

PROCESO DE PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Francisco Pando - GBIF España

Métodos y Calidad en la Digitalización de Colecciones
CENPAT Puerto Madryn, Argentina 2012



PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

- Metadatos
- Formatos de imágenes de alta resolución: TIFF Piramidal y JPEG 2000
- Protocolo IIP
- Qué es un servidor de imágenes. Soluciones en el mercado
- Visualizadores web de imágenes de alta resolución
- Modelos de trabajo propuestos
- Proyectos y bibliotecas digitales con información sobre biodiversidad en Internet
- Servicio de Alojamiento y Publicación de imágenes de GBIF.es:
 - -Integración del Servicio en la infraestructura de GBIF.es y de GBIF (protocolos, estándares, portales)
 - -Protocolo de uso del Servicio de Alojamiento y Publicación de imágenes de GBIF.es
 - -Ejemplo de uso
- Casos Prácticos:
 - -Caso Práctico I. Integración de imágenes digitalizadas por los alumnos con el visualizador zoomyfer y publicación en nuestra web
 - -Caso Práctico II. Integración de imágenes digitalizadas por los alumnos en el Servicio de Alojamiento y publicación de imágenes de GBIF.ES y publicación en nuestra web

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Metadatos

- **Metadatos:**

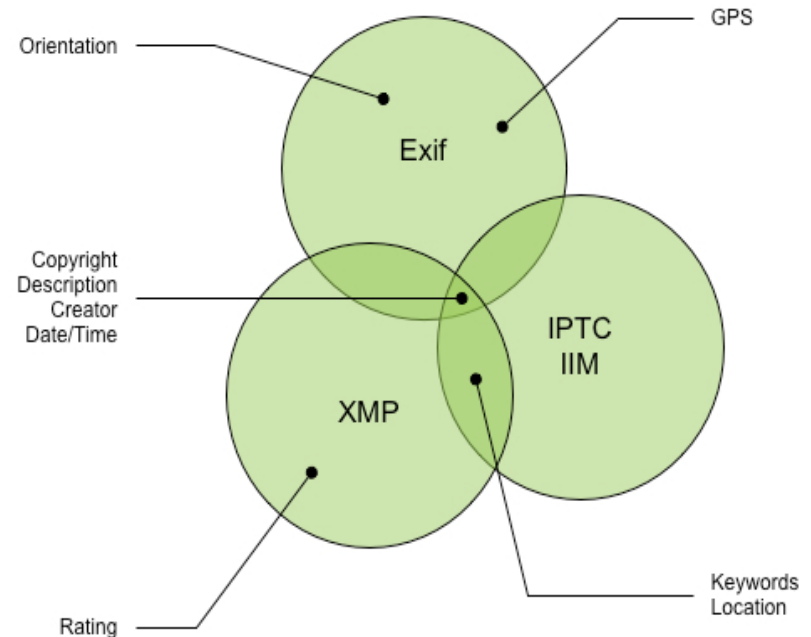
- Organización & Búsqueda
- Qué + Dónde + Cuándo + Cómo + Quién
- embebidos vs externos

- **Distintas especificaciones (EXIF, IFDs, Adobe XMP, PSIR, IPTC-IIM)**

- Reglas cómo almacenar + ordenar + codificar
- Cadenas de caracteres, números, arrays
- Leídos y modificados vs leídos
- Solapamiento información
- Compatibilidad

- **Iniciativas de trabajo. “Metadata Working Group”.**

- 2006 by Adobe Systems, Apple, Canon, Microsoft and Nokia. Sony se unió en in 2008.
- <http://www.metadataworkinggroup.org/>
- interoperabilidad
- disponibilidad



Fuente img: "Guidelines For Handling Image Metadata" - Metadata Working Group

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Metadatos

- EXIF
 - En TIFF, JPEG y PSD files
 - No soportado por PNG o GIF y
 - Estandarizado en XML para JPEG 2000 según ISO TC42 (12234-1.4.)
 - Date & time, camera settings (camara model, orientation, aperture...), thumbnail
 - Usado por cámaras fotográficas.
 - version 2.3 Abril 2010
- IPTC-IIM (*Information Interchange Model*)
 - En JPEG, TIFF
 - No soportado por GIF, PNG
 - <http://www.iptc.org>
 - "IPTC headers" ampliamente adoptado por software para edición
- XMP (*Extensible Metadata Platform*)
 - <http://www.adobe.com/products/xmp>
 - XML y RDF
 - Información de procesos por los q la imagen pasa. Dispositivos y software
 - Dublin core +Elementos extensibles + propios
 - En PDF, JPEG, JPEG 2000, GIF, PNG, HTML, TIFF, Adobe Illustrator, PSD, MP3, MP4,...
 - "IPTC Core Schema for XMP"

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Metadatos

Metadatos biodiversidad

TDWG. Image Interest Group-IIG

<http://www.tdwg.org/activities/img/charter/>
<http://wiki.tdwg.org/Image>

Estándar MRTG:

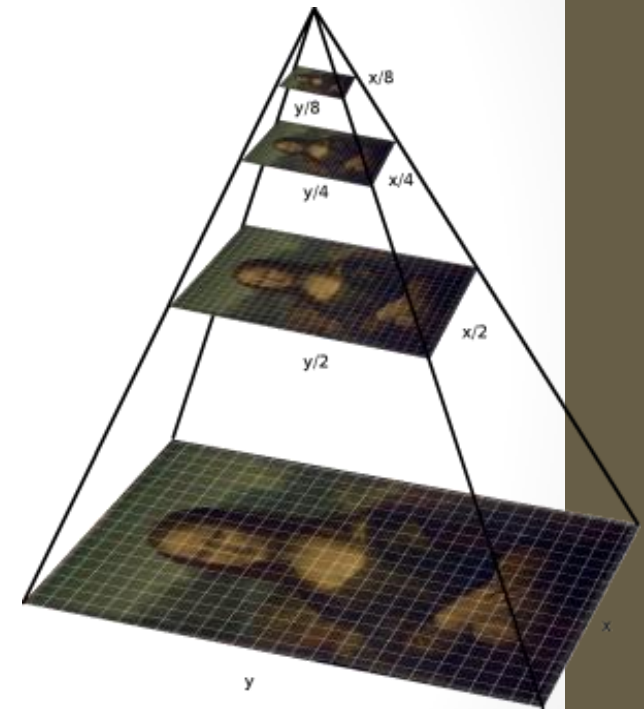
- Estado: Draft
- http://www.keytonature.eu/wiki/MRTG_v1.0 (AudubonCore v1.0)
- Vocabularios para describir metadatos asociados a multimedia en fuentes de datos biodiversidad
- Dublin Core(DC), Darwin Core (DwC), XMP, IPTC, Natural Collections Schema (NCD), etc.

Biodiversity
Information
Standards
T D W G

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

TIFF piramidal

- Formato multi-resolución, datos para múltiples resoluciones y lectura de porciones en distintos niveles de resolución a gran velocidad
- *Almacenamiento en secciones rectangulares.* (múltiplo de 16 pixeles; mismo tamaño (ancho y alto) todas las secciones)
- Comprimidas individualmente (JPEG,LZW,PAckBits, Huffman)
- Para imágenes de alta resolución
- Especificaciones IIP limitan el tamaño de las secciones a 64×64 pixeles



Fuente de la imagen:

<http://iipimage.sourceforge.net/documentation/images/>

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

TIFF piramidal

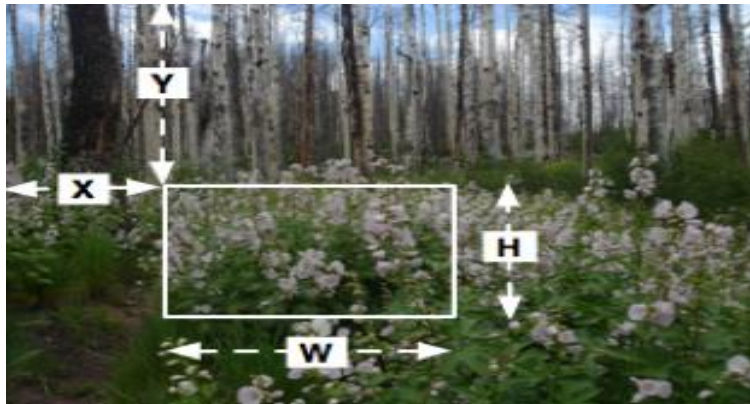
- Espacios de color de CMYK y CIELAB
- Perfiles ICC embebidos, transparencia, capas de Photoshop
- 16-bit por canal de datos
- Procesadores:
 - VIPS image processing suite <http://vips.sourceforge.net>
 - ImageMagick <http://www.imagemagick.net/script/index.php>
 - Implementaciones: JMagick, IMagick, PythonMagick
 - Desde la línea de comandos.
 - Automatización
- Metadatos: EXIFF, IPTC y XMP

TIFF									
TIFF IFD0									
34665	Exif								
700	XMP								
33723	IPTC								
34377	PSIR								
	<table border="1"><tr><td>1061</td><td>IIM Hash</td></tr><tr><td>1028</td><td>IPTC</td></tr><tr><td>1058</td><td>Exif</td></tr><tr><td>1060</td><td>XMP</td></tr></table>	1061	IIM Hash	1028	IPTC	1058	Exif	1060	XMP
1061	IIM Hash								
1028	IPTC								
1058	Exif								
1060	XMP								

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

JPEG 2000

- Segunda versión del estándar JPEG (compresión y formato). ISO/IEC 29 WG01
- <http://www.jpeg.org/jpeg2000/index.html>
- Algoritmo DWT (wavelet technology)
- Compresión
 - 20%+
 - menor degradación
 - lossless vs lossy
- Multiresolución.
 - Code/Decode por regiones: visualización y transmisión a diferentes calidades & navegación por la imagen
 - Rápida accesibilidad. Transmisión progresiva en:
 - Precisión (pixel)
 - Resolución



Fuente img: <http://www.dlib.org/dlib/september08/chute/09chute.html>

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
JPEG 2000

- **Contenedores; XML embebido; pdf**
- **Metadatos: XMP, IPTC, GML (OpenGIS), DublinCore**
- **Recomendado + de 50 megapixel**
- **También para sonido y vídeo**
- **Conversores para JP2:**
 - **Kadaku. <http://www.kakadusoftware.com/> (encoding, transcoding, viewing, metadata & file structure display) (Windows, Mac, Linux)**
 - **Jasper (Open source) <http://www.ece.uvic.ca/~mdadams/jasper/>**
 - **ExpressView browser plug-in (todavía no hay soporte nativo en navegadores) http://www.lizardtech.com/download/dl_options.php?page=viewers**

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Internet Imaging Protocol (IIP)

- Digital Imaging Group (DIG)
- Cliente-Servidor
- Sobre HTTP (imágenes y metadatos)
- Trabaja sobre “multiresolution tiled images” (TIFF piramidal, FlashPix, JPEG 2000)
- Descarga por cliente de porción de imagen (a determinada resolución)
- Servidas y comprimidas dinámicamente en JPEG
- Actualmente version 1.05.
- Especificaciones: <http://iipimage.sourceforge.net/IIPv105.pdf>
<http://iipimage.sourceforge.net/documentation/protocol/>

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Qué es un Servidor de imágenes

- **Software especializado en servir imágenes. (creadas por demanda basadas en una petición específica)**
- **Arquitectura cliente-servidor (API's públicas)**
- **Almacenamiento (formato base) y generación dinámica de x formatos (tamaños/resolución).**
- **Herramientas para navegación en tiempo real por imágenes de tamaño mega y gigapixel.**
- **Optimización tráfico cliente – servidor (ancho de banda) y eficiencia en gestión de la memoria del servidor**
- **Procesos extra: gestión de metadatos, ajustes de color, usuarios y grupos para control de acceso etc**
- **Ajuste dinámico:**
 - **tamaño de imagen - resolución navegador.**
 - **formatos - tipo/versión del navegador**
- **Formatos base: TIFF (tiled) piramidal, FlashPix/OpenPix, JPEG2000**

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Servidores de imágenes: Software

IIPImage



<http://iipimage.sourceforge.net/> Open Source

Servidor: C++

Varios clientes: AJAX, flash, JAVA applet

Imágenes 8/16 bit

Imágenes alta resolución (mega, gigapixel size), multiespectrales, objetos 3D

Formato generado: JPEG

Soporte JPEG2000 (Kakadu JPEG2000 library) desde enero 2010

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Servidores de imágenes: Software

DJATOKA

Open source. <http://sourceforge.net/projects/djatoka/>

Java

Version 1.1 2009

Formato base: JPEG 2000

Kakadu JPEG 2000 libreria (Extracción dinámica en múltiples resoluciones de regiones de archivos jpeg 2000)

Formatos de entrada/salida: BMP, GIF, JPG, PNG, PNM, TIF, JPEG 2000

Service framework, basado en “OCLC OpenURL Resolver”

Caché configurable

Linux/UNIX/MacOS-X/Windows

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Servidores de imágenes: Software

<http://www.yawah.com/> (Requiere licencia) (eRez)



ADOBE® SCENE7®

<http://billeder.silkeborgkommune.dk/erez4/html/index.html>

Servidor: Java

Cliente: Flash, AJAX, Javascript

Gestión de metadatos asociados (IPTC y EXIFF). Interfaz de búsqueda basada en metadatos

Formato base TIFF piramidal, FlashPix o PSD

Carga de imágenes (upload) en diferentes formatos: JPEG, EPS, GIF PDF, AI, SVG y CAD y otros implementando plug-ins. También .zip

Formatos generados: JPEG, GIF, PNG, Flash, PDF y TIFF a través de http requests

Gestión de directorios, usuarios y roles

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Servidores de imágenes: Software

Otros :

- iSeeMedia: <http://www.iseemedia.com/>

The logo for iSeeMedia, featuring the text "i see" in a white sans-serif font, a circular icon with a stylized eye, and the word "media" in a white serif font, all set against a dark blue rectangular background.

- MediaRich Server:
<http://www.equilibrium.com/Internet/Equil/Products/MediaRich/Product+Tour/index.htm>

The logo for MediaRich Server, with "mediarich." in a blue sans-serif font and "SERVER" in a smaller blue sans-serif font below it.

- Adobe Scene 7: <http://www.scene7.com/>
(antes TrueSpectra, eREZ)

The logo for Adobe Scene 7, with "ADOBE" in a black sans-serif font and "scene7" in a bold, italicized black sans-serif font with a red outline.

- FSI-Server: <http://www.fsi-viewer.com/>

The logo for FSI Server, with "FSI" in a white sans-serif font and "SERVER" in a black sans-serif font, both inside a rounded rectangular frame with a red-to-white gradient background.

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

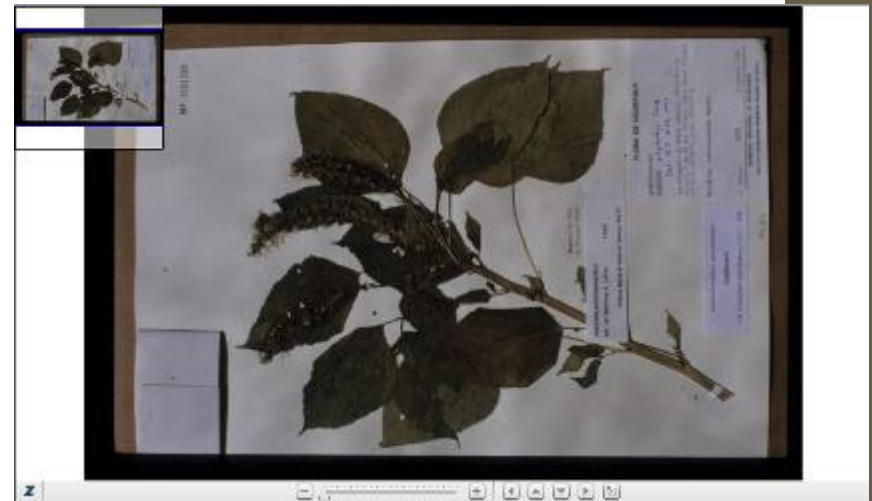
Visualizadores web de imágenes alta resolución

Software para navegación web con imágenes digitales de alta resolución:
Zoom, desplazamiento, rotación, plug-ins (medidas, pase de diapositivas, etc)

Integración a nivel cliente.

Zoomify <http://www.zoomify.com/>

- Sin servidor de imágenes ni TIFF piramidal
- Repositorios de imágenes preprocesadas
- web sites, CDs, or DVDs interactivo
- Multiplataforma: Windows, MacOS, Linux
- Express/Design/Flash/Enterprise
- Zoomify Express: Versión gratuita:
 - Procesador de imágenes + Visor (flash)
- .Para imágenes JPEG (JPEG2000 y TIFF¿implementado?)



PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Visualizadores web de imágenes alta resolución

FSI Viewer <http://www.fsi-viewer.com/>

- Con servidor de imágenes FSI-Server
- TIFF piramidal
- Requiere licencia
- Flash
- Imágenes 2D/3D
- Diversos Add-on y plug-in

IIPMooViewer

<http://iipimage.sourceforge.net/download/>

- Con servidor de imágenes IIPImage
- TIFF piramidales
- Código abierto
- Ajax Javascript



PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Modelos de trabajo propuestos

Estudio y análisis de las necesidades del proyecto. A evaluar entre otros:

Dimensión y arquitectura hardware y software del proyecto

Presupuesto disponible (open source/licencia)

**Estado y operaciones necesarias a realizar sobre las imágenes
imágenes originales (procesamiento)**

**Cualificación técnica de personal (manejo de software de imágenes y
desarrollo de scripts y/o uso e integración de cliente en sitio web)**

“Cloud computing” opción (BHL-Proyecto piloto)

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Modelos de trabajo propuestos

Modelo I.

Base de datos (datos biodiversidad: taxonómicos, nomenclaturales, especímenes/observaciones, especies, etc + (metadatos))

+

Repositorio de imágenes (procesadas y en n formatos y resoluciones)

+

Sitio web (Interacción con bases de datos + Visor para imágenes+ Metadatos e imágenes)

Más simple, más espacio en disco, p.e.; con Zoomify

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Modelos de trabajo propuestos

Modelo II.

Base de datos (datos biodiversidad:taxonómicos, nomenclaturales, especímenes/observaciones, especies etc + (metadatos))

+

Servidor de imágenes (imágenes, una única imagen formato base)

+

Sitio web (Interacción con base de datos + Cliente- servidor imágenes+Metadatos)

Más complejo de montar, un solo fichero por imagen

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Modelos de trabajo propuestos

Modelo III.

**Base de datos (datos biodiversidad:taxonómicos, nomenclaturales,
especímenes/observaciones, especies etc)**

+

Repositorio de imágenes (solo formato base)

+

**Sitio web (Interacción con base de datos + Desarrollo con API's de
procesadores para generación automática de imágenes (n formatos y
resoluciones) y automatizado de procesos+ Metadatos)**

Un solo fichero por imagen, montar “a pelo” servidor de imágenes. obsoleto

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Proyectos y bibliotecas digitales con información sobre biodiversidad en Internet

- Sin servidor de imágenes. Zoomify:

Flora brasiliensis:

<http://florabrasiliensis.cria.org.br/>

- Con servidor de imágenes:

The Biodiversity Heritage Library

<http://www.biodiversitylibrary.org/>

(Djatoka + IIPMooViewer)

Herbario virtual de Santiago de Compostela – SANT

<http://herbario.usc.es/virtual>

(IIP Image ???)



Fuente img:<http://www.biodiversitylibrary.org/>

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Proyectos y bibliotecas digitales con información sobre biodiversidad en Internet

- *Con servidor de imágenes:*

Virtual Herbarium Berolinense <http://ww2.bgbm.org/Herbarium/default.cfm>
(TrueSpectra+FSI Viewer)

Proyecto de digitalización de los dibujos de la Real
Expedición Botánica del Nuevo Reino de Granada
(1783-1816), dirigida por José Celestino Mutis

<http://www.rjb.csic.es/icones/mutis>

(Servicio de Alojamiento y Publicación GBIF.es:
eREZ+FSI Viewer)

Tropicos,
Botanical information system at the
Missouri Botanical Garden:
<http://www.tropicos.org/>
(Djatoka (JPEG2000 image server))



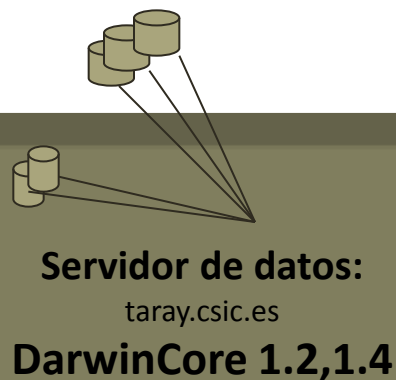
Fuente img: <http://www.tropicos.org/>

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET
Servicio de Alojamiento y publicación de imágenes de GBIF.ES

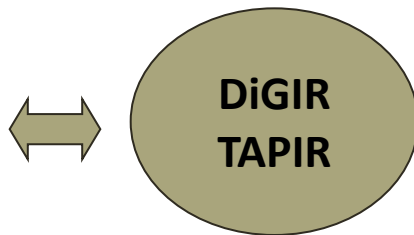
- **La Unidad de Coordinación de GBIF.es ofrece a los proveedores de datos de la red de GBIF la posibilidad de alojar y/o publicar imágenes de alta resolución en Internet.**
- **Condiciones de uso**
 - **Ser proveedor de datos de GBIF**
 - **Sincronizar la actualización de datos e imágenes (imgs siempre asociadas a datos publicados)**
- **Condiciones técnicas:**
 - **Formato base para las imágenes TIFF piramidal**
 - **Datos necesarios (DarwinCore + Datos imágenes + Metadatos EXIFF e IPTC)**

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Integración del Servicio de Alojamiento y publicación de imágenes en la infraestructura de GBIF.es y de GBIF (protocolos, estándares, portales)




Servidor de datos:
taray.csic.es
DarwinCore 1.2,1.4



Visualización de datos y metadatos:

www.gbif.es/datos
data.gbif.org
es.mirror.gbif.org

(url/registro)



Servidor de imágenes:
imágenes.gbif.es
**TIFF piramidal
EXIFF, IPTC**

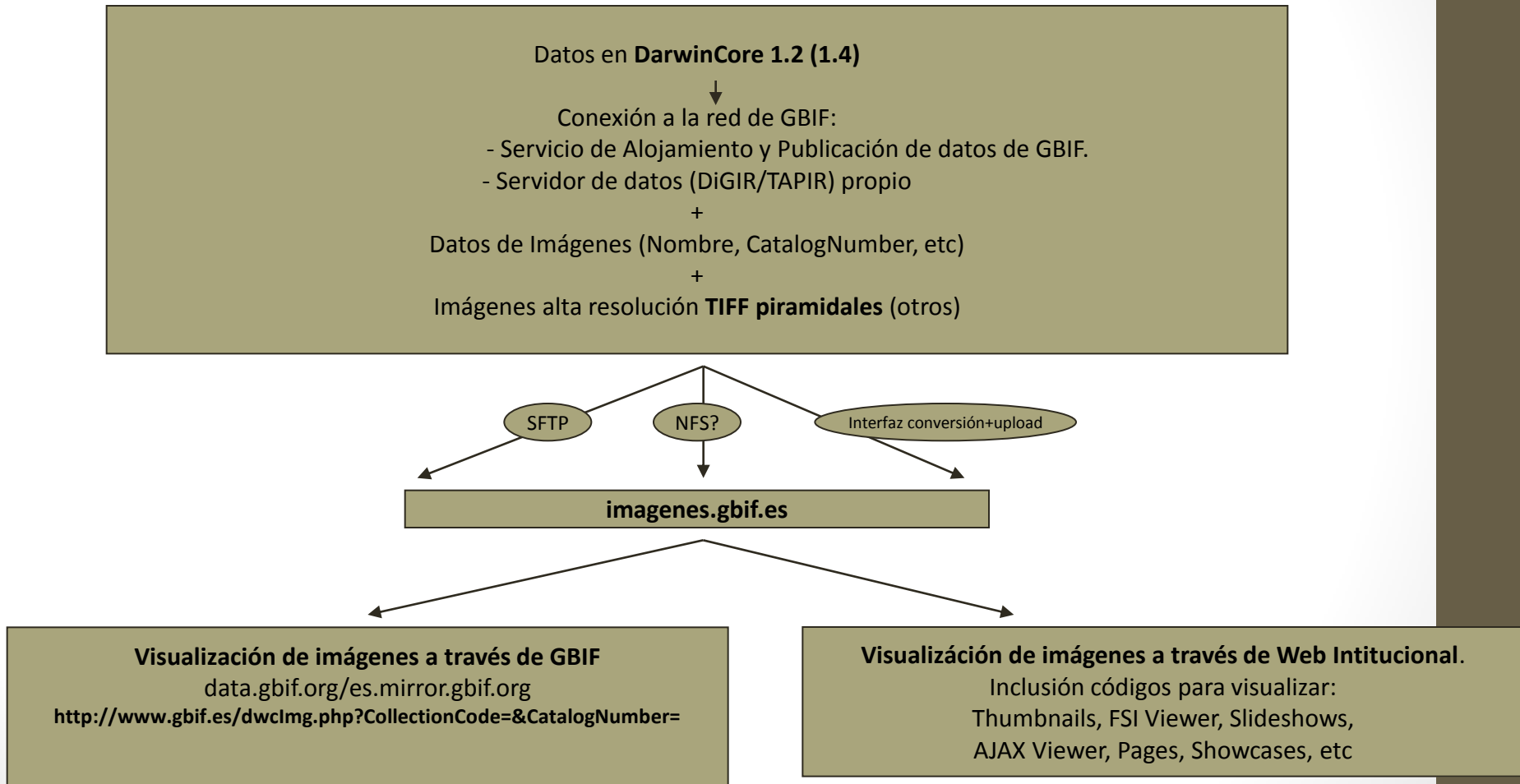


Visualización de imágenes
de alta resolución:

[http://www.gbif.es/dwclmg.php?CollectionCode=&CatalogNumber=\(N imágenes/CatalogNumber\)](http://www.gbif.es/dwclmg.php?CollectionCode=&CatalogNumber=(N imágenes/CatalogNumber))

PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Protocolo de uso del Servicio de alojamiento y publicación de imágenes e integración en la red de GBIF

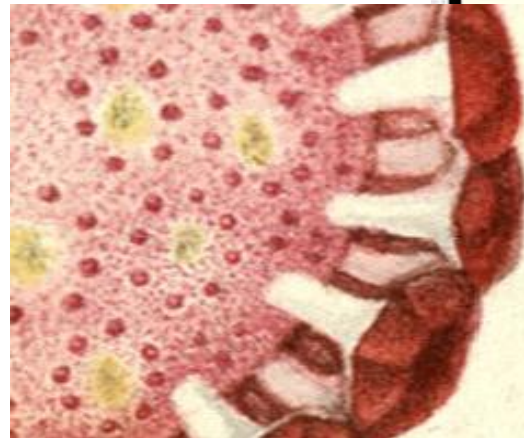


PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Ejemplo de uso

"Proyecto de digitalización de los dibujos de la Real Expedición Botánica del Nuevo Reino de Granada (1783-1816), dirigida por José Celestino Mutis"

- Datos DarwinCore1.2/1.4
Conectados a través del Servicio de Alojamiento y Publicación de datos GBIF.es
- Imágenes. Obtención. Procesado.
- Visualización de datos
<http://data.gbif.org/datasets/resource/9092/>
<http://es.mirror.gbif.org/datasets/resource/9092/>



PUBLICACIÓN DE IMÁGENES DIGITALES DE BIODIVERSIDAD EN INTERNET

Ejemplo de uso

"Proyecto de digitalización de los dibujos de la Real Expedición Botánica del Nuevo Reino de Granada (1783-1816), dirigida por José Celestino Mutis"

Visualización de imágenes de alta resolución:

Imágenes en la web del Proyecto:

<http://www.rjb.csic.es/icones/mutis>

Imágenes en el Servicio de Alojamiento y Publicación de imágenes de GBIF.es:

<http://www.gbif.es/dwclmg.php?CollectionCode=Arch-Div-III&CatalogNumber=A00088->

