Accessing WoSIS data

Niels Batjes, Eloi Ribeiro, Ad van Oostrum

ISRIC World Soil Information

Eloi Ribeiro

2017-05-17, Wageningen

World Soil Information Service (WoSIS)

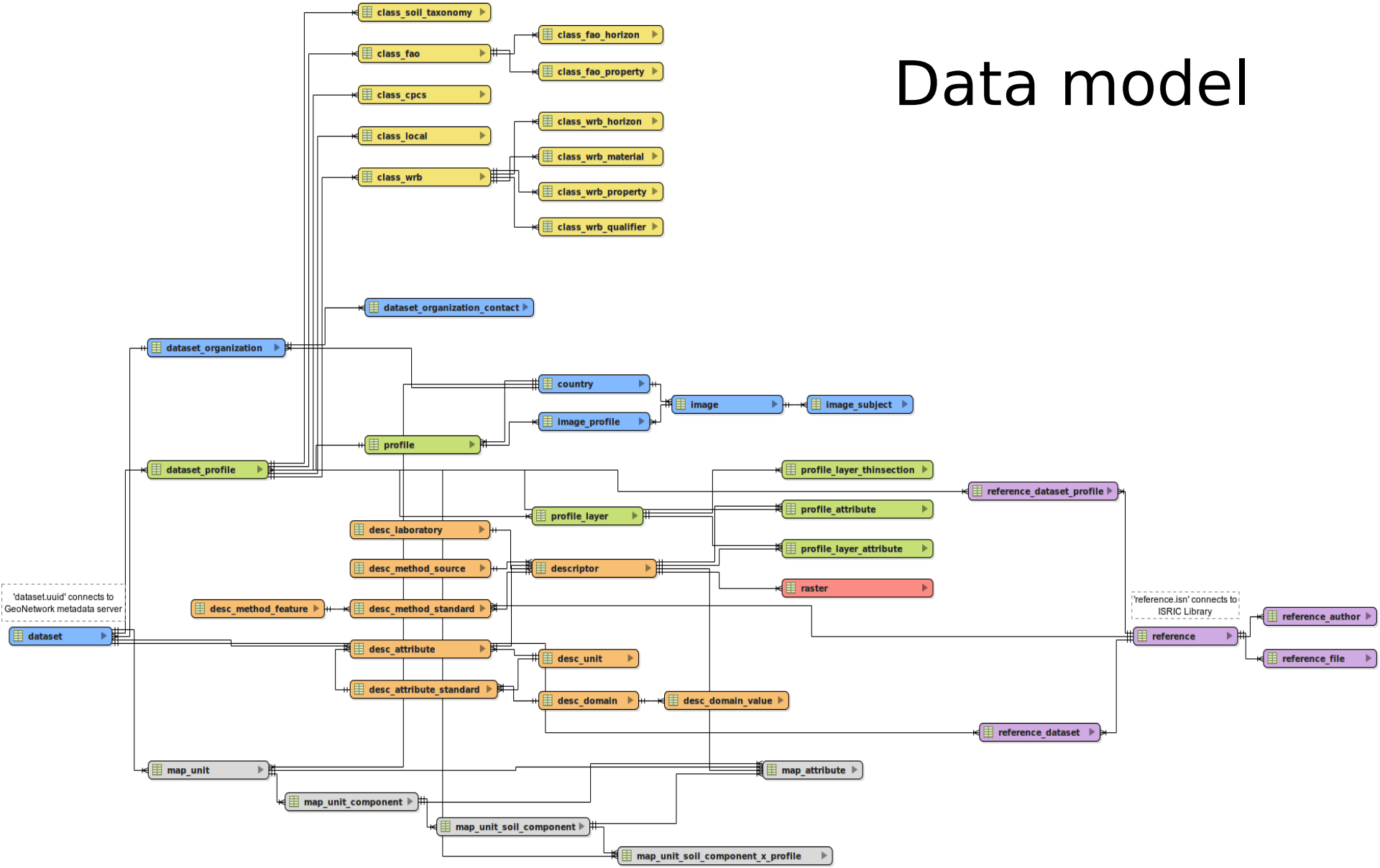
The aims of the WoSIS are:

- Safeguard world soil data 'as is'

- Share soil data (point, polygon) upon their standardization and harmonization

- Provide quality-assessed soil data for digital soil mapping and a range of environmental applications

Data model

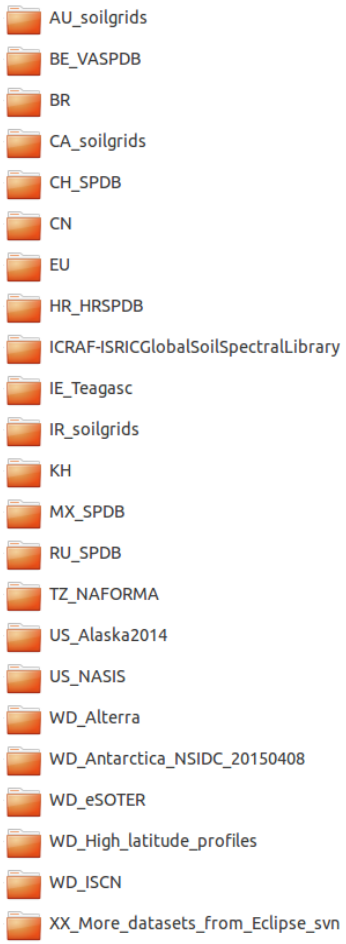

class_soil_taxonomy
class_fao
class_fao_horizon
class_fao_property
class_cpcs
class_local
class_wrb
class_wrb_horizon
class_wrb_material
class_wrb_property
class_wrb_qualifier
dataset_organization_contact
dataset_organization
country
image
image_subject
image_profile
profile
dataset_profile
profile_layer_thinsection
reference_dataset_profile
profile_attribute
profile_layer
desc_laboratory
profile_layer_attribute
desc_method_source
descriptor
raster
'dataset.uuid' connects to GeoNetwork metadata server
desc_method_feature
desc_method_standard
'reference.isn' connects to ISRIC Library
reference_author
dataset
desc_attribute
desc_unit
reference
reference_file
desc_attribute_standard
desc_domain
desc_domain_value
reference_dataset
map_unit
map_attribute
map_unit_component
map_unit_soil_component
map_unit_soil_component_x_profile

Import steps

- Select the dataset to be imported, using criteria like:
  - type of license
  - number of profiles
  - spatial distribution

- Convert the dataset into PostgreSQL format. Not into WoSIS! But to a separate schema, respecting the source data structure.

- Import the data into WoSIS.

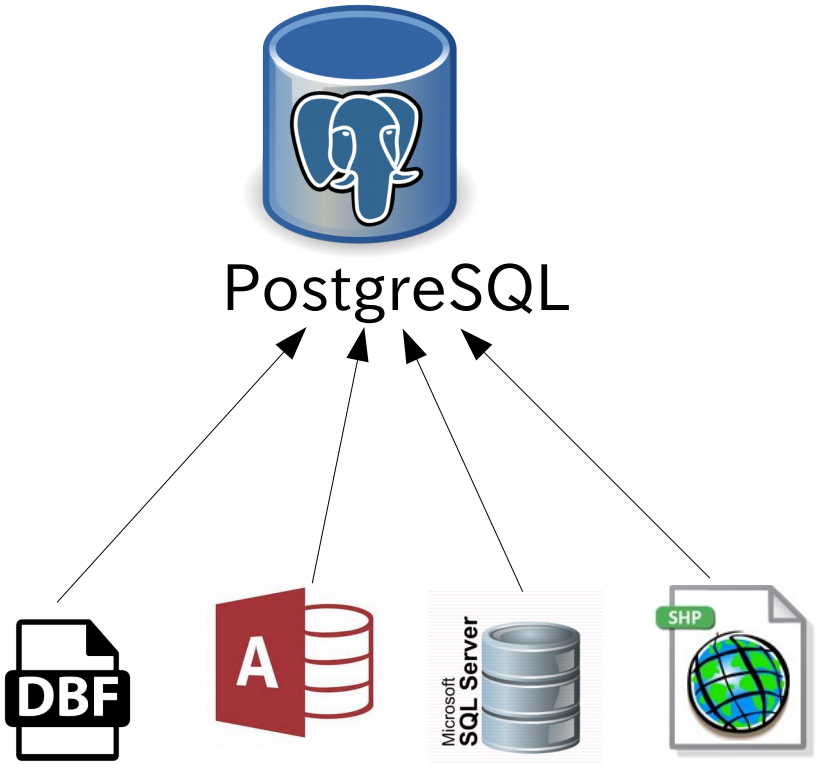


Import steps

- Select the dataset to be imported, using criteria like:
  - type of license
  - number of profiles
  - spatial distribution

- Convert the dataset into PostgreSQL format. Not into WoSIS! But to a separate schema, respecting the source data structure.

- Import the data into WoSIS.



Import steps

- Select the dataset to be imported, using criteria like:
  - type of license
  - number of profiles
  - spatial distribution

- Convert the dataset into PostgreSQL format. Not into WoSIS! But to a separate schema, respecting the source data structure.

- Import the data into WoSIS.

Import steps

- Select the dataset to be imported, using criteria like:
  - type of license
  - number of profiles
  - spatial distribution

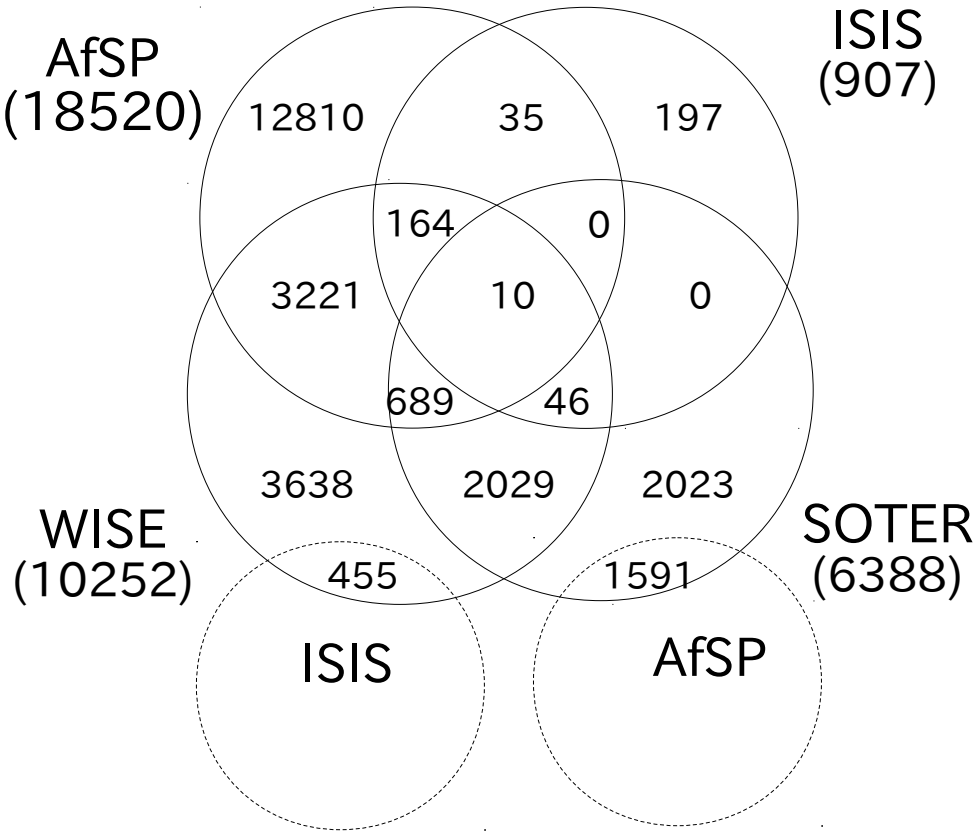
- Convert the dataset into PostgreSQL format. Not into WoSIS! But to a separate schema, respecting the source data structure.

- Import the data into WoSIS.

- Add primary keys
- Add foreign keys
- Grant permissions
- Change Data types
- Create geometry (Point,4326)
- Add comments on objects
- Fix referential integrity violations
- Fix repeated values
- Check codes and domains
- Fix Typos
- Backup

Standardization steps

- Identify repeated profiles
- Attribute names
- Units
- Conversion factors
- Measured values
- Analytical methods



Standardization steps

- Identify repeated profiles
- Attribute names
- Units
- Conversion factors
- Measured values
- Analytical methods

AfSP-1.2	Sand	% weight/weight	Sand total	g/100g	*1.0
CanSIS	PS SAND T	% weight/weight	Sand total	g/100g	*1.0
Caturrita-BR	Sand	g/kg	Sand total	g/100g	/10.0
ISIS	Total sand	% weight/weight	Sand total	g/100g	*1.0
NCSS	sand tot psa		Sand total	g/100g	*1.0
SMaria-BR	Sand	g/kg	Sand total	g/100g	/10.0
SOTER-AR	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-CAF	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-CN	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-CU	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-EUR	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-KE	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-KET	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-LAC	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-MW	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-NP	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-SAF	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-SN&GM	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-TN	Total sand	% weight/weight	Sand total	g/100g	*1.0
SOTER-ZA	Total sand	% weight/weight	Sand total	g/100g	*1.0
WISE	Sand content	% weight/weight	Sand total	g/100g	*1.0

Standardization steps

- Identify repeated profiles
- Attribute names
- Units
- Conversion factors
- Measured values
- Analytical methods

source_value text	standard_value numeric(6,2)
104	10.40
116	11.60
144	14.40
145	14.50
149	14.90
155	15.50
163	16.30
164	16.40
165	16.50
169	16.90
170	17.00
175	17.50
179	17.90
182	18.20
184	18.40
186	18.60
187	18.70
189	18.90
191	19.10
192	19.20
196	19.60
197	19.70
198	19.80
200	20.00
201	20.10

Standardization steps

- Identify repeated profiles
- Attribute names
- Units
- Conversion factors
- Measured values
- Analytical methods

Coarse sand text	Middle sand text	Fine sand text	Coarse silt text	Middle silt text	Fine silt text	Clay text	total_sand numeric	total_silt numeric	total numeric
1.01	2.87	19.62	12.95	19.02	15.05	29.4	23.50	47.02	99.99
1.44	6.77	15.62	14.46	18.61	16.36	26.7	23.83	49.43	100.01
0.84	2.68	16.19	13.55	17.77	15.03	33.9	19.71	46.35	100.00
2.55	3.52	18.23	11.38	20.24	20.05	24.0	24.30	51.67	100.00
3.65	7.43	22.95	17.06	20.80	11.74	16.3	34.03	49.60	100.00
4.85	15.93	25.91	15.35	19.99	6.67	11.2	46.69	42.01	99.99
2.07	10.02	31.97	13.14	17.34	10.44	15.0	44.06	40.92	99.99
30.26	14.29	9.71	4.59	8.47	12.27	20.4	54.26	25.33	100.00
25.84	17.45	10.35	5.03	9.47	12.33	19.5	53.64	26.83	99.99
27.43	15.65	12.33	5.35	9.45	12.74	17.0	55.41	27.54	100.02
26.26	17.62	11.19	1.98	8.33	14.09	20.5	55.07	24.40	100.00
41.55	14.78	11.04	2.67	6.50	11.22	12.2	67.37	20.39	100.00
0.28	0.89	6.86	7.51	19.29	21.89	43.2	8.03	48.69	100.00
0.34	1.13	6.82	9.49	19.54	21.71	40.9	8.29	50.74	100.00
0.12	0.95	6.60	8.86	20.76	22.37	40.3	7.67	51.99	99.99
0.27	0.88	6.86	9.89	19.78	21.85	40.4	8.01	51.52	100.00
0.68	1.59	9.52	9.34	18.83	19.06	40.9	11.79	47.23	100.00
1.40	2.44	13.89	13.57	20.46	13.75	34.4	17.73	47.78	99.99
1.80	2.76	15.14	13.03	17.41	13.82	36.0	19.70	44.26	100.01
3.01	3.45	15.77	13.65	18.45	12.85	32.8	22.23	44.95	100.00
4.20	14.61	18.46	7.30	10.08	11.11	34.2	37.27	28.49	100.00
4.22	15.83	18.63	6.72	7.68	9.89	37.0	38.68	24.29	100.00
2.67	15.32	21.23	6.71	9.58	11.36	33.1	39.22	27.65	100.00

Standardization steps

- Identify repeated profiles
- Attribute names
- Units
- Conversion factors
- Measured values
- Analytical methods

attribute_agg_name [PK] character varyin	feature_name [PK] character	feature_option [PK] text
Sand	dispersion	Ammonium [NH4]
Sand	dispersion	no dispersion
Sand	dispersion	Sodium hexametaphosphate [(NaPO3)6] - Calgon ty
Sand	dispersion	Sodium hydroxide [NaOH]
Sand	dispersion	unknown
Sand	instrument	analyzer
Sand	instrument	field hand estimate
Sand	instrument	hydrometer
Sand	instrument	sieve
Sand	instrument	unknown
Sand	pretreatment	Hydrogen peroxide [H2O2] plus Hydrochloric acid
Sand	pretreatment	Hydrogen peroxide [H2O2] plus mild Acetic acid
Sand	pretreatment	no pretreatment
Sand	pretreatment	pretreatment, deferration included
Sand	pretreatment	unknown
Sand	size	0.02 - 2 mm
Sand	size	0.05 - 0.1 mm
Sand	size	0.05 - 1.7 mm
Sand	size	0.05 - 1 mm
Sand	size	0.05 - 2 mm
Sand	size	0.06 - 2 mm
Sand	size	0.063 - 2 mm
Sand	size	0.10 - 0.25 mm
Sand	size	0.2 - 2 mm
Sand	size	0.25 - 0.5 mm
Sand	size	1 - 2 mm
Sand	size	unknown

Standardization steps

- Identify repeated profiles
- Attribute names
- Units
- Conversion factors
- Measured values
- Analytical methods

descriptor [PK] smalli	standard_a [PK] text	feature_name [PK] text	feature_option [PK] text	confidence character va
11585	Sand	pretreatment	pretreatment, deferration included	high
11585	Sand	size	unknown	high
11586	Sand	dispersion	Sodium hexametaphosphate [(NaPO3)6]	high
11586	Sand	instrument	sieve	high
11586	Sand	pretreatment	pretreatment, deferration included	high
11586	Sand	size	unknown	high
11587	Sand	dispersion	Sodium hexametaphosphate [(NaPO3)6]	high
11587	Sand	instrument	sieve	high
11587	Sand	pretreatment	no pretreatment	high
11587	Sand	size	unknown	high
11588	Sand	dispersion	Sodium hexametaphosphate [(NaPO3)6]	high
11588	Sand	instrument	sieve	high
11588	Sand	pretreatment	no pretreatment	high
11588	Sand	size	unknown	high
11589	Sand	dispersion	Sodium hexametaphosphate [(NaPO3)6]	high
11589	Sand	instrument	sieve	high
11589	Sand	pretreatment	no pretreatment	high
11589	Sand	size	unknown	high
11590	Sand	dispersion	Sodium hexametaphosphate [(NaPO3)6]	high
11590	Sand	instrument	sieve	high
11590	Sand	pretreatment	no pretreatment	high
11590	Sand	size	unknown	high
11591	Sand	dispersion	Sodium hexametaphosphate [(NaPO3)6]	high
11591	Sand	instrument	sieve	high
11591	Sand	pretreatment	no pretreatment	high
11591	Sand	size	unknown	high

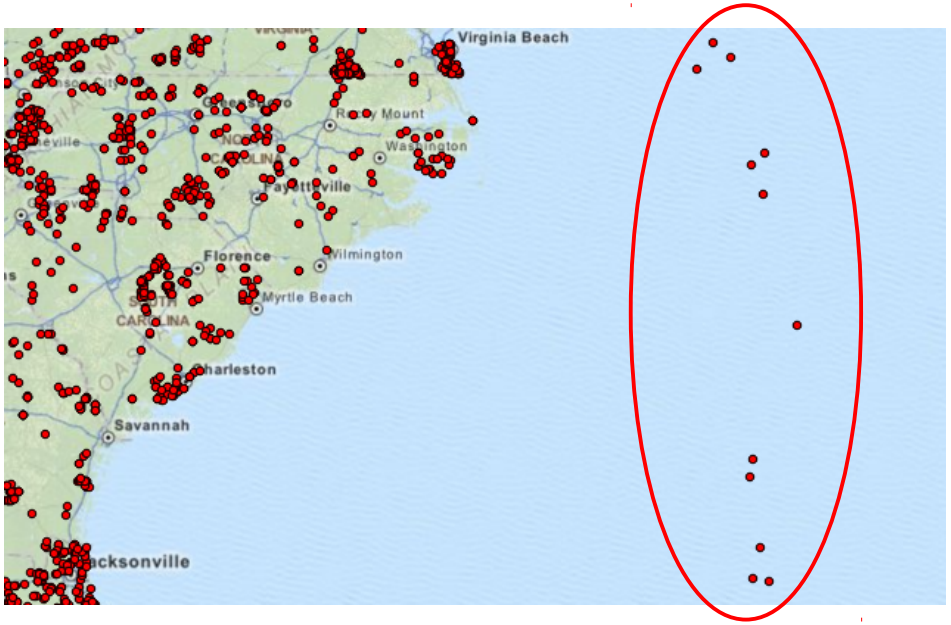
Standardization steps

- Identify repeated profiles
- Attribute names
- Units
- Conversion factors
- Measured values
- Analytical methods

source_analytical_method_name text	standard_analytical_method_name text
Sand fraction (0.053 to 2 mm) determined	dispersion = Sodium hydroxide [NaOH], instrument = sieve,
Pipette method, with appropriate dispers	dispersion = unknown, instrument = sieve, pretreatment =
NaOH 1N,shaking witho.remov.OM&carbonate	dispersion = Sodium hydroxide [NaOH], instrument = unknow
TE07-/- Bouyoucos. Sand= 0.2-2mm plus 50	dispersion = unknown, instrument = sieve, pretreatment =
Pipette method, with appropriate dispers	dispersion = unknown, instrument = sieve, pretreatment =
Calgon, shaking,without remov.OM&carbona	dispersion = Sodium hexametaphosphate [(NaPO3)6] - Calgon
Pipette method, full dispersion (c<0.001	dispersion = unknown, instrument = sieve, pretreatment =
TAMIZADO	dispersion = unknown, instrument = sieve, pretreatment =
Fine earth fraction; 0.25-0.10 mm by sie	dispersion = unknown, instrument = sieve, pretreatment =
Pipette method, with appropriate dispers	dispersion = Ammonium [NH4], instrument = sieve, pretreat
Hydrometer method, with dipsersion treat	dispersion = unknown, instrument = hydrometer, pretreatme
Pipette method, with appropriate dispers	dispersion = unknown, instrument = sieve, pretreatment =
Pipette method, with appropriate dispers	dispersion = unknown, instrument = sieve, pretreatment =
Pipette method, with appropriate dispers	dispersion = unknown, instrument = sieve, pretreatment =
Fine earth fraction; 1.0-0.5 mm by sievi	dispersion = unknown, instrument = sieve, pretreatment =
Pipette method, with appropriate dispers	dispersion = Ammonium [NH4], instrument = sieve, pretreat
TE07-/- Sand= 0.2-2mm plus 50% of 0.02-0	dispersion = unknown, instrument = sieve, pretreatment =
Pipette method, with appropriate dispers	dispersion = unknown, instrument = sieve, pretreatment =
Hydrometer method, with dispersion treat	dispersion = unknown, instrument = hydrometer, pretreatme
Pipette method, with appropriate dispers	dispersion = unknown, instrument = sieve, pretreatment =
Pipette method, with appropriate dispers	dispersion = unknown, instrument = sieve, pretreatment =
Fine earth fraction; 0.10-0.05 mm by sie	dispersion = unknown, instrument = sieve, pretreatment =
Hydrometer method, with dispersion treat	dispersion = unknown, instrument = hydrometer, pretreatme

Last steps

- Filter profiles (water, wrong country,...)
- Filter measurements
- Refresh tables for GeoServer



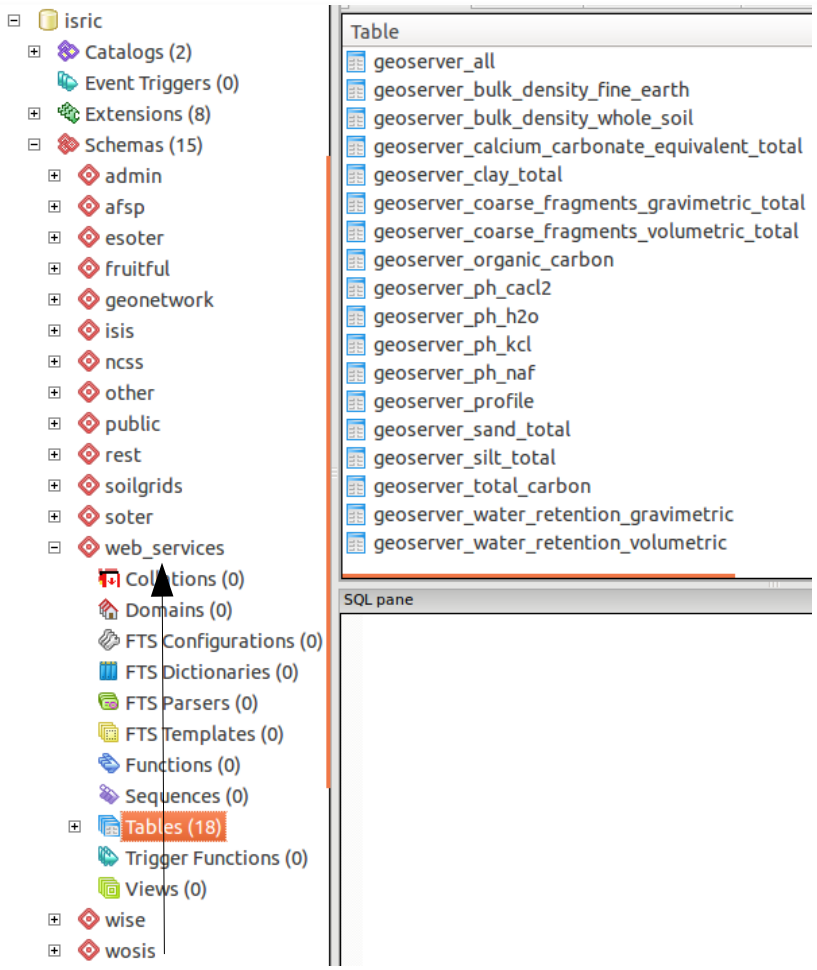
Last steps

- Filter profiles (water, wrong country,...)
- Filter measurements
- Refresh tables for GeoServer

desc_attribute_standard_id character varying(100)	minimum numeric(6,2)	maximum numeric(6,2)
Bulk density fine earth	0.05	2.70
Bulk density whole soil	0.05	3.60
Calcium carbonate equivalent total	0.00	1000.00
Clay total	0.00	100.00
Coarse fragments gravimetric total	0.00	100.00
Coarse fragments volumetric total	0.00	100.00
Organic carbon	0.00	1000.00
pH CaCl2	1.50	13.00
pH H2O	1.50	13.00
pH KCl	1.50	13.00
pH NaF	1.50	13.00
Sand total	0.00	100.00
Silt total	0.00	100.00
Total carbon	0.00	1000.00
Water retention gravimetric	0.00	100.00
Water retention volumetric	0.00	100.00

Last steps

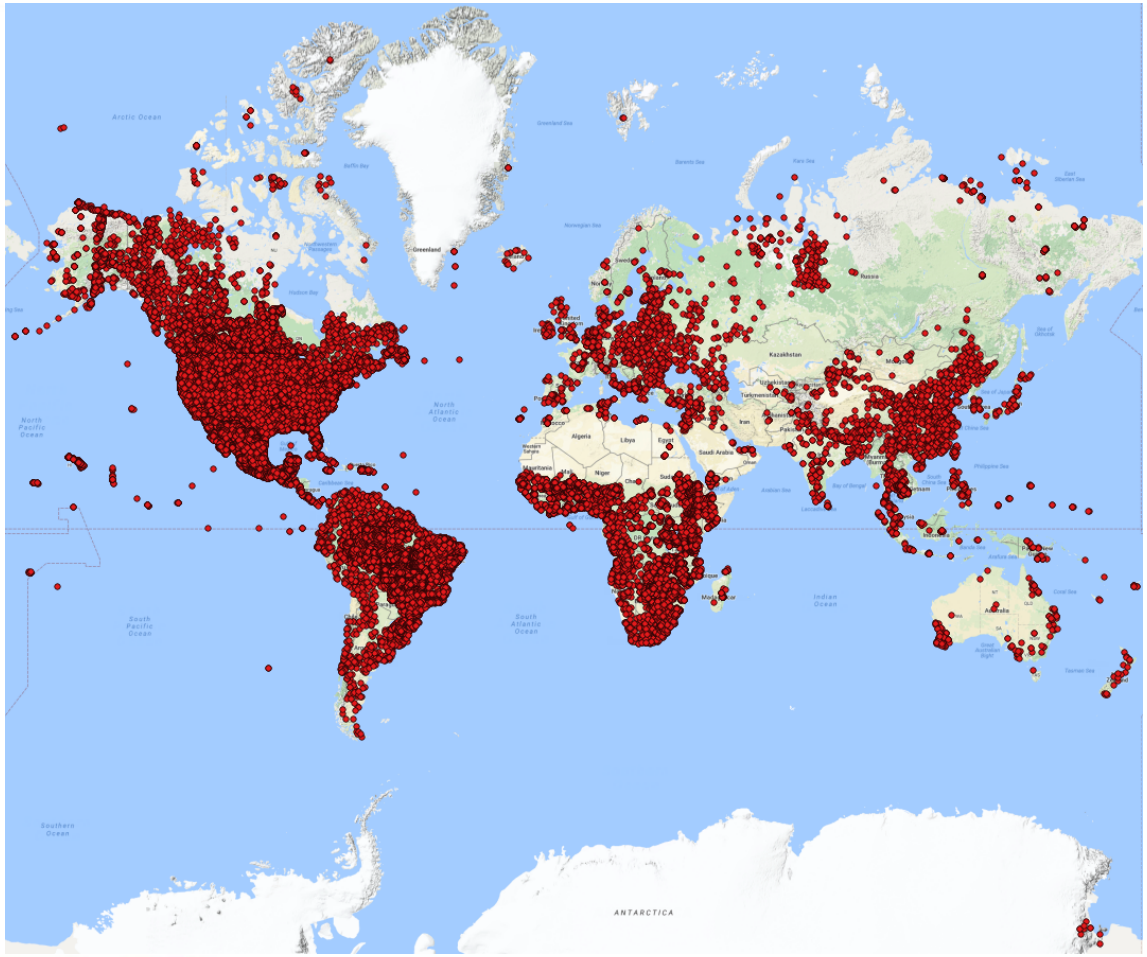
- Filter profiles (water, wrong country,...)
- Filter measurements
- Refresh tables for GeoServer



Imported datasets

dataset_id character varying(20)	dataset_title character varying(220)	n_profiles integer	n_layer_rows_inserted integer
US-NCSS	National Cooperative Soil Survey, National Cooperative Soil Cl	64049	24263841
AF-AfSP-1.2	Africa Soil Profiles Database	18514	1619478
MX-INEGI	Conjunto de Datos Edafológicos Alfanuméricos, Serie I & II, Co	12535	656423
WD-WISE	World Inventory of Soil Emission Potentials	10242	701669
CA-CanSIS	Canadian Soil Information Service	7177	781624
BR-Cooper	National Soil Profile Database for Brazil Available to Interna	5616	186148
BR-Bernoux	Radam Brazil Project - Exploratory Soil Surveys / Projeto Rada	3795	406487
KH-CambSoils	Cambodia Soils Data Base	1712	55933
LAC-SOTER	Soil and Terrain Database for Latin America and the Caribbean	1672	117614
CN-SOTER	Soil and Terrain Database for China	1430	58242
SAF-SOTER	Soil and Terrain Database for Southern Africa	969	74702
WD-ISIS	ISRIC Soil Information System	907	323319
EU-SOTER	Soil and Terrain Database for Central and Eastern Europe	661	44192
ZA-SOTER	Soil and Terrain Database for South Africa	619	48929
KE-SOTER	Soil and terrain database for Kenya	440	35709
BR-Samuel-Rosa-2015	Alessandro Samuel-Rosa PhD thesis data from Santa Maria, Rio (	410	9638
MW-SOTER	Soil and Terrain Database for Malawi	375	19879
BE-UplandsI	Soil sampling in the belgian uplands I	305	21779
AR-SOTER	Soil and Terrain Database for Argentina	222	10008
CAF-SOTER	Soil and Terrain Database for Central Africa	165	21421
NP-SOTER	Soil and Terrain Database for Nepal	156	8787
NL-Kempen	Bas Kempen data from Drenthe province in The Netherlands	150	1253
KET-SOTER	Soil and Terrain Database for Upper Tana River Catchment	140	9479
SN&GM-SOTER	Soil and Terrain Database for Senegal and the Gambia	96	5467
BT-Caspari	Thomas Caspari PhD thesis data from Bhutan	80	13492
TN-SOTER	Soil and Terrain Database for Tunisia	56	5317
BR-Samuel-Rosa-2011	Alessandro Samuel-Rosa data from Caturrita, Santa Maria, Rio (	48	863
CU-SOTER	Soil and Terrain Database for Cuba	30	3363

WoSIS in numbers



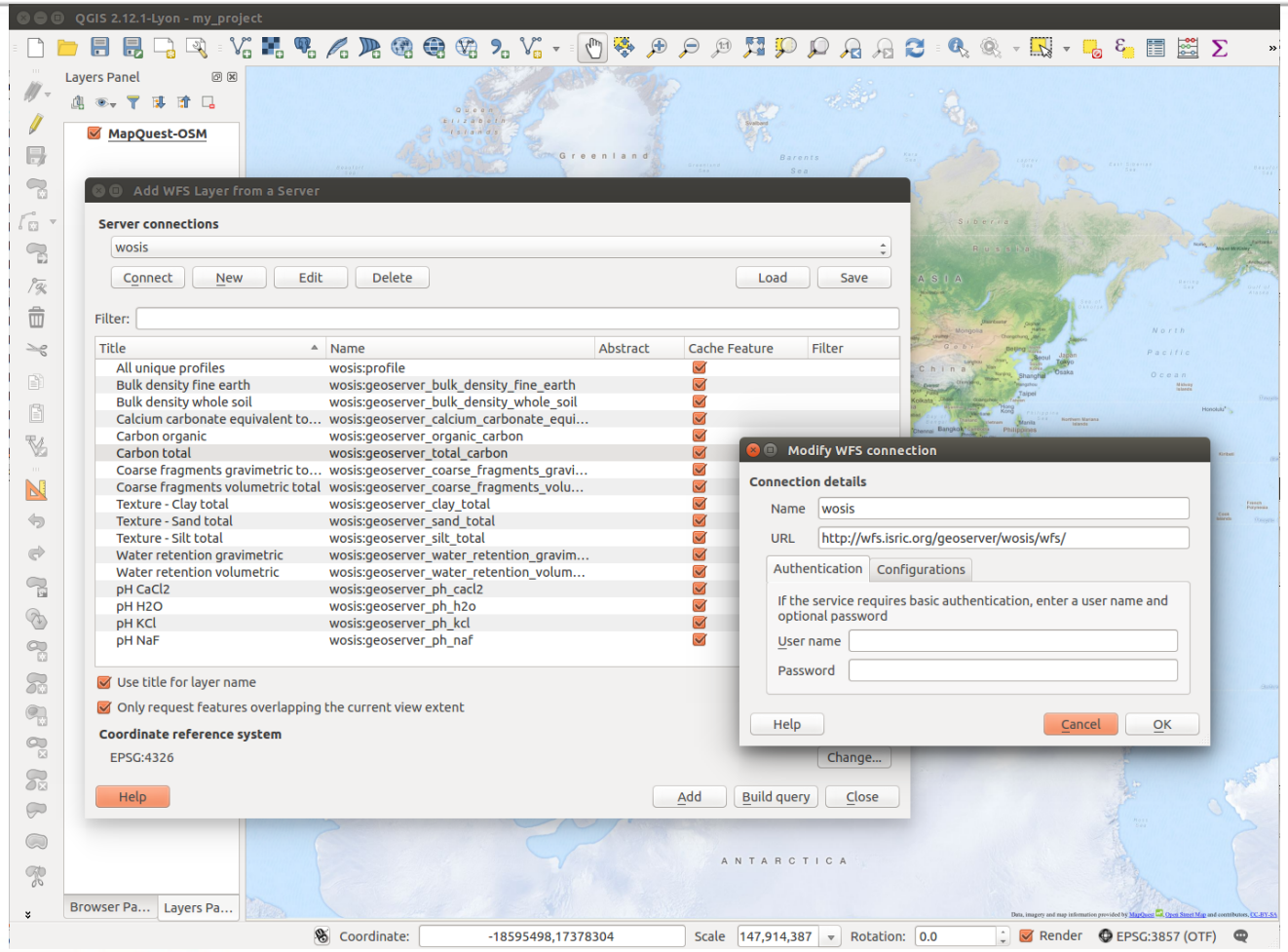
- **126 589** unique profiles
- **110 897** profiles with geometry
- **30 million** soil properties measurements
- **4 million** soil properties measurements being served through WFS
- **10** soil properties being served through WFS:
  - Bulk density
  - Calcium carbonate
  - Carbon (Total / Organic)
  - Coarse fragments
  - pH
  - Water retention
  - Texture (Sand, Silt, Clay)
  - Cation exchange capacity
  - Electrical conductivity
  - Classification: FAO, WRB, S. taxonomy

WoSIS derived datasets

Latest (dynamic) dataset: This dataset contains the most recent complement of standardised soil data served from WoSIS. Being dynamic, the dataset will grow once new point data are standardised, additional soil properties are considered, and/or when possible corrections are required. The latest dataset is served through an OGC-compliant WFS (Web Feature Service). The point data can be accessed from GIS or R.

Snapshot (static) dataset: These are a representation of the complement of standardised data available at a given moment (e.g. July 2016). Each snapshot is given a unique name and digital object identifier (doi) for consistent citation purposes. So far, one snapshot has been released. Served as a zipped txt files. Methodological details are provided in a paper in Earth System Data Journal.

WoSIS latest



- http://wfs.isric.org/geoserver/wosis/wfs/
- One central place for data gathering by GIS applications
- Free to use and download
- Start your research immediately with standardized and validated soil data
- More info at isric.org/explore/wosis

WoSIS snapshots

Were: http://geonode.isric.org/layers/geonode:wosis_201607_profiles
Download Layer
Zipped Shapefile

Or in the command line

wget http://geonode.isric.org/uploaded/documents/WoSIS_2016_July.zip
unzip WoSIS_2016_July.zip

What: 32 MB zipped file
4 files in it
ReadmeFirst_WoSIS_2016.pdf (148 KB)
wosis_201607_attributes.txt (4 KB / 23 records)
wosis_201607_layers.txt (679 MB / 455 820 records)
wosis_201607_profiles.txt (9 MB / 94 442 records)

How: Using R

WoSIS snapshots

 # load data

```
setwd("/home/eloi/Downloads/")
attributes = read.table("wosis_201607_attributes.txt", sep="\t",quote = "",
header=TRUE)
profiles = read.table("wosis_201607_profiles.txt", sep="\t",quote = "",
header=TRUE)
layers = read.table("wosis_201607_layers.txt", sep="\t", quote="",
header=TRUE)

dim(attributes)
dim(profiles)
dim(layers)

colnames(attributes)
colnames(profiles)
colnames(layers)
```

WoSIS snapshots

 # display data
attributes[1:22, 1:5]
profiles[1:10, 1:6]
layers[1:10, 1:6]

 # merge profiles with layers
mer <- merge(x = profiles, y = layers, by = "profile_id", all = TRUE)
sel <- mer[mer$country_name=="Brazil" & !is.na(mer[,"orgc_value_avg"]),]
sel <- sel[order(sel$profile_id, sel$top, sel$bottom),]
dim(sel)

 # plot data
hist(log(sel[,"orgc_value_avg"]))
plot(sel$longitude, sel$latitude)

WoSIS snapshots

 # plot point data

```
install.packages("leaflet", dependencies = TRUE)
library(leaflet)
mymap <- leaflet(sel) %>% addProviderTiles(providers$CartoDB.Positron)
%>% addMarkers(~longitude, ~latitude, popup = ~paste(as.character(top),
as.character(bottom), as.character(orgc_value_avg), sep=" | "), label =
~paste(as.character(top), as.character(bottom),
as.character(orgc_value_avg), sep=" | "), clusterOptions =
markerClusterOptions())
mymap
```

More about the use of Leaflet in R: https://rstudio.github.io/leaflet/

Thank you

ISRIC World Soil Information

Niels Batjes
Eloi Ribeiro
Ad van Oostrum
Johan Leenaars

isric.org

CC BY