**Consultoría**

**Diagnóstico, identificación de vacíos de información, brechas y oportunidades de mejora en la sistematización y gestión de datos nacionales de tráfico de vida silvestre en Ecuador**

**Producto 3:  
Análisis comparativo de fuentes de datos**

**Elaborado por:** Mat. Javier Núñez, MSc.  
**Fecha de entrega:** 05/05/2025  
**Proyecto y número POA:** GEF / POA 3.3.3

1. **Resumen Ejecutivo**

En el presente informe se describe el desarrollo del producto 3 perteneciente a la Fase 2 de la consultoría “Diagnóstico, identificación de vacíos de información, brechas y oportunidades de mejora en la sistematización y gestión de datos nacionales de tráfico de vida silvestre en Ecuador”.

Wildlife Conservation Society (WCS) promovió esta consultoría con la finalidad de diseñar y desarrollar un plan nacional de recopilación de información para recolectar, organizar, analizar y gestionar datos de manera eficiente, con el fin de apoyar la toma de decisiones informadas y coordinar esfuerzos entre múltiples actores involucrados en acciones de combate al tráfico de vida silvestre en el país.

Luego de la revisión de 15 archivos en formato .xlsx en los que constaban un total de 37 bases de datos de las instituciones Fiscalía General del Estado, Consejo de la Judicatura, Ministerio de Ambiente, Agua y Transición Ecológica (MAATE), Parque Nacional Galápagos (PNG) y Unidad Nacional de Policía de Protección del Medioambiente (UPMA), se procedió a realizar recomendaciones, tanto de fondo como de forma sobre cada base de datos.

Las recomendaciones tienen como primer objetivo facilitar el uso de las bases de datos, proponiendo nombres adecuados para las distintas variables, junto con una delimitación clara de su dominio de valores. Como segundo objetivo, se busca ampliar y mejorar la generación de información mediante recomendaciones acerca de la forma de alimentar la base de datos, enmarcadas en evitar el uso de celdas combinadas, definiendo previamente si cada registro en la base de datos es por evento o espécimen.

1. **Desarrollo**

**2.1 Actualización de diccionarios de variables**: En esta sección se detalla los cambios realizados sobre los diccionarios de variables de generados en el producto anterior de esta consultoría, en pos de facilitar su comprensión, uso y accesibilidad a grupo más amplio de usuarios. En la siguiente tabla se enumeran los diccionarios de variables generados por Fuente de información.

|  |  |  |  |
| --- | --- | --- | --- |
| Orden | Fuente | Nombre del archivo | Bases de datos en el archivo |
| 01 | Fiscalía General del Estado | Informe\_estadistico\_2025012322001116 FIscalia.xlsx | 1 |
| 02 | Consejo de la Judicatura | CJ 0936 Causas Art 247 corte agosto 2024(1).xlsx | 3 |
| 03 | MAATE/BIODOVERSIDAD | BD\_2021\_VS (1).xlsx | 2 |
| 04 | MAATE/BIODOVERSIDAD/CITES | Informe CITES 2024.xlsx | 3 |
| 05 | MAATE/BIODOVERSIDAD/CITES | Matriz permisos CITES 2022 y 2023.xlsx | 2 |
| 06 | MAATE/BIODOVERSIDAD | Consolidado 2020\_Retenciones y Rescates\_Zonales.xlsm | 2 |
| 07 | MAATE/BIODOVERSIDAD | Consolidado trafico 2014\_2019-1 FINAL.xls | 1 |
| 08 | MAATE/BIODOVERSIDAD | Rescate\_2022 (1).xlsx | 1 |
| 09 | MAATE/BIODOVERSIDAD | Rescates y Retenciones CONSOLIDADO NACIONAL\_2024.xlsx | 2 |
| 10 | MAATE/BIODOVERSIDAD | Retencion\_2022 (1).xlsx | 1 |
| 11 | MAATE/PROYECTO | BDD\_PSNCFVS\_2022\_2023\_2024\_WCS.xlsx | 3 |
| 14 | PNG | Procesos Judiciales Trafico de Vida Silvestre.ods | 1 |
| 16 | UPMA | FAUNA SILVESTRE CON CARNE DE MONTE.xlsx | 6 |
| 17 | UPMA | FORMATO DGO UPMA FLORA 2022, 2023 Y 2024.xlsx | 3 |
| 18 | UPMA | RESCATE DE FAUNA SILVESTRE 2019 AL 2024.xlsx | 6 |
| Total | | **15** | **37** |

Sobre cada uno de los diccionarios de variables se realizaron los siguientes procesos:

* Tratamiento del campo “Nombre de la variable” en función a las recomendaciones establecidas en la norma ISO/IEC 11179 (más detalles en la siguiente sección).
* Eliminación del campo “Unidad de medida”, esto debido a que en la gran mayoría de las bases de datos el valor para este campo constaba como *No aplica*, a excepción de las variables de ubicación geográfica. Para estos casos, se sugiere describir el Sistema de Referencia de Coordenadas como parte del campo *Descripción*.
* Generación de un código de color en los diccionarios en función al tratamiento sugerido para cada variable, donde:
  + Rojo: eliminar.
  + Amarillo: falta contexto o información para acabar de definir todas sus propiedades.
  + Verde: aumentar variable con la finalidad de resolver ambigüedades o mejorar la identificación única de cada registro de evento/espécimen.

**Tratamiento del campo “Nombre de la variable”**: la Parte 5 de la norma ISO/IEC 11179 incluye lineamientos específicos sobre cómo deben nombrarse las variables de una base de datos, estos pueden resumirse en los siguientes principios clave:

1. Estructura del nombre: se recomienda que los nombres de los elementos de datos sigan una estructura semántica y sean coherentes, usualmente basados en componentes como:

* Clase de objeto: el tipo de registro se está describiendo.
* Propiedad: qué atributo se está midiendo.

Ejemplo: *especie\_nombre*, *persona\_edad*.

1. Reglas para los nombres: algunas de las reglas recomendadas para los nombres son:

* Deben ser claros, precisos y no ambiguos.
* Deben reflejar el significado del dato.
* Se sugiere evitar abreviaciones poco conocidas.
* No deben contener información redundante.
* Deben permitir el uso automatizado evitando duplicidad, espacios o caracteres especiales.
* Pueden usar formatos del tipo CamelCase o snake\_case, manteniendo la consistencia en toda la base.

1. Identificadores vs. Nombres: la norma distingue entre:

* Nombre: nombre semántico legible, por ejemplo *nombre\_provincia*.
* Identificador: código único (a veces alfanumérico) que no cambia aunque el nombre cambie en el tiempo, por ejemplo *pro*.

Alternativa:

Nombres, Descripción e Identificador de variables.

Nombre: nombre semántico legible, por ejemplo “Nombre de provincia”.

Descripción: explicación de la variables, por ejemplo “Nombre de la provincia donde se realizó el rescate”.

Identificador: código único (a veces alfanumérico) que no cambia aunque el nombre cambie en el tiempo, por ejemplo pro, npro, i01, provincia.

Considerando todo lo expuesto anteriormente, se renombra el campo “Nombre de la variable” a “Nombre actual” y se añade el campo “Nombre propuesto” a todos los diccionarios de variables, en función a las recomendaciones descritas en la norma. Como resultado se obtiene la versión 2.0 del diccionario de variables, la cual consta de las siguientes variables:

|  |  |  |
| --- | --- | --- |
| Elemento | Descripción | Ejemplo |
| Nombre actual | Nombre presente en la base de datos | Especie |
| Nombre propuesto | Nombre propuesto para el campo analizado | especie |
| Definición | Explicación clara del concepto | Nombre científico de la especie traficada |
| Formato de Datos | Tipo de dato y restricciones | Cadena |
| Fuente de Datos | Origen de la variable | Base de datos CITES |
| Clasificación | Categorización de la variable según su función (identificación, geolocalización, transacción) | Variable de identificación |
| Ejemplo de valor | Muestra cómo se vería un dato real | *Panthera onca* (Jaguar) |
| Notas adicionales | Información extra relevante. | Si no se conoce el nombre científico usar la categoría "especie no identificada" |

**2.2. Análisis comparativo por fuente de información**: En esta sección se detallas las observaciones generales a ser implementadas sobre cada una de las bases de datos analizadas por fuente de información. Se ser necesario, se incluyen capturas de pantalla para proporcionar el contexto adecuado.

**Fiscalía general del estado**

1. Archivo: Informe\_estadistico\_2025012322001116 FIscalia.xlsx

* Base de datos cuenta con una buena estructura, los nombres de las variables son consistentes, presentando una misma estructura de construcción para todas las variables. Se recomienda revisar los dominios de valores de las variables categóricas, con la finalidad de estandarizar categorías (de ser necesario) y asegurarse que sean completos.

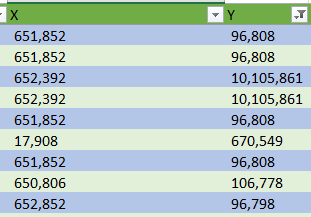
**Consejo de la Judicatura**

1. Archivo: CJ 0936 Causas Art 247 corte agosto 2024(1).xlsx

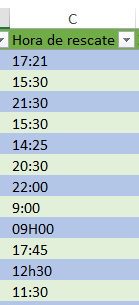
**Ministerio de Ambiente, Agua y Transición Ecológica**

1. Archivo: BIODIVERSIDAD\BD\_2021\_VS (1)

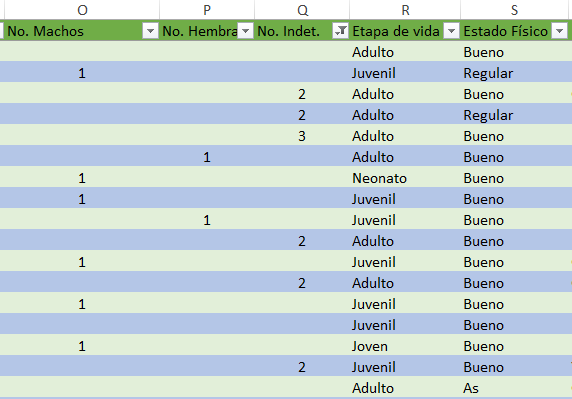
* Los nombres de la base de datos presentan espacios, caracteres especiales como puntos y espacios. Se sugiere seguir las recomendaciones asociadas a sus nombres de variables.
* No se especifica el Sistema de Referencia de Coordenadas (SRC) de las coordenadas presentes. Se sugiere incluir esa información en el campo *Definición* del diccionario de variables.



* Diferente formato al registrar la hora de rescate.



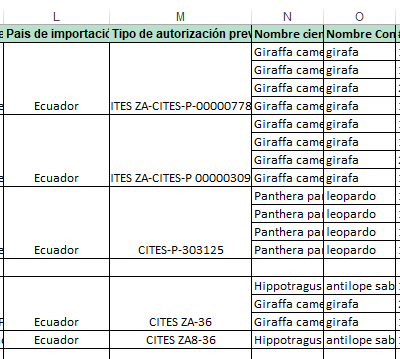
* En primera instancia, la base de datos estar a nivel de evento (rescate), sin embargo, existen variables dirigidas a espécimen, tales como “Etapa de vida” o “Estado Físico”. Es recomendable definir el nivel de la base de datos, y registrar la información acorde al nivel especificado.



* Se recomienda revisar los dominios de valores de las variables categóricas, con la finalidad de estandarizar categorías (de ser necesario) y asegurarse que sean completos.
* En los campos tipo cadena se observan caracteres no reconocidos en Excel (rombo negro con un signo de interrogación dentro), se sugiere definir un formato adecuado que impida la pérdida de información.

1. Archivo: BIODIVERSIDAD\CITES\ Informe CITES 2024
2. Archivo: BIODIVERSIDAD\CITES\Matriz permisos CITES 2022 y 2023

* Base de datos cuenta con una buena estructura, sin embargo los nombres de las variables presentan espacios y caracteres especiales. Se sugiere seguir las recomendaciones asociadas a sus nombres de variables.
* La base de datos presenta múltiples columnas combinadas



* Se recomienda revisar los dominios de valores de las variables categóricas, con la finalidad de estandarizar categorías (de ser necesario) y asegurarse que sean completos.

1. Archivo: BIODIVERSIDAD\Consolidado 2020\_Retenciones y Rescates\_Zonales
2. Archivo: BIODIVERSIDAD\Consolidado trafico 2014\_2019-1 FINAL
3. Archivo: BIODIVERSIDAD\Rescate\_2022 (1)
4. Archivo: BIODIVERSIDAD\Rescates y Retenciones CONSOLIDADO NACIONAL\_2024
5. Archivo: BIODIVERSIDAD\Retencion\_2022 (1)
6. Archivo: PROYECTO\BDD\_PSNCFVS\_2022\_2023\_2024\_WCS

**Parque Nacional Galápagos**

1. Archivo: CJ 0936 Causas Art 247 corte agosto 2024(1).xlsx

**Unidad Nacional de Policía de Protección del Medioambiente**

1. Archivo: FAUNA SILVESTRE CON CARNE DE MONTE
2. Archivo: FORMATO DGO UPMA FLORA 2022, 2023 Y 2024
3. Archivo: RESCATE DE FAUNA SILVESTRE 2019 AL 2024
4. **Conclusiones y Recomendaciones**

* La estructura de las bases de datos revisadas varía considerablemente dependiendo de la fuente de la información a la que pertenece. Por ejemplo, las bases de la Fiscalía General del Estado y Consejo de la Judicatura están mejor estructuradas, mientras que las bases del MAATE y UPMA necesitarán cambios tanto de forma como de fondo.

En la segunda versión de los diccionarios de variables se recogen recomendaciones para facilitar su comprensión, uso y accesibilidad a grupo más amplio de usuarios. Se sugiere acogerlas con la finalidad de establecer adecuadamente su nivel de integración.

* En este informe se presentan los primeros resultados del análisis de consistencia y duplicidad de las bases de datos, centrándose especialmente en la consistencia de la información presente en ellas, puesto que al ser llevadas en Excel, presentan celdas combinadas/vacías que dificultan su lectura, tratamiento y análisis. La determinación de duplicidad en las bases de datos no se puede determinar de manera clara, ya que, no se determina si los registros de las bases deben ser por ejemplo el evento (rescate) o cada uno de los individuos.
* Para lograr una estandarización de nombres y dominio de valores (categorías de las variables) comunes entre las diferentes bases de datos se sugiere utilizar las categorías descritas en CITES para las variables XXXX. Mientras que para las variables de ubicación *provincia*, *cantón*, *parroquia*, *distrito, circuito y subcircuito* se recomienda utilizar la información presente en la base XXXX.XXX que contiene la Distribución Político Administrativa del Ecuador al año 2022 y la división en distritos, circuitos y subcircuitos.
* Es imperativo mantener reuniones con las instituciones involucradas con la finalidad de discutir ciertos aspectos expuestos en este informe, tales como el definir el nivel de las bases de datos revisadas (evento o espécimen), al igual que en los diccionarios de variables propuestos.

SOLO DEBEMOS DEJAR COMO UNA CONCLUSIÓN O COMO UN PUNTO A DAR UN SIGUIENTE PASO.

1. **Próximos pasos**

* Es imperativo mantener reuniones con las instituciones involucradas con la finalidad de discutir ciertos aspectos expuestos en este informe, tales como el definir el nivel de las bases de datos revisadas (evento o espécimen), al igual que en los diccionarios de variables propuestos.