

Definición de Perfiles y creación de ficheros XML de metadatos para imágenes de Teledetección, según la normativa ISO, utilizando la aplicación IME (ISO Metadata Editor) desarrollada en el Servicio de Teledetección del INTA.

Amaro Cormenzana, Alberto¹
Dominguez Barroso, Nines²
Prado Ortega, Elena³

Instituto Nacional de Técnica Aeroespacial (I.N.T.A.) (España) <mailto:amaroca@inta.es>¹, <mailto:dominguezba@inta.es>², <mailto:pradooe@inta.es>³

Resumen: *El objetivo de la presentación es mostrar los pasos seguidos en el Servicio de Teledetección del INTA para incorporar ficheros XML de metadatos a las imágenes de teledetección, respetando la normativa ISO y las recomendaciones del Nucleo Español de Metadatos (N.E.M.).*

Utilizando la aplicación IME (ISO Metadata Editor) desarrollada en el INTA, se ha creado y depurado un perfil de metadatos que resulte aplicable a los datos de las imágenes hiperespectrales. Para ello se ha seguido la norma ISO19115 completándolo con las recomendaciones del NEM, así como con extensiones de aplicación específica a imágenes adquiridas con sensores aeroportados.

Esta herramienta también permite conocer y entender la jerarquía de metadatos que se define en la ISO19115:2003, por lo que resulta interesante mostrar sus posibilidades para los usuarios que necesiten acceder a esta información.

Una vez definido el perfil, se ha trabajado con las posibilidades de edición de metadatos que permite el software IME para generar plantillas (en inglés y castellano) que servirán como referencia para futuros proyectos.

Como última fase de esta metodología se han generado ficheros XML según los requerimientos del último borrador de la especificación ISO19139. Estos ficheros de metadatos han sido distribuidos a los usuarios para documentar las últimas campañas de vuelo realizadas (Doñana, Bélgica, Albacete y Rosarito).

INTRODUCCIÓN

La directiva europea INSPIRE que pretende regular y armonizar la distribución de Información Geográfica en Europa, ha dado lugar a la creación de las Infraestructuras de datos Espaciales (IDE) en los distintos países de la Unión Europea. En el Servicio de Teledetección del INTA, como productor de datos de Teledetección tanto espacial como aeroportada, se ha querido, desde el primer momento, incorporar la nueva normativa sobre distribución de Información Geográfica, para lo cual se ha desarrollado un perfil de metadatos lo más estandarizado posible para los productos distribuidos.

El objetivo ha sido generar un perfil que sea de utilidad para proporcionar información a los usuarios de nuestros datos, tanto de aquellos proporcionados por CREPAD (Centro de REcepción, Proceso, Archivo y Distribución de datos de observación de la Tierra, situado en la Estación Espacial de Maspalomas) como de los datos aeroportados producidos en el Servicio de Teledetección (LABTEL). Además, este perfil tiene que cumplir los requisitos exigidos por la ISO19115, en cuanto a contenido, y de la ISO19139 en cuanto a estructura.

METODOLOGÍA

Para mantener la coherencia con otros organismos y con la IDEE (Infraestructura de Datos Espaciales de España) que lidera el IGN (Instituto Geográfico Nacional), se utilizó como punto de partida la norma internacional ISO19115 y se completó el perfil con la primera versión del NEM. Concretamente los pasos seguidos fueron:

1. Estudio de la ISO19115 y selección de los metadatos de interés.
2. Desarrollo de una aplicación de apoyo para la generación de perfiles de metadatos y ficheros XML de acuerdo con las normas ISO19115 e ISO19139.

3. Generación de un perfil de metadatos propio.
4. Comprobación de nuestro perfil con el NEM y generación del perfil definitivo.
5. Creación de plantillas.
6. Estudio de la ISO19139 y utilización del esquema XML propuesto hasta la fecha por dicha norma.

Como apoyo a todo este proceso se diseñó una herramienta de trabajo que, al mismo tiempo que facilitaba la comprensión de una norma tan compleja como la ISO19115, permitía ajustar y modificar el perfil del Servicio de Teledetección y obtener una salida en formato XML de los metadatos ya completados.

El software, denominado IME (ISO Metadata Editor) es propiedad del INTA y se distribuye gratuitamente a cualquier usuario que lo solicite o lo desee descargar desde el enlace: <http://www.crepad.rcanaria.es/info/metadatos/index.htm>

Esta herramienta servirá para mostrar, en esta presentación, los pasos seguidos.

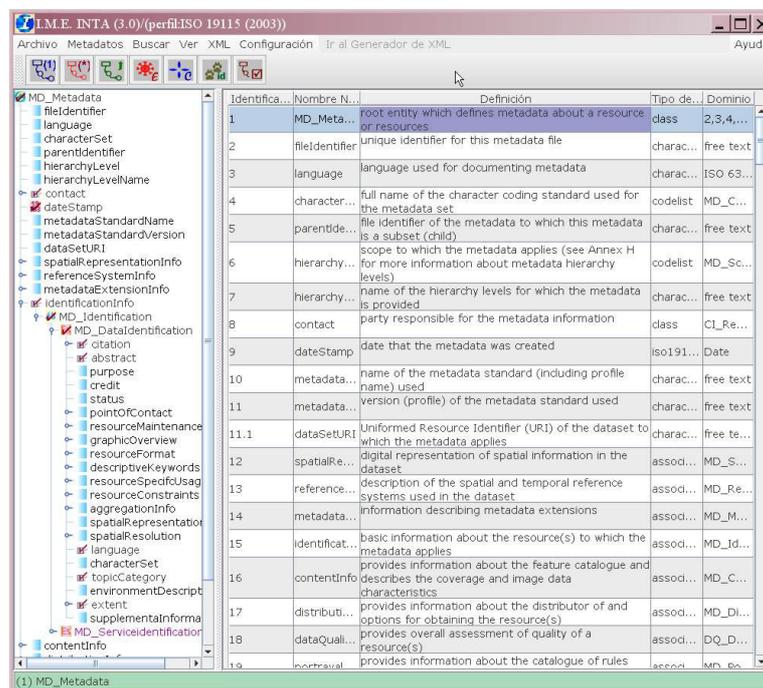


Figura 1: Ventana principal software IME version 3.0

Selección de metadatos de la ISO19115 para el perfil de los datos de Teledetección

Para confeccionar el perfil INTA se han seleccionado aquellas clases y elementos enumerados en la ISO19115 que tienen interés para describir el contenido de las imágenes aeroportadas producidas en el Servicio de Teledetección, así como las imágenes y productos de satélite proporcionados a través de CREPAD.

Para nuestro perfil se seleccionaron las siguientes entidades (figura 2):

MD_Metadata->*spatialRepresentationInfo* para proporcionar información sobre el número de filas y columnas de las imágenes.

MD_Metadata->*referenceSystemInfo* para establecer los parámetros de la proyección cartográfica de las imágenes cuando están georreferenciadas.

MD_Metadata->*identificationInfo* donde se describen características de los datos y su uso.

MD_Metadata->*contentInfo* que contiene una clase específica para descripción de imágenes y que permite incluir información sobre el parámetro contenido por los píxeles, sobre las bandas de la imagen y también sobre las condiciones de adquisición y el nivel de proceso.

MD_Metadata->*distributionInfo* que permite informar sobre el formato de distribución.

MD_Metadata->*dataQualityInfo* que incluye información sobre la precisión, así como sobre la fuente original de los datos y la descripción de cada uno de los procesos seguidos hasta llegar al nivel actual.

Además se han incluido aquellos elementos de la entidad *MD_Metadata* que permitan proporcionar información sobre los propios metadatos (nombre del fichero, versión de la norma, responsable de los metadatos, etc.).

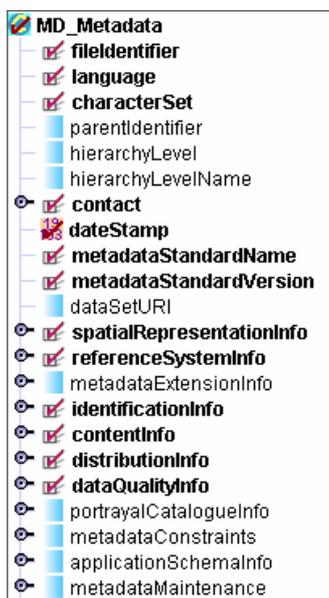


Figura 2: Metadatos y Entidades seleccionados

Desarrollo de la herramienta IME para la aplicación de la normativa ISO

Debido a la complejidad de la normativa ISO, se decidió crear una herramienta propia que facilitara el trabajo con un gran número de metadatos y permitiera mantener la jerarquía que define la norma. El objetivo fundamental era poder gestionar perfiles de metadatos sin necesidad de conocer todas las asociaciones y dependencias entre ellos que define la norma ISO19115.

La herramienta IME fue diseñada con dos premisas fundamentales:

- Ser un reflejo de la norma ISO19115 para permitir crear perfiles de metadatos diferentes según las necesidades de cada usuario.
- Funcionar en un entorno de uso extendido y ser fácilmente adaptable a los cambios en la normativa.

Entre las principales utilidades que ofrece se incluyen:

- Identificación del tipo de clase (especificada, abstracta, clase,...) mediante un icono (ver figura 3).
- Identificación de metadatos definidos en otras normas relacionadas con la ISO19115
- Visualización del camino completo para localizar un metadato (ver figura 4).
- Activación y desactivación de metadatos que forman el perfil (ver figura 5).
- Búsqueda de metadatos.
- Lectura y creación de ficheros de metadatos XML y plantillas de texto.

El diseño del software IME se realizó con la intención de facilitar a cualquier usuario la incorporación de modificaciones a la definición de los metadatos y de la salida en XML sin necesidad de modificar y recompilar el código fuente.

El software cuenta con ficheros editables de configuración que permiten realizar los siguientes cambios:

- Incorporación de nuevas entidades y metadatos.
- Modificación de las propiedades de los metadatos definidos.
- Modificación de los metadatos que aparecen activados al iniciarse la aplicación.
- Modificación de los espacios de nombres definidos en la ISO19139.

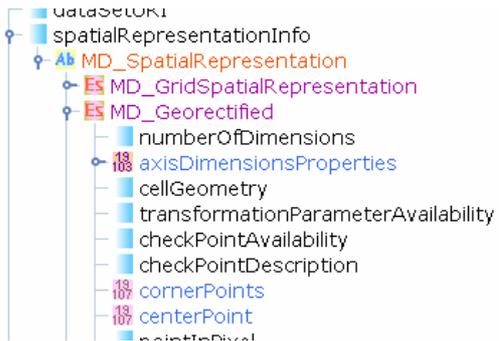


Figura 3: Iconos para identificación de Tipo de Dato

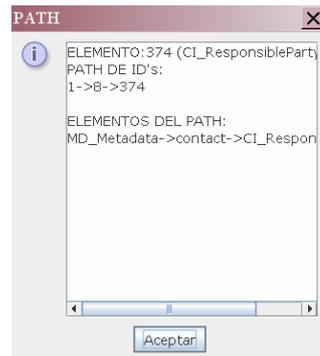


Figura 4: Path desde el elemento raíz hasta un metadato

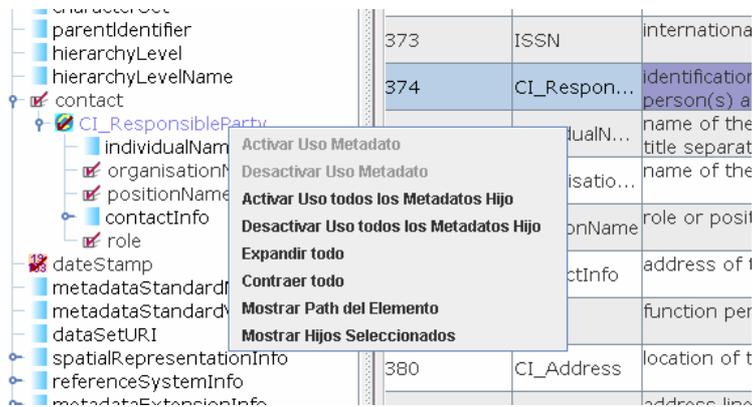


Figura 5: Menú de selección de metadatos para configuración de un Perfil

Generación de un perfil propio para datos de Teledetección

Además de las entidades de metadatos enumeradas anteriormente, fue necesario incluir otras que podían ser de interés para los usuarios de los datos aeroportados producidos en el Servicio de Teledetección. Se creó una nueva entidad siguiendo la metodología proporcionada por la ISO para la extensión de perfiles de metadatos.

Como se puede observar en la figura 6, se generó una nueva entidad denominada *acquisitionInformation* constituida, a su vez, por: *platformId*, para proporcionar información sobre la plataforma de adquisición; *instrumentId*, para describir información útil sobre el instrumento de adquisición de las imágenes; y *flightInfo*, que incluye datos sobre altura, dirección y velocidad de vuelo. Tanto los nombres de estas clases como las características de los metadatos se han basado en borradores de especificaciones ISO (19115-2, 19130, etc.) para facilitar su posterior adaptación a estas normas.

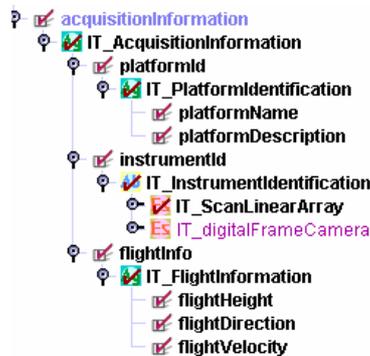


Figura 6: Extensión para datos de sensores aeroportados

Comparación del perfil de Teledetección con el NEM

En el Servicio de Teledetección se empezó a desarrollar un perfil aplicado exclusivamente a teledetección a finales del 2004. La posterior aparición del NEM resultó muy útil para comprobar que el perfil que se había desarrollado seguía las pautas correctas y para ampliar aspectos de los metadatos que no se habían considerado, tales como la información sobre el *dataset* agregado. También sirvió para entender el significado de algunas de las clases y elegir la que más se adaptaba a nuestros datos.

Debido a que el NEM ha sido pensado para datos de distinta naturaleza a la de las imágenes de teledetección, encontramos diferencias en varias entidades. Las más importantes se describen a continuación.

El perfil del Servicio de Teledetección incluye las clases *MD_SpatialRepresentation* y *MD_ContentInformation*, no contempladas en el NEM, con el objetivo de describir los datos de teledetección correctamente.

La entidad *MD_SpatialRepresentation* (fig.7) incorpora la información sobre la representación espacial de los datos. Dentro de las cuatro entidades que ofrece *spatialRepresentation*, se ha elegido la clase específica *MD_GridSpatialRepresentation* porque es la que más se adapta de manera general a todas las imágenes y productos que proporcionamos. La información fundamental aquí, se refiere a la descripción de los ejes, que se denominan dimensiones. El número de dimensiones espaciales (*numberOfDimensions*) en las imágenes es 2, correspondientes a filas y columnas, y las propiedades asociadas a estas dimensiones serían nombre (*dimensionName*) y tamaño o número de unidades (*dimensionSize*). Lo más correcto sería elegir *MD_GridSpatialRepresentation* para las imágenes raw; *MD_Georectified* para las imágenes corregidas y *MD_Georeferenceable* para las imágenes que se acompañan de ficheros para su georreferenciación. Sin embargo, esto supondría crear tres perfiles en lugar de uno para los datos aeroportados, lo que se traduciría también en tres plantillas. En la versión actual del perfil se decidió poner *0-No*, en *transformationParameterAvailability*, para indicar que la imagen no es georreferenciable ni está georreferenciada, y *1-Yes*, para indicar que la imagen es georreferenciable o está georreferenciada. Además, para indicar que está georreferenciada existe un elemento en *referenceSystemInfo* donde se indica la proyección de la imagen. Por último, existe la clase *MD_AggregateInfo* donde se puede incluir información sobre los ficheros auxiliares de georreferenciación, en caso de que existan.

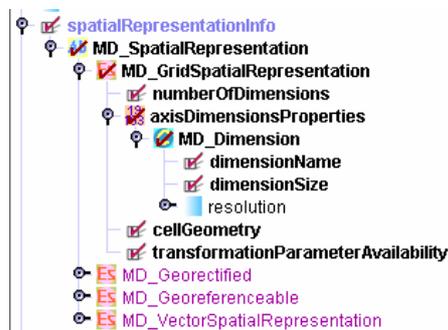


Figura 7: MD_SpatialRepresentation

En el caso de *MD_ContentInformation* (figura 8) se proporciona información sobre el contenido de los datos. Contiene una sola clase abstracta que a su vez puede ser sustituida por cuatro clases específicas, siendo la más interesante la denominada *MD_ImageDescription* que permite incluir información sobre las bandas del sensor (*MD_Band*), cobertura nubosa (*cloudCoverPercentage*) y nivel de proceso (*processingLevelCode*).

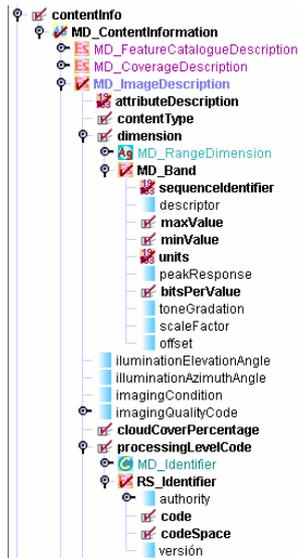


Figura 8: MD_ContentInformation

Como breve referencia al resto de las entidades:

- En *referenceSystemInfo* sólo se incluye la descripción del sistema de referencia espacial mediante la entidad *MD_CRS*.
- Se utiliza *identificationInfo*, al igual que en el NEM, para incluir toda la información relativa a los datos en cuanto a restricciones, palabras clave, referencias de contacto, uso, extensión espacial y temporal, etc...
- En *distributionInfo* se proporciona información sobre la distribución de los datos, es decir, cómo y quién los distribuye. Sin embargo, dado que la mayor parte de esta información ya ha sido proporcionada en otras secciones del perfil, en este caso sólo se ha añadido información del formato.
- En *dataQuality* también se encuentran diferencias respecto al NEM, en la figura 9 se pueden ver los metadatos con los que se ha trabajado en este perfil:

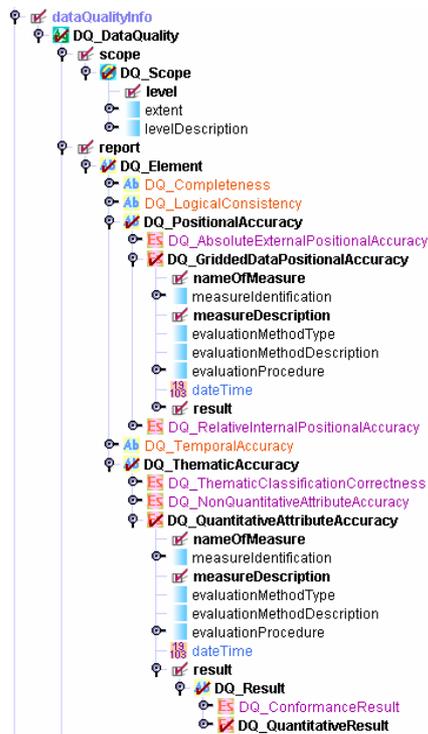


Figura 9: DQ_DataQuality

Por último, como ya se ha mostrado en la figura 6, ha sido necesario añadir una extensión a los metadatos de la ISO19115 .

Creación de Plantillas y generación de ficheros XML

Una vez definido el perfil, se utilizó el editor de metadatos de la herramienta IME para la introducción de los valores y creación de las plantillas. Esta utilidad permite la edición de texto libre en cualquier metadato editable y la visualización de las opciones disponibles en los "codelists" que, para otros metadatos, definen los únicos valores posibles (ver figuras 11 y 12).

También se incluyeron algunas utilidades para facilitar el duplicado de metadatos (con verificación previa de la norma), borrado y copiado de nodos entre varias sesiones de IME abiertas simultáneamente (ver figura 10).

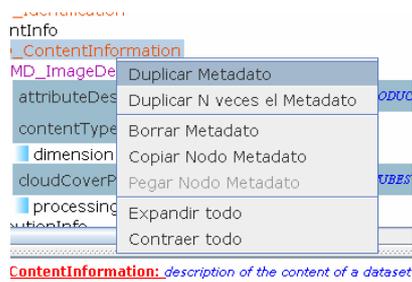


Figura 10: Submenú del editor de metadatos

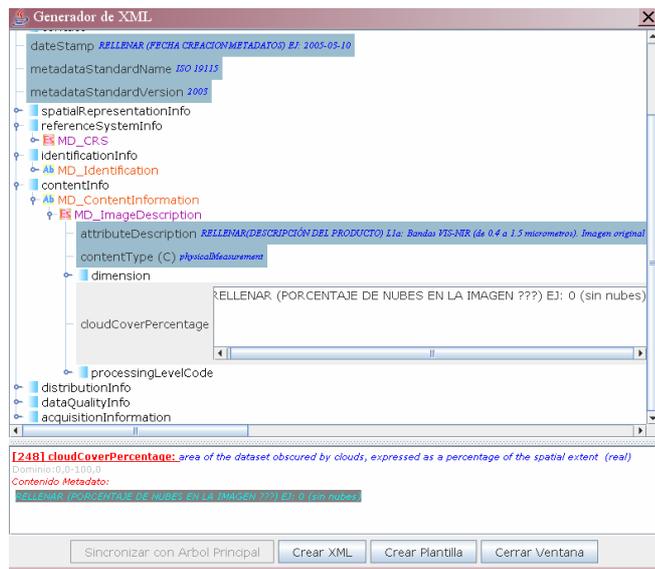


Figura 11: Edición de metadatos con texto libre

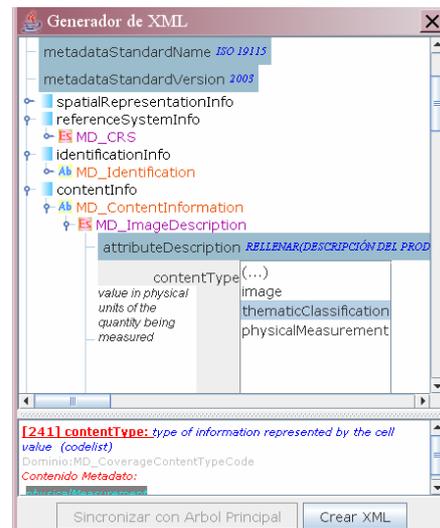


Figura 12: Selección de opciones en codelists

Para completar la plantilla, se decidió:

- Escribir directamente en este Editor los valores que se mantienen constantes en diferentes campañas de vuelo.
- Incluir comentarios que faciliten la localización de los metadatos que varían en cada campaña.

Al mismo tiempo se definió un formato de fichero (plantilla de texto, figura 13) en el que se incluyera toda la información del perfil y que fuera fácilmente editable. De este modo un usuario puede incluir la información manualmente y puede acceder a ella desde otra aplicación para automatizar el proceso del relleno de los datos.

```

TPL_AHS_PTALL_spa.tpl
# (1) fileIdentifier [2]
# MD_Metadata->fileIdentifier
# Dominio:free text
RELLENAR (NOMBRE FICHERO METADATOS) EJ: MD_AHS_040921_DOANA_PO5AD_L10020_PT1.XML
# (2) language [3]
# MD_Metadata->language
# Dominio:ISO 639-2 or others.
spa
# (3) characterSet [4]
# MD_Metadata->characterSet
# CodeList: (Ucs2,Ucs4,Utf7,Utf8,Utf16,8859part1,8859Part2,8859Part3,8859Part4,8859part5,88
(reserved for future use),8859part13,8859part14,8859part15,8859part16,Jis,ShiftJIS,EucJP,us
8859part1
# (4) individualName [375]
# MD_Metadata->contact->individualName
# Dominio:free text
Edwards de Miguel

```

Figura 13: Formato de plantilla de metadatos generada con IME

El paso final fue la obtención de un fichero XML utilizando la misma aplicación.

A la espera de la publicación definitiva de la norma ISO19139, se ha buscado una aproximación lo más fidedigna posible al último borrador disponible siguiendo la estructura del esquema XSD (Schema Definition Language) propuesto. Esta ha sido una tarea ardua ya que el esquema no sigue la estructura exacta de la ISO19115 en cuanto a distribución de las entidades y de los elementos dentro de ellas.

Los principales problemas encontrados han sido:

- La norma ISO 19139 todavía está en desarrollo y aún no se ha podido contar con un esquema de implementación definitivo.
- Por otra parte se ha comprobado que en algunos casos el esquema actual disponible no se adecúa perfectamente a la norma ISO19115, por ejemplo en el orden que debe seguir una secuencia de metadatos o en el número de veces que puede aparecer un metadato.

RESULTADOS: DISTRIBUCIÓN DE LOS METADATOS DESDE EL SERVICIO DE TELEDETECCIÓN

Las imágenes adquiridas por los sensores aerportados del Servicio de Teledetección (AHS, AMDC y ATM) son procesadas de acuerdo a un flujo de trabajo desarrollado e implementado en el INTA. Estos datos se caracterizan por un nivel de proceso geométrico y radiométrico que da lugar a una serie de productos bien definidos.

Desde que el perfil fue completado, los productos imagen generados para cada proyecto se acompañan de ficheros XML de metadatos que son proporcionados a los usuarios. Actualmente se utilizan una serie de rutinas que permiten obtener, de forma automática, parte de la información contenida en la cabecera de nuestras imágenes y productos e incorporarla a los ficheros XML.

En la siguiente tabla se muestra un resumen de los proyectos en los que la entrega de imágenes hiperespectrales se ha completado con ficheros de metadatos adjuntos. Estas campañas se han realizado durante el año 2005 para diferentes usuarios.

ID Campaña	Marco de ejecución	Productos entregados (*)	Propósito	Volumen datos	Comentarios
ROSARITO	----	Imágenes nivel L1a Imágenes nivel L1c	Análisis biofísicos sobre láminas de agua	12 ficheros	----
DOÑANA	Convenio INTA-IGME 2005	Imágenes nivel L1a Imágenes nivel L1b	Análisis cuantitativos y cartografía temática multitemporal sobre humedales	90 ficheros	Se han realizado 3 campañas similares de adquisición de datos en diferentes fechas (2004-2005)
SEN2FLEX	Campaña SEN2FLEX (financiada por	Imágenes nivel L1a Imágenes nivel L1b	Análisis biofísicos estudio de la	126 ficheros	Datos adquiridos en diferentes campañas, el proyecto cuenta con

	ESA a través de la Univ. Valencia)		fluorescencia de la vegetación		múltiples usuarios
VITO	Contrato INTA-VITO para la adquisición de imágenes, Plan Hiperespectral Belgica 2005	Imágenes nivel L1a Imágenes nivel L1b	----	162 ficheros	Múltiples usuarios y zonas test con propósitos de investigación diferentes
HYPERVAL	Proyecto de investigación financiado por el Plan Nacional de I+D (AGL-FOR2002 04017 C03 C1)	Imágenes nivel L1a Imágenes nivel L1b Imágenes nivel L1c	Análisis biofísicos sobre vegetación	42 ficheros	Varias campañas de adquisición sobre diferentes zonas test
GUADA	Convenio INTA-INIA 2005	----	Estudios cuantitativos y cartografía temática sobre un área afectada por un incendio forestal	21 ficheros.	Imágenes en proceso

De momento no se ha recibido ninguna apreciación por parte de los usuarios sobre la facilidad o dificultad de acceder a estos datos entregados en formato XML. En nuestra opinión, es pronto aún para que realicen valoraciones sobre los metadatos proporcionados porque la interoperabilidad entre las aplicaciones no será efectiva mientras no se publique la recomendación definitiva de la ISO19139 . No obstante, sí han mostrado interés en el uso del software IME y en el conocimiento de la normativa.

REFERENCIAS

1. INSPIRE. Propuesta de DIRECTIVA DEL PARLAMENTO EUROPEO Y DEL CONSEJO por la que se establece una infraestructura de información espacial en la Comunidad (INSPIRE).(http://inspire.jrc.it/proposal/ES.pdf)
2. IDEE. Portal de Infraestructura de Datos Espaciales de España. (www.idee.es)
3. ISO/TC211 (Cómite Técnico 211 de la ISO)
4. ISO19115. Geographic Information-Metadata (edición 2003-05-01)
5. ISO1939. Geographic Information-Metadata-Implementation Specification (2004-06-30 estado actual: " CDstudy/ballot initiated")
6. XML. Extensible Markup Language 1.0 (Third Edition. Recommendation 04-02-2004) http://www.w3.org/XML
7. XML Schema Definition v.1.1 (http://www.w3.org/XML/Schema)
8. Web Software IME v.2.0. (Documentacion y Descarga) www.crepad.rcanaria.es/info/metadatos/index.htm