D3.5: User requirements for the IT infrastructure

Project No. 869200

Version 1.0
May 28, 2021

Deliverable: Soils4Africa_D3.5_Requirements_v1.0

Contact details

Director of Coordinating Institute – ISRIC: Rik van den Bosch
Project Coordinator: Mary Steverink-Mosugu
Address: Droevendaalsesteeg 3, 6708 PB Wageningen (Building 101), The Netherlands
Postal: PO Box 353, 6700 AJ Wageningen, The Netherlands
Phone: +31 317 48 7634
Email: mary.steverink-mosugu@isric.org

Project details

Project number	862900
Project acronym	Soils4Africa
Project name	Soil Information System for Africa
Starting date	01/06/2020
Duration	In months 48
Call (part) identifier	H2020-SFS-2019-2
Topic	SFS-35-2019-2020 Sustainable Intensification in Africa

Document details

Work Package	WP3: Design of the SIS and methods for field and laboratory
Deliverable number	D3.5: User requirements for the IT infrastructure
Version	1.0
Filename	Soils4Africa_D3.5_v1.0
Type of deliverable	Report
Dissemination level	Public
Lead partners	ISRIC
Contributing partners	ARC, MetaMeta, FARA, ICRAF, IITA
Authors	Luís M. de Sousa (ISRIC), Ulan Turdukulov (ISRIC), Bas Kempen (ISRIC)
Contributors	Abraham Abhishek (MetaMeta), Meklit Chernet (ICRAF), Oluwole Fatunbi (FARA), Jeroen Huising (IITA), Jorge M. de Jesus (ISRIC), Johan Leenaars (ISRIC), Garry Paterson (ARC), Pieter Pypers (ICRAF), Andrew Sila (ICRAF), Mary Steverink-Mosugu (ISRIC), Elvis Weullow (ICRAF), Leigh Winowiecki (ICRAF)
Due date	31 May 2021
Submission date	31 May 2021

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

Contents

This project has received funding from the European Union Horizon 2020
Research and Innovation Programme under Grant Agreement 862900

1 Abbreviations

- ARC - Agricultural Research Council
- DSM - Digital Soil Mapping
- FAIR - Findable, Accessible, Interoperable and Reusable
- FNSSA - Food and Nutrition Security and Sustainable Agriculture
- ICRAF - International Council for Research in Agroforestry
- IITA - International Institute of Tropical Agriculture
- OGC - Open Geospatial Consortium
- PTF - Pedo-Transfer-Function
- SIS - Soil Information System
- UML - Unified Modelling Language
- WP - Work Package

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

2 Introduction

2.1 Project Overview

The Soils4Africa project aims to provide an open-access Soil Information System (SIS) hosting a set of key indicators of soil quality. These indicators are to based on field data collected from 20 000 sampling sites spread across the African continent, according to a sound methodology for repeated soil monitoring. This soil information system will become part of the knowledge and information system of the EU-Africa Partnership on Food and Nutrition Security and Sustainable Agriculture (FNSSA) and will be hosted by an African organisation with the requisite capacity to manage the system. This system will inform decision making and other activities towards sustainable agricultural intensification in Africa and facilitate future monitoring and evaluation and it will enhance the performance of other land uses.

Activities of the project include:

i define use cases and indicators in consultation with stakeholders;

ii make a functional design of the soil information system;

iii develop detailed procedures and tools for the field activities based on the LUCAS methodology and collect 20 000 soil samples;

iv develop detailed procedures for laboratory work and analyse the collected soil samples at one reference laboratory located in Africa; and

v develop the technical infrastructure for the soil information system and serve the results as open data linked with open EO data.

The Soils4Africa project is composed by seven interlinked work packages. Work package one (WP1) deals with project coordination, communication and dissemination. WP2 is dedicated to stakeholder engagement and the identification of the relevant soil indicators. WP3 covers the development of a soil information system, of which the present report makes part. A continent wide sampling campaign is to be conducted within WP4, that also aims at capacity building on the topic. In WP5 the soil analyses activities are conducted, using the data collected in the sampling campaign. Finally WP6 shall deal with capacity building around the system, powering users to make the best of the information it will provide. In addition, WP7 safeguards ethical aspects of the project, tackling items such as data privacy and environmental protection.

2.2 Goals of this deliverable

This report documents the expectations of users towards the Soils4Africa soil information system with a collection of formal requirements. It is does not describe how the system will ultimately come to be, or how it should be implemented. It rather informs on the activities users expect the system to support, the data it is expected to host and how users intend to interact with it. Some of these requirements might be too ambitious and not feasible with the resources available to the project. Others may yet be too crude and difficult to translate into concrete implementation actions. However, the requirements gathered herein provide a systematic overview of the understanding users and stakeholders have of the system at the onset of the project.

These user requirements provide a guide for the development of the system. In first place by listing the functionalities expected and rendering an early sketch of data flows. Secondly since it ranks requirements by importance, outlining which aspects of the system are essential and should therefore be implemented first. Finally, both to stakeholders and project/task managers, it furnishes an early assessment of risks to the successful realisation of the system.

Beyond being an early guide to system development in WP3, this report is also relevant to WP6. By systematising users and stakeholders expectations, requirements offer a preview of the concrete ways in which users expect to employ the system. In this regard it is therefore a relevant addendum to deliverable D2.1 [Fatunbi and Abhishek, 2020].

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

2.3 Document Structure

This document gathers the requirements specifying the SIS. They are presented in three different ways: (i) through descriptive text; (ii) recurring to graphical diagrams with the Unified Modelling Language (UML), and (iii) in a systematised set of requirements records. The textual description and the graphical diagrams are most useful during the inception phase, simplifying the communication between the development team and project stakeholders. The requirements records play their role later, in the development phase, providing for a quick assessment of system progress.

Section 3 outlines the methodology employed to gather and register the requirements. A textual account the system, as description by project stakeholders and potential system users is provided in Section 4. This section also attempts to synthesise the most relevant processes with activity diagrams. Section 5 presents the domain model, that formalises the information concepts building up the SIS. Basic interaction requirements are laid out in Section 6 following the UML philosophy of Actors and Use Cases. Section 7 ends the textual description by outlining those requirements still needing further refinement before being translated into concrete engineering tasks.

The document ends with the list of requirements records in Appendix A. Each record can be seen as a file or card that describes and details various aspects of a particular requirement. In the text description (Sections 5 to 7), each requirement record is referenced with its identifier in square brackets, e.g. [**R101**]. These records are stored and managed in a small information system, from which the contents of Appendix A are automatically generated.

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

3 Methodology

The requirements for the SIS were gathered and developed through the analysis method proposed by Dick et al. [2017]. This approach was selected considering that project partners and stakeholders were relatively well identified at the start of the project, are geographically dispersed (across two different continents) and have particular expectations about the system. Whereas in normal circumstances this work would have started with requirements development sessions, involving as many stakeholders as possible, the COVID-19 pandemic prevented the realisation of physical project meetings in which such sessions could take place. Thus the requirement gathering process relied primarily on stakeholder surveys and virtual interviews.

A first group of stakeholders and prospective end users of the SIS, was addressed within work package two (WP2). This WP aims to structure an effective engagement of stakeholders for the development and delivery of a usable SIS. In WP2 an engagement process is being conducted in order to facilitate the stakeholder interaction and identify specific needs.

In the first months of the project an online questionnaire was developed within WP3 and sent to more than 560 preselected stakeholders involved in agricultural research, soil science, extension services, academics, fertilizer manufacturers and policy makers. The questionnaire contained 20 questions on key issues, informing on the design and content of the SIS. A total of 184 stakeholders completed the questionnaire. The results of this questionnaire are gathered in deliverable D2.1 of the project [Fatunbi and Abhishek, 2020].

The questionnaire was followed by the identification of key informants with broad knowledge, and/or a leadership role in well- known initiatives or statutory functions, to take part in the development of use cases[1] expressing in more detail the role of a soil information system in various practical activities. To this end a series of interviews was conduct with these key informants.

The use cases developed with key informants include: (i) soil information use in integrated landscape management; (ii) soil data use for a sustainable intensification program in African farming systems; (iii) soil data use in agricultural extension and advisory services; (iv) use of soil information in public land resource conservation; and (iv) use of soil information by fertiliser producers and suppliers. These use case examples characterise practical interventions requiring the development of relevant soil quality indicators, which determine the soil parameters to be included in the SIS. These use cases showed further the prospective interaction between the end user and the system itself.

Another group of stakeholders was addressed in parallel through structured interviews: key project partners that will to be using (or benefiting from) the SIS. These interviews comprised partners involved in work packages 2, 4 and 5, form ARC, IITA, ICRAF, FARA, MetaMeta and ISRIC. These stakeholders not only have specific expectations towards the SIS but also critical knowledge of the practical application of its information. As joint development sessions were not possible, these stakeholders were interviewed various times, in an iterative process. From rough requirements at an earlier phase, more refined and concrete requirements were thus progressively obtained. The interviews (listed in Table 1) went through the following topics:

- processes engaged to collect and derive information,
- life-cycles of the different information assets,
- user interfaces for decision support,
- user roles and access levels,
- integration of the SIS with other systems.

Beyond gathering user expectations towards the system in natural language, this deliverable also performs an early formalisation of requirements into engineering assets. This formalisation is a first step bridging the human understanding of what the SIS is expected to do into a concrete implementation. Accompanying the requirements records is a set of diagrams developed with the UML [OMG, 2005] catering a more abstract view of the requirements put forth. Three types of diagrams are used, structuring requirements into three different perspectives on the system. Activity diagrams serve to systematise the

[1] Note that the term “use case” used in deliverable D2.1 is diverse to the use case concept found in the UML.

Date	Participants	Topic(s)
05-11-2020	FARA; MetaMeta; ISRIC	Stakeholders questionnaires and interviews
10-11-2020	ARC; ISRIC	Laboratory work
03-12-2020	ARC; ICRAF; ISRIC	Laboratory work
08-12-2020	ISRIC	Decision support
15-01-2021	IITA; ISRIC	Sampling campaign
26-01-2021	ISRIC	Data for decision support
04-02-2021	ICRAF; ISRIC	Laboratory work - spectral analyses
05-02-2021	IITA	Sampling campaign
04-03-2021	IITA; ISRIC	Sampling campaign
22-03-2021	FARA; MetaMeta; ISRIC	Decision support (deliverable D2.1)

Table 1: Stakeholder interviews conducted towards the of system requirements.

processes described by stakeholders. These diagrams are usually the most similar to natural language in which stakeholders describe the system. Activity diagrams help identifying information life-cycles and too omissions in the processes described by stakeholders. Class diagrams are used to gather a first architecture of the information to be managed by the system. They are a first step towards the data model eventually implemented by the system. Finally are Use Case diagrams, furnishing a more functional perspective of the system. To note that "use case" has a diverse meaning in information science to that employed in more colloquial discourse (as is the case in deliverable D2.1). In this context, "use case" identifies an action a user can perform on the system, i.e. a form of system use.

It is important to note that by coming at such an early stage in the project, the delivery of this report does not set an end to the requirements analysis process. The requirement analysis will continue during software prototyping, with further refinements of stakeholders needs and system functionalities. New feed back on system requirements will emerge during the implementation phase; as early versions of the SIS become available, the stakeholders' understanding of their own needs and the role of the SIS within the project is likely to evolve. The description of the system requirements is to be kept up to date as part of the documentation.

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

4 Main Processes

4.1 Preamble

This section provides a descriptive account of users' expectation towards the system. It is the closest to the actual user description of the SIS uttered in natural language. Cross-references link to the respective requirements record in Appendix A.

The requirements are sub-divided in three sub-domains, or system facets, reflecting different moments in the project and data use. First comes the **Sampling** campaign, concerned with the collection, validation and storage of field data and meta-data. Secondly is the **Laboratory** analysis corresponding to the processes that derive measurements of soil properties from the data and samples collected in the field. And lastly is the **Decision Support** facet, the interface to end users.

4.2 Sampling

In order for survey tasks to be assigned and later tracked by survey managers, surveyors must be able to register with the system, providing their contact details [**R102**]. The system must be able to identify the surveyor (or surveyors) responsible for a sample. This information is relevant not only to manage the survey but also later during the laboratory analyses. It might be necessary to contact the surveyor to clarify doubts with the sample, or even to repeat it [**R101**]. Survey managers must be able to contact surveyors, therefore their contacts must be stored by the system and be accessible to managers [**R103**]. Surveyors must be associated with a regional or national institution. The survey will be made operational through contract agreements with the institution and not with individual persons. This for the sake of countability, sharing responsibility and having multiple channels of communication [**R104**].

The assignment of surveyors to sampling sites must be managed or assisted by the system. Surveyors must be able to apply through the system for the site(s) they are able to sample [**R105**]. Whether the assignment of sampling sites is based on an auction mechanism or some simpler process, the assignment must always be vetted by a survey manager before sampling can take place [**R116**].

Before surveying a particular site, the surveyor must know a unique site identifier beforehand. This unique identifier is likely to be encoded as a QR code used to label soil sample vessels and other tangible assets associated with the site survey. These labels are later used in the laboratory to identify the provenance of soil samples. The unique site identifier is also associated with the site and profile descriptions stored digitally in the system. Yet to be determined how the system should interact with label printers and label scanners [**R111**]. Each soil sample must be fully identifiable at any time. The sample identifier can be a composite including the site identifier [**R121**].

Surveyors must be provided with a set of documents that aid their work on the field. These can be maps, ortho–photo maps or other to help identify the sampling site. Other document types may be also necessary [**R112**]. The system must provide functionality for survey managers to associate aid documents with the sampling site. Later to be used by the surveyors [**R113**].

Functionality must be present for the surveyors to register and update planned survey dates [**R115**]. Survey managers must be informed by the system on the date(s) a surveyor plans to go on the field. This allows for checks afterwards if no data and/or samples are submitted [**R114**]. The system must use the expected survey date to identify delayed field work. Once a determined time span passes over the expected date the system must warn survey managers of the delay [**R125**].

It is possible that during field work the system is not reachable (e.g. lack of internet coverage). The site meta-data and description must thus be collected in a simple way and later loaded into the system. Vendor services like ODK and Kobo, provide mobile phone applications that have been used with a degree of success in similar projects and are thus strong candidates for this purpose. The system must thus include the functionality to ingest the meta-data collected with these services [**R123**]. The mobile phone applications upload the data collected in the field to a vendor service (ODK, Kobo, etc). The system should be able to retrieve and ingest these forms on a daily basis in order for the survey managers to keep track of the evolution of field work [**R122**]. It is possible that during field work the system is not reachable. In such cases the meta-data must be collected in a simple spreadsheet and later loaded into the system [**R127**].

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

The description and meta-data of a site submitted to the system by a surveyor (or automatically ingested) must be subject to an automated validation procedure. It uses the site meta-data to check for inconsistencies, missing information and eventually remove redundancies. A similar validation procedure was implemented in the OCP project that can provide further insight [**R124**]. Upon retrieval from the vendor server (e.g. ODK, Kobo) the system must immediately run automated validation procedures. If these checks fail survey managers must be warned to act upon eventual problems [**R126**].

During the sampling campaign two types of sites are considered: standard and reference. Standard sites total 20 000, for which soil samples are collected at fixed depths and the site description is recorded. Reference sites will be 250, for which the same data and samples will be collected as for normal sites. In addition, a detailed profile description shall also be recorded for reference sites, plus further properties yet to be determined [**R137**]. Detailed soil profile descriptions will be recorded for the reference sites by the surveyor. This includes horizon characterisation plus the observations of a set of properties, also yet to be determined [**R134**]. For the reference sites certain characteristics of the site must be collected too. These include soil depth, soil colour, stoneniness, erosion features (e.g. drainage) and more. A concrete list of these site characteristics must be devised [**R131**]. Beyond site description and meta-data, surveyors will also collect photographs of the site and possibly of the soil itself. The system must be able to store these photographs properly and associate them correctly with the respective site [**R132**].

A collection of meta-data must be recorded for each site surveyed. It includes the identification of the surveyor, land cover, land degradation and information cues to assist the validation of site descriptions. An overlap between site meta-data and site description may exist, that must be carefully considered. It is also necessary to ascertain the nature of the information cues required for validation [**R133**].

Sampling sites undergo at least four different states: 1 - Unassigned: waiting for a surveyor to be selected from those that applied. At this stage a unique site identifier is already known and stored in the system. 2 - Waiting survey: a surveyor has been assigned to the site, but site description and soil samples are yet to be collected. 3 - Waiting validation: the surveyor has submitted the site description and meta-data (and profile description for reference sites), but it has not yet been validated. 4 - Validated: the site information surveyed was successfully validated. To clarify which entities must be validated for the site to switch to the Validated state: meta-data, description, photographs, or a combination? Also yet to be determined into which state the site goes back to in case validation fails [**R135**].

Sample meta-data is not accepted "as is" once collected in the field. It must undergo a quality control procedure before samples can be analysed in the laboratory. Moreover, the samples undergo a preparation process before being shipped to the laboratory. This "sample status" must be recorded with the meta-data. The system must support this life-cycle and associated user roles [**R136**]. Laboratory staff need access to the sample meta-data and identification to correctly match samples to analyses [**R152**].

Upon the completion of a sampling assignment the surveyor is paid. The system should keep track of which surveyors have been paid or not [**R141**]. It would be convenient if the system itself could process the payments to surveyors automatically. This requirement is understood as ambitious, considering the different electronic payments systems in existence across the continent [**R142**].

Figure 1 synthesises the overall Sampling process with an activity diagram.

4.3 Laboratory procedures

The wet chemistry analyses conducted in the laboratory undergo a verification process that ultimately requires approval by a responsible project leader. The system is not expected to support this process, but it must be able to trace back the responsible project leader (and the respective laboratory), for instance to clarify doubts with the results or even repeat the analysis if necessary [**R202**]. Laboratory instruments and processes produce unstructured spreadsheets as outputs. These must be ingested into the system, keeping track of the respective samples [**R201**]. End users of chemical analysis results expect to receive these data in the form of spreadsheets. A result spreadsheet might be the result of a particular query. This requirement might need to be split in two: i) spreadsheet generation ii) data querying [**R203**].

The laboratory runs analyses with various instruments, each covering a particular segment of the spectrum. This means there will be various different results for the same sample [**R215**]. The spectroscopy analysis produces a discrete spectrum for each sample analysed. A discrete spectrum is a list (or array) of pairs [wavelength, value]. ICRAF currently runs a Django based information system as a spectral

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

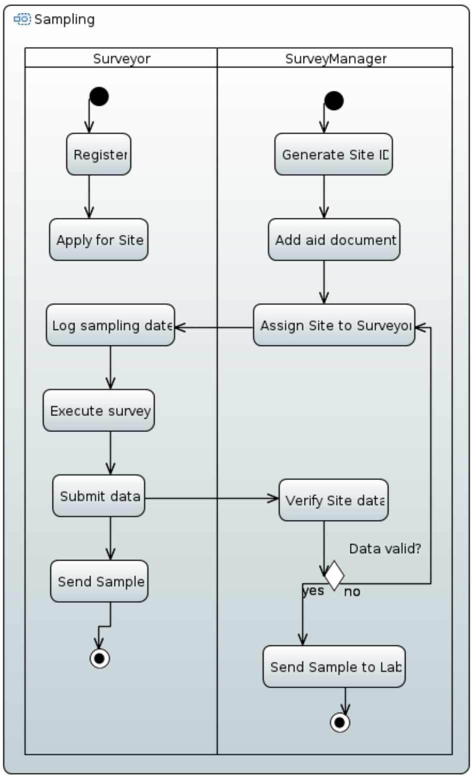

Figure 1: Activity diagram for the Sampling facet of the system.

library. ISRIC has investigated the storage of spectra in relational databases [**R221**]. Spectroscopy instruments produce raw files (aka machine files) with the spectra read from the samples. These files are relevant assets in themselves that need to be stored (or at least indexed) alongside the processed spectra themselves [**R211**]. Users must be able to register spectral analysis instruments with the system, so they can later associate those records with individual spectra [**R223**]. Associated to each spectral analysis result a set of characteristics describing the instrument used in the analysis must be recorded [**R214**].

Once stored in the system spectra must undergo a verification procedure. Within this procedure a manual overview is carried out by one or more users [**R213**]. Spectra loaded from raw files may be inconsistent. This could be due to the spectral analysis procedure itself. The system must run automatic checks on the spectra to identify these possible inconsistencies. ICRAF currently has an automatic verification process in place within their Django based spectral library system. If the intellectual property rights of ICRAF's system so allows, it could be used as a template to implement a similar procedure within the Soils4Africa system [**R212**].

Spectra go through various stages after being loaded into the system. They undergo at least two verification steps: a manual overview and an automatic consistency check. This can ultimately lead to the rejection of a spectrum and the repetition of analysis. There are thus at least four life stages: (i) loaded, but unverified; (ii) verified only automatically, (iii) verified automatically and manually, (iv) rejected. Transition between stages can be dependent on user role/profile. It is possible the system run by ICRAF already implements this life cycle [**R218**].

Soil properties are derived from spectra by prediction models. The system must be able to store these derived data, keeping track of the respective model and related spectrum/spectra [**R216**]. These models are relevant in themselves and must be in some way stored or indexed by the system. Most likely models will be not be stored in a relational database. ICRAF maintains spectroscopy models of this nature, in a procedure that provides cues on how they can be managed within the Soils4Africa project. ICRAF develops its models in the R language and curates them in a code versioning system. However, no code marking mechanisms are used (e.g. tags, branches), models are just stamped with a date to identify a particular

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

state of the repository. The ICRAF models are applicable worldwide, therefore the date is effectively a sufficient marker. In the Soils4Africa project this procedure will likely need to be more refined, especially if the system needs to manage models developed in different laboratories [**R217**].

Meta-data is necessary to trace the origin of each spectrum plus associated instruments and lab procedures. An inventory of the necessary meta-data must be gathered. ICRAF already gathers such meta-data in their Django based system [**R222**].

End users are primarily interested in the properties or characteristics derived from spectra than on the spectra themselves. Currently, the system used at ICRAF provides these derived data in the form of spreadsheets [**R231**]. Survey managers are interested in reviewing the results of laboratory analysis conducted on the samples. This could help identify issues with the survey or meta-data inconsistencies [**R232**].

Figure 2 synthesises the overall Laboratory process with an activity diagram.

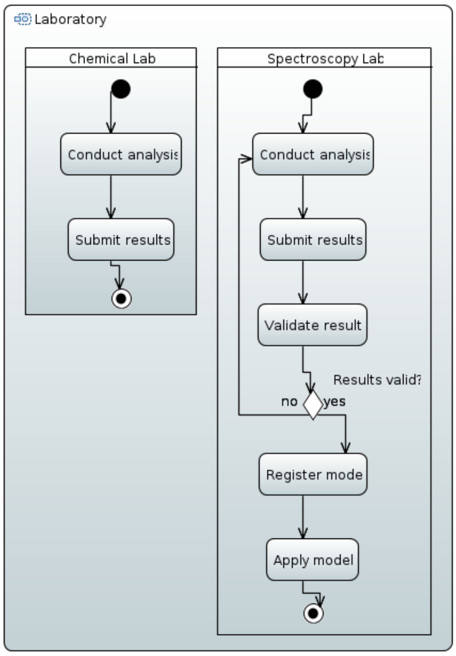

Figure 2: Activity diagram for the Laboratory facet of the system.

4.4 Decision Support

The results of laboratory and spectroscopy analyses must be available to users. This includes physical, chemical and biological properties plus the descriptions of analyses methods [**R302**]. Users expect to access maps of soil properties through the system. These are the properties assessed in laboratory for a sub-set of sampling sites and through spectroscopy analyses for all the sites [**R303**]. Soil descriptions collected in the field must be accessible to users through the system [**R304**]. Beyond soil properties, users are interested in knowing soil quality parameters. These parameters can be obtained from base soil properties through pedo-transfer-functions (PTFs). To clarify whether or not these are part of the indicators managed outside the system [**R305**].

Users expect to access published books through the system [**R306**]. Maps of soil classes must be accessible to users through the system. These are required both in raster and polygon formats [**R311**].

Maps of soil indicators are one class of outputs from the project. These maps will be created through a machine learning DSM procedure that is fully external to the system. However, the system must be able to store these maps and also serve them [**R312**]. The system must provide the data necessary for the DSM process producing the soil indicator maps. These data are primarily the soil properties obtained from laboratory analyses [**R301**].

European Commission | Horizon 2020 European Union funding for Research & Innovation
This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

Users have a precise expectation on the resolution of the maps provided by the system. A scale of 1:250 000 is required, which using Waldo Tobbler's rule translates into a raster cell side of 125 meter [**R314**]. Maps produced at a scale of 1:25 000 are also required, to allow planning and decision support at the scope of villages. This translates into a raster cell side of 12.5 meters [**R315**].

The data stored in the system must be downloadable in a format characterised as "basic" by stakeholders. Users are thus interested in working directly with the data in their own environments. While it has not been possible to determine exactly the meaning of "basic", open standards, like those issued by the OGC, provide for ease of access from open source software and are possibly the best candidates. An example is the GeoPackage standard, but there could be others [**R321**].

The system must also serve other (continental) environmental data layers relevant for sustainable agriculture in Africa. Some of these layers were already obtained by ISRIC and can be served directly (agro-ecological zones, farming systems, WRB soil class map from the Soil Atlas of Africa, land cover layer, etc.). Besides these relevant layers served by third parties must also be considered (see also Task 2.4 in the proposal) [**R322**].

The system is expected to be interoperable with existing soil information systems. So far no requirements have been set out for the inclusion of external data sources, so this requirement is considered at this stage to exclusively concern data publication. To this end the web services specified by the OGC should be especially suited [**R323**].

Data stored in the system should be made available through the Google Earth Engine platform. Some users are expected to perform analyses on that platform. This requirement might raise issues regarding intellectual property that should be investigated [**R324**].

The most important party is the H2020 AfriCultures project[2], this project is explicitly mentioned in the project proposal for this purpose. The project is already running for some years and it has developed various data services related to agriculture and food security that are served with OGC web services. Relevant data from this project must also be accessible via the system. The 'what' and 'how' needs to be further explored with the project stakeholders [**R325**].

In page 2 of the Grant Agreement it is stated: "All data will be available as open data (and according to FAIR (Findable, Accessible, Interoperable and Reusable) principles through user-friendly web services allowing data to be reviewed and downloaded." Especially in relation to the "reusable" aspects, it implies data to be harmonised and standardised (for current and future reuse) [**R326**]. In order to meet the requirement for FAIR data the system must make use of standardised codelists when serving the data collected during the sampling campaign. GloSIS is a natural candidate to be the source of these codelists, but a concrete choice it is yet to be established. Codelists of soil properties and other characteristics should be well established before the sampling campaign starts, so that surveyors already use the same codelists later supporting the data by the system [**R327**].

Beyond download, users expect to explore the data stored in the system within a graphical interface. This facet of the system is also referred by users as a "database analytics page". This is yet a crude requirement, but it might be possible to meet with available data visualisation technologies for the web. Notwithstanding, further refinement of this requirement into more specific functionalities will be necessary at a later stage [**R331**]. The graphics provided by the system must be digital and machine readable. This kind of graphics is also referred as "smart graphics" [**R332**].

Users should be able to access an analytics function provided by the system. This is a key aspect to meet the needs of stakeholders in terms of decision support. No further details are yet known on the mechanics of this function. Further refinement is necessary, involving not only end users but also soil scientists [**R333**]. The system must provide to users algorithms to predict the evolution of the status of the soil against the specification of land use patterns and land management regimes. Yet to be clarified whether the algorithms in question must be defined within the project or if they can be sourced from existing literature [**R334**].

The system should provide guidelines for the interpretation of soil data in an easily available form [**R335**].

The data stored in the system must also be accessible via a mobile phone application. This is yet the only requirement regarding the form factor of the user interface [**R341**].

The system must provide means for users to record their feedback on the workings of the system

[2] `http://africultures.eu/`

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

itself. This feedback will be used to conduct the improvement of the system and also track user satisfaction [**R342**].

A thorough process composed by a sequence of activities can not at this stage be identified in the Decision Support facet of the system. Users and stakeholders have essentially described specific functionalities expected from the SIS. Therefore no activity diagram is presented in this section. Section 6 provides a more formal perspective on decision support functions.

5 Domain Model

5.1 Preamble

This section sketches a first domain model for the system envisioned by the users. It collects the most relevant information entities and the relationships between them. Some of these information entities are still very crude, requiring substantial work towards refinement. At this early stage they provide stakeholders with an overview of the data the system will eventually store, thus allowing the identification of gaps and inconsistencies.

Notwithstanding, this draft domain model provides an early template for the data model to be implemented by the system. Most of the information entities identified at this stage should be recognisable in some form in the final realisation of the system.

5.2 Sampling

In the facet of the system directed at the Sampling campaigns the `Surveyor` is a key information entity. The identity and contacts of these persons are required throughout the campaign, in the assignment of sites, the record of field data and in the management of sample statuses. Surveyors may need to be contacted to clarify issues regarding field data and even to repeat the sampling if necessary. Even the `Surveyor` entity is entirely managed in Sampling facet, its information may also be relevant in the laboratories. Surveyors may grouped within an `Institution`, which will broadly correspond to a geographic area within which it has some authority.

The central information entity in the Sampling facet is the `Site`. It is simply a geographic location, but connects together most of the other entities. During the preparatory phase there is the `AidMaterial` entity, documents like maps, aerial imagery or other that survey managers associate with a site to guide field work.

During field work surveyors collect data on various information entities. The `SiteMeta` entity represents the meta-data referring to the sampling site itself (e.g. date of recording). `SiteDescription` gathers the soil relevant data, a set of pre-determined properties with controlled thesauri characterising the site (mostly related with the surface and involving terrain). The entity `Photograph` represents the pictures taken in the field by surveyors to support the site description. These deserve an explicit entity in the model their different nature as documents (*vis à vis* the textual description).

A special case of `Site` is the `ReferenceSite`, for which a `ProfileDescription` is associated. The soil profile is also described through a set of pre-defined properties, but these more focused on the soil itself. A profile description may be segmented in soil horizons or depth intervals, although this aspect is still subject to consolidation.

Finally is the `Sample` entity itself, representing the actual soil samples collected in the field. Each individual sample will be assign a unique identifier (possibly including a site identifier) that will be used to tag and track soil vessels up to their analysis in the laboratory. `SampleMeta` encapsulates the meta-data recorded for each sample by the surveyor. The actual contents of this entity are yet to be defined.

Figure 3 presents these information entities in a class diagram with relevant associations.

5.3 Laboratory

The `Sample` entity has again a central role in the facet of the system supporting work in the laboratory. Physical soil samples collected in the field are stored in labelled vessels that are sent to different laboratories to undergo chemical and spectroscopy analysis. The laboratory staff must have access to the sample identifiers and also the meta-data created by surveyors to complete their work.

The chemistry laboratories produce for each sample a document captured in the `WetChemistryResult`. In broad terms, this document matches numerical estimates assayed in the laboratory to one or more instances of `SoilProperty`. The analyses conducted on the first field campaign (within the Soils4Africa project time-frame) are to be lead by the Agricultural Research Council (ARC) in South Africa. The ARC already has data structures in places to store wet chemistry results, that can serve as guide to further refine the `WetChemistryResult` information entity if necessary. The `ResponsiblePerson` entity represents the

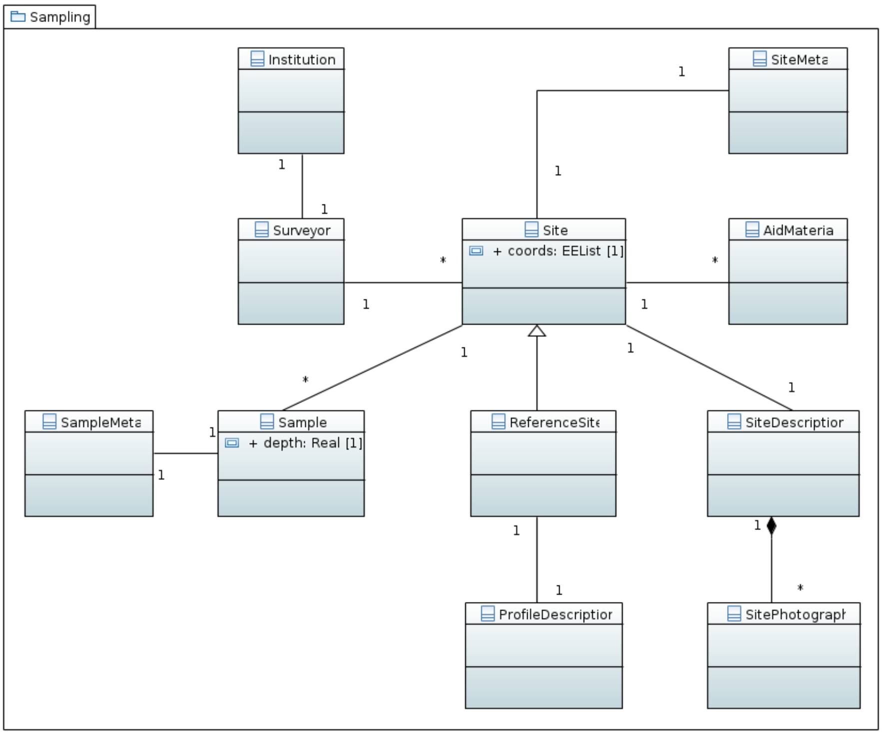

Figure 3: Sketch domain model for the Sampling facet of the system.

need to identify an individual laboratory staff member as responsible for a particular wet chemistry result document. The detail necessary to characterise the `ResponsiblePerson` entity is also yet to be determined.

Beyond wet chemistry analysis, soil samples are also subject to spectral analyses. The result of this process is a `Spectrum`, in essence an array of wavelength and reflectance pairs. As with the wet chemistry results, a spectrum is submitted to the system in the form of a document. For the same sample various spectra will be obtained, each with a different `Instrument`. Characteristics of the instrument are relevant information that must be stored by the system. Furthermore, a `SpectralModel` can be applied to a spectrum to obtain a `DerivedResult`, an estimate of one or more soil properties. These models are computer programmes curated in a code versioning system, they are not stored in the system itself, but relevant versions must be identifiable. Within the Soils4Africa project spectroscopy analyses are to be lead by International Centre for Research in Agroforestry (ICRAF) in Kenya. As with ARC, ICRAF also has an information system in place to store soil spectra that can serve as guide for further refinement of spectral analysis entities.

Spectra can be collected into a `SpectralLibrary` for ease of reference. More details on the contents and purpose of this entity must be gathered. Figure 4 relates the different information entities in the Laboratory facet of the system.

5.4 Decision Support

In the Decision Support facet of the system there are few information entities beyond those already identify in the other facets. This reflects the nature of decision support, more reliant on functionality than on the creation of new data. Figure 5 presents the entities identified in this facet of the system and their relationships.

Users are interested in accessing a type of asset that is not considered in other facets of the SISs: the `Map`. They are created outside the system, but system must provide the means for access. There two distinctive types of maps: the `IndicatorMap` and the `ThirdPartyMap`. The former represents the maps of soil indicators that will be created within the Soils4Africa project using the information gathered in the

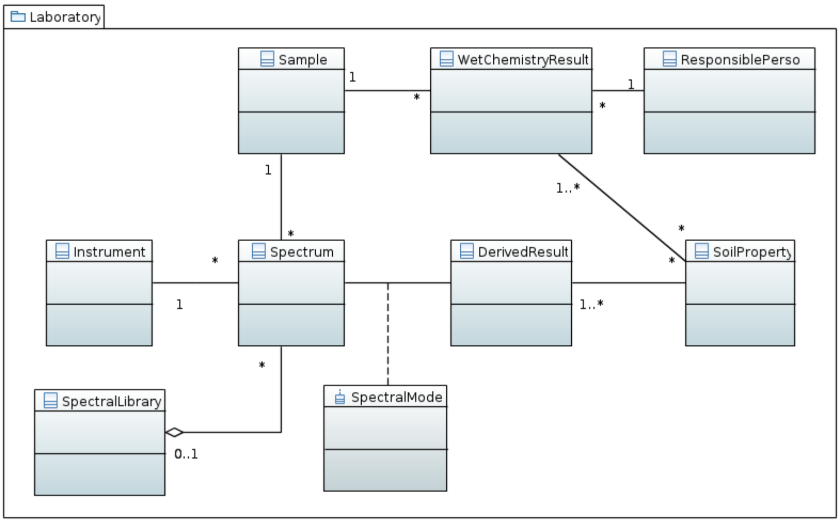

Figure 4: Sketch domain model for the Laboratory facet of the system.

field and in laboratories. The latter captures maps created by third parties that users add to the system. Finally is the `Guidelines` entity, enclosing information on how the maps provided by the system can by interpreted by end users.

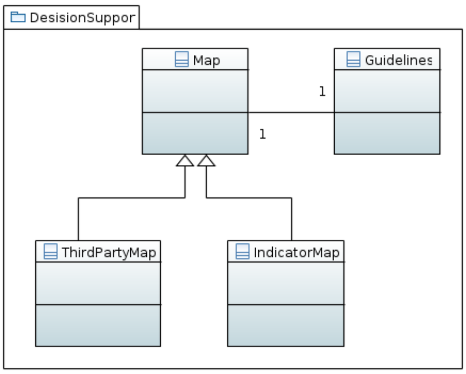

Figure 5: Sketch domain model for the Decision Support facet of the system.

This project has received funding from the European Union Horizon 2020
Research and Innovation Programme under Grant Agreement 862900

6 User Profiles and Use Cases

6.1 Preamble

This section provides yet another perspective on the SIS, this time from a functional viewpoint. Here the UML Actors and Use Cases diagram is used to formalise specific functionalities the system is expected to perform (use cases in UML). Each of these functionalities is associated with a specified user role (or user profile, called actors in UML). A rough outline of access permissions is thus too identified.

6.2 Sampling

The Sampling facet of the system is perhaps the most profuse in terms of functionality and actions performed by users. Two user profiles are identified: the `Surveyor` and the `Survey Manager`. The former stands for the staff conducting field work, displacing themselves to sampling sites to collect soil samples, profile descriptions and meta-data. The latter are staffers responsible for organising the sampling campaign and guarantee data quality. Figure 6 synthesises the use cases available to these two user profiles.

The `Surveyor` must in first place register with the system, providing the system with contact details and other information relevant for the campaign. This user profile also applies for the sites that they are willing or able to sample. After one or more sites have been assigned the `Surveyor` then register an expect date with the system for the site visit. Finally the `Surveyor` submits the data and meta-data collected for the site and respective samples (termed "sampling records" in Figure 6 for simplicity). This use case might in reality take place in a more automated way, but it is included here due to its relevance.

The `Survey Manager` corresponds to a profile that interacts with the "back office" of the system, meaning that these users have access to system functionalities that are not available to general users (e.g. surveyors). This user profile can add aid materials related to a site to facilitate field work. It is also responsible for assigning sampling sites to surveyors, according to the applications by the latter. A survey manager must also be able to monitor the status of a site, particularly during field work. To this end the recorded site visit date is crucial, in order to detect delays in field work. Once the sampling data has been submitted, the system performs various consistency checks, but the survey manager must validate those data explicitly. Finally the survey manager records payments to surveyors upon a successful sampling task.

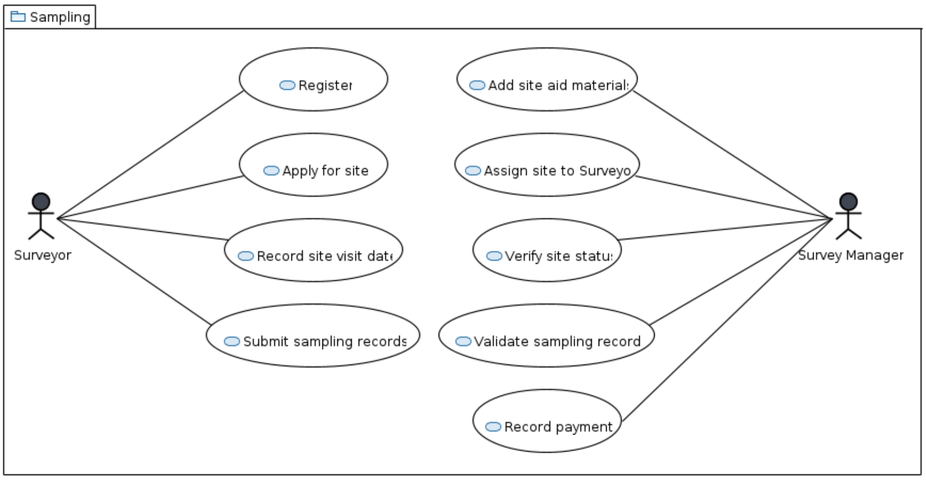

Figure 6: Actors and use cases diagram for the Sampling facet of the system.

6.3 Laboratory

In the Laboratory facet there are fewer use cases, all accessible to the `Laboratory Technician` profile. This essentially corresponds to the users that are able to submit the results of laboratory activities to the system. Figure 7 summarises uses cases in this facet, also portraying the role of `Laboratory Supervisor`. At this stage the supervisor role does not have access to specific use cases, thus it might not actually feature in the final system. However, considering its identification must be recorded in relation to wet chemistry results it could acquire a more specific role in the future.

The `Laboratory Technician` can in first place submit results of wet chemistry or spectroscopy analysis to the system, expectedly the most frequent activities in this facet. Associated with the spectroscopy results comes the management of instruments and spectral models. Results derived from spectra with spectral models can also be submitted to the system by the laboratory staff. Finally, laboratory technicians will conduct oversight on the spectroscopy results, after automated checks performed by the system.

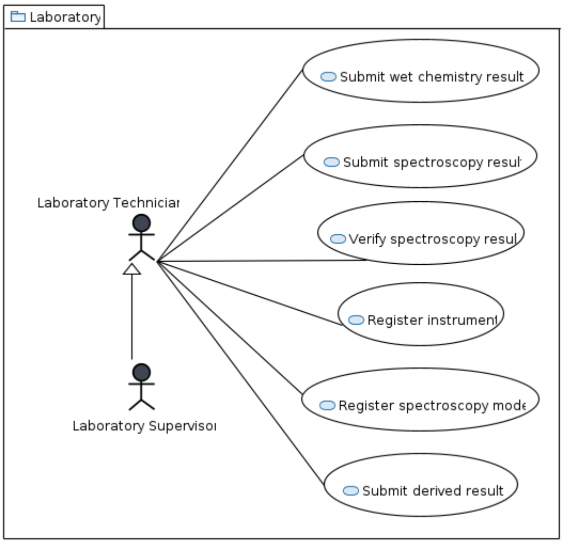

Figure 7: Actors and use cases diagram for the Laboratory facet of the system.

6.4 Decision Support

Figure 8 sketches user profiles and use cases in the Decision Support facet. Most use cases are associated with the `User`, the stakeholders interested in the data generated in the Soils4Africa project and stored by the system. These may be policy makers, academics, researchers, industry professionals, farmers and the general public. An additional profile, `User Manager`, captures the need to manage some of the contents made available to the general public.

Users require in first place functionalities to download the data stored by the system, the data records collected during the sampling campaign, the results of laboratory analysis and the consequent soil indicator maps. The system must also provide guidelines for the interpretation and use of these data. The manager role is responsible for adding and updating these guidelines.

Then come use cases provide more objective decision support. Users expect to consult through the system smart graphics that synthesise data in certain way. A further use case is provided by an analytics function assessing probabilities and trends of soil status evolution in the future. Lastly is the functionality of predictive algorithms, from which users expect to obtain impact estimates of changes in land use or management practices. These use cases are understood by stakeholders as yet poorly defined, in some cases even requiring active research. They must be further refine in the course of project and implemented according to available resources.

This project has received funding from the European Union Horizon 2020
Research and Innovation Programme under Grant Agreement 862900

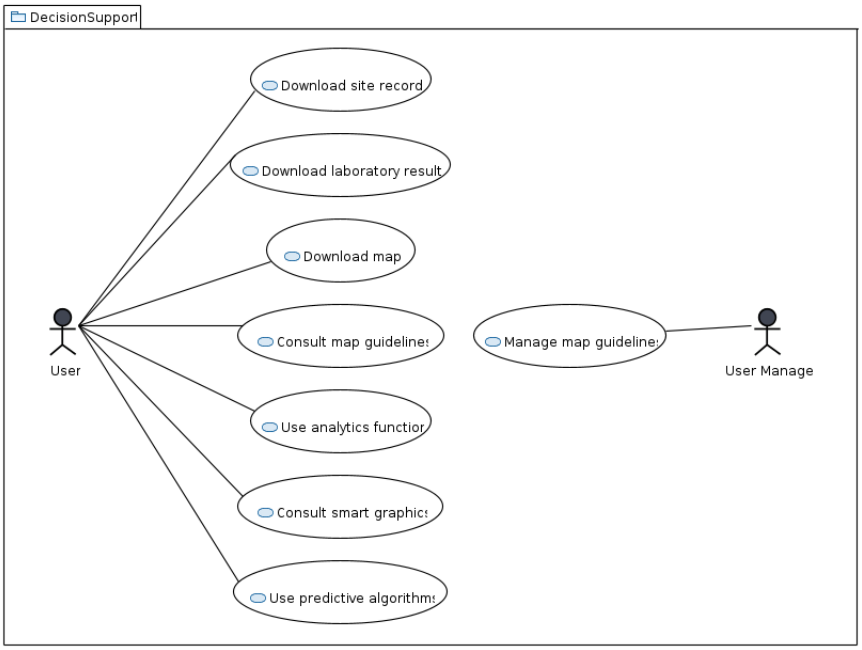

Figure 8: Actors and use cases diagram for the Decision Support facet of the system.

This project has received funding from the European Union Horizon 2020
Research and Innovation Programme under Grant Agreement 862900

7 Further refinements

The end goal of a requirements engineering exercise is to provide a set of verifiable actions that can unambiguously be put to practical effect. This goals is not always fully achieveable, constraints might prevent the complete detailing of requirements, in other circumstances the perception stakeholders and users have of the system may not be entirely clear (particularly at the early stages of a project). Such has been the case in the production of this report. This section reviews the most salient points needed further clarification, where possible indicate ways forwards.

Starting with the Sampling aspect, the interaction with labelling and scanning hardware was not yet determined [**R111**]. The system needs to be the source of site and sample identifiers and these must somehow reach the hardware unequivocally.

Some indefinity remains on the nature of the aid documents associated with each sampling site. It is possible unforeseen document types may be necessary that have a relevant impact on system design [**R112**]. Joint work with survey managers is likely to clarify this question satisfactorily.

Several possible mechanisms have been referred to assign sampling sites to surveyors, but none yet with enough precision [**R116**]. Eventually, this assignment activity can be fully manual, but with 20 000 sites to sample this may not be realistic.

Site description and meta-data need to undergo an automated validation procedure that is not well defined. It will likely depend upon a set of rules [**R124**] plus a set of cues derived from meta-data [**R133**]. These procedure can likely be developed iteratively with survey managers, making use of sound collection of unit tests.

At the time of writing a concrete list of site characteristics to survey are yet to be defined [**R131**]. The soil properties to observe in reference sites are also unknown [**R134**]. Both of these are also likely to change with time, especially since further sampling campaigns can take place in the future. From an implementation point of view, the best approach might be to consider these lists of properties and characteristics as fully dynamic, i.e. managed by the system users themselves. Therefore a new range of requirements is likely to emanate from this issue, focused on user interaction.

There are two important information entities that undergo relevant life cycles: the `SamplingSite` [**R135**] and the `Spectrum` [**R218**]. Both of these life cycles could the described with some detail, but not fully and thus require further elaboration. To this end, the spectral library system currently in use by ICRAF can provide important insight. It will be important to keep track of user roles involved in state changes, possibly with more specific requirements emerging.

The results of chemical analyses must be consultable through a query mechanism that is not yet established [**R203**]. As it seems, no similar mechanisms are currently in use by projects partners like ARC or ICRAF. A detailed survey of software in use can shed some light on this requirement, but further stakeholder interviews are likely necessary.

Spectroscopy models used by ICRAF are tracked simply with their commit date [**R217**]. This might need to be expanded, especially if models from more than one institution are involved. The future role of spectroscopy models in a continental scale project like Soils4Africa must be better understood to fully consolidate this requirement.

The exact nature of the "soil parameters" required by end users is yet to be established [**R305**]. It is also necessary to understand the relationship of these parameters with the soil properties derived from sampling and laboratory analyses, and to the soil indicator maps.

During stakeholder interviews it was not possible to determine the kind of file formats users require for data download. These are simply dubbed as "basic" formats in deliverable D2.1 [Fatunbi and Abhishek, 2020] but were not further elaborated. It will be important to identify early on the implications in terms of software dependencies stemming from this requirement.

As most of the data stored by the system will be available online [**R321**], and some of it even deployed to third party platforms [**R324**], a thorough intellectual property policy should be designed. Legal frameworks in each country must be considered in this matter, authorisations or intellectual property agreements might be required with land owners.

The requirement for compliance with FAIR principles [**R326**] has deep implications that must be considered from the onset of the project. Soil information systems in general do not comply with such principles. Primarily for the lack of a widely accepted and established soil data ontology, but also due

European Commission | Horizon 2020 European Union funding for Research & Innovation
This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

to difficulties in implementing interoperability effectively. This requirement must be refined in further steps that will need to be conducted at each phase of the project, particularly considering the "re-usable" principle in FAIR. Related to FAIR principles is the selection of code-lists supporting the data and metadata collection by surveyors [**R327**]. These code-lists must be identified and select from an existing or upcoming soil information ontology.

The relationship with the H2020 AfricaCultures project needs to be detailed further. Some level of interoperability with the data outputs of the project is necessary. To determine which of those data will be used by the system and how, i.e. loading and serving mechanisms.

The definition of the "database analytics page" [**R331**] could not be developed past a elusory description. The development of this facet of the system will likely require an iterative process of prototyping, during which end users provide feedback on successive early versions of this web page.

Also not yet fully defined is the analytics function [**R333**], deemed a core functionality of the system. In this case an iterative prototyping process might not be most suitable, as the scientific background is not yet established. Meeting requirement **R333** possibly demands in first place a feasibility assessment from a soil science view point and only later further refinement with end users.

The prediction of soil status evolution [**R334**] is yet another functionality likely requiring background work in soil science before moving into an engineering phase. A critical aspect is the nature of the predictive algorithms themselves. It might be possible to source them from literature, but it could also be necessary (or preferable) to develop bespoke algorithms for the system. Together, requirements **R333** and **R334** are those demanding more attention to be fully developed into actionable engineering tasks and should thus be comprehensively approached.

References

J. Dick, E. Hull, and K. Jackson. *Requirements Engineering*, chapter Engineering in the Problem Domain, pages 109–130. Springer-Verlag, 2017.

O. A. Fatunbi and A. Abhishek. D2.1: A set of use cases plus supporting soil quality indicators, 2020.

OMG. Unified Modelling Language 2.0. Standard, Object Management Group, Milford, USA, July 2005.

A Requirements records

This appendix sorts out the requirements outlined in the text into a congruent list of records. These records are composed by a number of attributes described below.

- Name - A one sentence description of the requirement.
- Priority - the reletive degree of importance of the require ment to the overall success of the project, following the MoSCoW scale:
  - Must,
  - Should,
  - Could,
  - Would.
- Type - the type of requirement, classified in:
  - Functional,
  - Performance,
  - Quality factor,
  - Interface
  - Constraint.
- Subject - the subject from which the requirement stems:
  - Domain,
  - User profiles,
  - Processes.
- Life phase - the phase of the project at which the requirement is expected to be ready:
  - Concept,
  - Development,
  - Integration,
  - Deployment.
- Risk level - the risk of failling to meet the requirement, expressed in three level:
  - High,
  - Medium,
  - Low.
- Rationale - a textual justification for the requirement (when necessary).
- Verifiable - a boolean indicating whether the requirement is verifiable or not.
- Verification - a description outlining how to verify the requirement.

R101 : Surveyors must be identifiable	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	Technical documents ARC
Rationale:	The system must be able to identify the surveyor (or surveyors) responsible for a sample. This information is relevant not only to manage the survey but also later during the laboratory analyses. It might be necessary to contact the surveyor to clarify doubts with the sample, or even to repeat it.
Verifiable:	True
Verification:	Data must be accessible through the user interface

R102 : Surveyors must be able to register with the system	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 05-02-2021
Rationale:	In order for survey tasks to be assigned and later tracked by survey managers, surveyors must be able to register with the system, providing their contact details.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R103 : Surveyors contacts must be known to the system	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 05-02-2021
Rationale:	Survey managers must be able to contact surveyors, therefore their contacts must be stored by the system and be accessible to managers.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

R104 : Surveyors must be associated with an institution	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 15-01-2021
Rationale:	Surveyors must be associated with a regional or national institution. The survey will be made operational through contract agreements with the institution and not with individual persons. This for the sake of countability, sharing responsibility and having multiple channels of communication
Verifiable:	True
Verification:	The institution can be consulted in the system.

R105 : Surveyors must be able to apply for sampling sites	
Priority:	Should
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 15-01-2021
Rationale:	The assignment of surveyors to sampling sites must be managed or assisted by the system. Surveyors must be able to apply through the system for the site(s) they are able to sample.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R111 : Unique site identifiers must be provided by the system to surveyors	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 15-01-2021
Rationale:	Before surveying a particular site, the surveyor must know a unique site identifier beforehand. This unique identifier is likely to be encoded as a QR code used to label soil sample vessels and other tangible assets associated with the site survey. These labels are later used in the laboratory to identify the provenance of soil samples. The unique site identifier is also associated with the site and profile descriptions stored digitally in the system. Yet to be determined how the system should interact with label printers and label scanners.
Verifiable:	True
Verification:	Functionality can be tested

R112 : Aid documents must be made available to surveyors	
Priority:	Should
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	High
Source:	T3.7 meeting 15-01-2021
Rationale:	Surveyors must be provided with a set of documents that aid their work on the field. These can be maps, ortho–photo maps or other to help identify the sampling site. Other document types may also be necessary.
Verifiable:	True
Verification:	Aid assets for each site must be available to surveyors

R113 : Survey managers must be able to add aid documents	
Priority:	Should
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 15-01-2021
Rationale:	The system must provide functionality for survey managers to associate aid documents with the sampling site. Later to be used by the surveyors.
Verifiable:	True
Verification:	Functionality can be tested

R114 : Survey managers must know the expected dates of field work for each site	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 05-02-2021
Rationale:	Survey managers must be informed by the system on the date(s) a surveyor plans to go on the field. This allows for checks afterwards if no data and/or samples are submitted.
Verifiable:	True
Verification:	Information must be accessible through the user interface

R115 : Surveyors must register the expected dates of field work for each site it is assigned	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 05-02-2021
Rationale:	Functionality must be present for the surveyors to register and update planned survey dates.
Verifiable:	False
Verification:	System must make this functionality available

R116 : Survey managers must be able to vet the assignment of sampling sites to surveyors.	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 05-02-2021
Rationale:	Whether the assignment of sampling sites is based on an auction mechanism or some simpler process, the assignment must always be vetted by a survey manager before sampling can take place.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R121 : The system must track samples collected in the field	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Deployment
Risk level:	Low
Source:	T3.7 meeting 05-11-2020
Rationale:	Each soil sample must be fully identifiable at any time. The sample identifier can be a composite including the site identifier.
Verifiable:	True
Verification:	Information must be accessible through the user interface

R122 : ODK/Kobo forms should be ingested into the system on daily basis	
Priority:	Should
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 05-02-2021
Rationale:	The mobile phone applications upload the data collected in the field to a vendor service (ODK, Kobo, etc). The system should be able to retrieve and ingest these forms on a daily basis in order for the survey managers to keep track of the evolution of field work.
Verifiable:	True
Verification:	Functionality can be tested

R123 : The system must be able to ingest sample meta-data from ODK forms	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Deployment
Risk level:	Medium
Source:	T3.7 meeting 05-11-2020
Rationale:	It is possible that during field work the system is not reachable (e.g. lack of internet coverage). The site meta-data and description must thus be collected in a simple way and later loaded into the system. Vendor services like ODK and Kobo, provide mobile phone applications that have been used with a degree of success in similar projects and are thus strong candidates for this purpose. The system must thus include the functionality to ingest the meta-data collected with these services.
Verifiable:	True
Verification:	Functionality can be tested

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R124 : Site description must be automatically validated	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	High
Source:	T3.7 meeting 15-01-2021
Rationale:	The description and meta-data of a site submitted to the system by a surveyor (or automatically ingested) must be subject to an automated validation procedure. It uses the site meta-data to check for inconsistencies, missing information and eventually remove redundancies. A similar validation procedure was implemented in the OCP project that can provide further insight.
Verifiable:	True
Verification:	Functionality can be tested

R125 : System must alert survey managers of delayed field work	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 05-02-2021
Rationale:	The system must use the expected survey date to identify delayed field work. Once a determined time span passes over the expected date the system must warn survey managers of the delay.
Verifiable:	True
Verification:	Warning must be present in the user interface in case of delay.

R126 : The system must warn surveyors of inconsistencies with field work data	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 05-02-2021
Rationale:	Upon retrieval from the vendor server (e.g. ODK, Kobo) the system must immediately run automated validation procedures. If these checks fail survey managers must be warned to act upon eventual problems.
Verifiable:	True
Verification:	Functionality can be tested

R127 : The system must be able to ingest sample meta-data from spreadsheets	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Deployment
Risk level:	Low
Source:	T3.7 meeting 05-11-2020
Rationale:	It is possible that during field work the system is not reachable. In such cases the meta-data must be collected in a simple spreadsheet and later loaded into the system.
Verifiable:	True
Verification:	Functionality can be tested

R131 : Site description must be collected for all sites	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 15-01-2021
Rationale:	Detailed profile descriptions are only collected for the reference sites, but certain characteristics of the site must be collected. These include soil depth, soil colour, stoneniness, erosion features (e.g. drainage) and more. A concrete list of these site characteristics must be devised.
Verifiable:	True
Verification:	Functionality to add and retrieve plot or site data must be present.

R132 : Site photographs must be stored by the system	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 15-01-2021
Rationale:	Beyond site description and meta-data, surveyors will also collect photographs of the site and possibly of the soil itself. The system must be able to store these photographs properly and associate them correctly with the respective site.
Verifiable:	True
Verification:	Data must be accessible through the user interface

R133 : Site meta-data must be recorded by the system	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 15-01-2021
Rationale:	A collection of meta-data must be recorded for each site surveyed. It includes the identification of the surveyor, land cover, land degradation and information cues to assist the validation of site descriptions. An overlap between site meta-data and site description may exist, that must be carefully considered. It is also necessary to ascertain the nature of the information cues required for validation.
Verifiable:	True
Verification:	Data must be accessible through the user interface

R134 : System must be able to store detailed soil profile descriptions	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 15-01-2021
Rationale:	Sample sites are split in two kinds: standard (the vast majority) and reference (in the order of hundreds). Detailed soil profile descriptions will be recorded for the reference sites by the surveyor. This includes horizon characterisation plus the observations of a set of properties yet to be determined.
Verifiable:	True
Verification:	Soil profile description must be retrievable from the system

R135 : Sampling sites have a life cycle	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 15-01-2021
Rationale:	Sampling sites undergo at least four different states: 1 - Unassigned: waiting for a surveyor to be selected from those that applied. At this stage a unique site identifier is already known and stored in the system. 2 - Waiting survey: a surveyor has been assigned to the site, but site description and soil samples are yet to be collected. 3 - Waiting validation: the surveyor has submitted the site description and meta-data (and profile description for reference sites), but it has not yet been validated. 4 - Validated: the site information surveyed was successfully validated. To clarify which entities must be validated for the site to switch to the Validated state: meta-data, description, photographs, or a combination? Also yet to be determined into which state the site goes back to in case validation fails.
Verifiable:	True
Verification:	It must be possible to identify (and modify) the status of a sampling site through the user interface.

R136 : Sample meta-data have a life-cycle	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Development
Risk level:	Medium
Source:	T6.1 meeting 14-01-2021
Rationale:	Sample meta-data is not accepted "as is" once collected in the field. It must undergo a quality control procedure before samples can be analysed in the laboratory. Moreover, the samples undergo a preparation process before being shipped to the laboratory. This "sample status" must be recorded with the meta-data. The system must support this life-cycle and associated user roles.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

R137 : Two types of sites must be identifiable: standard and reference	
Priority:	Must
Type:	Constraint
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	T3.7 meeting 15-01-2021
Rationale:	During the sampling campaign two types of sites are considered: standard and reference. Standard sites total 20 000, for which soil samples are collected at fixed depths and the site description is recorded. Reference sites will be 250, for which the same data and samples will be collected as for normal sites. In addition, a detailed profile description shall also be recorded for reference sites, plus further properties yet to be determined.
Verifiable:	True
Verification:	System must be able to distinguish between the two types of sites.

R141 : Payments to surveyors must be registered by the system	
Priority:	Should
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 05-02-2021
Rationale:	Upon the completion of a sampling assignment the surveyor is paid. The system should keep track of which surveyors have been paid or not.
Verifiable:	True
Verification:	Information must be accessible through the user interface

European Commission | Horizon 2020 European Union funding for Research & Innovation

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R142 : The system should be able to execute payments to surveyors	
Priority:	Would
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	High
Source:	T3.7 meeting 05-02-2021
Rationale:	It would be convenient if the system itself could process the payments to surveyors automatically. This requirement is understood as ambitious, considering the different electronic payments systems in existence across the continent.
Verifiable:	True
Verification:	Functionality can be tested

R152 : Sample meta-data must be available to laboratory staff	
Priority:	Should
Type:	Interface
Subject:	Sampling
Life phase:	Deployment
Risk level:	Medium
Source:	T3.7 meeting 10-11-2020
Rationale:	Laboratory staff need access to the sample meta-data and identification to correctly match samples to analyses.
Verifiable:	True
Verification:	Some kind of user interface for laboratory staff must be available

R201 : Outputs from wet chemistry analyses must be ingested into the system	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Medium
Source:	T3.7 meeting 10-11-2020
Rationale:	Laboratory instruments and processes produce unstructured spreadsheets as outputs. These must be ingested into the system, keeping track of the respective samples.
Verifiable:	True
Verification:	Functionality can be tested

R202 : Responsible for Laboratory analyses must be known to the system	
Priority:	Should
Type:	Functional
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Medium
Source:	T3.7 meeting 03-12-2020
Rationale:	The wet chemistry analyses conducted in the laboratory undergo a verification process that ultimately requires approval by the responsible project leader. The system is not expected to support this process, but it must be able to trace back the project leader (and the respective laboratory), for instance to clarify doubts with the results or even repeat the analysis if necessary.
Verifiable:	True
Verification:	Laboratory project leader identifiable in each analisys result.

R203 : Chemical analysis results must be made available as spreasheets	
Priority:	Should
Type:	Interface
Subject:	Laboratory
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 10-11-2020
Rationale:	End users of chemical analysis results expect to receive these data in the form of spreadsheets. A result spreadsheet might be the result of a particular query. This requirement might need to be split in two: i) spreadsheet generation ii) data querying.
Verifiable:	True
Verification:	Functionality can be tested

R211 : Raw outputs from spectroscopy instruments must be stored (or tracked) by the system.	
Priority:	Should
Type:	Functional
Subject:	Laboratory
Life phase:	Integration
Risk level:	Low
Source:	T3.7 meeting 03-12-2020
Rationale:	Spectroscopy instruments produce raw files (aka machine files) with the spectra read from the samples. These files are relevant assets in themselves that need to be stored (or at least indexed) alongside the processed spectra themselves.
Verifiable:	True
Verification:	Raw spectra file must be accessible from the system.

This project has received funding from the European Union Horizon 2020
Research and Innovation Programme under Grant Agreement 862900

R212 : Spectra must be automatically verified for inconsistencies once loaded into the system	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Medium
Source:	T3.7 meeting 03-12-2020
Rationale:	Spectra loaded from raw files may be inconsistent. This could be due to the spectral analysis procedure itself. The system must run automatic checks on the spectra to identify these possible inconsistencies. ICRAF currently has an automatic verification process in place within their Django based spectral library system. If the intellectual property rights of ICRAF's system so allows, it could be used as a template to implement a similar procedure within the Soils4Africa system.
Verifiable:	True
Verification:	Functionality can be tested

R213 : Spectra must be overseen by users	
Priority:	Must
Type:	Interface
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Low
Source:	T3.7 meeting 03-12-2020
Rationale:	Once stored in the system spectra must undergo a verification procedure. Within this procedure a manual overview is carried out by one or more users.
Verifiable:	True
Verification:	The system interface must allow for user inspection of spectra.

R214 : Spectral instrument characteristics must be known	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Concept
Risk level:	Low
Source:	T6.1 meeting 04-02-2021
Rationale:	Associated to each spectral analysis result a set of characteristics describing the instrument used in the analysis must be recorded.
Verifiable:	True
Verification:	Information must be accessible through the user interface.

R215 : Multiple spectral instruments can be used to analyse a sample	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Concept
Risk level:	Low
Source:	T6.1 meeting 04-02-2021
Rationale:	The laboratory runs analyses with various instruments, each covering a particular segment of the spectrum. This means there will be various different results for the same sample.
Verifiable:	True
Verification:	The user interface should be able to display/retrieve multiple spectroscopy results for the same sample.

R216 : Soil properties derived from spectra must be stored in the system	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Low
Source:	T3.7 meeting 03-12-2020
Rationale:	Soil properties are derived from spectra by prediction models. The system must be able to store these derived data, keeping track of the respective model and related spectrum/spectra.
Verifiable:	True
Verification:	Functionality to load and retrieve derived soil properties must be present in the system

R217 : Spectral prediction models must be stored (or tracked) by the system	
Priority:	Should
Type:	Functional
Subject:	Laboratory
Life phase:	Deployment
Risk level:	High
Source:	T3.7 meeting 03-12-2020 and T6.1 meeting 04-02-2021
Rationale:	Information such as soil properties can be derived from spectra by specific prediction models. These models are relevant in themselves and must be in some way stored or indexed by the system. Most likely models will be not be stored in a relational database. ICRAF maintains spectroscopy models of this nature, in a procedure that provides cues on how they can be managed within the Soils4Africa project. ICRAF develops its models in the R language and curates them in a code versioning system. However, no code marking mechanisms are used (e.g. tags, branches), models are just stamped with a date to identify a particular state of the repository. The ICRAF models are applicable worldwide, therefore the date is effectively a sufficient marker. In the Soils4Africa project this procedure will likely need to be more refined, especially if the system needs to manage models developed in different laboratories.
Verifiable:	True
Verification:	Functionality to load and retrieve prediction models must be present in the system

R218 : Spectra have a life cycle	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Medium
Source:	T3.7 meeting 03-12-2020
Rationale:	Spectra go through various stages after being loaded into the system. They undergo at least two verification steps: a manual overview and an automatic consistency check. This can ultimately lead to the rejection of a spectrum and the repetition of analysis. There are thus at least four life stages: (i) loaded, but unverified; (ii) verified only automatically, (iii) verified automatically and manually, (iv) rejected. Transition between stages can be dependent on user role/profile. It is possible the system run by ICRAF already implements this life cycle.
Verifiable:	True
Verification:	Each spectrum must be associated with a clearly identifiable life state. Funcionality to transition between life states must be present.

R221 : The system must be able to store a spectral library	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Medium
Source:	T3.7 meeting 03-12-2020
Rationale:	The spectroscopy analysis produces a discrete spectrum for each sample analysed. A discrete spectrum is a list (or array) of pairs [wavelength, value]. ICRAF currently runs a Django based information system as a spectral library. ISRIC has investigated the storage of spectra in relational databases.
Verifiable:	True
Verification:	Functionality to create and retrieve spectra must be available in the system.

R222 : Meta-data of spectral library must be stored in the system	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Medium
Source:	T3.7 meeting 03-12-2020
Rationale:	Meta-data is necessary to trace the origin of each spectrum plus associated instruments and lab procedures. An inventory of the necessary meta-data must be gathered. ICRAF already gathers such meta-data in their Django based system.
Verifiable:	True
Verification:	Functionality to load and retrieve spectrum meta-data must be present in the system

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R223 : System must record spectral instrument characteristics	
Priority:	Must
Type:	Functional
Subject:	Laboratory
Life phase:	Concept
Risk level:	Low
Source:	T6.1 meeting 04-02-2021
Rationale:	Users must be able to register spectral analysis instruments with the system, so they can later associate those records with individual spectra.
Verifiable:	True
Verification:	Functionality must be present in the system.

R231 : Soil properties derived from spectra must be available to end users	
Priority:	Must
Type:	Interface
Subject:	Laboratory
Life phase:	Deployment
Risk level:	Low
Source:	T3.7 meeting 03-12-2020
Rationale:	End users are primarily interested in the properties or characteristics derived from spectra than on the spectra themselves. Currently, the system used at ICRAF provides these derived data in the form of spreadsheets.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

R232 : Survey managers must be able to access laboratory analyses results	
Priority:	Must
Type:	Functional
Subject:	Sampling
Life phase:	Concept
Risk level:	Low
Source:	T3.7 meeting 05-02-2021
Rationale:	Survey managers are interested in reviewing the results of laboratory analysis conducted on the samples. This could help identify issues with the survey or meta-data inconsistencies.
Verifiable:	True
Verification:	Data must be accessible through the user interface

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R301 : Provide data for Soil Indicator mapping	
Priority:	Must
Type:	Functional
Subject:	End user interface
Life phase:	Integration
Risk level:	Low
Source:	T6.1 meeting 14-01-2021
Rationale:	The system must provide the data necessary for the DSM process producing the soil indicator maps. These data are primarily the soil properties obtained from laboratory analyses.
Verifiable:	True
Verification:	Soil property data, including profile structure and location, must be accessible to a foreign programme.

R302 : User access to laboratory results	
Priority:	Must
Type:	Functional
Subject:	Decision Support
Life phase:	Concept
Risk level:	Low
Source:	Deliverable D2.1
Rationale:	The results of laboratory and spectroscopy analyses must be available to users. This includes physical, chemical and biological properties plus the descriptions of analyses methods.
Verifiable:	True
Verification:	Data must be accessible through the user interface

R303 : User access to maps of soil properties	
Priority:	Must
Type:	Functional
Subject:	Decision Support
Life phase:	Concept
Risk level:	Low
Source:	Deliverable D2.1
Rationale:	Users expect to access maps of soil properties through the system. These are the properties assessed in laboratory for a sub-set of sampling sites and through spectroscopy analyses for all the sites
Verifiable:	True
Verification:	Data must be accessible through the user interface

R304 : User access to profile descriptions	
Priority:	Must
Type:	Functional
Subject:	Decision Support
Life phase:	Concept
Risk level:	Low
Source:	Deliverable D2.1
Rationale:	Soil descriptions collected in the field must be accessible to users through the system.
Verifiable:	True
Verification:	Data must be accessible through the user interface

R305 : User access to soil quality parameters	
Priority:	Must
Type:	Functional
Subject:	Decision Support
Life phase:	Concept
Risk level:	Low
Source:	Deliverable D2.1
Rationale:	Beyond soil properties, users are interested in knowing soil quality parameters. These parameters can be obtained from base soil properties through pedo-transfer-functions (PTFs). To clarify whether or not these are part of the indicators managed outside the system.
Verifiable:	True
Verification:	Data must be accessible through the user interface

R306 : User access to published books	
Priority:	Should
Type:	Functional
Subject:	Decision Support
Life phase:	Concept
Risk level:	High
Source:	Deliverable D2.1
Rationale:	Users expect to access published books through the system.
Verifiable:	True
Verification:	Books must be accessible through the user interface

R311 : Provide maps of soil classes	
Priority:	Must
Type:	Functional
Subject:	Decision Support
Life phase:	Concept
Risk level:	Low
Source:	Deliverable D2.1
Rationale:	Maps of soil classes must be accessible to users through the system. These are required both in raster and polygon formats.
Verifiable:	True
Verification:	Data must be accessible through the user interface

R312 : Store and serve Soil Indicator maps	
Priority:	Must
Type:	Functional
Subject:	End user interface
Life phase:	Development
Risk level:	Low
Source:	T6.1 meeting 14-01-2021
Rationale:	Maps of soil indicators are one class of outputs from the project. These maps will be created through a machine learning DSM procedure that is fully external to the system. However, the system must be able to store these maps and also serve them.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

R314 : Provide maps at 1:250K scale	
Priority:	Must
Type:	Quality factor
Subject:	Decision Support
Life phase:	Concept
Risk level:	Medium
Source:	Deliverable D2.1
Rationale:	Users have a precise expectation on the resolution of the maps provided by the system. A scale of 1:250 000 is required, which using Waldo Tobbler's rule translates into a raster cell side of 125 meters.
Verifiable:	True
Verification:	Maps of the required resolution must be made accessible

R315 : Provide maps at 1:25K scale	
Priority:	Must
Type:	Quality factor
Subject:	Decision Support
Life phase:	Concept
Risk level:	High
Source:	Deliverable D2.1
Rationale:	Maps produced at a scale of 1:25 000 are also required, to allow planning and decision support at the scope of villages. This translates into a raster cell side of 12.5 meters.
Verifiable:	True
Verification:	Maps of the required resolution must be made accessible

R321 : Data must be available for download in ”basic” format(s)	
Priority:	Must
Type:	Functional
Subject:	Decision Support
Life phase:	Concept
Risk level:	Low
Source:	Deliverable D2.1
Rationale:	The data stored in the system must be downloadable in a format characterised as “basic” by stakeholders. Users are thus interested in working directly with the data in their own environments. While it has not been possible to determine exactly the meaning of “basic”, open standards, like those issued by the OGC, provide for ease of access from open source software and are possibly the best candidates. An example is the GeoPackage standard, but there could be others.
Verifiable:	True
Verification:	Functionality can be tested

R322 : The system must serve environmental data layers	
Priority:	Must
Type:	Interface
Subject:	End user interface
Life phase:	Deployment
Risk level:	Medium
Source:	E-mail Bas Kempen 04-12-2020
Rationale:	The system must also serve other (continental) environmental data layers relevant for sustainable agriculture in Africa. Some of these layers were already obtained by ISRIC and can be served directly (agro-ecological zones, farming systems, WRB soil class map from the Soil Atlas of Africa, land cover layer, etc.). Besides these relevant layers served by third parties must also be considered (see also Task 2.4 in the proposal).
Verifiable:	True
Verification:	Data must be accessible through the user interface

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R323 : System must be interoperable with other systems	
Priority:	Must
Type:	Functional
Subject:	Decision Support
Life phase:	Integration
Risk level:	Low
Source:	Deliverable D2.1
Rationale:	The system is expected to be interoperable with existing soil information systems. So far no requirements have been set out for the inclusion of external data sources, so this requirement is considered at this stage to exclusively concern data publication. To this end the web services specified by the OGC should be especially suited.
Verifiable:	True
Verification:	Functionality can be tested

R324 : Data availability through Google Earth Engine	
Priority:	Should
Type:	Constraint
Subject:	Decision Support
Life phase:	Deployment
Risk level:	Medium
Source:	Deliverable D2.1
Rationale:	Data stored in the system should be made available through the Google Earth Engine platform. Some users are expected to perform analyses on that platform. This requirement might raise issues regarding intellectual property that should be investigated.
Verifiable:	True
Verification:	Data must be accessible in the platform

R325 : The system must ingest spatial data provided by third parties: points, polygons and grids	
Priority:	Must
Type:	Functional
Subject:	End user interface
Life phase:	Integration
Risk level:	Medium
Source:	E-mail Bas Kempen 04-12-2020
Rationale:	The most important party is the H2020 AfriCultures project (http://africultures.eu/), this project is explicitly mentioned in the project proposal for this purpose. The project is already running for some years and it has developed various data services related to agriculture and food security that are served with OGC web services. Relevant data from this project must also be accessible via the system. The 'what' and 'how' needs to be further explored with the project stakeholders.
Verifiable:	True
Verification:	Functionality can be tested

R326 : All data must be published according to FAIR principles	
Priority:	Must
Type:	Quality factor
Subject:	Decision Support
Life phase:	Concept
Risk level:	High
Source:	Grant Agreement
Rationale:	In page 2 of the Grant Agreement it is stated: "All data will be available as open data (and according to FAIR (Findable, Accessible, Interoperable and Reusable) principles through user-friendly web services allowing data to be reviewed and downloaded." Especially in relation to the "reusable" aspects, it implies data to be harmonised and standardised (for current and future reuse).
Verifiable:	True
Verification:	FAIR principles are somewhat open to interpretation, but a framework can be established to evaluate this requirement. Advice from experts may be of help.

R327 : Use codelists to meet FAIR requirements.	
Priority:	Should
Type:	Quality factor
Subject:	Sampling
Life phase:	Concept
Risk level:	Medium
Source:	Grant Agreement
Rationale:	In order to meet the requirement for FAIR data the system must make use of standardised codelists. GloSIS is a natural candidate to provide them, but the exact source is yet to be determined.
Verifiable:	True
Verification:	It is possible to verify whether the system uses standardised codelists or not.

R331 : Data analytics page	
Priority:	Must
Type:	Interface
Subject:	Decision Support
Life phase:	Development
Risk level:	Medium
Source:	Deliverable D2.1
Rationale:	Beyond download, users expect to explore the data stored in the system within a graphical interface. This facet of the system is also referred by users as a "database analytics page". This is yet a crude requirement, but it might be possible to meet with available data visualisation technologies for the web. Notwithstanding, further refinement of this requirement into more specific functionalities will be necessary at a later stage.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

R332 : Provide smart graphics	
Priority:	Should
Type:	Interface
Subject:	Decision Support
Life phase:	Development
Risk level:	Medium
Source:	Deliverable D2.1
Rationale:	The graphics provided by the system must be digital and machine readable. This kind of graphics is also referred as "smart graphics".
Verifiable:	True
Verification:	Data must be accessible through the user interface

R333 : Provide analytics function	
Priority:	Must
Type:	Interface
Subject:	Decision Support
Life phase:	Development
Risk level:	Medium
Source:	Deliverable D2.1
Rationale:	Users should be able to access an analytics function provided by the system. This is a key aspect to meet the needs of stakeholders in terms of decision support. No further details are yet known on the mechanics of this function. Further refinement is necessary, involving not only end users but also soil scientists.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

R334 : The system must be able to predict the evolution of soil statuses	
Priority:	Must
Type:	Interface
Subject:	Decision Support
Life phase:	Development
Risk level:	Medium
Source:	Deliverable D2.1
Rationale:	The system must provide to users algorithms to predict the evolution of the status of the soil against the specification of land use patterns and land management regimes. Yet to be clarified whether the algorithms in question must be defined within the project or if they can be sourced from existing literature.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

This project has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement 862900

R335 : Provide interpretation guidelines	
Priority:	Should
Type:	Interface
Subject:	Decision Support
Life phase:	Concept
Risk level:	Low
Source:	Deliverable D2.1
Rationale:	The system should provide guidelines for the interpretation of soil data in an easily available form.
Verifiable:	True
Verification:	Functionality can be tested in the user interface

R341 : Data access through mobile phone application	
Priority:	Must
Type:	Interface
Subject:	Decision Support
Life phase:	Development
Risk level:	High
Source:	Deliverable D2.1
Rationale:	The data stored in the system must also be accessible via a mobile phone application. This is yet the only requirement regarding the form factor of the user interface.
Verifiable:	True
Verification:	Application can be tested

R342 : Record user feedback on the system	
Priority:	Should
Type:	Interface
Subject:	Decision Support
Life phase:	Concept
Risk level:	Medium
Source:	Deliverable D2.1
Rationale:	The system must provide means for users to record their feedback on the workings of the system itself. This feedback will be used to conduct the improvement of the system and also track user satisfaction.
Verifiable:	True
Verification:	Functionality can be tested in the user interface