

Servicio de Datos de la Unidad de Tecnología Marina (UTM-CSIC)

REVISTA **MAPPING**
Vol. 28, 194, 32-42
marzo-abril 2019
ISSN: 1131-9100

Marine Technology Unit Data Service (UTM-CSIC)

Juan Luis Ruiz Valderrama, Susana Diez Tagarro, Alberto Hernández Jiménez,
Jordi Sorribas Cervantes

Resumen

La Unidad de Tecnología Marina del Consejo Superior de Investigaciones Científicas (UTM-CSIC) se encarga de la gestión de buques oceanográficos de investigación. Durante las campañas oceanográficas se adquieren una gran cantidad de datos espaciales que son gestionados por el Servicio de Datos de la UTM, desde datos que se adquieren de forma automática y continua en una trayectoria, como la temperatura y salinidad superficial del mar, a datos que se adquieren en puntos o áreas concretas de estudio, como sería un perfil de velocidad del sonido en la columna de agua o una batimetría para cartografiar el fondo marino.

El propósito del Servicio de Datos de la UTM es dar a conocer qué datos existen y dónde, cómo y cuándo han sido adquiridos y además, dar accesibilidad a esos datos. Para ello, se dispone de una IDE nacional que consta de un Catálogo de Campañas Oceanográficas llevadas a cabo en los buques que gestiona desde 1991, con más de 500 campañas, y de un Geoportal que permite crear mapas combinando capas.

La UTM es un National Oceanographic Data Center (NODC) de la Infraestructura Distribuida Europea de Datos Marinos SeaDataNet, a la que aporta metadatos (Cruise Summary Report -CSR- y Common Data Index -CDI-, ISO19139) y datos públicos para ser compartidos..

Abstract

The Marine Technology Unit of the Spanish Research Council (UTM-CSIC) is responsible for the management of oceanographic research vessels. During the oceanographic cruises, a large amount of spatial data is acquired and managed by the UTM Data Service, from data that is automatically and continuously acquired in a trajectory, such as the temperature and salinity of the sea, to data that they are acquired in specific points or areas of study, such as a sound velocity profile in the water column or a bathymetry to map the seabed.

The purpose of the UTM Data Service is to disclose what data exists and where, how and when it has been acquired and, in addition, to provide access to that data. For this, a national SDI is available, consisting of a Catalogue of Oceanographic Cruises carried out on ships that it has managed since 1991, with more than 500 cruises, and a Geoportal that allows the creation of maps combining layers. The UTM is a National Oceanographic Data Center (NODC) of the Distributed European Sea Data Infrastructure SeaDataNet, to which it provides metadata (Cruise Summary Report -CSR- and Common Data Index -CDI-, ISO19139) and public data to be shared.

Palabras clave: Datos marinos, metadatos, oceanografía, Geonetwork, Geonode, SeaDataNet.

Keywords: Marine data, metadata, Oceanography, Geonetwork, geonode, SeaDataNet.

Unidad de Tecnología Marina (UTM-CSIC)

jlruiz@utm.csic.es

sdiez@utm.csic.es

ahernandez@utm.csic.es

sorribas@utm.csic.es

Recepción 08/01/2019
Aprobación 24/01/2019

1. INTRODUCCIÓN

Los orígenes de la UTM se remontan al año 1992 cuando se constituye la Unidad de Gestión de Buques Oceanográficos (UGBO), que cambió su denominación en el año 2001 a Unidad de Tecnología Marina (UTM, 2018)). La Unidad nace con el objetivo de proporcionar al Ministerio el apoyo tecnológico y logístico necesario para realizar la investigación marina y polar promovida por el Plan Nacional de I+D+i, colaborando en la coordinación y soporte de las actuaciones relativas a los buques científicos y las bases antárticas. Asume inicialmente el mantenimiento y operación de los buques oceanográficos BIO Hespérides (2018) y B/O García del Cid (2018) y, tras liderar su diseño y construcción, incorpora en su gestión al nuevo buque B/O Sarmiento de Gamboa (2018), operativo desde 2008. Por último, en 2013 asumió el encargo de gestionar el B/O Mytilus (2018), botado en 1997.

Durante las campañas oceanográficas de investigación se trabaja 24 horas al día, por lo que se adquiere una gran cantidad de datos espaciales de diferente naturaleza que representan las propiedades del océano, desde observaciones puntuales a datos ráster, pero también perfiles verticales en la columna de agua y trayectorias a través de la superficie oceánica.

Desde sus inicios hasta fechas recientes, la implicación de la UTM sobre estos datos se ha limitado a la custodia pasiva y entrega de copias a personas autorizadas por el Investigador Principal (IP) del proyecto como propietario de los datos, una vez finalizada la campaña. La necesidad de los organismos gestores de optimizar el uso de los recursos, así como las directivas europeas relativas a la liberación de datos financiados con fondos públicos, lleva a la UTM a aumentar su compromiso con la gestión posterior de la información generada durante las expediciones científicas realizadas en los buques que opera.

En una fase inicial, se ha realizado una prospección de los recursos disponibles para hacer efectivo ese compromiso, participando en foros nacionales e internacionales relativos a la gestión de datos como son los organizados por el servicio de datos abiertos del CSIC (Digital.CSIC, 2018), colaborando en el diseño inicial de la IDE para el Centro Nacional de Datos Polares del Instituto Geológico Minero de España (CNDP-IGME, 2012) y participando en los proyectos internacionales del *Open*

Geospatial Consortium (OGC, 2018), *Ocean Data Interoperability Platform* (ODIP, 2018) y la infraestructura paneuropea de datos marinos SeaDataNet. Finalmente, la UTM ha apostado con determinación por la creación de una nueva área transversal con entidad propia, especializada en la gestión de datos de investigación oceanográfica, haciendo un esfuerzo considerable por asignar los recursos humanos y materiales necesarios para ofrecer un servicio profesional y de calidad.

La Comisión de Coordinación y Seguimiento de las Actividades de los Buques Oceanográficos (COCSABO, 2018), adscrita al Ministerio de Ciencia, Innovación y Universidades, es el órgano colegiado de propuesta y coordinación de las actuaciones de carácter científico y técnico realizadas por los buques oceanográficos. En el ejercicio de sus funciones, en 2017 confeccionó un nuevo Protocolo de Acceso a la Flota de Investigación Oceanográfica del que se pueden extraer los siguientes requerimientos relativos a la cesión de datos y metadatos y la citación en publicaciones derivadas de las campañas:

1. Todos los datos generados en campañas oceanográficas financiadas parcial o totalmente por el Ministerio deben terminar siendo públicos y accesibles.
 - a. El plazo de cesión para los datos elaborados por parte del personal científico embarcado es de dos años desde la finalización de la campaña o desde la finalización del proyecto.
 - b. Los datos de carácter básico adquiridos de forma automática y aquellos necesarios para el ajuste técnico o calibración instrumental son propiedad del organismo gestor; por lo que son susceptibles de ofrecerse públicamente sin ninguna restricción temporal.

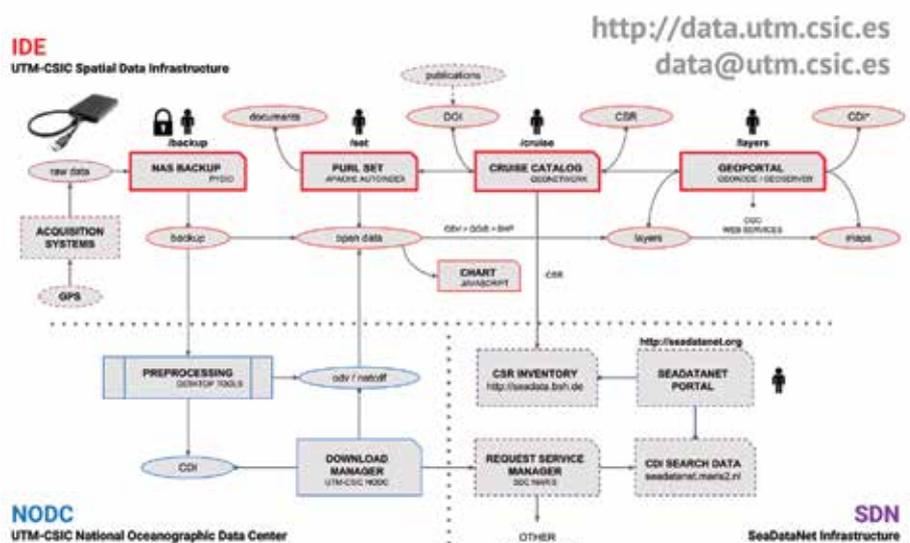


Figura 1. Diseño conceptual del Servicio de Datos UTM-CSIC

2. Junto al informe final de campaña, se deberá enviar también al operador del buque en formato digital, la lista de metadatos obtenidos durante la campaña en formato ISO-19139.
3. En las publicaciones generadas con datos obtenidos en campañas oceanográficas deberá hacerse mención explícita en los agradecimientos al buque empleado.
4. El IP del proyecto deberá enviar al operador del buque una copia de los documentos publicados como consecuencia de la participación en una campaña.
5. Cada operador mantendrá actualizado un registro de actuaciones de I+D de cada buque, en el que recogerá todas aquellas publicaciones o resultados obtenidos en proyectos desarrollados a bordo de cada buque.

Estas y otras condiciones relacionadas con la gestión interna de la UTM, productos y servicios de datos reclamados por la comunidad científica y el ajuste a los estándares de los proyectos internacionales a los que la UTM se ha suscrito, han marcado las líneas de actuación y diseño del nuevo Servicio de Datos.

2. ADQUISICIÓN Y CUSTODIA DE DATOS

Cada uno de los grupos especializados del área técnica que conforman la UTM se hace responsable del correcto funcionamiento, gestión y operación de los equipos y sistemas necesarios para la recogida de muestras y adquisición de datos durante las campañas, siguiendo los requerimientos previos reflejados en el plan de campaña y atendiendo las indicaciones *in situ* del equipo científico embarcado.

De este modo, se opera sobre el equipamiento para proporcionar muestras y datos para una amplia variedad de parámetros:

- Físicoquímicos: temperatura del agua, salinidad, oxígeno disuelto, clorofila...
- Dinámicos: corrientes, oleaje, mareas...
- Biológicos: muestreo por observación y captura...
- Geológicos y geofísicos: batimetrías, reflectividad, magnetometría, gravimetría, testigos de sedimento, imágenes de estructuras suboceánicas...
- Meteorológicos: temperatura del aire, humedad, presión, radiación solar, velocidad y dirección del viento...

Los datos adquiridos automáticamente de forma continuada, como son los de navegación, profundidad, parámetros físicoquímicos de la superficie del agua y meteorología, además de ser registrados localmente por sus respectivos equipos, son distribuidos en tiempo real a toda la red del barco y a los servidores de las oficinas centrales en Barcelona haciendo uso de la conexión vía satélite. Este Sistema de Adquisición de Datos Oceanográficos (SADO) (García, Hernández, Ruiz, Sorribas, et al., 2016), además de ser una herramienta de registro distribuida, permite implementar otras aplicaciones como las de monitorización gráfica de datos en tiempo real (RTP) (Sorribas, Hernández, Ruiz, et al., 2009) programada por el Servicio de Datos en Javascript, disponible tanto en la intranet de los buques como en el portal web de la UTM (2018).

El equipo científico tiene la opción de trabajar durante el transcurso de la campaña sobre las muestras y datos brutos gracias a los laboratorios multidisciplinares y servicios informáticos y de comunicaciones dispo-

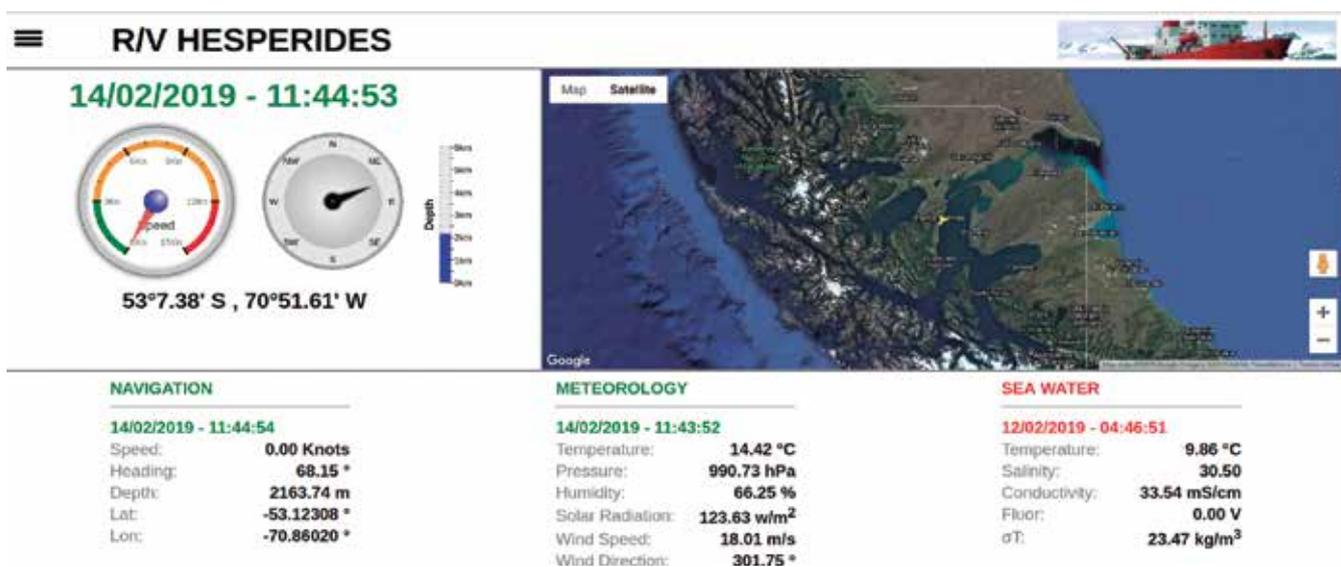


Figura 2. Panel de visualización de datos en tiempo real vía satélite

nibles en el buque. Al finalizar la campaña, el personal de informática embarcado realiza una copia final de todos los datos y metadatos generados para ser entregados al Investigador Principal (IP) junto al informe técnico detallado de las actividades realizadas. Asimismo, como operador del buque, la UTM guarda una copia de seguridad para ser custodiada por el Servicio de Datos en las oficinas centrales.

Desde 1991, esta custodia se ha realizado guardando la copia de seguridad sobre el soporte más adecuado disponible en cada momento: cintas, discos ZIP, CD, DVD o HD. Actualmente, con la adquisición de nuevos equipos *hardware* y asignación de personal específico para la gestión de datos, se ha iniciado la unificación de toda la información en un sistema de almacenamiento en red (NAS) con posibilidad de ponerla a disposición de los usuarios a través de servicios y aplicaciones en línea siguiendo la política de accesos y privilegios convenida según normativa o la voluntad expresa de los propietarios.

Durante el año 2012, se trabajó en la unificación de soportes, consiguiendo recuperar la información desde diferentes formatos de las campañas posteriores a 2001 y replicarlas a discos duros USB. Actualmente, el Servicio de Datos tiene a su disposición un servidor con doce discos duros de 10 TB dispuestos en LVM con dos grupos en RAID 5 con *spare*, lo que supone una capacidad de almacenamiento final de 70 TB. Con los trabajos de replicación del histórico que se han realizado hasta el momento, se estima un volumen aproximado de 1 TB anual por buque.

De los buques B/O García del Cid y B/O Mytilus no se han conservado copia de los datos históricos, por lo que únicamente están disponibles los adquiridos en las plataformas principales: BIO Hespérides y B/O Sarmiento de Gamboa. El volumen provisional de datos históricos previsto desde 2001 a 2018 es de unos 30 TB, perfectamente asumible con los recursos actuales.

El acceso a este repositorio para administradores y

usuarios autorizados desde la red local o red privada virtual (configurada con todos los buques y sedes periféricas), se puede realizar haciendo uso de protocolos seguros sobre SSH, como son SFTP o, si se prefiere montar en el propio sistema de ficheros, mediante SSHFS. El acceso predeterminado para el exterior y el resto de usuarios es a través del protocolo seguro de transferencia de hipertexto HTTPS. La aplicación en línea escogida para gestionar estas peticiones es Pydio (2018): un *software* abierto de sincronización y uso compartido de ficheros de código abierto ejecutado en servidor.

De este modo, se ofrece fácilmente al IP o a personas autorizadas el acceso remoto y seguro al *dataset* completo o parcial de una campaña determinada.

3. GENERACIÓN DE METADATOS

El uso de metadatos facilita la indexación y búsqueda de conjuntos de datos, describe en detalle su naturaleza (qué) y amplía la información relativa a las condiciones de adquisición: quién las adquiere, a quién pertenecen, quién las guarda, cómo, cuándo, dónde y porqué se han adquirido...

Desde que en 2012 la UTM inició su colaboración con la Infraestructura Paneuropea de Datos Marinos SeaDataNet, ha adoptado el formato de metadatos definidos por ésta para la descripción de los diferentes datos de adquisición y de las propias campañas oceanográficas que los agrupa:

- *Cruise Summary Reports* (CSR) informa sobre expediciones y campañas oceanográficas. Proporciona un inventario de primer nivel de mediciones y muestras recolectadas en el mar. La información, como la trayectoria del barco, los objetivos de la campaña y los datos de contacto del investigador principal también se incluyen entre otros elementos de metadatos de utilidad probada en la práctica de la comunidad marina.
- *Common Data Index* (CDI) es un descriptor de datos marinos que incluye elementos adicionales para bibliografía, referencias a CSR e información sobre control de calidad.

Ambos utilizan el esquema normalizado ISO-19139 que forma parte de las Normas de Ejecución de la Directiva INSPIRE de la UE, por lo que se facilita la interoperabilidad con iniciativas y comunidades internacionales, tales como OGC, EMODnet (2018)²¹ y GEOSS (2018)²², así como un alcance internacional más allá de *SeaDataNet*.

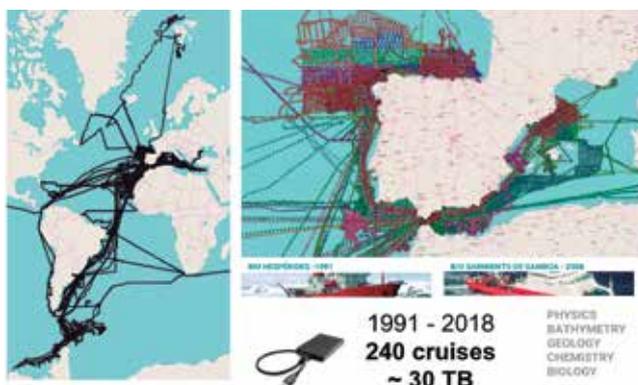


Figura 3. Campañas con copias de seguridad custodiadas por la UTM

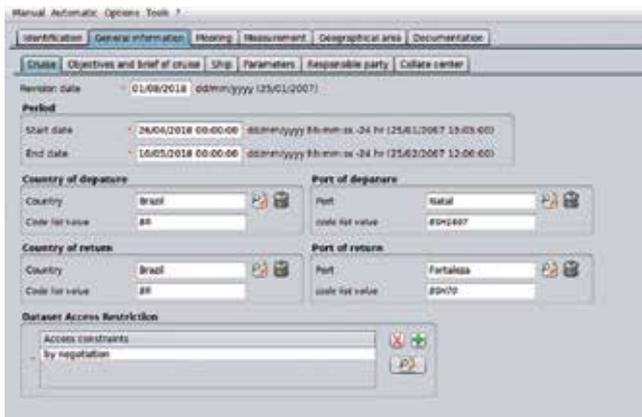


Figura 4. Aplicación Mikado para la generación de metadatos SeaDataNet

Además, incorporan referencias al Directorio Europeo de Organizaciones Marinas (EDMO, 2018) y a los Vocabularios Comunes de SeaDataNet, gestionados y alojados técnicamente por el *British Oceanographic Data Center* (BODC, 2016), un prerrequisito imprescindible para la coherencia y la interoperabilidad de los metadatos. Los vocabularios comunes consisten en listas de términos estandarizados que cubren un amplio espectro de disciplinas de relevancia para la comunidad oceanográfica. El uso de conjuntos de términos estandarizados resuelve el problema de las ambigüedades asociadas con el marcado de datos y también permite que los registros sean interpretados por algoritmos programados.

La herramienta que ofrece SeaDataNet para generar estos metadatos es Mikado (2018), una aplicación de escritorio multiplataforma que utiliza los últimos esquemas XML ISO-19139 de CSR y CDI además de proporcionar la última versión de los vocabularios comunes mediante conexiones de sincronización en línea con el servicio web del BODC.

Con el objetivo de facilitar la edición semiautomática de metadatos durante las campañas, el Servicio de Informática y Comunicaciones de la UTM ha programado un paquete de aplicaciones web en Java instaladas en los servidores de los buques que permite al equipo científico confeccionar una lista de CDI, o eventos de adquisición, especificando exclusivamente la fecha de inicio y fin, un identificador único de estación o línea y el código del instrumento usado. Posteriormente, un técnico con privilegios de administración en la aplicación, genera automáticamente los metadatos y otros productos derivados. Para ello, la aplicación hace uso de los datos de navegación registrados en el SADO, plantillas CDI preconfiguradas con Mikado para describir cada instrumento y la plantilla del CSR de la campaña previamente modificada con la información proporcionado por el IP para la campaña en curso.

4. CATÁLOGO DE CAMPAÑAS

Un elemento principal de la Infraestructura de Datos Espaciales de la UTM a nivel nacional es el Catálogo de Campañas Oceanográficas (2018). Los metadatos de campaña generados (CSR) conforman el catálogo que actualmente consta de aproximadamente 500 campañas, llevadas a cabo desde 1991 en los buques que gestiona la UTM. Los metadatos básicos de cada campaña oceanográfica se completan con una serie de documentación y recursos en línea añadiendo a la infraestructura enlaces adicionales que atienden tanto a los usuarios en general como a los gestores de las plataformas nacionales.

Como primer paso, es fundamental ofrecer una manera eficiente de facilitar la cita del conjunto de datos adquiridos en una campaña oceanográfica. La forma más extendida que se está imponiendo en este campo es el empleo del sistema *Digital Object Identifier* (DOI): un identificador y una dirección de referencia, o URL permanente, que permite localizar un recurso sin que cambie con el paso del tiempo, aunque éste sea reubicado en una aplicación o dominio distinto. DOI facilita la correcta cita de los datos, su uso y reutilización, así como su búsqueda y la posibilidad de poder contar con datos estadísticos.

DataCite (2018) es una de las principales agencias de registro de DOI. El CSIC, adscrito a esta organización como *full member* desde 2015, ofrece a la UTM la cobertura necesaria para generar sus propios identificadores bajo el prefijo 10.20351, de manera que cada DOI o enlace permanente lleva a un ficha de campaña (*landing page*), implementada sobre Geonetwork.

Geonetwork es un entorno de *software* libre y código abierto de gestión y catalogación para recursos referenciados en el espacio geográfico. Tras una transformación de adaptación técnica de los CSR mediante XSLT (eXtensible Stylesheet Language Transformation), se incorpora a Geonetwork una entrada por campaña con el esquema ISO-19139, ofreciendo así una ficha con

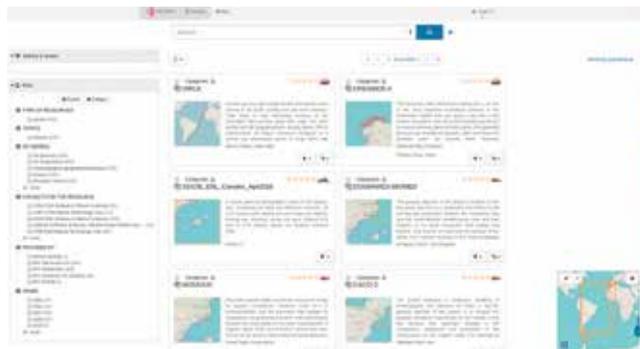


Figura 5. Portal de búsqueda del Catálogo de Campañas UTM-CSIC

diferentes representaciones o parseos: simplificada, completa y XML.

Una vez creada cada ficha a partir del CSR, el administrador del sistema las edita añadiendo una imagen icónica con el mapa de derrota (*Cruise Track*), una restricción (*resource constraint*) que indica cómo se debe citar el *dataset* haciendo uso del DOI y diversos enlaces a recursos en línea relacionados con la campaña, como son:

- DOI: Identificador de la campaña con un enlace a la ficha de citación en DataCite.
- SCIENTIFIC PLAN: Documento en PDF con el plan científico de campaña.
- CRUISE MAP: Enlace al mapa de campaña del Geoportal UTM.
- OPEN DATASET: Directorio con los datos abiertos en diferentes formatos.
- UNDERWAY CHART: Aplicación desarrollada por el Servicio de Datos que muestra gráficas lineales interactivas de los parámetros adquiridos en continuo (SADO).
- TECHNICAL REPORT: Documento en PDF con el informe técnico de campaña.

Cada uno de estos enlaces, a excepción del primero, corresponde a una entrada del directorio ofrecido por el servidor web de la IDE en SET (2018). Este directorio tiene como objetivo crear referencias permanentes y simplificadas a enlaces locales (o pURL). Para ello hace uso del módulo *autoindex* de Apache y comandos PHP de redireccionamiento. De este modo, se unifica el formato de los enlaces a recursos y permite la sustitución de aplicaciones en línea sin que ello afecte a su

dirección de referencia; lo que garantiza su validez a lo largo del tiempo, lo que facilita la cita de esos enlaces en publicaciones o la configuración en DataCite del landing page del DOI.

Como ejemplo, se puede observar cómo la dirección de las fichas de campaña que tiene el formato asociado a Geonetwork:

<http://data.utm.csic.es/geonetwork/srv/eng/catalog.search#/metadata/urn:SDN:CSR:LO-CAL:29HE20170612>, pasa a ser:

<http://data.utm.csic.es/set/hes/20170612/doi>.

Accediendo a la raíz <http://data.utm.csic.es/set> también es posible navegar por las diferentes plataformas, campañas y recursos del directorio con la ayuda de textos descriptivos.

Actualmente, el acceso a los datos abiertos sólo contempla una parte de los que pertenecen al operador del buque según el Protocolo de Acceso a la Flota de la COCSABO: los adquiridos de forma automática y continua por el SADO (underway) y los perfiles de velocidad del sonido para la calibración de las sondas mediante los batitermógrafos desechables (XBT). Estos ficheros son publicados tras los trabajos de preprocesado a cargo del Servicio de Datos, donde se transforman y añaden parámetros de control de calidad (QC) haciendo uso de estándares internacionales.

Los procedimientos de QC de datos son importantes para detectar valores atípicos, duplicados o valores que en realidad se corresponde a información no registrada. Para no modificar los puntos de datos observados el QC consiste en asignar a cada valor

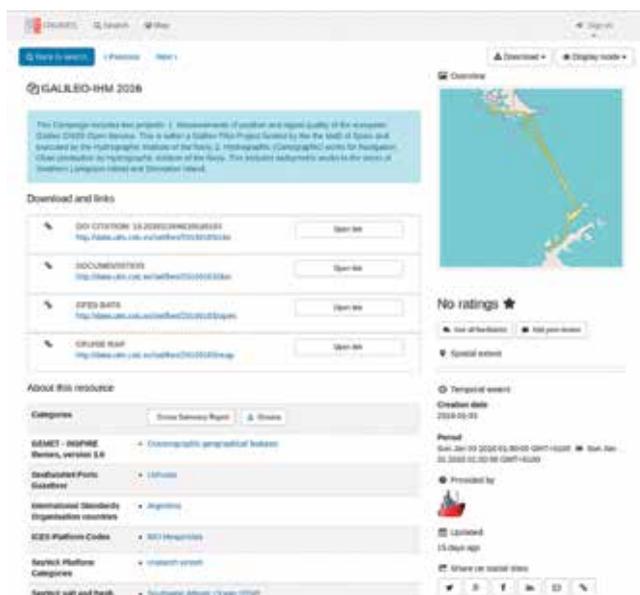


Figura 6. Registro del Catálogo de Campañas UTM-CSIC

Index of /set/hes/20160103/open/underway

Name	Last modified	Size	Description
Parent Directory			
csv/	2018-01-30 15:58		ASCII-CSV Non-stander format without QC
csv/	2018-01-30 15:57		ASCII-TXT Open Data Viewer format with QC flags
csv/	2018-01-30 15:57		ASCII-CSV with QC filtered
chart/	2018-01-30 15:57		Data graphs
cdi/	2018-01-30 15:57		Common Data Index metadata

Figura 7. Directorio de datos de campaña públicos

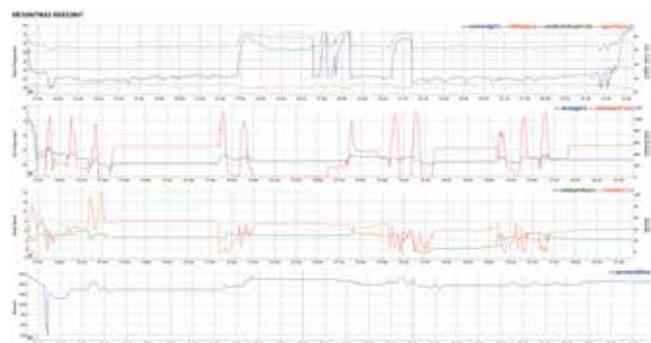


Figura 8: Gráficas lineales interactivas de los parámetros adquiridos en continuo

numérico un indicador de calidad según una escala definida y normalizada de *flags* de QC.

Los datos están disponibles en varios formatos dependiendo del nivel de control de calidad llevado a cabo. De esta manera se ofrecen los datos brutos, en formato ASCII-CSV sin QC, los datos en formato ASCII-ODV con QC basado en *flags* y por lo tanto manteniendo todos los valores, y los datos en formato ASCII-CSV filtrados atendiendo al QC llevado a cabo. Con estos últimos se confecciona el graficado lineal interactivo como producto de datos adicional.

5. GEOPORTAL DE DATOS ABIERTOS

A pesar de que Geonetwork incorpora el tratamiento de mapas, desde el Servicio de Datos se ha buscado una alternativa que se ajuste a las expectativas de calidad de un Sistema de Información Geográfica en línea o un GeoPortal avanzado; incluyendo herramientas para el diseño dinámico de estilos, medidas geográficas, búsqueda por atributos, interactividad con capas y mapas, catalogación propia ajustada a las necesidades, y con servicios incorporados que cumplan los requisitos de interoperabilidad internacionales.

Geonode es una aplicación y plataforma basada en la web para desarrollar sistemas de información geoespacial y desplegar infraestructuras de datos espaciales. Incorpora los requisitos deseados y, al ser de código abierto y estar bien documentada, está pensada para ser modificada y ampliada.

En los metadatos CSR ofrecidos por el Catálogo de Campaña aparece una lista completa de todos los sistemas de adquisición y equipos de captación de muestras utilizados, así como la descripción de los parámetros registrados digitalmente o que el equipo científico pretende obtener tras el análisis de muestras o procesado de datos puros. En cambio, este Geoportal alternativo cubre un subconjunto de esta información, centrándose exclusivamente en aquellos datos públicos descargables por el usuario. De este modo se convierte en la herramienta principal de la IDE para la búsqueda avanzada y descarga de datos abiertos.

El portal permite la búsqueda de recursos por:

- Título: compuesto por el identificador de campaña, identificador de instrumento y acrónimo de campaña.
- Categoría: *underway, profile, bathymetry...*
- Palabras clave: reservadas para describir el instru-

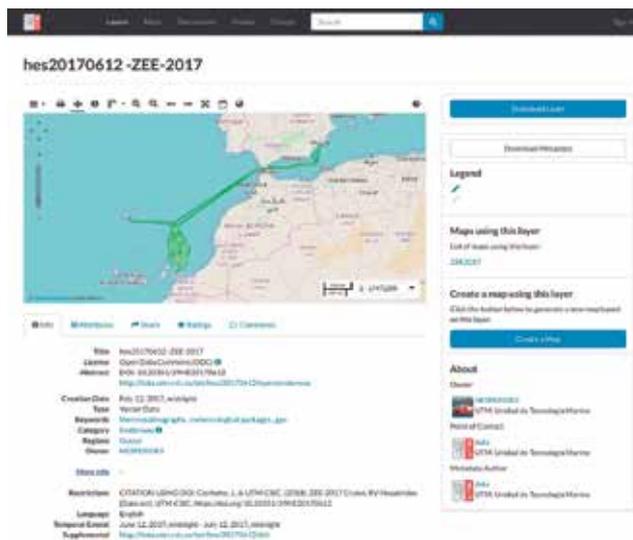


Figura 9. Ficha del GeoPortal UTM-CSIC para capa underway

mento o sistema de adquisición: *meteorological packages, thermosalinographs, xbt...*

- Dueños: donde se indica el buque donde se ha realizado la campaña: Hespérides, Sarmiento...
- Rango de fechas de inicio y fin de campaña.
- Nombre de región marítima.
- Extensión (*bounding box*) en un mapa interactivo.

La información que figura en los metadatos de cada ficha contiene, además de los campos usados en los criterios de búsqueda, información adicional como:

- Licencia: se ha escogido Open Data Commons (2018) para todos los datos liberados.
- Resumen: con el código DOI y un enlace al directorio de descarga de los archivos de datos en diferentes formatos desde el directorio de recursos SET.
- Restricciones: citación mediante DOI con el formato recomendado como muestra.

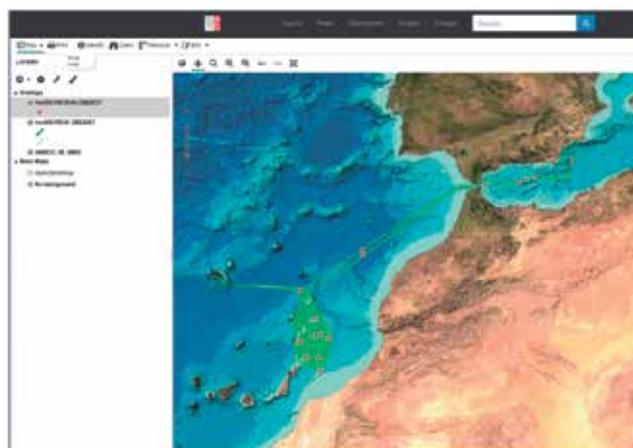


Figura 10. Mapa de campaña con capa underway, puntos de muestreo XBT y capa base batimétrica global del servicio WMS de GEBCO (2018)

- Información adicional: enlace al registro en el Catálogo de Campañas en Geonetwork.

Las capas en formato *shapefile* del *underway* se crean a partir del archivo ASCII-CSV con QC cargados en la aplicación SIG *desktop* de *software* libre QGIS (2018). La resolución habitual de estas capas es de una muestra por minuto. Si a la información geográfica le sumamos los datos de adquisición en forma de atributos, el resultado es un gran volumen de información que puede llegar a colapsar o hacer impracticable el uso en línea del GeoPortal. El Servicio de Datos ha optado por simplificar la información convirtiendo la nube de puntos muestreados en una línea continua de derrota. Con esta transformación se pierden los atributos espacio-temporales y los parámetros, pero se gana en ligereza y velocidad al realizar las búsquedas y descargas de *datasets* voluminosos de alta resolución; se reserva la visualización y análisis detallado a las aplicaciones *GIS desktop* locales una vez descargados los datos por el usuario.

Las capas de perfiles verticales aparecen en el mapa como puntos, por lo que no existe posibilidad de visualización. En su lugar se ofrece un enlace para la descarga directa del archivo de datos de la estación mediante uno de los atributos de la capa, accesible en el mapa de forma interactiva.

Geonode, a través de los servicios de Geoserver (2018) que le dan soporte (incluidos en la instalación estándar de Geonode), puede ofrecer los metadatos de las capas en los formatos ISO 19115, FGDC, eBRIM, Dublin Core, DIF y Atom; y las capas de información geográfica en Tiles, KML, GeoJSON, Excel, CSV, GML y

Shapefile, además de imágenes en PNG, PDF y JPEG.

Geoserver posibilita también la interoperabilidad mediante los diferentes servicios de mapas que cumplen los estándares OGC (OWS) como Web Map Service (WMS), Web Coverage Service (WCS), Web Feature Service (WFS) o Web Map Tile Service (WMTS), entre otros.

Geonode también incluye en su instalación un Servicio de Catálogo Web (CSW), un estándar diseñado por OGC que define una interfaz común para el descubrimiento, búsqueda y consulta de metadatos relacionados con datos, servicios y recursos de tipo geográfico. Este servicio se ofrece a través de la aplicación PyCSW (2013).

6. DISEMINACIÓN INTERNACIONAL

En 2016 la UTM se integró como *partner* de pleno derecho en el proyecto del programa H2020 de la UE SeaDataCloud (2018), que tiene como objetivo mejorar considerablemente los servicios de la infraestructura de datos SeaDataNet adoptando la tecnología en la nube. SeaDataNet es una infraestructura paneuropea distribuida y estandarizada para administrar los conjuntos de datos recopilados por las flotas oceanográficas y los sistemas de observación automática. SeaDataNet conecta actualmente 110 centros de datos de 34 países, con el objetivo de preservar y poder reutilizar datos marinos de diferentes ámbitos.

Esta participación de la UTM se materializa con la creación del UTM-CSIC *National Oceanographic Data Center* (NODC) como nuevo nodo a la red que configura la infraestructura del proyecto.

Como NODC la UTM aporta por un lado los metadatos de las campañas oceanográficas (CSR) que pasan a formar parte del catálogo *Cruise Summary Report*

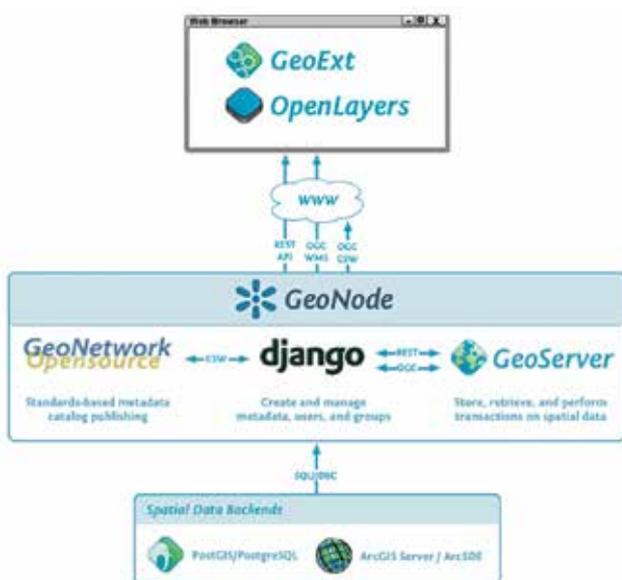


Figura 11. Arquitectura de Geonode (2017)

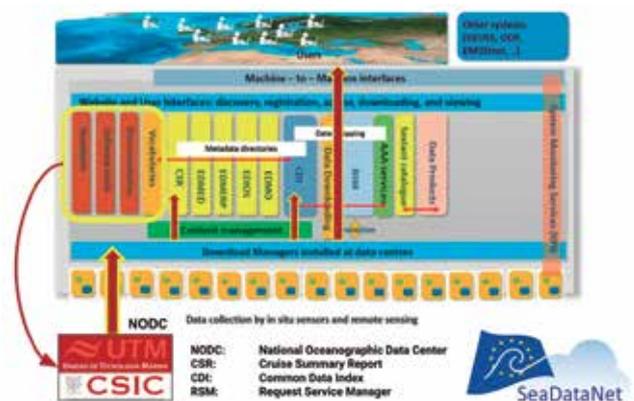


Figura 12. Infraestructura de datos SeaDataNet

Inventory (2018). Este portal abarca campañas desde 1873 hasta la actualidad de más de 2000 buques de investigación con casi 53 000 expediciones, en todas las aguas europeas y océanos. La interfaz de consulta permite buscar por varios criterios y recuperar el informe completo.

Por otro lado, se aportan los metadatos de los datos (CDI) de carácter básico adquiridos en las campañas, que por tanto son de acceso público, que se incluyen en el catálogo *Common Data Index (2018)* proporcionando una interfaz única de búsqueda por un conjunto de criterios permitiendo recuperar los metadatos completos para cada conjunto de datos. Asimismo, el servicio permite descargar conjuntos de datos armonizados de NODC distribuidos en toda Europa. Este proceso se controla a través del servicio *Request Status Manager (RSM)*, que se comunica con los NODC, que a su vez están conectados a la infraestructura SeaDataNet mediante la herramienta *Replication Manager (RM)*, la cual replica los datos de libre acceso de cada nodo en la nube albergada en los servidores de EUDAT: la infraestructura europea de servicios para datos y recursos de apoyo a la investigación.

Los conjuntos de datos son aportados por todos los NODC en los formatos estándar de SeaDataNet, de manera que estos sean interoperables. Se han definido los siguientes formatos de transporte de datos (2018): SeaDataNet ODV4 ASCII y SeaDataNet NetCDF para perfiles, series temporales y trayectorias, SeaDataNet MedAtlas como formato adicional opcional y NetCDF para datos de observación 3D.

Para estandarizar los datos, al mismo tiempo que proporcionar a los usuarios medios comunes para realizar el control de calidad y presentar datos y productos de datos, se han diseñado herramientas *software* (SeaDataNet Software, 2018) comunes para centros de datos y usuarios, como:

- NEMO para la conversión de cualquier formato ASCII al SeaDataNet ODV4 ASCII.
- OCTOPUS para conversión entre diferentes formatos SDN.
- *Ocean Data View (ODV)* para el análisis, visualización y control de calidad de conjuntos de datos en formato ODV.

SeaDataNet es una infraestructura consolidada que permite encontrar, acceder, usar y reutilizar datos marinos interoperables y de calidad provenientes de plataformas como las flotas oceanográficas. Esta infraestructura ha sido adoptada por otros proyectos como el portal de datos batimétricos EMODnet Bathymetry (2018), desarrollado en el marco de la Red Europea de Observaciones y Datos Marinos (EMODnet, 2018).

La constitución de la UTM como National Oceanographic Data Center es clave para poder compartir la información y los datos de las campañas oceanográficas, así como productos como las batimetrías, también a nivel internacional.

7. RETOS Y PERSPECTIVAS DE FUTURO

El reto actual del Servicio de Datos de la UTM es consolidar la Infraestructura de Datos Espaciales Marinos de Buques Oceanográficos gestionados por el CSIC, tanto a nivel nacional como internacional.

Para ello es necesario, en primer lugar, la incorporación de la gran cantidad de datos libres históricos almacenados, lo que implica un esfuerzo importante de armonización, control de calidad y generación o revisión de metadatos. Asimismo, se podrían incluir en esta infraestructura datos o productos derivados que la comunidad científica de forma voluntaria decida liberar.

En segundo lugar, la aportación constante de nuevos datos y metadatos es primordial, tanto de los derivados de la realización de nuevas campañas oceanográficas como de los datos que según normativa se irán liberalizando paulatinamente.

La inclusión de los datos del B/O García del Cid y del B/O Mytilus e incluso de otras plataformas como *Autonomous Underwater Vehicles (AUV, 2018)* también es un objetivo a cumplir de forma inmediata.

En definitiva, el nuevo Servicio de

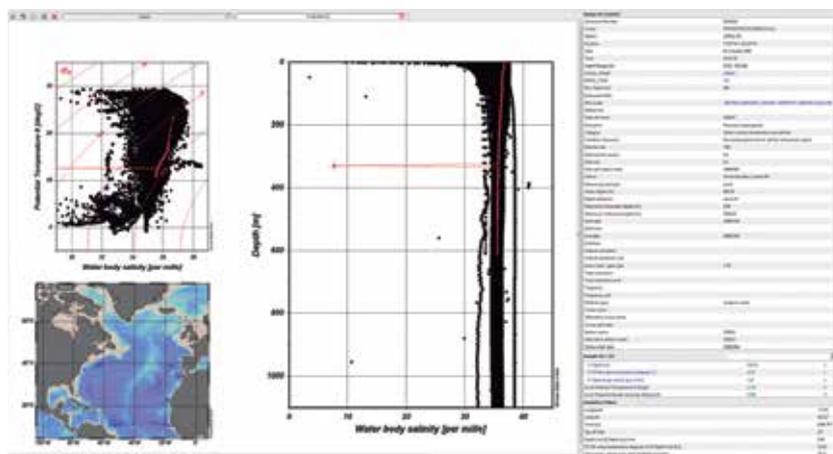


Figura 13. Aplicación Ocean Data View

Datos de la UTM pretende consolidar y reforzar su infraestructura con el objetivo de afrontar nuevos retos, aunando esfuerzos con organismos afines, para de este modo constituirse a medio-corto plazo como un recurso esencial para la oceanografía española, europea e internacional.

REFERENCIAS

- AUV (2018). Vehículos Submarinos Autónomos de la UTM. Recuperado de: <http://www.utm.csic.es/es/servicios/auv>
- B/O García del Cid (2018). Recuperado de: <http://www.utm.csic.es/es/instalaciones/gdc>
- B/O Mytilus (2018). Recuperado de: <http://www.utm.csic.es/es/instalaciones/myt>
- B/O Sarmiento de Gamboa (2018). Recuperado de: <http://www.utm.csic.es/index.php/es/instalaciones/sdg>
- BIO Hespérides (2018). Recuperado de: <http://www.utm.csic.es/es/instalaciones/hes>
- BODC (2016). British Oceanographic Data Center. National Environment Research Council, NERC Vocabulary Server. Recuperado de: https://www.bodc.ac.uk/resources/products/web_services/vocab
- Catálogo de Campañas Oceanográficas de la UTM (2018). Recuperado de: <http://data.utm.csic.es/cruise>
- CNDP-IGME (2012). Centro Nacional de Datos Polares del Instituto Geológico y Minero de España. Recuperado de: <http://hielo.igme.es>
- COCSABO (2018). Comisión de Coordinación y Seguimiento de las Actividades de los Buques Oceanográficos. Recuperado de: <http://www.ciencia.gob.es/portal/site/MICINN/COCSABO>
- CSIC (2018). Consejo Superior de Investigaciones Científicas. Recuperado de: <http://www.csic.es>
- DataCite (2018). Organisation that provides persistent identifiers (DOIs) for research data. Recuperado de: <https://www.datacite.org/mission.html>
- Digital CSIC (2018). Repositorio del Consejo Superior de Investigaciones Científicas. Recuperado de: <https://digital.csic.es>
- EDMO (2018). European Directory of Marine Organisations. Recuperado de: <https://www.seadatanet.org/Metadata/EDMO-Organisations>
- EMODnet (2018). European Marine Observation and Data Network. Recuperado de: <http://www.emodnet.eu>
- EMODnet bathymetry portal (2018). Recuperado de: <http://www.emodnet-bathymetry.eu>
- García, O.; Hernández, A; Ruiz, J.L.; Sorribas, J. et al. (2016). Data Acquisition in Oceanographic Research Vessels. In: Third International Conference on Advances in Mesh Networks, Venice, Italy. DOI: 10.1109/MESH.2010.25
- GEBCO (2018). General Bathymetric Chart of the Oceans. Recuperado de: <https://www.gebco.net>
- Geonetwork (2018). Open-source catalog application to manage spatially referenced resources. Recuperado de: <https://geonetwork-opensource.org>
- GeoNode (2012). Open-source geospatial content management system. Recuperado de: <http://geonode.org>
- GeoNode (2017) como Sistema Base para la Implementación de una Infraestructura de Datos Espaciales. Recuperado de: <http://www.geoidep.gob.pe/mapeandoperu/presentaciones2017/03-02-SEG-DI-geonode.pdf>
- Geoserver (2018). Open source server for sharing geospatial data. Recuperado de: <http://geoserver.org>
- GEOSS (2018). Group on Earth Observations. Recuperado de: <https://www.earthobservations.org>
- INSPIRE (2018). Infrastructure for spatial information in europe. Recuperado de: <https://inspire.ec.europa.eu>
- Mikado (2018). Software tool for SeaDataNet metadata edition. Recuperado de: <https://www.seadatanet.org/Software/MIKADO>
- ODC (2018). Open Data Commons. Recuperado de: <https://opendatacommons.org>
- ODIP (2018). Ocean Data Interoperability Platform. Recuperado de: <http://www.odip.eu>
- OGC (2018). Open Geospatial Consortium. Recuperado de: <http://www.opengeospatial.org>
- PyCSW (2013). OGC CSW server implementation written in Python. Recuperado de: <http://pycsw.org>
- Pydio (2018). Open-source file-sharing running on server. Recuperado de: <https://pydio.com>
- QGIS (2018). Aplicación GIS desktop de software libre. Recuperado de: <https://www.qgis.org>
- SeaDataCloud H2020 Project (2018). Recuperado de: <https://www.seadatanet.org/About-us/SeaDataCloud>
- SeaDataNet (2018). Pan-european infrastructure for ocean & marine data management. Recuperado de: <https://www.seadatanet.org>
- SeaDataNet Common Data Index Catalog V3 (2018). Recuperado de: http://seadatanet.maris2.nl/v_cdi_v3/search.asp
- SeaDataNet Cruise Summary Report Inventory (2018). Recuperado de: http://seadata.bsh.de/Cgi-csr/retrieve_sdn2/start_sdn2.pl

SeaDataNet Data Transport Formats (2018). Recuperado de: <https://www.seadatanet.org/Standards/Data-Transport-Formats>

SeaDataNet Software (2018). Recuperado de: <https://www.seadatanet.org/Software>

Servicio de Datos de la UTM. <http://www.utm.csic.es/es/dataservice-data@utm.csic.es>

Sorribas, J.; Hernández, A.; Ruiz, J.L. et al. (2009). Real-time fleet ship monitoring system using satellite broadband communications and google earth. In:

Advances in Satellite and Space Communications. SPACOMM, Colmar, France. DOI: 10.1109/SPACOMM.2009.44

UTM (2018). Unidad de Tecnología Marina. Recuperado de: <http://www.utm.csic.es>

UTM (2018). Visualización en tiempo real de los datos de adquisición continua. Recuperado de: <http://www.utm.csic.es/es/instalaciones/sdg/realTimeData>

SET (2018). Directorio de recursos de campaña de la UTM. Recuperado de: <http://data.utm.csic.es/set>

Sobre los autores

Juan Luis Ruiz Valderrama

Licenciado en Ingeniería Electrónica por la Universidad de Barcelona (UB), trabaja desde el año 2001 en la Unidad de Tecnología Marina (UTM) del Consejo Superior de Investigaciones Científicas (CSIC) en el ámbito de las Tecnologías de la Información y las Comunicaciones. Ha participado en más de 20 expediciones en buques oceanográficos y bases antárticas dando apoyo a proyectos en diferentes disciplinas científicas como geofísica, oceanografía, climatología, biología, cartografía y glaciología. Desde 2016 trabaja en el diseño y puesta en marcha de la IDE para el nuevo Servicio de Datos de la UTM y colabora en proyectos nacionales e internacionales relacionados con la gestión de datos de investigación. Es coautor de 13 comunicaciones a congresos y publicaciones sobre trabajos técnicos en plataformas de investigación oceanográfica y polar, 3 de ellas como primer autor.

Susana Díez Tagarro

Licenciada en Físicas por la Universidad de Barcelona (UB), trabaja desde el año 2000 en la Unidad de Tecnología Marina (UTM) del Consejo Superior de Investigaciones Científicas (CSIC) en la adquisición, control de calidad y procesamiento de datos oceanográficos. Actualmente coordina el Servicio de Datos de la UTM encargado de la gestión y disseminación de metadatos y datos adquiridos en campañas oceanográficas. Desde 2015 es representante del CSIC en el Grupo de Trabajo de Cartografía Marina (GTCM) perteneciente a la Comisión Interministerial de Estrategias Marinas (CIEM) y en el Grupo de Trabajo Técnico 15 y 16 (GTT1516) para implementar la normativa INSPIRE. Desde 2018 también forma parte del Grupo de Trabajo de Datos Marinos (GTDM) de la CIEM. Ha participado en 15 campañas

oceanográficas a bordo de diferentes buques de investigación nacionales y extranjeros y es autora o coautora de 24 publicaciones.

Alberto Hernández Jiménez

Ingeniero en Telecomunicaciones por la Universidad de Alcalá (UAH) y Master en Software Libre por la Universitat Oberta de Catalunya (UOC), trabaja desde el año 2008 en la Unidad de Tecnología Marina (UTM) del Consejo Superior de Investigaciones Científicas (CSIC) en el ámbito de las Tecnologías de la Información y las Comunicaciones. Ha participado en más de 25 expediciones en buques oceanográficos y bases antárticas dando apoyo a proyectos en diferentes disciplinas científicas como geofísica, oceanografía, climatología, biología, cartografía y glaciología. Es autor o coautor de 14 publicaciones técnicas y divulgativas.

Jordi Sorribas Cervantes

Titulado Superior en Geociencias por la Universidad de Barcelona (UB) y DEA en Telemática por la Universidad Politécnica de Catalunya (UPC). En la actualidad, director de la Unidad de Tecnología Marina del CSIC. Desempeñando desde 1997 responsabilidades en la implantación de los sistemas de información en los buques oceanográficos que gestiona el CSIC, impulsando entre otros proyectos el empleo de sistemas de comunicación de banda ancha a bordo de los buques de investigación, la gestión automática de adquisición de datos oceanográficos de propósito general, introduciendo a los buques en el catálogo paneuropeo de plataformas de generación de datos marinos (SeaDataNet) como buques de referencia que ofrecen a la comunidad científica mundial datos del ámbito marino en tiempo real. Entre 2006 y 2013 autor principal y coautor en una quincena de comunicaciones a congresos y publicaciones sobre tecnologías de Información en plataformas de investigación oceanográfica y polar.