A background document to discuss developing a cyber-infrastructure for *GlobalSoilMap.net*

Compiled by Bob MacMillan
on behalf of the *GlobalSoilMap.net* Consortium
15/04/2010

Background

The *GlobalSoilMap.net* project is primarily an information technology exercise. It has a goal of producing global maps of soil properties at fine resolution (90 m) that will be freely available and widely accessed. In order to meet this goal, the project obviously needs to design, develop and implement an effective cyber-infrastructure to support both initial production of the proposed maps and subsequent delivery of the map information through a web-accessible platform.

Under the present budget and work plan, the responsibility for developing and implementing cyber-infrastructure is formally assigned to the African node of the project (AfSIS) under Objective 2 of the funding arrangement with the Bill and Melinda Gates Foundation (BMGF). While this is nominally the case, the reality is that the AfSIS project is mainly focussed on the cyber-infrastructure and support needed for the African Soil Information System. A project website http://www.africasoils.net has been set up and is hosted at CIESIN. A rich query interface to permit scientists to access the raw data collection (for Africa) is being developed and a user interface fully integrated with the core AfSIS data system is planned.

So, while planning and implementation are well along the way to completion for the African node, there has been almost no discussion about the cyber-infrastructure needs for the overall *GlobalSoilMap.net* project. Apart from a few general statements about how the data will be freely accessible over the web there are very few specifics about what data will be delivered over the web, how it will be delivered and who will deliver it. Equally important, there has been almost no discussion about what is needed, in terms of cyber-infrastructure, to support the process of producing the maps of global soil properties.

This document is intended stimulate discussion, amongst project participants, about what is needed, in terms of cyber-infrastructure, to support production of soil property maps in the short term and dissemination of final output maps in the longer term. This will be anon-going discussion, but it has to start somewhere and sometime; hence this document.

A proposal for setting up a task group on Cyber-infrastructure

It is proposed that a task group be struck for the purpose of developing both a long term vision for what the GlobalSoilMap.net project will need in terms of cyber-infrastructure and a short term implementation plan for how to satisfy immediate cyber-infrastructure needs for capabilities for hosting and sharing data, programs, ideas, discussions and outputs among scientists who elect to contribute to the digital soil mapping comparison project.

The following individuals are identified as having appropriate skills and experience to contribute to this task group, Most have already been involved in some way in initial discussions about the cyber-infrastructure needs of the project. Other volunteer members would be welcome.

Sonya Ahamed	AfSIS and CIESIN
Daniel van Kraalingen	WUR on behalf of ISRIC
Peter Wilson	NRCS – NCSS
Andy Jarvis	CIAT
Panos Panagos	JRC

Questions and challenges for the task group

As an initial starting point for discussion the following questions and challenges are posed to this task group.

1. Produce recommendations for an overall design and accompanying implementation plan for cyber-infrastructure for the *GlobalSoilMap.net* project.
    a. What is our vision?
    b. Who will implement it?
    c. How will it be implemented
        i. Will it be centralized or decentralized?
        ii. Will it be hosted by a participating institute (e.g. ISRIC) or by Google?
    d. What will be the scope of its ambitions?
    e. How will it be paid for and by whom?
2. Produce recommendations and decisions about the project's short terms needs for functionality to support collaborative efforts to produce maps for pilot areas immediately and larger areas in the longer term.
    a. Do we try to create a central facility to store and serve common data sets?
    b. Can we benefit from a platform that supports collaboration and sharing of data, ideas, programs, outputs and plans amongst the various partners in the *GlobalSoilMap.net* project?

   c. If so, where should this platform be hosted and who should build and service it (CIESIN, ISRIC, CIAT, NRCS, JRC, Google)?
   d. What are the components of such a platform and do we have the skills and expertise available within the project to build and maintain it?
   e. How fast can we have one up and running to support the DSM comparison activities?
3. Produce recommendations and decisions about the project's long terms ambitions for eventual end-user delivery of final end products.
   a. What data will be included in a web based delivery platform? Will it consist solely of the 9 properties for the six depths identified in the specifications document or will it consist of a much broader range of spatial and non-spatial products?
   b. Will there be a single central delivery system or will the *GlobalSoilMap.net* project and ISRIC simply provide a point of entry to direct users to node-level delivery systems?
   c. Does the *GlobalSoilMap.net* project intend to simply produce data (maps) and deliver this data to end users or will it offer an ability to interact with and analyse the soils data on-line?
      i. Do we need to plan only for a platform that is capable of supporting web based discovery, visualization and download of data?
      ii. Or do we need to plan for a much more elaborate system capable of supporting web-based interactive analysis and query in addition to discovery, visualization, and download?
      iii. If interactive query and analysis are required, who will establish what kinds of queries and analyses are supported and who will design and implement the systems required to support them?