



SMIS USE CASE DOCUMENTATION

V1.0





TABLE OF CONTENTS

Revision History	3
Abbreviations	3
1. Introduction	4
1.1. Purpose	4
1.2. Scope	4
2. Overview	5
2.1. Actors	5
2.1.1. Operator	5
2.1.2. Administrator	6
2.2. Types of use cases	6
2.2.1. Exploratory	6
2.2.2. Analytical	7
2.2.3. Administrative	7
3. Use cases	7
3.1. Browse and export grower data	7
3.2. Browse and export experimental data	8
3.3. Browse and export literature data	9
3.4. View rule bases	9
3.5. Query rule base	10
3.6. Established queries	11
3.6.1. Assess compaction risk: an example of an established query	11
3.7. Import grower data	12
3.8. Extend grower data	12
3.9. Import experimental data	13
3.10. Import literature data	13



REVISION HISTORY

Version	Description	Date	Author
1.0	Original SMIS use case report. “Established query” (see section 3.6) use cases will be updated as their viability is tested and their interfaces are implemented.	9/10/2017	Tomasz Kurowski

ABBREVIATIONS

API	Application Programming Interface
NG	National Grid
OS	Ordnance Survey
REST	Representational State Transfer
SMIS	Soil Management Information System



1. INTRODUCTION

1.1. PURPOSE

The following document describes the 'use cases' defined for the SMIS Analytics Toolkit software developed as part of the SMIS project (AHDB CP107D). The document's primary purpose is to inform the end user about the defined uses of the system and the interaction flows required to achieve particular goals within its scope. On the developer side, the document will also serve as an implementation guide for the Analytics Toolkit, in particular for the design, development and installation of its front-end interface.

Additionally, as the inclusion of some use cases (see Section 3.6) will depend on validation of their viability during the remaining duration of the project, future versions of this document will effectively catalogue any additional use cases which are found to be viable and will therefore be included in the SMIS Analytics Toolkit.

1.2. SCOPE

This document describes the SMIS use cases without specifying implementation details beyond the overall flow of actor interactions the system is designed to allow. In particular, this means that both the back-end database, analytics and API functionalities are not discussed, and front-end interface details are not presented. Data gathering and manual curation activities undertaken within the scope of the SMIS project are also not discussed.

However, as the delineation of user roles (i.e. Operator and Administrator actors), as well as the classification of use cases used in this document result directly from certain system design and implementation decisions, a brief overview of the SMIS software design is provided in Section 2.

Additionally, a certain class of use case, the Established Query (described in Section 3.6) effectively consists of special cases of a more general use case, each intended to be used with its own custom interface and a more narrowly defined set of goals and user inputs. The list of Established Queries provided in the current version of this document is not intended to be exhaustive. These use cases and their associated interfaces will continue to be added to during the duration of the project as the viability (which is strongly dependent on gathered data) of various queries is verified and their usefulness validated. The document will be updated to include them as development continues.



2. OVERVIEW

The SMIS Analytics Toolkit software system can be considered to consist of several subsystems:

- Parsing suite¹ – a set of command-line tools developed for the purpose of importing external datasets (e.g. Grower Data, Experimental / Research Project Data, Literature Data) into the SMIS database, capturing their internal relationships and ensuring their integrity. This component interfaces with the database back-end directly and is the primary route of interacting with the system for Administrator actors.
- Database back-end¹ – a non-relational database responsible for storing the parsed data gathered within the scope of the SMIS project, as well as the results of analyses conducted by the Analytics back-end. This will be available to the front-end and Analytics back-end through a REST API.
- Analytics back-end – a set of scientific computing tools developed for generating analyses and visualisations for the SMIS system. Other than relatively trivial analyses like summarising available data, this subsystem primarily focuses on machine learning algorithms designed to identify relationships within the gathered datasets or their subsets. This is available to the front-end through a REST API.
- Web application front-end – a web application which provides the primary interface for the delivery of the main SMIS use cases. This accepts user input and accesses the Database back-end and Analytics back-end through a REST API in order to present (or generate) results of user queries. This is the primary route for interacting with the system for Operator actors, and as such it represents the main SMIS Analytics Toolkit user interface in general.

2.1. ACTORS

2.1.1. OPERATOR

The 'Operator' actor is the primary user role defined for the SMIS Analytics Toolkit. The role represents users whose primary goal is the exploration and extraction of useful information from the SMIS database. This can consist of browsing and searching through the raw (and curated) datasets, visualising their contents, exploring "rule bases" which represent the information derived from the stored datasets using machine learning, constructing queries which can be used to generate more

¹ See the *SMIS Database Architecture Technical Documentation* for details.



specific rule bases, or using pre-defined “Established Queries” to explore problems which have been identified as being both useful and possible to address, based on the data collected during the course of the SMIS project.

2.1.2. ADMINISTRATOR

The ‘Administrator’ actor is a user role responsible for populating the SMIS database. During the development stage, this role can therefore be fully identified as the developer. However, the data parsing suite is intended to be generic and well-documented, capable of being used to update and expand the SMIS database beyond the scope of initial development, if further data is acquired. At the same time, the wholly separate interface (command-line scripts on the SMIS server) and very different input and server access requirements clearly delineate the role of an Administrator as separate from a “normal” user or Operator.

It should be noted that the SMIS database is generally intended for periodic, rather than continuous, updates. This is because the addition of new data may invalidate the results of previous analyses, requiring them to be re-computed. This means that the Administrator actor is intended to be active relatively infrequently.

An additional point is that as the SMIS Analytics Toolkit has been designed with the assumption that it will be used internally within an organisation rather than made directly available to the public, certain functionalities which would normally be assumed to depend on an “Administrator” role (such as user account management or access control), are not part of the design and are therefore not covered by this document.

2.2. TYPES OF USE CASES

Three types of use cases have been defined. The first two (Exploratory and Analytical use cases) overlap to an extent, due to being available through a common Web interface and used by the same actor.

2.2.1. EXPLORATORY

Exploratory use cases are ones which involve the use of the SMIS Analytics Toolkit Web application to browse raw and curated datasets stored in the SMIS database. They do not lead to the generation or storage of new information and do not involve the use of the Analytics back-end.



2.2.2. ANALYTICAL

Analytical use cases are ones which involve the use of the SMIS Analytics Toolkit Web application to view summaries and visualisations of datasets stored in the SMIS database, as well as to create queries which result in the creation of novel rule bases. These use cases employ the Analytics back-end to generate summaries, visualisations and machine learning models and store this new data in the database. The primary purpose of storing the results is to avoid the need to re-compute results for repeated queries, allowing for easy exploration of previously generated rule bases.

2.2.3. ADMINISTRATIVE

Administrative use cases are ones which involve the parsing of novel data, resulting in either the addition of whole new datasets, or appending new data to existing ones. These cases involve the use of the SMIS parsing suite by an Administrator.

3. USE CASES

3.1. BROWSE AND EXPORT GROWER DATA

Primary Actor: Operator

Type: Exploratory, Analytical (extensions)

Description: The Operator browses the SMIS Grower datasets and filters them by an arbitrary selection of the available data fields. The contents of the database for a selected set of conditions can be inspected and visualised, allowing the Operator to assess the size and potential usefulness of the gathered data for a given set of conditions. The visualisations and raw data can be exported.

Main success scenario:

1. The Operator accesses the SMIS Analytics Toolkit application through a Web browser.
2. The Operator navigates to the Grower data view.
3. A paginated table containing grower data is displayed.
4. The Operator selects (for qualitative fields) or types search terms for each column of interest.
5. Database contents are filtered and displayed in a new paginated table.
6. The Operator inspects the filtered data.
7. (Optional) The Operator presses a button to export the filtered data as a spreadsheet and the generated visualisations (if any) as a PDF report.

Extensions:

- 4a. The Operator can hide extraneous or empty table columns for clarity.
- 6a. **(Analytical)** Timeline view.



6a1. The Operator selects the Timeline view option.

6a2. A paginated list of farm fields and associated scrollable timelines of filtered field operations and applications are displayed in a rotational context. Field operations are coloured depending on an assessment as to whether they were undertaken in or outside of LandIS generated workability days.

6b. **(Analytical)** Summary view.

6b1. The Operator selects the Summary view option.

6b2. A summary report describing the filtered data is displayed, including the dates for which data is available, soil types and crop hectarages (including diagrams showing the relative hectarages of various crops and varieties).

6c. By selecting a data row, the Operator may navigate to a **Rule Base view** (see Section 3.4.) filtered to display rule bases which make use of this data point.

3.2. BROWSE AND EXPORT EXPERIMENTAL DATA

Primary Actor: Operator

Type: Exploratory

Description: The Operator browses the SMIS Experimental datasets, selecting the specific Experimental datasets to be viewed and inspecting the collected data in a tabular form, with the option to filter it by an arbitrary list of conditions and, optionally, export any filtered data subset.

Main success scenario:

1. The Operator accesses the SMIS Analytics Toolkit application through a Web browser.
2. The Operator navigates to the Experimental data view.
3. A list of available Experimental datasets is displayed along with short summaries of the size and content of each.
4. The Operator selects an Experimental dataset to view.
5. A paginated table containing Experimental data is displayed.
6. The Operator selects (for qualitative fields) or types search terms for each column of interest.
7. Database contents are filtered and displayed in a new paginated table.
8. The Operator inspects the filtered data.
9. (Optional) The Operator presses a button to export the filtered data as a spreadsheet.

Extensions:

4a. Alternatively, the Operator may navigate to a **Rule Base view** (see section 3.4.) filtered to display rule bases which make use of this experimental data set.

5a. The Operator can hide extraneous or empty table columns for clarity.

9a. By selecting a data row, the Operator may navigate to a **Rule Base view** (see section 3.4.) filtered to display rule bases which make use of this data point.



3.3. BROWSE AND EXPORT LITERATURE DATA

Primary Actor: Operator

Type: Exploratory

Description: The Operator browses the SMIS Literature datasets, viewing and inspecting the manually curated data summaries in a tabular form, with the option to filter them by an arbitrary list of conditions and, optionally, export any filtered data subset.

Main success scenario:

1. The Operator accesses the SMIS Analytics Toolkit application through a Web browser.
2. The Operator navigates to the Literature data view.
3. A paginated table containing manually curated summaries of Literature data is displayed.
4. The Operator selects/enters search terms for each column of interest.
5. Database contents are filtered and displayed in a new paginated table.
6. The Operator inspects the filtered data.
7. (Optional) The Operator presses a button to export the filtered data as a spreadsheet.

Extensions:

- 4a. The Operator can hide extraneous or empty table columns for clarity.
- 6a. By selecting a literature entry, the Operator may navigate to a **Rule Base view** (see section 3.4.) filtered to display rule bases which make use of this literature entry.

3.4. VIEW RULE BASES

Primary Actor: Operator

Type: Exploratory

Description: The Operator can view the soil management and yield “rule bases” derived from the SMIS database using machine learning analyses. These rule bases can be viewed either in tabular form, as lists of issues vs. causes/solutions (for soil management rule bases) or yield vs. factors-affecting-yield (for yield rule bases) pairs along with their associated weights, or in graph form, rendering the same pairwise relationships in a graphical manner.

Main success scenario:

1. The Operator accesses the SMIS Analytics Toolkit application through a Web browser.
2. The Operator navigates to the Rule Base view.
3. The Operator chooses to view either Soil Management Rule Bases or Yield Rule Bases.
4. A paginated list of available Soil Management / Yield rule bases is displayed alongside queries used to generate them. Rule bases still in the process of being generated are greyed out and display a progress bar with an estimate of time to completion.
5. The Operator selects a rule base to view.



6. A graphical view of the rule base is displayed. This consists of a set of nodes and edges connecting them. The nodes represent soil management issues, causes and solutions (for soil management rule bases) or crop yield and factors affecting crop yield (for yield rule bases). Edges connecting the nodes represent relationships identified based on the SMIS database contents according to a specific query. The thickness of edges depends on the strength and the degree of confidence of the relationship (weight).
7. The Operator inspects specific relationships by clicking on edges which display the list of Experimental, Grower and Literature evidence for the relationship along with their associated weights in absolute and relative terms.
8. (Optional) The Operator presses a button to export the entire rule base (or a selected set of nodes/edges) in tabular form as a spreadsheet.

Extensions:

6a. Tabular view.

6a1. The Operator selects the Tabular view option.

6a2. A paginated table containing a list of issues vs. causes/solutions (for soil management rule bases) or yield vs. factor-affecting-yield (for yield rule bases) pairs along with their associated weights is displayed.

6a3. The Operator types in search terms for each column of interest in the table.

6a4. The rule base contents are filtered and displayed in a new paginated table.

7a. By selecting a piece of Experimental, Grower or Literature evidence from the list, the Operator may navigate to their respective database browse views (see sections 3.1, 3.2, and 3.3), filtered to display the selected entry.

3.5. QUERY RULE BASE

Primary Actor: Operator

Type: Analytical

Description: The Operator constructs a query used to select data used by a machine learning algorithm to generate a novel, specific rule base.

Main success scenario:

1. The Operator accesses the SMIS Analytics Toolkit application through a Web browser.
2. The Operator navigates to the Rule Base view.
3. The Operator chooses to view either Soil Management Rule Bases or Yield Rule Bases.
4. The Operator constructs a query by selecting available categories (data fields from Grower and Experimental data and curated Literature keywords), optionally combining them with a manually entered search term. A query can contain an arbitrary number of such category/search term pairs.
5. The Operator selects the "Run Query" option.



6. Database entries are filtered according to the query and the Analytics back-end begins the process of generating a rule base based on the filtered data.
7. The Operator is redirected to the rule base view (see section 3.4), with the new query and its progress bar highlighted.

Extensions:

6a. If no data stored by SMIS matches the entered query the Operator is informed of this by a pop-up message and the query is aborted.

6b. If a rule base related to the same or equivalent query already exists in the database, the Operator is informed of this by a pop-up message and navigates directly to a view of that rule base (see section 3.4).

3.6. ESTABLISHED QUERIES

Established Queries can be considered “special cases” of rule base queries as used in Section 3.5. They rely on the same database and Analytics capabilities, but instead of allowing for the construction of arbitrary queries, they have their particular input and output interfaces, each of which are tailored to only one very specific query, selected during the SMIS project based on their potential usefulness and the confirmed availability of necessary data. One example is the assessment of soil compaction risk.

3.6.1. ASSESS SOIL COMPACTION RISK: AN EXAMPLE OF AN ESTABLISHED QUERY

Primary Actor: Operator

Type: Analytical

Description: The Operator assesses the soil compaction risk for a specific field (identified by NG code / OS Map Sheet) at a particular range of dates, optionally associated with a specific crop.

Main success scenario:

1. The Operator accesses the SMIS Analytics Toolkit application through a Web browser.
2. The Operator navigates to the Established Queries view.
3. The Operator selects the Assess Soil Compaction Risk query.
4. The Operator enters the NG Code and Map Sheet for the field of interest.
5. The low-resolution LandIS data stored by SMIS is accessed to provide the most likely soil type for the field. This is displayed in a drop-down menu and can be changed manually.
6. The Operator enters a start and end date for the Soil Compaction Risk assessment and (optionally) selects the crop of interest from a drop-down menu.
7. A colour-coded timeline of workability days based on soil type and weather data is displayed and inspected by the Operator.

Extensions:



7a. By selecting the assessment results, the Operator can navigate to the associated rule base (see Section 3.4) and view possible solutions to soil compaction, if any were identified.

3.7. IMPORT GROWER DATA

Primary Actor: Administrator

Type: Administrative

Description: The Administrator imports standard Grower datasets (e.g. Gatekeeper XML / spreadsheet format) into the SMIS database.

Main success scenario:

1. The Administrator accesses the SMIS server through a command line interface and launches the Grower Data Import script.
2. The Administrator selects the option to import a new dataset and provides a path to a Grower data file.
3. The Grower data is imported on a row-by-row basis. Novel data fields or qualitative field values require confirmation, modification or rejection by the user (see the *SMIS Database Architecture Technical Documentation* for details).
4. A data import report is displayed and can be saved as a text file.

3.8. EXTEND GROWER DATA

Primary Actor: Administrator

Type: Administrative

Description: The Administrator imports tabular Grower data not stored in a standard format such as Gatekeeper XML. Such data needs to be manually associated with datasets already present in the SMIS datasets and serve to add extra fields not covered by Gatekeeper datasets imported earlier.

Main success scenario:

1. The Administrator accesses the SMIS server through a command line interface and launches the Grower Data Import script.
2. The Administrator selects the option to expand an existing dataset and provides a path to a Grower data file.
3. Data fields present among Grower datasets stored in the SMIS database, as well as those identified in the provided file, are listed.
4. The Administrator selects one or more pairs of data fields to be considered equivalents to be used for joining the new dataset with data present in the database.
5. The Grower data is imported on a row-by-row basis. Novel data fields or qualitative field values require confirmation, modification or rejection by the user (see the *SMIS Database Architecture Technical Documentation* for details).
6. A data import report is displayed and can be saved as a text file.



3.9. IMPORT EXPERIMENTAL DATA

Primary Actor: Administrator

Type: Administrative

Description: The Administrator imports novel Experimental datasets into the SMIS database using project-specific parsing modules grouped under a common interface.

Main success scenario:

1. The Administrator accesses the SMIS server through a command line interface and launches the Experimental Data Import script.
2. The Administrator provides a path to a data file.
3. The Experimental data is imported by a project-specific parser. Unexpected data fields or field values may require confirmation, modification or rejection by the user (see the *SMIS Database Architecture Technical Documentation* for details).
4. A data import report is displayed and can be saved as a text file.

3.10. IMPORT LITERATURE DATA

Primary Actor: Administrator

Type: Administrative

Description: The Administrator imports spreadsheets containing manually curated Literature data gathered during the SMIS project.

Main success scenario:

1. The Administrator accesses the SMIS server through a command line interface and launches the Literature Data Import script.
2. The Administrator provides a path to a data file.
3. The Experimental data is imported on a row-by-row basis
4. A data import report is displayed and can be saved as a text file.