

Foundation Infrastructure for Quality Assurance of Soil Management (SIKB)

Protocol for the digital exchange of soil data

SIKB 0101

**version 10.2.0 [English translation]
SIKB Central Board of Experts**

**PLEASE NOTE: this version is for information purposes only.
The Dutch version of this text is, with BRL SIKB 0100, the basis for certification of
software, indicating the ability of the software to exchange data through the
XML-format of this protocol. The XML-format is available in Dutch only.**

Contents

1	SCOPE.....	3
1.1	Background and objectives.....	3
1.2	Data exchange between systems.....	4
1.3	Scope.....	5
1.4	Processes and data sets.....	5
2	BASIC DATA EXCHANGE PREMISES.....	7
2.1	Site definition and contours.....	7
2.2	Data ownership and coding.....	8
2.3	Expired codes.....	9
2.4	Method for reading a Labeldelivery file.....	9
3	DESCRIPTION OF THE DATA MODEL.....	10
4	REFERENCES.....	11
	APPENDIX 1: CONVENTION FOR THE DESCRIPTION OF COMPONENTS.....	12

1 Scope

1.1 Background and objectives

Digital soil quality data is widely used in the Netherlands. This includes field data (e.g. drilling logs), laboratory analysis results, historical and general data about locations, progress data about current and past soil surveys and remediation projects, etc. This data is already being exchanged in digital form to a greater or lesser extent. Data is exchanged between:

- consultants and laboratories, for defining and issuing lab orders;
- laboratories and consultants, for reporting on the results of analyses;
- consultants and their clients (government bodies, businesses and individuals);
- data managers.

Here, 'consultants' includes other organisations undertaking investigations, such as municipal engineering departments, etc. Similarly, clients may receive laboratory results directly, without an intermediary. However, the principle of the exchange remains the same.

Unambiguous digital data exchange will also improve data management at the site or organisation level. Examples include industrial estates (site management), and organisations responsible for a number of sites such as SBNS, NAM, etc.

The structure of the protocol is such that all data exchange needs can be met with a uniform XML protocol. The protocol was designed to be flexible enough to allow future development. The composition of the data set which can be exchanged is determined by the contents of the relevant systems. This project specifically did not address what data should be exchanged.

The objective of the protocol is to provide a standard exchange format for digital soil data for the sector as a whole, to prevent errors due the inability to exchange digital data unambiguously.

1.2 Data exchange between systems

A major principle is that the data is exchanged between systems containing data. In the Netherlands several system levels are distinguished:

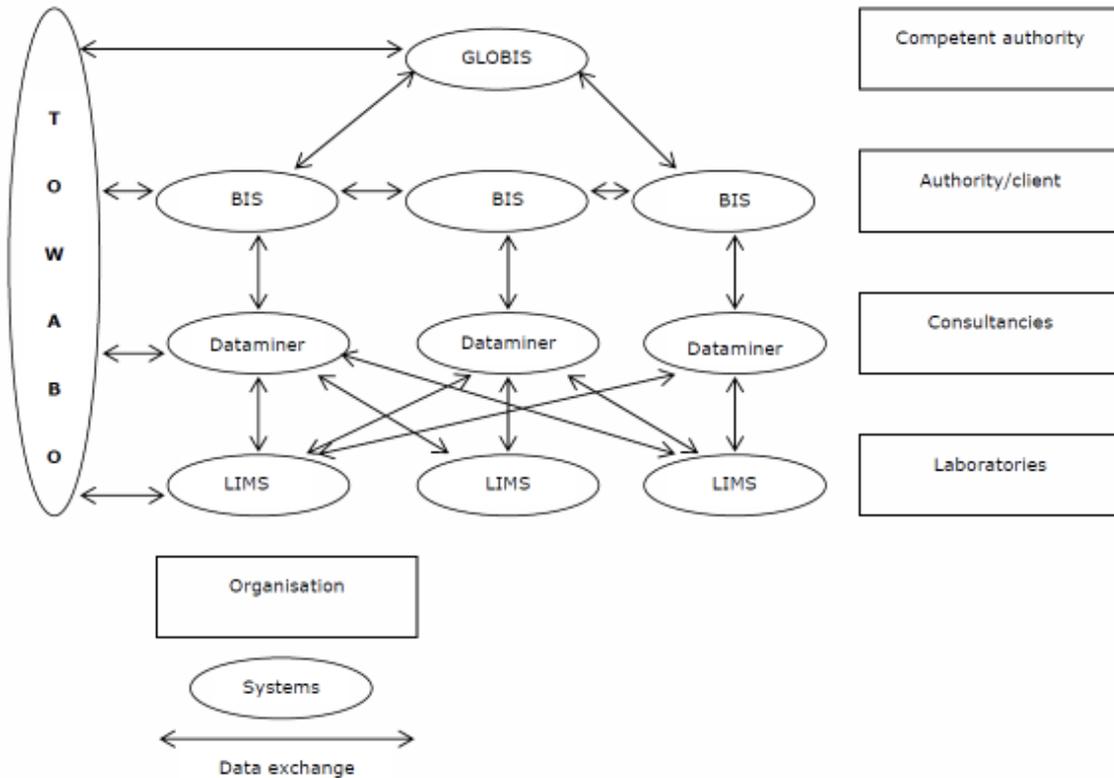
- laboratory systems;
- data miners;
- local and regional soil information systems (Dutch abbreviation: BIS);
- national soil information system (GLOBIS);
- watercourse sediment assessment system (Towabo).

The laboratory systems primarily have an internal role within the laboratory concerned. External data exchange occurs when a laboratory receives an order and when it reports the results of the analyses. Data miners are systems primarily used by consultants when they send orders to laboratories, process field data, and process and interpret the results of the analyses, possibly together with site data. Local and regional government bodies, as well some larger private sector organisations have soil information systems. These systems largely store data at the site and project level. Data is stored at different levels, but all systems always contain various elements of the site data, analysis data for projects and sometimes drilling data. The data normally has a geographical component. Finally, in the Netherlands, there is GLOBIS, the nationwide information system which primarily contains aggregated data about the contamination status and progress of any soil remediation efforts. GLOBIS does not contain data at the analysis or bore level. The system is used by the competent authorities under the Soil Protection Act, however it is not used by Rijkswaterstaat (Directorate-General of Transport, Public Works and Water Management) for national watercourses.

Instead, Towabo is used for the collection and assessment of watercourse sediment data.

The protocol aims to standardise the digital exchange of data. This will greatly improve the opportunities to exchange data in the digital domain. It should be emphasised that the protocol does not refer to the composition and completeness of the data set, or the quality of the data.

The relationships are shown in the diagram below:



1.3 Scope

The draft protocol addresses the exchange of data concerning the environmental quality of soils or sediments. With respect to sediment, the protocol is limited to data exchange with TOWABO. Once there is greater standardisation in the field of watercourse sediments, the protocol will be extended appropriately.

1.4 Processes and data sets

The SIKB protocol can support a number of processes which are associated with specific data sets. The processes to be supported include:

- orders sent to laboratories;
- exchange of the analysis results between laboratories and consultants/clients;
- data exchange further to national policies;
- monitoring soil remediation programmes;
- importing data into municipal soil information systems, for example to produce soil quality maps.

A number of different data sets are exchanged between collecting the soil sample and publishing the monitoring report. We can distinguish:

- results of field work (drilling logs, raw data);
- laboratory orders (analysis results, raw data);
- consultant's reports (basic data, raw data);
- site data (decisions by the authorities, monitoring, etc., interpreted data);
- geographical data and related information (e.g. contours on a soil quality map and the relevant percentile values).

2 Basic data exchange premises

To exchange soil data we have to know which site the data relates to, what contours (isocontours) can be exchanged, who owns the data, and which data set we are dealing with.

2.1 Site definition and contours

For the definition of the site and contours to be exchanged this protocol follows the definition and description given in the **Handboek Monitoring Bodemsanering** (Soil remediation monitoring handbook) of November 2001:

A site is a unit (or aggregation of basic units) with suspected or proven soil contamination, which is dealt with as a single entity.

The following basic units can be distinguished:

1. Contours of suspect sites, point sources and linear sources. This may be an address, a line (backfilled drainage ditch) or an area (landfill determined by aerial photography).
2. Contour of the area under investigation (e.g. the business estate which was investigated), this is a comprehensive contour based on one or more reports, which the conclusions relate to.
3. Contamination contour (intervention limit contour).
4. Remediation contour.

The site always consists of one or more of these units and is delineated by the outer contour. The contour should be related to the status fields concerning the investigation and remediation. This will normally correspond to the contour of the area under investigation (2).

The protocol includes the following geo-objects for the data exchange:

- geo-object site;
- geo-object suspect activity;
- geo-object investigation;
- geo-object contamination contour;
- geo-object remediation contour;
- geo-object borehole.

Note: in XML spatial information is stored in a GeoObject.

2.2 Data ownership and coding

It is relatively straightforward to add soil investigation data for a new site to an existing data set. However, it is more difficult to update and supplement existing data. This requires agreements about who owns the data (so that data cannot be modified at several locations) and about the coding of the data (so that data can be updated and linked). This is relevant to systems where the submitted data may have to be updated in batches in the future, without data loss, and for systems which want to transfer the submitted data, including their origin.

The premise behind coding soil investigation data is that it is not feasible to introduce a unified national coding system. To be able to exchange data, systems have to be aware of the internal codes of the systems they are exchanging data with (similar to MDB). This is relevant to updates of modified data and the addition of new data. Initially this information will be contained in the metadata. Where necessary, the various soil information systems will have to add a field to each table for the internal code of the exporting (source) system. This code is a combination of the origin (code indicating the municipality or province), system (Glo, Str, Naz, B4A, etc.) and an internal code. It will have to be determined for each system if it should exchange the internal codes for all tables or if just the site and report codes would suffice.

First, one must agree about the location to which supplied data has to be added. If this is done, data can be added. How to deal with new data is described on the following web page:

<http://www.bodemloket.info/lib/downloads/uitgangspuntennotitie.doc>

It has to be agreed upon who is authorised to modify data. We propose that only the data owner is authorised to modify it, to prevent data being changed in two locations.

This approach is effective for reporting and analysis data. However, a problem may arise with site data as several sources may supply digital reports and conclusions at the site level for a single site. At the site level it will only be possible to exchange digital data fully automatically for new sites. Bilateral agreements would have to be concluded for any updates to data.

Coding and exchange:

- All BIS soil information systems and GLOBIS will continue to use their own internal codes.
- When data is exchanged, a code is added to identify the municipality, province or watercourse manager, as well as a code to identify the system (Glo, Str, Naz, B4A, etc.), when combined they should form a unique code.

- Where necessary, each system adds a field at the table level to identify the owner of the data. This may present a problem at the site level (1:N if there is more than one data supplier).
- Only data owners are authorised to modify data.
- It will always have to be agreed if data is added to an existing site or added as a new site. Where possible this should be determined using the GIS.

Agreements about the ownership of the data will have to be made for data exchange processes where the origin of the data has to remain traceable and where it may not be modified. This will have to be indicated in additional fields in the soil information systems. Only the data owner is authorised to modify these fields.

2.3 Expired codes

At the moment that in a particular version of the exchange protocol SIKB0101 a value in the domain table has received the status "Cancelled", that value from that version may not be exchanged between applications.

In the domain table this value remains visible, so the history is preserved.

When changes are approved, granting the status "Cancelled" of a domain value, the applicant should indicate what the replacement value will be or that the domain value has no substitute. The end user or software vendor should take precautions in his application to prevent that those expired values can be selected or reside in the database.

This prevents that in an exchange of soil data it keeps necessary to maintain "Expired" domain values. New participants do not need to look further into the history. In addition, this methods does not frustrate any further coordination with other domain tables of information models.

2.4 Method for reading a Labdelivery file

Before reading a new labdelivery file, the receiving system must first remove all data belonging to the client id's of the new labdelivery file, from the tables where storage of labdelivery data takes place.

3 Description of the data model

Until version 6.1.0 this document contained also a global description of the different datasets. From 7.1.0 these are available as HTML-pages. At the time of the release of version 7.1.0, these can be found on the website of www.sikb.nl: Datauitwisseling – Richtlijnen en Protocollen – BRL SIKB0100.

There you can find a global description in HTML-format of:

- soil data (Protocol (XSD))
- notification data (Melding (XSD))
- laboratory delivery data (Labaanlevering (XSD))
- laboratory assignments (Labopdracht (XSD))
- laboratory results (Labresultaat (XSD))

For the content of the data tables refer to the Lookup XML and XSD file.

The changes for each version of the different XSD's since version 1.0, can be found in document "Historie SIKB0101.doc"

The changes for each version of the lookup-XML can be found in document "Historie codetabellen SIKB0101.doc".

Appendix 1 lists all code tables. For the contents of the tables you are referred to the Lookup-XML and XSD files.

4 References

- Handboek Monitoring Bodemsanering: Afspraken over indicatoren en gegevens, def. versie, oktober 2001
- Concept Circulaire Landsdekkend beeld, Fase 1, Nulmeting Werkvoorraad landbodems bodemverontreiniging (versie 4.01, oktober 2001)
- GLOBIS versie 1.1.1, oktober 2001
- VKB protocol 2017 van SIKB
- NEN 5104
- Datamodel Miras BST, versie 1.9, oktober 2001
- Datamodel Boormanager 4.0, oktober 2001
- Implementatie Laboverleg, versie 1.2, 21-09-2001
- Datamodel Strabis 9.3.5, oktober 2001
- Datamodel Nazca 3.0, november 2000
- Datamodel Bis4all, versie 4.1.006, 5-10-2000
- Resultaten Laboverleg versie 2.0 30-10-2002

Appendix 1: Convention for the description of components

1. Component names start with capital letters, e.g. Dichloromethane. Suggestion: also use capital letters for appended trivial names, see item 4.
2. Write abbreviations in capitals without full stops, e.g. DDT.
3. Spell Greek letters out in full, in lower case e.g. alpha-Endosulfan. Suggestion: also use lower case letters for other prefixes, e.g. iso, ortho, trans, etc.
4. Add common trivial names in brackets e.g. Trichloromethane (Chloroform).
5. Components with multiple substituents should be unambiguously defined, e.g. 1,1,1-Trichloromethane.
6. Use numerals for substituents e.g. 4,4-DDT, however if letters are also commonly used then they should be appended, in brackets e.g. 4,4-DDT (para,para-DDT).
7. Place suffixes after the component name in brackets, e.g. Asbestos (brown) and Cyanide (total). Suggestion: separate multiple suffixes by commas, in brackets, e.g. Asbestos (brown, Amosite).
8. Write elements in full, place abbreviations from the periodic table in square brackets, e.g. Antimony [Sb].
9. Abbreviations with numbers should be unambiguous, e.g. PCB-101.
10. Where appropriate, indicate what the concentration is based on, e.g. Phosphate (as PO4).
11. Indicate grain size fractions unambiguously, e.g.: fraction < 2 um.
12. Indicate column leaching tests unambiguously: column leaching test L/S=1.
13. Indicate BOD and number of days as follows: BOD-5.
14. Indicate mineral oil fractions as follows: Mineral oil C6-C10.
15. Add 'total' in brackets to indicate an addition if it is clear which components are included, e.g. Dichlorobenzenes (total). Example Dichlorobenzenes (total), includes 1,2-, 1,2- [translator's note: presumably '1,2-, 1,3-'] and 1,4-dichlorobenzene. Add 'total' in brackets followed by details to indicate an addition if it is not immediately clear which components are included, use '+' as separators. Example: Dichlorobenzenes (total 1,2 + 1,4) includes 1,2- and 1,4-dichlorobenzene, but not 1,3-dichlorobenzene. Add 'total' in brackets to indicate an addition if two components with different names have been determined together, use '+' as separators. Example: alpha-Endosulfan + Endosulfansulphate (total).
16. Specify Organic matter and Argillaceous material separately.
17. Place suffixes in brackets in the description when dealing with different components since different analyses are carried out. Example: Ammonium (as N) or Ammonium (as NH4).
18. If there are several analytical methods available for a components, but they all essentially determine the same, then one description should be selected, without specifying the method. In this case, components determined using different methods are distinguished by the analytical technique field.
19. Exception: Oil (IR).
20. Oil is reported including the fraction, however preliminary treatment (e.g. with Florisil) forms part of the analytical technique part and does not require a separate component.

STICHTING INFRASTRUCTUUR KWALITEITSBORGING BODEMBEHEER (SIKB)

21. Information provided in the report which does not concern components is included in a separate category 'Report' (insert R in front of the code.). e.g.: R4, R107, Cascade tests etc.
22. Specify grain fractions as follows: Fraction < 2 um; Fraction 1000 - 2000 um.