

MetaMapa: Sistema de gestión de mapeos colaborativos



Trabajo Práctico Anual Integrador
-2025-

Contexto general

Contexto

La inteligencia colectiva permite solucionar problemas a partir del aporte de cada individuo de información de la que dispone. Sin embargo, para que esta sea confiable, es importante poder ordenarla, etiquetarla y ubicarla en tiempo y espacio. Esto permitirá, por ejemplo, en una ubicación determinada, reportar focos de incendios y prevenirlos, visibilizar causas, identificar focos de contaminación o aportar información valiosa para resolver desapariciones forzadas. Para estos fines, contar con información multimedia adecuadamente organizada en tiempo y geográficamente puede ser fundamental.

Nuestro sistema

MetaMapa será un Sistema de código abierto encargado por una ONG para la recopilación, visibilización y mapeo colaborativo de **información**. *MetaMapa* estará diseñado para que otras ONG, universidades u organismos estatales puedan instalarlo en sus servidores, gestionarlo y ofrecerlo para sus comunidades, en forma de **instancias particulares**. El sistema debe rendir cuentas¹ a la sociedad toda, maximizando la disponibilidad y veracidad de la información, protegiendo la identidad de quienes lo visitan y suben información. Por eso, *MetaMapa*, más allá de estadísticas básicas de uso, no recopilará datos de visitantes, es decir, quienes ingresan al Sistema únicamente para consultar la información disponible.

Con el objetivo de garantizar la disponibilidad de información y mitigar fallas de infraestructura y posibles ataques deliberados, esta se almacenará de forma descentralizada: cada **instancia** de *MetaMapa* podrá contar con una o más **fuentes de datos**, servidas desde diferentes nodos, en las que se almacenará la información disponible en sí. Esta información podrá ser redundante, y es uno de los desafíos de *MetaMapa* proveer herramientas de **etiquetado** que permitan vincular **la información** y generar consenso sobre la misma.

Algunas de estas **fuentes de datos** serán **estáticas**, es decir, proveerán información obtenida directamente de **lotes de datos** (*datasets*) bien conocidos. Sin embargo, *MetaMapa* también permitirá que cualquier persona mayor de edad pueda subir piezas de información (llamadas **hechos**), ya sean descripciones textuales, imágenes, audios o videos. Para esto, también se desarrollarán fuentes de datos **dinámicas**, que permitirán que las personas carguen colaborativamente nuevas piezas de información.

En este marco, diseñaremos e implementaremos:

- una **f fuente estática**: un servicio de solo lectura para publicar información estática, utilizando *datasets* que provengan de archivos u otras fuentes de datos provistas por otras ONGs,
- una **f fuente dinámica**: un servicio propio para la subida colaborativa de información, por perfiles anónimos y/o registrados,
- una **f fuente proxy**: integraciones con servicios de **fuentes de datos** y *datasets* provistos por otras ONGs.
- un servicio de **agregación**, que se encargará de consultar las distintas **fuentes de datos** y combinar sus datos,
- un servicio de **visualizaciones y estadísticas**, que podrán ser consultadas tanto online a través de una interfaz Web como consumidas por otras ONGs,
- un módulo para la **recepción de denuncias** sobre el contenido y protección de datos personales.

¹ https://www.researchgate.net/publication/221248162_Designing_for_accountability

A modo orientativo, se tiene este diagrama de despliegue (Fig. 1) como contexto general del sistema. El mismo no es definitivo y sufrirá modificaciones conforme avance el desarrollo del sistema. Además, no todas las instituciones tendrán por qué seguir esta misma estructura de despliegue, ya que la podrán ajustar a sus necesidades y a las de sus comunidades. Por ejemplo, algunas instituciones podrían exponer múltiples fuentes estáticas o consumir múltiples datasets, o sólo exponer una fuente dinámica.

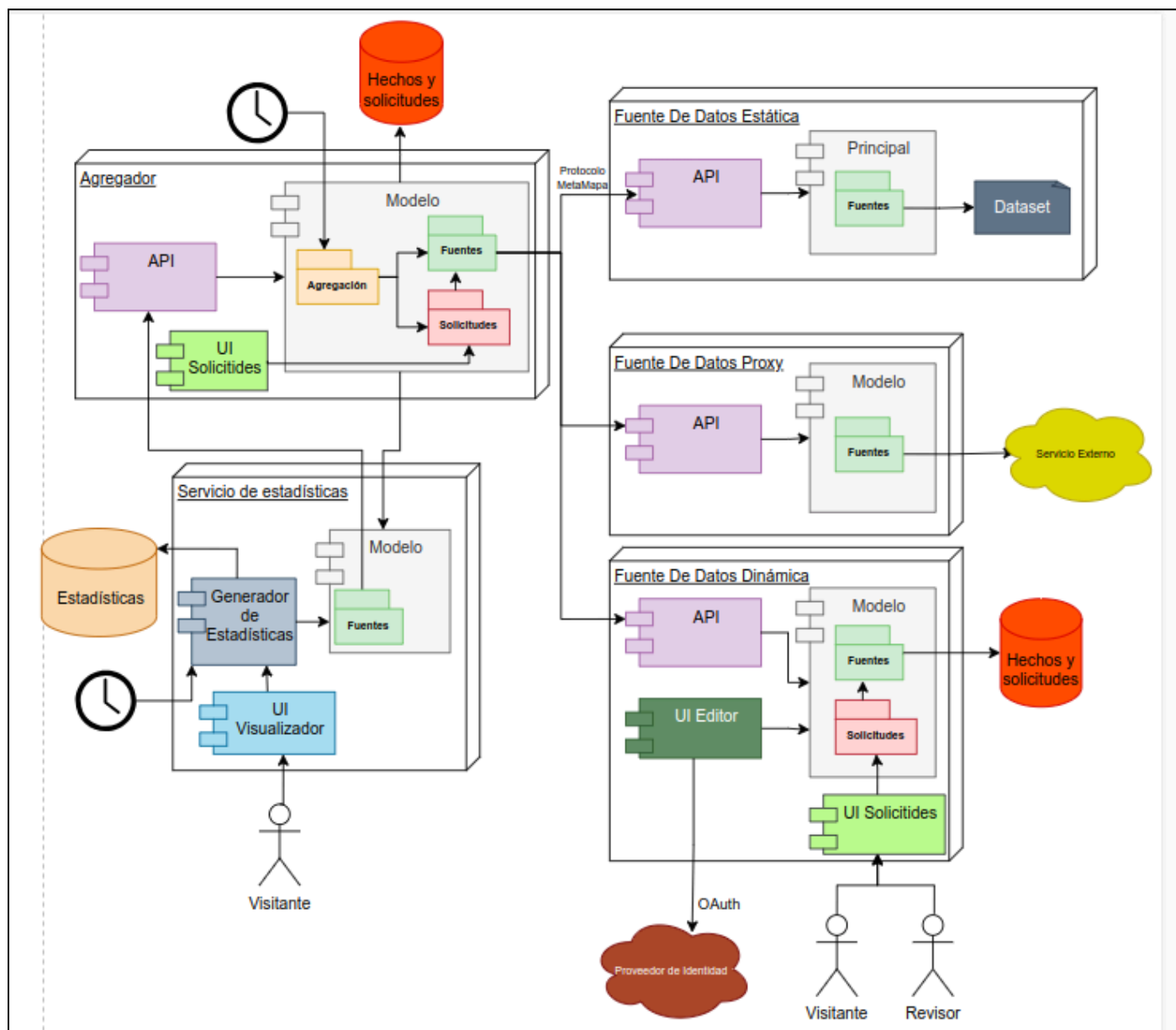


Figura 1 - Diagrama de Despliegue Inicial.



Entregas

El trabajo se encuentra estructurado en 6 entregas, algunas orientadas específicamente a la inclusión de funcionalidades, mientras que otras se abocarán a la inclusión de algunos aspectos del diseño y herramientas tecnológicas para la implementación del mismo.

Las entregas previstas se muestran a continuación, aunque pueden sufrir algunas modificaciones en su alcance o fechas: (*)

Nro.	Título	Semana de entrega propuesta
1	Arquitectura y Modelado en Objetos, Parte I	Semana del 1 de Mayo
2	Arquitectura, y Modelado en Objetos, Parte II	Semana del 1 de Junio
3	Integración, y Modelado en Objetos, Parte III	Semana del 7 de Julio
4	Persistencia y Análisis de Datos	Semana del 14 de Septiembre
5	Arquitectura Web MVC y Maquetado de Interfaz de Usuario	Semana del 21 de Octubre
6	Despliegue, Observabilidad y Seguridad	Semana del 28 de Noviembre

(*) - Estas fechas son a modo de referencia, consultar con el cronograma del curso asignado.

Glosario

- **Colecciones:** conjuntos de hechos organizados bajo un título y descripción, creados y gestionados por administradores. Son públicas y no pueden ser editadas ni eliminadas manualmente.
- **Contribuyentes:** personas que suben hechos a la plataforma, ya sean anónimas o registradas con datos personales.
- **Dataset:** lote de datos bien conocidos y usualmente en formato de archivo CSV desde el cual la plataforma extrae hechos.
- **Datos dinámicos:** información proveniente de contribuciones colaborativas, como descripciones textuales, imágenes, audios o videos, que crece en cantidad a lo largo del tiempo.
- **Datos estáticos:** información proveniente de archivos o bases de datos conocidas, que usualmente no cambian a lo largo del tiempo.
- **Etiqueta:** metadato asignado a un hecho para facilitar su organización y filtrado.
- **Fuentes de datos:** orígenes de la información en el sistema, pudiendo ser estáticas (datasets), dinámicas (aportadas por usuarios) o intermediarias (**proxy**, integraciones con otras instancias de MetaMapa u otros sistemas)
- **Hecho:** pieza de información registrada en el sistema, con un título, descripción, categoría, ubicación, fecha y origen.
- **Instancia:** un caso particular de despliegue de los servicios de MetaMapa en los servidores de una ONG, estado o universidad. Este despliegue podría realizarse a través de uno o más nodos, siguiendo la arquitectura orientativa de la Figura 1, o bien podría seguir otra estructura, según las necesidades de la institución.
- **Nodos:** cada una de las computadoras que procesan, exponen y almacenan información en el sistema.
- **Servicios:** cada uno de los procesos software que procesan, exponen y almacenan información en el sistema y se ejecutan en un nodo.
- **Pieza de información:** cada una de las unidades discretas de información que una fuente expone (y que en el caso de las fuentes dinámicas se pueden cargar colaborativamente). En MetaMapa, todas las piezas de información representan hechos ubicados en un tiempo y espacio determinado.
- **Servicio de agregación:** servicio encargado de combinar datos de distintas fuentes para generar una visión integrada de la información. Los servicios de agregación exponen la información siguiendo los mismos protocolos y formatos que las demás fuentes, por lo que también pueden ser utilizados como fuentes.
- **Sistema descentralizado:** arquitectura que almacena información en distintos nodos para mejorar la disponibilidad y seguridad de los datos.

ENTREGA 1: Arquitectura y Modelado en Objetos - Parte I

Objetivos de la entrega

- Entrar en contacto con el dominio y sus principales abstracciones.
- Incorporar de forma paulatina conceptos y principios de Diseño.
- Familiarizarse con el entorno de desarrollo y las principales tecnologías a ser aplicadas a lo largo del Trabajo Práctico.
- Familiarizarse con la arquitectura del sistema.

Unidades del Programa Vinculadas

- Unidad 2: Herramientas de Concepción y Comunicación del Diseño
- Unidad 3: Diseño con Objetos
- Unidad 6: Diseño de Arquitectura
- Unidad 8: Validación del Diseño

Alcance

- Fuentes estáticas.
- Colecciones.
- Hechos: listado, filtrado y búsqueda.
- Visualizadores y Contribuyentes de hechos.
- Solicitudes de eliminación.

Dominio

Fuentes estáticas

Tal como se mencionó en el detalle de nuestro sistema, se tienen fuentes de datos de tipo estáticas, provenientes de lotes de datos bien conocidos. Para esta iteración, se requiere diseñar e implementar el componente que posibilite la lectura de estos datasets y que extraiga los hechos de los mismos. En esta primera iteración estaremos incorporando un lote de datos estático de tipo archivo .csv.

Colecciones

Las **colecciones** representan conjuntos de **hechos**. Las mismas pueden ser consultadas por cualquier persona, de forma pública, y no pueden ser editadas ni eliminadas manualmente (esto último, con una sola excepción, ver más adelante).

Las **colecciones** tienen un título, como por ejemplo “Desapariciones vinculadas a crímenes de odio”, o “Incendios forestales en Argentina 2025” y una descripción. Las personas administradoras pueden crear tantas **colecciones** como deseen.

Las **colecciones** están asociadas a una **fuentes** y tomarán los **hechos** de las mismas: para esto las colecciones también contarán con un **criterio de pertenencia** configurable, que dictará si un hecho pertenece o no a las mismas. Por ejemplo, la colección de “Incendios forestales...” deberá incluir

automáticamente todos los **hechos** de **categoría** "Incendio forestal" ocurrido en Argentina, acontecido entre el 1 de enero de 2025 a las 0:00 y el 31 de diciembre de 2025 a las 23:59.

Hechos

Cada **hecho** representa una pieza de información, la cual debe contener mínimamente: título, descripción, categoría, contenido multimedia opcional, lugar y fecha del acontecimiento, fecha de carga y su origen (carga manual, proveniente de un dataset o provisto por un **contribuyente**). Idealmente, todos estos campos podrán ser utilizados por cualquier **visitante** como filtros de búsqueda desde la interfaz gráfica y como **criterios de pertenencia** a una colección.

Hasta el momento, existen dos tipos de hechos: de texto y con soporte para contenido multimedia. Para esta primera iteración, solo se requerirá dar soporte al primero, pero es importante contemplarlo en el diseño de la plataforma.

Cualquier **hecho** provisto por **contribuyentes** de la plataforma podrá ser sometido a etapas de etiquetado y revisión manual por parte de las personas administradoras (lo cual será desarrollado en entregas venideras).

Obtención de hechos en lote por archivos CSV

Si bien una **fuentes de datos** estática podría venir de cualquier tipo de dataset, se pide dar soporte al acceso de hechos mediante archivos .csv. Las piezas de información provenientes de este tipo de archivos se leerán directamente y no se traerán a memoria. En esta etapa, se considerará que un **hecho** está repetido si el "título" es el mismo. De ser así, se pisarán los atributos del existente. La ubicación del hecho vendrá determinada por su latitud y longitud, para poder representarla en un mapa en entregas posteriores.

Se deberá leer este tipo de archivos teniendo en cuenta este formato:

Título	Descripción	Categoría	Latitud	Longitud	Fecha del hecho

Visualizadores y contribuyentes de hechos

Se considera **visualizador** a toda persona humana anónima que desee ingresar a la plataforma para visualizar y utilizar la información disponible en el Sistema. No requerirá identificarse, y podrá subir hechos si así lo quisiera manteniendo su anonimato. No se recopilarán (ni mostrarán) datos personales.

En cambio, si lo desearan podrían darse a conocer y subir hechos a los que se les asocie algunos datos personales, como nombre, apellido y edad. Únicamente el nombre sería el campo obligatorio. En esta etapa, tampoco se mostrarán de forma pública datos personales de las personas identificadas, pero sí se almacenarán y quedarán a disposición de las personas administradoras.

Consideraremos a las personas humanas (ya sean anónimas o identificadas) que han aportado alguna pieza de información como **contribuyentes**.

Solicitudes de eliminación

Cualquier **hecho** del sistema debe admitir solicitudes de eliminación, para aquellas situaciones adecuadamente fundadas (por ahora, mediante un texto de al menos 500 caracteres) en que se deba eliminar del sitio la información, aún cuando esté en una fuente.

Cuando estas **solicitudes** sean aceptadas, el hecho se mantendrá internamente en el sistema, pero ya no se mostrará en ninguna colección y no se lo incorporará nuevamente si una fuente lo vuelve a retornar.

Naturalmente muchas **solicitudes** serán *spam*, por lo que las mismas quedarán pendientes hasta que una persona administradora las acepte o rechace (eliminándose por completo).

Requerimientos

Requerimientos de dominio

1. Como persona **administradora**, deseo crear una colección.
2. Como persona **administradora**, deseo poder importar hechos desde un archivo CSV.²
3. Como persona **visualizadora**, deseo navegar todos los hechos disponibles de una colección.
4. Como persona **visualizadora**, deseo navegar los hechos disponibles de una colección, aplicando filtros.
5. Como persona **contribuyente**, deseo poder solicitar la eliminación de un hecho.
6. Como persona **administradora**, deseo poder aceptar o rechazar la solicitud de eliminación de un hecho.

Entregables

1. **Modelo del Dominio:** diagrama de clases inicial que contemple las funcionalidades requeridas.
2. **Diagramas de Arquitectura:** diagrama de despliegue, componentes y/o cualquier otro tipo de diagrama que refleje la arquitectura física y lógica de la entrega actual.
3. **Justificaciones de Diseño Iniciales.**
4. **Diagrama General de Casos de Uso.**
5. **Implementación** de los requerimientos.

Fuentes Ejemplo

- [Hechos de Desastres Naturales](#)
- [Hechos de Incendios Forestales en España](#)
- [Hechos y víctimas de muertes viales en la República Argentina](#)
- [Hechos de personas asesinadas por el estado en la República Argentina](#)

² Se debe utilizar una fuente que considere más de 10.000 hechos en cada importación. Las fuentes propuestas son de ejemplo, se puede proponer una propia.