



Informe sobre la adaptación/integración de los datos de la web Armador (Visor de Mareas) de la Secretaría General de Pesca en la base de datos de entrada al Geoport

Proyecto GEOCAP

Este informe corresponde a la fuente de verificación FV2 de la actividad A2 del proyecto GEOCAP y ha sido realizado por:

Instituto de Investigaciones Marinas (CSIC)

con la participación de:

Fundación Pública Gallega Centro Tecnológico de Supercomputación de Galicia (CESGA)

Sealab Solutions S.L.

Instituto Español de Oceanografía (IEO)

Cofradía de pescadores de Sanlúcar

Organización de Productores Pesqueros de Lugo (OPP-07)

Organización de Productores Pesqueros de Carboneras (OPP-66)

Este estudio se ha desarrollado en el marco de la actividad 2 del proyecto GEOCAP cofinanciado por el FEMP



VICEPRESIDENCIA
CUARTA DEL GOBIERNO
MINISTERIO
PARA LA TRANSICIÓN ECOLÓGICA
Y EL RETO DEMOGRÁFICO



Unión Europea
Fondo Europeo Marítimo y
de Pesca (FEMP)



1.- Introducción

En el presente informe se detallan las actividades realizadas para conseguir la transferencia efectiva de datos de capturas desde la web Armador a la base de datos de entrada en el Geoportal.

Por una parte, se desarrolló una interfaz de programación de aplicaciones (en adelante DEO) para capturar los datos de las mareas de los barcos de las entidades pesqueras colaboradoras. Tras ello, se seleccionaron aquellos campos que se determinaron como relevantes para conseguir los objetivos del proyecto y se almacenaron en un fichero de datos con formato csv.

Por otra parte, se desarrolló una nueva estructura de la base de datos de entrada al Geoportal que permite cargar directamente los datos del fichero csv.

2.- Descripción de la herramienta DEO

El software DEO, en sus iniciales DIARIO ELECTRÓNICO PARA ORGANIZACIONES, es una herramienta que capacita a las Organizaciones de Productores para que, de manera automática y por medios digitales, la Organización acceda de manera rápida a informes sobre:

- Control de las capturas de los miembros de la OPP.
- Consulta de mareas realizadas.
- Control de las cuotas de cada stock tanto por parte de la OPP como por parte de cada armador individual a través de una App.

A nivel de arquitectura de programación, el software consta de un conjunto sofisticado de aplicaciones, las cuales se representan a continuación:

Arquitectura DEO

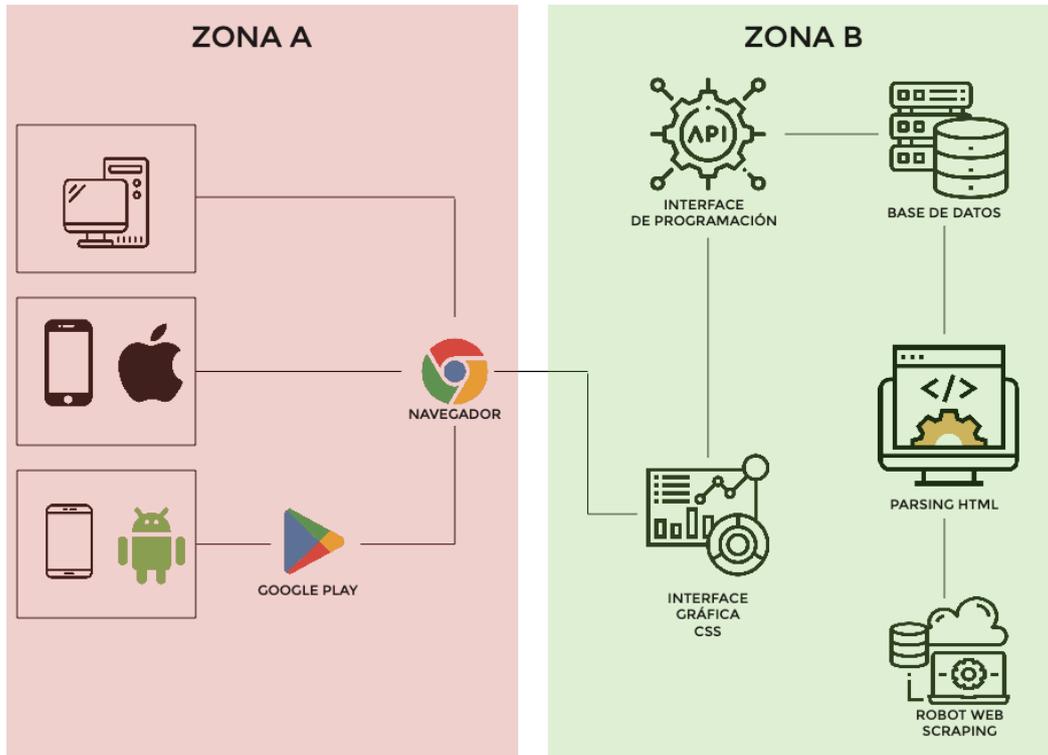


Ilustración 1. Representación de las aplicaciones del software

En la ZONA A básicamente el desarrollo es sencillo, es decir, se trata del hardware que se puede utilizar para conectarse a la 'interface gráfica CSS'. Tanto en el caso de PC, como iPhone, como Android, no existe ninguna aplicación nativa, todo se visualiza a través de navegador. Esto quiere decir que:

El PC NO UTILIZA SOFTWARE DE SOBREMESA, sino que se conecta a través de internet al resto de software (ZONA B) alojado en dos servidores diferentes debido a sus requerimientos técnicos.

En términos sencillos podríamos decir que al software se accede a través de una página web común al resto de usuarios, en la cual la OPP se logea para que se cargue SU APLICACIÓN ESPECÍFICA.

Este portal de acceso, representado en el gráfico como 'interface gráfica CSS', es el punto de acceso común para todas las Organizaciones en las que se ha desarrollado el programa, razón por la cual, el año que figura en la web es el año en la que se creó.



En los tres casos de uso, la puerta de acceso (ZONA A) es idéntica para todos los usuarios del DEO, razón por la que las fechas corresponden al momento en que se creó el acceso.

La ZONA B se compone de diferentes herramientas software pero unidas. El proceso comienza con el robot de scrapeo que se descarga las mareas ubicadas en 'Visor de Mareas' en formato html a través de la tecnología 'web scraping'. Las mareas que se descargan son las pertenecientes a los barcos de las OPPs adscritas al servicio, lo cual se hace a través de login. A continuación, se deben "parsear" todas las mareas, para su integración en la base de datos, para que a continuación se pueda trabajar con los datos.

En el presente caso del proyecto GEOCAP, se consulta de la base de datos del DEO la información de capturas y descartes geolocalizables de todas las mareas realizadas por los barcos de las organizaciones pesqueras colaboradoras del proyecto. Una vez seleccionados los datos de interés para el proyecto, éstos se pasan a un archivo CSV (valores separados por comas) que queda disponible en una dirección web a la que se accede a través de unas credenciales.

Con respecto al contenido del archivo csv, en la siguiente tabla vemos un ejemplo:

cod_buque	arte	cod_marea	puerto_llegada	fecha_inicio	fecha_fin	fecha_actividad	cod_bao	peso	num_espumplero	latitud	longitud	peso_bajo_talla	operacion
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	16/06/2020 15:30	BSH	3850.00	9999 42 6666666666666666	28.016666666666667	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	16/06/2020 15:30	SWQ	90.00	9999 42 6666666666666666	28.016666666666667	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	13/06/2020 16:15	SWQ	315.00	9999 42 4333333333333333	-35.88333333333333	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	13/06/2020 16:15	SMA	170.00	9999 42 4333333333333333	-35.88333333333333	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	13/06/2020 16:15	BSH	4043.00	9999 42 4333333333333333	-35.88333333333333	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	10/06/2020 20:35	BSH	7450.00	9999 42 0333333333333333	-44.016666666666666	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	10/06/2020 20:35	SWQ	1600.00	9999 42 0333333333333333	-44.016666666666666	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	10/06/2020 20:35	LEC	34.00	9999 42 0333333333333333	-44.016666666666666	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	10/06/2020 20:35	SMA	380.00	9999 42 0333333333333333	-44.016666666666666	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	09/06/2020 19:30	SWQ	810.00	9999 41 95	-44.416666666666664	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	09/06/2020 19:30	BSH	6480.00	9999 41 95	-44.416666666666664	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	09/06/2020 19:30	BET	82.00	9999 41 95	-44.416666666666664	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	08/06/2020 19:05	BSH	4480.00	9999 42 0666666666666667	-44.633333333333333	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	08/06/2020 19:05	SWQ	830.00	9999 42 0666666666666667	-44.633333333333333	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	08/06/2020 19:05	SMA	230.00	9999 42 0666666666666667	-44.633333333333333	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	07/06/2020 18:05	SMA	180.00	9999 42 2	-45.25	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	07/06/2020 18:05	BSH	1649.00	9999 42 2	-45.25	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	07/06/2020 18:05	SWQ	668.00	9999 42 2	-45.25	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	07/06/2020 18:05	LEC	60.00	9999 42 2	-45.25	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	06/06/2020 18:00	BSH	3929.00	9999 42 0833333333333336	-44.783333333333333	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	06/06/2020 18:00	SMA	350.00	9999 42 0833333333333336	-44.783333333333333	0.00	captura		
8e4d7e1591576b5002688bc1398f	Palangrero de Superficie	7a159eece54c80da3768b1c70f929841	Vigo (ESPAÑA)	29/04/2020 15:00	06/06/2020 18:00	SWQ	936.00	9999 42 0833333333333336	-44.783333333333333	0.00	captura		

Ilustración 2. Ejemplo de datos del fichero csv que contiene los datos de las operaciones

Los datos contenidos en ese fichero CSV son los siguientes:

- **cod_buque:** Identificador del barco que se obtiene encriptando el nombre de este, debido a que en el visor de mareas no es obligatorio rellenar el CFR de este.
- **arte:** Lo define la organización de productores según el buque, no proviene del visor de mareas.
- **cod_marea:** Identificador de marea también encriptado, se extrae del campo *Informe de Marea* en el visor.
- **puerto_llegada:** Campo de *Puerto (País)* en la declaración de Desembarque en el visor de mareas. En caso de no declararse desembarque aparecerá como *None*.



- **fecha_inicio** y **fecha_fin**: Fecha y hora de inicio y de final de una marea que figuran al principio del documento del visor de mareas. En caso de tratarse de una marea en curso, la fecha_fin aparecerá vacía.
- **fecha_actividad**: Fecha y hora de la declaración de la operación a la que corresponde.
- **cod_fao**: Código de especie AL3 que figura en la declaración de captura o descarte en el visor.
- **peso**: Peso declarado en la *Información de las Especies* del visor.
- **num_ejemplares**: Número de ejemplares capturados o descartados en *Información de las Especies*. En caso de estar vacío en el visor, aparecerá como -9999.
- **latitud** y **longitud**: Coordenadas decimales de la operación.
- **peso_bajo_talla**: Peso de ejemplares bajo talla de una captura. Aparecerá como -9999 en caso de estar vacío o tratarse de un descarte.
- **operacion**: Identificador de la operación. Diferencia entre **descarte** o **captura**.

3.- Creación/adaptación de la base de datos de entrada al Geoportal

Una vez definidos los campos que se van a transferir desde la web Armador al Geoportal, y a partir de la versión preliminar de estructura de la base de datos del Geoportal presentado en un informe anterior, se modificó dicha estructura para adaptarla a los datos que finalmente estarán presentes en el fichero de transferencia de formato CSV generado a través de la aplicación DEO. La estructura del modelo de datos se presenta en la ilustración 3.

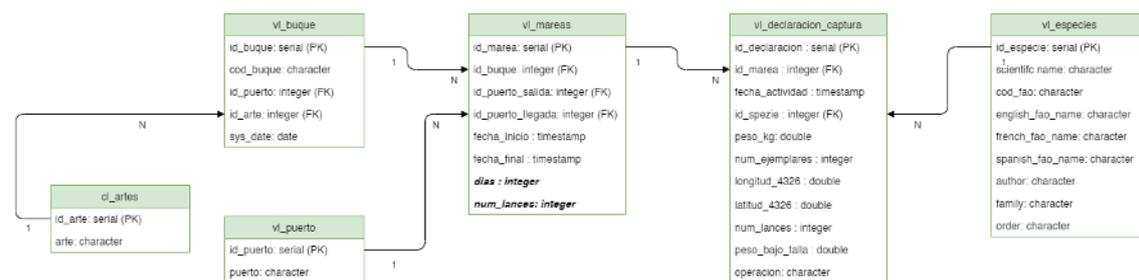


Ilustración 3. Segunda versión del modelo de datos de GEOCAP

Por tanto, se implementó en el Geoportal del CESGA la nueva base de datos para recoger la información contenida en el archivo CSV suministrado por Sealab a través del DEO con la estructura de datos (coincidente con la descrita en la sección anterior) descrita en las siguientes tablas:

Nombre de variable	Tipo	Descripción
cod_buque	String	Identificador del buque anonimizado
arte	String	Arte que se asocia al buque
cod_marea	String	Identificador de la marea anonimizada
puerto_llegada	String	Puerto de llegada de la marea
fecha_inicio	Timestamp	Fecha de inicio de la marea
fecha_fin	Timestamp	Fecha de final de la marea
num_lances	Integer	Número de lances (En realidad cada registro es un lance)
fecha_actividad	Timestamp	Fecha de la operación
cod_fao	String	Código de la FAO que identifica la especie
peso	Double	Peso de la operación en Kg
num_ejemplares	Integer	Número de ejemplares declarados en la operación
latitud	Double	Coordenadas de latitud en DD y EPSG: WGS84
longitud	Double	Coordenadas de longitud en DD y EPSG: WGS84
peso_bajo_talla	Double	Peso de la operación bajo talla en Kg
operacion	String	Tipo de operación: captura y descartes

En base a los datos disponibles se ha optado por el siguiente modelo de datos¹.

1.1.- Entidades

- vl_buque. (E1)
- vl_mareas. (E2)
- vl_declaracion_capturas. (E3)
- vl_especies (E4)
- vl_puerto (E5)
- cl_artes (E6)

1.2.- Relaciones

- Los buques tienen asignado un código de arte. (R1)
- Los buques realizan mareas (R2)
- Las mareas tienen un puerto de llegada (R3)
- Las mareas tienen declaraciones de capturas (R4)
- Las declaraciones de capturas identifican especies (R5)

Las entidades tienen las siguientes variables:

vl_buque (E1)		
Nombre de variable	Tipo	Descripción
id_buque	Serial	Identificador del buque (PK)
cod_buque	String	Nombre del buque encriptado
id_puerto	Integer	Identificador de puerto (FK)
id_arte	Integer	Identificador de arte (FK)
sys_date	Date	Fecha de la carga del dato

¹ El diseño ha sido elaborado siguiendo el modelo Entidad/Relación Extendido, según está definido en *Ramez Elmasri, Shamkant B. Navathe. Sistemas de bases de datos: Conceptos fundamentales. Addison-Wesley Iberoamericana. 1997*

vl_mareas (E2)		
Nombre de variable	Tipo	Descripción
id_marea	Serial	Identificador de marea (PK)
cod_marea	String	Código de marea encriptado
Id_buque	Integer	Identificador de buque (FK)
id_puerto_salida	Integer	Identificador de puerto de salida (FK)
id_puerto_llegada	Integer	Identificador de puerto de llegada (FK)
Fecha_inicio	Timestamp	Fecha inicio marea
Fecha_final	Timestamp	Fecha final marea
Días	Integer	Duración en días de la marea
Num_lances	Integer	Número de lances de la marea
sys_date	Date	Fecha de la carga del dato

vl_declaracion_capturas (E3)		
Nombre de variable	Tipo	Descripción
id_declaracion	Serial	Identificador de declaración (PK)
id_marea	Integer	Identificador de marea (FK)
Fecha_actividad	Timestamp	Fecha de la actividad
Id_espezie	Integer	Identificador de especie (FK)
Peso_kg	Double	Peso en kg de la captura
Num_ejemplares	Integer	Número de ejemplares
Longitud_4326	Double	Coordenadas de longitud
Latitud_4326	Double	Coordenadas de latitud
Num_lances	Integer	Número de lances de la marea
Peso_bajo_talla	Double	Peso bajo talla
Operacion	String	Tipo de operación captura/descarte
Sys_date	Date	Fecha de la carga del dato

vl_especies (E4)		
Nombre de variable	Tipo	Descripción
id_especie	Serial	Identificador de especie (PK)
Scientific_name	String	Nombre científico de la especie
Cod_fao	String	Código FAO de la especie
English_fao_name	String	Nombre FAO en inglés de la especie
French_fao_name	String	Nombre FAO en francés de la especie
Spanish_fao_name	String	Nombre FAO en castellano de la especie
Autor	String	Referencia de la especie
Family	String	Familia
Order	String	Orden

vl_puertos (E5)		
Nombre de variable	Tipo	Descripción
Id_puerto	Serial	Identificador de puerto (PK)
puerto	String	Etiqueta del arte

cl_artes (E6)		
Nombre de variable	Tipo	Descripción
Id_arte	Serial	Identificador del arte (PK)
arte	String	Etiqueta del arte



4.- Conclusiones

Desde la web Armador se capturan los datos de cada marea de cada barco y, mediante la herramienta DEO se cargan en una base de datos. Una vez seleccionados los campos de interés para el proyecto, los datos se cargan en un fichero CSV que se pone a disposición del CESGA para su recogida y carga de esos datos a la base de datos del Geoportal.

Se ha desarrollado una nueva base de datos para la recogida de los datos pertinentes desde el fichero CSV.