Spatial information Research: An Aarhus In-Situ Thematic and Operational Immersion

Dr Saviour Formosa

Immersion Research Activities

25-31 January 2009

Joint Research Centre, Ispra
Milan

MT/06/IB/EN/01 Further Capacity Building in the Environment Sector

EXPERT REPORT

BC expert/trainee
Dr. Saviour FORMOSA
Malta Environment and Planning Authority (MEPA)

Traineeship duration
25.-31.01.2009

Host country
Italy

MS hosting institution
Joint Research Centre, Ispra

Objectives/contents of the study visit

1. Structure of spatial data
2. Data cycle process employed for environmental and land use data
3. Prediction modelling using spatial data
4. Dissemination processing such as web portals and on-line analytical/querying systems
5. SEIS implementation at JRC
6. Dynamics between environmental, social and land-use spatial data
7. Overview of JRC functioning

Main objective is get achieve overarching knowledge of the way the JRC SDIU works and to review how MEPA can organise its spatial data and units in terms of consolidation of DIKA process as well as operational organisation of flows and staff. The workings of JRC and well as knowledge on the types of work carried out will help MEPA get a clear idea of what JRC is or is not and how it relates to spatial data. In turn, MEPA will be able to use such knowledge in its quest to enhance its data process and be able to launch the SF_Mon project outputs within a reasonable operational structure.

Detailed work programme

Day: 26th January 2009
Time: 09:00
Meeting with Dr. Stephen Peedell
Objective Number: 7
Objectives of the meeting: To acquire an overview of the JRC and the IES Institute and of SDIU. In addition to draft a detailed work programme for the study visit.

i)   Topics discussed:
      a. Overview of the Community Spatial Information Network (COSIN) Action (11601)
      b. Agenda setting of the traineeship exercise.

ii)   The COSIN Action aims to port the National-level INSPIRE to the Community Spatial Information Network
      a. Documentation was provided with main comments on review were based on the following:
           i. Use of keyword approach to be taken up by MEPA: to establish a set of keywords that would be used across all Actions and Malta activities: the same issue is that which needs to be streamlined across the JRC Actions to enable harmonisation of the keywords which are currently termed differently
          ii. The issue of the Action which proposes that "the Commission will, if appropriate, come forward with a legislative proposal regulating the establishment of the components of INSPIRE within the relevant Community Institutions and bodies": an interesting proposal considering that INSPIRE regulates States but not EC bodies
          iii. The different objectives were reviewed with main comments from a non-JRC insider focusing on the grading used at JRC in its level of involvement (1 to 3? – highest signifance?), OUTPUT Category 1 to 5?)
      b. One interesting item highlighted was the use of SAP processing for MT
iii)   Reviewed the Unit Structure:
      a. 1. Structure of spatial data (Steve Peedell, SDIU)
      b. 2. Data cycle process employed for environmental and land use data (Steve Peedell, SDIU + other relevant Units)
      c. 3. Prediction modelling using spatial data and how they bring it together for dissemination purposes such as web portals and on-line analytical/querying systems (Steve Peedell, SDIU + other relevant Units for prediction modelling (EFFIS/EFFAS/Soil Erosion, otehrs)
      d. 4. SEIS implementation at JRC (Max Craglia, Michel Millot, Stephen Peedell) (SEIS Pilots) + LMNH for Data Centres)
      e. 5. Dynamics between environmental, social and land-use spatial data

Day: 26th January 2009
Time: 12:00

Meeting with Dr. Ing. Harald Scholtz, Dr. Paul Smits
Objective Number: 7
Objectives of the meeting: To acquire an overview of JRC as an entity and the different activities carried out by the IES.

iv)   Topics discussed:
a.   Official overview of JRC and its structures

v)   Presentation given by Dr. Scholtz covered the JRC administrative structure and how each Unit fits within each Institute, in this case the Institute for Environment and Sustainability.

vi)   Reviewed the list of topics that SF indicated were of interest and the relative experts were identified
a.   SESIS – Harmon
b.   Remote sensing
c.   Risk mapping
d.   Infrastructure and networks
e.   Geoportals
f.   Metadata technologies
g.   INSPIRE and IRs
h.   Cognition

Day: 26th January 2009
Time: 11:00
Meeting with Dr. Stephen Peedell and Dr. Paul Smits
Objective Number: 2
Objectives of the meeting: To acquire knowledge of datacycles, data exchange and transfers.

8.   Data cycle process employed for environmental and land use data

vii)   Topics discussed:
a.   Differentiation between environmental (thematic) and landuse-data
i.   Who gathers what?
ii.   Data transfer and ownership
iii.   Inter-departmental exchange system (costings and buying of material/services) between units
iv.   Models that exist which can help process data across the disciplines:
1.   lineage
2.   iso procedures

b.   MoUs
i.   Intra and inter-organisational MoUs
ii.   Inter-regional
iii.   Inter-national
iv.   GEO-level

Day: 26th January 2009
Time: 13:30
Meeting with Dr. Bertrand de Longueville
Objective Number: 4
Objectives of the meeting: To acquire an overview of webservices and the systems being used by JRC to disseminate spatial information.

    viii)   Topics discussed:
        a. Web-services
            i. Client-side Ajax base
                1. Use of opensource:
                    a. MapFish
                        i. Integrates OpenLayers inside another framework that allows user to create an elaborated interface
                        ii. MapFish builds on OpenLayers
                        iii. ArcGIS to OpenLayers to MapFish
                    b. OpenLayers
                    c. Javascript libraries
            ii. server-side
                1. webmap services
                    a. wcs (data)
                    b. wfs (raster)
        b. Geo-processing services: use fat-server:thin-client approach
            i. Geoservers include mapserver (allowing tile-caching or ArcGIS server using proprietary software) – comes with an administrative server tool
            ii. To check if geoserver accesses SQL
            iii. Wms server send the jpg and the client generates the map for the user
                1. However to precompute the required maps and caches them the best method is to create all the maps in a caching system as is employed by Google and users browse down between the layers and shows the user the layer closest to their request
                2. used in ArcGIS Desktop (allows users to chache)
                3. publish as a web-service
                4. reads the javascript as a .wmc file as in a txt file
        c. issues of replication at ISP
            i. scale issue
                1. as different scales are required at the different zoning issues, have to create a series of maps of different detail
                2. also need to upload the mxd file
        d. website naming: geonames.org
            i. list of placenames for websites
            ii. challenge is to try to emulate the RSS from Reuters onto a geomap as is done by JRC

Day: 27th January 2009
Time: 10:00
Meeting with Dr. Gianluca Luraschi
Objective Number: 4
Objectives of the meeting: To achieve knowledge of geoportal technologies and the Processes behind the Metadata Editor.

ix)   Topics discussed:
   a. Discussed the issue of a prototype of a discovery view client for metadata both in JRC and in Malta
   b. One option would be to use Mapinfo or Arcinfo to export the metadata and then make it INSPIRE compliant
   c. There is potential to have one metadata client across for all Europe, in addition creating a 2-way system, where the national clients feed the European client
   d. Two options also exist in transmission where the metadata xml is uploaded to one repository (url) or else to create a csw system
   e. MT can take the option to upload all its metadata to the JRC folder as in the case of the EEA CDR function
   f. We can put in an iframe of the JRC client in the MEPA website
   g. Regarding the Metadata Editor (MEditor) review the MEditor website and review the catalogue
      i.   Can also search by map area in the geoportal
      ii.   Created a geochannel that allows a query tool – georss field and then use tools such as Outlook or Mozilla to view the georss
      iii. Will be available externally from JRC in June09

Day: 27th January 2009
Time: 15:00
Meeting with Dr. Max Craglia, Dr. Robin S. Smith, Dr. Michele Campana
Objective Number: 5
Objectives of the meeting: To acquire an overview of SEIS and HARMON as well as to discuss the new metadata outputs. The integration of SF_Mon in line with SEIS outputs to be covered by the discussion.

x)   Topics discussed:
   a. discussed the issues of socio-economic impacts and planning
   b. discussed the SEIS implementation and overview through the new JRC project HARMON that is being initiated at the present time
xi)   RELIT: Michele Campana gave an overview of the current projects (ex RELIT – 2005-2008) underway in Lombardia on planning and environment particularly those assessing of the impacts of SDI on spatial e-gov services. These concentrated on regional SDIs and the building of of a regional service that picks up those small areas

(comuni) that are too small to manage such systems - similar to MT local councils and SDI at district level should MT need to decentralise, though not advisable as MT NUTS 3 is still small enough to create regional SDIs

a. Cost and benefits impact assessments are carried out as based on their first project in Catalunya SDI

b. The RELIT project created a topographic database from the large to smaller scales as comuni had 1:1000 and 1:2500 and the regione needed 1:10000. A comprehensive run was undertaken (50:50 expenditure base) and new shp-based database was created that also resulted in a geoportal (services: catalogue, discovery, view, download)

c. The socio-economic impact reviewed the following parameters
  i. Cost – technological / production / data
    1. Internal and external costs benefits
  ii. Efficiency
  iii. Efficacy
  iv. Democracy (access to different catalogues, citizens, business
    1. communi and wider to the region

d. spatial planning activities and documentation were investigated

e. Methodology tools employed varied on the projects run in the Spanish and Italian regions. One can use such methodology form MT Structure Plan Review and Local Plans. The drive is to see whether the process derives benefits for finding data and information more easily based on SDI implementation
  i. Catalunya CBD
  ii. eGEP – egov: information and infrastructure evaluation
  iii. eSDNET

f. the cost-benefit analysis look only at the catalogueing and publishing sides of the project and not the actual data creation/gathering itself as this is prohibitive (since data will anyway be created as part of requirements and may be sourced elsewhere)

g. SEIS and INSPIRE metadata converter for the xls version
  i. Discussed the possibility of existence of tools that would QA/QC the xls input to verify the content: the metadata editor is INSPIRE-compliant but there is no way to verify INSPIRE-compliance in the xls version. Max Craglia stated that this has not been done yet. In the previous meeting with Gianluca Luraschi I suggested that one could convert the xls to xml and then the ME (metadata editor) would bring it in and the use can then review the pages and confirm whether the metadata is compliant. To investigate such a possibility. The other issues lie with the other tools that produce metadata that is not INSPIRE-compliant such as ESRI, MapInfo, etc – one needs to check whether there are third party tools or updates that will do this – JRC/MS are still in discussion at this stage.

xii) Harmon: Robin Smith gave an overview of thye new project funded by DG Env with JRC as the managing agency

a. Project is monitoring SEIS and other projects such as NESIS for harmonisation of works and is reviewing and running:
  i. Baseline assessment procedures
  ii. Comparative analysis

       iii.   Propose alternative policy options (inclusive of cost-benefits)

       iv.   It looks solely at monitoring strategy and infrastructure

b.   The project leads to SEIS and will guide the EU to new policy or maybe even a do-nothing approach

c.   Projects it is monitoring and reviewing the following projects. The result will also integrate all three projects' outcomes.

       i.   Environment monitoring Facilities (coming to a close) – partners UBA, France, UK & Slovakia – ask Johannes Mayer for info

       ii.   Café

       iii.   NESIS – ICT side of INSPIRE

d.   Costs are being reviewed – case studies are currently covering air and water quality, inclusive of atmospheric air quality data flows

e.   UK case study was run based on the ERFF –Environmental Research Funding Forum organisation that developed a database of all the UK research programmes/projects

f.   Check the iGoogle mode of information transmission

g.   SF tasked to review the current questionnaire that is being proposed for dissemination to the MS and which will have to be filled in by each country for each project. Main comments

       i.   It is currently too large to maintain

       ii.   May be too complex for countries (technical-wise)

       iii.   However it is easy flowing and understandable in terms of flow and explanations on aims by section

       iv.   Main and detailed comments are found in the file entitled HARMON_Content_AQ-UK_SF.doc

       v.   It takes an average of 5 hours to review and comment and may take similar time-periods to fill in and carry out the necessary find-outs to complete

h.   SEIS implementation at JRC

     i.   Current progress

     ii.   What will be done through the EEA visit and post-EEA visit

     iii.   Any new progress

     iv.   INSPIRE and SEIS

       1.   Malta's system based on the xml outputs

       2.   How can this be enhanced?

       3.   What is JRC planning post-INSPIRE implementation by the states

Day: 28th January 2009
Time: 10:00
Meeting with Mr. Martin Tuchyna
Objective Number: 5

Objectives of the meeting: To review the Annex I Data Specifications testing Phase.

xiii)   Topics discussed:
   a.   discussed the issues of INSPIRE data Specification progress
xiv)   The outputs of the data specifications should be presented in Brussels in June as based on Annex I data specifications
   a.   They are based on OO
   b.   Based on 9 themes covering both the reference and thematic aspects, thus it has been difficult to draft such specifications
   c.   The current process aims to build up a conceptual framework for a common framework that brings in all the 9 themes together
   d.   The process is based on 8 steps and JRC s currently at the 7th step: the testing phase
xv)   Members states do not need to change the data they hold to OO but need to ensure that transfer mechanisms are in place
   a.   As an example an MS may have a cadastre or base map
   b.   The EU will ask for the first reporting in 2 years
   c.   The MS, in its report needs to state that the data is available to the EU in the shape specified in the data specifications
   d.   MS has to ensure that tools exist to transform the object into the specifications required
xvi)   the deadline for the testors on data specifications is the 20th February (for feedback)
   a.   testors are reviewing the Object-Oriented structures to help them improve daily data management
   b.   MS may need UML experts in order to understand and manage the result and implement the system
   c.   There is an issue with the language itself due to network services issues on data format
      i.   This will be in GML, currently in v3.1.1 but JRC is insisting on v3.2.1
xvii) to identify the MEPA contact point for data specifications and see if we sent out feedback
xviii) the result of the findings will be presented in summer 2009
xix)   the final transmission may be finalised as a Decision not as a Regulation (which is different from the Metadata outcome)
xx)   it is still not clear who will 'own' the GML files, whether JRC or Eurostat

Day: 28th January 2009
Time: 11:00
Meeting with Dr. Andrea Baraldi
Objective Number: 2
Objectives of the meeting: To review the e-cognition technologies available for use

xxi)   Topics discussed:
    a. Methodologies
    b. Current and future tools

xxii)   The debate centred on the need for new technologies that automate the image interpretation process, through space-time

xxiii) The main issue centres on the fact that all algorithms require ground-truthing

xxiv)   The system requirements for a system to be operations should be based on the following:
    a. Ease of use
    b. Efficiency
    c. Effectiveness
    d. Scalability and robustness

xxv)   Imagery is no longer a real-problem as availability for all is increasing rapidly through free satimages
    a. Seabers – available since 2007 (low quality)
    b. Landsat – available since October 2008 (high quality)
    c. Sentinel – planned for 2012 (GMES – GEO)

xxvi)   The method developed at JRC is based on a heuristic approach that is dependent on segmentation iterations
    a. The process creates primitives with no semantics
    b. The process starts with colour as subdivided into strata where each stratum is a category set
        i.   Eg: separates trees from grass (add then a height factor)
        ii. The process moves from signal to semantics and the results are very similar
        iii. The system is symbolic throughout
        iv. System should be fully functional within the new 4th generation intelligent EO systems being developed
    c. A tool will be available in July called Envi 4.3
        i.   Works on well-bahaved data
        ii. Decision rules are can be inserted at the beginning of the process
        iii. Based on a hierarchical approach
            1. first-level – generic
            2. second-level – detailed – example oil spills and Corine
            3. moves from colour to shape to spatial relationships (each is not fundamental and finally the sum is unique
        iv. uses a convergence of evidence approach (or constructive reasoning)
            1. eg: both trees and cars are green
            2. only a tree can be taller than 5 meters
    d. The tool can be used to set up a collaboration to test such issues as oil spills, fire damage, hi-res urban analysis, structure plan-level landuse analysis and detaile local-plan analysis, etc
        i.   Requires the establishment of metadata
        ii. Need data on ground truthing to set up a comparative approach in order to execute a validation of the system
        iii. Can be run under an IDEA project

    e. Policy issues:
        i.   Who vets and forwards data?
        ii. Who is involved in verifying data
        iii. QA/QC issues?
        iv. Standards employed
        v.   INSPIRE implications

Day: 28th January 2009
Time: 15:00
Meeting with Dr. Andreas Skouloudis
Objective Number: 6
Objectives of the meeting: To discuss progress and difficulties encountered with the Ferrara monitors.

    b. JRC-Ferrara Air monitoring sensors
        i.   Reviewed the main data emanating from the sensors
        ii. Analysed the latest data gathering processes
        iii. Reviewd the issue of installation of second monitor since the first one was corroded
        iv. Data is ssen as reliable but needs to be cross analysed with the current station output

Day: 29th January 2009
Time: 11:00
Meeting with Dr. Jean Dusart
Objective Number: 1
Objectives of the meeting: To acquire an overview of the infrastructure and data structures employed at JRC.

9. Structure of spatial data

    xxvii) Topics discussed:
        a. Note was made that ArcGIS server is OGC compliant and not INSPIRE compliant
        b. Services available are wms and wcs
        c. Wms as based on WGS84
        d. Caching should only be carried out for stable datasets such as aerial imagery
        e. ArcGIS server will create the schemas whilst ArcGIS desktop creates the tiles
        f. Creates images 512x512
        g. Png24 is the best for topographic maps (256 col)
        h. Jpg for orthos as it has more colour

     i.   Then use WMS client such as GAIA 3,2 (opensource) – supports only 111, whilst OpenLayers supports 130

     j.   Use the new ArcExplorer since it gives the same outputs as Google

     k.   One can combine ArcGIS layers with Google Maps or MS Virtual Earth: to download the link code from Google for the former where one can also add geoprocessing tools such as print, etc as well as geostatistical tools

Day: 29th January 2009
Time: 10:00
Meeting with Dr. Nicole Ostlaemder
Objective Number: 3
Objectives of the meeting: To achieve knowledge of distributed geoprocessing and how it can be employed in the new datasets that will be acquired by Malta from the SF_Monitoring project.

10. Prediction modelling using spatial data and how they bring it together for dissemination purposes such as web portals and on-line analytical/querying systems

    xxviii)   Topics discussed:
        a.   Reviewed WSDL – webservice description language
        b.   Reviewed SOAP
        c.   Use WPS – Webprocessing service for transformation processes
        d.   Reviewed ORCHESTRA which is a service oriented architecture (FP6) for risk management – more of a method than a structure
        e.   Each data provider sends the data in WFS format – adds local schemas into a common schema
        f.   Handles both raster and vector transformations
        g.   MapAlgebra was used for NUTS3 and NUTS5 simulations
    xxix)   Distributed geoprocessing
        a.   Data intensive
        b.   Data light
    xxx)   Service composition
        a.   Client controlled
        b.   Cascading service changing (WPS)
        c.   Engine controlled services (BPEL engine – business process executive language)
        d.   Uses OASIS standard
        e.   Active BPEL software is available through opensource – engine is ORACLE
        f.   Running BPEL with WPS can become tedious as one needs to run xml
    xxxi)   Access services employed
        a.   FAS – feature access service
        b.   MAS – map access service (with stats)
        c.   SCAS – service chain access service
    xxxii)   Existing WPS interface exists in GRASS
    xxxiii)   One can use opaque chaining so that the users only view a single service
    xxxiv)   Use SOAP as an Interface representation (SOAP SCOPE)

xxxv)   Export videos using Camtasia for VLC player

Day: 30th January 2009
Time: 11:30
Meeting with Dr. Christian Giovando
Objective Number: 6
Objectives of the meeting: To achieve knowledge on risk-mapping outputs.

xxxvi)   Topics discussed:
a.   Review of the EFFIS
b.   Other risks mapping projects: water and drought
xxxvii) An overview was given of the EFFIS (European Forest Fire Information System) through the use of the EFFIS geoportal which output is based on the FWI (Fire Water Index) methodology and has data of the last ten years. The outputs are resultant from MODIS with daily updates in summer (fire season)
a.   The system has both archival data and predictive outputs
xxxviii) Issues highlighted:
a.   The resolution for this product ranges on a 30-50Km pixel size which in effect makes it difficult at local (Maltese level) to analyse since it produces a cell for Gozo instant and two for Malta (one North and one South)
b.   The system results in polygons of burn areas that are larger than 40Ha
xxxix)   Operationally two meetings are held every year: one before and one after the fire season
xl)   In the case of water and flood risk analysis, this is handled by a specific unit, and the issue revolves around a restricted system for flood alerts. Draughts are also monitored. The contact points are Jurgen Vogt, Paolo Barroso and Stefan Niemeyer

Day: 30h January 2009
Time: 12:30
Meeting with Dr. Dennis Sarigianni
Objective Number: 6
Objectives of the meeting: To review cross-discipline issues such as environment and health and the use of modelling.

xli)   Topics discussed:
i.   Dynamics between environmental, social and land-use spatial data
1.   Models that analyse across the disciplines
a.   Economic - physical
b.   Socio-economic – landuse
c.   Risk-based modelling
d.   Single-variable vs multi-variable analysis

General Recommendations/Conclusions:

This training was instrumental in reviewing and learning about new hi-tech tools and methodologies used for spatial, metadata and the whole datacycle focusing on the spatial infrastructure integration.

The main outputs from this training will be implemented within the Structural Funds project and technologies will be employed by the Information Resources Unit in its services to the Environment and Planning Directorates within MEPA as well as the core services provided by the Corporate Se4rvices for MEPA and other MT organizations.

Main issues

- Main Issue 1: The variety and complexity of GI-related tools employed by JRC indicate the linkages that are required between the different phases of the datacycle employed for environmental and land use data. The tools' progression show that JRC is committed to the dissemination of data through web-portals using hi-end and preferably opensource software. It is imperative that the GI specialists in MEPA become consonant with the tools identified in this document.

- Main Issue 2: This training was of a highly technical nature and the issues covered included an analysis of the spatial data employed for INSPIRE implementation particularly the dissemination processing such as web portals and on-line analytical/querying systems The system also looked prediction modelling tools and ecognition tools using spatial data structures and methodologies which will enhance MEPA's analytical construct for both planning and environment. The training however reviewed the linkages between the social-physical and environmental constructs.

- Main Issue 3: SEIS implementation at JRC is highly advanced and this training served as a basis for carrying out a hi-level input for HARMON project for JRC consumption. This project output will enable the harmonisation of data structures, metadata and dissemination as well as technical, operational and informational constructs.

References

Official Journal of the European Union, Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC (Aarhus). L 041 , 14/02/2003 P. 0026 – 0032 (28 January 2003)

Official Journal of the European Union, 17 November 2003b, Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, L 345 , 31/12/2003

Official Journal of the European Union, Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). L108, Volume 50 (25 April 2007)