

SaludTech de los Alpes (STA), basado en Domain- Driven Design (DDD)

Grupo: [Error404](#)

Integrantes:

- Adrián Francisco Becerra Arias
- Camilo Lesmes
- Javier Estupiñan
- Anderson Rodriguez

DOCUMENTACIÓN DE DOMINIOS Y SUB-DOMINIOS

Visión global del negocio y dominios

SaludTech de los Alpes (STA) es una empresa cuyo modelo de negocio se centra en la recolección, anonimización y distribución de imágenes médicas y diagnósticos a organizaciones y desarrolladores de IA. Para estructurar la complejidad, el dominio principal de STA es la logística y distribución de imágenes y diagnósticos a compañías de IA y desarrolladores y se distribuye en los siguientes dominios y subdominios:

1. **ProcessingImageDiagnosticDomain (Dominio principal):** Logística y distribución de imágenes y diagnósticos
 - a. CollectSubdomain (Subdominio): Recaudo de imágenes y diagnósticos.
 - b. ProcessSubdomain (Subdominio): Procesamiento de imágenes y diagnósticos.
 - c. StructuringSubdomain (Subdominio): Canonización de imágenes, diagnósticos y tokenización de pacientes.
 - d. DistributeSubdomain (Subdominio): Distribución de imágenes, diagnósticos e información adicional.
 - e. UserSalesAssistanceSubdomain (Subdominio): Servicio de atención de ventas y usuarios.

ContextMapper: <https://github.com/elle184/SaludTechDeLosAlpes-Error404/blob/master/src/main/cml/domain/processing-image.cml>.

Dominio principal

(ProcessingImageDiagnosticDomain) Logística y distribución de imágenes y diagnósticos para compañías de IA y Desarrolladores

Subdominios

Vision statement:
Gestiona todo el proceso y estrategia de recaudo de las imágenes y diagnósticos médicos.
Tipo: Support

(CollectSubdomain) Recaudo de imágenes y diagnósticos

Vision statement:
Procesamiento de la información de las imágenes y diagnósticos.
Tipo: Núcleo

(ProcessSubdomain) Procesamiento de imágenes y diagnósticos

Vision statement: Mapeo y estructuración de los datos anonimizados junto con la tokenización de pacientes.
Tipo: Núcleo

(StructuringSubdomain) Canonización de imágenes, diagnósticos y tokenización de pacientes.

Vision statement: Gestión de toda la logística de distribución de las imágenes, diagnósticos e información adicional a los usuarios registrados.
Tipo: Núcleo

(DistributeSubdomain) Distribución de imágenes, diagnósticos e información adicional

Vision statement:
Gestión de usuarios y clientes que contratan los servicios de STA.
Tipo: Generic

(UserSalesAssistanceSubdomain) Servicio de atención de ventas y usuarios.

HISTORIAS DE USUARIO

HU1 – Recopilación y etiquetado de datos

Como Ingeniero de Datos

Quiero recolectar y etiquetar información proveniente de múltiples fuentes (APIs, archivos, bases de datos)

Para unificarla en un repositorio único con estructura estandarizada y facilitar su posterior análisis

Criterios de Aceptación:

- Debe ser posible configurar cada fuente de datos y su método de extracción.
- El sistema debe mostrar un reporte de la calidad de las etiquetas generadas (ej. porcentaje de campos completos, precisión de la anotación).
- Los datos deben ser accesibles para otros servicios únicamente cuando la etiquetación esté completa.

HU2 – Normalización e integración de datos

Como Data Scientist

Quiero normalizar e integrar los datos provenientes de distintos orígenes

Para asegurar que se encuentren en un formato uniforme y listo para su uso en proyectos de analítica e IA

Criterios de Aceptación:

- El sistema debe permitir definir reglas de normalización (unidades, formato de fechas, codificación).
- Se debe generar un log que registre cualquier conflicto o incompatibilidad de datos.
- El proceso de integración debe asegurar la consistencia entre diferentes tablas/columnas.

HU3 – Anonimización y tokenización

Como Responsable de Seguridad

Quiero anonimizar y tokenizar la información sensible (p. ej. datos de pacientes)

Para cumplir con las normativas de protección de datos y salvaguardar la privacidad de las personas

Criterios de Aceptación:

- Los campos sensibles deben ser transformados en tokens que no permitan la identificación directa de la persona.
- El proceso debe cumplir con estándares de cumplimiento (HIPAA, según corresponda).

HU4 – Verificación y control de calidad

Como Auditor de Datos

Quiero verificar la calidad y completitud de los datos anonimizados

Para asegurar que cumplan los criterios mínimos requeridos y evitar inconsistencias en modelos de IA

Criterios de Aceptación:

- El sistema debe permitir configurar reglas de validación (valores esperados, rangos, formatos).
- Se debe generar un informe que detalle los registros o campos que no cumplan con los criterios de calidad.
- Solo se aprobarán los datos que superen el umbral mínimo de calidad establecido.

HU5 – Acceso a datos anónimos

Como Usuario del Servicio Data Partnership

Quiero acceder a un repositorio con datos anónimos y ya normalizados

Para utilizarlos en proyectos de desarrollo o investigación de IA sin exponer información sensible

Criterios de Aceptación:

- El repositorio debe requerir autenticación y autorización basadas en roles.
- El sistema debe registrar en un log todas las descargas y consultas realizadas.
- El acceso a los datos debe otorgarse de forma granular según las políticas de la organización.

HU6 – Solicitud de cuentas en STA

Como Cliente

Quiero solicitar la creación de cuentas en la plataforma STIA para mi equipo de desarrolladores

Para que puedan tener acceso controlado a los conjuntos de datos requeridos y a las herramientas de IA

Criterios de Aceptación:

- La plataforma debe tener un flujo de aprobación de cuentas (pendiente, aprobado, rechazado).
- Se deben enviar notificaciones al solicitante cuando la cuenta esté lista o rechazada.

- La plataforma debe registrar el responsable que aprueba o rechaza la solicitud.

HU7 – Gestión de credenciales y aprobaciones

Como Administrador de la Plataforma

Quiero gestionar las credenciales de usuario y las aprobaciones de acceso

Para garantizar que únicamente el personal autorizado pueda descargar y utilizar los datos en STIA

Criterios de Aceptación:

- Debe existir un panel de administración para asignar y revocar permisos.
- El sistema debe registrar fecha, hora y razones de cada cambio en los permisos.
- Solo los administradores con un nivel de autorización específico pueden aprobar accesos a datos sensibles.

HU8 – Búsqueda y descarga de dataset

Como Data Scientist en una empresa de IA (cliente)

Quiero filtrar imágenes de resonancias cerebrales con diagnóstico de “tumor benigno”

Para descargar un dataset equilibrado para mi modelo de clasificación

Criterios de Aceptación:

- La búsqueda debe permitir filtrar por modalidad (MRI), región anatómica (cerebro) y patología (benigno).
- El sistema debe notificar cuando el dataset (ZIP o JSON) esté listo para su descarga.
- El dataset debe indicar la proporción de casos (balance) y un reporte de metadatos (número de imágenes, resolución, etc.).

HU9 – Notificaciones de disponibilidad de datos

Como Administrador del Sistema

Quiero notificar por correo electrónico o mensaje interno al cliente/desarrollador cuando sus datos estén listos

Para que puedan iniciar su proceso de entrenamiento o análisis de inmediato

Criterios de Aceptación:

- El sistema debe enviar un correo/mensaje con el enlace de descarga y credenciales si aplica.

- Se debe registrar la fecha y hora en que se envía la notificación.
- El mensaje debe incluir instrucciones breves sobre cómo acceder y usar los datos.

HU10 – Entrenamiento de modelos con datos anonimizados

Como Especialista en IA

Quiero entrenar mis modelos con datos ya anonimizados y verificados

Para respetar la privacidad de las personas y generar resultados confiables y reproducibles

Criterios de Aceptación:

- Los datos disponibles para entrenamiento deben haber pasado por los procesos de anonimización y verificación.
- El sistema debe permitir integrar estos datos con librerías o plataformas de IA (TensorFlow, PyTorch, etc.).
- Debe existir un registro de qué datasets se han utilizado en cada experimento para trazabilidad y auditoría.

LENGUAJE UBICUO

El lenguaje ubicuo se trabajó sobre dos grandes flujos de trabajo.

- Servicio “Data Partnership”:

Comienza con la preparación y etiquetado de datos de distintas fuentes (incluyendo revisión, integración y limpieza). Luego se realizan procesos como normalización, anonimización y verificación de datos para salvaguardar la privacidad (por ejemplo, anonimizar pacientes o tokenizar información sensible). Finalmente, estos datos (ya estandarizados y seguros) se ponen a disposición para su uso en otros entornos, como entrenamientos de modelos o análisis avanzados.

- Servicio “Desarrolladores de IA”:

Parte de la solicitud del cliente para acceder a distintos tipos de datos (pueden ser datos para entrenamiento inicial, datos enterprise, etc.). Se describe la creación de cuentas y el flujo de aprobación para que los desarrolladores obtengan los datos requeridos de manera segura y conforme a las políticas establecidas. Incluye pasos de interacción con el cliente (enviar correos, compartir enlaces de descarga, proporcionar credenciales, etc.), hasta llegar a la ejecución de modelos o pruebas, y la notificación de resultados.

Para su definición se trabajó haciendo uso de MIRO como herramienta:

https://miro.com/app/board/uXjVLLSoZks=

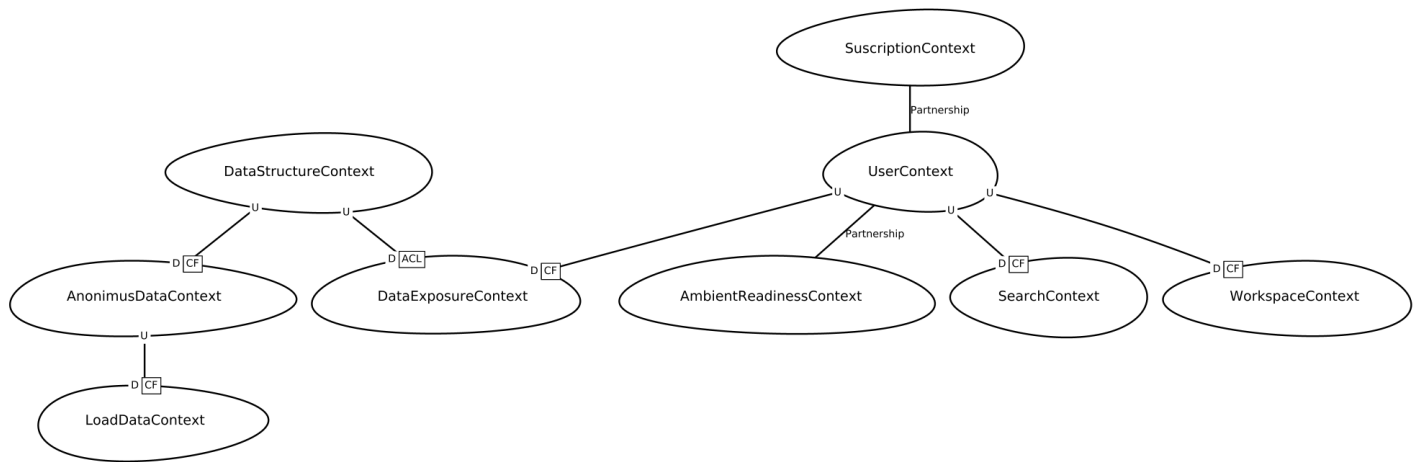
Lenguaje Ubicuo



CONTEXTOS ACOTADOS

A partir de los resultados obtenidos, se han identificado y delimitado diversos Contextos Acotados junto con sus interacciones, para lograr una comprensión más clara y organizada de la arquitectura de nuestro sistema. En las secciones siguientes se describen estos contextos haciendo uso de la herramienta ContextMapper y su DSL, con el fin de brindar una visión integral y coherente de la solución.

AS-IS: https://github.com/elle184/SaludTechDeLosAlpes-Error404/blob/master/src/main/cml/domain/SaludTechContext_AS-IS.cml



TO-BE: <https://github.com/elle184/SaludTechDeLosAlpes-Error404/blob/master/src/main/cml/domain/SaludTechContext.cml>

