarlas>

Fernández, Y. (2022, 4 marzo). *Qué es Alexa, qué puedes hacer con él y qué dispositivos son compatibles*. Xataka. <https://www.xataka.com/basics/que-alexa-que-puedes-hacer-que-dispositivos-compatibles>