

Digital Platform for Soil Health Monitoring and Management: Data Model Documentation

Sukanya Burry*, Gilbert Rozario, Mukund Patil

*Corresponding author: Sukanya.burry@icrisat.org

Abstract:

This Data Model Documentation is a comprehensive guide describing the structure, relationships, and specifications of a database system used to develop the digital soil health monitoring platform. It serves as a blueprint for database design and implementation, ensuring that all stakeholders—including business analysts, database developers, and system administrators—have a shared understanding of how data is organized and managed.

Table of Contents

Introduction	4
Purpose.....	4
Database Design Overview.....	4
Conceptual data model	4
Logical data model	5
Master tables.....	6
Transactional tables	7
Entity relationships.....	8
Physical Data Model	9

List of tables

Table 1 Higher-level description of entities to be used in the platform	5
Table 2 Details of master tables	6
Table 3 Details of transactional tables	7
Table 4 Details of entity relationships.	8
Table 5 States table	9
Table 6 Districts table	9
Table 7 Blocks table	9
Table 8 Villages table	9
Table 9 Crops table	10
Table 10 Cultivars table	10
Table 11 Projects	10
Table 12 Mapping table for projects and states.....	10
Table 13 National agriculture research partners	10
Table 14 Mapping table for NARS and states.....	10
Table 15 User roles	11
Table 16 Users table.....	11
Table 17 Users and Projects Mapping table	11
Table 18 Farmers table	11
Table 19 Farmers and project mapping table	12
Table 20 Farms table.....	12

Table 21 Soil sample table.....	12
Table 22 Data table for soil samples lots.	13
Table 23 Mapping table between LOT and Soil Sample	13
Table 24 Soil sample dispatch details	13
Table 25 Mapping Table between transport and LOTs	13
Table 26 Mapping Table between transport and reference photos	13
Table 27 Soil History Table	13
Table 28 Soil analysis data.....	14
Table 29 Table to store the information about digital soil map and model.....	14

Digital Platform for Soil Health Monitoring and Management: Data Model Documentation

Short name	MSOIL Data Model
Version	1.0
Date	November 2024
Authors	Sukanya Burry, Gilbert Rozario, Mukund Patil
Status	Complete

Introduction

Purpose

This document describes the development of data model from a conceptual data model to a physical data model for the MSOIL – a digital platform for soil health monitoring and management. The MSOIL platform is aimed at digitizing soil health monitoring by leveraging digital data sourcing and analytical tools to provide actionable insights to farmers, agronomists, and policymakers. A well-structured and comprehensive database forms the backbone for the proposed digital platform. The scope of this document is to define the database structure, attributes of tables, and relationships between the various data tables required for developing the proposed digital platform.

Database Design Overview

Conceptual data model

A Conceptual Data Model (CDM) is the highest level of abstraction in database design. It provides a simplified, business-oriented representation of the data elements and their

relationships in a system. The CDM focuses on the ‘what’ of the data rather than how it will be implemented, serving as a communication tool between stakeholders such as business analysts, data architects, and developers.

Table 1 Higher-level description of entities to be used in the platform

Sr. No.	Entities	Relationships
1	Farmer	Farmers → belongs to → a village → owns a farm
2	Farm	Farms → have → many soil samples
3	Soil sample	Soil sample → belongs to → a farm → has → soil analysis results → has → real-time tracking status
4	Users	User → has → a role → assign to → multiple projects
5	Village	Village → belong to → a block → belongs to → a district → belongs to → a state
6	State	State → has → many districts → has → may → blocks → has → many villages
7	Crop	Crop → has → many cultivars
8	NARS	NARS → has → jurisdiction over multiple states or districts
9	Fertilizer recommendations	Fertilizer recommendations → by → NARS → for → many crops
10	Soil model/maps	Soil model → belongs to → block/district

Logical data model

A Logical Data Model (LDM) is a conceptual representation of the structure of a database. It is an intermediary step between the Conceptual Data Model (which focuses on high-level business entities) and the Physical Data Model (which focuses on implementation details). The LDM outlines the entities, attributes, and relationships in the system while being independent of any specific database management system (DBMS).

The data model is divided into master tables (Table 2) and transactional tables (Table 3). Further, an Entity-Relationship (ER) model presented in Table 4 is used in database design to represent the structure of data and the relationships between entities in a system.

Master tables

Master tables are foundational tables in a relational database that store relatively static or reference data. They act as a single source of truth for key entities and attributes used throughout the system. Their primary purpose is to ensure data consistency, standardization, and reusability across transactional and other related tables. Following are the master tables essential for the platform.

Table 2 Details of master tables

Sr.No	Entities	Description
1	<code>states*</code>	The state names of the project area will be stored in this table. It is assumed that the information stored in this table belongs to the administrative Level-1 (States, Provinces, and equivalent).
2	<code>districts*</code>	The district names will be stored in this table with a relationship with the <code>states</code> table. This is equivalent to administrative Level-2 (counties, districts, etc.)
3	<code>blocks*</code>	The block names will be stored in this table with a relationship with the <code>districts</code> table. This is equivalent to administrative Level-3 (communes, Municipalities, sub-districts, etc.)
4	<code>villages*</code>	The village names will be stored in this table with a relationship with the <code>blocks</code> table.
5	<code>crops</code>	The crop names are stored in this table. List of crops displayed in select box while collecting soil samples and getting the fertilizer recommendations.
6	<code>nars</code>	The knowledge repository will gather information from various knowledge-generating institutes. The National Agricultural Research System (NARS) is authorized to provide the package of cropping practices. The names of these NARS institutes will be stored in this table, which will be used for getting site-specific advisories.
7	<code>roles</code>	The role names will be stored in this table with a relationship with the <code>users</code> table.
8	<code>projects</code>	The project names will be stored in this name with a relationship with <code>farmers</code> , the <code>users</code> table
9	<code>farmers</code>	Basic information about farmers will be stored in this table. This table stores various information on the farmer's profile. The minimum information required in this table are farmer's name, contact details, and location information.
10	<code>farms</code>	Details of farm parcels linked to individual farmers. These farm records will be referred, while collecting soil samples and providing advisories on fertilizer recommendations.

11	users	The user tables provide secure access to the database, allowing stakeholders to interact with various tools on the platform. However, access to the platform is controlled through assigned roles and projects.
----	-------	---

*The administrative levels used in this platform are specific to India but may be customized to any other country/region.

Transactional tables

Transactional tables are database tables designed to store dynamic, event-driven data that records specific actions or operations performed within a system. This data reflects real-time changes and is typically used to capture and manage system activity or user interactions. For example, soil sample data, soil analysis data, etc. All the transactional tables are linked with master tables through foreign keys. Following is the list of transactional tables being used in the platform.

Table 3 Details of transactional tables

Sr.No.	Entities	Description
1	soilSamples	This table holds information on the geocoordinates of the soil samples along with data related to cropping patterns and soil type.
2	soilAnalysis	This table stores soil analysis data submitted by the laboratory users.
3	projectUser	This table is used to map users with specific projects so that access to data is restricted to assigned projects only.
4	projectFarmers	This table stores the relationship mapping between projects and farmers
5	sampleLot	This table stores information on lots created for transporting.
6	sampleLotMap	This table stores information on soil sample lots and the associated soil samples.
7	soilTransport	This table stores information related to soil sample dispatch details.
8	transportLots	This table stores the mapping details between soilTransport and sampleLots. One record in the soilTransport may be associated with multiple records in sampleLots.
9	transportPhotos	This table stores the paths of photos related to records of soilTransport.
10	sampleHistory	This table stores the tracking history of soil samples from sampling till the analysis.
11	models	The model table stores information on the model file path and its performance matrix.

Entity relationships

Table 4 Details of entity relationships. .

SNo	Entities	Relationship
1	states	states (1:n) districts
2	districts	districts (1:n) blocks
3	blocks	blocks (1:n) villages
4	villages	villages (1:n) farmers villages (1:n) users villages (1:n) farms villages (1:n) soilSamples villages (1:n) projects
5	crops	crops (1:n) fertRec
6	cultivars	crops (1:n) cultivars
7	projects	-
8	projectStateMap	projects (m:n) states
9	nars	-
10	stateNarsMap	nars (m:n) states
11	roles	-
12	users	users (1:1) roles users (1:n) projects users (1:n) farmers users (1:n) farms users (1:n) soilSamples users (1:n) sampleLot users (1:n) sampleTransport users (1:n) sampleStatus users (1:n) soilAnalysis
13	projectUsersMap	projects (m:n) users
14	farmers	farmers (1:1) villages
15	projectFarmersMap	projects (m:n) farmers
16	farms	farms (1:1) farmers farms (1:n) projects farms (1:1) villages farms (1:n) soilSamples
17	soilSamples	soilSamples (1:1) farms soilSamples (1:1) projects
18	sampleLot	-
19	sampleLotMap	sampleLot (1:n) soilSamples
20	transportLot	-
21	transportPhotos	transportLot (1:n) transportPhotos
22	soilTransportMap	transportLot (1:n) sampleLot
23	sampleHistory	sampleHistory (1:1) soilSample
24	soilAnalysis	soilSamples (1:1) soilAnalysis
25	soilMaps	districts (1:n) soilMaps

Physical Data Model

A Physical Data Model (PDM) is a detailed representation of how data is physically stored in a database system. It defines the technical implementation of a database, including tables, columns, relationships, indexes, data types, and constraints. The PDM is derived from the Logical Data Model and is tailored to suit the specific database management system (DBMS) being used. Table 5 – 17 provide details of structures for different data tables used in the platform.

Table 5 States table

Attribute name	Data type	Constraints	Description
stateId	INT	Primary Key, Auto Increment	Unique identifier for state.
stateName	VARCHAR	NOT NULL	Full name of the state.
status	INT	NOT NULL	1: active , 0 : inactive

Table 6 Districts table

Attribute name	Data type	Constraints	Description
districtId	INT	Primary Key, Auto Increment	Unique identifier for district.
stateId	INT	NOT NULL	A foreign key to link <i>states</i> table.
districtName	VARCHAR	NOT NULL	Full name of the district
status	INT	NOT NULL	1: active , 0 : inactive

Table 7 Blocks table

Attribute name	Data type	Constraints	Description
blockId	INT	Primary Key, Auto Increment	Unique identifier for block.
stateId	INT	NOT NULL	A foreign key to link <i>states</i> table.
districtId	INT	NOT NULL	A foreign key to link <i>districts</i> table.
blockName	VARCHAR	NOT NULL	Full name of the block.
status	INT	NOT NULL	1: active , 0 : inactive

Table 8 Villages table

Attribute name	Data type	Constraints	Description
villageId	INT	Primary Key, Auto Increment	Unique identifier for village.
blockId	INT	NOT NULL	A foreign key to link <i>blocks</i> table.
stateId	INT	NOT NULL	A foreign key to link <i>states</i> table.
districtId	INT	NOT NULL	A foreign key to link <i>districts</i> table.
villageName	VARCHAR	NOT NULL	Village name
status	INT	NOT NULL	1: active , 0 : inactive

Table 9 Crops table

Attribute name	Data type	Constraints	Description
cropId	INT	Primary Key, Auto Increment	Unique identifier for crop.
cropName	VARCHAR	NOT NULL	Crop name
status	INT	NOT NULL	1: active , 0 : inactive

Table 10 Cultivars table

Attribute name	Data type	Constraints	Description
cultivarId	INT	Primary Key, Auto Increment	Unique identifier for cultivar.
cropId	INT	NOT NULL	A foreign key to link with <code>crops</code> table
cultivarName	VARCHAR	NOT NULL	Cultivar name
status	INT	NOT NULL	1: active , 0 : inactive

Table 11 Projects

Attribute name	Data type	Constraints	Description
projectId	INT	Primary Key, Auto Increment	Unique identifier for projects.
projectName	VARCHAR	NOT NULL	Name of the project.
projectDescription	TEXT	NULL	Description of project.
status	INT	NOT NULL	1: active , 0 : inactive

Table 12 Mapping table for projects and states.

Attribute name	Data type	Constraints	Description
mappingID	INT	Primary Key, Auto Increment	Unique identifier for projects.
projectId	INT	NOT NULL	A foreign key referring to projects table
stateId	INT	NOT NULL	A foreign key referring to states table
status	INT	NOT NULL	1: active , 0 : inactive

Table 13 National agriculture research partners

Attribute name	Data type	Constraints	Description
narsId	INT	Primary Key, Auto Increment	Unique identifier for NARS.
stateId	INT	NOT NULL	A foreign key to link with <code>states</code> table.
narsName	VARCHAR	NOT NULL	Full name of the institute.
status	INT	NOT NULL	1: active , 0 : inactive

Table 14 Mapping table for NARS and states.

Attribute name	Data type	Constraints	Description
mappingID	INT	Primary Key, Auto Increment	Unique identifier for projects.
narsId	INT	NOT NULL	A foreign key referring to nars table
stateId	INT	NOT NULL	A foreign key referring to states table
status	INT	NOT NULL	1: active , 0 : inactive

Table 15 User roles

Attribute name	Data type	Constraints	Description
roleId	INT	Primary Key, Auto Increment	Unique identifier for roles.
roleName	VARCHAR	NOT NULL	Name of the access role.
roleDesc	TEXT	NULL	Description of provilages.
status	INT	NOT NULL	1: active , 0 : inactive

Table 16 Users table

Attribute name	Data type	Constraints	Description
userId	INT	Primary Key, Auto Increment	Unique identifier for users.
userName	VARCHAR	NOT NULL	Username to be used for login.
firstName	VARCHAR	NOT NULL	First name of the user.
lastName	VARCHAR	NOT NULL	Last name of the user.
email	VARCHAR	NULL	Email address of user for verification.
phoneNumber	VARCHAR	NULL	Phone number of users.
password	VARCHAR	NOT NULL	Password to be set for login.
roleId	INT	NOT NULL	A foreign key to link with <code>roles</code> table
stateId	INT	NOT NULL	A foreign key to link with <code>states</code> table
districtId	INT	NOT NULL	A foreign key to link with <code>districts</code> table
blockId	INT	NOT NULL	A foreign key to link with <code>blocks</code> table
villageId	INT	NULL	A foreign key to link with <code>villages</code> table
status	INT	NOT NULL	1: active , 0 : inactive

Table 17 Users and Projects Mapping table

Attribute name	Data type	Constraints	Description
mappingId	INT	Primary Key, Auto Increment	Unique identifier for mapping.
userId	INT	NOT NULL	A foreign key to link with users table
projectId	INT	NOT NULL	A foreign key to link with project table
status	INT	NOT NULL	1: active , 0 : inactive

Table 18 Farmers table

Attribute name	Data type	Constraints	Description
farmerId	INT	Primary Key, Auto Increment	Unique identifier for users.
firstName	VARCHAR	NOT NULL	First name of the farmer.
lastName	VARCHAR	NOT NULL	Last name of the farmer.
email	VARCHAR	NULL	Email address of farmer (if available).
phoneNumber	VARCHAR	NULL	Phone number of farmer (if available).
stateId	INT	NOT NULL	A foreign key to link with <code>states</code> table
districtId	INT	NOT NULL	A foreign key to link with <code>districts</code> table
blockId	INT	NOT NULL	A foreign key to link with <code>blocks</code> table
villageId	INT	NULL	A foreign key to link with <code>villages</code> table
status	INT	NOT NULL	1: active , 0 : inactive

userId	INT	NOT NULL	A foreign key to link with <code>users</code> table.
--------	-----	----------	--

Table 19 Farmers and project mapping table

Attribute name	Data type	Constraints	Description
mappingId	INT	Primary Key, Auto Increment	Unique identifier for mapping.
farmerId	INT	NOT NULL	A foreign key to link with <code>farmers</code> table.
projectId	INT	NOT NULL	A foreign key to link with <code>projects</code> table.
status	INT	NOT NULL	1: active , 0 : inactive

Table 20 Farms table

Attribute name	Data type	Constraints	Description
farmId	INT	Primary Key, Auto Increment	Unique identifier for users.
projectId	INT	NOT NULL	A foreign key to link with project table.
farmName	VARCHAR	NOT NULL	Name of the farm for easy identification.
farmerId	INT	NOT NULL	A foreign key to link with <code>farmers</code> table.
surveyNumber	VARCHAR	NULL	Land parcel number.
farmType	VARCHAR	NULL	Farm type is to know if it is own or lease for cropping..
soilType	VARCHAR	NULL	Type of Soil like Black, Red, and Sandy.
landType	VARCHAR	NULL	Type of land based on irrigated or rainfed.
Area	DECIMAL	NOT NULL	Area of farm in acres.
latitude	DECIMAL	NULL	Latitude in degree decimals.
longitude	DECIMAL	NULL	Longitude in degree decimals.
villageId	INT	NOT NULL	A foreign key to link with <code>villages</code> table
status	INT	NOT NULL	1: active , 0 : inactive
userId	INT	NOT NULL	A foreign key to link with <code>users</code> table.

Table 21 Soil sample table

Attribute name	Data type	Constraints	Description
soilSampleId	INT	Primary Key, Auto Increment	Unique identifier for users.
projectId	INT	NOT NULL	A foreign key to link with project table
farmId	INT	NOT NULL	A foreign key to link with <code>farms</code> table.
farmerId	INT	NOT NULL	A foreign key to link with <code>farmers</code> table.
soilLabId	VARCHAR	NOT NULL	User-defined alpha-numeric code provided by laboratory.
previousCropId	VARCHAR	NOT NULL	A foreign key to link with <code>crops</code> table.
nextCropId	VARCHAR	NOT NULL	A foreign key to link with <code>crops</code> table.
soilType	VARCHAR	NOT NULL	Soil type
dateOfSampling	DATE	NOT NULL	Date of soil sampling.
latitude	DECIMAL	NULL	Latitude in degree decimals.
longitude	DECIMAL	NULL	Longitude in degree decimals.
status	INT	NOT NULL	1: active , 0 : inactive
userId	INT	NOT NULL	A foreign key to link with <code>users</code> table.

Table 22 Data table for soil samples lots.

Attribute name	Data type	Constraints	Description
LotId	INT	Primary Key, Auto Increment	Unique identifier for Lot.
LotNumber	VARCHAR	NOT NULL	A System Generator Lot Number
status	INT	NOT NULL	1: active , 0 : inactive
userid	INT	NOT NULL	A foreign key to link with <code>users</code> table.

Table 23 Mapping table between LOT and Soil Sample

Attribute name	Data type	Constraints	Description
LotMappingId	INT	Primary Key, Auto Increment	Unique identifier for Lot Mapping.
LotId	INT	NOT NULL	A foreign key to link with Lot table.
sampleId	INT	NOT NULL	A foreign key to link with Soil sample table.
status	INT	NOT NULL	1: active , 0 : inactive
userid	INT	NOT NULL	A foreign key to link with <code>users</code> table.

Table 24 Soil sample dispatch details

Attribute name	Data type	Constraints	Description
dispatchId	INT	Primary Key, Auto Increment	Unique identifier for dispatch.
SourceOfTransport	VARCHAR	NOT NULL	Source of transporting LOTs.
dateOfDispatch	DATE	NOT NULL	Date of Dispatching the LOTs
expectedDateOfDelivery	DATE	NOT NULL	Expected Date of Delivery to Lab
Message	VARCHAR	NOT NULL	User can provide any message regarding transport
status	INT	NOT NULL	1: active , 0 : inactive
userid	INT	NOT NULL	A foreign key to link with <code>users</code> table.

Table 25 Mapping Table between transport and LOTs

Attribute name	Data type	Constraints	Description
transportLotMappingId	INT	Primary Key, Auto Increment	Unique identifier for mapping.
transportId	INT	NOT NULL	A foreign key to link with transport table.
lotId	INT	NOT NULL	A foreign key to link with LOT table
status	INT	NOT NULL	1: active , 0 : inactive
userid	INT	NOT NULL	A foreign key to link with <code>users</code> table.

Table 26 Mapping Table between transport and reference photos

Attribute name	Data type	Constraints	Description
transportPhotoMappingId	INT	Primary Key, Auto Increment	Unique identifier for mapping.
transportId	INT	NOT NULL	A foreign key to link with transport table.
photo	INT	NOT NULL	A reference photo either vehicle or receipt in case of courier. Maximum of 3 Pics allowed. This is optional field.
status	INT	NOT NULL	1: active , 0 : inactive
userid	INT	NOT NULL	A foreign key to link with <code>users</code> table.

Table 27 Soil History Table

Attribute name	Data type	Constraints	Description
----------------	-----------	-------------	-------------

historyId	INT	Primary Key, Auto Increment	Unique identifier for mapping.
soilId	INT	NOT NULL	A foreign key to link with Soil sample table.
lotId	INT	NOT NULL	A foreign key to link with Lot table.
status	INT	NOT NULL	1: active , 0 : inactive
userId	INT	NOT NULL	A foreign key to link with <code>users</code> table.
updatedAt	Date	NOT NULL	Date of update

Table 288 Soil analysis data

Attribute name	Data type	Constraints	Description
soilAnalysisId	INT	Primary Key, Auto Increment	Unique identifier for users.
projectId	INT	NOT NULL	A foreign key to link with project table
soilId	INT	NOT NULL	A foreign key to link with <code>soilSample</code> table.
soilLabId	VARCHAR	NOT NULL	User-defined alpha-numeric code provided by laboratory.
analysisDate	DATE	NOT NULL	Date of soil analysis.
ph	DECIMAL	NULL	Soil pH
ec	DECIMAL	NULL	Electrical conductivity (mS)
oc	DECIMAL	NULL	Organic carbon (%)
p	DECIMAL	NULL	Available-phosphorous (mg/kg)
k	DECIMAL	NULL	Exchangeable potassium (mg/kg)
ca	DECIMAL	NULL	Exchangeable calcium (mg/kg)
mg	DECIMAL	NULL	Exchangeable magnesium (mg/kg)
S	DECIMAL	NULL	Available sulfur (mg/ka)
zn	DECIMAL	NULL	Available zinc (mg/kg)
b	DECIMAL	NULL	Available boron (mg/kg)
fe	DECIMAL	NULL	Available iron (mg/kg)
cu	DECIMAL	NULL	Available copper (mg/kg)
mn	DECIMAL	NULL	Available manganese (mg/kg)
Status	INT	NOT NULL	1: active , 0 : inactive
userId	INT	NOT NULL	A foreign key to link with <code>users</code> table.

Table 29 Table to store the information about digital soil map and model

Attribute name	Data type	Constraints	Description
soilMapId	INT	Primary Key, Auto Increment	Unique identifier for soil maps.
mapLevel	INT	NOT NULL	1:state, 2:district
mapLocationId	INT	NOT NULL	A foreign key for state or district.
mapPath	VARCHAR	NOT NULL	File system path to raster file
soilParameter	VARCHAR	NOT NULL	Name of soil parameter
mapDate	DATE	NOT NULL	Date of map generation.
userId	INT	NOT NULL	A foreign key to link with the <code>users</code> table.
rmse	DECIMAL	NOT NULL	Root mean squared error for the DSM
r2	DECIMAL	NOT NULL	R ² values for the DSM
lccc	DECIMAL	NOT NULL	Lin's concordance correlation coefficient for DSM
mse	DECIMA	NOT NULL	Mean squared error for DSM
mapModel	VARCHAR	NOT NULL	File system path to model file.



This publication has been prepared as an output of **CGIAR Research Initiative on [Digital Innovation](#)**, which researches pathways to accelerate the transformation towards sustainable and inclusive agrifood systems by generating research-based evidence and innovative digital solutions. This publication has not been independently peer-reviewed. Any opinions expressed here belong to the author(s) and are not necessarily representative of or endorsed by CGIAR. In line with principles defined in [CGIAR's Open and FAIR Data Assets Policy](#), this publication is available under a [CC BY 4.0](#) license. © The copyright of this publication is held by [IFPRI](#), in which the Initiative lead resides. We thank all funders who supported this research through their contributions to [CGIAR Trust Fund](#).