**Documento de requerimientos de software**

***Habla Conmigo***

***Fecha: 25/09/2019***

**Tabla de contenido**

Historial de Versiones 3

Información del Proyecto 3

Aprobaciones 3

1. Propósito 4

2. Alcance del producto / Software 4

3. Referencias 4

4. Funcionalidades del producto 5

5. Clases y características de usuarios 5

6. Entorno operativo 6

7. Requerimientos funcionales 6

7.4 Identificación de configuraciones en base a letras y palabras simples en base a la detección de señas 7

7.5 Desplegar palabras y letras sencillas 8

7.6 Dictado de texto creado 9

8. Reglas de negocio 9

9. Requerimientos de interfaces externas 10

9.1. Interfaces de usuario 10

9.2. Interfaces de hardware 10

Especificaciones del equipo 18

9.3. Interfaces de software 18

9.4. Interfaces de comunicación 18

10. Requerimientos no funcionales 18

11. Otros requerimientos 19

12. Glosario 19

# Historial de Versiones

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fecha** | **Versión** | **Autor** | **Organización** | **Descripción** |
| **25/09/2019** | **1** | **Jose Ricardo Rea Aguilar/David Fernando** | **Edge Tech** | **Primera versión de Requerimientos** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

# Información del Proyecto

|  |  |
| --- | --- |
| Empresa / Organización | Edge Tech |
| Proyecto | Habla conmigo |
| Fecha de preparación | 25/09/2019 |
| Cliente | Lamia Hamdan Medina - GPS |
| Patrocinador principal | Heineken, LALA, Peñoles |
| Gerente / Líder de Proyecto | Liliana Gallegos Ruvalcaba |
| Gerente / Líder de Análisis de negocio y requerimientos | José Ricardo Rea Aguilar |

# Aprobaciones

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Nombre y Apellido** | **Cargo** | **Departamento u Organización** | **Fecha** | **Firma** |
| **Liliana Gallegos Ruvalcaba** | **Lider de Proyecto/Scrum Master** | **Edge Tech** | **25/09/2019** |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

# Propósito

El proyecto “Habla Conmigo” versión 0.1, es la primera liberación del software para traducir lenguaje de señas de la lengua mexicana. Cuenta con un módulo de reconocimiento de objetos que por medio de inteligencia artificial en proceso de aprendizaje puede identificar diferentes señas e interpretarlas como dictado de letras para la construcción de palabras o frases sencillas. Este documento abarca la parte inicial del proyecto donde se tienen definidos en primera iteración los mockups de la versión pre alfa de la interfaz de usuario.

# Alcance del producto / Software

Desarrollo de una aplicación de escritorio, capaz de interpretar señas y configuraciones de mano a través de una cámara y traducirlas a letras y palabras, para una reinterpretación leída, o dictada por la misma aplicación, brindando la capacidad de comunicar gente sordomuda o que solo habla lenguaje de señas con cualquier persona que sea capaz de leer el texto generado o escuchar el dictado de la aplicación, rompiendo limitantes de comunicación, agilizando este proceso y permitiendo ser incluyentes con estas personas.

# Referencias

|  |  |
| --- | --- |
| [1] | E. Chacon, D. Aguilar y F. Saenz, «Desarrollo de una Interfaz para el Reconocimiento Automático del Lenguaje de Signos,» de *Desarrollo de una Interfaz para el Reconocimiento Automático del Lenguaje de Signos*, vol. 4, Universidad de las Fuerzas Armadas, 2014, pp. 14-21. |
| [2] | J. Garcia Mejia, L. Tinoco Monroy y G. Gonzalez Hernandez, «Modelado e implementación de unsistema de enseñanza de Lenguaje deSeñas Mexicano,» *Red de Revistas Científicas de América Latina, el Caribe, España y Portugal,* vol. 14, nº 2, pp. 185-190, 2007. |
| [3] | L. B. F. Mintz, «Verbos y espacios mentales en la Lengua de Señas Mexicana,» de *Verbos y espacios mentales en la Lengua de Señas Mexicana*, Colima, Universidad de Colima, INAH-SEP, 1996. |
| [4] | I. Orduña Claudio y L. G. Ruiz Serrano, «Morfemas clasificadores en la Lengua de Señas Mexicana,» *Signos lingüísticos,* vol. 6, nº 11, pp. 61-84, 2010. |

# Funcionalidades del producto

* Captura y almacenamiento de imágenes en memoria RAM mediante web cam.
* Detección de objetos (Manos) en las imágenes capturadas.
* Categorización de señas y configuraciones para interpretación en letras y palabras.
* Redacción de las letras y palabras interpretadas en una interfaz gráfica.
* Dictado del texto generado por la aplicación.

# Clases y características de usuarios

El software está enfocado para los usuarios de la LSM los cuales no les es posible comunicarse con el resto de la población, a través del software podrán traducir su método de comunicación a texto el cual será fácilmente entendido por la población general. Las principales características de uso del software en este grupo de usuarios son:

* Detección de objetos (Manos) en las imágenes capturadas.
* Redacción de las letras y palabras interpretadas en una interfaz gráfica.

El segundo grupo de usuarios que utilizan el software es la población general, la cual no sufre de ninguna discapacidad auditiva. El principal uso de este grupo de usuarios es el de entrenarse y familiarizarse con la LSM para una lectura más fluida y tener mejor entendimiento de las señas en el lenguaje. Las características que usaran del software son:

* Categorización de señas y configuraciones para interpretación en letras y palabras.
* Dictado del texto generado por la aplicación.

El tercer grupo de usuarios son los que ayudan a la inteligencia artificial a aprender por medio del uso de lenguaje de señas, este tipo de usuarios se dedicará a mostrar al programa señas y patrones por los cuales la inteligencia artificial podrá aprender con “Live Data” para poder analizar los diferentes objetos (manos) en distintos tipos de condiciones (iluminación, tamaño, color, etc.). Las funcionalidades a utilizar son:

* Captura y almacenamiento de imágenes.
* Detección de objetos en las imágenes capturadas.
* Categorización de señas y configuraciones para interpretación en letras y palabras.

# Entorno operativo

La aplicación ejecuta una GUI en pantalla que muestra una caja de texto, que desplegara letras y palabras, contara con un botón para efectuar el dictado del texto contenido y la opción de desplegar la imagen captada (con leve retraso) por la web cam, almacenara las imágenes en memoria RAM para efectuar su análisis despliegue y por fin su desecho y liberación de memoria.

Requisitos para su ejecución:

* Windows 7 o superior
* Python 2.7.5
* TensorFlow
* .NET Framework 4.5

# Requerimientos funcionales

* Aplicación de escritorio
* Recolección de imágenes o video
* Detección de señas de manos en imágenes recolectadas
* Identificación de configuraciones en base a letras y palabras simples en base a la detección de señas
* Desplegar palabras y letras sencillas
* Dictado de texto creado

## 7.1 Aplicación de escritorio

Descripción: Se requiere una aplicación de escritorio para Windows que se ejecute en java.

Prioridad: Alto

Acciones iniciadoras y comportamiento esperados: deberá contar con un archivo ejecutable que dispare la aplicación.

REQ–1: Contener el texto generado por la interpretación.

REQ–2: Contener y desplegar las imágenes capturadas

## 7.2 Recolección de imágenes o video

Descripción: Se requiere acceso a una cámara web conectada al equipo para recolectar la secuencia de imágenes a analizar.

Prioridad: Medio

Acciones iniciadoras y comportamiento esperados: Al iniciar la aplicación será consultado si existe un dispositivo conectado para continuar la interacción, si no se cuenta con él se notificara y se cerrara la aplicación.

REQ–1: Detección de un dispositivo existente.

REQ–2: Validar el acceso a la captura de datos mediante este dispositivo.

REQ–3: Almacenamiento de imágenes en un buffer que permita acceder a estas durante la ejecución de la aplicación (Borradas o desechadas al cerrar la aplicación, o al ser procesadas).

## 7.3 Detección de señas de mano en imágenes recolectadas

Descripción: Se requiere acceso al buffer que recolecta la captura de imágenes que realiza la aplicación, se procesan las imágenes para detectar las manos en las imágenes y lograr minimizar el tamaño de la imagen analizada y agilizar el proceso.

Prioridad: Medio

Acciones iniciadoras y comportamiento esperados: Al recolectar imágenes en el buffer este proceso comienza a identificar las manos en las imágenes capturadas y recorta las imágenes para colocarlas en un buffer que será usado después.

REQ–1: Acceso al buffer de captura.

REQ–2: Acceso a la GPU para agilizar cálculos.

REQ–3: Acceso a un buffer de imágenes pre procesadas.

## 7.4 Identificación de configuraciones en base a letras y palabras simples en base a la detección de señas

Descripción: se requiere que el software analice por medio de la cámara web las señas y patrones de movimientos que se le presenten para identificar las letras y/o configuraciones de palabras creadas.

Prioridad: Alto

Acciones iniciadoras y comportamiento esperados: a partir de la inicialización de la sección de detección de señas, el software estará escuchando constantemente al dispositivo de entrada (cámara web) para identificar señas y patrones, en base al conocimiento adquirido para la detección, al identificar como positiva una seña el valor de esta será almacenado en memoria y regresará a la fase de detección; esto ocurrirá sucesivamente hasta que el usuario de la instrucción de finalización.

Requerimientos funcionales:

REQ–1: detección acertada de señas, se requiere que el software sea capaz de reconocer las señas mostradas a través de la cámara.

REQ-2: almacenamiento de los valores significativos del reconocimiento, se deben guardar los valores correspondientes a las señas detectadas previamente para su posterior utilización o despliegue.

REQ-3: confirmación de detección de seña, el software debe mostrar al usuario cuando una seña o configuración ha sido reconocida correctamente y pasara a la identificación de una nueva seña.

REQ-4: notificación de fallo en la detección, si el software no es capaz de reconocer las señas indicadas, este lo hará saber mediante una notificación en pantalla.

## Desplegar palabras y letras sencillas

Descripción: el software por medio de la interfaz gráfica despliega la información de las letras y/o palabras que se van construyendo en base a la detección de señas.

Prioridad: Alto

Acciones iniciadoras y comportamientos esperados: una vez que el software haya indicado la detección de una nueva seña, luego de la notificación en pantalla, el valor de esta será desplegada en pantalla para que el usuario pueda monitorear la palabra o frase que va construyendo por medio de las señas. Esto sucederá cada vez que se indique una nueva detección para que el usuario tenga una buena retroalimentación por parte del programa.

Requerimientos funcionales:

REQ-1: despliegue de valores indicados, el software debe indicar en pantalla los valores de cada seña que detecte durante el proceso de reconocimiento.

REQ-2: identificación errónea de una seña, en caso de que el software interprete equivocadamente una seña, el usuario podrá indicar la regresión en el proceso de manera que se elimine la última detección y pueda volver a intentar reconocer lo indicado.

## Dictado de texto creado

Descripción: una vez finalizada la etapa de reconocimiento, el software tiene la opción de dictar la frase o texto creado a partir de las configuraciones reconocidas; dando así, un aspecto auditivo para una comunicación más completa por parte del traductor.

Prioridad: Baja

Acciones iniciadoras y comportamiento esperado: después de finalizada la etapa de detección, en caso de así indicarlo, el software hará el dictado del texto por medio de un sintetizador de voz.

Requerimientos funcionales:

REQ-1: anuncio de inicio de dictado, una vez que finalice la detección de señas, el software indicara que empezara a dictar el texto producido.

REQ-2: dictado de texto producido, por medio del sintetizador de voz se reproducirá el texto creado a partir de las señas detectadas anteriormente.

REQ-3: anuncio de fin de dictado, al igual que en el inicio, se hará un anuncio cuando el dictado por parte del software haya finalizado.

# Reglas de negocio

La aplicación debe ser accesible para cualquier usuario, no debe contar con bloqueos ni acceso, para ser manejada con facilidad, al iniciar se debe mostrar una caja de texto que contendrá el texto interpretado que deberá dar la opción de mostrar la imagen en (pseudo) tiempo real que se está capturando e interpretando.

Al no detectar dispositivo de captura, será alertado el usuario esperando su confirmación en una ventana de dialogo, la cual al ser cerrada procederá a cerrar la aplicación. Mostrará un botón para cerrar la aplicación, el cual predirá una confirmación y al intentar cerrar se desplegará una ventana de confirmación la cual deberá ser contestada.

# Requerimientos de interfaces externas

## Interfaces de usuario

Al inicio de la aplicación, se muestra un loading por unos segundos hasta que la aplicación esté completamente cargada, en la ilustración 1, podemos observar el logo de la aplicación.



Ilustración 1 Loading de la aplicación

Después de que la aplicación de escritorio ha sido cargada exitosamente como nos muestra en la ilustración 2, entraremos a la interfaz gráfica principal, la cual solo podrá reconocer 3 botones, ayuda, iniciar y cambio de tema de luz.

Con el fin de una experiencia más sencilla y evitando errores en su uso, en cada ventana se le asignó su botón de ayuda correspondiente, por ejemplo, en la ilustración 3 nos muestra lo que el usuario podrá ver si da click al botón de ayuda.

Para el cambio de luz o conocido mundialmente como ‘dark mode’ o ‘modo oscuro’, es una funcionalidad que contará la aplicación, ya que cuenta con los siguientes beneficios:

* Puede reducir el consumo de energía en una cantidad significativa (según la tecnología de pantalla del dispositivo).
* Mejora la visibilidad para los usuarios con visión reducida y aquellos con sensibilidad a la luz brillante.
* Facilita el uso de un dispositivo en un entorno con poca luz

Primero ilustraremos la interfaz gráfica en modo oscuro a partir de la ilustración 1 a la 9 y para concluir el modo automático, mostrado en la ilustración 10 a 13

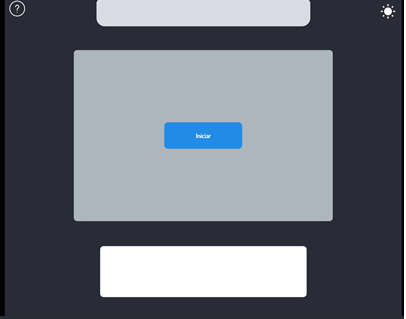


Ilustración 2 Pantalla principal

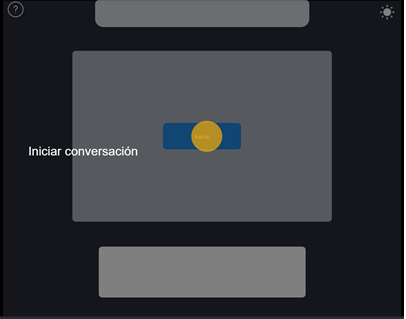


Ilustración 3 Parte de la animación del botón ayuda de pantalla principal

La siguiente acción del usuario, nos llevará a una parte muy importante de la aplicación (ilustración 4), la selección del puerto en la que está ubicada la cámara web del usuario, es importante esta ventana, ya que la cámara web será el lector de los datos de entrada necesarios para el funcionamiento de la aplicación, de nueva cuenta, se agregó el botón de ayuda ubicado en la esquina superior izquierda (ilustración 5).

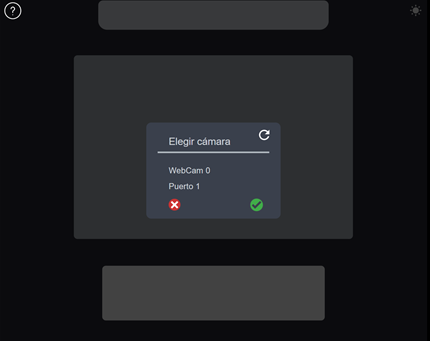


Ilustración 4 Venta para selección de puerto.

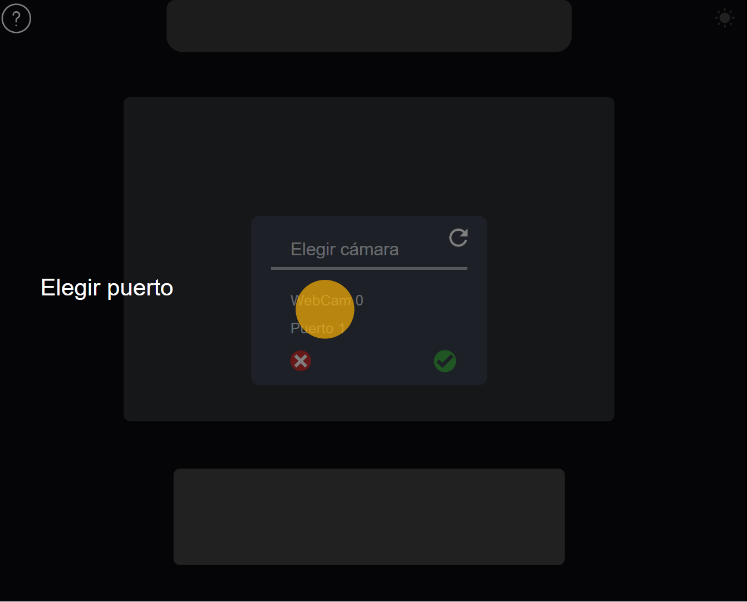


Ilustración 5 Parte de la animación del botón de ayuda de la ventana Selección de Puerto

A continuación, después de la elección del puerto donde está ubicado la cámara web, por automático se mostrará la ilustración 6, contiene la imagen de entrada junto a un menú en la parte superior, consiste en tres opciones, leer en voz alta la conversación obtenida, visualizar/ ocultar la imagen de entrada de la webcam (mostrada en la ilustración 8) y guardar conversación a un archivo de texto.

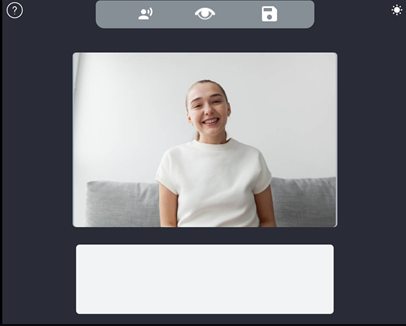


Ilustración 6

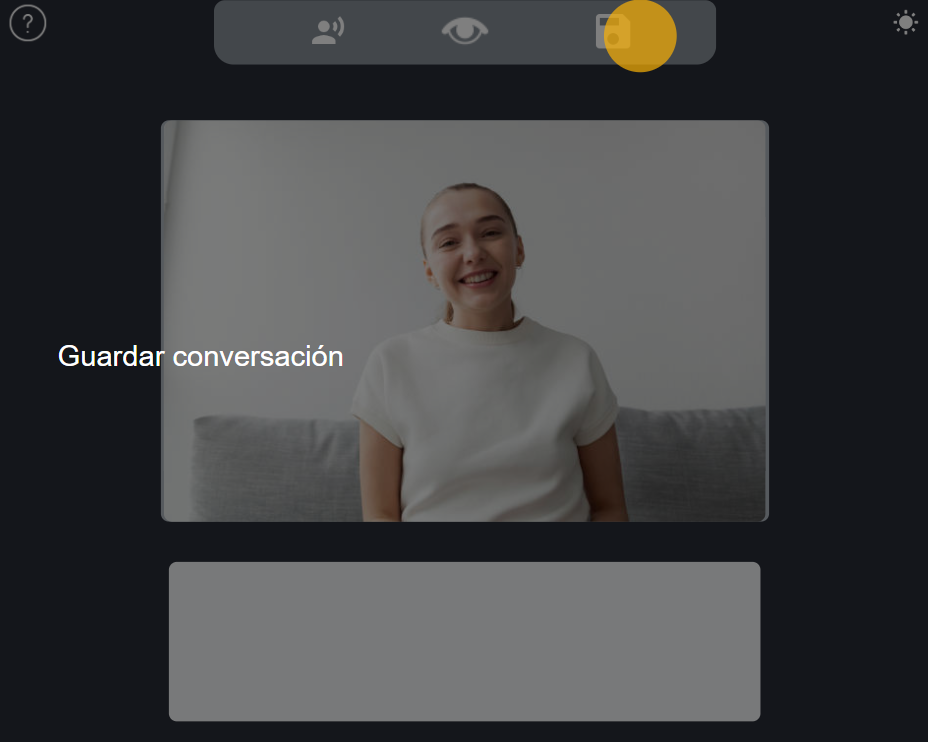


Ilustración 7 Parte de la animación del botón de ayuda de la ilustración 6

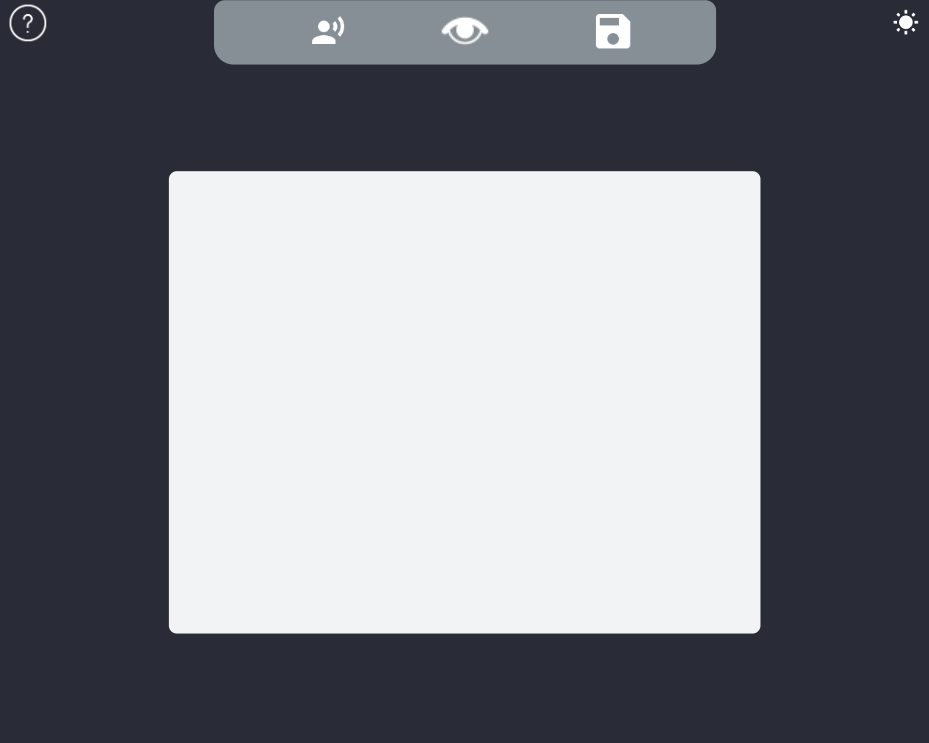


Ilustración 8 respuesta de la acción visualizar/ ocultar imagen de entrada

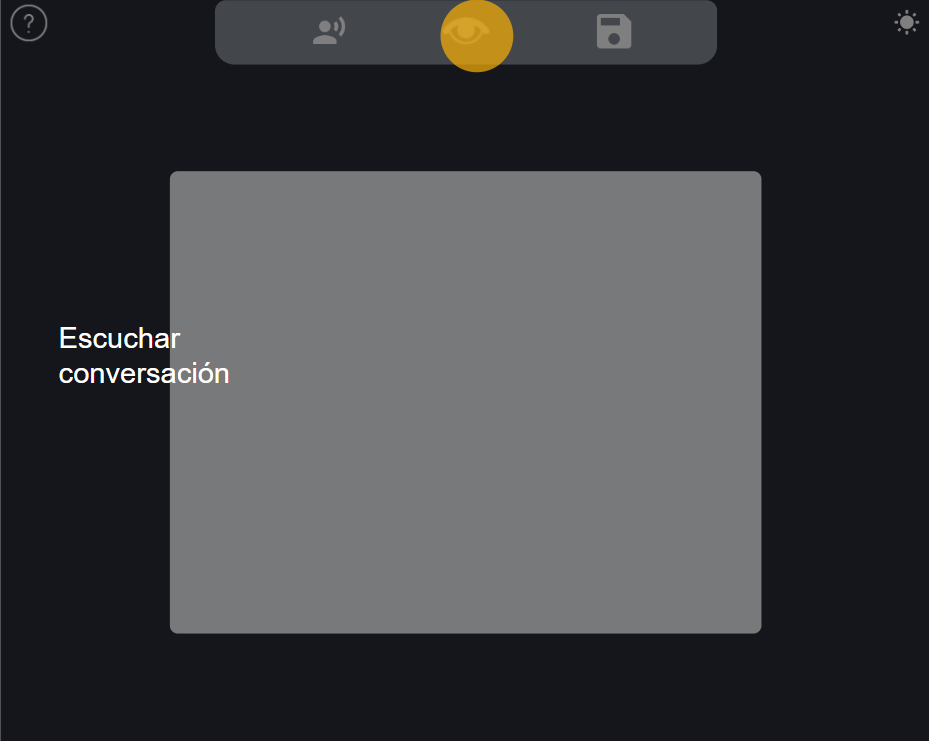


Ilustración 9 parte de la animación de la ilustración 8

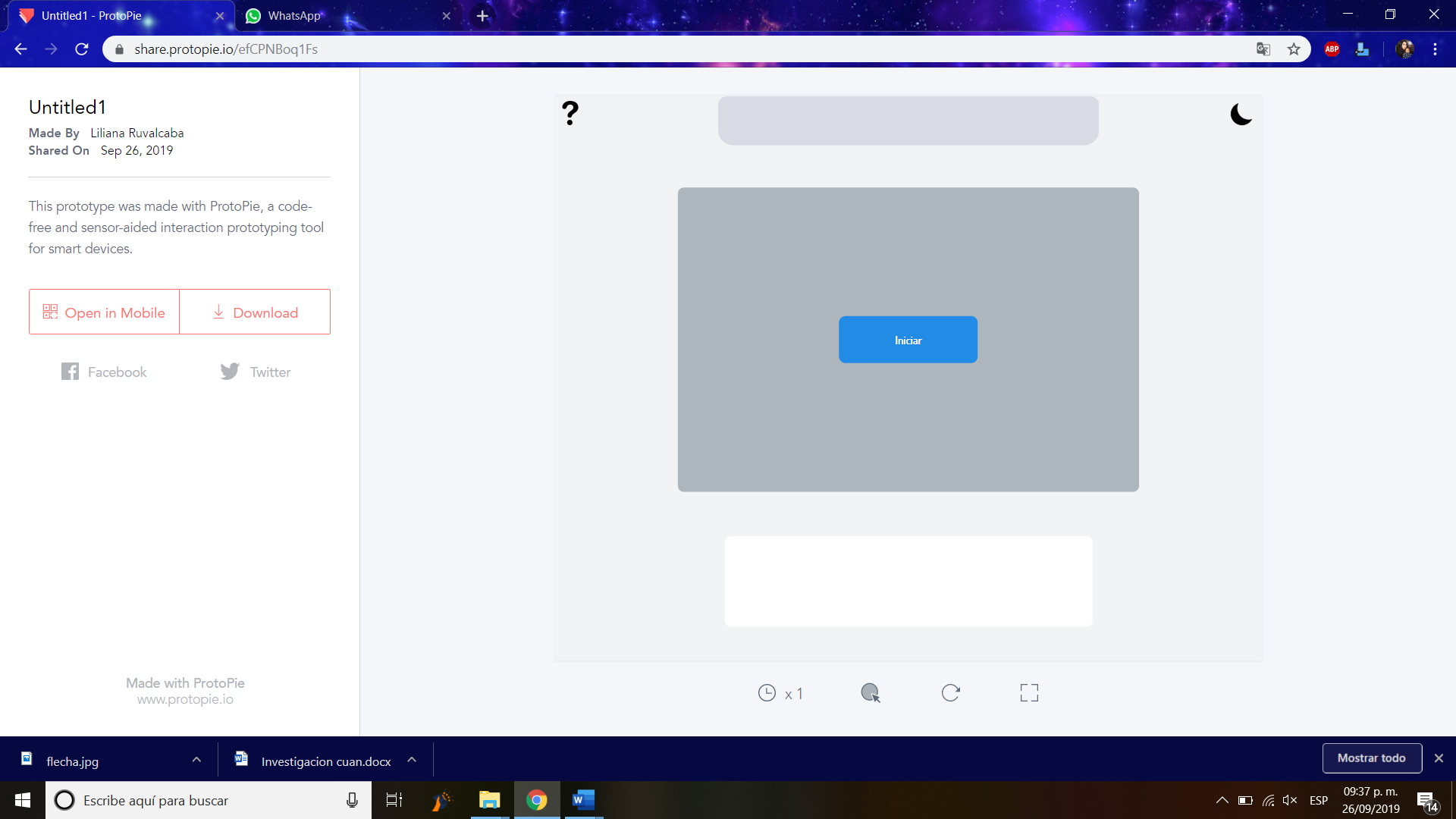


Ilustración 10 modo automático de pantalla principal

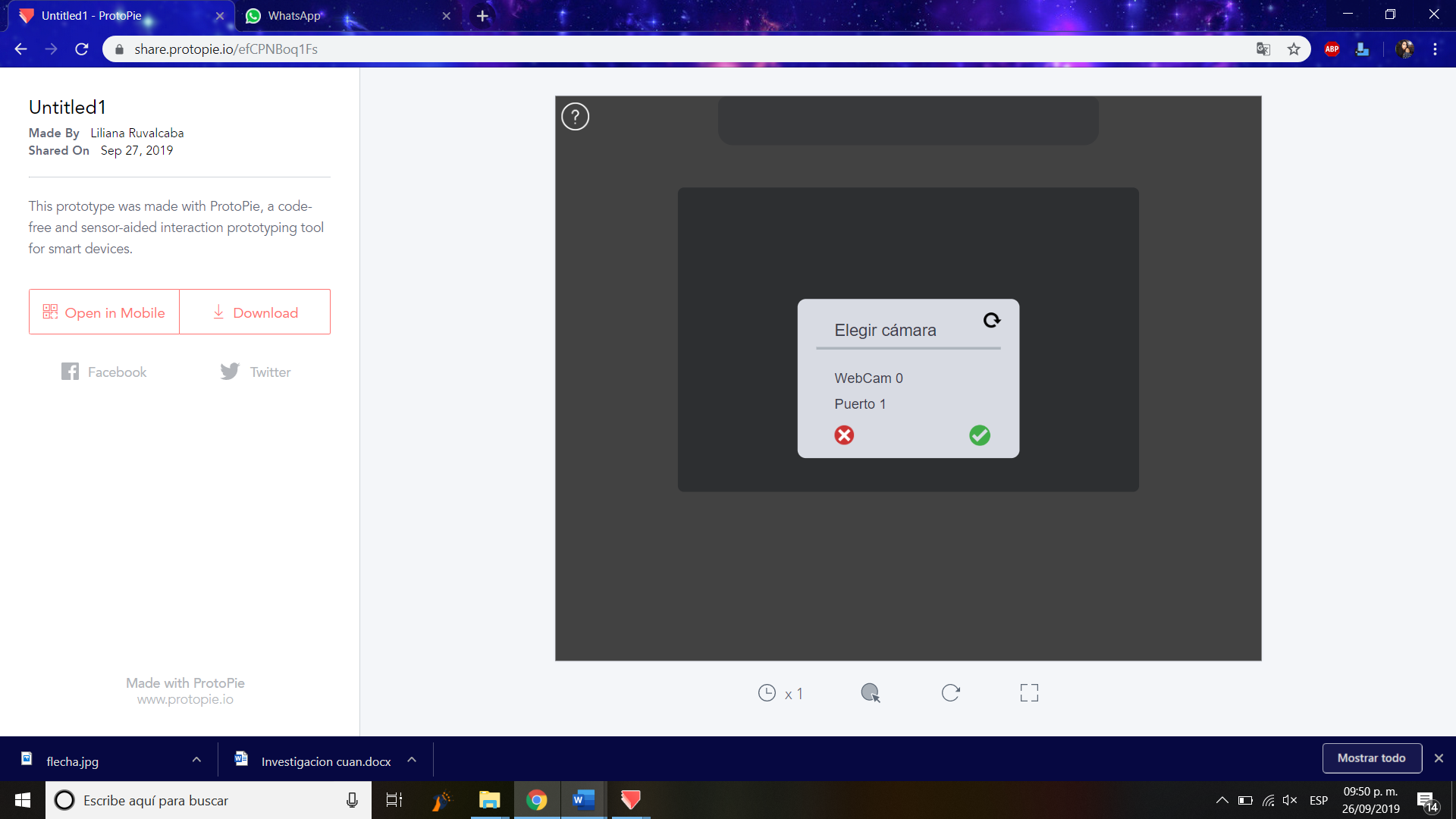


Ilustración 11

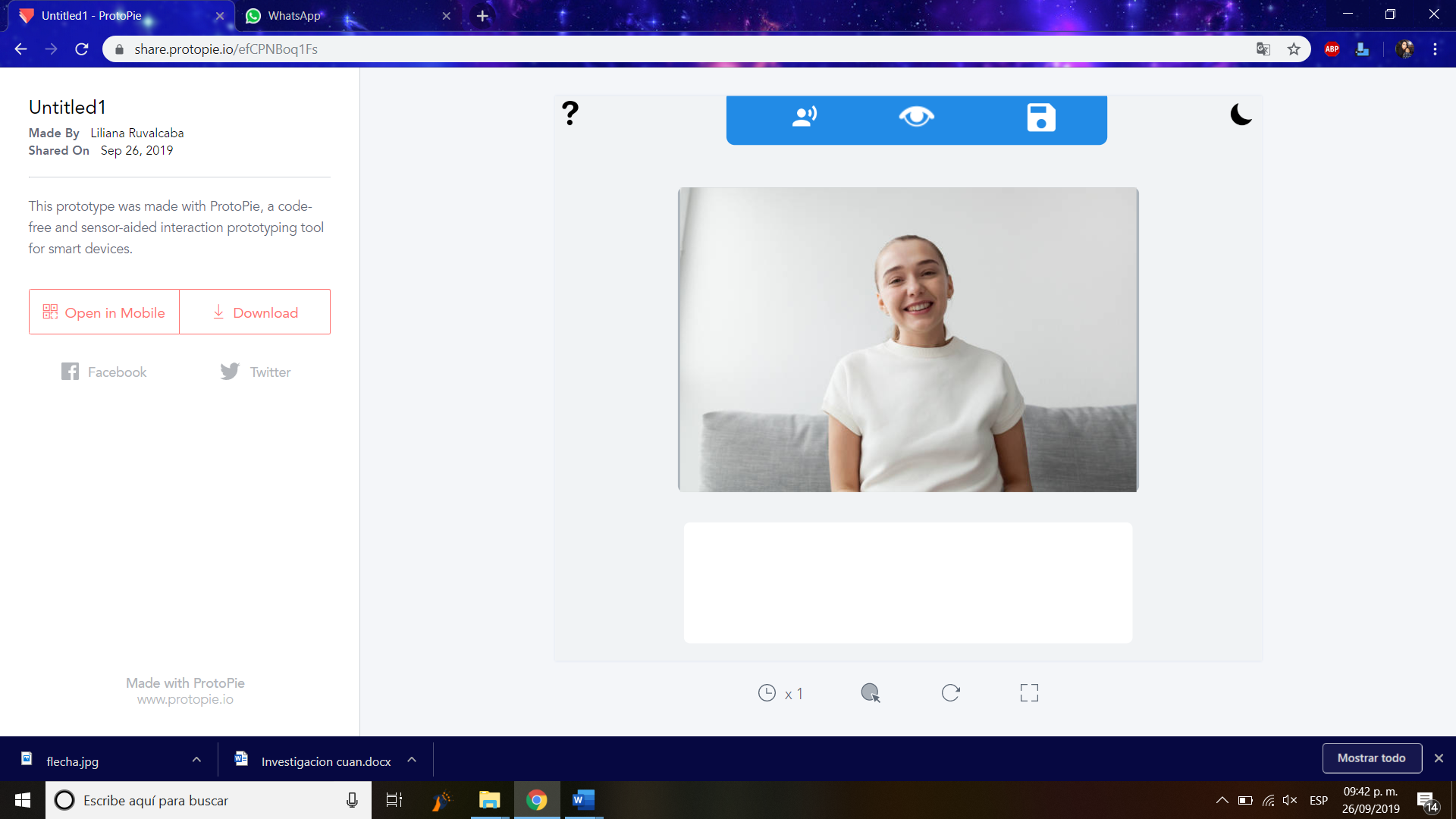


Ilustración 12

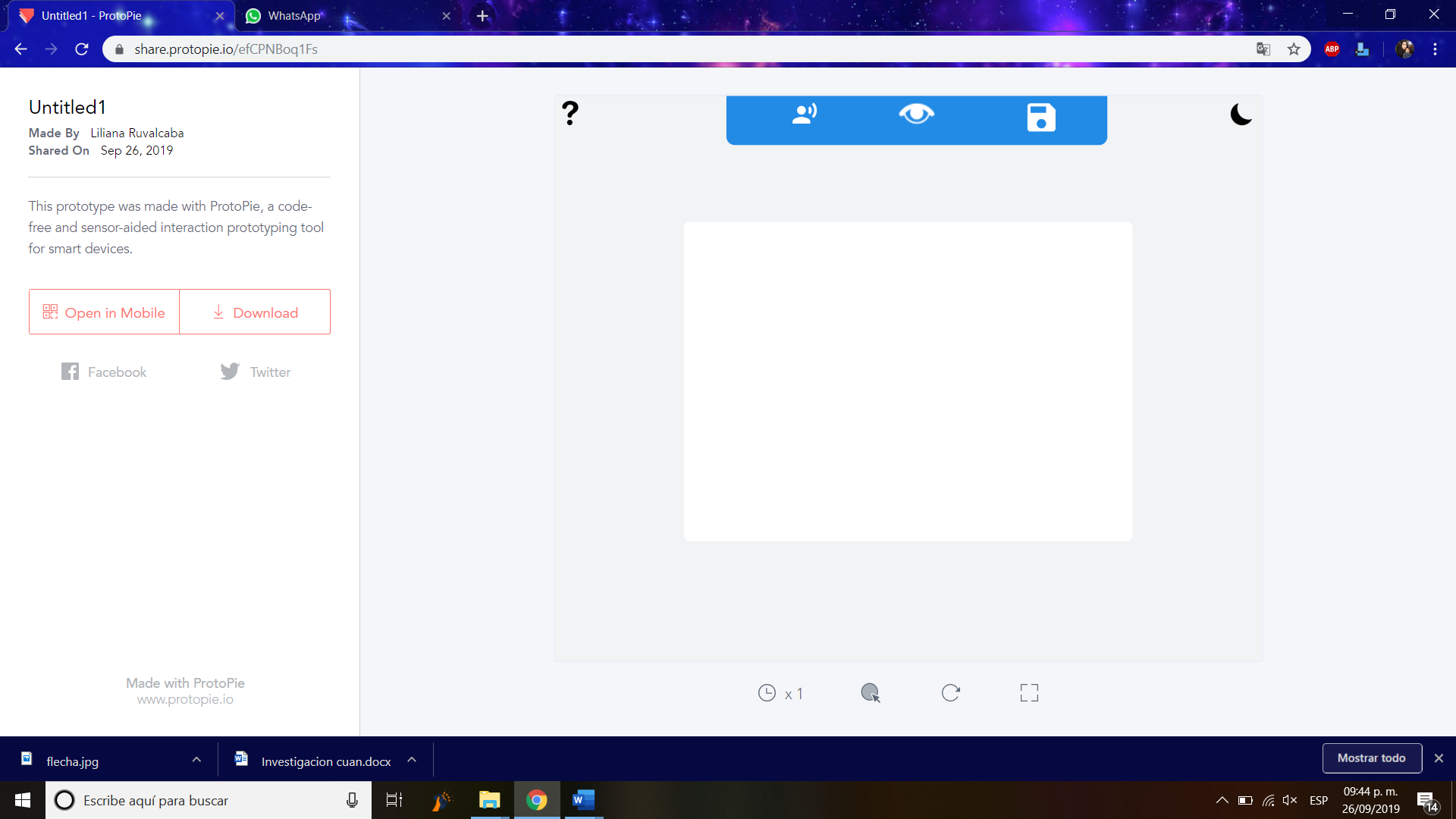


Ilustración 13

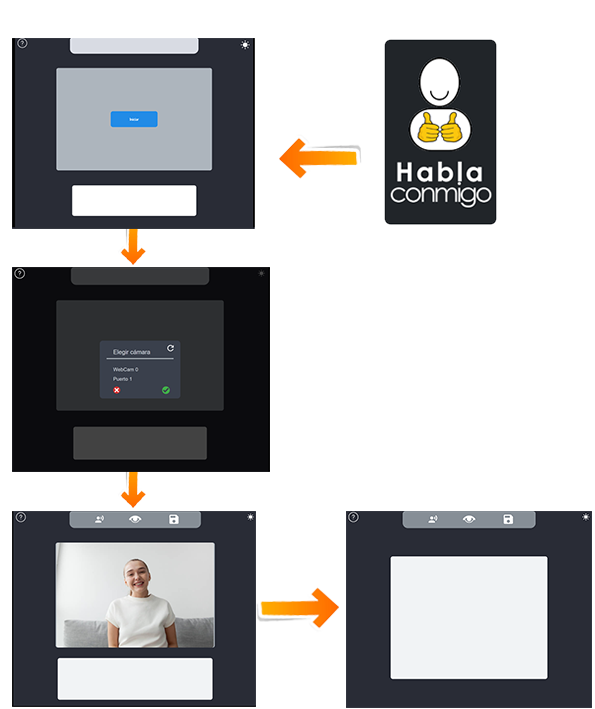


Ilustración 14 Diagrama básico de acciones de la aplicación

## Interfaces de hardware

## Especificaciones del equipo

* Computador con procesador Core i3
* 4GB RAM
* 2 GB de espacio en disco duro
* Windows 7 o superior
* GPU

**Protocolos de comunicación**

* Protocolo de comunicación USB (Para uso de WEBCAM)
* DSP (Uso de tarjeta de audio)
* AVX2 (Uso de GPU)

## Interfaces de software

Interfaz gráfica de Windows, driver de audio, driver de cámara web, API CUDA (Uso de CPU).

## Interfaces de comunicación

Protocolo stdio para intercomunicar a python con JAVA.

# Requerimientos no funcionales

La aplicación debe:

• Ser agradable a la vista.

• Tener un diseño bonito y moderno.

• Tener textos claros y legibles.

• Tener maro.

• Poder ser usada por niños y adultos.

• No debe fallar.

• Ser eficiente.

• Visualización fluida de video.

• Consumir pocos recursos.

# Otros requerimientos

REQ-1: API de sintetizador de voz, se requiere de una librería que ayude al dictado por medio de voces artificiales que sean claras y entendibles.

REQ-2: API de machine learning, se requiere componentes de software que ayuden al aprendizaje de la inteligencia artificial encargada de reconocer las configuraciones.

REQ-3: Certificación CONADIS, para lograr ser un software reconocido y comercial se necesita tener una constancia de certificación por parte de un organismo del gobierno mexicano, en este caso CONADIS (Consejo Nacional para el Desarrollo y la Inclusión de personas con Discapacidad), que avale que lo que está promoviendo el software cumple con las normas de la LSM.

# Glosario

Configuraciones:

Posición que pueden tomar las manos para indicar una palabra o letra.

LSM:

Lengua de señas mexicanas.

API:

Application Program Interface, conjunto de rutinas que provee acceso a funciones de un determinado software.