



# Near-term metadata challenges

**Michael Gould**, *Universitat Jaume I, ES*

**Jorge Rocha**, *Univ. Minho, PT*

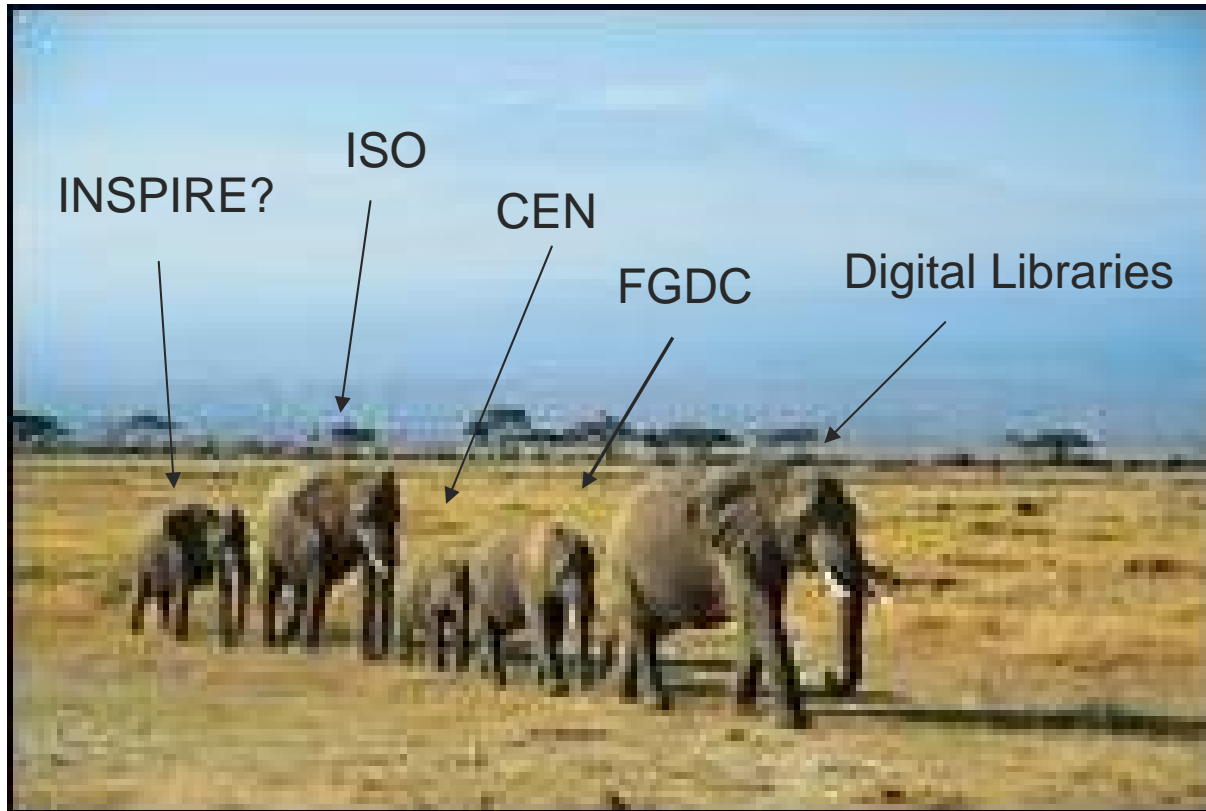
**Stefano Nativi**, *CNR-IMAA, IT*

**Javier Nogueras**, *Univ. Zaragoza, ES*

**Miguel Manso**, *Univ. Politécnica Madrid, ES*



# [ Geo-metadata: Status quo ]



# Research point of view

- Status quo (SDI 1.0) is not perfect, but becoming stable
  - Profiles of 19115 + DC
- ESDI needs providers to create+publish standardized metadata; no need to wait
- In parallel, design and test new md methodologies (SDI 2.0)



# Bulterman article

- IEEE Multimedia, Oct-Dec 2004.
- “Is it time for a moratorium on metadata?”
- Critical of past decade of metadata committee work
  - More standards, however less metadata being created and used
  - Google effect: more hits, without need for user-generated metadata (as semantic web asks)
- Conclusion: need to rethink md for application areas



# [ MD challenges (near term) ]

1. Cartographic --> informatics viewpoint
2. Automated extraction/production
3. Separation discovery + description
4. Linking MD to services
5. Multilinguality
6. Imagery + Earth science MD
7. Test the geodata commons



# [ Cartographic → Informatic view ]

- Digital resources, services, experiences (rl)
- From map (finished product) to feature collections (*ad hoc* compositions, presented cartographically or not)
- WFS is feature search filter
  - Presentation is separate
- These database objects require another style (non-digital-library) of metadata



# Automatic extraction

- Geodata should always be documented (content+md)
- Your GIS should help you produce md, like your camera does!
  - Extent, dates, data type/format, creator, many other elements
  - Human user supplies free text



# Automatic extraction (2)

- Work of Manso (ec-gis&gis 2005)
  - 50 GI data formats analysed
  - 9 of 22 ISO core elements extracted
- Further experiments to be conducted within development of gvSIG (see paper at 14:30)



# Separation discovery-description

- Of the metadata classes identified (description, discovery, use, evaluation, maintenance, system,...) discovery stands alone.
- SDI interoperability depends on discovery
- Bulterman: md is mostly about discovery; other descriptive elements should be encapsulated in the data

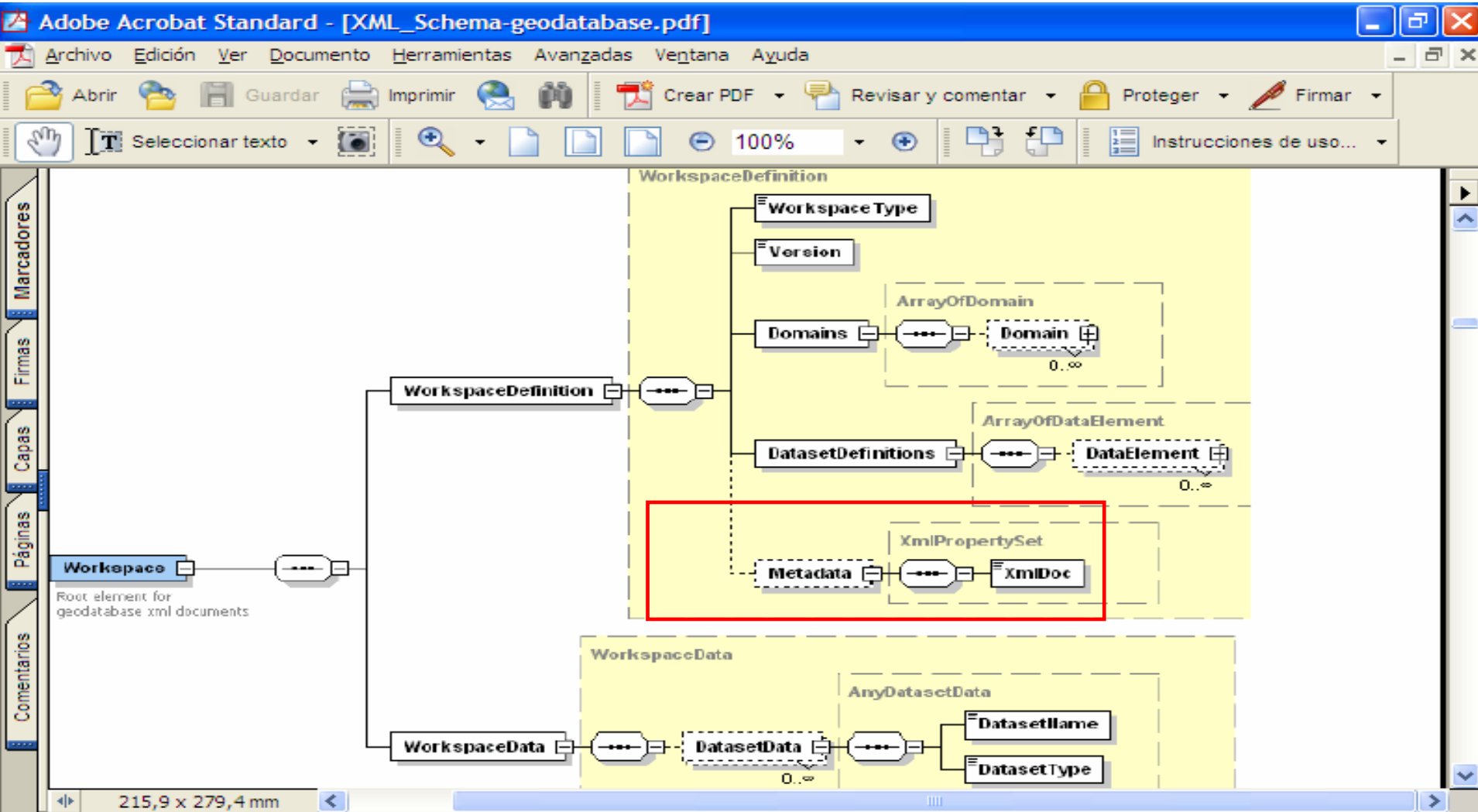


# Shapefile: package of data

- .shp = geometry
- .shx = index on data
- .dbf = attributes
  
- ...but they forgot to include metadata (1997)



# Geodatabase: package of data



# Linking MD and services

- Gap present between 19115 (weak on services) and 19119 (weak on data served)
- Data resources (in catalogues) should link to servers able to consume them
  - And services, point to consumable data



# Multilinguality

- Describe data/services in users' language
  - Advertising, yellow pages
- Exploit thesauri
  - GEMET, UNESCO, EUROVO, AGROVOC
- On-line translation
  - Experience from SDIGER pilot



# Imagery+Earth science md

- Imagery not well served by 19115
  - 19115-2 helps, but is not enough
- Compare these elements to typical image header metadata
- GMES-Inspire need to be harmonised
- ES fields work on other dimensions, time scales, spatial ref systems,...
- How is time modeled, are ES semantics well captured, how is capture process documented?



# Testing Geodata commons

- Published URISA Journal: Onsrud and others
- Semi-automated extraction and publishing of metadata
- Data freely published to commons
  - Internal code, which evolves each time dataset moves and is updated



# [ SDI ver. 2.0 ]

- Extract metadata from geodata source
  - At data collection (like GeoTIFF,...)
  - During processing (in GIS)
  - And automation of md publication (commons?)
- Encapsulate geodata+metadata
  - Exposing only Discovery metadata in registry
  - Common data formats (GML, Geodatabase) support this
- Discover geodata resource, and ask it to describe itself



# Conclusion

- Current md situation becoming stable: use it!
- However, no need to be complacent
- Plenty of near-term R&D work to be done in md field
- FP7, other programmes



# [ Thank you for your attention ]

- Any/all authors may be contacted via: [gould@uji.es](mailto:gould@uji.es)

