

# Extracció semiautomàtica de metadatos: hacia los metadatos implícitos

*L. Díaz, M. Gould, C. Granell, A. Beltrán, A. Llaves*

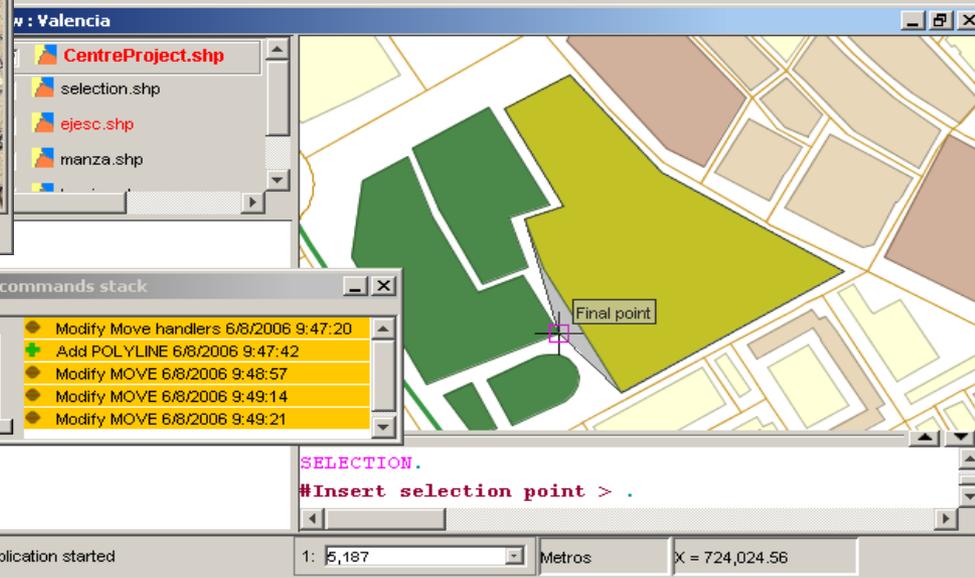
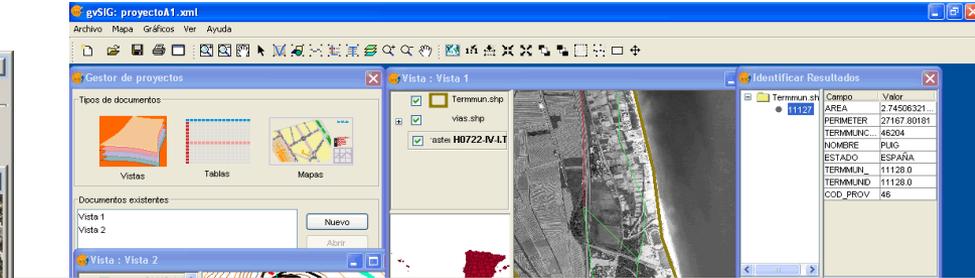
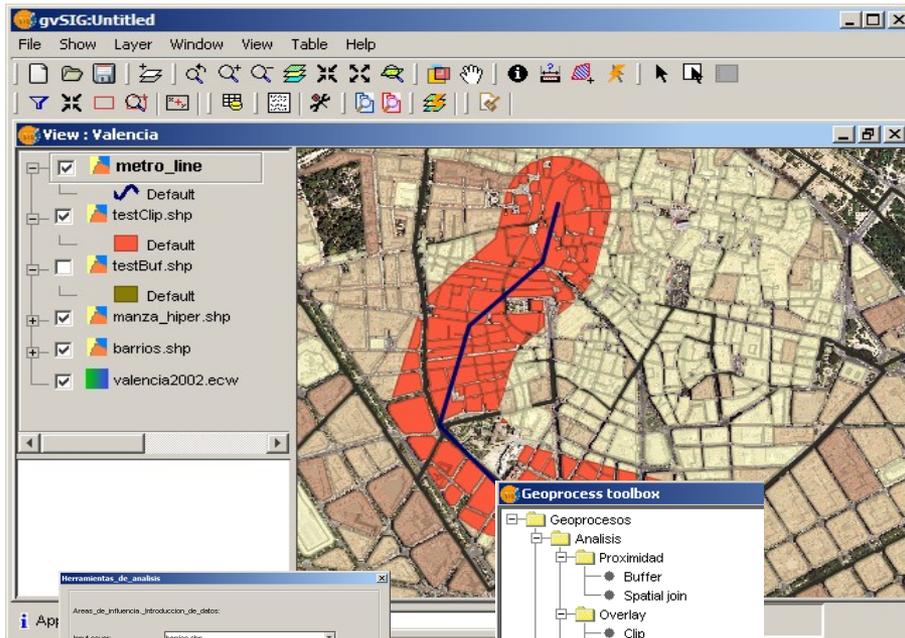
*Centro de Visualización Interactiva (CEVI)*

*Universitat Jaume I, Castellón*

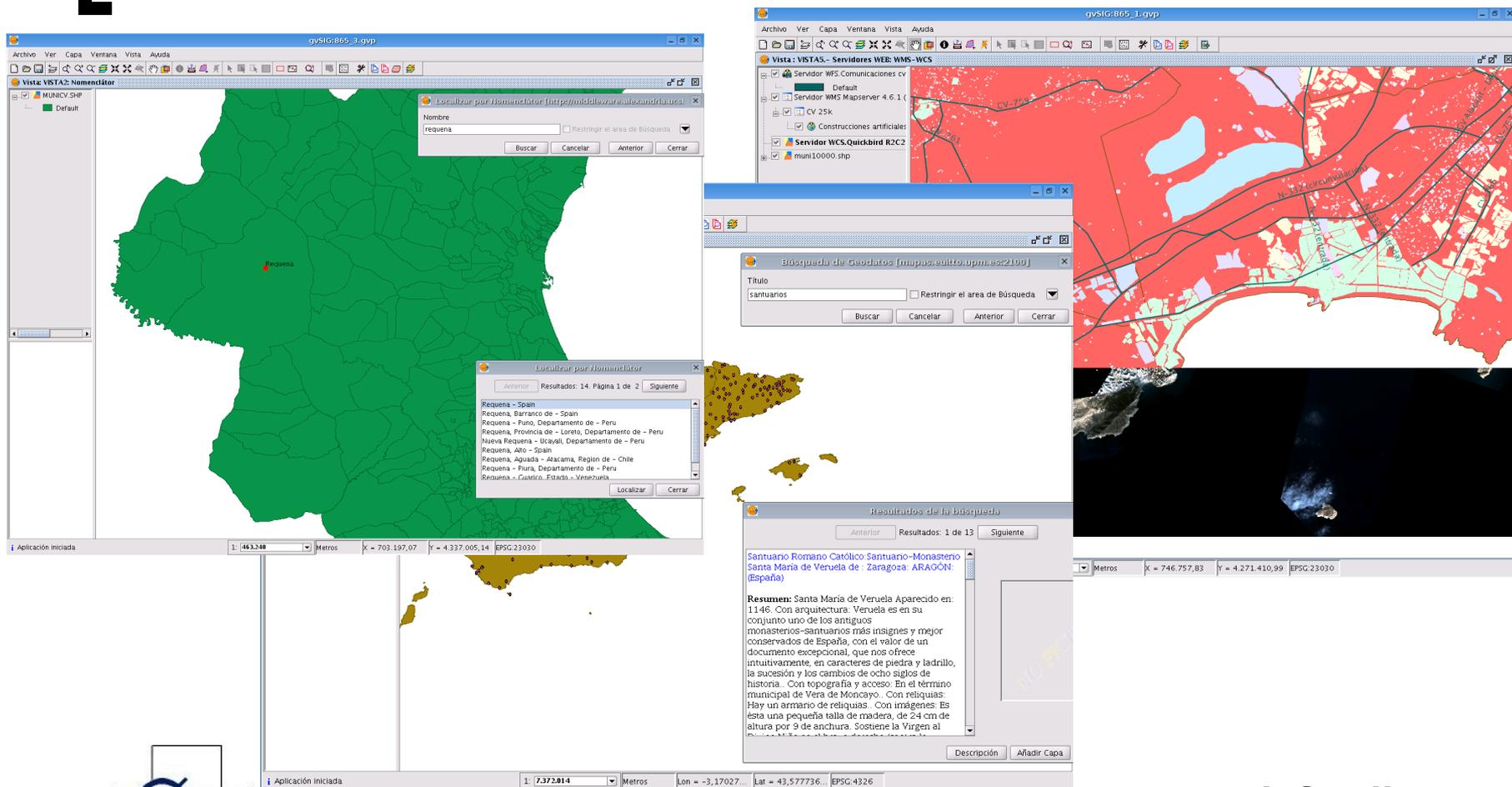
# Contenido

- Contexto del trabajo
- Metadatos explícitos vs implícitos
- Piloto Gestor de Metadatos
  - Extracción de metadatos
  - Publicación en Servicio de Catálogo
- Refactorización de gvSIG
  - Metadatos internos y externos
- Formato de intercambio (dato + MD)
- Conclusiones

# Contexto: proyecto gvSIG

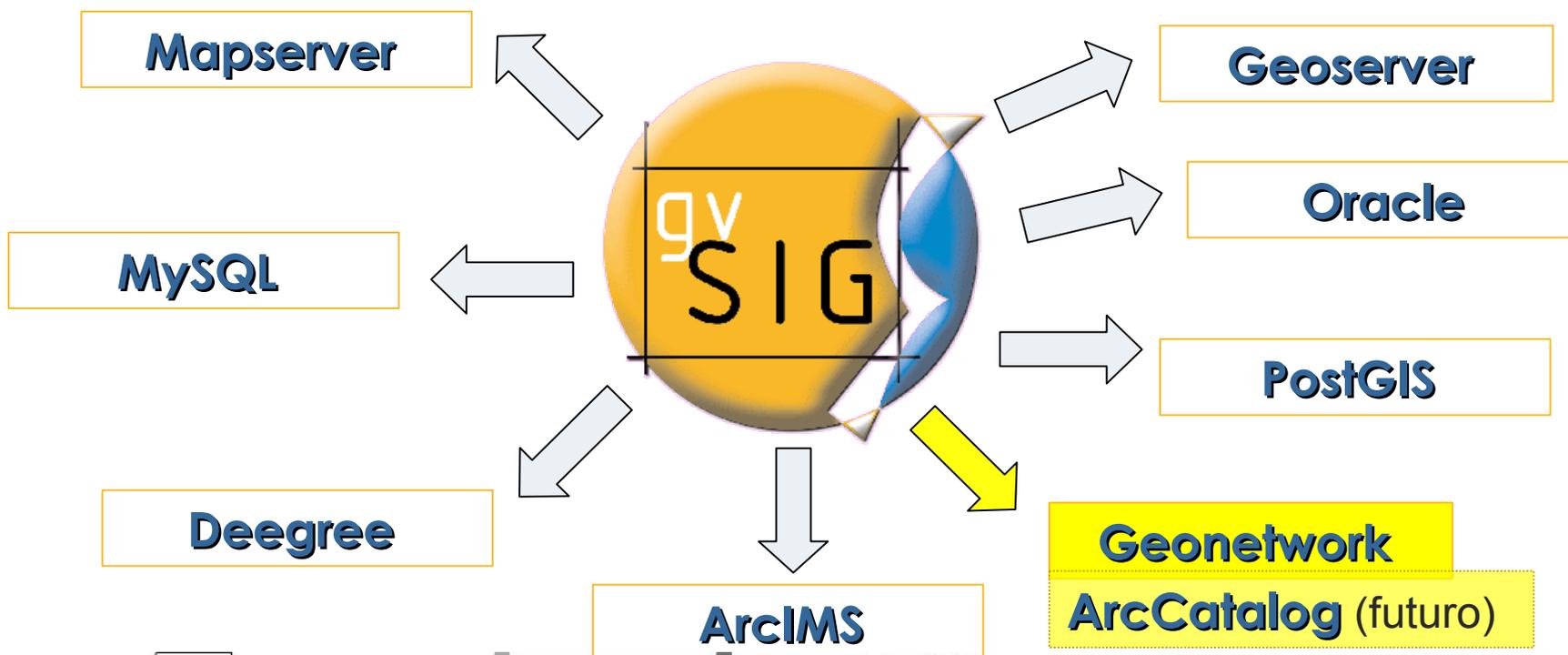


# gvSIG: Cliente GIS y de IDE



# gvSIG: Cliente IDE

## Conectividad a servicios diversos



# gvSIG: Extracción de Metadatos y Publicación (primer piloto 2007-08)

The screenshot displays the 'Editor de metadatos' window in gvSIG. The main interface includes a file tree on the left with 'espana\_munitotal.shp', 'municipios\_poblacion.shp', and 'horta.shp'. The central area is the 'Editor' for 'Ver XML', showing fields for 'mdFileID' (horta.shp), 'mdLang' (languageCode: aa), 'mdChar' (CharsetCd: ucs2), 'mdParentID' (padre), 'mdHrLv' (attribute), 'mdHrLvName', 'mdContact' (rpIndName: mi nom, rpOrgName: organ, rpPosName: first), 'role' (RoleCd: owner), and 'mdDateSt' (fecha). A 'Publicación de Metadatos' dialog box is open, with fields for 'Servidor' (http://localhost:8080/geonetwork), 'Usuario' (admin), 'Contraseña' (admin), 'Grupo' (2), and 'Esquema' (iso19115). A 'Mensaje' dialog box is also present, displaying the message 'Metadato publicado correctamente.' with an 'Aceptar' button. At the bottom of the main window, there is a 'Cambiado' status indicator and buttons for 'Aplicar', 'Aceptar', and 'Cancelar'.

# Hasta ahora, estilo tradicional

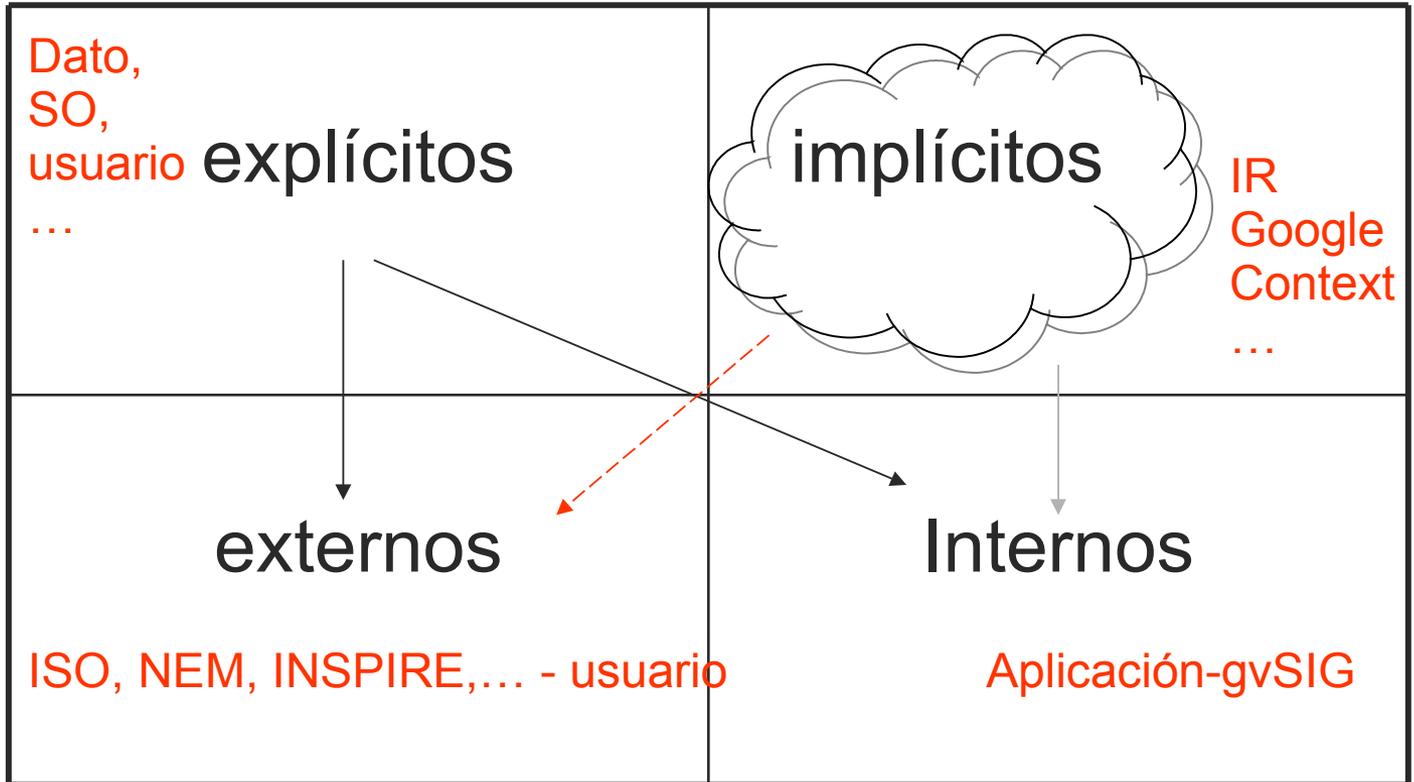
- Extraer algunos MD del recurso (p.e. shapefile), otros del S.O.
- Provisión de editor MD a través de nuevo menú gvSIG; usuario selecciona formato (p.e. ISO, NEM) para generar nuevos elementos MD, y luego exportarlos en XML a un servidor de catálogo; luego busca y los encuentra
- Ciclo completo demostrado (Jornadas gvSIG en noviembre 2007)
- Vinculado a MD (IDE, ie catalogables)

# Metadatos en el futuro

- Extracción (semi)automática
  - Explícitos (leyendo fichero, o con ayuda manual)
  - **Implícitos** (intuición / probabilidades)
- Gestión de metadatos *integrada* en el cliente IDE
  - Editores de texto: artefacto histórico
- Publicación a Catálogos también integrada
- Mayor sincronización (geo-recursos + sus metadatos)
  - + Actualización, + movilidad, todo de forma conjunta
  - Falta formato de intercambio común: estándar de facto

# Explicítos vs. Implícitos

Técnicas de extracción



Uso/Visibilidad

# MD explícitos

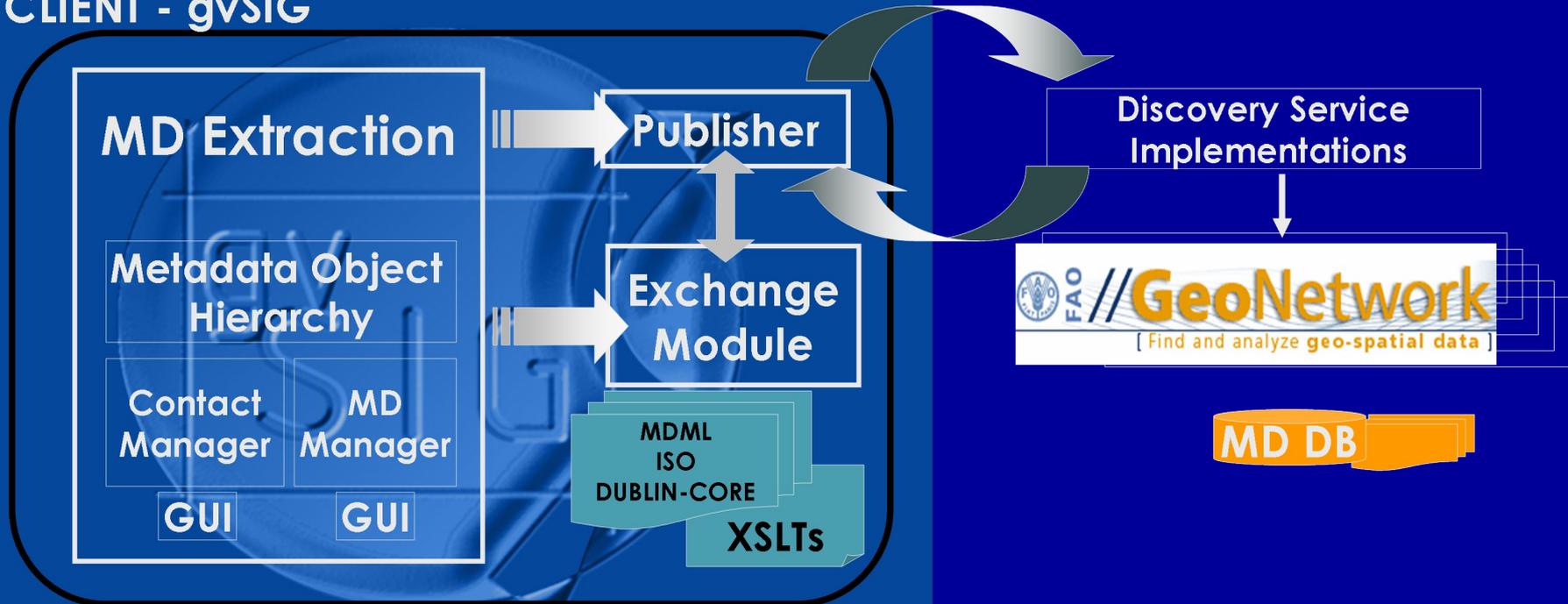
- Introducidos a mano
  - Texto libre del *abstract*
  - Selección de 4 coordenadas BBOX, empleando interfaz con mapa
- O sacados fácilmente del S.O o fuente de datos.
  - Nombre de archivo, fecha de última actualización
- En fin, metadatos normales (hoy en día)

# MD implícitos

- Minería de datos, métodos avanzados de la recuperación de información (IR)
- Ejemplo: incluir en el *abstract* (automaticamente) los lugares de cobertura por el BBOX → nomenclator → topónimos
- Otro ej: inferir época=invierno, por los valores espectrales de una imagen de tal zona
- Métodos 'estilo Google': geo-crawler + indexación
- Métodos no perfectos, pero a lo mejor 80% eficacia con 20% del esfuerzo
- Menos tiempo planteando el formato perfecto, y más tiempo creando+explotando metadatos!!!

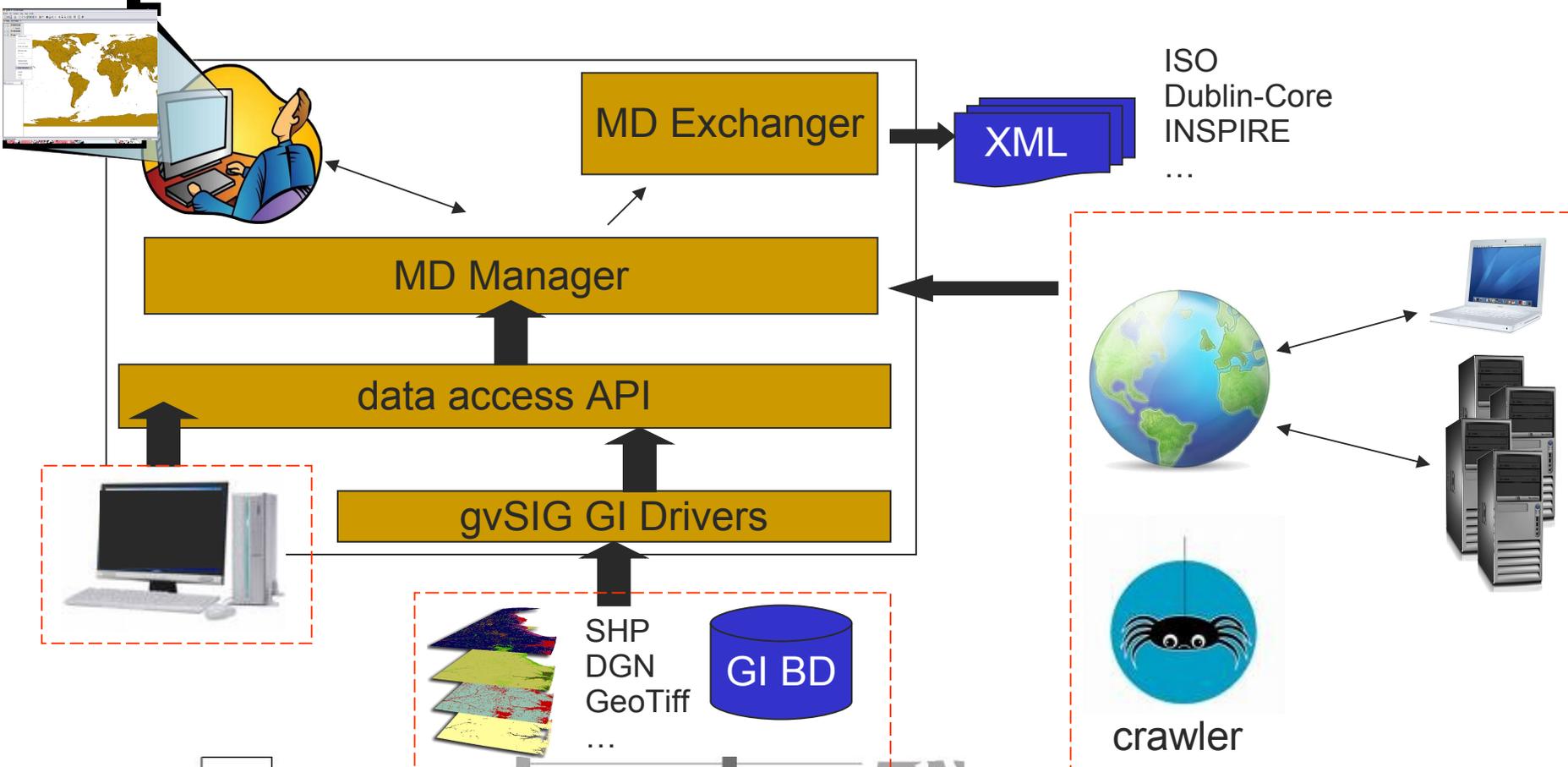
# Arquitectura: Gestor de Metadatos, versión anterior

SDI CLIENT - gvSIG



OPEN SOURCE FRAMEWORK

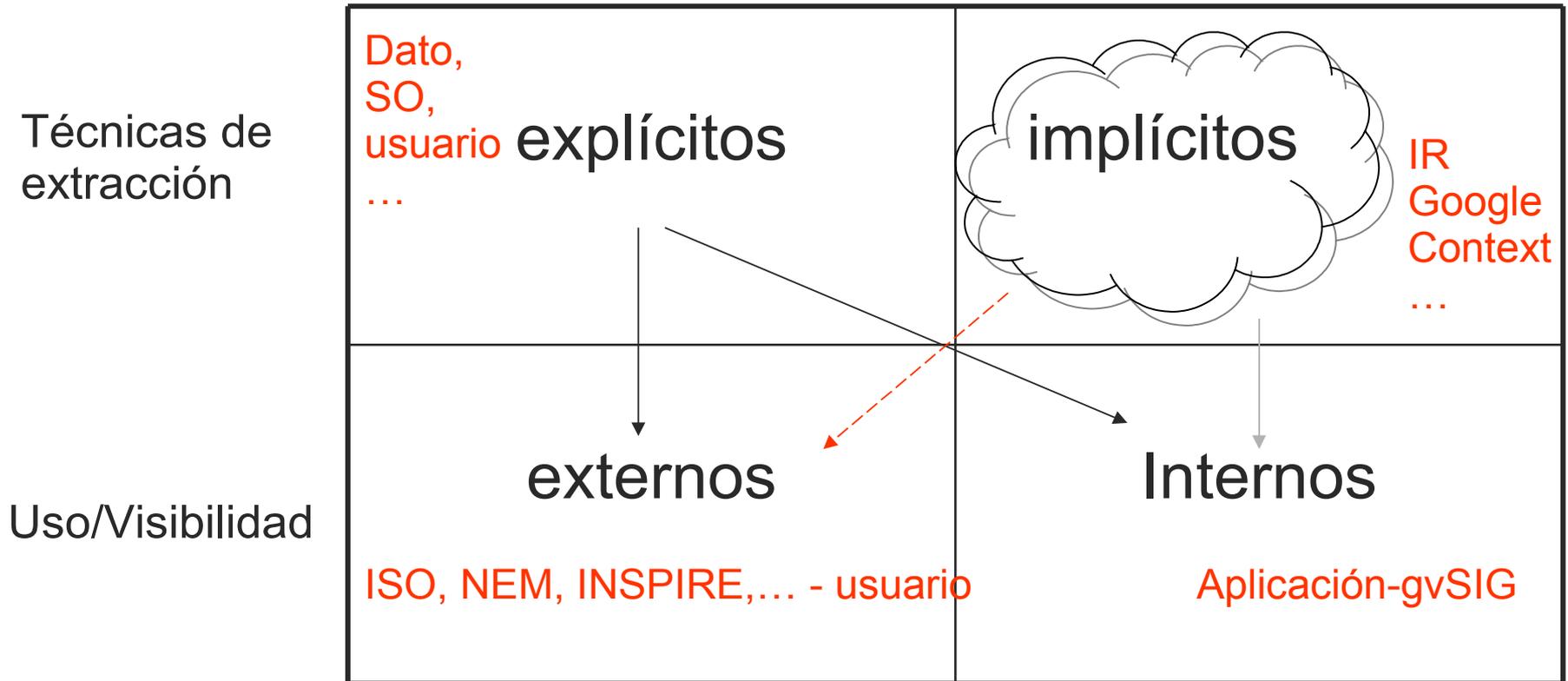
# Extracción MD



# Refactorización gvSIG

- Cambios arquitectónicos 2008
  - integrar gestión de metadatos (internos) en el núcleo de gvSIG (i.e. no como extensión)
  - Redefinimos el concepto de editor central de metadatos: se convierte a asistente para el proceso de publicación de MD
    - Mostrar campos obligatorios que faltan cuando se quiere compartir o publicar MD
    - Fijar parámetros de publicación
  - Además proporcionar y fomentar el uso de metadatos internos + externos

# Externos vs. Internos



# Metadatos internos

- Descripciones que ayudan a los usuarios a explotar la aplicación de una forma más eficiente
- Consultar fácilmente a un proyecto, una capa, un feature
  - ¿Cuándo fue creado?
  - ¿Por quien? Etc.
- Almacenamiento de metadatos internos en una base de datos interna (H2)

# MD internos siempre accesibles

gvSIG:Untitled

File Show Window Help

Vista: big project

cv\_300\_todo

Layer properties

General Symbols

Name: cv\_300\_todo\_def.dgn

Use spatial index

Scale range

To show always

Don't show the layer when the scale

To be under: (Minimum scale)

To be over: (Maximum scale)

Properties

ig\CV300 (dgn) \cv\_300\_todo\_def.dgn

Hyperlink

Field: ID Extension

Predefined action: Link to image files

Accept Apply Save legend... Load legend... Close

Application started | Meters | X = 712,158.37 | Y = 4,369,266.87 | EPSG:23030

CEVI VISUALIZACIÓN INTERACTIVA

## MD Externos: Cuando se necesite una IDE

- No todos los usuarios tienen acceso a una IDE: motivos técnicos, políticos, etc.
- Cuando es necesario y/o se decide publicar para compartir (datos+MD) añadimos la extensión de Publicación
- Combinado (algunos) metadatos internos con externos necesarios para cumplir un estándar elegido
  - ISO, INSPIRE, Dublin Core, etc.
- XML de MD se exporta a un catálogo.

# Cómo usamos MD internos junto a los externos

- Metadatos internos (o propiedades internas) son útiles dentro de la aplicación
  - No tanto fuera de ella
- Es necesario “empaquetar” tanto los metadatos internos como los externos para incrementar la interoperabilidad
- Exportándolos de forma conjunta a un estándar

# Formato de intercambio

- Empaquetar recurso GI + metadatos GI + derechos acceso/uso + otros MD
- como el ejemplo shapefile (.shp, .shx, .dbf) ... Extendamos esta idea
- Interesante la extensión del formato **MEF** (Metadata Exchange Format) de **GeoNetwork**
- gvSIG: en el futuro próximo leer y usar MEF
- Interés en producir una extensión; MEF ver 2.0

# Conclusiones

- Se necesitan soluciones para facilitar la creación de metadatos, su uso y su mantenimiento
  - Necesidad de nuevas técnicas de extracción para minimizar el esfuerzo (MD implícitos)
- Importante la unificación del recurso (dato) + metadato
- Importante una sincronización Global ( id único)
- La cooperación internacional es necesaria para asegurar la interoperabilidad

# ¿Nos vemos en Diciembre?

- **4as Jornadas gvSIG** (>400 asistentes)  
inscripción gratuita
- Junto al **OGC Technical Committee meeting** (150-180 más ....)
- Todo en Valencia. Palacio de Congresos
- 01-05 Diciembre 2008.



[ Gracias por su atención ]

**Nuestro grupo**  
<http://www.geoinfo.uji.es/>  
**Proyecto gvSIG**  
<http://www.gvsig.gva.es>

