National Research FLAGSHIPS
Sustainable Agriculture
CSIRO

Advancing online delivery of Australian soil data and information

Peter Wilson
CSIRO Manager, National Soil Information
Coordinator, Australian Collaborative Land Evaluation Program

May 2012 (V2.0)

for Department of Agriculture, Fisheries and Forestry

www.csiro.au

aclep
Australian Collaborative Land Evaluation Program

Enquiries should be addressed to:

Peter Wilson, Manager National Soil Information
CSIRO Land & Water, Black Mountain Laboratories

GPO Box 1666 Canberra ACT 2601

Distribution list

Michele Barson	Department of Agriculture, Fisheries and Forestry (DAFF)
Mike Grundy	CSIRO, Sustainable Agriculture Flagship, Landscape Systems and Trends Theme

National Committee on Soil and Terrain

Contents

List of Figures

EXECUTIVE SUMMARY

The Australian Collaborative Land Evaluation Program (ACLEP) promotes a vision that "*natural resource management in Australia is underpinned by appropriate soil and land resource information and knowledge to ensure sustainable economic and environmental systems*". To support this, ACLEP delivers the Australian Soil Resource Information System (ASRIS), a collation of the best available, nationally consistent soil data and information for the country.

ASRIS data is provided by numerous data owners and custodians, mostly jurisdictional government agencies, through a cooperative approach overseen by the National Committee on Soil and Terrain (NCST). Variation in the data management systems (hardware and software) implemented by data providers and minor variation in data formats and content provide significant and time consuming hurdles to efficient national data collation.

With technological advancement and improved network connectivity and capacity is it now plausible to develop and implement mechanisms for controlled, online data delivery and update. ACLEP has been working to devise a plan for improved online data delivery to and from ASRIS which will streamline future activity and provide mechanisms for interaction with other providers and users of digital soil data – such as may be required for example for soil carbon data as part of a carbon farming initiative.

The plan involves a standards based approach and a web service oriented architecture. This means that transferred data packets, with known structures and content, can be validated against agreed information models and supported by controlled vocabularies, such that users can easily understand and use the content. Data will be made accessible in a variety of formats including as web and mobile services, providing exciting opportunities for increased access and use of soil data in a wide variety of web tools, online models and mobile device applications.

Implementation of ASRIS data in this manner will make it compatible with developing national systems, such as the National Plan for Environmental Information (NPEI). It will make it more accessible and useful for modellers, including for example, Department of Climate Change and Energy Efficiency FullCam national carbon accounting, the emerging Terrestrial Ecosystems Research Network (TERN) facilities and biodiversity modelling through the Atlas of Living Australia.

Recommendations for moving forward

- ACLEP should further progress the development of a standard information model for Australian soil and landscape information
- The information model should be compatible with international developments in this area being coordinated by the International Union of Soil Science working group on soil information standards, and more broadly with related environmental domains (such as geology, water and biodiversity)
- ACLEP should progress a series of demonstrator projects, in conjunction with the Bureau of Meteorology NPEI, and the DAFF wind erosion monitoring projects, to test and trial the implementation of a standards and services based system architecture

- NCST members should encourage jurisdictional soil agencies to implement online, standards based soil data delivery to progress the collation of soil site data within ASRIS and to make it available for other users and client applications
- ACLEP should implement delivery of standard national soil data products from ASRIS as a set of services, including Open Geospatial Consortium web services, mobile device services and Google Earth interface files
- ACLEP should explore options for more efficient storage and delivery of large national soil data sets in collaboration with TERN Facilities, including Soil and Landscapes, Ecoinformatics and AusCover
- ACLEP should further develop a mobile device application to promote open access and use of standardised Australian soil data.

1. INTRODUCTION

The Australian Collaborative Land Evaluation Program (ACLEP)[1] has as one of its key deliverables, the Australian Soil Resource Information System (ASRIS)[2]. ASRIS is a unique national soil information infrastructure that is underpinned by a centralised collation of the best available, consistent soil data from around the country. Collation and improvements to this data have been occurring since 1998. The recent focus has been on the provision of available mapped soil information and detailed site data supporting soil samples held within the CSIRO National Soil Archive[3]. A national collation of all available site based soil data is being progressed.

Historic and current processes for national soil data collation are onerous and tedious, often requiring a range of data and file formats to be manipulated and consolidated in a non-automated process. Ability to undertake rigorous data quality and error checks are limited, depending on availability of the centralised data custodian's resources, and relying on data providers delivering well structured data in a timely manner. Re-provision of corrected or enhanced data often results in a repeat of the manual collation process.

To improve the efficiency of this national collation and provide mechanism for standardised data transfers between providers and many users, a system of information models, transfer standards and federated web services is proposed. Implementation of this system will allow ASRIS data to be provided as web services and other on-line formats directly from the data custodians, allowing innovative use of authoritative soils data in an ever increasing range of user defined applications.

2. PROBLEM STATEMENT

The current soil data delivery process requires each jurisdiction to provide a physical copy of a data dump from their database. These need to be provided to the central database custodian (CSIRO) and reformatted and uploaded to the ASRIS databases.

This process is inefficient, often requires jurisdictional resources to be diverted from 'core' activities and results in out-of-date data being available to clients.

The alternative is to provide data directly from the data custodians, in a consistent and expected format through web services. The data could be consumed from these services into the centralised ASRIS databases, or visualised within the ASRIS web mapping client and other 3rd party clients (such as NPEI, ALA and TERN) as and when required.

There are no major technical impediments to establishing this federated, service-oriented spatial data infrastructure. The Australian AuScope[4] initiative and the Canadian Groundwater Information Network[5] have both established production systems using such an infrastructure.

[1] http://www.clw.csiro.au/aclep/
[2] http://www.asris.csiro.au/index_ie.html
[3] http://www.clw.csiro.au/aclep/archive/index.htm
[4] http://www.auscope.org.au
[5] http://gw-info.net/

The issues that need to be resolved are mainly institutional and relate to establishing a national soil data sharing collaboration supported by agreed national governance (Wilson 2012). The solution requires an agreed soil 'information model', the commitment of data custodians to establish the necessary web-based services and modifying client services, such as the ASRIS data portal, to consume these services.

3. THE ROLE OF ACLEP

ACLEP has a long standing role in the development of soil related guidelines and standards for Australia, including the Australian Soil and Land Survey Handbook series (e.g. NCST 2009), the Australian Soil Classification (Isbell 2002) and ASRIS Technical Specifications for mapped soil data (McKenzie et al. 2005). These documents provide significant material for the development of soil data and information standards and models which aim to define the soil and land features of interest, their relationships and the attribution vocabularies used to describe them.

It is important that information standards and models are authoritative and well managed. ACLEP provides a proven history and mechanism, through the National Committee on Soil and Terrain (NCST)[6] to deliver the required governance and on-going management of Australian soil information standards. ACLEP also provides a mechanism for delivery and maintenance of a reference implementation of national soil information standards through the Australian Soil Resource Information System. This ensures that conceptual information models, with their accompanying definitions of soil features, attributes and vocabularies, are sensible in an operational environment. ACLEP manages a repository for nationally defined attributes, such as jurisdiction codes and project identifiers, which are required to ensure uniqueness of individual data records.

Soils are an important component of ecosystems and as such cannot be considered in isolation in most applications. Soil data and information does not therefore operate in isolation and ACLEP provides a link to cross-sectoral environmental data integration initiatives, such as through the Bureau of Meteorology National Plan for Environmental Information (NPEI) or the Terrestrial Ecosystems Research Network (TERN).

ACLEP operates in the national arena (and more recently international) and is well placed to undertake multi-jurisdictional analyses (such as the north Australian land suitability assessment – Wilson et al. 2009) by utilising the consistent soil data collated within ASRIS. It is also well placed to develop and deliver national soil data products, such as interpretation of clay content in the top 30cm or water holding capacity for the top 1m. ACLEP can release these data via ASRIS for use by others, either as direct data downloads or as on-line, on-demand services accessible to a range of applications (such as Geographic Information Systems or Google Earth) without the need for users to connect to multiple services from different jurisdictions. However, neither the role of ACLEP nor the technology would prevent users directly accessing services provided by each of the jurisdictions if available.

[6] http://www.clw.csiro.au/aclep/contacts.htm

4. NATIONAL DATA COLLATION IN ASRIS

As mentioned above, the current process for national data collation is hampered by variation in the data management systems (hardware and software) implemented by data providers and minor variation in data formats and content provided to ASRIS. This makes for an inefficient process, which often has to be repeated manually if data sets are re-supplied, such as when errors are corrected or when new data becomes available. The manual nature of the process means that error checking cannot easily be implemented as part of the procedure, so data quality is dependent on data providers.

National data collation involves considerable volumes of data, both spatial and attribute. Maintaining unique identity linkages between the two is essential to ensure correct representation of data and integration of attributes within spatial information systems and models, such as in the case of digital soil mapping.

ACLEP is currently planning a new national collation of soil site data, to complement the existing map based data and for use in national digital soil mapping, such as will be undertaken by the TERN Soil and Landscape Facility[7]. Considerable amounts of new soil data are also anticipated to come to ASRIS in the near future, through activities related to national soil carbon and soil condition monitoring projects (wind erosion, pH and carbon) and TERN rangelands ecological sampling. This provides an imperative to greatly improve mechanisms for data delivery and collation.

In addition, requests for nationally consistent data products are becoming more frequent, such as by the Department of Climate Change and Energy Efficiency for national carbon accounting, or the Atlas of Living Australia for biodiversity modelling. The development of an information model and standards based data transfer mechanisms will work just as well for delivery of national data from ASRIS to a wide range of users, as it will for transfer of data to ASRIS and other data users by data providers.

The development of other essential components of a national soil information infrastructure, such as vocabulary services and registry services, will allow data providers to validate their own soil data base and service implementations against agreed national soil information standards. It will also allow users of these data services to discover their existence and bind to them with clear understanding of their content and behaviours. This is critical for the delivery of integrated and interoperable soil data and information in the future.

5. VISION FOR ONLINE DATA DELIVERY

ACLEP supports a vision that *"natural resource management in Australia is underpinned by appropriate soil and land resource information and knowledge to ensure sustainable economic and environmental systems".*

From an ASRIS technological viewpoint this vision will be delivered by the provision of well known, accessible and authoritative, national soil data and information services supported by a nationally managed and governed soil information infrastructure.

[7] http://www.tern.org.au/Soil-and-Landscape-Grid-of-Australia-pg17731.html

5.1 Stakeholders and linkages

Development of a functioning national soil information infrastructure involves linkages between many domains of expertise. Most notably the soil science and information delivery aspects need to be well integrated to ensure that the content of data services is both technically correct and meaningful, as well as properly managed and controlled to provide consistency and repeatability for users.

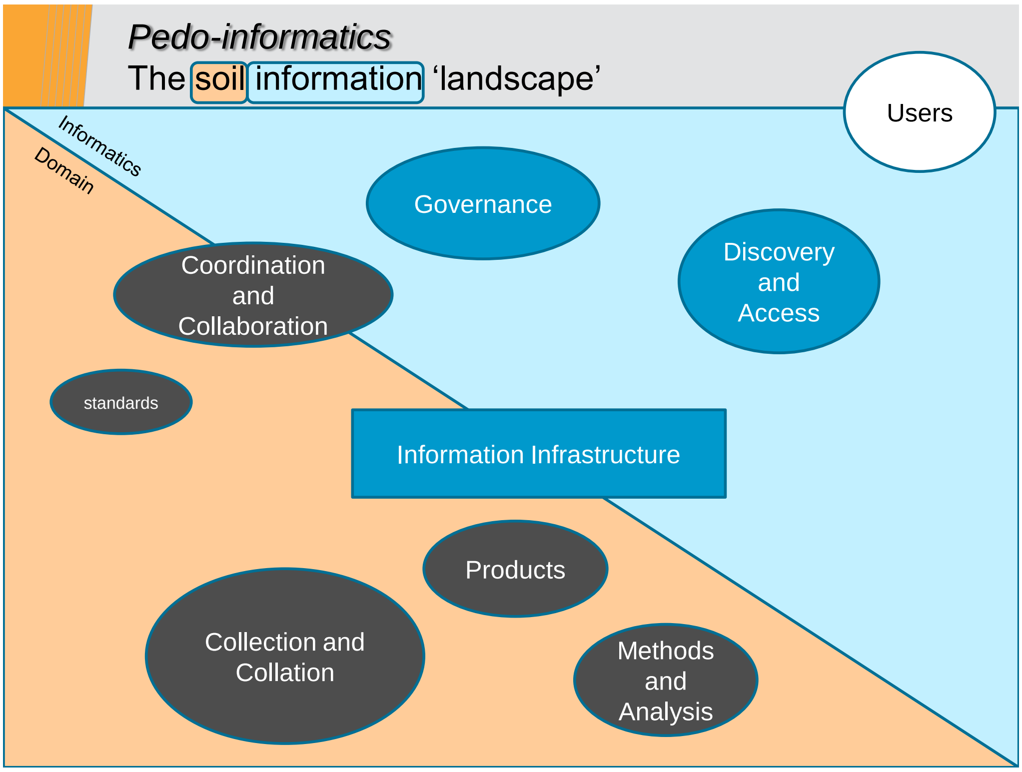

Figure 1 *Pedo-informatics* - interaction between soil science and information systems

There are many stakeholders in the delivery of national soil information, each operating within different spheres of influence and interest. These include the development of standards for the collection of primary source data as well as development of analytical methods and tools which provide a level of interpretation into soil information products. Information managers and systems designers are involved from the perspectives of information models and data base implementations as well as encoding information standards and tools to allow publication of valid and consistent soil data and information as useable products and services. The soil community, as well as broader interests are involved in the governance of soil data elements and the requirements for integration with other domains, particularly for whole ecosystem process analyses. Users of soil data operate at local to national, regional and international levels and have varied interest in accessibility and integration at those levels.

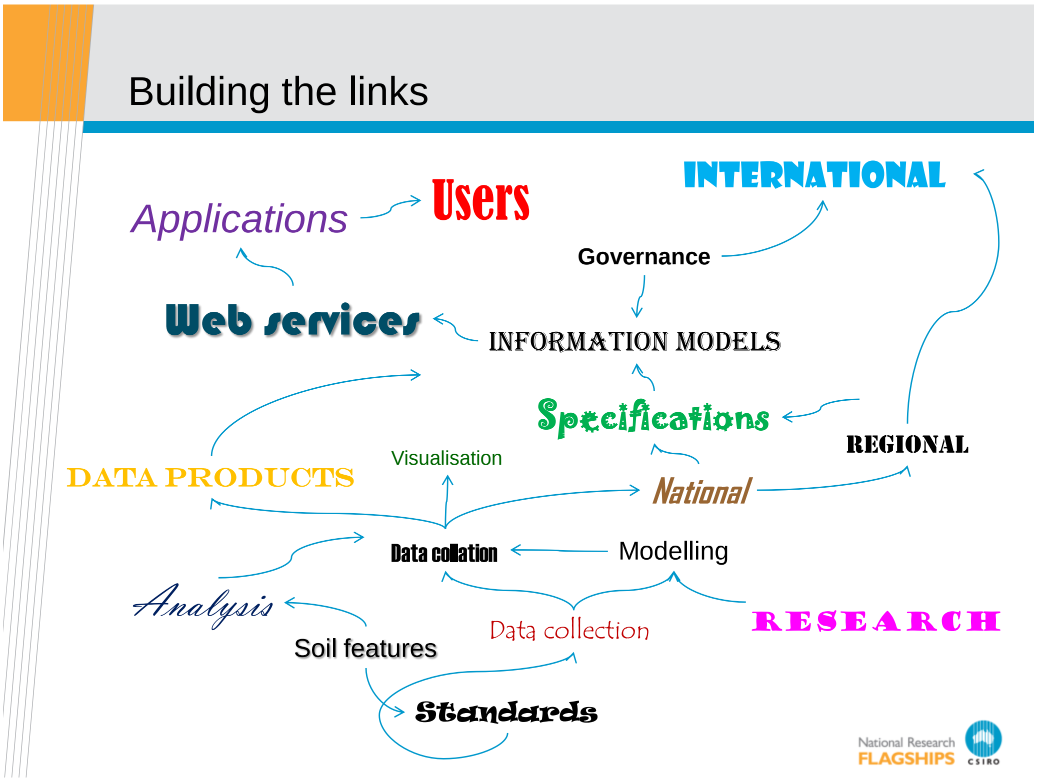

Figure 2 Soil information infrastructure linkages and interactions

Whilst it is important for ACLEP activities to be aware for all these stakeholders and various interests, it is essential, at least in the early stages, that the national soil infrastructure is focussed on delivery of Australian specified standards for soil data and information products. In the first instance this is likely to include delivery of site specific data according to the Australian Soil and Land Survey Field Handbook, mapped soil data according to existing specifications for ASRIS and development of new models and specifications for grid/raster representations of specific soil attributes (in consultation with emerging requirements from the TERN Soil and Landscape Facility, and Australia's role within the Oceania Node of GlobalSoilMap.net[8]). System components which will be applicable across environmental domains, such as meta-data catalogues, vocabulary and registry services will need to be developed in collaboration with other domains, through engagement with the National Plan for Environmental Information.

[8] http://www.globalsoilmap.net/

5.2 System architecture

The components required for delivery of a national soil information system can be conceptualised as a simple stack of interconnected layers, each providing control and functionality of an integral part of the total system.

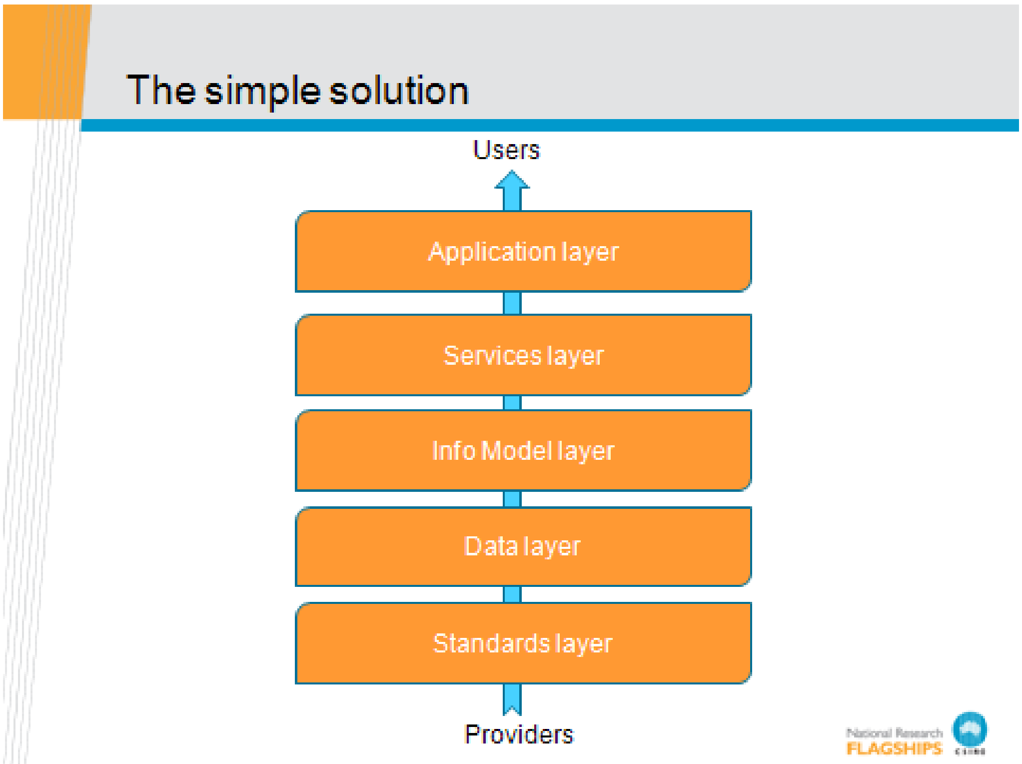

Figure 3 Components of a standards and services based soil information infrastructure

These layers provide seamless connectivity between data providers and users. They also allow flexible interaction at many levels of the system, such that anyone can potentially deliver as many of the components as their business model requires. For example, a state agency may just want to deliver a particular data set, while another may want to implement their data as an online service compliant with the underlying information model. Applications consuming the available data services may be developed by the provider agencies, a national collaborative initiative such as ACLEP, or by third party private enterprises.

The following sections provide a description of the individual components and a summary of relevant activity and the current status.

5.2.1 Standards Layer

The standards layer is established by the soil domain. It defines the soil features of interest, their relationship, their attributes and the vocabulary with which they are described and recorded. Specification of these features should consider the real world, tangible features that are the responsibility of this particular domain – i.e. they should be real world soil and landscape features.

A domain should refrain from defining and describing features from other domains as this will potentially lead to inconsistency and misinterpretation. For example, rock types are of interest to the soil domain as a parent material for soil development. However, rocks are of special interest to the geology domain and have already been described in detail by that community. If the soil domain was to 'invent' new definitions of rocks then they would be soil domain specific, not authoritative and not broadly useable by others. Similarly, geologists interested in hard rock geology should not be tempted to define overlying soil features and characteristics.

In Australia, the soil features of interest have largely been identified and agreed through the development of the Australian Soil and Land Survey Series Field Handbook (NCST 2009), and other volumes concerning chemical and physical analysis methods(Rayment and Lyons 2011, McKenzie *et al*. 2002). Real world soil and land features are defined at many scales of interest, from tracts or areas of land (surface features), through soil profiles and horizons (vertical depth features) to soil structures such as peds, cutans and voids (morphological features). Being a continuous substance, with no real boundary limits and variable and changing attributes, soil is difficult to clearly define in terms of spatial and temporal characteristics. Attempts to put identifiable spatial constraints on soil volumes of interest have led to terms such as the pedon and poly-pedon (Soil Survey Division Staff 1993) and soil individual (McKenzie *et al*. 2002) although these still recognise internal variance of soil attributes and are somewhat arbitrary in their dimensions. Similarly the spatial dimensions at which different features are expressed mean that a variable spatial reference can be attributed to any set of soil and landscape observations. For example, land surface (site) dimensions may be large (300m) when considering broad landscape attributes such as landform pattern, smaller (25m) for recording attributes such as landform element or slope, or be constrained to a few square metres or a discrete location for recording soil surface conditions. Location, dimension and time therefore become important attributes of soil feature delineation and description. Information models need to account for this variability and ensure recording of spatial and temporal dimensions where appropriate.

Published standards for soil feature description and attribution, such as the Handbook series, and standards for soil type classification, such as the Australian Soil Classification (Isbell 2002) provide a ready reference for Australian soil data codification. Many surveys and field investigations have been undertaken using these standards and future studies should be encouraged to make use of them. Existing standards should form the basis of the Australian soil information infrastructure. Re-interpretation of historic soil surveys into these agreed standard approaches may be required in some instances. Adoption of other internationally recognised standards, such as the World Reference Base classification (IUSS Working Group WRB. 2006) may require extension of soil information models or translation from Australian schemes and terminology.

5.2.2 Data Layer

The data layer provides the physical implementation of data storage. This may be in a variety of formats and systems, more recently as digital data sets, but also commonly as paper based field records or printed products (e.g. reports and maps). For the future Australian soil information infrastructure it is assumed that data will be encoded into digital systems.

Individual soil agencies and others (industry, research, community or private) have variously implemented their own data management environments over the years. These now represent significant investments and have been developed to deliver specific business requirements of organisations. There is no expectation that a new approach to individual data system implementation will be required. The proposed infrastructure will provide mechanisms whereby individual system implementations can provide independent representations of their data holdings as standards compliant packages, which can then be directly used or translated into user specific formats.

ASRIS is currently implemented as a collection of SQL Server, MS Access and ArcGIS spatial and attribute data sets. The current ARSIS data model (McKenzie *et al.* 2005) allows data at a range of scales, from site based records through to local and national resolution map products. For national consistency ASRIS data is implemented in a 5 layer depth model based on functional soil layer concepts. Data complying with the ASRIS model needs to be translated into this depth model for national consistency. In future implementations, particularly as new grid/raster data is developed, attributes values may be recorded at 6 standard depth (slices) or as functions of a spline, such that any depth of interest may be calculated. Accounting for recognised soil horizons or other profile features in a standard depth model needs to be further considered.

5.2.3 Information Model Layer

An information model provides a conceptual framework for recording the attributes of identified soil features. It defines the relationship of features and the vocabularies used to describe the attributes or characteristics. An information model is software agnostic such that it can be implemented, in various forms, in many data base packages, depending on specific user requirements. Depending on the defined feature’s relationships and dependencies, a user may implement only parts of an information model. In other cases a user may extend the model beyond agreed boundaries, if their particular use case requires such additions. In either case, the community of practice, or domain, needs to specify certain minimal requirements so that data from any source complies at least with a central, well used concept.

Previously, an information model was commonly encapsulated within a specific data base implementation, such as the ASRIS SQL server instance. The feature definitions and relationships were maintained within an entity relationship diagram and implicit rules and conditions were managed within accompanying documentation, or not documented at all. This makes use of such data bases, and transferability of the implied relationships between various software platforms very difficult. The preferred method for documenting an information model is now to use an independent application language such as Unified Modelling Language (UML). A UML model contains all the classes of features, their associations and relationship to each other and explicit links to controlled vocabularies and rule sets.

Using a UML approach, a user can implement a data base solution in any preferred software and use validation routines to test that their data are compliant with the agreed information model. Data can be encoded in a variety of formats (such as an XML or CSV file) for transfer to other users or applications and can be readily understood by re-applying the agreed information model.

5.2.4 Services Layer

As suggested above, data may be transferred between providers and users in a variety of formats. All that is required is for the user to be able to read the file and convert it for use in their particular application. This may require some form of middle-ware for file conversion but this would be useable for any data provided in suitable formats if it is compliant with the agreed information model. This provides a great benefit as data from many sources can readily be ingested into user applications.

Publishing data as a web enabled service is becoming more commonplace. A number of published standards for web services now exist for different types of data, such as those developed and supported by the Open Geospatial Consortium (OGC)[9]. These standards define the user interaction with the service (a query-response relationship) and the formats for the returned package. Data may be returned, for example, as an image file (e.g. a .png or .tif) in some services, such as with an OGC web map service (WMS) or a Google Earth keyhole markup language (KML) file. Alternatively full data records may be transferred by some web services, such as through an OGC web feature service (WFS) which delivers data as an XML encoded response. User applications need to be able to send a request to a web service in an expected, explicit form and be able to interpret the returned file. Web service standards define the machine to machine interaction of query-response processes but they do not (generally) define the content of the returned files. An agreed, published information model is therefore required to allow users to fully interpret the returned files.

A selection of ASRIS data layers are already available in a number of web service formats, including as OGC WMS/WFS and some web coverage service (WCS), Google Earth KML and more recently as ArcGIS mobile data services. The existence of these services is not generally published as they are not well supported within the current capacity of ACLEP, have minimal documentation, or are subject to change or removal. The capacity of the existing ASRIS data servers to support multiple users of these services is also largely untested. In addition the content of the data services is dictated by the current ASRIS data base implementation and is not accompanied by an agreed information model beyond the ASRIS technical specifications and layer meta-data. In the future, as new data layers become available and with development and management of a domain information model, this form of open publishing of soil data and information services will increase. Also, once standards for these services have been established, other providers of soil data, such as state agencies, will be able to publish compliant services of their own data.

[9] http://www.opengeospatial.org/

5.2.5 Application Layer

An application is a software tool that consumes and value adds to a data input. It may be a data base or geographic information system that allows users to interrogate and analyse data, or may be a visualisation tool that allows soil data to be displayed as context with other data, such as in the Atlas of Living Australia.

Commonly, users acquire a copy of a data resource and incorporate it within their own data management implementation. This allows users to manipulate input data and provides efficiencies in processing where large volumes of data are analysed. It may mean that data is separated from its authoritative source and care should be taken to ensure that users have access to the latest and best available data sets. Many examples exist of data that has been acquired in the past for one application, been modified within an organisation and then passed around to multiple users. In such instances the original source of the data may become unknown and is possibly not acknowledged, users do not get notified of modifications to the data (such as error corrections), or data providers are not contacted by future users to see if updated data sets are available.

A new approach is for applications to make a live connection to data published as web services, as discussed above. This allows for an application to make an on-demand, on-line call to an input data source and access the most recent available data inputs. Data updates and corrections made by the authoritative custodian of data resources are immediately reflected in the available web service ensuring that users access the best available data.

When data is published as a web service, compliant to an openly available information model, then multiple applications can be built which can access and consume the data through known query-response relationships. In this way the “publish once, use many-times” paradigm is implemented. Applications will be able to be developed by any user, often unknown to the data service provider, allowing an explosion of previously unknown use of soil data and information. The only danger for application developers in this scenario is the online availability of data services to which they connect. For business critical applications it is strongly advised that users develop some kind of service provision understanding or agreement with service providers. They should also implement back-up caching of services so that their application continues to function should connectivity to a service be unavailable.

6. ISSUES AND OPPORTUNITIES

6.1 Current activities and developments

6.1.1 Data collation

Over the last 5 years, significant effort has been directed to the collation of the best available mapped soil data across the country. The national ASRIS data base now has a considerable collection of map data at various levels, attributed according to standardised technical specifications. However, a number of gaps occur both in terms of spatial coverage and attribute completeness and ongoing activity is required to fill these where data is known to exist.

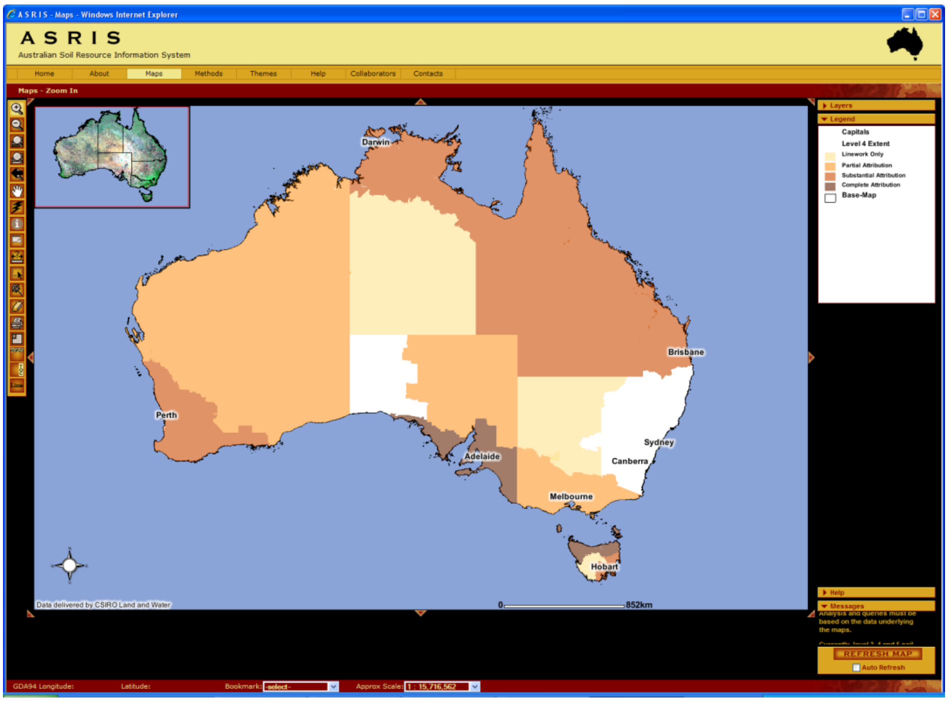

Figure 4 ASRIS data coverage for Level 4

ACLEP has also been working to improve a number of continental framework data sets, including additional interpretation of the Atlas of Australian Soils (Northcote et al.1960-68, McKenzie and Hook 1992, McKenzie et al. 2000); a revision of Physiographic Regions (Pain et al. 2011) and development of 250m resolution estimates[10] of key soil properties (such as clay% and carbon % n the top 30cm) based on the best available spatial soil data. These data are valuable as surrogates for missing data and as inputs and frameworks for further, more intensive studies.

With the recent improvements in digital soil mapping and assessment techniques it is now possible to make estimates of a number of soil attributes at fine resolution across the country

[10] http://www.asris.csiro.au/themes/NationalGrids.html

and to quantify the errors and uncertainties of those estimates. To progress this activity a new national collation of available soil site data and analytical results is anticipated. Existing ASRIS site data models have been refined to accommodate a range of site based data (including traditional soil survey results, new carbon quadrat sampling data and a range of soil condition data collected from transects, and quadrats of variable dimensions).

Further advances in information modelling, online data transfer and validation tools is expected to greatly improve future national data collation efforts.

6.1.2 Information modelling

The development of a conceptual information model for soil data and information is a relatively new activity. Initial progress has built on related activities in the geological and water communities. However, the required skills and capacity for information model development are scarce and it has been difficult to attract skilled operators to the soil domain.

Some basic soil information model development has occurred under the banner of the GlobalSoilMap.net project. This has also led to the formation of an International Union of Soil Science Working Group on Soil Information Standards[11], which seeks to coordinate a number of project activities in this area. Australian, and particularly CSIRO input and leadership in these activities ensures that they are aligned where possible to the needs of ASRIS. Considerable activity in Europe associated with development of information models supporting the Inspire Directive, and some ISO soil quality standards development have provided a good foundation for internationally applicable information models.

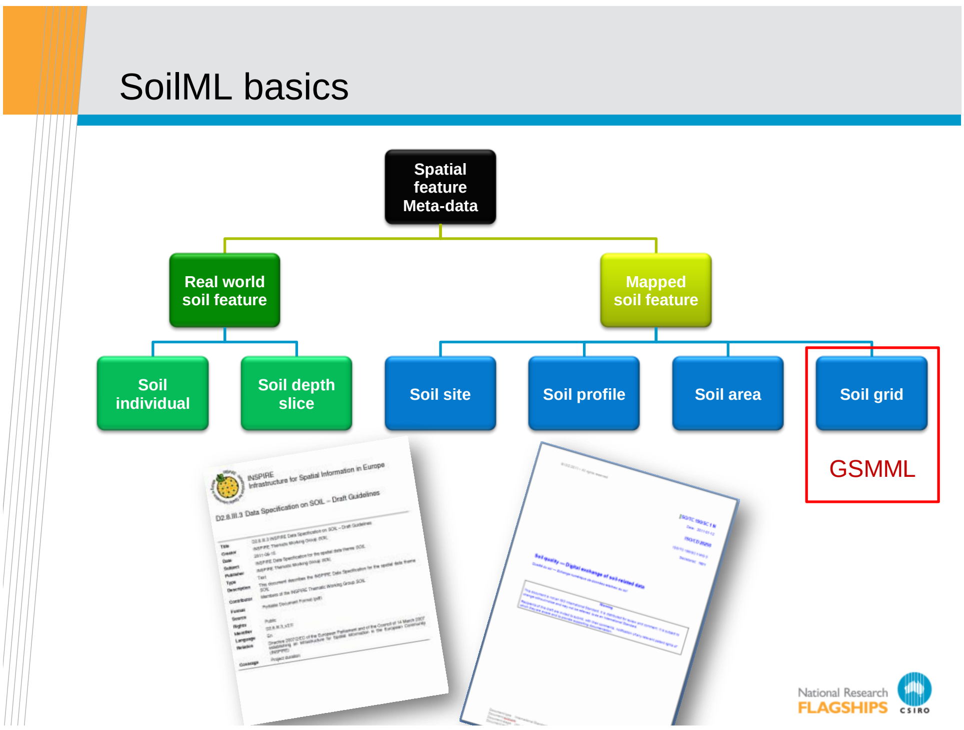

Figure 5 SoilML information model features of interest and international standards developments

11 http://www.soilinformationstandards.org

Australian "OzSoilML" Information modelling

ACLEP information modelling activity has focused on the development of an information model, called "OzSoilML" , based on Australian soil guidelines, standards and classifications. The strength of this model lies in the explicit identification of the real world soil and landscape 'features' about which the description of attributes and analytical results maintained in databases belong.

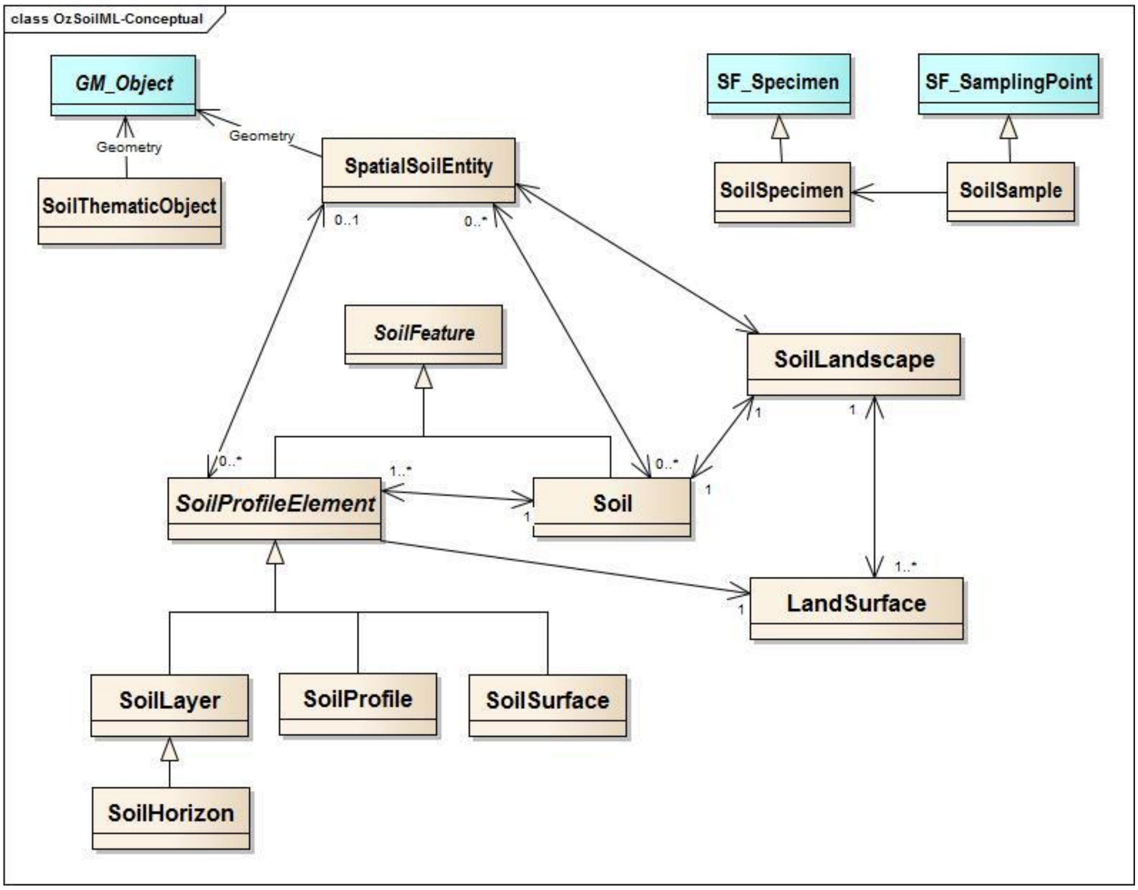

Figure 6 The high level conceptual features specific to the Soils Community included in OzSoilML.

OZSoilML is now being used in an implementation trial to map the contents of the ASRIS site database, NatSoil, to the conceptual model for publishing of standardised web services. These services will be made available to the National Plan for Environmental Information as part of a demonstration of online access to a range of thematic data layers, including soil, climate, water and marine.

As the OzSoilML information model matures, ACLEP will consider its use for supporting whole data base exchange, such as will be required by the national site data collation, or for delivery of standardised soil data products from ASRIS and the TERN Soil and Landscape Facility.

OzSoilML will be further progressed and integrated with related information modelling activities in New Zealand aimed at delivery of standardised national soil data products compliant to the specifications of GlobalSoilMap.net.

6.1.3 Services delivery, data, vocabulary, registry

Soil data web services have been available for ASRIS related data for some time. They represent the current content and data model of the ASRIS data base and as such are not broadly applicable as data transfer mechanisms. They do provide a link to nationally consistent data and have been accessed by a number of users through various applications. However, they are somewhat experimental at this stage and do not represent a supported source of national soil data. As a robust information model is developed, it is anticipated that ASRIS data will be transformed to comply with this community agreed model. Additional tools and software development will be required to allow on-the-fly publishing of ASRIS data in this way. Similarly, possible alterations to the underlying ASRIS data structures and content may be required to support all components of the new model.

In addition to development of data services, a number of supporting web service based functions will be required. These include an online soil vocabulary service, which allows data coding and terminology to be clearly interpreted and applied, both by human users and potential on-line machine-to-machine based applications. Another necessary system component is the registry service, which allows discovery and connectivity of available data services that have been registered. Such a service would typically also deliver meta-data regarding the available service and the content of the data within it, allowing users to make value judgments on the fitness for purpose of particular services.

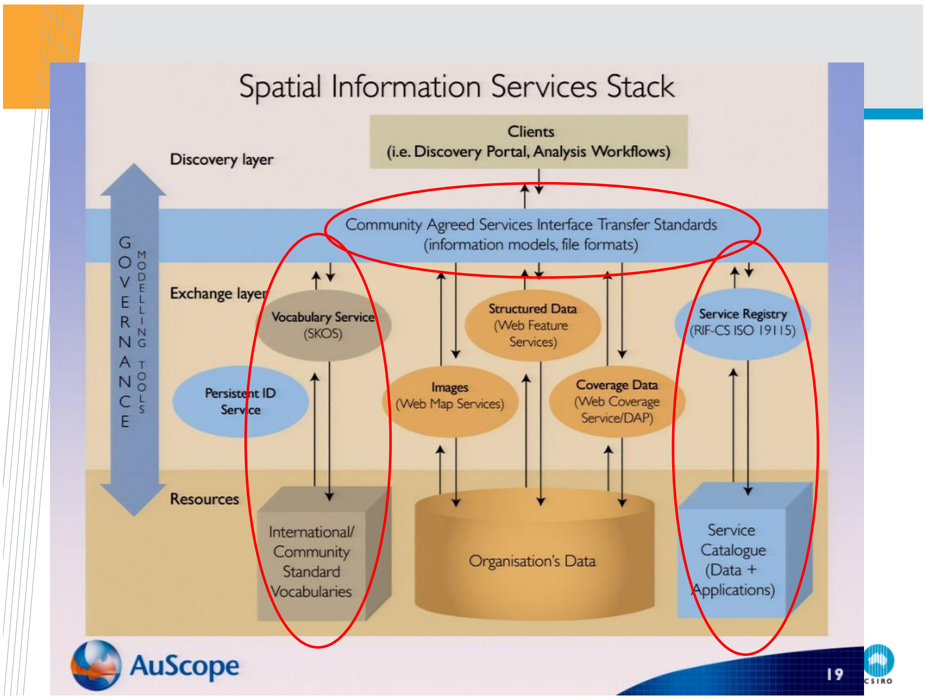

Figure 7 Spatial Information Services Stack (from AuScope[12])

Components such as vocabulary and registry services do not necessarily need to be developed or hosted by individual community domains. In fact, these services may be best developed and maintained by a centralised function, or at least through a federated approach. In this way services from multiple providers could be discovered and accessed. System efficiencies would also result from single implementations of vocabulary and registry services supporting multiple

[12] http://siss.auscope.org

domains. An initiative such as the Bureau of Meteorology's National Plan for Environmental Information could be well placed to provide these centralised services in the future.

6.2 Capacity for implementation

Many initiatives and agencies have developed centralised data repositories and web mapping interfaces over the last few years. Examples include the Australian Natural Resources Atlas, many state agency atlases and recent initiatives such as the Atlas of Living Australia. Except where these are developed and maintained within a single agency, primarily serving data which that agency has dedicated custodial responsibility for, the long term delivery of these initiatives is challenging. This is mostly due to the fact that initial data acquisition and re-formatting requirements take considerable project resources (typically 80% of project funds) and these processes cannot be repeated as new data becomes available if limited project funds are exhausted. Data acquired from various sources quickly looses connectivity to the original source, and in fact may become the default source for that data for future use within the acquiring agency. This means that users can be unaware of data updates and improvements, may utilise data which has been incorrectly interpreted and may not appropriately attribute the data to the original source.

Implementation of an on-line, on-demand data access system, such as a standards and services based infrastructure will eliminate many of the current issues for both data providers and users. Many agencies are now in a position to implement web service based data delivery, and many already do (eg WA SLIP[13], Information Qld[14]). Similarly many software platforms and applications (e.g. ESRI ArcGIS, Google Earth) are now able to consume web services, particularly mature OGC services (such as WMS and WFS) and other common file formats such as ESRI shapefiles and Google Earth KML files. Software (such as GeoServe, MapServer or 52North) for translation of native data base content and file formats into community information model compliant web services is also becoming more commonly implemented. Significant advancement has occurred in the geological information and water related domains and these learnings are readily transferable to associated soil and other environmental data, often within the same responsible agencies or through cross-agency information technology support. Such cross-domain adoption of common standards will be necessary for the NPEI vision to be realised.

The cessation of the National Land and Water Resources Audit in 2008 created a gap in communication and coordination of environmental data collation and sharing activities across multiple domains. New initiatives such as the Atlas of Living Australia are attempting to fill the consolidated data repository and delivery role, but suffer from short term project funding and a lack of mandate for cross domain coordination. The National Plan for Environmental Information may provide a solution in the future, but it will have to balance a centralised service delivery model against a distributed, community model.

[13] https://www2.landgate.wa.gov.au/web/guest
[14] http://www.information.qld.gov.au

7. RELATED ACTIVITIES

ACLEP activities, and particularly CSIRO supported initiatives, aimed at national and global soil data and information coordination provide a significant boost to development and implementation of national soil data standards and web service based data delivery.

At the international level, involvement in the GlobalSoilMap.net project is helping to deliver specifications for globally consistent soil data at fine resolutions with improved statement of error and uncertainty limitations. Leadership of the Oceania Node and direct involvement in many coordination and technical committees, regionally and internationally, ensure that Australia is well placed to delivery available data into this arena. Involvement at key levels in GlobalSoilMap.net also aims to guide international specifications and standards towards operationally implemented systems within Australia, such as ASRIS. Activity through GlobalSoilMap have also developed links to other key international organisations such as the Open Geospatial Consortium, International Union of Soil Science, European Inspire Directive, International Standards Organisation and the proposed United Nations Global Soil Partnership. All of these are relevant to the development of online soil information systems.

At the national level, interaction with the TERN Soil Facility is particularly important. This Facility has available resources to further develop and apply methods for improved estimation of key functional soil attributes across the country. This will greatly improve the national data products available through the current ACLEP workplan. The TERN project will be largely dependent on the existing ASRIS data and will look to ACLEP to facilitate data analysis and delivery of the final TERN data products. This is a great opportunity to progress the information models and web services proposed for the future enhancement of the national soil information infrastructure.

7.1 NPEI demonstrator

Of particular interest, and to fit soil information initiatives within a broader environmental information framework, is the progression of a National Plan for Environmental Information. A number of discussions have been had with the BoM to ensure that the proposed future online delivery mechanisms for soil data are compatible with the developing plan. Particular interaction has focussed on soil data governance issues and opportunities to develop supporting components of the proposed system architecture through NPEI activity. The latter will consider those components not specific to the soil domain, such as vocabulary and registry services. A demonstration pilot including discovery, access and use of national soil data published as web services in now being progressed.

7.2 Soil app development

An exciting advancement in recent years is the increasing popularity and availability of mobile data devices, such as smart phones and tablet computers (e.g. the IPad). These devices provide an excellent platform for accessing and visualising mapped data such as soil information. Coupled with in-built GPS capability, users can also expect to access specific data relevant to their location.

Recent discussions with a number of potential project proponents have identified a considerable interest in development of a soil app, delivering standardised and specific soil data for Australia.

Initially this will be targeted at agricultural farm advisors and could progress to include field validation of delivered soil data, modification of on-site soil attributes and submission to on-line crop modelling applications.

A pilot project proposal has been progressed initially through CSIRO application coding in consultation with a GRDC project and reference group based in South Australia. Adopting standardised information content and web service architectures means that the app should be applicable to all ASRIS data, potentially covering all Australia. As international soil information models are developed the app could be modified to work globally. There is already interest through the GlobalSoilMap.net project in the soil app concepts and particularly for delivery in Africa where the first sets of new GlobalSoilMap data are about to be delivered.

As an initial step in the development a sub-set of ASRIS data has been made available as a series of ArcGIS mobile data services. These include data for the Atlas of Australian Soil, and ASRIS Level 4 data depicting Australian Soil Classification and pH values. Using a recently released ArcGISApp from ESRI, this data can now be accessed through an IPad.

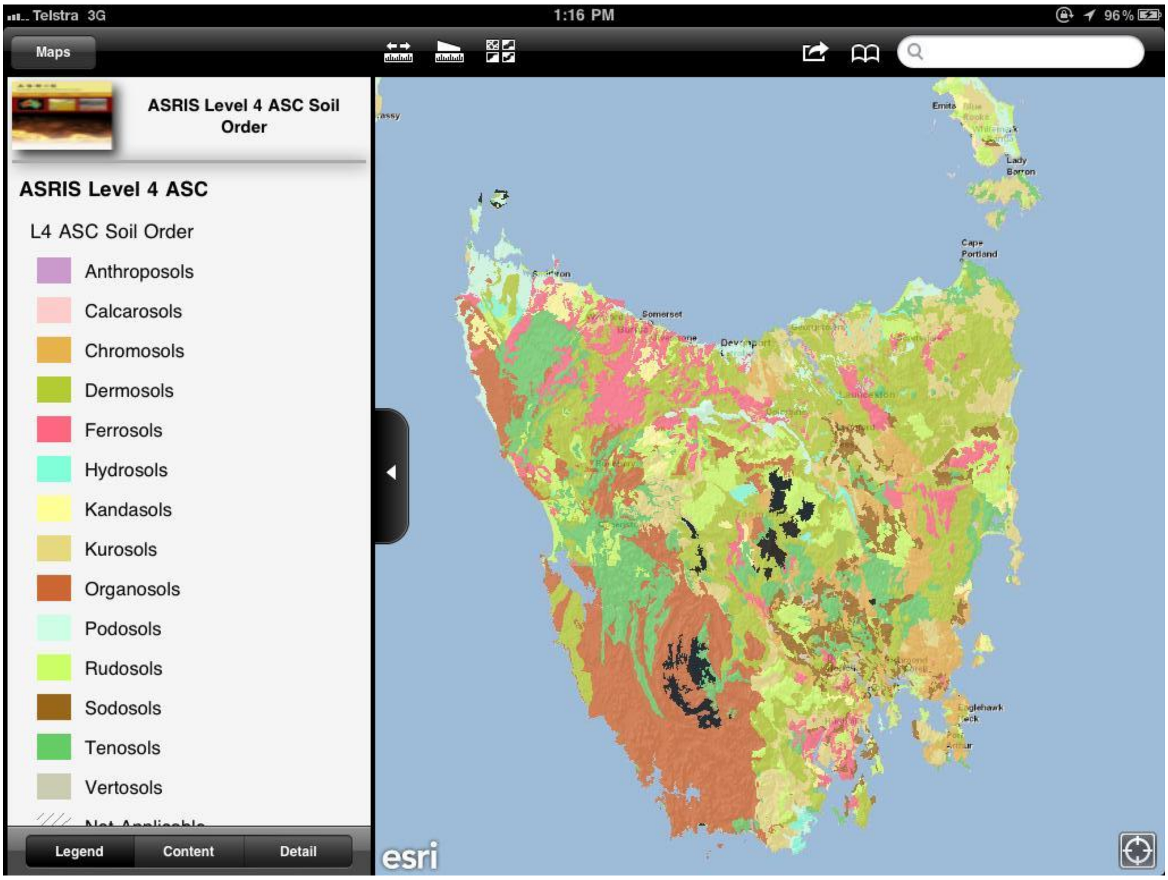

Figure 8 ASRIS Level 4 data accessed on-line through an IPad App.

Development of the pilot app, known as "SoilMapp" has now progressed significantly to include functionality for display and location based query of ASRIS map, site and CSIRO Soil Archive site data sets. Inclusion of soil moisture data from APSoil characterised sites has also been possible. A public release of "SoilMapp" is planned for the second half of 2012 and will demonstrate the value of providing soil data through online, on-demand web services.

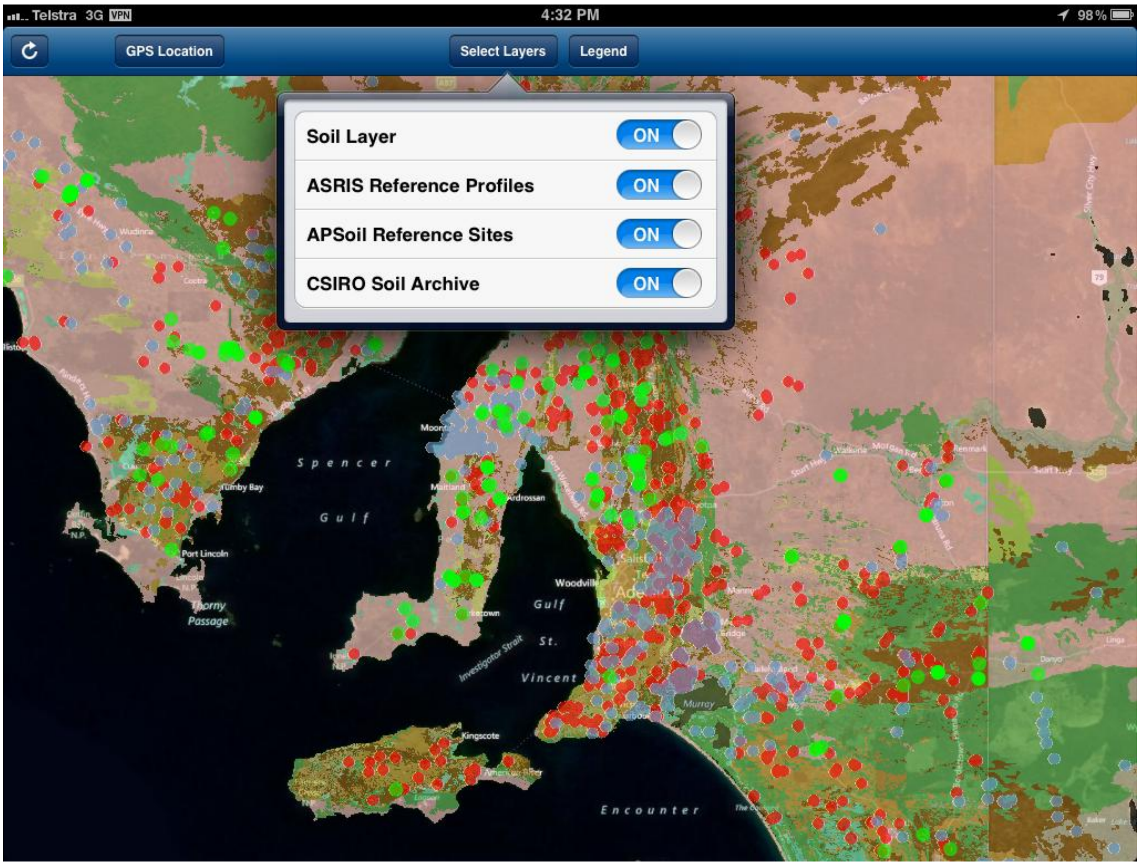

Figure 9 "SoilMapp" pilot IPad application showing different soil data sources from ASRIS, CSIRO Soil Archive and APSoil web services

8. MOVING FORWARD

Technological advancements in on-line information infrastructures provide a number of opportunities for the soils community. The rapid rate of development presents a challenge to skills and capacity, but implementation will eventually provide significant system efficiencies for both data providers and users.

For the Australian (and international) soil community to advance on-line delivery of soils data and information a number of component activities are required. These include development of:

- An agreed soil/landscape feature catalogue
- Integrated models for site, vector and grid data
- Conceptual soil information model
- Controlled vocabulary and service
- Data transfer standards and tools
- Data base implementations
- National data collations
- NCST endorsed national soil data products

ACLEP is committed to progressing development of a new generation, national soil information infrastructure. It will continue to improve the national data collation in ASRIS, including collation of available soil site data and generation of improved estimates of key soil attributes for the whole country. ACLEP will develop, in consultation with the soil community, an agreed conceptual information model for soil features of interest and work to embed Australian requirements within developing international standards. Through engagement with the National Plan for Environmental Information, the required system including data governance structures, licensing and service stack components will be progressed. ACLEP will develop and publish national soil data sets as standard compliant web services and will develop applications to further facilitate the broader access and use of Australian soil data and information.

REFERENCES

Isbell, R. F. (2002). The Australian Soil Classification. Revised Edition. CSIRO Publishing, Melbourne.

IUSS Working Group WRB (2006) World reference base for soil resources 2006. A framework for international classification, correlation and communication. World Soil Resources Reports No. 103. FAO, Rome.

McKenzie, N. J. and Hook, J. (1992). Interpretations of the Atlas of Australian Soils. Consulting Report to the Environmental Resources Information Network (ERIN). CSIRO Division of Soils Technical Report 94/1992.

McKenzie, N.J., Jacquier, D.W., Ashton L.J. and Cresswell, H.P. (2000) Estimation of Soil Properties Using the Atlas of Australian Soils. CSIRO Land and Water Technical Report 11/00.

McKenzie, N.J., Coughlan, K. and Cresswell. H. (2002) Soil Physical Measurement and Interpretation for Land Evaluation. CSIRO Publishing.

McKenzie NJ, Jacquier DW, Maschmedt DJ, Griffin EA and Brough DM on behalf of the National Committee on Soil and Terrain Information (2005). Australian Soil Resource Information System, Technical specifications Version 1.5. Australian Collaborative Land Evaluation Program ACLEP.

National Committee on Soil and Terrain (2009) Australian Soil and Land Survey Field Handbook Third Edition. Australian Soil and Land Survey Handbooks Series 1. CSIRO Publishing.

Northcote, K. H. with Beckmann, G. G., Bettenay, E., Churchward, H. M., Van Dijk, D. C., Dimmock, G. M., Hubble, G. D., Isbell, R. F., McArthur, W. M., Murtha, G. G., Nicolls, K. D., Paton, T. R., Thompson, C. H., Webb, A. A. and Wright, M. J. (1960-1968). Atlas of Australian Soils, Sheets 1 to 10. With explanatory data (CSIRO Aust. and Melbourne University Press: Melbourne).

Pain, C., Gregory, L., Wilson, P. and McKenzie, N. (2011) The physiographic regions of Australia – Explanatory notes 2011. Australian Collaborative Land Evaluation Program and National Committee on Soil and Terrain.

Rayment, G.E. and Lyons, D.J. (2011), Soil Chemical Methods - Australasia . Australian Soil and Land Survey Handbooks Series. CSIRO Publishing.

Soil Survey Division Staff (1993) Soil Survey Manual, United States Department of Agriculture Handbook No. 18.

Wilson P.L. (2012) Improving Australian Soil Data and Information Governance. Australian Collaborative Land Evaluation Program.

Wilson P. L, Ringrose-Voase, A, Jacquier, D., Gregory, L., Webb, M., Wong, M.T.F., Powell. B., Brough, D., Hill , J., Lynch, B., Schoknecht, N. and Griffin. T. (2009) Land and soil resources in northern Australia Chapter 2 - 1 in Northern Australia Land and Water Science Review *full report.* Northern Australia Land and Water Taskforce.

Contact Us
Phone: 1300 363 400
+61 3 9545 2176
Email: enquiries@csiro.au
Web: www.csiro.au/flagships

CSIRO and the Flagships program

Australia is founding its future on science and innovation. Its national science agency, CSIRO, is a powerhouse of ideas, technologies and skills. CSIRO initiated the National Research Flagships to address Australia's major research challenges and opportunities. They apply large scale, long term, multidisciplinary science and aim for widespread adoption of solutions.