CatCh AI – Traducción Automática de Voz a Lengua de Señas Chilena (LSCh)

CatCh AI – Sprint 1: Desarrollo del Pipeline Funcional (Voz → Texto → Seña)

**Sede:** Duoc UC – Puerto Montt  
**Profesora Guía:** Giocrisrai Godoy  
**Integrantes:** Eyleen Collado (Product Manager), Marcos Bombalas, Marcos Hernández, Luis Gómez  
**Duración:** 1 al 21 de septiembre de 2025

**Versión:** 1.0

**Tabla de Control de Cambios**

| **Versión** | **Fecha** | **Autor** | **Descripción del Cambio** |
| --- | --- | --- | --- |
| 1.0 | 01/09/2025 | Eyleen Collado | Creación del documento de Sprint 1 |
| 1.1 | 10/09/2025 | Eyleen Collado | Actualización tras validación del pipeline de voz a texto |
| 1.2 | 22/09/2025 | Eyleen Collado | Inclusión de retroalimentación de profesora guía y pruebas del diccionario |

# Sprint Planning

Objetivo General  
Desarrollar y validar la primera versión funcional del pipeline de traducción de voz a texto y su integración con el módulo de generación de señas, utilizando el diccionario de animaciones iniciales y el flujo de pruebas locales.

## Historias de Usuario Seleccionadas

* HU-004: Como usuario, quiero ingresar audio desde micrófono o archivo para obtener la traducción en texto.
* HU-005: Como sistema, debo interpretar las palabras procesadas y vincularlas con las animaciones de señas almacenadas en el diccionario.
* HU-006: Como equipo, necesito definir la estructura base del diccionario LSCh y las reglas de enlace entre texto y seña.

## Criterios de Aceptación

* El sistema genera texto legible con precisión ≥80%.
* El avatar muestra la animación correcta asociada a cada seña registrada.
* El pipeline completo se ejecuta sin interrupciones (voz → texto → seña).

Estimaciones de Esfuerzo  
Duración total: 3 semanas (30 horas estimadas por integrante).

## Riesgos Identificados

* Falta de cobertura del diccionario en palabras cotidianas.
* Problemas de sincronización entre audio y animación.
* Sobrecarga de procesamiento en equipos sin GPU.

## Mitigaciones

* Incremento progresivo del diccionario según demanda.
* Ajuste manual de tiempos de animación en Godot.
* Uso de modelos Whisper tiny/base según hardware disponible.

## Definition of Done (DoD)

* Pipeline funcional probado con ejemplos propios.
* Diccionario operativo con al menos 20 palabras base.
* Documento técnico actualizado con estructura y enlace de señas.

# Sprint Execution

## Tareas ejecutadas

| **Tarea** | **Responsable** | **Estado** | **Observaciones** |
| --- | --- | --- | --- |
| Desarrollo del pipeline voz → texto | Luis Gómez | Completada | Whisper local validado con 85% precisión promedio |
| Integración del módulo de texto → seña | Marcos Hernández | Completada | Se mapearon palabras clave al diccionario inicial |
| Generación de animaciones base en Blender | Marcos Bombalas | En curso | 15 animaciones exportadas en formato GLB |
| Pruebas de importación en Godot | Eyleen Collado | Completada | Validación visual correcta con avatar neutral |
| Documentación técnica y estructura de carpetas | Equipo | Completada | Versión inicial del README funcional publicada |
| Desarrollo del pipeline voz → texto | Luis Gómez | Completada | Whisper local validado con 85% precisión promedio |
| Integración del módulo de texto → seña | Marcos Hernández | Completada | Se mapearon palabras clave al diccionario inicial |
| Generación de animaciones base en Blender | Marcos Bombalas | En curso | 15 animaciones exportadas en formato GLB |
| Pruebas de importación en Godot | Eyleen Collado | Completada | Validación visual correcta con avatar neutral |
| Documentación técnica y estructura de carpetas | Equipo | Completada | Primer diagrama de flujo del pipeline completo |

Observaciones Generales  
Se consiguió que el pipeline procesara correctamente audios de menos de 10 segundos en entorno local, entregando resultados legibles en texto y activando la animación correspondiente en Godot. El diccionario base contempla palabras como “hola”, “gracias”, “ayuda”, “bien”, “día”, entre otras.

# Sprint Review

## Entregables

* Pipeline funcional local (voz → texto → seña).
* Diccionario LSCh base con 20 términos y animaciones asociadas.
* Registro de pruebas y videos demostrativos.

## Retroalimentación (22/09/2025, Prof. Giocrisrai Godoy)

* Priorizar pruebas con usuarios no técnicos para validar comprensión de señas.
* Mejorar documentación del diccionario y esquema de metadatos.
* Incluir trazabilidad de errores del pipeline en los informes técnicos.

## Acciones derivadas

* Se agregaron comentarios de validación externa (3 observadores).
* Se actualizó la documentación del diccionario con formato JSON propuesto.
* Se comenzó el registro de errores y tiempos de respuesta por cada ejecución.

# Sprint Retrospective

## Qué funcionó

* Primer flujo funcional completado.
* Coordinación efectiva entre IA y animación.
* Integración estable entre Blender y Godot.

## Qué mejorar

* Estandarizar los nombres y rutas del diccionario.
* Automatizar los enlaces texto-seña.
* Definir protocolo de prueba uniforme.

## Compromisos para Sprint 2

* Implementar el procesamiento continuo (RNF.03).
* Expandir el diccionario a 50 términos.
* Iniciar pruebas con interpretación de frases.

# 5. Seguimiento de Avance (Burndown Chart)

| **Semana** | **Tareas planificadas** | **Tareas completadas** | **% Avance** |
| --- | --- | --- | --- |
| Semana 1 | 7 | 4 | 57% |
| Semana 2 | 8 | 7 | 87% |
| Semana 3 | 5 | 5 | 100% |

Conclusión  
El sprint logró consolidar el núcleo funcional del sistema CatCh AI. Se demostró la viabilidad del pipeline de traducción y la integración modular entre los componentes. La siguiente fase se centrará en optimización, validación de usuarios y ampliación del diccionario.