

Acta de Reunión – INVIAS (Arquitectura y Datos)

Código: ACTA-INVIAS-2025-09-17

Proyecto: INVIAS – Arquitectura de datos y despliegue (Mapa de Vulnerabilidad Faunística)

Fecha: miércoles, 17 de septiembre de 2025

Hora: 11:00 – 12:00

Lugar: Reunión virtual (Google Meet)

1. Propósito de la reunión

Alinear la arquitectura tecnológica, flujos de datos y responsabilidades para el desarrollo del sistema de análisis y visualización del Mapa de Vulnerabilidad Faunística, incluyendo el uso de ArcGIS Enterprise del INVÍAS, definición de endpoints por modelo (MaxEnt como inicial), aprovechamiento de Google Earth Engine (GEE) como fuente secundaria de datos satelitales, y la integración de fuentes internas (SUKUBUN) y externas (NOAA, SIATA, IDEAM).

2. Asistentes

- Luis Esteban Gómez Cadavid
- Andrés Felipe Hernández Marulanda
- Jairo Iván Coy Coy
- Nelson Aníbal Miranda Ríos
- Jasmín Marín Perez

3. Desarrollo de la reunión (narrativa)

La sesión se centró en precisar la arquitectura sobre la cual correrán los procesos de datos y la publicación cartográfica. Se ratificó que, por requerimiento del cliente, la infraestructura de despliegue será ArcGIS, utilizando específicamente el ArcGIS Enterprise y su Portal, de modo que los geoprocursos puedan ejecutarse en la nube y verse en ArcGIS Online. Se dejó constancia de que las credenciales deben ser provistas por INVÍAS para garantizar alineación y gobernanza sobre los recursos y servicios.

En cuanto al pipeline, se planteó una estructura multicapa donde el equipo de arquitectura entrega un mapa base y la contraparte alimenta el sistema por medio de endpoints. Se discutió si consolidar resultados en un único endpoint o segmentarlos. Dado que los insumos clave son matrices/raster pesados (imágenes satelitales, índices

espectrales, modelos de terreno), se acordó que lo más eficiente es definir endpoints separados por modelo o proceso para evitar cuellos de botella de rendimiento.

Se acordó que el primer modelo operativo será Máxima Entropía (MaxEnt), apoyado en insumos satelitales producidos/automatizados con GEE y datos de SUKUBUN (avistamientos y atropellamientos) procesados por el equipo de modelado. GEE se usará para extraer y preprocesar colecciones (índices espectrales, temperatura superficial, radar), depositando resultados en repositorios organizados por zona/periodo; las capas estáticas (p. ej. coberturas IDEAM y MDT) se gestionarán aparte. El resultado esperado hacia el cliente puede publicarse en raster (GeoTIFF/PNG) o vectorial (SHP/GeoJSON) según defina el equipo de modelado, manteniendo interoperabilidad con ArcGIS.

Un punto operativo clave fue la disciplina de gestión en GitHub: cada necesidad técnica (librerías, conectores, tokens) debe registrarse como issue en el tablero de 'Librería de modelos y procesamiento'. Se reiteró la necesidad de crear una cuenta específica para GEE y de listar librerías requeridas (incluyendo compatibilidad para R, a solicitud de modelado). También se mencionó la verificación de fuentes externas (NOAA, SIATA) y la confirmación del vínculo a SUKUBUN, sistema propiedad de INVÍAS.

Finalmente, se acordó el esquema de sprints y checkpoints. Un micro-sprint de seguimiento los lunes a las 10:00 (15 min) y una reunión general los jueves a las 10:00 con periodicidad 'cada 8 días'. Se dejó como tarea presentar el lunes 22 de septiembre un script o, en su defecto, una esquematización del flujo de procesos y datos dentro de la arquitectura propuesta. Se registrará y dará trazabilidad en el tablero de gestión de equipos.

4. Agenda simplificada

- Arquitectura de despliegue: ArcGIS Enterprise/Portal y ArcGIS Online (credenciales INVÍAS).
- Diseño del pipeline y estrategia de endpoints por modelo (prioridad: MaxEnt).
- Fuentes de datos: GEE (dinámico), SUKUBUN (eventos), NOAA/SIATA/IDEAM (externas), capas estáticas.
- Resultados y formatos: raster (GeoTIFF/PNG) y/o vector (SHP/GeoJSON).
- Gestión en GitHub: issues por librerías, conectores y tokens; cuenta específica para GEE.
- Plan de sprints y fechas próximas (lunes 22/09 10:00; jueves 'cada 8 días' 10:00).

5. Revisión de compromisos pasados

- Instalación de librerías solicitadas para R/Python por el equipo de modelado: en curso; compromiso de cierre al finalizar la semana posterior a esta reunión.

- Tablero de 'Gestión de equipos' y 'Librería de modelos y procesamiento' en GitHub: operativo; pendiente que cada integrante cree issues con sus requerimientos específicos.

6. Próximos pasos y fechas

- INVÍAS: habilitar acceso y credenciales de ArcGIS Enterprise/Portal para el equipo (urgente).
- Arquitectura (Jairo Iván Coy Coy): capacitación acelerada en Google Earth Engine; presentar script o esquema del flujo de procesos el lunes 22/09/2025 a las 10:00.
- Modelado (Nelson Aníbal Miranda Ríos y Jasmín Marín Perez): confirmar oficialmente el uso de MaxEnt como único modelo inicial y detallar insumos mínimos (capas, ventanas temporales, buffer de vías).
- Datos satelitales (Luis Esteban Gómez Cadavid): automatizar extracción con GEE y depositar en repositorio por zona/mes; documentar colecciones y parámetros.
- Integración: definir formatos de intercambio por endpoint (p. ej., GeoTIFF/PNG para raster; SHP/GeoJSON para vector).
- Gestión GitHub (Andrés Felipe Hernández Marulanda): solicitar a todo el equipo registrar issues de librerías/tokens y fuentes (NOAA, SIATA, SUKUBUN).
- Ritmo de trabajo: micro-sprint semanal los lunes 10:00 (15 min); reunión general los jueves 10:00 'cada 8 días'.

7. Hitos y conclusiones

- Se utilizará la infraestructura ArcGIS Enterprise/Portal del INVÍAS; las credenciales deben ser provistas por el cliente.
- Decisión técnica: endpoints separados por modelo/proceso para evitar saturación y mejorar rendimiento.
- Modelo inicial oficial: MaxEnt, con arquitectura abierta para futuros modelos.
- GEE se adopta como fuente secundaria de datos satelitales y motor de preprocesamiento automatizado.
- Confirmada integración con SUKUBUN (eventos) y consideración de NOAA, SIATA e IDEAM.
- Establecido esquema de sprints y entrega temprana de script/esquema el lunes 22/09/2025 a las 10:00.

8. Análisis y recomendaciones

- La decisión de separar endpoints por modelo es adecuada para cargas matriciales intensivas; reduce latencia y riesgos de 'throttling'. No obstante, se recomienda normalizar nomenclatura y convenciones (nombres de capas, versionado, calendario de cortes) y fijar límites de tamaño por endpoint.
- Para ArcGIS Enterprise, conviene gestionar roles/perfiles y cuotas (almacenamiento, CPU) y habilitar logging de servicios. Definir contrato de interfaz por tipo de salida (raster/vector) y metadatos mínimos (CRS, resolución, fecha de corte, ventana temporal).
- En GEE, documentar colecciones, filtros temporales y máscaras; parametrizar fechas (END=fecha de sistema; START=END-1 año) cuando aplique, y separar capas estáticas en repositorios controlados. Automatizar exportes y control de calidad.
- Respecto a SUKUBUN, al tratarse de un sistema de INVÍAS, acordar mecanismo de acceso (API vs. descarga), esquema de atributos y frecuencia de actualización. Verificar licencias de uso y protección de datos.
- En la gestión de proyecto, sostener la práctica de issues por requerimiento y CI/CD de scripts. Un diagrama ETL-E2E y una guía de 'ambiente mínimo' (Python/R, librerías, tokens) facilitarán la repetibilidad y el on-boarding.