

Universidad Siglo 21



Trabajo Final De Grado. Prototipado Tecnológico

Carrera: Licenciatura en informática

Plataforma de Traducción de Audio a Lenguas de Señas Argentinas

Autora: Angie Evelyn Manassero

Legajo: VIN06599

Córdoba, julio de 2021

Índice

Resumen	4
Abstract	5
Título	6
Introducción	6
Antecedentes	6
Descripción Del Área Problemática	7
Justificación	15
Objetivo General del Proyecto	16
Objetivos Específicos del Proyecto	16
Marco Teórico Referencial	16
Dominio del problema	16
TICs	18
Competencia	21
Diseño Metodológico	22
Relevamiento	26
Relevamiento Estructural	26
Relevamiento Funcional	27
Procesos de negocios	29
Diagnóstico y Propuesta	31
Objetivo, Límites y Alcance del Prototipo	36
Objetivo del prototipo	36
Límites	36
Alcances	36
Descripción Del Sistema	37
Product Backlog	37

Historias de usuario	37
Sprint Backlog	41
Estructura de datos	43
Prototipo de interfaz de pantalla	47
Diagrama de arquitectura	49
Diagrama de componentes	53
Seguridad	55
Análisis de Costos	55
Análisis de Riesgos	58
Conclusiones	63
Demo	64
Referencias	65

Índice de Tablas

Tabla 1 - Comparación entre distintas soluciones	22
Tabla 2 - Product Backlog	37
Tabla 3 – Historia de usuario S01	37
Tabla 4 - Historia de usuario S02	38
Tabla 5 - Historia de usuario S03	38
Tabla 6 - Historia de usuario S04	39
Tabla 7 - Historia de usuario S05	39
Tabla 8 - Historia de usuario S06	40
Tabla 9 - Historia de usuario S07	40
Tabla 10 - Historia de usuario S08	41
Tabla 11 - Sprint Backlog	41
Tabla 12 - Costos de desarrollo	56
Tabla 13 - Análisis de transferencias	56
Tabla 14 - Costos variables	57

Tabla 15 - Costos Fijos	57
Tabla 16 - Costo total de infraestructura	58
Tabla 17 - Riesgos identificados del proyecto	58
Tabla 18 - Análisis cualitativo de riesgos	59
Tabla 19 - Relación entre parámetros cualitativos y cuantitativos	60
Tabla 20 - Asignación de parámetros cuantitativos	60
Tabla 21 - Matriz de probabilidad e impacto.	61
Tabla 22 - Análisis cuantitativo y grado de exposición	61
Tabla 23 - Plan de contingencias	62

Índice de Imágenes

Imagen 1 – Diagrama de Gantt	25
Imagen 2 - Organigrama de la organización modelada	26
Imagen 3 - Proceso 1: Generación de contenido audiovisual.	29
Imagen 4 - Proceso 2: Interpretación de información general	30
Imagen 5 - Diagrama entidad-relación	44
Imagen 6 - Diagrama de clases	45
Imagen 7 - Prototipo de interfaz de pantalla: aplicación de ejemplo.	47
Imagen 8 - Prototipo de interfaz de pantalla: aplicación de ejemplo.	48
Imagen 9 - Análisis de ratio de contraste de la interfaz	49
Imagen 10 - Diagrama de infraestructura	52
Imagen 11 - Diagrama de componentes	54
Imagen 12 - Gráfico de Pareto	61

Resumen

La ausencia de mecanismos concretos que faciliten una integración efectiva de la población sorda a través de su lengua materna (Lenguas de Señas Argentinas) se transforma en un severo limitante que deriva en factor de exclusión y en una ampliación de las brechas de desigualdades en todos los ámbitos de manera transversal. En los espacios digitales, no existen herramientas masificadas que den respuesta a las necesidades puntuales de la población sorda, por lo tanto, esta situación deriva en el acceso desigualitario al contenido y al acceso restringido al entorno digital. Por esta razón, en este proyecto se desarrolló una plataforma de software libre encargada de recibir audios desde diversas fuentes y generar la traducción simultánea en Lenguas de Señas Argentinas, a través de una representación de animaciones tridimensionales. El objetivo principal fue construir una solución que puede ser implementada fácilmente dentro de cualquier proyecto de software existente en internet, para favorecer la adopción global progresiva y acompañar a la comunicación de manera síncrona, haciendo del entorno digital, un espacio más accesible y diverso.

Palabras Clave: Accesibilidad, Inclusión, Tecnologías de apoyo, Población Sorda, Lenguaje de Señas Argentinas.

Abstract

The absence of specific mechanisms to facilitate the effective integration of the deaf population through their mother tongue (Argentine Sign Languages) becomes a severe constraint that leads to exclusion and a widening of the gap of inequalities in all areas in a transversal way. In digital spaces, there are no mass tools that respond to the specific needs of the deaf population, therefore, this situation leads to unequal access to content and restricted access to the digital environment. For this reason, in this project was developed a free software platform that receives audios from multiple sources and generate a simultaneous translation in Argentine Sign Languages, through a representation of three-dimensional animations. The main objective was to build a solution that can be easily implemented within any existing software project on the internet, to promote progressive global adoption and accompany communication in a synchronous manner, making the digital environment a more accessible and diverse space.

Keywords: Accessibility, Inclusion, Support Technologies, Deaf Population, Argentine Sign Language.

Título

Plataforma de traducción de audio a Lenguas de Señas Argentinas

Introducción

La revolución digital trajo consigo un abanico de posibilidades que representa un acceso ilimitado al conocimiento y a nuevas formas de desarrollo, evolucionando la dinámica de interacción social. Pero también significa un nuevo desafío, donde la falta de acceso a las nuevas tecnologías implica la ampliación de la brecha de desigualdades.

En el caso de las personas con discapacidades, esta problemática se complejiza, donde no solo es una variable el propio acceso a la tecnología sino que también surge la necesidad de mecanismos de integración para lograr un acceso efectivo.

En el caso particular de las personas sordas, la falta de herramientas que les permita acceder al contenido en su lengua materna (Lenguas de Señas Argentinas) se transforma en un severo limitante y en un factor de exclusión que impactan a los afectados en la transversalidad: desde el entorno familiar al laboral, desde el acceso a la educación a sus derechos cívicos y humanos. Por lo tanto, este proyecto busca generar una solución que derribe las barreras comunicacionales a través de una plataforma de traducción de audio a lenguas de señas argentinas (LSA) en tiempo real, para favorecer el acceso integral de las personas sordas en espacios digitales.

Antecedentes

Según la Organización Mundial de la Salud, las personas que sufren una discapacidad permanente representan el 10% de la población mundial. Actualmente en nuestro país, el Estudio Nacional sobre el Perfil de las Personas con Discapacidad (Instituto Nacional de Estadísticas y Censos, 2018) estableció que la discapacidad auditiva corresponde al 20,8% del total de las discapacidades.

Por otro lado, la interpretación español-Lenguas de Señas Argentinas es una actividad profesional relativamente reciente, que nace bajo la necesidad de la población sorda a comunicarse y recibir la información a través de su lengua materna natural. La actividad fue ejercida históricamente por familiares y allegados de las personas sordas e

intérpretes autodidactas. Comenzó a formalizarse como profesión en los últimos años, por lo cual la formación académica es reciente, al igual que la regulación de la actividad. Sin embargo, no se cuenta con áreas de especialización (necesarias para la realización de interpretaciones en campos específicos fuera del ámbito cotidiano) y tampoco se cuenta con formaciones académicas en el campo de interpretación de producción audiovisual.

En nuestro país no se registran datos sobre la cantidad de hablantes LSA y tampoco sobre la cantidad de intérpretes per cápita ni su distribución, tampoco se registra la existencia de un registro nacional de intérpretes, y solo se cuenta con matriculación profesional en la provincia de Mendoza. Esta situación deriva en la inexistencia de mecanismos claros de accesibilidad que sean facilitadores de una inclusión efectiva.

Por otro lado, en entornos digitales nos encontramos con la práctica masificada de integración de subtítulos en el contenido multimedia o la utilización de *closed captions* (texto que proporciona información interpretativa adicional) y si bien la utilización de estos recursos son variables que facilitan la accesibilidad del contenido a personas con pérdidas parciales de la audición u otras dificultades auditivas, no es una solución que de respuesta a las necesidades de la población sorda no bilingüe, ya que requiere de una pedagogización idiomática previa (es decir, la adopción anterior y exhaustiva del idioma español) dejando de lado a gran parte de la población no alfabetizada.

Descripción Del Área Problemática

Las causas que derivan en la situación actual se dan por diversos factores que convergen entre sí, por lo tanto, el estudio de la problemática se lleva a cabo con una visión transversal para generar un entendimiento integral de la problemática.

Aspectos Psicológicos y Neurobiológicos: Los fundamentos del LSA.

La historia de la comunicación y la educación de las personas sordas fue mutando desde una visión clínica y oralista de la sordera, a una visión antropológica que valoriza al lenguaje de señas como el lenguaje identitario de las personas sordas, representando un nuevo paradigma de bilingüismo. Es decir, lenguaje de señas es la lengua materna de las

personas sordas, y además es un elemento referencial, comunicativo y social que manifiesta la identidad misma, favoreciendo el desarrollo de las competencias lingüísticas y cognitivas.

Desde una concepción psicolingüística, “El lenguaje es una facultad innata que supone un proceso en el que se ven involucrados aspectos lingüísticos diversos para lograr la comunicación efectiva” (Lightbown y Spada, 2006).

Según Oviedo (2003), el ser humano nace con la capacidad para adquirir la lengua que se habla a su alrededor, por lo tanto, el ser humano utiliza esta capacidad para apropiarse del idioma de su entorno.

Cuando alguien nace sordo, no puede escuchar lo que se habla en su entorno, y de allí que tampoco pueda usar esa capacidad natural para aprender la lengua que hablan quienes están alrededor suyo. La naturaleza se las ingenia entonces para suplir la falta del sentido del oído. Los sordos no oyen, pero ven, y se dan cuenta de que una enorme cantidad de información se comunica con las expresiones de la cara, con las posturas y movimientos corporales, y comienzan a hacer uso de esos recursos para expresarse.

Por otro lado, Juliarena (2012) afirma que la lengua natural de los sordos es la que se actualiza en el canal viso-espacial, dado que el canal sonoro está fuertemente interferido, como tal no puede apartarse de las características psicofisiológicas. En ese sentido, el lenguaje de señas -a diferencia del oralismo- se incorpora sin la necesidad de pedagogizarse.

Históricamente, la concepción social de los sordos estuvo relegada a una situación de enfermedad, es decir, en esta construcción de la figura del sordo, este es pensado como un ‘oyente fallado’; esto derivó a que en el pasado las lenguas de señas fueran relegadas a la prohibición de su uso en el ámbito educativo, y a pesar que la situación pedagógica haya cambiado, el método privilegiado prevaleciente es el oralismo:

El ideal de sus prácticas combinadas es que el niño sordo adquiera la lengua oral -conductas de oyente- para integrarlo en un marco de

“concepción igualitaria” a través del dominio del cambio corporal y conductual, mediante la acción rehabilitadora (medicina)-normalizadora (educación). Se trata de corregir el “desvío” (conducta)-falla (biología), es decir, recuperar la completitud de un cuerpo, que el sordo se asemeje al oyente, que se aproxime a la norma. (Rey, Famularo y Ringuelet, 2020)

Es desde este aspecto donde ocurre una imposición social sobre esta población en su manera de auto percibirse y relacionarse con el mundo, al no considerarse al LSA como lengua materna, identitaria y de alguna forma étnica, un sistema válido de comunicación. Es decir, concebimos la idea de una persona extranjera desde la identidad cultural que incluye a la lengua como un aspecto fundamental asociado (por ejemplo: argentino - hablante español, Brasileiro - hablante portugués), sin embargo, no ocurre lo mismo con la población sorda, puesto que en la concepción social su idioma nativo resulta irrelevante, y por lo tanto, se intenta forzar una comunicación que les resulta compleja y limitante con el objetivo de ‘estandarizarlos’ a la norma.

Es en este sentido donde las identidades no solo se ven violentadas, sino que se aprecian como sujetos de exclusión: sujetos incompletos, rotos, o sujetos que restan. Por lo tanto, es también en estos campos donde se registra y legitima el déficit cognitivo y social de los sordos, dentro de la marginalidad e invisibilidad cultural en el sistema de desigualdades.

Aspectos Jurídicos y Cívicos

Desde hace más de un siglo, la trayectoria de la construcción del derecho sobre los sordos se canaliza a través de la negación de su existencia como sujetos políticos y sociales. Aunque los sordos forman parte de la coexistencia regulada por la justicia, y la misma contiene normas expresas referidas a las personas sordas, en muchos casos, lejos de beneficiarlos los afectan notablemente. Por ejemplo, analizaremos la ley 340, en su artículo 54 del Código Civil Argentino establece: “Tienen incapacidad absoluta: 1ro. Las personas por nacer; 2do. Los menores impúberes; 3ro. Los dementes; 4to. Los sordomudos que no saben darse a entender por escrito” En este caso se concibe a las personas sordas desde un principio de inhabilitación en el ejercicio de sus derechos y la

ausencia del reconocimiento a su necesidad de desarrollar su autonomía. En su artículo 57 inciso 3, establece la representación de estos a cargo de un curador, los artículos 153 al 158 comparan la figura del sordo con la de demente, en el artículo 469 los considera incapaces de poseer bienes, los artículos desde el 1060 al 3078 les prohíbe realizar contratos, solicitar préstamos, dejar o recibir herencias, y otros.

Es decir, que desde la visión jurídica, la población sorda queda relegada de disfrutar de una ciudadanía plena, considerándolos a estos sujetos de asistencia y no sujetos con derecho a la autonomía y a la libre decisión. Esta concepción jurídica de la población sorda, genera una compleja limitante no solo en materias de derechos fundamentales (Como en el acceso a la vivienda, educación y el trabajo), sino también civiles (aspectos relacionados con su capacidad para casarse, reconocer a sus hijos, divorciarse, dejar un testamento, heredar, votar, realizar denuncias o declarar como testigo ante un delito y otros) desconociendo derechos fundamentales y ampliando las brechas de equidad que previamente analizamos.

Con respecto a la legislación relacionada al LSA, recién en 2008 la República Argentina ratificó la Convención Internacional de los Derechos de las Personas con Discapacidad, mediante la Ley N.º 26.378, donde se acepta y reconoce la utilización de la lengua de señas y la identidad cultural y lingüística de las personas sordas. En su artículo 2º establece que “Se entenderá por lenguaje tanto el lenguaje oral como la lengua de señas y otras formas de comunicación no verbal” e insta a los Estados parte a promover servicios y políticas públicas que tengan como objetivo asegurar la accesibilidad. En el artículo 24, punto 3, inciso b, referido a la educación, compromete a los Estados parte a "Facilitar el aprendizaje de la lengua de señas y la promoción de la identidad lingüística de las personas sordas". El artículo 30, punto 4 establece: “Las personas con discapacidad tendrán derecho, en igualdad de condiciones con las demás, al reconocimiento y el apoyo de su identidad cultural y lingüística específica, incluidas la lengua de señas y la cultura de los sordos.”

Como podemos observar en materia jurídica se pueden notar diversos desequilibrios y contradicciones, por lo tanto, queda en evidencia la compleja dualidad ante el entendimiento de los derechos fundamentales, y la necesidad de desarrollar

legislaciones que protejan íntegramente a esta población en su identidad. Con respecto a la evolución en materia de derechos (y sus garantías), Sánchez (2012) se pregunta:

¿Cómo podría formarse de manera adecuada en sus derechos a una población que como dijimos, en su inmensa mayoría no dispone de las herramientas intelectuales necesarias? La población sorda, por su limitado acceso a la educación no está en capacidad de conocer sus derechos ni de comprender la situación social de su comunidad. Para eso habría que dotar a las instituciones correspondientes la capacidad de hablarles a los sordos en su lengua natural. Porque en el momento actual nadie les habla en su lengua. En estas condiciones, ¿Cómo podría la población sorda tener conocimientos de las cosas que precisamente los afectan de manera prioritaria?

Accesibilidad

- Medios de comunicación:

El artículo 66 de la ley de servicios de comunicación audiovisual prevee regulaciones que tienen que ver con la accesibilidad en la población con discapacidad, mencionando particularmente al LSA:

Accesibilidad. Las emisiones de televisión abierta, la señal local de producción propia en los sistemas por suscripción y los programas informativos, educativos, culturales y de interés general de producción nacional, deben incorporar medios de comunicación visual adicional en el que se utilice subtítulo oculto (*closed caption*), lenguaje de señas y audio descripción, para la recepción por personas con discapacidades sensoriales, adultos mayores y otras personas que puedan tener dificultades para acceder a los contenidos

Sin embargo, la implementación de estas regulaciones no se ven garantizadas en la cotidianeidad, ya que del pequeño porcentaje de canales que utiliza traducciones en LSA, solo lo implementa en algunos programas.

- Internet:

Si bien internet es un espacio diverso, existen organizaciones que desarrollan protocolos, recomendaciones y estándares que buscan asegurar el crecimiento de internet a largo plazo, un ejemplo de ello es la *World Wide Web Consortium (W3C)* que se considera la principal organización de estándares en internet. Un subnúcleo de esta organización es la *W3C Web Accessibility Initiative (WAI)* que se encarga de la estandarización de internet en cuestiones de accesibilidad, desarrollando las *Web Content Accesibility Guidelines (WCAG)* que son un conjunto de directrices que especifican cómo crear contenido web accesible, orientado a personas con discapacidades, con la intención de favorecer el acceso a la audiencia más extensa posible. La versión actual de estas directrices es la WCAG 2.0, publicada en diciembre de 2008 y convertida en estándar ISO/IEC 40500:2012 en octubre de 2012. Estas normas establecen el nivel de accesibilidad de una web en tres distintas categorías, las cuales se denominan A, AA, y AAA, y corresponden a criterios mínimos de accesibilidad, extendidos y accesibilidad máxima. En el apartado 1.2.6 (W3.org, 2012) se menciona la importancia de contar con las interpretaciones del contenido audiovisual, pero debido a la dificultad logística y los costos que esto implica, se considera un requisito fundamental solo para lograr la certificación de calidad más alta (nivel AAA), siendo los subtítulos los únicos requisitos para las certificaciones de menor rango. Por lo tanto, en la actualidad las traducciones en LSA significan una variable de implementación no masificada, derivando en una internet poco accesible con la población sorda.

Educación

Con respecto a la educación de la población sorda y su relación con el LSA, repasaremos algunos hitos históricos explicados por Veinberg (1996) que nos ayudarán a construir un contexto desde donde analizaremos la educación vigente.

- 1857: Creación de la primera escuela para sordos en Buenos Aires. La segunda en Latinoamérica.
- 1871: La escuela de sordos cierra sus puertas debido a la fiebre amarilla.
- 1878: El congreso universal de París concluye el reconocimiento de la mímica natural para la educación de la población sorda.
- 1880: El congreso universal de Milán ignora la lengua de señas señalando al método oral como único método de enseñanza, por instrucción confesional y religiosa.
- 1882: En el Congreso pedagógico de Buenos Aires, el Dr. Antonio Terry (diputado, senador, ministro plenipotenciario de Chile, jurisconsulto) propone la reapertura del colegio de sordos tomando de modelo el Real instituto de Milán, con su pedagogía oralista.
- 1883: Se prohíben las lenguas de señas en todos los establecimientos educativos.
- 1885: Se firma la Ley 1662 por la cual se crea el Instituto Nacional de Sordomudos, sólo para varones.
- 1901: Se crea el Instituto Nacional de Niñas Sordomudas.

El hecho de que los sordos de distinto sexo se mantuvieran aislados, influyó poderosamente en el proceso de desarrollo de la LSA, ya que al estar formalmente prohibida su práctica en los establecimientos educativos, los alumnos practicaban el LSA para comunicarse en secreto entre ellos por las noches y durante los recesos, generando en ese entonces muchas de las señas hoy en vigencia. Siendo Terry hijo ya egresado del instituto Nacional de Sordomudos y afectado por la necesidad de comunicarse en su lengua nativa (LSA), fundó la Asociación de Sordos Mudos de Buenos Aires en 1912, que -además de contribuir considerablemente en la construcción del LSA- generó un espacio propio donde los sordos se sintieran cómodos señando. (Veinberg, 1996)

Es señalando este contexto que Skilar (2003) afirma:

No existió una razón suficiente en el pasado, y no existe razón alguna en la época actual que nos obligue a mantener artificialmente una separación de los sordos, respecto de todos los demás seres humanos, en el discurso y en la práctica educativa. Esa separación, esa frontera, constituye la primera discriminación dentro de un intrincado conjunto de discriminaciones hacia los sordos: la de impedir que la educación de los sordos discuta sus quehaceres educativamente; la de tener, como consecuencia, que refugiarse y avergonzarse como si se tratara de un tema menor, inferior, irrelevante, que aborda a sujetos menores, inferiores, irrelevantes. El hecho que la educación de sordos esté excluida del debate educativo, es la primera y más importante discriminación sobre la cual, después, se hilvanan sutilmente todas las demás discriminaciones -por ejemplo, las de índole civil, legal, laboral, cultural, entre otros-.

Skilar et. al (2013), al analizar la educación actual de la población sorda, resalta que uno de los pilares fundamentales de la deficiencia en el modelo educativo se perpetúa a través de la concepción reduccionista del tipo sujeto deficiente/escuela de bajas expectativas. Y afirma: “Esto elude un debate educativo profundo y así induce inevitablemente a las bajas expectativas pedagógicas; los fracasos escolares, entonces, son fácilmente atribuidos a las supuestas culpas naturales de los propios deficientes.”

Justificación

Por lo expuesto anteriormente, no podemos evitar notar una necesidad tangible de una solución que permita un acceso igualitario de las personas sordas no bilingües a la comunicación transversal desde un enfoque respetuoso no solo con su identidad sino también con su idioma de origen. Sin embargo, este proyecto aspira a no limitarse solo al campo digital: pretende, a través de su estructura, ser una solución inmersa en la cotidianidad a través de sus diversas implementaciones.

En el ámbito digital facilita el acceso de las personas sordas a todo tipo de contenido que no contaban previamente con traducciones, incluyendo también otros medios tradicionales como televisión o radio. También les permite acceder a charlas, videoconferencias y clases virtuales, estando estas últimas en auge por la pandemia del COVID-19.

En su funcionamiento fuera del entorno digital permite ilimitados usos: captar audio mediante un celular y poder entender a través de la traducción en pantalla posibilita, por ejemplo, poder recibir un parte médico.

Con respecto a sus usos en materia educativa, significa una ampliación del alcance de los materiales pedagógicos y una expansión significativa de los canales de comunicación con los alumnos, ya que facilita una disponibilidad continua a muy bajo costo, impulsando la promoción de las alfabetizaciones múltiples, la inclusión pedagógica y el respeto de su identidad individual, facilitando la incorporación de los contenidos a través de la comunicación en su lengua natural.

Es importante resaltar que por su estructura y por el hecho de ser una herramienta de software libre, no existen costos de licencias, permitiendo a la comunidad perfeccionar la herramienta, desarrollar nuevas implementaciones, introducir nuevas señas o lenguajes de otros países, impactando en todos los clientes que dependan de la aplicación de manera síncrona y actualizada.

El impacto sobre estas variables que generan el acceso a la información en su manera más íntegra, favorecerá la autonomía individual, significando un punto de inflexión hacia el desarrollo individual y las garantías de sus derechos humanos fundamentales.

Objetivo General del Proyecto

Desarrollar un servicio multiplataforma de conversión de audio y texto a videos y animaciones que representen la traducción del audio a Lenguas de Señas Argentinas en tiempo real.

Objetivos Específicos del Proyecto

- Lograr un análisis integral de la problemática para el desarrollo de los requerimientos.
- Comprender las bases lingüísticas y estructurales de la Lengua de Señas Argentinas
- Diseñar una solución que integre las variables de accesibilidad establecidas por las directrices de accesibilidad WCAG en su certificación máxima.
- Desarrollar una plataforma de traducción Español-Lengua de señas Argentinas.
- Generar documentación integral del sistema incluyendo guías básicas de uso y ejemplos de implementación.

Marco Teórico Referencial

Dominio del problema

Sobre el lenguaje de señas y su representación lingüística

Según Roso (2013) Las lenguas de señas son parte del espectro de la comunicación humana, siendo “Sistemas verbales que se diferencian de las lenguas orales solo por el hecho de manejar una materialidad acústica y no sonora”.

Así como las lenguas orales, el lenguaje de señas se conformó (y se conforma, como elemento cultural elástico) a través de la espontaneidad de la interacción social. Es

por esta razón que se estructura no universalmente, es decir, que cada país o cada región tiene su propia lengua de señas.

Peluso (2019) nos ilustra las formas constitutivas básicas del lenguaje de señas, siendo estas:

- a) La configuración de la mano -la forma que adopta la mano (por ejemplo, puño cerrado o palma abierta, entre otros).
- b) El movimiento.
- c) La orientación de la mano (por ejemplo, palma arriba, palma abajo).
- d) La ubicación - el lugar en el espacio donde se realiza la seña.
- e) La expresión facial.

Si bien es posible representar la lengua de señas en forma de imágenes (utilizado en los casos de la enseñanza del alfabeto), es una práctica que Köppel, Suchodolski y Zappala (2011) desaconsejan:

Para la producción orientada a la LSA se sugiere trabajar con la imagen en movimiento (videos), ya que la lengua de señas se compone también por rasgos no manuales de cara, ojos, cabeza y cuerpo, que al igual que las diferencias en la duración, tamaño y ritmo de las señas pueden cumplir funciones lingüísticas o emotivas.

Tecnologías adaptativas - Diseño Universal

Köppel, Suchodolski y Zappala, definen a las tecnologías de apoyo de la siguiente manera: “Son recursos para superar las barreras de acceso a las tecnologías digitales, que producen un impacto positivo en la mejora de la calidad de vida de las personas con discapacidad”. Según lo establecido por el programa de inclusión en TICs del Ministerio de Educación de la Nación (2011): “Para que la utilización de nuevas tecnologías se constituya como un elemento promotor de la educación inclusiva, deben considerarse algunos criterios propuestos desde el ideario del diseño universal, que tiene en cuenta las necesidades de todos los posibles usuarios” refiriéndose al diseño de soluciones

informáticas tanto adaptativas como de uso pedagógico, las cuales especifica que deben cumplir, preferentemente, con tres características principales:

Invisibilidad: Para que la tecnología pase inadvertida en el entorno.

Adaptabilidad: Que responda a las necesidades del individuo.

Ubicuidad: Que pueda utilizarse en cada ámbito donde el usuario se desenvuelva.

Estas máximas fueron propuestas con el objetivo que las tecnologías de apoyo en ámbitos pedagógicos promuevan un uso facilitador y no impliquen nuevas barreras de integración.

TICs

A continuación se detalla brevemente la descripción las tecnologías utilizadas en este proyecto

JavaScript

Es un lenguaje de programación que se utiliza principalmente para crear páginas web capaces de interactuar con el usuario. Desde el punto de vista técnico, Javascript es un lenguaje interpretado, eso significa que las instrucciones son analizadas en secuencia por el intérprete de Javascript del navegador web, de manera que la ejecución es inmediata a la interpretación. Gauchat (2012)

Typescript

Es una extensión de JavaScript que permite un desarrollo más sencillo de aplicaciones a gran escala. Mientras que cada aplicación de JavaScript es una de TypeScript, TypeScript ofrece un mejor sistema de módulos, clases, interfaces y una gran oferta de sistema de tipos graduales. (Bierman, 2014)

NodeJS

Es una implementación del motor de ejecución de javascript v8 que permite la creación de aplicaciones de networking enfocadas a la alta performance. Ideado como un entorno de ejecución de JavaScript orientado a eventos asíncronos, Node.js está diseñado para crear aplicaciones network escalables. (Nodejs.org, 2021)

SocketIO

“Es una librería que habilita las comunicaciones en tiempo real, bidireccionales y basadas en eventos entre el navegador y el servidor. Es multiplataforma, pudiendo implementarse en cualquier navegador o dispositivo”. (socket.io, 2021)

GoogleSpeechToText

“Convierte voz en texto de forma precisa con una API basada en las tecnologías de IA de Google.” (google.com/speech-to-text, 2021)

ThreeJS

Es una librería de renderizado tridimensional basada en los estándares de WebGL. Parisi (2012) define a WebGL como el estándar para los gráficos 3D en la web, con la que los desarrolladores pueden explotar al máximo la capacidad de renderizado de la computadora, utilizando solo Javascript, un navegador y un *stack* web estándar.

Webrtc

Con esta herramienta puede agregar capacidades de comunicación en tiempo real a su aplicación que funciona sobre un estándar abierto. Admite video, voz y datos genéricos que se envían entre pares, lo que permite a los desarrolladores crear potentes soluciones de comunicación de voz y video. La tecnología está disponible en todos los navegadores modernos, así como en clientes nativos para las principales plataformas. (webrtc.org, 2021)

AutoML

“Permite que los desarrolladores con experiencia limitada en el aprendizaje automático entrenen modelos de alta calidad y específicos para las necesidades de sus negocios. Compila un modelo de aprendizaje automático personalizado en minutos.” (google.com/automl, 2012)

Unreal Engine

Es una suite completa de herramientas de Desarrollo de renderizado para cualquiera que trabaje con tecnología en tiempo real. Desde el diseño de vistas y las experiencias cinemáticas al Desarrollo de juegos *high-quality* para PC, consolas, dispositivos móviles y realidad virtual. (unrealengine.com, 2021)

GitHub

“Es una plataforma de desarrollo colaborativo de software para alojar proyectos usando el sistema de control de versiones Git. El código se almacena de forma pública, aunque también se puede de forma privada, creando una cuenta de pago.” (Github.com, 2021)

Github Actions

Automatiza, personaliza y ejecuta tus flujos de trabajo de desarrollo de software directamente en tu repositorio con GitHub Actions. Puedes descubrir, crear y compartir acciones para realizar cualquier trabajo que quieras, incluido CI/CD, y combinar acciones en un flujo de trabajo completamente personalizado. (github.com, 2021)

Node Package Module

“El administrador de paquetes de Node.js (npm) es el administrador de paquetes predeterminado y más popular del ecosistema de Node.js, y se utiliza principalmente para instalar y administrar módulos externos de proyectos de Node.js” (Nodejs.org, 2021)

Google Memory Store

“Almacenamiento en memoria escalable, seguro y de alta disponibilidad para Redis y Memcached.” (cloud.google.com, 2021)

Google Cloud Firestore

Es un servicio de almacenamiento de objetos potente, simple y rentable construido para el escalamiento de Google. Los SDK de Firebase para Cloud Storage agregan la seguridad de Google a las operaciones de carga y descarga de archivos de las apps de Firebase, sin importar la calidad de la red. (cloud.google.com, 2021)

Kubernetes

“Es una plataforma portable y extensible de código abierto para administrar cargas de trabajo y servicios. Kubernetes facilita la automatización y la configuración declarativa.” (Kubernetes.io, 2021)

Docker

“Permite usar contenedores como máquinas virtuales extremadamente livianas y modulares. Además, provee flexibilidad con estos contenedores: puede crearlos, implementarlos, copiarlos y moverlos de un entorno a otro, lo cual le permite optimizar aplicaciones en la nube.” (redhat.com, 2021)

Competencia

Actualmente, podemos encontrar en el mercado algunas aplicaciones que poseen un objetivo similar a la que aquí se propone, sin embargo, existen diferencias considerables a la hora de la implementación. En el caso de las aplicaciones relevadas, todas ellas cuentan la funcionalidad de traducir texto a lengua de señas, pero ninguna con la funcionalidad de incorporar traducciones de audio. En los tres casos encontrados, se trabaja con español-LSE (Lenguas de señas españolas) que, como mencionábamos anteriormente, no es compatible con el sistema argentino.

Por otro lado, podemos señalar que solo una de ellas es multiplataforma, las otras dos se reducen a aplicaciones incrustadas en una web particular, siendo estas últimas de uso gratuito. La aplicación restante es multiplataforma, pero el desarrollador no especifica información concreta sobre si la aplicación genera las traducciones en tiempo real.

Tabla 1 - Comparación entre distintas soluciones

	Texto a señas	Audio a Señas	Multiplataforma	Multilinguaje	Gratuito	Idioma
TexToSign	✓	x	✓	x	x	LSE
Hetah	✓	x	x	x	✓	LSE
Singslator	✓	x	x	x	✓	LSE

Elaboración propia – Fuente: textosign.com, 2021 – Hetah.com, 2021 – singslator.com, 2021

Diseño Metodológico

Herramientas metodológicas

La metodología utilizada para el desarrollo de la plataforma es SCRUM, en palabras de Mariño y Alfonso (2014):

Scrum es un marco de trabajo iterativo e incremental para el desarrollo de proyectos y se estructura en ciclos de trabajo llamados Sprints. Estos son iteraciones de 1 a 4 semanas, y se suceden una detrás de otra. Al comienzo de cada Sprint, el equipo multifuncional selecciona los elementos (requisitos del cliente) de una lista priorizada. Se comprometen a terminar los elementos al final del Sprint. Durante el Sprint no se pueden cambiar los elementos elegidos. Al final del Sprint, el equipo lo revisa con los interesados en el proyecto, y les enseña lo que han construido.

Como método ágil:

- Es un modo de desarrollo adaptable, antes que predictivo.
- Está orientado a las personas, más que a los procesos.

- Emplea el modelo de construcción incremental basado en iteraciones y revisiones.

Herramientas de desarrollo

El principal objetivo de la plataforma es ser una herramienta que acompañe en la comunicación: al ser este un proceso que se produce de forma instantánea, es una necesidad tácita que la traducción acompañe al proceso de forma fluida, por lo tanto, se requiere que todo el stack tecnológico esté pensado y definido a través de la comunicación en tiempo real y la reducción de latencias como prioridad máxima. Es por esta razón que la elección de las tecnologías se realizó priorizando la capacidad de las herramientas de operar los datos en forma de tuberías, es decir, que cada uno de los elementos que transforman la información lo realicen en el menor tiempo posible y lo envíen al siguiente eslabón de transformación de forma transparente.

A continuación, haremos un recorrido a través de las tecnologías que integran el *stack*:

Nodejs y SocketIO nos facilitan la comunicación bidireccional con una latencia menor a un segundo entre los clientes y el servidor, permitiéndonos enviar contenido en formato binario (Como el audio y las acciones o animaciones requeridas por el renderizado) en formatos de codificación estándar.

A través de Webrtc, capturamos los dispositivos multimedia del usuario para luego codificarlos y enviarlos al servidor para su posterior procesamiento.

Google SpeechToText nos facilita el reconocimiento de voz para luego convertirlo a texto con un costo relativamente bajo y una escalabilidad sin competencia.

Con respecto a Automl, este fue elegido por la simplicidad en la creación de modelos matemáticos de inteligencia artificial para la conversión o traducción de textos de un lenguaje a otro, ya que nos provee *datasets* estructurados de forma simple para facilitarnos esta funcionalidad.

Desde el punto de vista del renderizado, necesitamos distintas calidades en función al tipo de salida a producir, por lo tanto, utilizamos dos librerías diferentes: por un lado Threejs fue elegida por ser una librería de renderizado tridimensional de muy bajos requerimientos técnicos (con pocos recursos de procesamiento se puede mostrar

una animación tridimensional), y ante la necesidad de una opción profesional con mayor calidad visual, utilizamos Unreal Engine que actualmente es el motor de renderizado fotorealista en tiempo real mas popular disponible en el mercado.

Recolección de datos

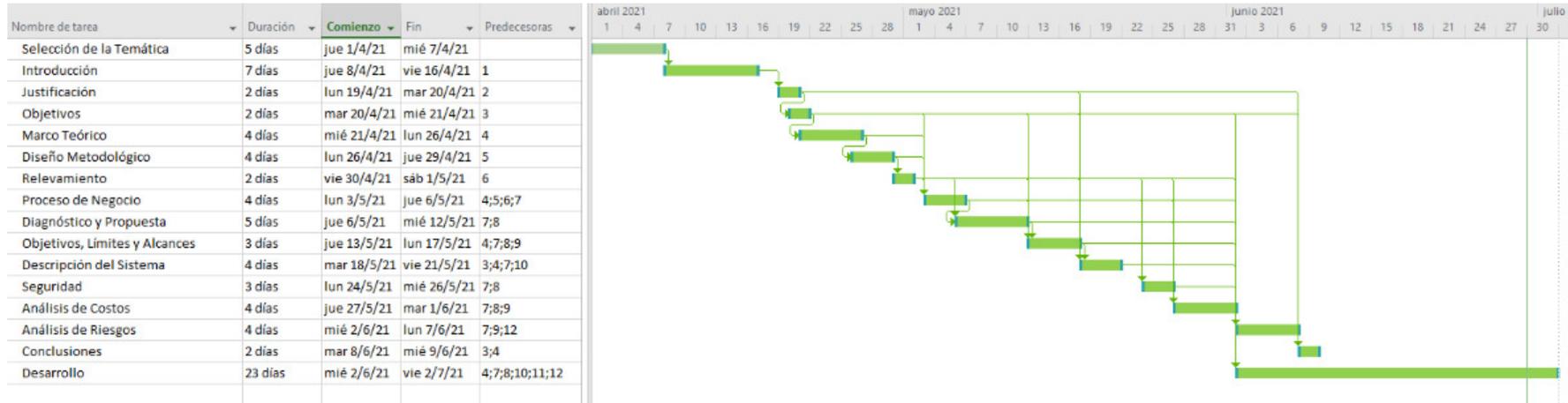
Para la recolección de los datos se emplearon diversas fuentes, desde el punto de vista técnico del lenguaje de señas, se realizaron entrevistas con dos intérpretes profesionales las cuales brindaron el análisis técnico en términos de viabilidad de las traducciones, y realizaron consideraciones lingüísticas y estructurales. Además, ambas profesionales sumarán traducciones durante el desarrollo continuo del proyecto.

Desde el punto de vista pedagógico y ante las necesidades mencionadas previamente en el apartado de tecnologías adaptativas y diseño universal, se realizaron entrevistas y validaciones de diseño con dos profesionales, psicopedagoga y fonoaudióloga, y el documento generado forma parte de las consideraciones claves en la documentación para el desarrollo de las implementaciones.

Planificación del proyecto

A continuación se indica los tiempos de desarrollo del proyecto incluyendo las pruebas.

Imagen 1 – Diagrama de Gantt



Fuente: Elaboración propia.

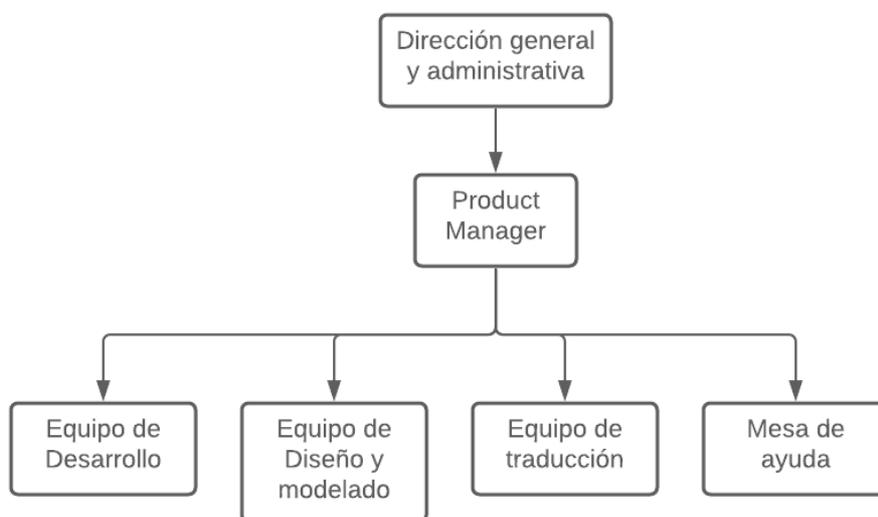
Relevamiento

Ya que el desarrollo de la plataforma es un proyecto independiente de software libre que pretende ser altamente escalable para permitir la futura integración de lenguajes de señas de otros países como también su mantenimiento colaborativo en la agregación de nuevas señas, es un desarrollo que trasciende la incumbencia de las organizaciones existentes y por lo tanto, el relevamiento se realizará sobre una organización modelada.

Relevamiento Estructural

A continuación se incluye el organigrama de la organización modelada dedicada a la traducción de contenido multimedia.

Imagen 2 - Organigrama de la organización modelada



Fuente: elaboración propia.

Por un lado, nos encontramos a la dirección general que toma las decisiones estratégicas relacionadas al producto y al proyecto, y a su vez, gestiona los aspectos financieros de la empresa. Luego se ubica el Product Manager, su rol principal es la

coordinación general de los distintos equipos, diseña, evalúa e implementa las estrategias de gestión y administra el desarrollo del producto en todas sus vertientes.

Los cuatro equipos siguientes consisten, por un lado, en el equipo encargado del desarrollo técnico de la herramienta, el equipo de traducción que cuenta con traductores profesionales de lenguaje de señas, el equipo de diseño y modelado el cual es el encargado de tomar las señas y representarlas en modelos tridimensionales, y la mesa de ayuda encargada de guiar a los clientes con las implementaciones.

Es importante señalar que al ser una herramienta que permite la implementación en diversas soluciones según las necesidades particulares de los desarrolladores cliente, el rol de estos últimos no es neutro, sino que también participan dentro de la configuración de desarrollo de la herramienta, y son actores fundamentales en la toma y la oferta de soporte.

Relevamiento Funcional

En la descripción de los procesos actuales podemos definir tres grupos donde se encuadran los actores principales:

1. Productor de contenido: Productores o creadores de material audiovisual, sea este en vivo o pregrabado, contemplando todos medios de comunicación.
2. Fuente de información: Abarca la comunicación unidireccional o bidireccional en español por parte de una o más personas, contemplando todos los contextos en los que se ejecuta la comunicación hablada (docentes impartiendo clases, llamadas, mensajes de audio en aplicaciones de mensajería, entre otros)
3. Intérpretes: Profesionales intérpretes y traductores LSA-español.
4. Persona Sorda: Se contemplan los casos donde la persona es bilingüe y no-bilingüe.

Nombre del proceso: Generación de contenido multimedia.

Roles: Productor de contenido - Intérprete

Pasos: 1 – El Productor recurre al intérprete LSA

2 – El intérprete se traslada al lugar de grabación

2.1 – Si el contenido es producido en vivo, el intérprete realiza la traducción en simultáneo a la transmisión.

2.2 – El productor transmite el contenido en simultáneo a la traducción.

3 – Si el contenido no es transmitido en vivo, se realiza la grabación de la traducción

3.1 – Se agrega la traducción durante el proceso de edición y se publica.

Nombre del proceso: Interpretación de información general.

Roles: Fuente de información – Persona sorda - Intérprete

Pasos: 1 – La fuente de información genera un mensaje en español

1.2 – Si posee subtítulos y la persona sorda es bilingüe, puede comprender el mensaje.

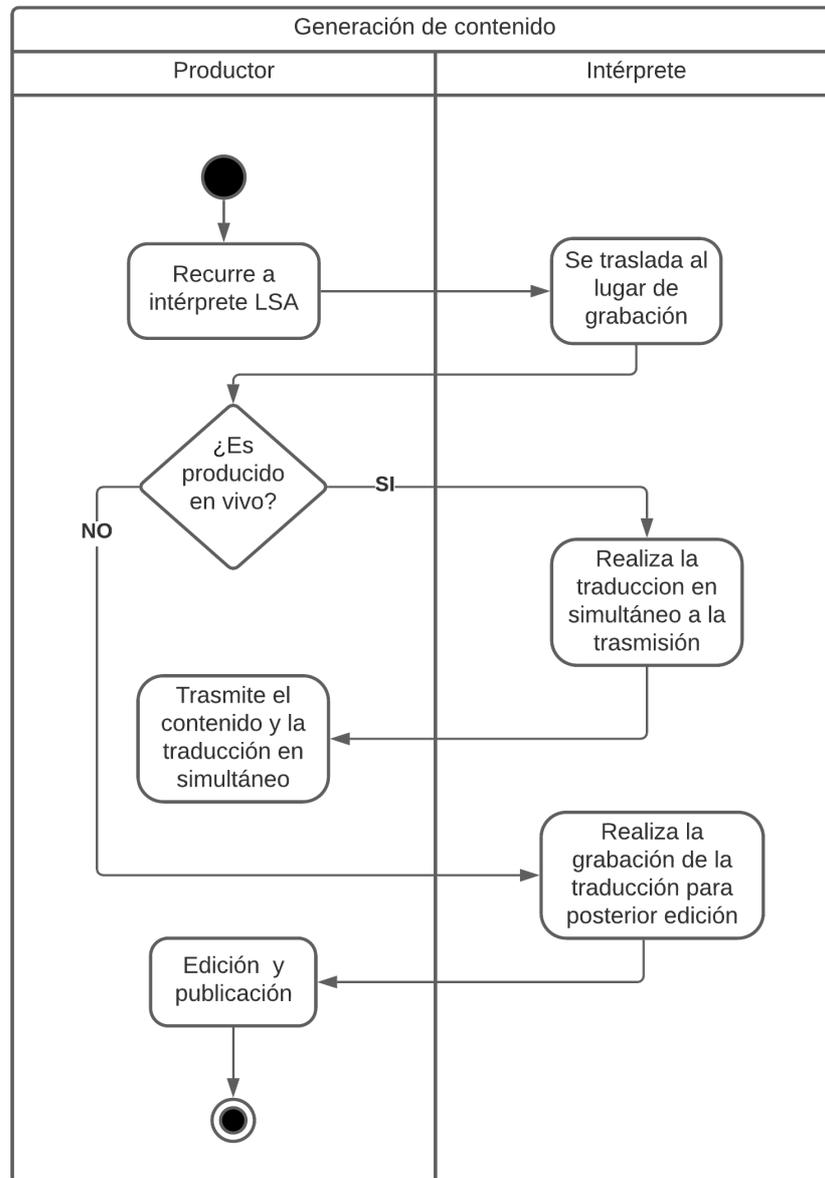
1.3 – Si no posee subtítulos o la persona sorda no es bilingüe, recurre a un intérprete.

1.3.1 - Si la persona sorda puede acceder al intérprete, este realiza la interpretación, por lo tanto comprende el mensaje.

1.3.2 – Si la persona sorda no puede acceder a un intérprete, no puede comprender el mensaje.

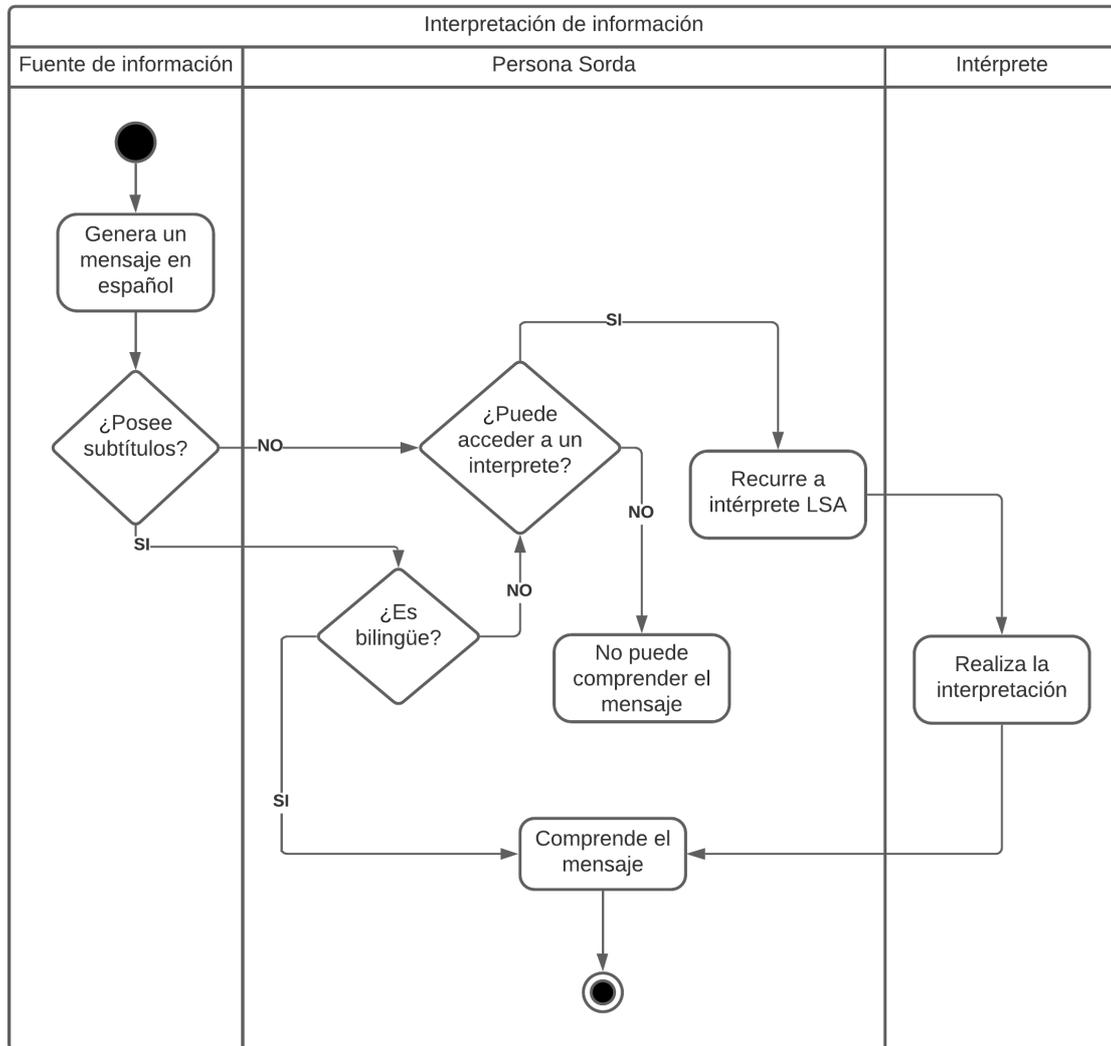
Procesos de negocios

Imagen 3 - Proceso 1: Generación de contenido audiovisual.



Fuente: elaboración propia

Imagen 4 - Proceso 2: Interpretación de información general



Fuente: Elaboración propia

Diagnóstico y Propuesta

Diagnostico

A continuación se detalla el diagnóstico en base a los procesos relevados. Es importante destacar que el inconveniente común está relacionado al acceso a los intérpretes, por lo tanto, cada ítem aquí dispuesto afecta a ambos procesos de forma igualitaria. Por esta razón, se omite referenciar al nombre de los procesos afectados en cada punto, con el objetivo de evitar redundancias.

Aspectos técnicos

Problema: Elevados costos de traducción.

Causas: - El factor humano implicado dentro del proceso de traducción encarece la accesibilidad en términos económicos.

- La producción de contenidos requiere un gran número de intérpretes involucrados.

Como se ha expuesto previamente, el acceso de la población sorda a los intérpretes resulta un factor indispensable para la inclusión en espacios no señantes. En el caso de quienes dependen de intérpretes para la ejecución de tareas de periodos sostenidos (tareas relacionadas al trabajo o al cursado de clases) deben enfrentarse con costos fijos demasiado elevados.

En el ámbito de la producción de contenidos, existen casos en los que se requiere un amplio número de intérpretes dada la extensión del material audiovisual, ya que la tarea de interpretación conlleva un trabajo físico y mental intenso. Estos factores incrementan los costos exponencialmente.

Problema: Planificación logística restrictiva y compleja

Causas: - La interpretación se ve restringida un horario disponible previamente pactado según disponibilidad.

- No es posible asegurar la disponibilidad del intérprete en situaciones de emergencia.

Las traducciones de material audiovisual se ven ligadas a la disponibilidad horaria de los intérpretes y restringidas a su realización en los periodos coordinados. En la mayoría de los casos es difícil contar con intérpretes a demanda para la realización de tareas de forma espontánea no prevista o en caso de emergencias. La disponibilidad de los intérpretes se reduce considerablemente en horarios no convencionales.

Problema: Dificultad en el acceso a intérpretes profesionales.

Causas: - Escasez de profesionales que brinden el servicio.

- Las variables geográficas y de traslado limitan el acceso.

La disponibilidad de los intérpretes en relación a su ubicación geográfica es otro factor que limita el acceso al servicio, siendo la causa principal la escasa oferta de profesionales en relación a la demanda. Esta variable impacta de manera considerable en los casos donde el servicio se demanda desde pequeñas zonas urbanas y/o rurales, donde generalmente no se cuenta con acceso. Como se mencionaba con anterioridad, si bien no existen registros nacionales de intérpretes y su distribución geográfica, al estudiar el caso de la provincia de Mendoza que cuenta con matriculación profesional, son 22 los profesionales matriculados donde su población asciende a 2.086.000 personas (plausibles a demandar parcialmente el servicio), de los cuales el 9 % presenta dificultades auditivas (propensos a necesitar del servicio de forma permanente), eso nos arroja una cantidad de 98.817 personas por cada intérprete profesional. Proyectando estos valores al resto del país, considerando un amplio margen de error y tomando en cuenta la distribución geográfica de la población, es apropiado concluir que en la mayoría de localidades medianas o pequeñas del país no se cuenta con la disponibilidad de profesionales que brinden el servicio.

En relación a lo previamente expresado, podemos señalar limitantes en las situaciones que se requiere de interpretación por el mismo sujeto en distintas zonas geográficas, como por ejemplo, en los casos donde se necesita un traslado interprovincial por motivos médicos.

Problema: Dificultad de acceder a la interpretación en campos especializados.

Causa: - No se cuenta con especializaciones formales.

Como señalamos con anterioridad, no se cuenta con especializaciones formales de LSA en el sector académico. Esto implica que, al requerirse intérpretes que realicen traducciones en campos de especialidad, no se puede garantizar la traducción contextual de términos concretos utilizados en estas áreas específicas.

Aspectos sociales

Problema: Accesibilidad restringida y desigualitaria al contenido.

Causa: - La mayoría de contenidos existentes no poseen traducciones en LSA.

Gran parte del contenido producido, tanto pedagógico o académico, informativo o de entretenimiento, no cuenta con sus respectivas traducciones en LSA y por lo tanto, no es accesible a la población sorda en general y a la población sorda no bilingüe en particular, generando una marcada brecha de desigualdad social.

Problema: Exclusión educativa y laboral.

Causa: - La ausencia de intérpretes impide estrictamente la comunicación.

Dada a la estricta necesidad de la población sorda de contar con intérpretes para lograr un entendimiento efectivo, la ausencia de estos actores genera como consecuencia un impacto negativo en la inclusión efectiva. Desde conferencias online, clases especializadas, seminarios, al entorno laboral, el acceso a estas oportunidades se ve seriamente restringida.

Problema: Internet poco accesible con las personas sordas.

Los costos, la disponibilidad y las variables relacionadas a la logística de los intérpretes limitan la traducción de los contenidos en ambientes digitales. La cantidad existente de idiomas señados complejiza exponencialmente el problema.

Propuesta

Se propuso el desarrollo de una solución que proporciona las herramientas necesarias para tomar una entrada de audio de un dispositivo y genera una animación en tiempo real con la traducción a Lenguas de Señas Argentinas.

Con un principal foco en la escalabilidad y flexibilidad de adaptación a distintos tipos de implementaciones sobre diferentes plataformas y hardware, se planteó una plataforma centrada en un servicio encargado de realizar el procesamiento de la información acompañado de librerías que permiten la comunicación en tiempo real, el manejo de las entradas de sonido y el renderizado de las animaciones y subtítulos.

Libre y colaborativo

Esta plataforma se presentó bajo los fundamentos del software libre, para poder desarrollarse en todo su potencial a través de la construcción comunitaria colaborativa, en beneficio de toda la población. Esta característica brinda la posibilidad de utilizarla sin costo, compartirla, estudiarla, mejorarla y adaptarla a las propias necesidades del usuario de manera libre, impactando en la comunidad de manera positiva. El objetivo es evitar restricciones de acceso sobre poblaciones en situación de vulnerabilidad, pero además mejorar la eficiencia del desarrollo y posibilitar a otras comunidades extender el proyecto para agregar soporte a nuevos lenguajes que necesiten en la región.

La flexibilidad y simplicidad de implementación de la solución permite ser integrada en diferentes herramientas ya existentes, facilitando una rápida adopción por parte de terceros, que nos permite impactar positivamente sobre la problemática en entornos digitales a nivel global.

La arquitectura del proyecto a demanda de las necesidades lingüísticas

Al ser el lenguaje un fenómeno cultural que evoluciona a través del tiempo de acuerdo a su uso y a otros factores sociológicos y antropológicos, es propenso de ser construido y deconstruido a través de la temporalidad. Es por ello que esta solución permite agregar nuevas traducciones a partir de la creación o el auge de las mismas de manera centralizada permitiendo que los cambios impacten en las aplicaciones cliente de forma síncrona y automática, logrando así acompañar a la comunicación de forma paralela a los cambios sociales del lenguaje sin necesidad de una re-implementación de la solución.

Por otro lado, mantener una arquitectura centralizada nos permite reducir latencia y conservar la calidad del servicio al escalar el mismo de forma lineal y distribuir los costos de procesamiento a demanda, en relación a los usuarios activos.

Inclusión en el respeto por las identidades

La traducción automática del español a través de la lengua materna de la población sorda es una variable de integración novedosa en el ámbito digital. Esta, a diferencia de la utilización de subtítulos y *closed captions*, permite una comunicación efectiva sin una necesidad de pedagogización previa, evitando un doble trabajo cognitivo en el caso de los sordos bilingües, y favoreciendo el acceso efectivo en el caso de los no bilingües.

La solución permite la integración transversal en ambientes donde no se cuenta actualmente con intérpretes, de manera espontánea y bajo demanda. Además, esta solución permite eliminar las complicaciones logísticas y los altos costos ligados a la contratación del servicio de intérpretes. En ambientes digitales, favorece la traducción de contenido de forma rápida y fluida, desde diversas fuentes y calidades de salida, de acuerdo a las necesidades particulares del usuario.

Accesibilidad: diseñada sobre los estándares máximos de la WCAG

Esta herramienta está diseñada para significar una solución que sea utilizable por la mayor cantidad de personas, independientemente de sus capacidades personales o las

características técnicas de sus dispositivos. Tomando como punto de partida las pautas del diseño universal para la generación de soluciones adaptativas y pedagógicas (ubicuidad, adaptabilidad, invisibilidad) extendimos estas normas sobre los modelos de renderizado y además contemplamos otras necesidades específicas de los usuarios en las interfaces según los estándares máximos de la WCAG 2.1, significando una herramienta amena con las personas que sufren disminución visual, epilepsia, u otras condiciones asociadas como trastorno del espectro autista y dislexia, entre otros. Por lo tanto, variables como la distingibilidad, legibilidad, predictibilidad, la compatibilidad y el uso consciente de colores y contrastes, fueron requisitos centrales sobre las salidas generadas por la solución.

Objetivo, Límites y Alcance del Prototipo

Objetivo del prototipo

Desarrollar herramientas modulares plausibles de ser implementadas a demanda que permitan la conversión de una entrada de audio a Lenguaje de Señas Argentinas.

Límites

Desde el desarrollador selecciona el tipo de módulo a implementar, hasta la generación de la salida predefinida por este.

Alcances

- Captura de audio de micrófono
- Trasmisión de audio en tiempo real
- Conversión de audio a texto en español
- Conversión de español a LSA
- Renderizado de animación en tiempo real

Descripción Del Sistema

Product Backlog

Tabla 2 - Product Backlog

ID	Historia de usuario	Prioridad	Puntos	Dependencias
S01	Comunicación bidireccional entre servidor y cliente	3	2	-
S02	Estructura de complementos	2	2	S01
S03	Traducción de audio a texto y señas	2	3	S04
S04	Captura de audio vía web API	2	1	S01
S05	Renderizado de subtítulos	1	1	S02 – S03
S06	Almacenamiento de animaciones	1	2	-
S07	Renderizado de animaciones	1	3	S01 – S06 – S03
S08	Implementación de ejemplo	1	3	S07 – S05 – S04

Historias de usuario

Tabla 3 – Historia de usuario S01

ID	S01	Nombre	Comunicación bidireccional entre servidor y cliente
Descripción			Como desarrollador quiero disponer de una librería que me abstraiga la conexión de mis clientes hacia el servidor del proyecto para enviar y recibir las informaciones necesarias para realizar el proceso de traducción.
Criterios de aceptación			<ul style="list-style-type: none"> Dado un problema de conectividad entre el servidor y la librería cliente cuando el cliente esté conectado, entonces, la librería generará un evento de error correspondiente. Dado un cliente, cuando este se conecte al servidor, entonces, el servidor generará un identificador único por cliente. Dada la librería cliente, cuando esté conectada al servidor, entonces, debe permitir el envío y recepción de objetos en formato JSON. Dada la librería cliente, cuando esté conectada al servidor, entonces, debe permitir el envío y recepción de objetos binarios sin una estructura definida.

Prioridad	Alta	Puntos de historia	8
-----------	------	--------------------	---

Tabla 4 - Historia de usuario S02

ID	S02	Nombre	Estructura de complementos
Descripción	Como desarrollador, quiero una librería base que permita ser extendida y abstraiga al resto de los complementos de la conexión contra el servidor cuando realice la implementación, para reducir los tiempos de trabajo.		
Criterios de aceptación	<ul style="list-style-type: none"> • Dada la librería cliente, cuando esta sea importada, entonces, debe proveer una clase base para ser extendida a fin de poder crear nuevas librerías (plugins) que extienden la funcionalidad del proyecto. • Dada la librería cliente, cuando la clase base sea extendida, entonces, debe heredar a la clase hija los métodos y tipos de datos requeridos para el envío y recepción de información contra el servidor. • Dada la librería cliente, cuando esta sea instanciada, entonces, debe recibir como parámetros la lista de complementos requeridos por la implementación y desencadenar una inicialización de los complementos proporcionados. 		
Prioridad	Alta	Puntos de historia	5

Tabla 5 - Historia de usuario S03

ID	S03	Nombre	Traducción de audio a texto y señas
Descripción	Como desarrollador, quiero que el servidor tome una fuente de audio, la traduzca a texto, y retorne la animación de la representación de una seña en LSA, para realizar la traducción.		
Criterios de aceptación	<ul style="list-style-type: none"> • Dada la librería cliente, cuando envíe una transmisión de audio al servidor, entonces, no debe negociar la frecuencia de muestreo y fijar este valor a 16000 muestras por segundo. • Dado el servidor, cuando detecte una palabra o frase en base al audio ingresado, entonces, debe emitir un evento hacia el cliente. • Dado el servidor, cuando encuentre una animación de LSA en la base de datos en base a la palabra o frase detectada, entonces, debe emitir un evento hacia el cliente. 		
Prioridad	Media	Puntos de historia	13

Tabla 6 - Historia de usuario S04

ID	S04	Nombre	Captura de audio vía web API
Descripción	Como desarrollador, quiero un complemento que tome el audio proveniente del micrófono del dispositivo del cliente y lo envíe al servidor, para reducir los costos de implementación al no tener que realizarlo internamente.		
Criterios de aceptación	<ul style="list-style-type: none"> • Dado el complemento, cuando necesite acceder a algún hardware de sonido del dispositivo, entonces, debe solicitar los permisos correspondientes al navegador. • Dado el complemento, cuando capture audio de un dispositivo, entonces, debe codificarlo en formato wav con una frecuencia de muestreo de 16000 muestras por segundo. • Dado el complemento, cuando finalice la codificación, entonces, debe utilizar los métodos establecidos por la librería cliente a fin de generar una transmisión del audio hacia el servidor. 		
Prioridad	Media	Puntos de historia	3

Tabla 7 - Historia de usuario S05

ID	S05	Nombre	Renderizado de subtítulos
Descripción	Como desarrollador, quiero un complemento que renderice sobre el DOM del navegador los textos emitidos por el servidor, para reducir los costos de implementación al no tener que desarrollarlo internamente.		
Criterios de aceptación	<ul style="list-style-type: none"> • Dado el complemento, cuando este sea inicializado, entonces, debe aceptar un elemento del DOM donde realizará la inyección. • Dado el complemento, cuando este sea inicializado, entonces, debe suscribirse a los métodos establecidos por la librería cliente a fin de poder recibir los eventos de subtítulos que emite el servidor. • Dado el complemento, cuando reciba los eventos de subtítulo del servidor, entonces, debe renderizarlos en un formato de lista HTML sobre el elemento indicado en la inicialización. 		
Prioridad	Baja	Puntos de historia	1

Tabla 8 - Historia de usuario S06

ID	S06	Nombre	Almacenamiento de animaciones
Descripción	Como organización, quiero disponer de una estructura de datos y el acceso a los mismos a fin de poder lograr un diccionario con un mapeo uno a uno entre palabras del idioma español y las animaciones que representan el LSA, para centralizar el proceso de traducción.		
Criterios de aceptación	<ul style="list-style-type: none"> • Dado un audio ingresado, cuando se detecte una palabra en español, entonces, debe permitir realizar una búsqueda heurística sobre el diccionario de palabras almacenado. • Dado un objeto de animación, cuando se almacene en la base de datos, entonces debe realizarse en formato blob junto con una clave única que lo identifique. 		
Prioridad	Baja	Puntos de historia	3

Tabla 9 - Historia de usuario S07

ID	S07	Nombre	Renderizado de animaciones 3D
Descripción	Como desarrollador, quiero un complemento que renderice un modelo 3D y que represente un intérprete el cual ejecute las animaciones emitidas por el servidor, para reducir el costo de implementación al no tener que desarrollarlo internamente.		
Criterios de aceptación	<ul style="list-style-type: none"> • Dado el complemento, cuando este sea inicializado, entonces, debe aceptar un elemento del DOM donde realizará la inyección del Canvas para renderizar el contexto tridimensional de WebGL. • Dado el complemento, cuando este sea inicializado, entonces, debe permitir utilizar diferentes modelos 3D usando el mismo sistema de control (rig). • Dado el complemento, cuando este sea inicializado, entonces, debe permitir configurar parámetros de calidad del renderizado a fin de lograr un frame rate adecuado según las necesidades del cliente y la disponibilidad de la potencia del hardware utilizado. • Dado el complemento, cuando este no disponga de animaciones a ejecutar, entonces, debe ejecutar en loop una animación de reposo. • Dado el complemento, cuando este reciba animaciones desde el servidor, entonces, debe almacenar estas en un buffer. • Dado el complemento, cuando la cantidad de animaciones a ejecutar en el buffer cambien anormalmente por un aumento en la velocidad y/o cantidad de eventos recibidos del servidor, entonces, debe aumentar la velocidad de reproducción de las animaciones a fin de reducir la latencia entre el audio proporcionado por el cliente y la reproducción de las señas de LSA 		

Prioridad	Baja	Puntos de historia	13
-----------	------	--------------------	----

Tabla 10 - Historia de usuario S08

ID	S08	Nombre	Implementación de ejemplo
Descripción	Como desarrollador, quiero una aplicación de ejemplo implementado las librerías existentes del proyecto donde se tome el audio del micrófono del cliente y muestre las animaciones de LSA junto con un subtítulo en español del mismo, para tener una referencia práctica sobre como desarrollar la implementación.		
Criterios de aceptación	<ul style="list-style-type: none"> • Dada la aplicación de ejemplo, cuando sea ejecutada en un navegador, entonces, debe funcionar en Google Chrome. • Dada la aplicación de ejemplo, cuando el usuario la inicialice en distintos dispositivos con distintas resoluciones, entonces, la interfaz debe renderizar correctamente todos los elementos disponibles. • Dado el código de la aplicación de ejemplo, cuando se ejecute, entonces, debe proveer únicamente los elementos de interfaz y los estilos de los mismos evitando agregar funcionalidad por sobre las provistas por las librerías y complementos del proyecto. 		
Prioridad	Baja	Puntos de historia	5

Sprint Backlog

A continuación se detalla la planificación del primer sprint. Cada sprint tiene la duración de dos semanas.

Tabla 11 - Sprint Backlog

Sprint	Historia de usuario	ID	Tareas	Prioridad	Tiempo Estimado	Estado
1	S01 Comunicación bidireccional entre servidor y cliente	T01	Implementar estructura del repositorio con Lerna	Alta	8 horas	Hecho
		T02	Crear estructura del paquete del servidor.	Alta	2 horas	Hecho
		T03	Crear estructura del paquete del cliente.	Alta	2 horas	Hecho

		T04	Definir configuración de TypeScript (tsconfig.json).	Alta	4 horas	Hecho	
		T05	Definir configuración de ESLint (eslintrc.js).	Alta	4 horas	Hecho	
		T06	Implementar Servicios Http y WebSocket en el servidor.	Alta	4 horas	Hecho	
		T07	Implementar Clase de Cliente y establecer la conectividad contra el servidor vía WebSocket.	Alta	4 horas	Hecho	
		T08	Generar un UUID cuando el cliente se conecte al servidor.	Media	2 horas	Hecho	
		T09	Implementar métodos de envío y recepción de objetos en formato JSON.	Media	4 horas	Hecho	
		T10	Implementar librería de manejo de transmisiones de binarios y los métodos correspondientes.	Media	16 horas	Hecho	
		S01 Comunicación bidireccional entre servidor y cliente	T11	Implementar métodos y configuraciones de depuración en el cliente	Baja	2 horas	Hecho
			T12	Implementar eventos de error, conexión y desconexión en el cliente	Baja	2 horas	Hecho
		1	S02 - Estructura de complementos	T13	Definir clase base para ser extendida por los complementos	Alta	2 horas
T14	Definir métodos de inicialización y gestión de componentes			Alta	2 horas	Hecho	

		T15	Implementar métodos de depuración de los componentes	Baja	2 horas	Hecho
		T16	Implementar métodos para el envío de binarios	Alta	2 horas	Hecho
		T17	Implementar métodos para recepción de textos	Alta	2 horas	Hecho
	S04 - Captura de audio vía web API		Crear estructura del paquete del complemento de grabación	Alta	2 horas	Hecho
			Implementar permisos de acceso de los dispositivos multimedia a través de getUserMedia	Alta	4 horas	Hecho
			Implementar captura de audio	Alta	4 horas	Hecho

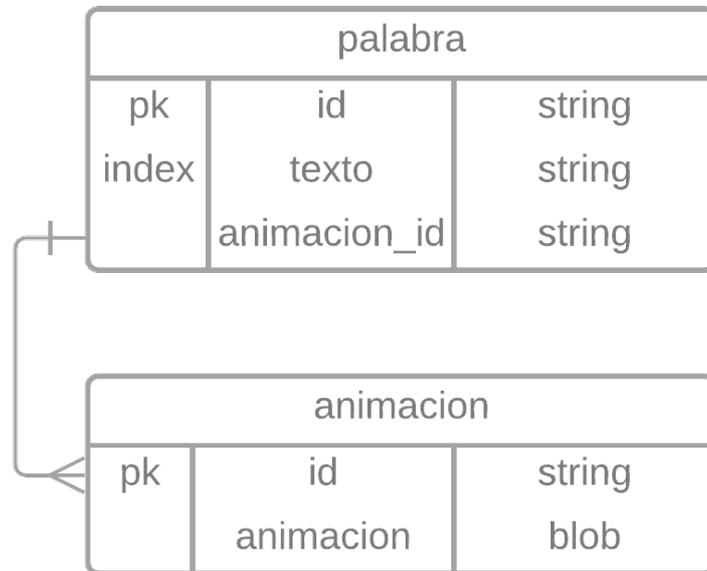
Estructura de datos

Debido al volumen de datos requeridos para realizar un modelo de aprendizaje automático supervisado por intérpretes de LSA, se opta por utilizar una estructura de datos de clave-valor a fin de poder realizar un mapeo uno a uno entre las palabras del lenguaje español de origen y las animaciones correspondientes al LSA.

Por otro lado, se utiliza otra colección para almacenar de forma centralizada los objetos binarios (blob) que representan los movimientos, en una base temporal de los nodos de control disponibles (*rig*) de los modelos tridimensionales, los cuales representan una animación.

Diagrama de entidad-relación

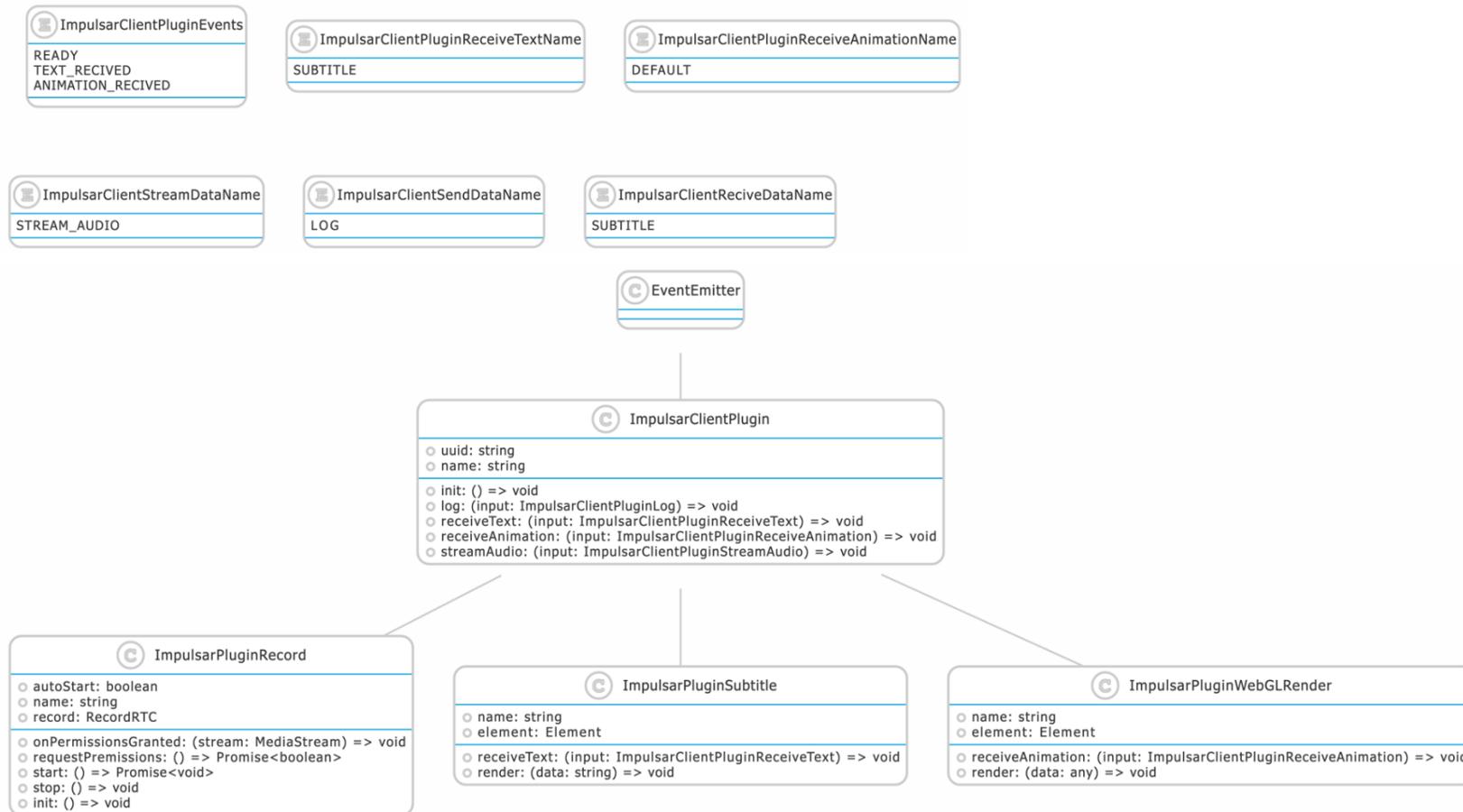
Imagen 5 - Diagrama entidad-relación



Fuente: Elaboración propia.

Diagrama de clases

Imagen 6 - Diagrama de clases



ImpulsarClient

- o debug: boolean
 - o plugins: ImpulsarClientPlugin[]
 - o host: string
 - o autoConnect: boolean
 - o #io: any
 - o #ioStream: SocketStream
 - o connected: boolean
-
- o #log: (out: string) => void
 - o #pluginsReady: ({ plugin }: { plugin: ImpulsarClientPlugin; }) => void
 - o #pluginsLog: ({ plugin, input }: { plugin: ImpulsarClientPlugin; input: ImpulsarClientPluginLog; }) => void
 - o #pluginsTextRecived: ({ plugin, input }: { plugin: ImpulsarClientPlugin; input: ImpulsarClientPluginReceiveText; }) => void
 - o #pluginsAnimationRecived: ({ plugin, input }: { plugin: ImpulsarClientPlugin; input: ImpulsarClientPluginReceiveAnimation; }) => void
 - o #pluginsStreamAudio: ({ plugin, input }: { plugin: ImpulsarClientPlugin; input: ImpulsarClientPluginStreamAudio; }) => void
 - o #pluginsInit: () => void
 - o #onIOConnect: () => void
 - o #onIODisconnect: () => void
 - o #onReciveSubtitle: ({ data }: { data: string; }) => void
 - o #onReciveDefaultAnimation: ({ data }: { data: any; }) => void
 - o #initIO: () => void
 - o #init: () => void
 - o connect: () => void
 - o disconnect: () => void
 - o stream: (streamData: ImpulsarClientStreamData) => boolean
 - o send: (sendData: ImpulsarClientSendData) => boolean

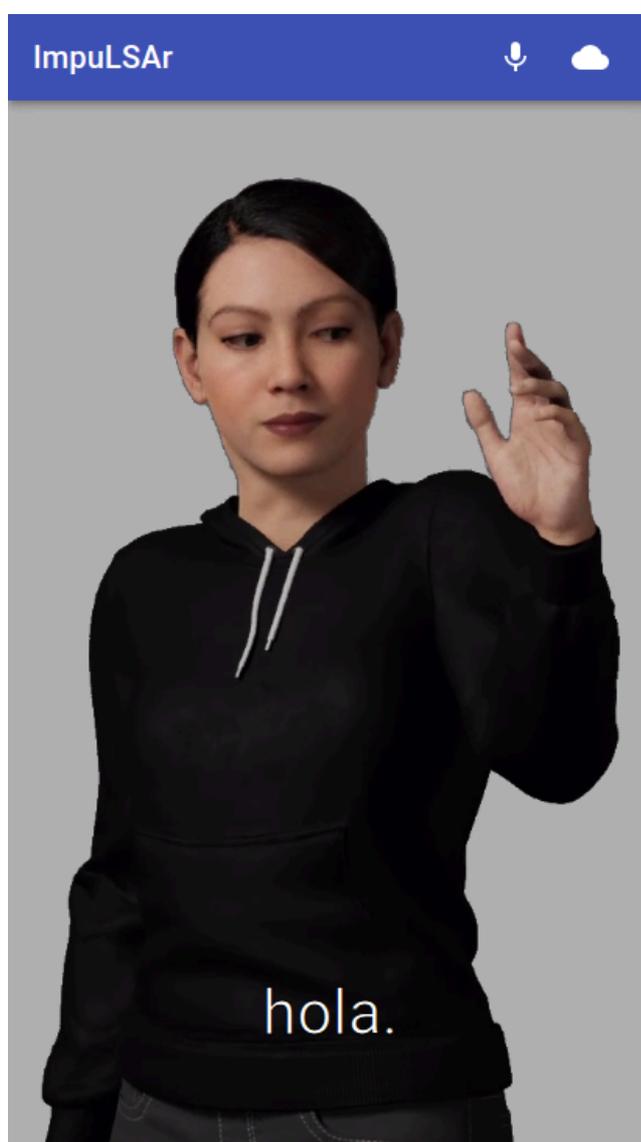
ImpulsarServer

- o app: any
 - o server: Server
 - o serverPort: number
 - o io: IOSever<DefaultEventsMap, DefaultEventsMap>
 - o ioStream: SocketStream
 - o sampleRate: number
 - o voskModel: any
 - o googleSpeechClient: SpeechClient
 - o googleSpeechStreamRecognitionConfig: google.cloud.speech.v1.IStreamingRecognitionConfig
-
- o defineHttpRoutes: () => void
 - o defineClientEvents: (client: Socket) => void
 - o onClientConnected: (client: Socket) => void
 - o onClientStreamAudio: (client: Socket, stream: any) => void
 - o defineSocketEvents: () => void
 - o debugAudioStream: (audio: any) => void
 - o recognizeTextStream: (text: any, cb: any) => void
 - o sanitizeTextStream: (text: any, cb: any) => void
 - o translateTextStream: (text: any, cb: any) => void
 - o transcribeAudioStream: (audio: any, cb: any) => void
 - o transcribeOfflineAudioStream: (audio: any, cb: any) => void
 - o init: () => void

Prototipo de interfaz de pantalla

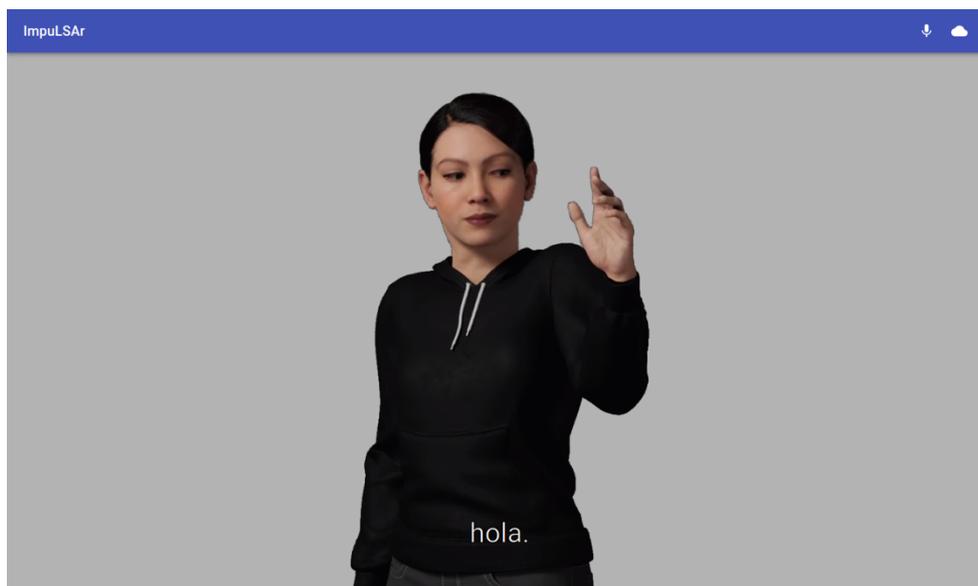
Al ser el LSA una lengua viso-espacial, manual y gestual, es fundamental generar una imagen sin interferencias de ninguna índole. Por lo tanto, el fondo consiste en un color plano sin irregularidades, y el modelo posee un rostro despejado sin accesorios que generen distracción. Los colores de la vestimenta permiten la visualización clara de las manos y los movimientos del tronco superior de la intérprete y del espacio señante.

Imagen 7 - Prototipo de interfaz de pantalla: aplicación de ejemplo.



Fuente: Elaboración propia

Imagen 8 - Prototipo de interfaz de pantalla: aplicación de ejemplo.



Fuente: Elaboración propia

Por otro lado, se presentan dos elementos gráficos que cumplen la funcionalidad de botones (conexión y micrófono activo), la elección de estos dos elementos en forma gráfica y no textual, está relacionada a las necesidades lingüísticas de los usuarios.

Distinguibilidad

Uno de los factores que contribuyen a que los elementos sean distinguibles está relacionado a la selección de colores de fondo y primer plano (pauta 1.4 de la WCAG 2.1), particularmente a la diferencia de los ratios de contraste. Contrariamente a la creencia popular de que este factor está dado por la diferencia de matices, (por ejemplo, amarillo sobre azul), es más efectivo conseguir un mejor contraste entre colores con matices de color similar pero diferente brillo. Esto ocurre ya que el ratio de contraste se debe a la diferencia de luminancia percibida.

Según el criterio 1.4.6. de la WCAG 2.1 de contraste mejorado (Para nivel AAA), la presentación de textos e imágenes de texto debe tener un ratio de contraste de al menos 7:1.

Imagen 9 - Análisis de ratio de contraste de la interfaz



Fuente: Elaboración propia

En la muestra izquierda podemos observar el color perteneciente a la barra superior, en combinación al correspondiente del logotipo y los botones.

Barra: hsla(231, 48%, 48%) - Botones: hsla(0, 100%, 100%).

El ratio de contraste es de 12,724:1. El contraste es suficiente para satisfacer el nivel AAA.

En la muestra derecha podemos señalar los colores correspondientes al fondo del modelo en contraste a la barra superior.

Fondo: hsla(0, 0%, 70%) – Barra: hsla(231, 48%, 48%)

El ratio de contraste es de 7,4854:1. El contraste es suficiente para satisfacer el nivel AAA.

Diagrama de arquitectura

Control de Versiones

El proyecto utiliza un esquema de repositorios del tipo mono-repo, donde todos los componentes de la solución se encuentran alojados en el mismo repositorio de Git en la plataforma de GitHub.

Integración Continua

Se utiliza el sistema de acciones brindado por Github para realizar las compilaciones de los componentes y las pruebas correspondientes para garantizar la

calidad del código, la funcionalidad del mismo y la integración conjunta de los diferentes componentes.

Despliegue Continuo

Una vez cumplidos todos los requerimientos de la integración se procede a distribuir los distintos componentes a sus correspondientes destinos:

- Las librerías se publican en el repositorio de Node Package Module (NPM) donde son utilizadas por desarrolladores de terceros para la implementación de la solución en sus proyectos.
- El servicio de procesamiento y traducción se publica en el repositorio de Docker Hub en forma de una imagen embebida con todos los requerimientos lista para ser utilizada por cualquier motor de ejecución de contenedores.
- La documentación del proyecto junto a la aplicación de ejemplo es desplegada a un disco de almacenamiento provisto por Google Cloud Storage.

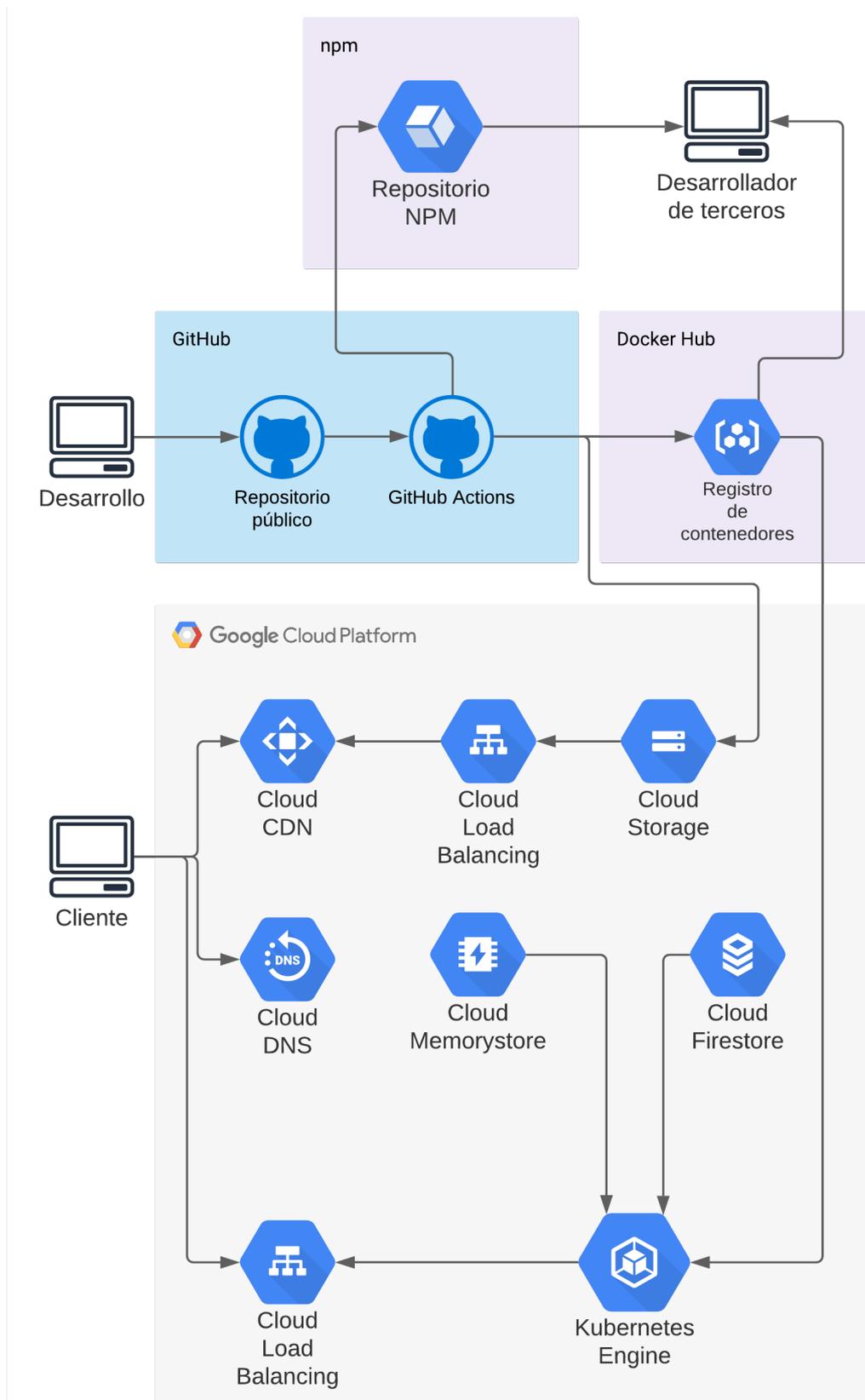
Entorno Productivo

El proyecto cuenta con tres servicios públicos expuestos a internet para ser consumidos por las diferentes soluciones de terceros que hacen uso del proyecto:

- Servidor de DNS: provisto por el servicio de Cloud DNS de Google, es el encargado de resolver los dominios del proyecto a las direcciones de IP correspondientes.
- Página Web: Partiendo del disco de almacenamiento en Google Cloud Storage donde se encuentran los archivos estáticos HTML, CSS y JS se utiliza el servicio de Google Load Balancing para proveer un servidor web que responda a la peticiones de los navegadores web y brinde los archivos correspondientes. Finalmente se utiliza Google Cloud CDN el cual nos provee una caché y distribución de estos contenidos en diferentes servidores alrededor del mundo para ofrecer la menor latencia posible de cara del usuario final, garantizando los más altos estándares de conectividad.
- Servicio de Procesamiento y Traducción: Se implementa un Cluster de Kubernetes provisto por Google Kubernetes Engine el cual nos permite una escalabilidad lineal virtualmente ilimitada, este utiliza la imagen alojada en

Docker Hub para proveer los contenedores. Estos contenedores acceden a dos bases de datos, por un lado Google Cloud Firestore para almacenar las clave valor de las palabras y su traducción a LSA como así también las animaciones correspondiente, y por otro lado Google Memory Store para el almacenamiento de caché, sincronización de variables entre los nodos del cluster y datos que necesiten una rápida velocidad de acceso y disponibilidad. Finalmente se dispone de un balanceador de carga provisto por Google Cloud Load Balancing que distribuye las peticiones entrantes entre los diferentes nodos del cluster.

Imagen 10 - Diagrama de infraestructura



Fuente: Elaboración propia

Diagrama de componentes

Servidor de Procesamiento y traducción (server)

Es el servicio principal del proyecto basado en WebSocket sobre NodeJS donde se nuclea la recepción de las transmisiones de audio de los clientes a las cuales vía diferentes vías de procesamiento. Esta información se transforma en distintas salidas:

- A texto en el lenguaje Español utilizando modelos de reconocimiento de voz del proyecto Google Speech Recognition.
- A animaciones representativas del lenguaje LSA utilizando una búsqueda heurística en una base de datos de clave-valor.

Librería Cliente (client)

Es la librería base del proyecto encargada de:

- Proveer la conectividad al servidor del proyecto.
- Establecer las interfaces necesarias para definir nuevas librerías que extiendan la funcionalidad de la misma en forma de Plugins.
- Establecer los estándares y canales de comunicación hacia el servidor.
- Gestionar la inicialización de los plugins requeridos y la comunicación bidireccional de estos con el servidor.

Librería de Captura de Sonido (plugin-record)

Este plugin utiliza las tecnología y estándares WebRTC para proporcionar el manejo de los permisos del navegador para acceder a los dispositivos de captura de sonido del hardware del cliente (por ejemplo, el micrófono), como así también la captura, codificación y transmisión del audio hacia la librería cliente.

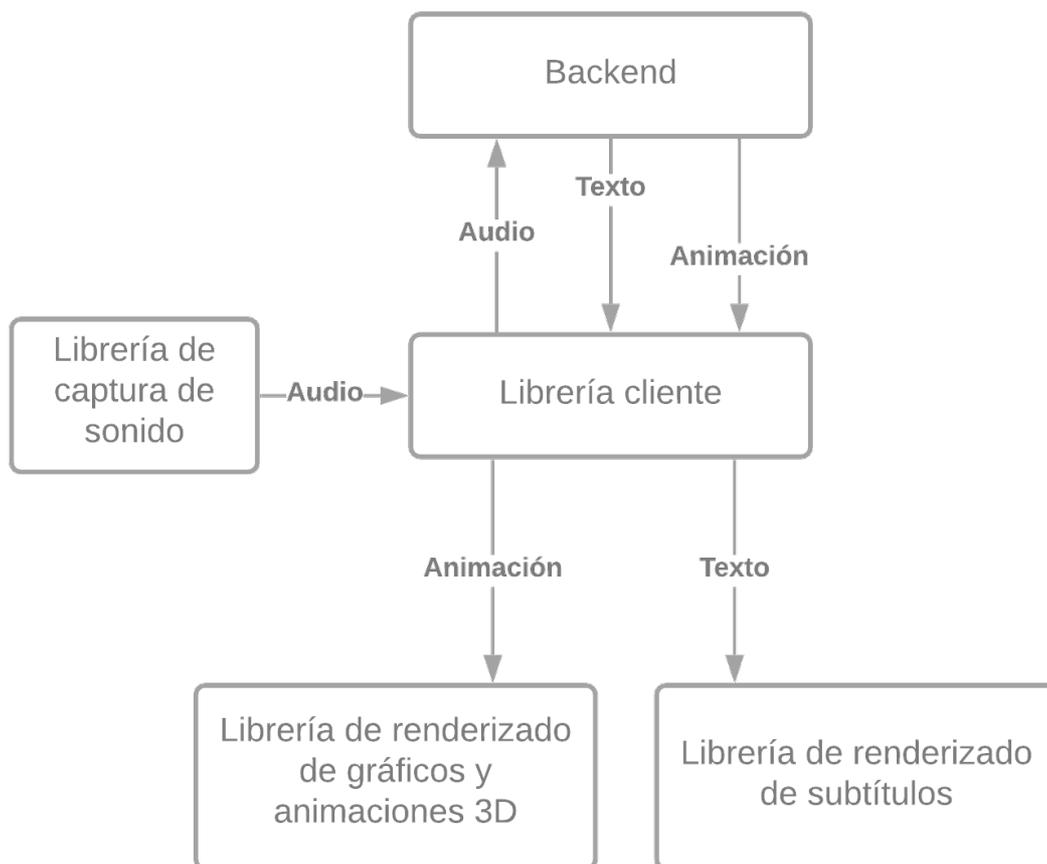
Librería de Renderizado de Subtítulos (plugin-subtitle)

Este plugin hace uso de una manipulación directa del modelo de objetos del documento (DOM) del navegador web para renderizar en tiempo real los textos en español que son recibidos desde la librería cliente.

Librería de Renderizado de gráficos y animaciones 3D (plugin-webgl-render)

Este plugin hace uso de las tecnologías de WebGL para renderizar un modelo tridimensional de un intérprete que se anima en tiempo real en base a las animaciones que son recibidas desde la librería cliente.

Imagen 11 - Diagrama de componentes



Fuente: Elaboración propia

Seguridad

Acceso a la aplicación

La solución se diseñó para ser utilizada libremente por cualquier usuario que lo requiera, la misma no brinda ningún tipo de control de acceso a los datos, ya que estos se encuentran disponibles públicamente y queda en responsabilidad de terceros si la implementación del proyecto lo requiere.

Políticas de respaldo de información

El proyecto basa su arquitectura de datos en documentos almacenados en el repositorio de Git, junto al código fuente de los distintos paquetes que componen la solución. En el proceso de despliegue, la base de datos (encargada de almacenar las animaciones y las relaciones contra el diccionario de las mismas) es recreada cada vez que se realiza un despliegue automático. Por lo tanto, toda la política de respaldo se da a través de la implementación del protocolo de Git para el manejo de versiones de código, y la fiabilidad de esta información es responsabilidad de GitHub. Sin embargo, por la misma naturaleza del protocolo, cada desarrollador del proyecto mantiene copia de todo el código fuente en el dispositivo en el cual realiza el desarrollo.

Análisis de Costos

A continuación se indican los costos de honorarios correspondientes al desarrollo de la aplicación. Tales costos fueron calculados en base a los honorarios provistos por el Consejo Profesional De Ciencias Informáticas de Córdoba, y por la Asociación Argentina de Traductores e Intérpretes, ambos consultados durante junio de 2021.

Tabla 12 - Costos de desarrollo

Profesional	Tiempo	Costo de hora	Costo total
Desarrollador	160 horas	\$4330	\$692.800
Desarrollador	160 horas	\$4330	\$692.800
Intérprete LSA	160 horas	\$3250	\$520.000
Intérprete LSA	160 horas	\$3250	\$520.000
Diseñador multimedia	480 horas	\$701	\$336.546
Total			\$1.377.931

Fuente: CPCIPC (cpcipc.org.ar, 2021) - AATI (aati.org.ar, 2021)

Costos de servicios e infraestructura

Para realizar el análisis de manera precisa es necesario tener en cuenta el tamaño aproximado de las transferencias entre cliente-servidor (y viceversa) a partir de las cuales realizaremos el cálculo de los costos.

Tabla 13 - Análisis de transferencias

Trasmisión de audio desde cliente a servidor	28,12 MB/h
Extra trasmisión desde servidor a cliente de subtítulos	5,6 MB/h
Extra trasmisión desde servidor a cliente de interpretación	14,06 MB/h
Descarga de archivos estáticos de la aplicación	5 MB

Costos variables

Los costos de infraestructura son detallados a continuación en pesos argentinos, en base al costo de los servicios dispuestos por Google Cloud en su página oficial, expresados en dólar estadounidense. Al momento de la estimación (junio 2021) la relación entre ambas divisas es de 1 dólar estadounidense = 95,13 pesos argentinos. Los costos variables refieren a servicios que se cobran por el proveedor de acuerdo al uso, correspondiente a la demanda de los usuarios.

Tabla 14 - Costos variables

Servicio	Parámetros de cobro	Unidad	Precio unitario	Uso Estimado	Costo por hora de servicio
SpeechToText	Tiempo de uso por bloques – Precio desde 60 minutos a 1 millón de minutos	15 segundos	\$0,006	3600 segundos	\$136,98
Cloud Load Balancing	Cantidad de GB utilizados por mes	1 GB	\$0,008	0,046 GB	\$0,03
Instancia de Kubernetes	CPU autopilot de GKE por hora	1 vcpu por hora	\$0,0445	1 hora	\$5,17
	Memoria del POD por hora por GB	2 GB por hora	\$0,0049	1 hora	
	Almacenamiento efímero por hora por GB	1 GB por hora	\$0,0000548	1 hora	
Costo total de procesamiento por hora					\$141,74

Costos extraídos de la página oficial de Google Cloud (Cloud.google.com, 2021)

Servicios con costos fijos

Tabla 15 - Costos Fijos

Servicio	Proveedor	Parámetro de cobro	Uso estimado	Costo por mes
Cloud Storage	Google	GB por mes	1 GB	\$1,90
Cloud Memory Store	Google	M1 (1 a 4 GB de capacidad) 3 GB/s rendimiento – GB/Hora.	720 horas por mes	\$4.383,59
Firestore	Google	GB. Gratuito por 1 GB	50 MB	Gratis
Procesamiento mínimo Kubernetes	Google	Contenedor 1vcpu – 2Gb RAM	720 horas por mes	\$3.725,29
Costo fijo total				\$8.110,78

Costos extraídos de la página oficial de Google Cloud (Cloud.google.com, 2021)

Costo de visita

Este apartado refiere a cada persona nueva que visita la aplicación de ejemplo. Como se mencionaba anteriormente, el peso total de la aplicación es de 5MB en archivos estáticos, por lo tanto 1000 visitas únicas equivalen a 5GB. El costo proyectado para mil visitas es de \$51,37, siendo el costo por cada visita de \$0,05.

Costo total estimado de servicios

Esta estimación responde al uso de 10 personas, con un promedio de uso de 1 hora por persona por día, durante el plazo de un mes.

Tabla 16 - Costo total de infraestructura

Servicios con costos variables	\$37.370,86
Servicios de costo fijo	\$8.110,78
Costo de visitas	\$15,41
Total por mes	\$46.353,04

Análisis de Riesgos

A continuación se realiza el análisis de los riesgos del proyecto, su exposición y su impacto estimado, posteriormente se indica el correspondiente plan de contingencia.

Tabla 17 - Riesgos identificados del proyecto

Tipo	Riesgo	Causas
Proyecto	Insolvencia	-No se cuenta con financiación para solventar los costos mensuales de la infraestructura del proyecto. -No se cuenta con financiación para solventar el costo de desarrollo.
Sistema	Corte en el servicio de infraestructura	-Se interrumpe el servicio de un componente de la infraestructura o de varios.

Proyecto	Corte en el repositorio de NPM	-Corte momentáneo del servicio
Proyecto	Corte de DockerHub	-Corte momentáneo del servicio
Proyecto	Baja de GitHub	-Corte momentáneo del servicio
Proyecto	Cambios en las políticas de SpeechToText	-El precio del servicio subió por encima del doble. -Cambios en las políticas de aceptación de streaming de audio. -El proyecto se abandona.
Proyecto	Traducciones poco fiables	-Errores de traducción por falta de señas o interpretación incorrecta

Tabla 18 - Análisis cualitativo de riesgos

ID	Tipo	Riesgo	Probabilidad	Impacto
R1	Proyecto	Insolvencia	Media	Muy Alto
R2	Sistema	Corte en el servicio de infraestructura	Muy Baja	Muy Alto
R3	Proyecto	Corte en el repositorio de NPM	Muy Baja	Medio
R4	Proyecto	Corte en DockerHub	Muy Baja	Muy Bajo
R5	Proyecto	Baja de GitHub	Muy Baja	Muy Bajo
R6	Proyecto	Cambios en las políticas de SpeechToText	Baja	Alto
R7	Sistema	Traducciones poco fiables	Alta	Muy Alto

Análisis cuantitativo de los riesgos

A continuación se indica el análisis cuantitativo de los riesgos y los parámetros establecidos para establecer las mediciones.

Tabla 19 - Relación entre parámetros cualitativos y cuantitativos

Probabilidad		Impacto	
Descripción	Valor	Descripción	Valor
Muy Baja	0.1	Muy bajo	0.05
Baja	0.3	Bajo	0.1
Media	0.5	Medio	0.2
Alta	0.7	Alto	0.4
Muy alta	0.9	Muy alto	0.8

Tabla 20 - Asignación de parámetros cuantitativos

Riesgo	Probabilidad	Impacto
Insolvencia	0.5	0.8
Corte en el servicio de infraestructura	0.1	0.8
Corte en el repositorio de NPM	0.1	0.2
Corte en DockerHub	0.1	0.05
Baja de GitHub	0.1	0.05
Cambios en las políticas de SpeechToText	0.3	0.4
Traducciones poco fiables	0.7	0.8

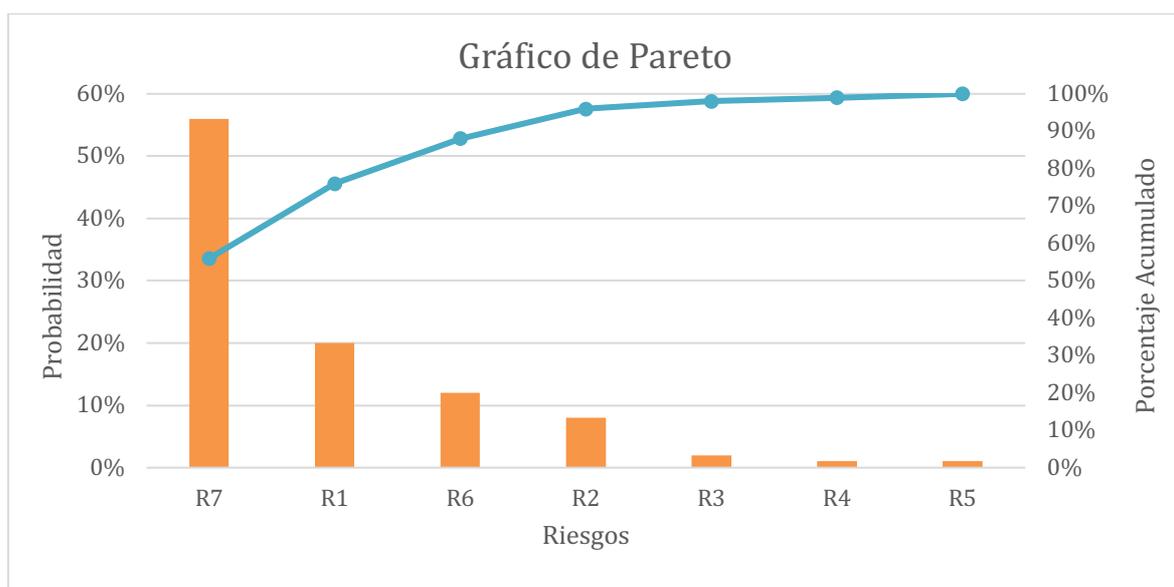
Tabla 21 - Matriz de probabilidad e impacto.

Probabilidad		Impacto				
		Muy Bajo	Bajo	Medio	Alto	Muy Alto
		0.05	0.10	0.20	0.4	0.8
Muy Alta	0.9	0.05	0.09	0.18	0.36	0.72
Alta	0.7	0.04	0.07	0.14	0.28	0.56
Media	0.5	0.03	0.05	0.10	0.20	0.40
Baja	0.3	0.02	0.03	0.06	0.12	0.24
Muy Baja	0.1	0.01	0.01	0.02	0.04	0.08

Tabla 22 - Análisis cuantitativo y grado de exposición

ID	Probabilidad	Impacto	Grado de exposición	porcentaje	porcentaje acumulado
R7	56,00%	0,56	0,8	56,00%	56,00%
R1	20,00%	0,2	0,4	20,00%	76,00%
R6	12,00%	0,12	0,4	12,00%	88,00%
R2	8,00%	0,08	0,8	8,00%	96,00%
R3	2,00%	0,02	0,2	2,00%	98,00%
R4	1,00%	0,01	0,1	1,00%	99,00%
R5	1,00%	0,01	0,1	1,00%	100,00%

Imagen 12 - Gráfico de Pareto



A través del gráfico anterior, podemos establecer un concreto orden de prioridades, discriminando entre los problemas fundamentales y los que impactan de manera menos crítica. Por lo tanto, podemos señalar que los riesgos R7, R1 y R6, acumulan el 80% de los inconvenientes y son estos los que deben atenderse de manera prioritaria.

Tabla 23 - Plan de contingencias

ID	Prioridad	Riesgo	Plan de contingencia
R7	Alta	Traducciones poco fiables	Realizar sesiones de entrenamiento de los datasets con moderación manual
R1	Alta	Insolvencia	<ul style="list-style-type: none"> -Buscar inversión de privados o realizar campañas de financiamiento colaborativo. -La infraestructura se convierte en responsabilidad del implementador. -Integrar publicidad al modelo de negocios. -Cobrar por el servicio a empresas que implementen las librerías en sus aplicaciones.
R6	Alta	Cambios de las políticas y/o precio de SpeechToText	Suplantar la integración de SpeechToText e implementar Vosk. Realizar un reentrenamiento de los datasets.
R2	Media	Corte en el servicio de infraestructura.	La solución puede implementarse en otro proveedor.

Conclusiones

Se llevó a cabo el desarrollo de una plataforma para la traducción de audio español a lenguas de señas argentinas. El desarrollo de esta herramienta se realizó con el objetivo de ser el puntapié inicial hacia el camino de la accesibilidad de la población sorda en el ámbito digital. Para ello, fue necesario sentar las bases técnicas para facilitar la adopción comunitaria para el desarrollo colaborativo. Desde este aspecto es que significó un gran desafío a la hora del análisis, el diseño y la selección de las tecnologías, ya que su requerimiento fundamental debía ser dotarla de una gran versatilidad en cuanto a la compatibilidad y escalabilidad de cada uno de sus componentes, para favorecer la adopción global y significar una herramienta robusta y de gran calidad técnica, plausible de ser implementada en todo proyecto en el que sea necesario.

No es casualidad la ausencia de este tipo de herramientas en la tecnología actual. A la complejidad técnica del lenguaje natural oralista, (siendo hoy un motivo de exploración científica en el campo del desarrollo de algoritmos), se le suma una complejidad considerable: la relación de éste con el lenguaje viso-espacial, aquel que se compone de sus propias reglas gramaticales, sus propios conceptos y componentes. Aquel que, lamentablemente, solo es dominado por una pequeña parte de la población mundial.

Para lograr los objetivos planteados, fue necesario realizar un recorrido transversal a través de diversos campos de conocimiento, para ello fue indispensable la interrelación de conceptos sociológicos, lingüísticos y antropológicos con los del campo de la tecnología. Por esta razón, significó un proceso sumamente enriquecedor, al tener la oportunidad de aplicar las herramientas obtenidas a lo largo de todo mi recorrido académico, como así también la necesidad de descubrir nuevas.

No puedo evitar señalar el gran desafío personal que significó la exploración de la problemática. En el marco de un trabajo sobre discapacidad, fue necesario establecer la responsabilidad y la empatía como valores rectores del marco de trabajo, no solo para evitar imponer nuevas barreras de integración sino que también para proponer desde el respeto por las identidades. Por lo tanto, la búsqueda del entendimiento significó un camino que se transitó con precaución y responsabilidad, los cuales considero que fueron

la clave para conseguir un diseño técnico eficaz en primera instancia, y alcanzar los objetivos planteados al momento de la primer versión de la herramienta.

Sin embargo, no puede omitirse la desigualdad que ha sido evidenciada con ímpetu. La realidad de las poblaciones con discapacidad en nuestro país es, cuanto menos, dolorosa. Es por ello que encuentro sumamente satisfactorio el haber tenido la oportunidad de aportar a través de mis conocimientos en la construcción de una sociedad más inclusiva y accesible, con el anhelo de vivir cada día en un mundo más justo, del que todos podamos sentirnos parte.

Demo

En el siguiente enlace se encuentra el repositorio principal de la plataforma, el cual contiene el código fuente y la documentación necesaria para su implementación.

Link: <https://github.com/LSAorg/impulsar>

Referencias

Bierman G., Abadi M. y Torgersen M. (2014). Entendiendo TypeScript. - ECOOP 2014 Programación Orientada a Objetos. ECOOP 2014. vol 8586.

BURAD, V. (2013) “Los derechos de la población sorda: trabajo y ciudadanía plena”. Recuperado el 09/04/2021 de: [https://cultura-sorda.org/wp-content/uploads/2015/02/Burad_Viviana_Derechos_poblacion_sorda_trabajo_ciudadania_plena_20131.pdf]

Gauchat, D. (2021). El Gran Libro De HTML, CSS3 y Javascript - Primera Edición

GitHub Actions – Página oficial de GitHub - recuperado el 21/04/2021 de [<https://docs.github.com/es/actions>]

GitHub – Página oficial - recuperado el 21/04/2021 de [<https://github.com/Hispano/Guia-sobre-Git-Github-y-Metodologia-de-Desarrollo-de-Software-usando-Git-y-Github>]

Google Cloud Store – Página oficial de Google firebase - recuperado el 20/04/2021 de [<https://firebase.google.com/docs/storage>]

Google Memory Store – Página oficial - recuperado el 20/04/2021 de [<https://cloud.google.com/memorystore/>]

INDEC (2018). Instituto Nacional de Estadística y Censos - Estudio Nacional sobre el Perfil de las Personas con Discapacidad: resultados definitivos 2018. - Ciudad Autónoma de Buenos Aires.

Juliarena, G. E. (2012). Bilingüismo en sordos [en línea]. IV Jornadas de Español como Lengua Segunda y Extranjera, 7 y 8 de noviembre de 2012, La Plata, Argentina. Experiencias, Desarrollos, Propuestas. En Memoria Académica. Recuperado

04/04/2021

de:

[http://www.memoria.fahce.unlp.edu.ar/trab_eventos/ev.2725/ev.2725.pdf]

Kubernetes – Página oficial – recuperado el 20/04/2021 de [<https://kubernetes.io/>]

Lightbown y Spada, (2006). *How Languages are Learned*. Oxford: Oxford University Press. ISBN 978-0-19-442224-6.

Mariño, S., y Alfonzo, P. (2014). Implementación de SCRUM en el diseño del proyecto del Trabajo Final de Aplicación. *Scientia Et Technica*, ISSN: 0122-1701. recuperado el 21/04/2021 de [<https://www.redalyc.org/articulo.oa?id=8493>]

Nodejs – Página Oficial de OpenJs foundation – recuperado el 20/04/2021 de [<http://www.NodeJs.org> – 2021]

NodeJS - Página oficial de Node Package Module - recuperado el 21/04/2021 de [<https://nodejs.org/api/packages.html>]

Organización Mundial de la Salud (2015). Sordera y pérdida de la audición. Recuperado 02/03/2021 de [<http://www.who.int/mediacentre/factsheets/fs300/es/>]

Oviedo, (2003). El estudio de la lengua de señas venezolana – Universidad de los Andes – recuperado el 04/04/2021 de: [https://cultura-sorda.org/wp-content/uploads/2015/03/Oviedo_Estructura_LSV_Parte1_de_6-1.pdf]

Parisi, T. (2012) *WebGL: iniciando y corriendo – Creando una web 3D* - O'Reilly Media, Inc. ISBN: 9781449323578

Peluso, L (2010). *Sordos y oyentes en un liceo común. Investigación e intervención en un contexto intercultural*. Montevideo: Psicolibros – Recuperado el 20/04/2021 de [<https://www.academica.org/000-051/16.pdf>]

- RedHat – Página oficial de RedHat - recuperado el 21/04/2021 de
[<https://www.redhat.com/es/topics/containers/what-is-docker>]
- Rey, M. I., Famularo, R. y Ringuelet, R. (2020). La alteridad compleja de la comunidad sorda argentina. *Revista de Psicología*, 19(1), 18-31. doi: 10.24215/2422572Xe050 Recuperado el 08/04/2021 de
[<https://revistas.unlp.edu.ar/revpsi/article/download/9779/8579/>]
- Roso (2013). Las lenguas de señas: una aproximación al conocimiento de sus características y propiedades ISSN 1851-8060 Secretaría de investigación, Ciencia y Técnica - FFYH - Universidad Nacional de Córdoba – Recuperado el 13/04/2021 de [<http://bit.ly/roso2013>]
- Skliar C (2003). La educación de los sordos – Recuperado el 12/04/2021 de
[https://cultura-sorda.org/wp-content/uploads/2015/03/Skliar_educacion_sordos-2003.pdf]
- Socket.io – Página Oficial de Open Source Collective - recuperado el 20/04/2021 de
[<http://www.socket.io> – 2021]
- Veinberg, S. (1996) Argentinien: Anfänge und Entwicklung der Erziehung Gehörloser (Argentina: Inicios y desarrollo de la educación del sordo). *Das Zeichen. Zeitschrift Zum Thema Gebärdensprache und Kommunikation Gehörlose*, Recuperado el 09/04/2021, de:
[<http://www.modalidadespecial.educ.ar/datos/recursos/pdf/argentina-inicios-y-desarrollo-de-la-educacion-de-los-sordos.pdf>]
- WebRTC 2021 – Página oficial de Google Developers – recuperado el 20/04/2021 de
[<http://www.webrtc.org>]
- WCAG (2018). How to Meet WCAG (Quick Reference) A customizable quick reference

to Web Content Accessibility Guidelines, requirements (success criteria) and techniques – Recuperado el 12/04/2021 de [<https://www.w3.org/WAI/WCAG21/quickref/#sign-language-prerecorded>]