

TEXTE

104/2022

Integration von im UBA vorhandenen Bodendaten in eine Fachanwendung zur bundesweiten Auswertung

Abschlussbericht

TEXTE 104/2022

Ressortforschungsplan des Bundesministerium für Umwelt,
Naturschutz, nukleare Sicherheit und Verbraucherschutz

Forschungskennzahl 3717 74 252 0

FB000696

Integration von im UBA vorhandenen Bodendaten in eine Fachanwendung zur bundesweiten Auswertung

Abschlussbericht

von

Matthias Müller
PIKobytes GmbH, Dresden

Im Auftrag des Umweltbundesamtes

Impressum

Herausgeber

Umweltbundesamt
Wörlitzer Platz 1
06844 Dessau-Roßlau
Tel: +49 340-2103-0
Fax: +49 340-2103-2285
buergerservice@uba.de
Internet: www.umweltbundesamt.de

 [/umweltbundesamt.de](https://www.facebook.com/umweltbundesamt.de)

 [/umweltbundesamt](https://twitter.com/umweltbundesamt)

Durchführung der Studie:

PIKobytes GmbH
Spenerstraße 14
01309 Dresden

Abschlussdatum:

Februar 2021

Redaktion:

Fachgebiet II 2.7 Bodenzustand, Bodenmonitoring
Dr. Marc Marx

Publikationen als pdf:

<http://www.umweltbundesamt.de/publikationen>

ISSN 1862-4804

Dessau-Roßlau, September 2022

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen und Autoren.

Kurzbeschreibung: Integration von im UBA vorhandenen Bodendaten in eine Fachanwendung zur bundesweiten Auswertung

Das Umweltbundesamt führt Bodeninformationen zum chemischen, physikalischen und biologischen Zustand der Böden in Deutschland. Insbesondere aus Sicht des Bodenschutzes sind diese Informationen von großem Interesse und es ist für viele Fragestellungen von großer Bedeutung, auf möglichst viele dieser Datenbestände für Auswertezwecke zurückgreifen zu können. Die zusammengesetzte Auswertung ist wichtig, weil dadurch zeitlich, räumlich und stofflich differenzierte Aussagen getroffen werden können, die aus den einzelnen Datenbanken nicht ableitbar sind. Diese differenzierten Aussagen sind notwendig, um geeignete Maßnahmen zur Wiederherstellung bzw. Erhaltung der Bodenfunktionen zu ermöglichen. Viele der hausintern vorhandenen Datensätze liegen jedoch in unterschiedlichen Systemen, Datenbanken und Dateien vor und sind untereinander nicht kompatibel, sodass zunächst keine effiziente quellübergreifende Datenauswertung möglich ist.

Gegenstand dieses Projektes ist daher die Zusammenführung einzelner Datenbestände, um künftig eine integrierte Auswertung zu ermöglichen. Hierfür wurden zunächst acht Datenbanken unterschiedlicher Größe, Struktur und Aktualisierungshäufigkeit hinsichtlich ihrer Zusammenführbarkeit untersucht. In einem zweiten Schritt wurden Ableitungsregeln definiert, die die Transformation der einzelnen Datenbestände in ein einheitliches Zielschema beschreiben. In einem dritten Schritt wurde ein Softwarewerkzeug erstellt, das aus diesen Regeln einen integrierten Datenbestand erzeugt.

Abstract: Integration of UBA soil data into a specialized application for nationwide evaluation

The Federal Environment Agency maintains soil information on the chemical, physical and biological condition of soils in Germany. Especially from the perspective of soil protection, this information is of great interest, and it is of great importance for many questions to be able to fall back on as many of these data stocks as possible for evaluation purposes. The composite evaluation is important because it enables statements to be made that are differentiated in terms of time, space and substance, which cannot be derived from the individual databases in this way, and thus enables appropriate measures to be taken to restore or maintain soil functions, respectively. However, many of the in-house data sets are available in different information systems, databases and files and are not compatible with each other, so that initially no efficient cross-source data evaluation is possible.

The aim of this project is therefore to merge individual data sets to enable integrated evaluation in the future. For this purpose, eight databases of different size, structure and update frequency were first examined regarding their mergeability. In a second step, derivation rules were defined that describe the transformation of the individual databases into a uniform target schema. In a third step, a software tool was created that generates an integrated dataset from these rules.

Inhaltsverzeichnis

Abbildungsverzeichnis.....	9
Tabellenverzeichnis	9
Abkürzungsverzeichnis.....	10
Zusammenfassung.....	11
Summary	13
1 Allgemeine Bestandsaufnahme.....	14
1.1 Projektanlass.....	14
1.2 Untersuchte Datenbanken.....	14
1.3 Kategorisierung der Datenbanken	15
1.3.1 Standortscharfe Probedatenbanken	16
1.3.2 Flächenhafte Bodendatenbanken.....	16
1.3.3 Toxizitäts- und Stoffdatenbanken.....	16
1.4 Festlegungen zur Schemadokumentation	17
2 Dokumentation der Datenbankschemata	18
2.1 Bodendatenbank eBIS.....	18
2.2 Dioxindatenbank.....	18
2.3 Bodendaten aus den bundesweiten Auen.....	19
2.4 Sonstige – Bodenzustandserhebung Wald II	19
2.5 Organische und anorganische Hintergrundgehalte	19
2.6 Umweltprobenbank des Bundes.....	20
2.7 Sonstige – Sickerwasser	21
2.8 Bodendaten aus der Klimafolgenforschung / Erosionsdaten	22
2.9 Stoffdatenbank STARS.....	23
2.10 Boden – Bodenorganismen.....	23
3 Vorgehensmodell zur Datenbankintegration.....	24
3.1 Standortscharfe Probedatenbanken.....	24
3.2 Flächenhafte Bodendatenbanken.....	25
3.3 Toxizitäts- und Stoffdatenbanken.....	25
3.4 Technologieauswahl und Umsetzung der Transformation.....	26
3.5 Datensatzübergreifende Integrationsaufgaben.....	26
3.5.1 Technische Aufgaben.....	26
3.5.1.1 Angleichung der Schreibweise von Tabellen- und Feldnamen.....	26
3.5.1.2 Händische Aufbereitung der Excel-Tabellen.....	26

3.5.1.3	Automatisierung der Import- und Integrationsworkflows	27
3.5.2	Inhaltliche Aufgaben	27
3.5.2.1	Erstellung von Mapping-Tabellen auf eBIS-Zieltabellen und -felder	27
3.5.2.2	Erstellung von Lookup-Tabellen auf eBIS-Schlüsselwerte	27
4	Importprozesse für die einzelnen Datensätze	29
4.1	Allgemeiner Importprozess in FME / SQL	29
4.1.1	Grobstruktur	29
4.1.2	Importprozess für standortscharfe Bodenmesswerte	30
4.1.3	Metadatensätze für standortscharfe Bodendatenbanken	31
4.1.4	Flächenhafte Bodendaten und Toxizitätsdaten	31
4.2	Importvorgänge für die einzelnen Datenquellen	31
4.2.1	Dioxindatenbank	32
4.2.2	Bodendaten aus den bundesweiten Auen (Auen2016)	32
4.2.3	Bodenzustandserhebung Wald II	32
4.2.4	Organische und anorganische Hintergrundgehalte	33
4.2.5	Umweltprobenbank des Bundes (UPB)	33
4.2.6	Sickerwasser	33
4.2.7	Bodendaten aus der Klimafolgenforschung / Erosionsdaten	34
4.2.8	Boden – Bodenorganismen	34
4.3	Auswertungs- und Abfragefunktionen	34
4.3.1	Dublettendetektion	35
4.3.2	Geografische Sichten	36
5	Projektergebnisse	38
5.1	Weiterentwicklung vorhandener Datenstrukturen	38
5.1.1	Zeitpunkt der Probenahme	39
5.1.2	Lagegenauigkeit der Standorte	40
5.1.3	Nutzung von EPSG-Codes für Koordinatenangaben	40
5.1.4	Ausgliederung extern beschaffbarer Stammdaten	41
5.1.4.1	Geografische Stammdaten	41
5.1.4.2	Stoffdaten	41
5.1.4.3	Klimadaten	41
5.1.5	Codierung von Horizontbezeichnungen	42
5.1.6	Abbildung von Herkunftsinformationen	42

5.1.7	Anpassung der Tabellenbezeichnungen an die einschlägigen Bezeichnungen in der Bodenkundlichen Kartieranleitung	43
5.1.8	Konsolidierung der Beprobungsmethoden.....	43
5.1.9	Flexibilisierung bei der Angabe des Verdichtungsraumes	43
5.1.10	Explizite Mitführung von Bestimmungsgrenzen.....	43
5.1.11	Eliminierung der Doppelstrukturen für Einzelmesswerte und Mittelwerte.....	43
5.1.12	Datenbereinigung	43
5.2	Weitere Erkenntnisse und Empfehlungen	44
5.2.1	Aufwände bei der Datenintegration	44
5.2.2	Erkenntnisse zur Datenpublikationen und möglichen OpenData-Aktivitäten des UBA	45
6	Quellenverzeichnis	47
A	Anhang Datenbankschemata	48
A.1	Datenbanktabellen eBIS.....	49
A.2	Dioxindatenbank	53
A.3	Datenbanktabelle Bodendaten aus den bundesweiten Auen	71
A.4	Datenbanktabelle Bodenzustandserhebung Wald II	78
A.5	Datenbanktabellen Organische und Anorganische Hintergrundgehalte.....	95
A.6	Tabellenblätter Umweltprobenbank	122
A.7	Tabellenblätter Sonstige – Sickerwasser	127
A.8	Tabellenblätter Bodendaten Klimafolgenforschung / Erosionsdaten	131
A.9	Datenbanktabelle Boden – Bodenorganismen	134
B	Sonstige Anlagen	142
B.1	Checkliste der Projektfragen.....	142

Abbildungsverzeichnis

Abbildung 1: Importprozess – Grobstruktur	29
Abbildung 2: Feinstruktur Importprozess „standortscharfe Bodenmesswerte“	30
Abbildung 3: Schematische Darstellung der Zieltabellen in eBIS für die Übernahme der Bodenmesswerte	30
Abbildung 4: Beispielansichten für kartografische Darstellungen.....	36

Tabellenverzeichnis

Tabelle 1: Tabellarische Übersichten der zu integrierenden Datensätze .	15
Tabelle 2: Tabellarische Übersichten zu Datenbanktabellen.....	17
Tabelle 3: Priorisierung von Datenfeldern	17
Tabelle 4: Datentabellen der Bodenzustandserhebung Wald II	19
Tabelle 5: Datentabellen der Datenbank Organika_HW.....	20
Tabelle 6: Tabellendaten der Umweltprobenbank	21
Tabelle 7: Excel-Tabellen „Sonstige-Sickerwasser“	21
Tabelle 8: Excel-Dateien der Erosionsdaten	22
Tabelle 9: Übernahmevarianten für eBIS-fremde Kontextdaten.....	24
Tabelle 10: Verknüpfungspunkte zwischen eBIS und den nicht- standortscharfen Daten.....	31
Tabelle 11: Datenbankobjekte für Auswertungs- und Abfragefunktionen.	34
Tabelle 12: Attribute der Sicht <i>messwerte_geo_view</i>	36
Tabelle 13: Anzahl übernommener Messwerte.....	38
Tabelle 14: Übersicht m_bodenmesswerte	49
Tabelle 15: Übersicht T_ANALYSEN_ERGEBNIS	53
Tabelle 16: Übersicht T_PROBE.....	54
Tabelle 17: Übersicht T_PROBENNAHME	58
Tabelle 18: Übersicht T_STANDORT.....	65
Tabelle 19: Übersicht Datenbanktabelle Auen2016	71
Tabelle 20: Übersicht Tabellen BZE_Organika_0_5_neu, BZE_Organika_5_10_neu, BZE_Organika_Auflage_neu	78
Tabelle 21: Übersicht Tabelle PCDDF_dl_PCB_Wald_Auflage.....	89
Tabelle 22: Übersicht Tabelle Access_LW	96
Tabelle 23: Übersicht Tabellen AggregierteWerte_Boden_Analysen.....	122
Tabelle 24: Übersicht „Kopfdaten“	127
Tabelle 25: Übersicht „KGV_BO“.....	127
Tabelle 26: Übersicht „KW_BO“.....	128
Tabelle 27: Übersicht „KAK_BO“	130
Tabelle 28: Datentabelle Abtrag_Ist_Naturraum_UBA – Datenfelder	131
Tabelle 29: Übersicht Tabelle „datenbank“	134

Abkürzungsverzeichnis

BfN	Bundesamt für Naturschutz
ETL	Extract-Transform-Load
FME	Feature Manipulation Engine, eine ETL-Software der Firma Safe
GIS	Geografisches Informationssystem
PFC	Engl. perfluorinated compound (Perflorierte Verbindung)
SQL	Structured Query Language
TEQ	Toxizitätsäquivalent
UBA	Umweltbundesamt, Dessau

Zusammenfassung

Das Umweltbundesamt führt Bodeninformationen zum chemischen, physikalischen und biologischen Zustand der Böden in Deutschland. Insbesondere aus Sicht des Bodenschutzes sind diese Informationen von großem Interesse und es ist für viele Fragestellungen von großer Bedeutung, auf möglichst viele dieser Datenbestände für Auswertungszwecke zurückgreifen zu können. Die zusammengesetzte Auswertung ist wichtig, weil dadurch zeitlich, räumlich und stofflich differenzierte Aussagen getroffen werden können, die aus den einzelnen Datenbanken nicht ableitbar sind. Viele der hausintern vorhandenen Datensätze liegen jedoch in unterschiedlichen Systemen, Datenbanken und Dateien vor und sind untereinander nicht kompatibel, sodass zunächst keine effiziente quellübergreifende Datenauswertung möglich ist.

Gegenstand dieses Projektes ist daher die Zusammenführung einzelner Datenbestände, um künftig eine integrierte Auswertung zu ermöglichen. Hierfür wurden zunächst acht Datenbanken unterschiedlicher Größe, Struktur und Aktualisierungshäufigkeit hinsichtlich ihrer Zusammenführbarkeit untersucht. In einem zweiten Schritt wurden Ableitungsregeln definiert, die die Transformation der einzelnen Datenbestände in ein einheitliches Zielschema beschreiben. In einem dritten Schritt wurde ein Softwarewerkzeug erstellt, das aus diesen Regeln einen integrierten Datenbestand erzeugt. Die Anwendung der Regeln erfolgt dabei weitgehend automatisch, sodass auch neuere Versionen der einzelnen Datenbestände „auf Knopfdruck“ in einer relationalen Datenbank zusammengeführt werden können.

Um künftig auf Änderungen in den Quelldatensätzen reagieren zu können, sind die teils sehr komplizierten Ableitungsregeln in Konfigurationsdateien bzw. in SQL hinterlegt und können im Bedarfsfall angepasst oder korrigiert werden. Des Weiteren wurden Sichten und Tabellen zur Qualitätssicherung erstellt, in denen z.B. Duplikate leicht erkennbar sind oder die eine kartografische Darstellung der Daten in einem Geoinformationssystem erleichtern.

Die Ergebnisse dieses Projektes sind in erster Linie die integrierten Daten sowie das Integrationswerkzeug, mit dem der integrierte Datenbestand aus den einzelnen Datenquellen jederzeit neu aufgebaut werden kann. Darüber hinaus konnten wichtige Erkenntnisse zu Stärken und Schwächen der aktuellen Erfassung, Dokumentation und Übernahme von bodenrelevanten Daten gewonnen werden. Insbesondere flossen erhebliche Zeit- und Arbeitsaufwände in die Beschaffung und Rekonstruktion von Datensatzdokumentationen. Erhebliche manuelle Aufwände entstanden auch bei der Ableitung der Transformationsregeln und Einzelfallentscheidungen zur Qualität und Zusammenführbarkeit einzelner Datensätze. Hier zeigt sich eine große Schwäche der aktuellen Datenhaltung: Abgesehen von unvollständigen Datensätzen sind Erfassungsvorschriften und Datenstrukturen in den einzelnen Quelldatenbanken so inhomogen, dass eine automatische Zusammenführung a priori nicht möglich ist. Diese Situation ist in erster Linie den gewachsenen Strukturen in den jeweiligen Systemen geschuldet. Die daraus entstandenen Nachteile werden erst bei einem Integrationsversuch offenbar. Während die einzelnen Datenquellen sehr unterschiedliche Schwerpunktsetzungen haben, ist aus Sicht einer integrierten, quellenübergreifenden Auswertung eine möglichst umfassende und deutschlandweit standardisierte Erfassungs- und Aufnahmevorschriften wünschenswert. In diesem Zielkonflikt müssen alle beteiligten Akteure dauerhaft zusammenarbeiten, um die Nachnutzbarkeit einmal erhobener Daten dauerhaft zu erhöhen. Auch für diese Aufgabe enthalten die Projektergebnisse erste Anregungen.

Dieser Projektbericht ist in fünf Hauptkapitel gegliedert. Kapitel 1 beschreibt kurz die Ausgangssituation, nennt die untersuchten Datenbestände und beschreibt grundsätzliche Festlegungen für die Zusammenführung der Daten. In Kapitel 2 werden die einzelnen Datenbestände genauer charakterisiert. Die dort referenzierten Datenschemata und

Ableitungsregeln für die Datenintegration sind in Anhang A aufgeführt. Der detaillierte Regelsatz für die Datenableitung des integrierten Datensatzes wurde aus Gründen der Lesbarkeit und der Vermeidung von Inkonsistenzen zusammen mit dem Transformationswerkzeug gepflegt.

In Kapitel 3 werden mögliche Ansätze zur Integration der Datenbestände beschrieben und hinsichtlich Ihrer Zweckmäßigkeit bewertet. Daraus werden im Nachgang Leitplanken für die softwaretechnische Umsetzung der Datenintegration abgeleitet. Der allgemeine Aufbau des Übernahmeprozesses sowie die jeweilige Umsetzung für die einzelnen Datensätze ist in Kapitel 4 beschrieben. Dort sind auch die zusätzlichen Datenstrukturen und Sichten für die Dublettendetektion und die Erstellung von Auswertefragen dokumentiert.

Die Darstellung der Projektergebnisse erfolgt in Kapitel 5. Hier finden sich auch Erkenntnisse und Hinweise zur künftigen Weiterentwicklung des Bodeninformationssystems „eBIS“ sowie der zur allgemeinen Weiterentwicklung von Schnittstellen und Erfassungsvorschriften für den Austausch von Bodendaten. Die dort ausgesprochenen Empfehlungen und Anregungen sollten bei der Weiterentwicklung der Datenhaltung und des Datenaustauschs im Bereich „Boden“ in künftigen Projekten berücksichtigt werden.

Summary

The German Environment Agency maintains soil information on the chemical, physical and biological condition of soils in Germany. Especially from the perspective of soil protection, this information is of great interest, and it is of great importance for many questions to be able to fall back on as many of these data stocks as possible for evaluation purposes. The composite evaluation is important because it enables statements to be made that are differentiated in terms of time, space and substance, which cannot be derived from the individual databases in this way, and thus enables appropriate measures to be taken to restore or maintain soil functions, respectively. However, many of the in-house data sets are available in different information systems, databases and files and are not compatible with each other, so that initially no efficient cross-source data evaluation is possible.

The aim of this project is therefore to merge individual data sets to enable integrated evaluation in the future. For this purpose, eight databases of different size, structure and update frequency were first examined regarding their mergeability. In a second step, derivation rules were defined that describe the transformation of the individual databases into a uniform target schema. In a third step, a software tool was created that generates an integrated dataset from these rules. The application of the transformation rules is largely automatic, so that newer versions of the individual data stocks can be merged into a single relational database "at the push of a button".

To be able to react to future changes in the source data sets, the derivation rules, some of which are very complicated, are stored in configuration files or in SQL and can be adapted or corrected if necessary. Furthermore, views and tables for quality assurance were created, in which e.g., duplicates are easily recognizable, and which facilitate a cartographic representation of the data in a geoinformation system.

The foremost results of this project are the integrated data base and the software tool, that compiles this data base from the individual data sources. In addition, important insights were gained into the strengths and weaknesses of the current collection, documentation, and transfer of ground-relevant data. Considerable time and effort also went into the acquisition and reconstruction of data set documentation. Considerable manual efforts also arose in the derivation of transformation rules and individual case decisions on the quality and mergeability of data sets. This reveals a major weakness of current data management: apart from incomplete data sets, acquisition rules and data structures in the individual source databases are so inhomogeneous that automatic merging is not straight forward. This situation is primarily due to the evolved structures in the respective systems. The resulting disadvantages only become apparent when an integration attempt is made. While the individual data sources have very different focuses and objects of investigation, from the point of view of an integrated, cross-source evaluation, it is desirable to have recording and recording regulations that are as comprehensive as possible and standardized throughout Germany. In this conflict of interests, all actors involved must cooperate on a permanent basis to permanently increase the reusability of collected data.

1 Allgemeine Bestandsaufnahme

In diesem Kapitel werden, bezogen auf den Projektanlass, die untersuchten Datensätze, die Kategorisierung der Datenbanken und Festlegungen zur Schemadokumentation beschrieben.

1.1 Projektanlass

Das Umweltbundesamt führt Bodeninformationen zum chemischen, physikalischen und biologischen Zustand der Böden in Deutschland. Die Bodeninformationen machen grundsätzlich die Ableitung von z.B. Vorsorge-, Prüf- und Maßnahmenwerte für die Bundes-Bodenschutzverordnung möglich. Auch kann es besonders bei neu auftretenden Fragestellungen von großer Bedeutung sein, kurzfristig auf möglichst viele, länderübergreifende Bodeninformationen zurückgreifen zu können. Dies erfordert eine effiziente Integration der umfangreichen hausinternen Datenbestände zum Thema Boden. Dazu gehören unter anderem:

1. Bodendaten aus der Klimafolgenforschung (Erosionsdaten),
2. Dioxin-Datenbank,
3. eBIS,
4. Humus- und Bewirtschaftungsdaten aus den Bundesländern,
5. organische und anorganische Hintergrundwerte,
6. Bodendaten aus den bundesweiten Auen,
7. Umweltdatenbank STARS
8. Umweltprobenbank des Bundes

Die Datenquellen können einzeln ausgewertet werden. Die Struktur der einzelnen Datenquellen ist aber untereinander nicht kompatibel. Das erschwert die gemeinsame, quellen-übergreifende Auswertung, weil dazu mehrere Schritte zur Datenharmonisierung und Qualitätssicherung nötig sind. Eine quellübergreifende Datenauswertung ist letztendlich ein Weg, um:

- ▶ Bodendaten in die Öffentlichkeit zu transportieren,
- ▶ Probleme des Bodenschutzes und Bodenbelastungen über geeignete Darstellungen für die Öffentlichkeit greifbar zu machen
- ▶ und damit schließlich dem Boden den Stellenwert eines wichtigen Umweltmediums in der Bevölkerung zu geben.

Die zusammengesetzte Auswertung ist wichtig, weil dadurch zeitlich, räumlich und stofflich differenzierte Aussagen getroffen werden können, die aus den einzelnen Datenbanken nicht ableitbar sind.

Anhang B enthält zudem eine kurze Zusammenstellung von Antworten zu den Fragestellungen, die zu Projektbeginn vom UBA gestellt wurden.

1.2 Untersuchte Datenbanken

In Abstimmung mit dem UBA wurden folgende relevante Datensätze für dieses Projekt identifiziert (Tabelle 1). In der Übersicht ist auch festgehalten, ob die Daten statisch (fix=J) sind oder in Abständen verändert oder fortgeschrieben werden (fix=N).

Tabelle 1: Tabellarische Übersichten der zu integrierenden Datensätze

Name	Erläuterung	fix
eBIS	Die primäre Datenbank zu Bodenmesswerten des UBA. Sie wird als PostgreSQL Dump bereitgestellt und periodisch aktualisiert.	N
Bodendaten aus den bundesweiten Auen	Messungen von verschiedenen anorganischen und organischen Stoffen und Begleitparametern in Auenböden. Statistischer Datensatz aus abgeschlossenem Projekt.	J
Organische und anorganische Hintergrundgehalte	Standortbezogene Daten zu organischen und anorganischen Hintergrundgehalten. Datensatz wird in großen Zeitabständen aktualisiert und wird in diesem Projekt als endgültig angesehen.	J
Bodendaten aus der Klimafolgenforschung / Erosionsdaten	Flächenhafte Daten zum Bodenabtrag. Die Daten sind regional modelliert, d.h. die getroffenen Aussagen beziehen sich auf Polygon-Geometrien einer naturräumlichen Gliederung Deutschlands.	J
Stoffdatenbank STARS	Alte Datenbank über Stoffdaten. STARS wird benutzt, wenn im Rahmen einer Abfrage Stoffeigenschaften, Referenzwerte oder ökotoxikologische Informationen ermittelt oder eingegrenzt werden sollen. Der Datensatz liefert Hinweise für Filterkriterien. Die in ihm enthaltenen Daten haben keinen Standortbezug. Die Datenbank ist statisch und wird derzeit aktualisiert, daher hat die technische Integration der Datenbank keine Priorität. Die in der Datenbank enthaltenen Informationen sollen eher für die Anreicherung der Standortdaten oder für die Generierung von Filterkriterien genutzt werden.	N
Umweltproben-Datenbank des Bundes	Die Probenahmegebiete der Umweltprobenbank vertreten die Hauptökosystemtypen Deutschlands und berücksichtigen die unterschiedliche Intensität der anthropogenen Beeinflussung und Flächennutzung. Die Gebiete sind dadurch in ihrer Gesamtheit für die Umweltsituation in Deutschland weitgehend repräsentativ. Die Umweltprobenbank enthält nicht nur terrestrische Proben, sondern auch limnische, marine und Humanproben.	N
Dioxindatenbank	Die Dioxindatenbank wird intern im UBA gepflegt und enthält in großen Teilen vertrauliche Daten. Die Datenbank läuft auf Oracle; Auszüge werden über die zuständige Fachabteilung bezogen. Für die Auswertung wurde mit einem stark reduzierten Abzug der Datenbank gearbeitet, der nur Bodendaten enthielt.	N
Boden- und Bodenorganismen	Dieser Datensatz enthält Informationen zur Schadstofftoleranz verschiedener Bodenorganismen. Die Daten stammen aus einem abgeschlossenen Forschungsprojekt des UBA.	J
Bodenzustandserhebung Wald II	Daten zu organischen Stoffen im Boden und verschiedenen Begleitparametern aus der zweiten Bodenzustandserhebung Wald.	J

1.3 Kategorisierung der Datenbanken

Die Datenbanken aus diesem Projekt werden in drei Kategorien aufgeteilt.

- Standortscharfe Probendatenbanken
- Flächenhafte Bodendatenbanken

- Toxizitäts- und Stoffdatenbanken

In diesem Kapitel werden die drei Kategorien kurz beschrieben und die Datenbanken zugeordnet.

1.3.1 Standortscharfe Probendatenbanken

Analysedaten von Bodenproben, die zu einer bestimmten Zeit an einem bestimmten Ort gezogen wurden, werden in Probendatenbanken abgelegt. Üblicherweise werden Bodenproben aus unterschiedlichen Tiefen nach unterschiedlichen Stoffen ausgewertet. Der einzelne Messwert zur festgestellten Stoffkonzentration ist grundsätzlich über die Geokoordinate der Messung, das Datum der Probennahme und die Bodentiefe räumlich und zeitlich referenziert.

Darüber hinaus enthalten Probendatenbanken eine Vielzahl an Kontextdaten zur Probennahme, Laborauswertung und Herkunft der Daten.

Vertreter:

- ▶ Dioxindatenbank
- ▶ Bodendaten aus den bundesweiten Auen
- ▶ Bodenzustandserhebung Wald II
- ▶ Organische und anorganische Hintergrundgehalte
- ▶ Umweltprobenbank des Bundes
- ▶ Sonstige – Sickerwasser

1.3.2 Flächenhafte Bodendatenbanken

Flächenhafte Bodendaten beziehen sich nicht auf Einzelstandorte, sondern auf Gebietseinheiten. Grundsätzlich können die bewerteten oder gemessenen Bodeneigenschaften, ähnlich wie bei den standortscharfen Probendaten, auf geografisch referenzierte Bezugsflächen, einen Bezugszeitraum und eine bestimmte Bodentiefe oder Bodenschicht bezogen werden. Auf Grund der deutlich größeren Aggregationseinheiten sind die flächenhaften Bodendaten in diesem Projekt fachlich deutlich weniger detailliert und beinhalten deutlich weniger Kontextdaten als die standortscharfen Probendatenbanken.

Vertreter:

- ▶ Bodendaten aus der Klimafolgenforschung / Erosionsdaten

1.3.3 Toxizitäts- und Stoffdatenbanken

Toxizitäts- und Stoffdaten dienen der Bewertung von (Schad-)Stoffkonzentrationen hinsichtlich ihrer Auswirkungen auf Lebewesen und Bodennutzung. Während einzelne Toxizitäts- und Stoffdaten sehr spezifisch auf das Fachgebiet „Boden“ zugeschnitten sind, werden auch fachübergreifende Datenbanken gebraucht, die allgemeine Kenngrößen für die jeweiligen Substanzen oder normative Grenzwerte für erlaubte Stoffkonzentrationen bereitstellen.

Vertreter:

- ▶ Boden – Bodenorganismen

► Stoffdatenbank STARS

1.4 Festlegungen zur Schemadokumentation

In den untersuchten Datensätzen sind die Bodendaten auf eine oder mehrere Tabellen, häufig in einer Datenbankstruktur, verteilt. Zur strukturierten Auswertung und Übersichtsdarstellung wird das in Tabelle 2 dargestellte Schema verwendet. Neben den Feldnamen, Typisierung und inhaltlicher Bedeutung sind auch Vollständigkeit und Priorisierung für die geplante Auswertung erfasst. Letztere helfen zu entscheiden, wie sinnvoll und wichtig die Integration dieser Informationen in ein gemeinsames Zielschema ist (Tabelle 3).

Tabelle 2: Tabellarische Übersichten zu Datenbanktabellen

Spalte	Erläuterung
Feld	Datenbankfeld (= Tabellenspalte)
Datentyp	Datentyp des Feldes
Verweis (foreign key)	Fremdschlüsselbeziehung, wenn vorhanden
Bedeutung	Bedeutung und Inhalt des Datenfeldes
Vollständigkeit	Vollständigkeit innerhalb aller Datensätze (von 0 bis 100%)
Priorität	Priorität für die Auswertung (siehe Tabelle 3)

Tabelle 3: Priorisierung von Datenfeldern

Priorität	Erläuterung
N	Irrelevant; nicht für die Auswertung benötigt
A	Wichtig; in jeder oder fast jeder Auswertung enthalten
B	Weniger wichtig; in einigen oder wenigen Auswertung von Interesse
C	Ergänzend; nicht Bestandteil der eigentlichen Auswertung, aber hilfreich für die Prüfung einzelner (z.B. auffälliger) Werte
U	Relevanz derzeit noch unklar

2 Dokumentation der Datenbankschemata

In diesem Kapitel werden die im Projekt berücksichtigten Datenbanken beschrieben. Hierbei wird auch abgegrenzt, welche Inhalte relevant für das Projekt sind und welche nicht.

2.1 Bodendatenbank eBIS

Die Bodendatenbank eBIS definiert im Rahmen des Projekts die maßgeblichen Leitplanken für das Zieldatenschema. Die Datenbank enthält neben den relevanten fachlichen Daten auch eine Reihe zusätzlicher Relationen, die für die Nutzerverwaltung und das Zugriffsmanagement eingesetzt werden. Des Weiteren enthält die Datenbank auch eine Reihe an Informationen, die für die Bearbeitung dieses Projektes irrelevant sind.

Ziel dieser eBIS-Dokumentation ist daher die Beschreibung der vorhandenen Daten in einem Maße, das für die Erfüllung der beabsichtigten Aufgaben sinnvoll und mit Blick auf die Handhabbarkeit des Abfragewerkzeugs angemessen ist. Die Bewertung der Relevanz einzelner Datenspalten für die Erstellung fachlich korrekter Auswertungen folgt den in Tabelle 3 definierten Prioritäten. Komplementär dazu wurde das Vorhandensein der jeweiligen Information über alle enthaltenen Datensätze erfasst.

Die Haupttabellen in eBIS sind *uf_titeldaten*, *horizonte*, *m_bodenmesswerte*, *m_pflanzenmesswerte*. Die Zuordnung folgt grundsätzlich der Systematik Untersuchungsflächen – Horizonte – Messwerte, wobei in dieser Reihenfolge jeweils 1:n-Beziehungen bestehen. Weitere auswertungsrelevante Einträge befinden sich in der Tabelle *aufnahmesituation*.

Die Zeitstempel der Bodenproben sind i.d.R. tagesscharf. Da einige Datenquellen ungenauere Datumsangaben enthalten, kann die zeitliche Genauigkeit von Einzelproben deutlich schlechter sein (Monat, Jahr). In der Regel wurde bei bekanntem Monat und fehlendem Tag der Zeitstempel auf den 1. Tag des Monats gesetzt. Ist dagegen nur das Jahr der Erhebung bekannt, wurde der erste 1. Tag des Jahres als Zeitstempel gesetzt.

Koordinatenangaben sind grundsätzlich sehr genau. Da einige Datenquellen möglicherweise (z.B. aus Gründen des Datenschutzes) Positionsangaben künstlich verschlechtern, kann die Lagegenauigkeit von Einzelproben deutlich schlechter sein. Üblicherweise wurden hier Nachkommastellen abgetrennt bzw. genullt.

Mit dem vorliegenden Versionsstand > 1.011 sind in *m_bodenmesswerte* auch die Messwerte aus der bundesweiten Humusdatenbank in *m_bodenmesswerte* integriert.

Die Bodenmesswerte befinden sich in der Datenbanktabelle *m_bodenmesswerte* im Schema *public*. Eine tabellarische Übersicht zu den verwendeten Datenfeldern findet sich im Anhang A.1. Die meisten Schlüssel Tabellen werden im Schema *lists* gepflegt. Die vorhandenen Schlüssel Listen sind weitgehend statisch und werden bei Bedarf auf Veranlassung des UBA inkrementell fortgeschrieben. Die häufigsten Aktualisierungen sind in den Schlüssel Listen *messgroessen* und *einheiten* zu erwarten.

2.2 Dioxindatenbank

Die POP-Dioxin-Datenbank (nachfolgend Dioxindatenbank) des Bundes und der Länder ist ein Archivierungs- und Informationssystem für die zentrale Dokumentation und Auswertung von Ergebnissen von Untersuchungsprogrammen zur Dioxinbelastung und Belastungen der Umwelt durch andere persistente organische Schadstoffe (POPs) wie beispielsweise polychlorierte Biphenyle (PCB) oder bromierte Flammschutzmittel.

Das Datenbankschema der Dioxindatenbank ist historisch gewachsen und besteht aus einer Vielzahl von Tabellen mit diversen Fremdschlüsselbeziehungen. Die Dioxindatenbank unterteilt die in ihr gespeicherten Proben nach sogenannten Kompartimenten, von denen für die Übernahme nach eBIS nur das Kompartiment „Boden“ relevant ist. Die Dokumentation der für die Datenübernahme relevanten Teile des Datenbankschemas befindet sich im Anhang A.2.

2.3 Bodendaten aus den bundesweiten Auen

Bodendaten zu den bundesweiten Auen wurden in einer Datenbank gepflegt. Erfasst werden verschiedene Bodeneigenschaften an unterschiedlichen Probenahmepunkten (Begleitparameter) sowie Stoffdaten, die bei der Beprobung gemessen wurden.

Als Datenbankformat kommt SQLite zum Einsatz. Maßgeblich ist die Tabelle „Auen2016“, dessen Datenschema in Anhang A.3 dokumentiert ist.

2.4 Sonstige – Bodenzustandserhebung Wald II

Im Rahmen der Waldzustandserhebung wurde bislang zweimal der Zustand des Waldbodens untersucht. Die Beprobung erfolgt in drei Tiefenstufen und die Proben wurden auf unterschiedliche Substanzen untersucht. Die gewonnenen Daten sind auf unterschiedliche Datenbanktabellen verteilt (Tabelle 4).

Tabelle 4: Datentabellen der Bodenzustandserhebung Wald II

Tabelle	Erläuterung
BZE_Organika_0_5_neu	Organische Schadstoffe, Schicht 0-5 cm im Mineralboden
BZE_Organika_5_10_neu	Organische Schadstoffe, Schicht 5-10 cm im Mineralboden
BZE_Organika_Auflage_neu	Organische Schadstoffe, Auflage auf dem Mineralboden
PCDDF_dl_PCB_Wald_Auflage	Vorkommen von PCDD und PCB in der Auflage

Die in den jeweiligen Tabellen vorhandenen Datenfelder sind im Anhang A.2 aufgeführt. Die Geokoordinaten der beprobten Standorte sind als Gauß-Krüger-Koordinaten in der Datenbank enthalten (Spalten RECHTS und HOCH). Es wird einheitlich die Gauß-Krüger Zone 4 verwendet.

Lookup-Tabellen für Toxizitätsäquivalente

Toxizitätsäquivalente sollten in einer Lookup-Tabelle gesammelt werden. Die Lookup-Tabelle besteht dann aus dem Tupel <Stoff-ID, Jahreszahl (gültig ab ...), Äquivalenzwert>.

Eine entsprechende Tabelle ist verfügbar und wird im Rahmen der Integrationsarbeiten über den Schlüssel der jeweiligen Messgröße mit eBIS verknüpft.

2.5 Organische und anorganische Hintergrundgehalte

Der Datensatz zu organischen und anorganischen Hintergrundgehalten stammt aus abgeschlossenen Projekten. Die Daten sind standortbezogen und ähneln schematisch den Datensätzen aus der Bodenzustandserhebung Wald II (Tabellen 4).

Die Beprobung erfolgte in unterschiedlichen Tiefenstufen und die Proben wurden auf unterschiedliche Substanzen untersucht. Die gewonnenen Daten sind auf unterschiedliche Datenbanktabellen verteilt (Tabelle 5).

Tabelle 5: Datentabellen der Datenbank Organika_HW

Tabelle	Erläuterung
Access_LW	Organische Schadstoffe, Tiefenstufe 0-30 cm in Böden unter Acker
BZE_Organika_0_5	Organische Schadstoffe, Tiefenstufe 0-5 cm im Mineralboden (Forst) ¹
BZE_Organika_5_10	Organische Schadstoffe, Tiefenstufe 5-10 cm im Mineralboden (Forst) ²
BZE_Organika_Auflage	Organische Schadstoffe, organische Auflage auf dem Mineralboden (Forst) ³
PCDD_F_dIPCB	Dioxine/Furane und dl-PCB in der Auflage (Forst) ⁴

Die in den jeweiligen Tabellen vorhandenen Datenfelder sind im Anhang A.5 aufgeführt. Die Geokoordinaten der beprobten Standorte sind als Gauß-Krüger-Koordinaten in der Datenbank enthalten (Spalten RECHTS und HOCH). Es wird einheitlich die Zone 4 verwendet.

Für viele Parameter sind dual Zielfelder für die Messwerte angegeben, z.B. *Acenaphtylen_µg kg* und *Acenaphtylen_µg kg_BG*. Das Feld mit dem Zusatz „BG“ wird nur dann verwendet, wenn der Messwert unterhalb der Bestimmungsgrenze liegt. Ansonsten liegt der Messwert im regulären Feld (ohne Zusatz BG). Für jeden Parameter wird immer nur das eine oder das andere Feld verwendet (das ungenutzte Feld enthält den Wert NULL).

Lookup-Tabellen für Toxizitätsäquivalente

Aus Gründen der leichteren Nachvollziehbarkeit sind in der aktuellen Datenbank die jeweils gültigen Toxizitätsgrenzwerte zusätzlich hinterlegt. Diese sollten im Rahmen der Übernahme in einer separaten Lookup-Tabelle als Tupel (z.B. Stoff-ID, Jahreszahl [gültig ab ...], Wert) gesammelt werden, da sie nicht zum eigentlichen Messwert gehören, sondern global gültige Zusatzinformationen sind.

Eine entsprechende Tabelle ist verfügbar und wird im Rahmen der Integrationsarbeiten über den Schlüssel der jeweiligen Messgröße mit eBIS verknüpft (s.o.).

2.6 Umweltprobenbank des Bundes

Die Umweltprobenbank des Bundes (UPB) wird durch das UBA gepflegt und bereitgestellt. Die Probenahmegebiete vertreten die Hauptökosystemtypen Deutschlands und berücksichtigen die unterschiedliche Intensität der anthropogenen Beeinflussung und Flächennutzung. Die Gebiete sind dadurch in ihrer Gesamtheit für die Umweltsituation in Deutschland weitgehend repräsentativ.

¹ Nach Sichtprüfung bereits enthalten in BZE Wald II, daher nicht nochmals zu importieren

² Nach Sichtprüfung bereits enthalten in BZE Wald II, daher nicht nochmals zu importieren

³ Nach Sichtprüfung bereits enthalten in BZE Wald II, daher nicht nochmals zu importieren

⁴ Nach Sichtprüfung bereits enthalten in BZE Wald II, daher nicht nochmals zu importieren

Die Umweltprobenbank enthält nicht nur terrestrische Proben, sondern auch limnische, marine und Humanproben. In Tabelle 6 sind die in diesem Projekt geprüften Bodendaten aus der UPB zusammengestellt.

Tabelle 6: Tabellendaten der Umweltprobenbank

Datei	Erläuterung
AggregierteWerte_Boden_Analysen	Aggregierte Analysedaten aus der UPB
AggregierteWerte_Boden_Protokoll_Probenbeschreibungen	(nicht benötigt)
Einzelwerte_Boden_Protokoll_Probenbeschreibung_ohne_Leitung	(nicht benötigt)
Boden_PNF_Ortsbezüge	Probenahmestandorte

Eine tabellarische Übersicht zu den vorhandenen Datenfeldern findet sich im Anhang A.6. Das Datum der Probenahme ist lediglich monats-scharf bekannt und muss bei der Interpretation und Filterung der Daten entsprechend berücksichtigt werden.

2.7 Sonstige – Sickerwasser

Der Datensatz zum Thema Sickerwasser stammt aus einem F&E Vorhaben und enthält Daten zu diversen, insbesondere anorganischen, Stoffen. Relevant für die Datenintegration ist nur die Zusammenfassung in einer Excel-Tabelle (GesamtDB_Auszug_SiWaP_korr_2011.xls; in Tabelle 7 fett dargestellt).

Tabelle 7: Excel-Tabellen „Sonstige-Sickerwasser“⁵

Datei	Erläuterung
Kopfdaten	Standortspezifische Angaben
Wasser	Elementkonzentrationen und Kennwerte des Sickerwassers (BW und SW-oGW), Koordinaten der Bohrpunkte (in diesem Projekt nicht zu integrieren)
KG_V_BO	Korngrößenverteilung des Mineralbodens
KAK_BO	Kationenaustauschkapazität des Mineralbodens
KW_BO	Königswasserextrahierbare Elementgehalte im Mineralboden
LF+pHBO+CNS_BO	pH-Werte, elektrische Leitfähigkeit, C-,N-,S-Gesamtgehalte des Mineralbodens (in diesem Projekt nicht zu integrieren)
KG_V_SEDI	Korngrößenverteilung des Bodenmaterials (Sediment) aus dem SW-oGW-Bereich (in diesem Projekt nicht zu integrieren)
KAK_SEDI	Kationenaustauschkapazität des Bodenmaterials (Sediment) aus dem SW-oGW-Bereich (in diesem Projekt nicht zu integrieren)

⁵ Im Rahmen dieses Projektes sind nur die typografisch hervorgehobenen Bodendaten interessant (Endung auf Kürzel „_BO“)

Datei	Erläuterung
KW_SEDI	Königswasserextrahierbaren Elementgehalten des Bodenmaterials (Sediment) aus dem SW-oGW- Bereich (in diesem Projekt nicht zu integrieren)
LF+pHBO+CNS_SEDI	pH-Werte, elektrische Leitfähigkeit, C-,N-,S-Gesamtgehalte des Bodenmaterials (Sediment) aus dem SW-oGW- Bereich (in diesem Projekt nicht zu integrieren)
ELUAT_2zu1	wasserelulierbaren Elementkonzentrationen im Wasser- /Feststoffverhältniss 2l/kg Bodenmaterials (Sediment) aus dem SW-oGW- Bereich (in diesem Projekt nicht zu integrieren)

Eine tabellarische Übersicht zu den vorhandenen Datenfeldern findet sich im Anhang A.7.

In der Tabelle KW_BO fehlten die Angaben zu den Horizonttiefen. Sie wurden manuell anhand der Reihenfolge und der Horizontbezeichnungen ergänzt. Die Horizontangaben in den Ergebnistabellen sind jedoch nicht immer eindeutig, und nicht immer waren die Tiefenangaben zu rekonstruieren.

Außerdem fehlen in den Ergebnistabellen die Zeitpunkte der Probenahme. Hierfür existieren zwei externe Quellen: Eine Access-Datenbank mit den Rohdaten sowie eine Übersicht aus dem veröffentlichten Ergebnisbericht (Duijnsveld, et al., 2008). Die Datumsangaben in beiden Quellen unterschieden sich zum Teil erheblich. Da der Rohdatenbestand kaum dokumentiert ist, wurden die Angaben aus dem veröffentlichten Ergebnisbericht entnommen. Die Kombination aus Standortname und -nummer für den Standort Hohenzenthen ist mehrdeutig. Die Zuordnung des Probenahmedatums wurde anhand der Standortnummer getroffen.

2.8 Bodendaten aus der Klimafolgenforschung / Erosionsdaten

Dieser Datensatz wurde einmalig in einem Forschungsprojekt erhoben und liegt als Sammlung von Excel-Dateien vor. Diese sind in Tabelle 8 beschrieben.

Tabelle 8: Excel-Dateien der Erosionsdaten

Datei	Erläuterung
Abtrag_Ist_Naturraum_UBA.xls	Verschiedene Szenarien modellierter Bodenabträge mit Flächenangaben naturräumlicher Einheiten; relevant ist nur das Tabellenblatt Abtrag_Ist Dokumentation in Anhang A.8
CRKLS_Naturraum_Diff_jahr.xls	Irrelevant; Hilfstabelle für Auswertung
C_Faktor_Naturraum_Diff.xls	Irrelevant; Hilfstabelle für Auswertung
RKLS_Naturraum_Diff.xls	Irrelevant; Hilfstabelle für Auswertung
RKS_Naturraum_Diff.xls	Irrelevant; Hilfstabelle für Auswertung
R_Faktor_Naturraum_Diff.xls	Irrelevant; Hilfstabelle für Auswertung
Zsfssg_WETTREG.xls	Irrelevant; Hilfstabelle für Auswertung
Zsfssg_WETTREG_Rstat.xls	Irrelevant; Hilfstabelle für Auswertung

2.9 Stoffdatenbank STARS

STARS ist eine Stoffdatenbank für bodenschutz- und umweltrelevante Stoffe. Die Stoffauswahl und die Informationstiefe orientieren sich an den praktischen Erfordernissen aus den Arbeitsgebieten Bodenschutz, Altlasten und Gewässerschutz.

Die derzeit aktuelle Version 4.2.1 wird zukünftig generalüberholt, sodass die Erstellung einer detaillierten Dokumentation zum Zwecke der fortlaufenden Datenintegration derzeit nicht sinnvoll erscheint. Die Prüfung einer engeren Kopplung von STARS an eBIS sollte Gegenstand einer künftigen Untersuchung sein.

2.10 Boden – Bodenorganismen

Dieser Datensatz enthält Informationen zur Stofftoleranz verschiedener Bodenorganismen, d.h. wieviel potenziell schädliche Stoffe können verschiedene Bodenlebewesen verkraften, bevor sie (teilweise) sterben. Die Stoffverfügbarkeit ist u.a. abhängig von den Bodeneigenschaften, d.h. welcher Bodentyp, Bodenart, pH-Wert usw. liegt vor. Inhaltlich hat der Datensatz eine gewisse Nähe zu Stoffdatenbanken wie STARS, indem er empirische Grenzwerte für die Bewertung von Bodenproben und den darin auftretenden Schadstoffen liefert.

Eine tabellarische Übersicht zu den vorhandenen Datenfeldern findet sich im Anhang A.9.

3 Vorgehensmodell zur Datenbankintegration

In diesem Kapitel werden nachfolgend die Integrations- und Lösungsansätze für die Zusammenführung der Datenbank beschrieben.

3.1 Standortscharfe Probandatenbanken

Alle Datenschemata standortscharfer Probandatenbanken haben eine ausreichend große Schnittmenge mit dem eBIS-Datenschema, sodass eine Zusammenführung grundsätzlich erfolgen kann. Je nach Datenbank sind die Daten bisweilen unterschiedlich normalisiert oder sind implizit referenziert (z.B. Tiefenschichtung Teil des Tabellennamens statt echtem Attribut), sodass in geeigneter Weise transformiert werden muss.

Weitere technische Unterschiede, wie z.B. die Nutzung unterschiedlicher Koordinatenreferenzsysteme oder die Verwendung abweichender Einheitenbezeichnungen lassen sich ebenfalls eindeutig lösen.

Große Unterschiede bestehen in den Kontextdaten, d.h. zusätzlichen Attributen, die den Messwert in seinem Aufnahmesituation / lokalem Umfeld charakterisieren oder die Messmethodik und Messunsicherheit besonders detailliert charakterisieren. Hier sind in einigen Fällen keine entsprechenden Felder in eBIS vorhanden.

Für den Umgang mit eBIS-fremden Kontextdaten wurde eine praktikable Lösung gesucht. Diese sollte zum einen leicht durchschaubar und wartbar sein, zum anderen sollte sie den Abfrageaufwand für die Recherche eBIS-fremder und damit nicht integrierbarer Datenfelder reduzieren. Dafür wurden die in Tabelle 9 dargestellten Übernahmevarianten erarbeitet.

Tabelle 9: Übernahmevarianten für eBIS-fremde Kontextdaten

Vorteile	Nachteile
A) Verzicht auf die Integration von Kontextdaten ohne Entsprechung in eBIS; Verweis auf die Originaldaten	
<ul style="list-style-type: none"> > Komplexitätsreduktion > Reinhaltung des Datenschemas > Übersichtlichkeit 	<ul style="list-style-type: none"> > Geringerer Integrationsgrad > zusätzlicher „externer“ Rechercheaufwand bei komplexen Fragestellungen
B) Nutzung von Pseudoattributen⁶ in eBIS	
<ul style="list-style-type: none"> > Komplexitätsreduktion für einfache Fälle > Reinhaltung des Datenschemas > Übersichtlichkeit, vollständige Sichtbarkeit aller relevanten Daten > integrierbar in SQL-Abfragen 	<ul style="list-style-type: none"> > Potenzielle „Überfrachtung“ des Freitext-/JSON-Datenfeldes > Einführung einer größeren Zahl zusätzlicher Lookup-Tabellen wahrscheinlich, wenn transitive Attribute mitgeführt werden, um Schlüsselwerte aus den Quelldaten zu referenzieren (Widerspruch zum Zweck der Pseudoattribute) > SQL-Abfragen auf zusammengesetzten Feldern sind komplex und aufwändiger zu erstellen
C) Erweiterung von eBIS um zusätzliche fachspezifische Datenfelder	

⁶ *various_text* und *various_json* dienen der Aufnahme von Bedeutung und Typ her variierender Informationen zur Reinhaltung des Datenbankschemas als Text der Form '<pseudoattributname>:<wert>'

Vorteile	Nachteile
<ul style="list-style-type: none"> > vollständige Sichtbarkeit aller relevanten Daten > integrierbar in SQL-Abfragen 	<ul style="list-style-type: none"> > Verunreinigung des eBIS-Datenschemas (Widerspruch zu Design-Zielen) > signifikante Komplexitätssteigerung im Datenschema

D) Erhalt der Originaldaten in separatem Schema (realisierte Variante)

<ul style="list-style-type: none"> > Komplexitätsreduktion > Reinhaltung des Datenschemas > Übersichtlichkeit > Vollständiger Informationserhalt > Hilfsstrukturen (Views) in den jeweiligen Schemata möglich, die eine tiefergehende Auswertung/Rückverfolgbarkeit unterstützen 	<ul style="list-style-type: none"> > Geringerer Integrationsgrad > Einbindung in SQL-Abfragen etwas komplexer gegenüber (C)
--	---

Für dieses Projekt wurde die Übernahmevariante (D) umgesetzt. Durch eine vollständige Kopie der Originaldaten werden die Quelldaten strukturell erhalten und Kontextattribute können im Bedarfsfall über SQL-Abfragen direkt eingebunden werden.

Hinsichtlich der Umsetzung und Pflege einer umfassenden Integrationslösung unter Erweiterung des eBIS-Datenmodells werden in den Kapiteln 5.1 und 5.2 Empfehlungen zur Weiterentwicklung gegeben.

3.2 Flächenhafte Bodendatenbanken

In den zu integrierenden Datensätzen gibt es lediglich einen flächenhaften Datensatz zur Bodenerosion. Dieser hat kaum inhaltliche oder strukturelle Gemeinsamkeiten mit der eBIS-Datenbank, sodass eine echte Integration hinfällig wird.

Aus Auswertungssicht erscheint es am sinnvollsten, diesen Datensatz neben den eBIS-Daten zu speichern, ggf. in einem eigenen Schema, mindestens jedoch in einem eigenen Set von Tabellen. So können die Daten – je nach Aufgabenstellung – mit geeigneten SQL-Abfragen oder grafischen Abfragewerkzeugen mit geografischen Zusatzfunktionen (beispielsweise mit Talend Spatial Data Integrator, FME, QGIS) leichter verknüpft werden.

3.3 Toxizitäts- und Stoffdatenbanken

Toxizitäts- und Stoffdaten werden in eBIS in größeren Lookup-Tabellen gepflegt. Eine echte Integration externer Systeme ist bisher nicht erfolgt. Im Zuge der Kopplung externer Datenbanken, teils mit deutlich engerem Fokus auf bestimmte Stoffgruppen (z.B. PAKs), würden langfristig von der Anbindung externer Stoffdatenbanken wie STARS profitieren. Hauptquelle neuer Stoffgruppen im Zuge der Datenbankintegration ist vermutlich die DIOXIN-Datenbank.

Hierfür sollten diese Datenbanken in dem für eBIS relevanten Umfang gespiegelt und in entsprechenden Tabellenstrukturen eingebettet werden. Grundsätzlich gibt es zwei unterschiedliche Fälle, die behandelt werden müssen:

- 1) Es müssen zusätzliche Zeilen in den vorhandenen eBIS-Schlüsseltabellen ergänzt werden. (unkritisch)
- 2) Es müssen neue Tabellen/Beziehungen erstellt oder vorhandene Schlüsseltabellen um zusätzliche Spalten ergänzt werden. (problematisch, da das Datenmodell von eBIS

künftig auch strukturell erweitert oder umgebaut werden kann und somit Schemakonflikte entstehen könnten)

Fall (1) ist in diesem Projekt direkt lösbar. Für den Fall (2) muss künftig eine entsprechende Abstimmung mit allen Beteiligten erfolgen. Dieses Projekt kann das nicht leisten, entsprechende Tätigkeiten jedoch vorbereiten, indem es solche Fälle dokumentiert und ggf. Lösungsvorschläge nennt.

Die Toxizitätsgrenzwerte aus dem Datensatz Boden-Bodenorganismen (vgl. Kapitel 2.10) sind nach Möglichkeit mit den bisher verwendeten Kennungen zu harmonisieren und sollten ebenfalls als zusätzliche Tabellenstruktur an das eBIS-Schema herangeführt werden. Eine vollständige Integration in eBIS (mit der notwendigen Abstimmung aller Nutzer/Betreuer) ist derzeit nicht vorgesehen.

3.4 Technologieauswahl und Umsetzung der Transformation

Nach mehrfacher Abstimmung mit dem UBA werden folgende Vorgaben für die Lösungsarchitektur getroffen. Zieldatenbank für die Zusammenführung aller Daten ist Postgresql in der Version 12, PostGIS in der Version 3. Diese Versionsstände spiegeln den Entwicklungsstand von eBIS Mitte des Jahres 2021.

Die Datenübernahme von den Quellformaten und die Konvertierung in die Zielformate soll nach Möglichkeit mit FME (Feature Manipulation Engine, Safe Software)⁷ erfolgen. Zusätzliche Transformationsaufgaben werden in SQL programmiert und nach Möglichkeit in die FME-Workflows integriert. Die Bedienung der FME-Workflows erfolgt auf einem Desktop-Rechner, der lokal Zugriff auf die zu integrierenden Quelldaten bietet.

3.5 Datensatzübergreifende Integrationsaufgaben

In diesem Kapitel werden Integrationsaufgaben, die für jeden der Datensätze durchgeführt werden müssen, beschrieben. Hierzu gehören sowohl technische als auch inhaltliche Aufgaben.

3.5.1 Technische Aufgaben

Technische Aufgaben beziehen sich auf strukturelle oder programmatische Tätigkeiten, die für die Zusammenführung der einzelnen Datensätze durchgeführt werden müssen.

3.5.1.1 Angleichung der Schreibweise von Tabellen- und Feldnamen

Tabellen- und Feldnamen der Quelldaten enthalten oft Sonderzeichen oder sind in Groß-/ Kleinschreibung formatiert oder innerhalb der Tabelle nicht eindeutig bezeichnet. Postgresql als Integrationsziel nutzt standardmäßig Kleinbuchstaben, vermeidet Sonderzeichen und verlangt bevorzugt bestimmte strukturelle Eigenschaften der Feld- und Tabellennamen. Im Zuge des Datenimports wird darauf hingewirkt, diese Vorgaben einzuhalten und auf ähnliche, sinnerhaltende Bezeichner umzustellen, wenn die ursprünglichen Bezeichner dem Postgresql-Standardformat nicht entsprechen.

3.5.1.2 Händische Aufbereitung der Excel-Tabellen

Einige Datentabellen, insbesondere Excel-Sheets, nutzen keine einheitlichen Spaltennamen oder verwenden Spaltentitel, die über mehrere Zeilen gehen. Hier sind manuelle Vorarbeiten

⁷ Die Umsetzung mit FME ist voll-funktional, hat aber grundsätzlich exemplarischen Charakter. Die damit durchgeführten Transformationsschritte können auch mit anderen Softwarewerkzeugen umgesetzt werden. Die Wahl fiel auf FME, weil das Produkt am UBA vorhanden ist.

unvermeidbar. Soweit erkennbar, sind davon ausschließlich abgeschlossene Projektdatensätze betroffen, für die künftig keine neuen Versionsstände importiert werden müssen.

3.5.1.3 Automatisierung der Import- und Integrationsworkflows

Es ist davon auszugehen, dass künftig einige Datensätze (insbesondere eBIS) in neuen Versionen re-importiert werden müssen. Weiterhin kann nicht ausgeschlossen werden, dass sowohl Quelldaten als auch Abbildungsvorschriften Fehler enthalten, die erst nach Ende der Projektlaufzeit offenbar werden.

Daher wird für den Aufbau der Import- und Integrationsworkflows ein hoher Automatisierungsgrad angestrebt, der alte Integrationsstände sicher und transparent ersetzt.

3.5.2 Inhaltliche Aufgaben

Inhaltliche Integrationsaufgaben beziehen sich auf die inhaltliche Zuordnung und Ergänzung aller Datenbestände. Sie erfordern technische Vorarbeiten, benötigen jedoch im Wesentlichen Expertenwissen zu bodenkundlichen Untersuchungen, Auswertungsmethodiken und Konventionen zur Nutzung von eBIS bzw. der übrigen Datenbanken. Hierzu gehört insbesondere das händische Erstellen von Abbildungsvorschriften, sog. Mapping-Tabellen, die Datenfelder der Quelldaten in die entsprechenden eBIS-Zielfelder überführen.

3.5.2.1 Erstellung von Mapping-Tabellen auf eBIS-Zieltabellen und -felder

Tabellen- und Feldnamen der Quelldaten sind anders formatiert und benannt als ihre Entsprechungen in eBIS. Es sind geeignete Abbildungsvorschriften („Mappings“) auf die entsprechenden eBIS-Felder zu erstellen. Diese werden als zusätzliche Spalten in der Schemadokumentation (Anhang A) gepflegt.

Da die Quelldaten ihrerseits oft unvollständig dokumentiert sind und nach zunächst unklaren Erfassungsvorschriften bzw. Kartieranleitungen erhoben wurden, war die Erstellung dieser Abbildungsvorschriften sehr zeit- und arbeitsintensiv. Die Qualität der vorhandenen Informationen musste zunächst durch umfangreiche Auswertungen festgestellt und anschließend von den Projektbeteiligten gemeinsam bewertet werden. Die Dokumentation der Abbildungsregeln in Anhang A ist das kompakte Ergebnis dieses langwierigen und aufwändigen Prozesses.

3.5.2.2 Erstellung von Lookup-Tabellen auf eBIS-Schlüsselwerte

Schlüsselwerte für bestimmte Parameter, Begleitparameter, Einheiten, Gebietsbezeichnungen etc. müssen einzeln geprüft und in die jeweiligen Schlüssel-Codes in eBIS übersetzt werden. Dieser Vorgang konnte nur für wenige Attribute teilautomatisiert erfolgen. In der Regel wurden die Lookup-Tabellen manuell erstellt.

In einigen Fällen ist es notwendig, zusätzliche Werte in die eBIS-Schlüssellisten einzupflegen. In diesen Fällen wurden durch das UBA die zusätzlich benötigten Schlüsselwerte in den Stammdatentabellen der primären eBIS-Datenbank nachgepflegt, sodass z.B. neu angelegt Stoffe und Einheitendefinitionen künftig in einheitlicher Form zur Verfügung stehen.

Da einige Quelldatenbanken von eBIS abweichende Erfassungs- und Kartieranleitungen nutzen, konnten in vielen Fällen keine eindeutigen Entsprechungen gefunden werden. In Abwägung zwischen möglichst umfassender Datenübernahme einerseits und Konformität zum Zielschema andererseits wurde i.d.R. zugunsten der Konformität entschieden. Da der Übernahmeprozesses vollständig automatisiert umgesetzt wurde, kann diese Entscheidung später auch noch

nachjustiert werden. In diesem Fall sind lediglich die Transformationskripte anzupassen, sodass zusätzliche Datenfelder in die jeweiligen eBIS-Tabellen ausgespielt werden.

4 Importprozesse für die einzelnen Datensätze

In diesem Kapitel werden die realisierten Integrationsprozesse für die einzelnen Quelldatenbanken beschrieben.

4.1 Allgemeiner Importprozess in FME / SQL

Die programmatische Umsetzung der Importprozesse erfolgt im Wesentlichen über FME Workflows und SQL-Skripte. Deren Aufbau wird nachfolgend genauer beschrieben.

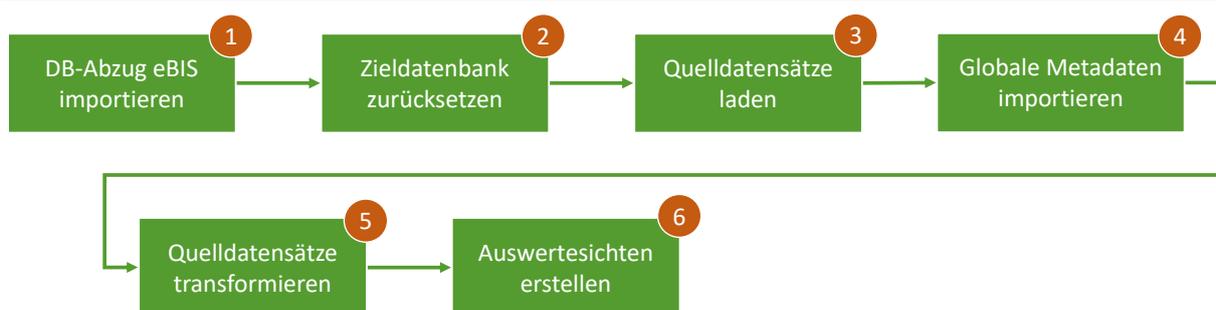
4.1.1 Grobstruktur

Der Aufbau einer integrierten Zieldatenbank über alle beschriebenen Datenquellen erfolgt nach dem in Abbildung 1 dargestellten Prozess. Die Zieldatenbank wird dabei auf einem aktuellen eBIS-Datenbankdump sowie den jeweiligen Versionsständen der übrigen Quelldaten komplett neu aufgebaut. Gegenüber einer inkrementellen Fortschreibung der Zieldatenbank ist dieser Prozess deutlich transparenter, robuster und sicher testbar. Sollten zu einem späteren Zeitpunkt Fehler in den Quelldaten oder in den Transformationsskripten festgestellt werden, kann nach diesem Prozess sehr schnell eine neue Datenbank mit entsprechenden Korrekturen erzeugt werden.

Des Weiteren bietet dieser Prozess die Möglichkeit, künftig weitere Datenquellen nach dem gleichen Vorgehen anzubinden oder aus den vorhandenen Datenquellen weitere auswertungsrelevante Daten zu extrahieren.

Konzeptionell entspricht dieser Prozess einem abgewandelten Extract-Load-Transform (ETL)-Prozess, in dem Teile der Transformationsschritte nach dem Laden erfolgen. Dadurch stehen für den Transformationsvorgang gut ausgebaute und sehr performante Datenbankfunktionen zur Verfügung. Durch das nahezu unveränderte⁸ Laden der Quelldaten stehen für spätere Nachkontrollen die Quelldaten in der gleichen Datenbank bereit.

Abbildung 1: Importprozess – Grobstruktur



Quelle: eigene Darstellung, PIKOBYTES

In Schritt **(1)** wird auf Basis eines aktuellen eBIS-Abzugs eine neue Datenbank erzeugt. Aus dieser Datenbank wird in Schritt **(2)** eine neue Zieldatenbank erzeugt, die Hilfsfunktionen und -tabellen für die nachfolgend importierten Datenquellen enthält. In Schritt **(3)** werden die übrigen Quelldatensätze unter Einsatz geeigneter FME-Transformatoren in die Zieldatenbank geladen. In Schritt **(4)** erfolgt der Import globaler (statischer) Metadaten, die zur Dokumentation der zusätzlichen Datenquellen im eBIS-Schema benötigt werden. Die Transformation der Quelldatensätze erfolgt dann in Schritt **(5)** über SQL-Skripte, die direkt auf

⁸ Es werden im Wesentlichen nur leichte Anpassungen von Feldnamen und Datentypen vorgenommen

der Zieldatenbank ausgeführt werden. Am Ende von Schritt (5) liegt eine erweiterte eBIS-Datenbank vor, die zusätzliche Datensätze aus den übrigen Datenquellen enthält. Um die Sichtung des Datenbestandes zu erleichtern und typische Abfragen zu vereinfachen, werden in Schritt (6) zusätzliche Datenbanksichten erzeugt.

Die folgenden Kapitel dokumentieren Teilaspekte der Importprozesse für unterschiedliche Arten von Quelldatensätzen.

4.1.2 Importprozess für standortscharfe Bodenmesswerte

Die Importprozesse für Datensätze zu standortscharfen Bodenmesswerten sind ähnlich aufgebaut und folgen einer einheitlichen Struktur, die in Abbildung 2 dargestellt ist.

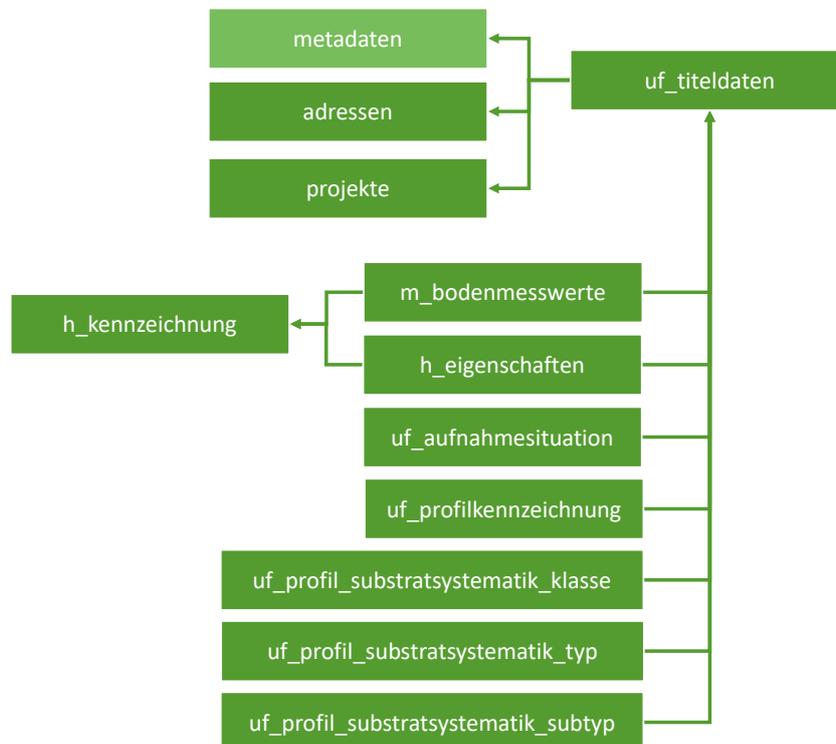
Abbildung 2: Feinstruktur Importprozess „standortscharfe Bodenmesswerte“



Quelle: eigene Darstellung, PIKobytes

Ausgehend von den im Vorfeld importierten Rohdaten werden im Schritt (2) für jeden Datensatz spezifische Schlüsseltabellen angelegt. Diese enthalten Abbildungsvorschriften von der Codierung der Datenquelle zur harmonisierten Codierung in eBIS. Mit diesen Schlüsseltabellen werden die Rohdaten in Schritt (3) transformiert und in Schritt (4) mit in die passenden eBIS-Tabellen kopiert. Die dafür relevanten eBIS-Zieltabellen sind in Abbildung 3 mit ihren referenziellen Abhängigkeiten dargestellt.

Abbildung 3: Schematische Darstellung der Zieltabellen in eBIS für die Übernahme der Bodenmesswerte



Quelle: eigene Darstellung, PIKobytes

Diese Teilschritte werden für jede Zieltabelle in ähnlicher Form durchlaufen. Die Strukturierung dient der besseren Lesbarkeit und Wartbarkeit der SQL-Skripte. Nachträgliche Änderungen (z.B. zur Einbindung neuer Zielfelder oder Korrekturen in den Schlüsseltabellen) können in den hinterlegten SQL-Skripten vorgenommen werden.

Für jeden Datensatz werden zwei Workflows bereitgestellt:

1. Ein FME-Workflow, der die Quelldatenbank in ein eigenes Schema in der eBIS-Postgresql-Datenbank „kopiert“
2. Ein SQL-Workflow, der die in (1) geladenen Quelldaten in das eBIS-Schema transformiert

4.1.3 Metadatensätze für standortscharfe Bodendatenbanken

Für jede importierte standortscharfe Probandatenbank werden die assoziierten Metadaten in einer globalen Metadatentabellen erfasst. Diese enthält den Namen der Quelle bzw. der Quelldatenbank und den zuständigen qualifizierten Ansprechpartner.

4.1.4 Flächenhafte Bodendaten und Toxizitätsdaten

Die zu übernehmenden nicht-standortscharfen Daten umfassen flächenhafte Bodendaten (Kapitel 3.2) und Toxizitätsdaten (Kapitel 3.3). Für diese Datensatzgruppen enthält das eBIS Schema nur wenige Anknüpfungspunkte. Diese sind in Tabelle 10 zusammengefasst.

Tabelle 10: Verknüpfungspunkte zwischen eBIS und den nicht-standortscharfen Daten

Datensatz	Zieltabelle(n)	Verknüpfung nach eBIS
Bodendaten aus der Klimafolgenforschung	in_erosion.abtrag_ist_src (die eigentlichen Erosionsdaten) in_erosion.nat_gld_meynen_schmithuesen_haupt (die Geodaten für die in <i>abtrag_ist_src</i> angegebenen räumlichen Bezugseinheiten)	Verknüpfung über die Geo-Koordinaten der in eBIS erfassten Proben (aus <i>uf_titeldaten</i>)
Äquivalenzfaktoren zur Toxizität (TEQs)	public.x_aequivalenzfaktoren	lists.messgroessen.id
Boden-Bodenorganismen	in_bodenorganismen.datenbank	lists.messgroessen.id

Für diese Datensätze entfällt eine stärkere Eingliederung in die eBIS-Tabellenstruktur. Sie werden über die dargestellten Verknüpfungen in Auswertungsabfragen direkt eingebunden. So lassen sich z.B. geschlüsselte Stoffmengen ad hoc in TEQs umrechnen oder mit Literaturangaben zur Toxizität für bestimmte Bodenorganismen in Beziehung setzen. Messdaten und Daten zur Bodenstruktur können anhand ihrer Standortkoordinaten mit den flächenhaft berechneten Erosionsdaten in Datenbankabfragen oder via GIS verknüpft werden.

4.2 Importvorgänge für die einzelnen Datenquellen

Anhand des allgemeinen Importprozesses für standortscharfe Bodenmesswerte (Kapitel 4.1.2) sind nachfolgend die quellenspezifischen Anpassungen dokumentiert. Die jeweiligen Abbildungsvorschriften finden sich im Anhang Datenbankschemata. Die detaillierte Umsetzung der Abbildung vom Quell- auf das Zielschema ist den FME-Workflows und den SQL-Skripten der Transformationswerkzeuge zu entnehmen.

4.2.1 Dioxindatenbank

Die Übernahme von Bodendaten aus der POP-Dioxin-Datenbank erfolgt aus mehreren, stark verzweigten Quelltabellen der Datenbank entsprechend des in Kapitel 4.1.2 beschriebenen Prozesses. Es werden folgende Datenbanktabellen in eBIS befüllt:

- ▶ adressen
- ▶ uf_titeldaten
- ▶ h_kennzeichnung
- ▶ h_eigenschaften
- ▶ m_bodenmesswerte
- ▶ uf_aufnahmesituation
- ▶ uf_profilkennzeichnung

4.2.2 Bodendaten aus den bundesweiten Auen (Auen2016)

Die Übernahme der Bodendaten erfolgt direkt aus der Quelltable Auen2016 entsprechend des in Kapitel 4.1.2 beschriebenen Prozesses. Es werden folgende Datenbanktabellen in eBIS befüllt:

- ▶ projekte
- ▶ uf_titeldaten
- ▶ h_kennzeichnung
- ▶ h_eigenschaften
- ▶ m_bodenmesswerte
- ▶ uf_aufnahmesituation
- ▶ uf_profilkennzeichnung
- ▶ uf_profil_substratsystematik_klasse
- ▶ uf_profil_substratsystematik_typ
- ▶ uf_profil_substratsystematik_subtyp

4.2.3 Bodenzustandserhebung Wald II

Die Messdaten der BZE sind auf mehrere Tabellen nach Tiefenstufen und Stoffgruppen aufgeteilt und dazugehörige Metadaten (z.B. Standorte, Klima etc.) teilweise redundant gespeichert.

Dieses Schema wird in normalisierter Form weiterverarbeitet:

- 1) Übernahme sämtlicher Metadaten aus der Tabelle *bze_organika_5_10_neu*
- 2) Zusammenführung aller Messdaten in einer neuen Tabelle *bze_messwerte_all*

Die Übernahme der zusammengeführten Messwerttabellen erfolgt entsprechend des in Kapitel 4.1.2 beschriebenen Prozesses.

Es werden folgende Datenbanktabellen in eBIS befüllt:

- ▶ uf_titeldaten

- ▶ h_kennzeichnung
- ▶ h_eigenschaften
- ▶ m_bodenmesswerte
- ▶ uf_aufnahmesituation
- ▶ uf_profilkennzeichnung

4.2.4 Organische und anorganische Hintergrundgehalte

Die Übernahme der Bodendaten erfolgt direkt aus der Quelltablelle *AccessLW* entsprechend des in Kapitel 4.1.2 beschriebenen Prozesses. Es werden folgende Datenbanktabellen in eBIS befüllt:

- ▶ projekte
- ▶ uf_titeldaten
- ▶ h_kennzeichnung
- ▶ h_eigenschaften
- ▶ m_bodenmesswerte
- ▶ uf_aufnahmesituation

4.2.5 Umweltprobenbank des Bundes (UPB)

Aus der UPB werden die aggregierten Probandaten aus der Quelltablelle *aggregierte_werte_boden_analysen* übernommen. Die Ortsbezüge werden aus der Tabelle *boden_pnf_ortsbezeuge* abgeleitet. Der Übernahmeprozess entspricht dem in Kapitel 4.1.2 beschriebenen Vorgehen. Es werden folgende Datenbanktabellen in eBIS befüllt:

- ▶ uf_titeldaten
- ▶ uf_profilkennzeichnung
- ▶ m_bodenmesswerte

4.2.6 Sickerwasser

Die Messwerttabellen aus der projektbezogenen Datenbank zum Thema Sickerwasser wurden händisch korrigiert und aufbereitet. Eine programmatische Transformation aus den Quelldaten brächte keine Vorteile, da im zugrunde liegenden Excel-Format relativ viele Eingabefehler/-Abnormitäten und doppelte Spaltennamen vorkommen, die zunächst nicht vollautomatisch behebbar sind. Des Weiteren handelt es sich um einen statischen Datensatz, der nicht fortgeschrieben wird, sodass eine einmalige händische Umwandlung akzeptabel ist.

Die Übernahme der nachbearbeiteten Daten erfolgt zunächst über einen FME-Workflow, der die Datentabellen KGV_BO, KAK_BO und KW_BO als Quelldaten in die Postgresql-Datenbank überspielt. Von dort werden die Tabellen dann nach eBIS übertragen.

Es werden folgende Datenbanktabellen in eBIS befüllt:

- ▶ uf_titeldaten
- ▶ uf_aufnahmesituation
- ▶ uf_profilkennzeichnung
- ▶ uf_profil_substratsystematik
- ▶ m_bodenmesswerte

In einigen Fällen konnten das Probenahmedatum oder die Entnahmetiefe nicht mehr zweifelsfrei rekonstruiert werden. In diesen Fällen wurden keine Daten übernommen.

4.2.7 Bodendaten aus der Klimafolgenforschung / Erosionsdaten

Die Erosionsdaten aus dem projektbezogenen Datensatz zur Klimafolgenforschung wurden über einen FME-Prozess in ein separates Schema in die eBIS-Datenbank integriert.

Die Werte in den Datensätze beziehen sich auf naturräumliche Gliederungseinheiten nach Meynen/Schmithüsen, die über einen Webdienst des BfN beziehbar sind⁹. Geschlüsselt sind die Daten über das Objektattribut „id“ bzw. „name“.

Für den Import wurden die Stammdaten zur naturräumlichen Gliederung einmalig vom Webdienst des BfN in ein Shapefile kopiert und zusammen mit den Projektdaten hinterlegt. Für Auswertungszwecke wurden die Projektdaten zusammen mit den zugehörigen Geometrieinformationen im Koordinatensystem ETRS89/UTM32 in der Datenbanktabelle *in_erosion.abtrag_ist* hinterlegt.

4.2.8 Boden – Bodenorganismen

Die Tabellen des Projektdatensatzes wurden über einen FME-Prozess in die Datenbanktabelle *in_bodenorganismen.datenbank* übernommen. Dabei werden leere Felder normalisiert, d.h. Felder, die nur aus Leerzeichen und Zeilenumbrüchen bestehen durch übliche NULL-Werte ersetzt. In einem nachgelagerten Verarbeitungsschritt innerhalb der Datenbank werden die in der Tabelle gelisteten Stoffe über SQL-Skripte auf die in eBIS definierten Schlüsselwerte für Messgrößen geschlüsselt.

4.3 Auswertungs- und Abfragefunktionen

Um künftig die Sichtung und Auswertung des Datenbestandes für typische Anwendungsfälle zu vereinfachen, wurden dem eBIS Schema weitere Funktionen, Sichten und Tabellen hinzugefügt. Ein Überblick der dafür angelegten Datenbankobjekte ist in Tabelle 11 dargestellt.

Tabelle 11: Datenbankobjekte für Auswertungs- und Abfragefunktionen

Datenbankobjekt	Erläuterung
<code>createPoint(rechtswert:double, hochwert:double, koordinatensystem_id:int) : geometry(point)</code>	<p>Die Funktion erzeugt aus den Angaben rechtswert, hochwert, koordinatensystem_id ein POSTGIS-Punktobjekt. Für die in der Tabelle <i>lists.koordinatensysteme</i> definierten Koordinatensysteme sind in der Funktion entsprechende Umrechnungsvorschriften definiert.</p> <p>Beispielaufrufe: -- automatische Umrechnung von Gauß-Krüger-Koordinaten mit führender Zonenkennung im Rechtswert <code>SELECT createPoint(3498500, 6000300);</code> -- Transformation aller Koordinaten in uf_titeldaten <code>SELECT createPoint(rechtswert, hochwert, koordinatensystem_id) FROM uf_titeldaten;</code> -- UTM-Koordinaten mit 6 und 8 Stellen im Rechtswert <code>SELECT createPoint(32498434, 5998386);</code> <code>SELECT createPoint(498434, 5998386, 2);</code></p>

⁹ https://geodienste.bfn.de/arccgis/rest/services/bfn_lnd

Datenbankobjekt	Erläuterung
horizontsymbol(vorsymbole:int[], hauptsymbol:int, nachsymbole:int[]): text	Die Funktion erzeugt aus den IDs der Vor-, Haupt- und Nachsymbole der jeweiligen Schlüssel Tabellen im Schema <i>lists</i> einen vollständigen Horizontsymbolstring in Textform.
uf_titeldaten_geo_view	Sicht auf uf_titeldaten mit zusätzlicher Geometriespalte
uf_titeldaten_geo	Abgeleitete Tabelle aus <i>uf_titeldaten_geo_view</i> . Enthält nur solche Titeldaten, die auch eine Geometrieinformation haben.
messwerte_geo_view	Erweiterte Sicht auf <i>m_bodenmesswerte</i> und ausgewählte Begleitdaten. Diese Sicht enthält die Messwerte mit ihren jeweiligen Messgrößen, Datumsangaben, Entnahmetiefen sowie spezifischen Angaben zum Boden und zum Standort der Entnahme. Geschlüsselte werden in lesbare Angaben umgewandelt (z.B. Einheiten, Horizontangaben u.a.)
m_bodenmesswerte_dupscan	Optimierte Tabelle für die Suche nach Duplikaten
cand_duplicates	Kontrolltabelle für die Dublettendetektion. Kandidaten für mögliche Dubletten werden zur Nachkontrolle in diese Tabelle geschrieben.
messwerte_begleitdaten_base_view	Basissicht für nachfolgende Auswertungen zu Messgrößen und ihren Begleitdaten.
messgroessen_bodenarten _letzter_messtermin_begleitdaten_view	Sicht auf ausgesuchte Messgrößen aus verschiedenen Jahren und Begleitdaten (z.B. Bodentyp, Nutzung) sowie die jeweils zuletzt gemessene (aktuelle) Bodenart (aufbauend auf <i>messwerte_begleitdaten_base_view</i>)
toc_vorrat_bodenarten _begleitdaten_view	Sicht auf den Kohlenstoff-Gehalt und die Messgrößen, die nötig sind, um aus diesem Gehalt den Kohlenstoff-Vorrat zu bestimmen (Trockenrohddichte, Lagerungsdichte, Skelettanteil) sowie die Probenahmetiefe. (aufbauend auf <i>messwerte_begleitdaten_base_view</i>)

Funktion und Anwendung von Datenbankobjekten zur Dublettendetektion und für einfache geografische Sichten werden nachfolgend genauer beschrieben.

4.3.1 Dublettendetektion

Für die Dublettendetektion auf der Messwerttabelle werden folgende Attribute herangezogen:

1. Probenahmedatum
2. Messgroesse
3. Messwert
4. Mittelwert
5. Geom (Punktgeometrie der Standortangabe)

Messwerte gelten als potenzielle Dublette, wenn die in den Eigenschaften 1-4 gleich sind und nicht weiter als ca. 5 Meter auseinanderliegen.

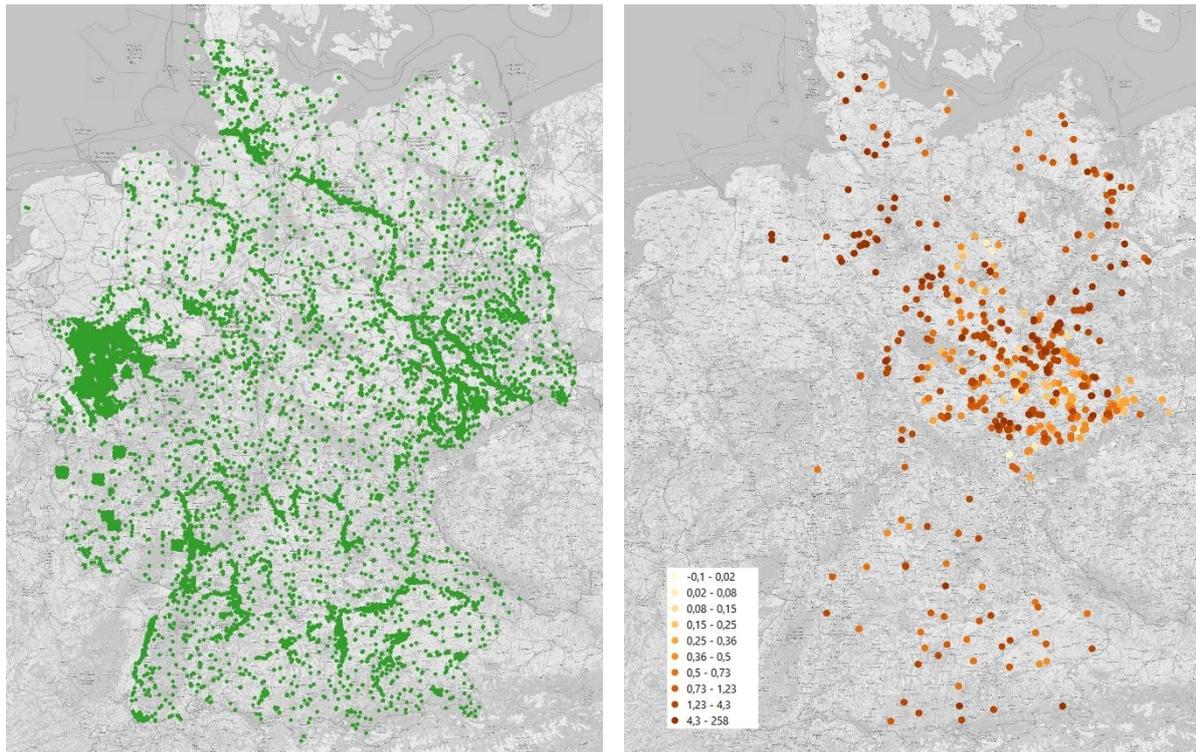
Die in der Tabelle *cand_duplicates* abgelegten Einträge dienen der manuellen Nachkontrolle potenzieller Duplikate. Die Bereinigung der Duplikate muss ggf. manuell erfolgen.

4.3.2 Geografische Sichten

Zur Darstellung und Auswertung der Messdaten in einem Geoinformationssystem sind zwei unterschiedliche Sichten realisiert worden:

- ▶ *uf_titeldaten_geo_view* enthält sämtliche Titeldaten und kann zur Visualisierung der in der Datenbank enthaltenen Standorte verwendet werden.
- ▶ *messwerte_geo_view* ist eine Auswertungssicht auf die Bodenmesswerte. Sie enthält die in Tabelle 12 aufgeführten Attribute.

Abbildung 4: Beispielansichten für kartografische Darstellungen



Quelle: eigene Darstellung, PIKobytes

Tabelle 12: Attribute der Sicht *messwerte_geo_view*

Attribut	Bedeutung
id	m_bodenmesswerte.id,
probenahmedatum	m_bodenmesswerte.probenahmedatum
entnahmetiefe_ok	m_bodenmesswerte.entnahmetiefe_ok
entnahmetiefe_uk	m_bodenmesswerte.entnahmetiefe_uk
Messgroesse	messgroessen.kurzwert
Messwert	m_bodenmesswerte.messwert oder m_bodenmesswerte.mittelwert

Messwertzusatz	messwertzusaetze.kurzwert
Einheit	einheiten.kurzwert
Extraktionsmittel	extrakte_boden.langwert
Labormethode	labormethoden.langwert
Horizont	vollst. Horizontsymbolstring gemäß Funktion <i>horizontsymbol()</i>
name_bzw_ort	uf_titeldaten_geo_view.name_bzw_ort
Geom	Standortkoordinaten der Probenahme

5 Projektergebnisse

Die Leitfrage für dieses Forschungsprojekt kann auf eine einfache Formel gebracht werden: Wie können im UBA vorhandene Bodendaten zusammen auswertbar gemacht werden?

Das UBA verfügt aus unterschiedlichen Forschungs- und Monitoringaktivitäten über heterogene Datenbestände, die Aussagen über den Zustand der Böden in Deutschland geben können. Diese Datensätze enthalten unterschiedliche Parameter und Angaben zur Belastung mit bestimmten Schadstoffen, zur Zusammensetzung des Bodens, zur Horizontierung, zur Durchwurzelung und vielen weiteren Größen. Dominant sind in jedem Fall Datensätze zu Schadstoffbelastungen und Bodenbeschreibung, die in vielfältigen systematischen und anlassbezogenen Untersuchungen gewonnen wurden.

Im Rahmen dieses Vorhabens wurden sieben verschiedene Datenquellen in einer Datenbank zusammengeführt. Dadurch konnte der bisher zur ad-hoc-Auswertung zur Verfügung stehende Datenbestand in der eBIS-Datenbank in Bezug auf die gemessenen Parameter um ca. 40% vergrößert werden (vgl. Tabelle 13). Darüber hinaus konnten für viele dieser Datensätze zusätzliche Begleitparameter zu den Bodeneigenschaften mit übernommen werden.

Ein wichtiges Zwischenergebnis sind die in manuell erstellten Dokumentationen und Ableitungsvorschriften der betrachteten Datenquellen (zu finden in Kapitel 2 und Anhang A).

Tabelle 13: Anzahl übernommener Messwerte

Quelldaten	Anzahl übernommener Messwerte (ca.)
eBIS	2.200.000
Dioxindatenbank	245.000
Auen2016	486.000
BZE Wald II	57.500
Organ. / anorgan. Hintergrund	44.400
Umweltprobenbank	7.200
Sickerwasser	10.800

5.1 Weiterentwicklung vorhandener Datenstrukturen

Im Rahmen dieses Projektes wurde damit begonnen, zusätzliche Auswertungsfunktionen und Stammdaten in die eBIS-Datenbank zu integrieren (vgl. Kapitel 4). Ausgangspunkte für diese Entwicklungen war das bestehende Datenbankschema. Dieses Schema erfuhr auch während der Projektlaufzeit immer wieder Änderungen auf Veranlassung des UBA mit dem Ziel das bestehende Schema zu verschlanken und die Ergänzung und Abfrage von Datensätzen robuster und einfacher zu gestalten.

Neben den standortscharfen Probandatenbanken wurden in die eBIS-Datenbank zusätzlich Daten zu Toxizitätsäquivalenten aufgenommen, die die Schadwirkung bestimmter Stoffe auf Biota nach verschiedenen einschlägigen Quellen beschreiben. Diese Daten haben den Charakter

von Stammdaten und sind als eigenständige Datenbanktabelle Tabelle *x_aequivalenzfaktoren* in Abfragen einbindbar¹⁰.

In gleicher Weise wurden flächenhafte Daten zur Bodenerosion in neuen Datenstrukturen eingepflegt, sodass sie gemeinsam mit den standortscharfen Probandaten über deren Geo-Koordinaten ausgewertet werden können.

Für wiederkehrende Aufgaben wurden im Rahmen des Projektes zudem neue Sichten und Funktionen generiert, die die Auswertung der Daten und die Sichtung und Kontrolle des Datenbestandes unterstützen (vgl. Kapitel 4.3). Neu sind insbesondere Funktionen zur Dublettendetektion und materialisierte Sichten zur Einbindung in GIS. Die bereitgestellten Sichten sind anpassbar (z.B. durch zusätzliche WHERE-Clauses), sodass im GIS nach bestimmten Stoffen oder Raumausschnitten gefiltert werden kann. Mit diesen Sichten ist im Vergleich zum Status quo ein effizienteres Arbeiten möglich, da lediglich Grundkenntnisse in SQL nötig sind, um die Datengrundlagen abzufragen.

Darüber hinaus hat die Integration der verschiedenen Datenbanken auch Schwächen in der eBIS-Struktur offenbart, die eine Datenintegration erschweren. Diese Erkenntnisse werden nachfolgend zusammengefasst und können Anhaltspunkte für künftige Weiterentwicklungen des Datenbankschemas sein.

5.1.1 Zeitpunkt der Probenahme

Jedem Messwert ist im eBIS-Schema ein Probenahmezeitpunkt – codiert als Datumstyp mit Jahr/Monat/Tag – zugeordnet. In der Behandlung dieses Attributes gibt es einen offensichtlichen Zielkonflikt. Einerseits sollen in eBIS nur solche Daten landen, die grundlegende Qualitätskriterien erfüllen, andererseits soll der Datenbestand möglichst umfangreich sein, um flächendeckende Auswertungen über das gesamte Bundesgebiet zu ermöglichen. Besonders zu weiter in der Vergangenheit liegenden Messwerten ist das Datum der Probenahme nicht bekannt und es kann häufig nur das Jahr festgestellt werden. In solchen Fällen wird das Probenahmedatum i.d.R. auf den 1. Januar des jeweiligen Jahres gesetzt, da eine Probenahme an diesem Tag als unwahrscheinlich gilt. Für Proben, bei denen nur Monat und Jahr der Probenahme bekannt sind, wird das Probenahmedatum auf den ersten Tag des Monats gesetzt. Beide Praktiken sind Versuche, mit der Unschärfe der vorhandenen Datumsangabe umzugehen. Für die Anbindung digitaler Abfragewerkzeuge sind solche Konventionen hinderlich, denn es ist unklar, wie die Unschärfe in adäquater Form ausgegeben werden kann.

Es wäre daher zu überdenken, ob die Erfassung des Probenahmedatums künftig auch für den Umgang mit Unsicherheiten ertüchtigt werden sollte. Eine Möglichkeit dafür ist die Angabe von Zeitintervallen, also die Ableitung eines frühestmöglichen bzw. spätestens möglichen Zeitpunktes der Beprobung. Bei jahresscharfen Angaben in der Quelle sind dann erster und letzter Tag des Jahres, bei monatscharfen Angaben erster und letzter Tag des Monats als frühester bzw. spätester Zeitpunkt anzugeben. Mit diesem Ansatz lassen sich auch beliebige andere Unsicherheiten gut abbilden, z.B. Probenahmezeitpunkte, bei denen nur die Jahreszeit bekannt ist oder die innerhalb eines mehrjährigen Projektes zu einem unbekanntem Zeitpunkt genommen wurden. Je kürzer der Unsicherheitszeitraum, desto besser ist die Qualität der Datumsangabe. Mit diesen Konventionen lassen sich Unsicherheiten auch gut nutzergerecht ausgeben oder in Form von Filterabfragen (in SQL) programmieren.

Unabhängig von den technischen Problemen bestimmt der Untersuchungsgegenstand die benötigte Genauigkeit des Probenahmedatums. Während z.B. für die Feststellung einer

¹⁰ Beispielhaft in der Sicht *messwerte_begleitdaten_base_view* realisiert

Belastung mit Schwermetallen ein jahresscharfes Probenahmedatum ausreicht, werden für Untersuchungen zum Düngemittleinsatz möglichst tagesgenaue Angaben benötigt.

5.1.2 Lagegenauigkeit der Standorte

Einige Standorte sind in ihren Geokoordinaten nicht exakt hinterlegt. Dafür gibt es zunächst zwei Kategorien von Ursachen:

- ▶ Der Standort ist bei der Probenahme nicht genau bestimmt worden oder das Wissen über den genauen Standort der Probenahme ist verlorengegangen.
- ▶ Der exakte Standort wurde bewusst verschleiert, um bestimmte Schutzanforderungen zu berücksichtigen (Datenschutz in Verbindung mit Gestattungsverträgen zur Probenahme, Schutz vor Vandalismus).

Da das eBIS-Schema ausschließlich exakte Koordinatenangaben aufnehmen kann, behilft man sich hier mit dem Abschneiden bzw. dem „Nullen“ von signifikanten Stellen. Da in der Vergangenheit allerdings auch wiederholt „Rasterbeprobungen“ (z.B. auf einem regelmäßigen Gitter mit 1..n Kilometern Rasterweite) stattgefunden haben, lassen sich die generalisierten Daten nur noch bedingt im Datenbestand identifizieren. Das Wissen zur Unterscheidung kann sich in einigen Fällen aus Hinweisen auf bestimmte Quellen oder Messkampagnen in den Metadaten rekonstruieren lassen. Zusätzlich wird allerdings häufig auch das Expertenwissen der Fachaufsicht der jeweiligen Projekte für die korrekte Interpretation benötigt. Die nachträgliche Identifikation generalisierter Positionsangaben wird zusätzlich erschwert, wenn im Prozess auf ein anderes Koordinatensystem gewechselt wurde. In diesem Fall enthalten X- und Y-Koordinaten nicht mehr die charakteristischen Nullen am Ende, sondern führen dort beliebige Ziffern.

Um unscharfe Standortangaben künftig besser abbilden zu können, bestehen mehrere Möglichkeiten:

- ▶ Die Angabe der signifikanten Stellen der X- und Y-Koordinaten.
- ▶ Die Erweiterung der bisher zulässigen Punktgeometrie auf allgemeinere Geometrieobjekte, d.h. Punkte, Linien und Flächen.

Welcher Ansatz zielführend ist, muss genauer untersucht werden. Grundsätzlich ist auch eine Kombination beider Varianten denkbar – allerdings um den Preis zusätzlicher Komplexität.

5.1.3 Nutzung von EPSG-Codes für Koordinatenangaben

Standortkoordinaten sind immer an ein bestimmtes Erdkoordinatensystem gebunden. In eBIS werden die verwendeten Koordinatensysteme unter eigenen Bezeichnern in einer eigenen Tabelle geführt. Für die Codierung von Koordinatensystemen hat sich in Verwaltung und Industrie mittlerweile die Verwendung sog. EPSG-Codes¹¹ durchgesetzt. Diese werden extern gepflegt und sind auch in Deutschland für alle Koordinatensysteme der amtlichen Vermessung verfügbar.

Es wird empfohlen, eBIS künftig auf die Nutzung von EPSG-Codes weiterzuentwickeln. Damit kann die Pflege von Koordinatensystemen innerhalb des eBIS-Schemas entfallen und es kann

¹¹ <https://epsg.io/>

mit vorhandenen Datenbankfunktionen¹² zwischen den Koordinatensystemen transformiert werden.

Insbesondere für flächenhafte Auswertungen und GIS-basierte Analysen ist die Umrechnung der vorhandenen Koordinatenangaben auf ein einheitliches Zielkoordinatensystem notwendig. Zu diesem Zweck wurde im Rahmen dieses Projektes eine entsprechende Umrechnungsfunktion *createPoint()* implementiert.

In eBIS sind bisher auch nicht-EPSSG-konforme Angaben zu den Koordinatensystemen zulässig (z.B. unterschiedliche Koordinatenbreiten für UTM oder die Angabe einer ganzen Gruppe von Koordinatensystemen – bspw. „Gauß-Krüger“). Diese Sonderfälle werden von der Umrechnungsfunktion in geeigneter Form behandelt. Im Zuge einer vollständigen Umstellung auf EPSSG-Codes lassen sich auf Basis dieser Funktion auch die neuen Koordinatensysteme für die vorhandenen Datenbankeinträge bestimmen.

5.1.4 Ausgliederung extern beschaffbarer Stammdaten

5.1.4.1 Geografische Stammdaten

Einige Metadaten sind rein geografischer Natur und können aus vorhandenen Standortkoordinaten abgeleitet werden. Zu diesen ableitbaren Angaben gehören z.B. naturräumliche Haupteinheiten, Bundesländer, Naturräume und Verwaltungseinheiten. Sofern in eBIS entsprechende Stammdaten integriert werden, können Auswertungsfunktionen auf diese Angaben direkt zugreifen. Im Zuge der Datenerfassung lassen sich damit auch inkonsistente Eingaben detektieren und als Eingabefehler abfangen.

5.1.4.2 Stoffdaten

Des Weiteren wäre es grundsätzlich sinnvoll, ein konsolidiertes Stoffverzeichnis einzubinden, welches konsequent genutzt wird. Aktuell besteht keine brauchbare Systematik, in die neu aufgenommene Stoffe eingebettet werden können. Dadurch entstehen immer wieder gelegentliche Inkonsistenzen in der Benennung und beim Abgleich der Stoffnamen, die nachträglich manuell behoben werden müssen. Geeignete Stammdatenbanken müssten dafür noch identifiziert werden.

Grundsätzlich kommen hierfür bestehende Kataloge, wie z.B. STARS oder die CAS-Registry¹³ infrage. Zunächst muss jedoch geklärt werden, ob diese Kataloge alle benötigten Stoffe und Stoffgruppen enthalten. Diesbezügliche Aktivitäten sollen im Rahmen einer einheitlichen Digitalisierungsstrategie innerhalb des UBA koordiniert erfolgen, da Stoffdaten auch von anderen Abteilungen und IT-Systemen verarbeitet werden und so ein Datenaustausch zwischen den Systemen ideal vorbereitet werden kann.

5.1.4.3 Klimadaten

Das Mitführen von Klimadaten in eBIS hat in der Praxis nur geringen Nutzen, weil die zum einen nur sehr grobgranular als Integralwert für eine bestimmte Klimaperiode vorliegen und das Stationsnetz des DWD, der Länder und Kommunen sowie privater Akteure ständig um- und ausgebaut wird. Deshalb kann erst zum Zeitpunkt der Abfrage bzw. der jeweiligen Auswertung entschieden werden, welche Klimadaten in welcher Auflösung für den Auswertezweck am besten geeignet sind. Weiterhin sind für viele Einzelproben keine Klimadaten erfasst.

¹² Die dafür notwendige Erweiterung PostGIS ist in der eBIS-Datenbank bereits vorhanden

¹³ <https://www.cas.org/cas-data/cas-registry>

Aus diesem Grund wird empfohlen, auf die Klima-Metadaten in eBIS zu verzichten und die Beschaffung adäquater Klimadaten in den Analyseprozess zu verlagern. Geeignete Stationen und Klimadatensätze können z.B. anhand der Standortkoordinaten und des Probenahmedatums identifiziert werden.

5.1.5 Codierung von Horizontbezeichnungen

Horizontbezeichnungen werden in eBIS redundant abgespeichert:

1. Zerlegt in Vor-, Haupt- und Nachsymbole
2. Unzerlegter Text (als Zeichenkette, nicht zerlegt und nicht klassifiziert)

Grundsätzlich verlangt eBIS eine Mindestqualität der Daten, sodass unklassifizierte Horizontbezeichnungen nicht vorkommen dürften. Es besteht ein gewisser Zielkonflikt zwischen den Anforderungen an die Datenqualität und dem Anspruch auf eine möglichst hohe Abdeckung des Bundesgebietes.

Fehlerhafte oder nicht-konforme Horizontbezeichnungen in Bezug auf die Bodenkundliche Kartieranleitung (2005, KA 5) waren in den zu integrierenden Datensätzen überaus häufig. In vielen Fällen musste die ursprüngliche Horizontbezeichnung deshalb verworfen werden. Es ist daher zu prüfen, ob die redundante Speicherung fortgeführt werden sollte oder eine KA 5-konforme Erfassung vorausgesetzt werden kann.

5.1.6 Abbildung von Herkunftsinformationen

Da die Einspielwege in die eBIS-Datenbank von der Erhebung bis zur Integration höchst unterschiedlich sein können, besteht auch eine Diversität in den Herkunftsinformationen, die sich derzeit nicht adäquat abbilden lassen. Probleme bei der Angabe der Herkunft bzw. Datenquelle treten insbesondere bei Sekundärquellen zutage: Zum einen gelten diese selbst als Datenquelle und müssen entsprechend referenziert werden. Zum anderen enthalten die dort enthaltenen Datensätze wiederum auch wieder Herkunftsinformationen, die zumindest näher an der Primärquelle liegen. Beide Angaben sind grundsätzlich relevant, denn:

- ▶ Es könnten Übernahmefehler von der Übertragung von der Primär- in die Sekundärquelle aufgetreten sein.
- ▶ In der Sekundärquelle könnten Datentransformationen oder sogar Korrekturen auf den Primärdaten erfolgt sein.
- ▶ Die Systematik zur Bezeichnung der Primärquelle in der Sekundärquelle und in eBIS können sich stark unterscheiden und sind durch das UBA bzw. eBIS nicht beeinflussbar.

Eine Überarbeitung der Strukturelemente zur Herkunftsdokumentation sollte anhand einer größeren Stichprobe erfolgen. Aus Sicht dieses Projektes sind mindestens zwei Angaben erforderlich:

- ▶ Die Herkunftsangabe zur Sekundärquelle.
- ▶ Die für den jeweiligen Datensatz in der Sekundärquelle hinterlegte Herkunftsangabe. (Diese ist entweder die Primärquelle oder eine weitere „zwischengeschaltete“ Sekundärquelle.)

5.1.7 Anpassung der Tabellenbezeichnungen an die einschlägigen Bezeichnungen in der Bodenkundlichen Kartieranleitung

In der KA5 werden für bodenkundliche Untersuchungen einschlägige Begrifflichkeiten verwendet. Diese sollten auch in eBIS konsequent genutzt werden, da sie die Lesbarkeit des Datenschemas erhöhen. Ein Beispiel ist die Tabelle *lists.humusanteile*, die besser *lists.humusklassen* heißen sollte.

5.1.8 Konsolidierung der Beprobungsmethoden

Daten zur Beprobungsmethode werden in mehreren Tabellen geführt: *lists.probengewinnungen*, *lists.probenahmegeräte* und *lists.beprobungsmethoden_mixe*. Die Aufspaltung der Beprobungsmethoden auf mehrere Tabellen in unterschiedlicher Granularität erschwert Datenpflege und Auswertung. Die Konsolidierung in einer Tabelle sollte erwogen werden.

5.1.9 Flexibilisierung bei der Angabe des Verdichtungsraumes

Viele Datenquellen enthalten Angaben zum Verdichtungsraum. Die Angabe reicht von einfachen binären Angaben (Verdichtungsraum ja/nein) über grobgranulare Typisierungen bis hin zu einer kleinräumigen Taxonomie der Verdichtungsräume. In eBIS wird derzeit nur die Angabe kleinräumiger Verdichtungsräume unterstützt. Auswertungsrelevant ist häufig aber nur die binäre Angabe oder die großräumige Einordnung. Es wird daher empfohlen, die Angabe des Verdichtungsraums zu flexibilisieren.

5.1.10 Explizite Mitführung von Bestimmungsgrenzen

Bestimmungsgrenzen werden in eBIS über einen Messwertzusatz („<=“) gekennzeichnet. Kennzeichnungen dieser Art sind gängige Praxis in Laborprotokollen. Für Einzelproben ist diese Konvention eindeutig. Bei Mischproben, in die auch Einzelproben unterhalb der Bestimmungsgrenze eingehen können, sind eventuell einfließende Bestimmungsgrenzen (insbesondere die größte Bestimmungsgrenze) nicht mehr erkennbar. Auch bei Einzelwertproben kann es u.U. interessant sein, Substanzkonzentrationen nahe der Bestimmungsgrenze zu erkennen. Aus diesen Gründen wird empfohlen, die explizite Mitführung der Bestimmungsgrenze für jeden Mess- bzw. Mittelwert zu erwägen.

5.1.11 Eliminierung der Doppelstrukturen für Einzelmesswerte und Mittelwerte

Einzelwerte und Mittelwerte werden derzeit in getrennten Feldern geführt. Für gängige Auswertungen werden beide i.d.R. wieder in einem Ausgabefeld zusammengeführt. Ein alternatives Modell für die Datenspeicherung ist die Ablage in einem einzigen Messwertfeld und eine zusätzliche Kennzeichnung zur Art des Messwertes (Einzelwert, Mittelwert) in einem weiteren Feld, über das die Unterscheidung nachträglich getroffen werden kann (De-Pivotisierung).

5.1.12 Datenbereinigung

Im Zuge der Datenintegration und der Nachführung von Metadaten- und Schlüsselwerttabellen sind in den vorhandenen eBIS-Tabellen einige Inkonsistenzen aufgefallen, die ggf. Gegenstand eines allgemeinen Data Clearings sein könnten:

- ▶ Die neu eingeführte Funktion zur Dublettendetektion findet sehr viele doppelt vorhandene Messwerte. Die Duplikate befinden sich sowohl im originären eBIS-Datenbestand als auch in den neu integrierten Daten. Hier sind entsprechende Algorithmen zur automatisierten

Fehlerbereinigung zu definieren, z.B. indem bestimmte Quellen priorisiert werden. Da die Dublettendetektion nur Messwerte, Zeitpunkte und Standorte berücksichtigt, müssten hier künftig auch noch die Begleitdaten verglichen und ggf. zusammengeführt werden.

- ▶ In der Tabelle *uf_bodenmesswerte* werden insbesondere die Felder *name; beschreibung; kennung; name_bzw_ort* mit sehr unterschiedlichen und schwer vergleichbaren Angaben befüllt
- ▶ Die Feldnamen in den Datentabellen passen manchmal nicht zu Schlüssellisten im Schema *lists*, z.B. in *uf_bodemesswerte* das Feld *zusatz_id*. Der Name der korrespondierenden Stammdatentabelle heißt *messwertzusaeetze*. Um die Navigation zu erleichtern, sollte das Datenfeld besser *messwertzusatz_id* heißen, sodass der Name der Stammdatentabelle mit dem Feldnamen weitestgehend übereinstimmt. Ein ähnlicher Fall besteht in der Tabelle *h_eigenschaften* in den Feldnamen *hauptgruppen_id, gruppe_id, untergruppe_id*. Diese verweisen auf die Schlüsseltabellen zur Substratgenese mit anders lautenden Namen. Diese steht zwar in der Kommentierung der Felder, sollte jedoch idealerweise aus dem Feldnamen selbst hervorgehen.
- ▶ Angaben zu Raumordnungsstrukturen bzw. Verdichtungsräumen müssten klarer definiert und ansprechbar sein.
- ▶ Die Stammdatentabellen zur naturräumlichen Gliederung (*naturraeume, naturraeumliche_gliederungen_grob, naturraeumliche_gliederungen_gruppen, naturraeumliche_grossregionen*) sollten navigierbar sein. Das bedeutet, dass beispielsweise zu einem gegebenen Naturraum der korrespondierende Naturraum in der Grobgliederung direkt bestimmt werden kann. (ggf. im Wechselspiel mit 5.1.4.1 neu zu konzipieren).
- ▶ In der Stammdatentabelle *uf_aufnahmesituation* fehlt eine Möglichkeit zur Angabe der dominierenden Baumart.

5.2 Weitere Erkenntnisse und Empfehlungen

In diesem Kapitel werden weitere Erkenntnisse und Empfehlungen aus dem Projekt vorgestellt, z.B. zu den Aufwänden der Integration und die Möglichkeit Daten zu verwerten bzw. zu veröffentlichen. Schlussendlich müssten dafür auch die bestehenden Verwaltungsvereinbarungen¹⁴ entsprechend neu ausgehandelt und formuliert werden.

5.2.1 Aufwände bei der Datenintegration

Die Integration externer Datenbestände in eBIS ist ein arbeitsintensiver Prozess mit einem hohen Grad an manueller, fallweiser Recherche-, Dokumentations- und Programmierarbeit. Bei der Zusammenführung unabhängiger Datenquellen ist dieses Problem zunächst unvermeidlich.

Dennoch besteht die Möglichkeit Vorgaben und Freiheitsgrade für regelmäßige Datenlieferanten sowie für Projekte und Forschungsprojekte mit Beteiligung des UBA hinsichtlich der Erfassung von Bodendaten stärker zu vereinheitlichen. Dadurch können Aufwände für die

¹⁴ Anhang II. 4 „Austausch von bodenschutzrelevanten Daten“ sowie Annex 3: „Daten von Bodendauerbeobachtungsflächen“ zur Verwaltungsvereinbarung zwischen Bund und Ländern über den Datenaustausch im Umweltbereich

Datenübernahme erheblich reduziert sowie Fehler und Informationsverluste bei der Datenübernahme minimiert werden.

Großen Aufwand bei der Datenintegration hat die Wiederbeschaffung oder gar Rekonstruktion von Dokumentationsunterlagen für projektbezogene Datensätze verursacht. Einige Angaben aus den Quelldatensätzen mussten verworfen werden, da die Bedeutung nicht zweifelsfrei geklärt werden konnte und sie damit später nicht sinnvoll auswertbar geworden wären. Insbesondere die Rekonstruktion von Datensatzdokumentationen und die mehrfache Klärung und Abwägung unscharf definierter Datenfelder haben in der Projektdurchführung sehr viel Zeit gekostet.

Im Rahmen des Möglichen wird daher empfohlen, für künftige Datenerhebungen und -übernahmen an einer einheitlichen Erfassungsleitlinie für Bodendaten zu arbeiten, die zunächst die wichtigsten Angaben zu Messwerten, Horizontabfolge und Bodengefüge klar und in der benötigten Qualität definiert.

Im Zuge der strategischen Weiterentwicklung von eBIS sollte auch der wiederholt offenbar gewordene Zielkonflikt zwischen hoher Datenqualität und der möglichst vollständigen Aufnahme vorhandener Bodendaten nochmals diskutiert werden. Die Arbeit an einer einheitlichen Erfassungsleitlinie ist damit untrennbar verbunden.

5.2.2 Erkenntnisse zur Datenpublikationen und möglichen OpenData-Aktivitäten des UBA

Die Ressource Boden ist kostbar und hat in den vergangenen Jahren einen deutlichen Zuwachs an Aufmerksamkeit erhalten. Dies zeigt sich insbesondere auch an aktuellen und geplanten Maßnahmenbündeln der Europäischen Kommission, die im Rahmen mehrerer Initiativen den Schutz unserer Böden verbessern sollen.¹⁵ Parallel dazu wird auch an einer besseren Datenbasis gearbeitet, die der Planung und Überwachung dieser Maßnahmen dienen.¹⁶ Insgesamt steigt damit die Nachfrage nach bodenbezogenen Daten, die sowohl den aktuellen Zustand als auch die Entwicklungsdynamik unserer Böden aufzeigen.

Auf Bundesebene werden u.a. die Daten der Boden-Dauerbeobachtung zentral beim UBA in der Bodendatenbank eBIS gesammelt. Aufgrund unklarer Datenschutz- und Nutzungsbestimmungen kann dieser flächendeckende Datensatz allerdings vorerst nur intern genutzt werden. Das UBA erfasst viele der Daten nicht selbst, sondern bezieht sie aus Forschungsprojekten und von den Bundesländern. Dort liegt dann auch das Wissen, welche Datensätze ggf. Schutzrechte von Personen und Betrieben berühren oder welche Attribute eines Datensatzes besonders schutzwürdig sind, um z.B. Vandalismus zu verhindern und v.a. den Datenschutz, der sich aus den Gestattungsverträgen zur Probenahme ergibt, zu wahren.

Diese Schutzrechte werden im Status quo auf Ebene der Datensätze nicht dokumentiert und geprüft. Folglich sind die Gestattungsverträge für die Übernahme der Daten besonders defensiv nach dem Infektionsprinzip ausgelegt: Besteht Grund zur Annahme, dass auch nur ein sehr kleiner Teil einer Datenlieferung oder ein einzelnes Attribut schutzwürdige Daten enthält, wird die gesamte Datenlieferung als „vertraulich“ eingestuft und kann nicht öffentlich weitergegeben werden. Unter Umständen werden damit die Ausgabe und Weitergabe von bodenbezogenen Daten stark eingeschränkt.

¹⁵ Montanarella, L., & Panagos, P. (2021). The relevance of sustainable soil management within the European Green Deal. *Land Use Policy*, 100, 104950. <https://doi.org/10.1016/j.landusepol.2020.104950>.

¹⁶ <https://ec.europa.eu/jrc/en/eu-soil-observatory>

Im Zuge der Erarbeitung einheitlicher Erfassungsleitlinien können datenschutzrechtlich relevante Datenfelder und Aufnahmesituationen bereits sehr früh mitbedacht und eine spätere Datenveröffentlichung inhaltlich und organisatorisch vorbereitet werden.

6 Quellenverzeichnis

Wesentliche Quellen für dieses Forschungsprojekt sind die jeweiligen projektspezifischen Datensätze und deren Dokumentationen. Weitere Literaturquellen waren für die Integration, Transformation und Auswertung des Datenbestandes nicht nötig. Der Vollständigkeit halber sind die wenigen externen Quellangaben nachfolgend zusammengefasst.

Ad-Hoc-AG Boden (2005): Bodenkundliche Kartieranleitung. (KA 5). E. Schweizerbart'sche Verlagsbuchhandlung, Hannover.

Anhang II. 4 „Austausch von bodenschutzrelevanten Daten“ zur Verwaltungsvereinbarung zwischen Bund und Ländern über den Datenaustausch im Umweltbereich.

Annex 3: Daten von Bodendauerbeobachtungsflächen zu Anhang II.4 „Austausch von bodenschutzrelevanten Daten“ zur Verwaltungsvereinbarung zwischen Bund und Ländern über den Datenaustausch im Umweltbereich.

Duijnsveld, et al. (2008): Ermittlung flächenrepräsentativer Hintergrundkonzentrationen prioritärer Schadstoffe im Bodensickerwasser, UBA, Dessau.

Montanarella, L., & Panagos, P. (2021). The relevance of sustainable soil management within the European Green Deal. *Land Use Policy*, 100, 104950. <https://doi.org/10.1016/j.landusepol.2020.104950>.

The European Soil Observatory <https://ec.europa.eu/jrc/en/eu-soil-observatory> (26.08.2021)

A Anhang Datenbankschemata

A.1 Datenbanktabellen eBIS

Tabelle 14: Übersicht m_bodenmesswerte

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
id	int		Unique ID	100	A
ido	int		ID des realen Objekts über Zeiträume und Revisionen hinweg gleichbleibend	100	A
state	int		Zustand (0 - aktuell, 1 - vorheriger ...) (Wert ist immer „0“)	100	U
owner_id	int	users.id	Eigentümer des Objekts	100	U
creator_id	int	users.id	Ersteller des Objekts	100	U
operation_id	int	operations.id	Operation, die diese Zeile erzeugte (Nutzung für Rechtemgmt. Beim Datenzugriff)	100	U
from_date	varchar		Zeitpunkt des Gültigkeitsbeginns, Format "YYYY-mm-dd HH:MM:SS"	0	U
to_date	varchar		Zeitpunkt des Gültigkeitsendes, Format "YYYY-mm-dd HH:MM:SS"	0	U
process	int		Interner Rollback-Identifizier	76	U
create_date	varchar		Zeitpunkt der Erst-Erstellung, konstant über Zeitraum und Zustand, Format "YYYY-mm-dd HH:MM:SS"	100	U
change_date	varchar		Zeitpunkt der Erstellung dieser Zeile = Ztp. der letzten Änderung, Format "YYYY-mm-dd HH:MM:SS"	100	U
metadaten_ido	int	metadaten.ido		100	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
uf_titeldaten_ido	int	uf_titeldaten.ido		99	A
t_tiefenstufen_ido	int			1	A
h_kennzeichnung_ido	int	h_kennzeichnung.ido		54	B
horizont_id	varchar			54	B
probenummer	varchar		Nummer der Probe	7	C
labornummer	varchar		Labornummer	3	C
probenehmer	varchar		Name des/der Probenehmers/-In	7	C
probenahmedatum	varchar		Datum der Probenahme	100	A
analysedatum	varchar		Datum der Analyse	64	C
probengewinnung_ido	int	probengewinnungen.ido	Art der Gewinnung der Probe (Punkt vs. Fläche)	49	C
teilflächen	varchar		Angabe der Teilfläche(n), die für die Ermittlung dieses Wertes herangezogen wurden.	38	B
sondierungen	varchar		Angabe der Sondierung(en), die für die Ermittlung dieses Wertes herangezogen wurden.	20	B
entnahmetiefe_ok	int		Entnahmetiefe, Oberkante [cm]	97	A
entnahmetiefe_uk	int		Entnahmetiefe, Unterkante [cm]	97	A
beprobungsmethode_ido	int	beprobungsmethoden_mixe.ido	Angabe der Beprobungsmethode	19	C
probenahmegefaess_ido	int	probengefaessmaterialien.ido	Kurzzeichen zur Kennzeichnung des Materials des Probenahmegefäßes	5	C
probenahmegeeraet_ido	int	probennahmegeeraete.ido	Kurzzeichen zur Kennzeichnung des Probenahmegerätes	4	C
probenvorbehandlung_ido	int	probenvorbehandlungen.ido	Kurzzeichen zur Kennzeichnung der Probenvorbehandlung	1	B

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
probenlagerung_ido	int	probenlagerungen.ido	Kurzzeichen zur Kennzeichnung der Probenlagerung	5	C
labor	varchar		Angabe des Labors und/oder des/der jeweiligen Bearbeiter/-In	0	C
messgroessengruppe_ido	int	messgroessengruppen.ido	Messgroessengruppe, denen diese Messgroesse angehört	65	C
messgroesse_ido	int	messgroessen.ido	Referenz auf die Messgröße (ehem. Parameter)	99	A
labormethode_ido	int	labormethoden.ido	Angabe der Labor- bzw. Analyseverfahren für den jeweiligen Parameter	68	A
messverfahren_ido	int	mesverfahren.ido	Angabe des jeweiligen Mess- bzw. Elementbestimmungsverfahrens der entsprechenden Messapparatur	22	A
zusatz_ido	int	messwertzusaeetze.ido	Referenz auf dne Messwertzusatz (wenn, dann meist "<" i.S.v. unterhalb Bestimmungsgrenze; Oder NN für Werte < Nachweisgrenze)	49	A
anzahl_einzelwerte	int		Wenn Mittelwert, dann hier bitte Anzahl der zu Grunde liegenden Einzelwerte	20	C
mittelwert	double		Mittelwert der Verteilung	20	A
messwert_kennung	varchar		Kennung des Messwerts	0	U
messunsicherheit	int		Messunsicherheit des Messwertes	0	C
median	double		Median der Verteilung	20	C
min	double		Minimalwert der Verteilung	20	C
max	double		Maximalwert der Verteilung	20	C
messwert	double		Messwert (Hoffentlich ist das dann auch ein Einzelwert gewesen. Sonst bitte Mittelwert nutzen.)	79	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
einheit_ido	int	einheiten.ido	Einheit des Parameters	99	A
code_sequenz_a	varchar			0	U
code_sequenz_b	varchar			0	U
code_sequenz_c	varchar			0	U
code_sequenz_d	varchar			0	U
code_medium	varchar			0	U
code_parameter	varchar		z.Zt. nicht genutzt. (Mit dem Methodencode soll eine Probe künftig einmal eindeutig beschrieben werden.)	0	U
bemerkung	varchar		Bemerkung Analyse	5	C
konfidenz_algorithmus_ido	int	konfidenzinformationen.ido		0	C
konfidenz_experte_ido	int	konfidenzinformationen.ido		0	C
konfidenz_gesamt_ido	int	konfidenzinformationen.ido		0	C
berechnet	bool		(niemals „false“)	0	C
external_key1	varchar		1. Externer Schlüssel (Hauptschlüssel) zur Identifikation des Wertesprungs aus einer Fremdquelle. Bitte nicht "BY" oder so etwas nutzen. Das haben wir schon durch die metadaten_ido und ihre quellen_ido	77	A
external_key2	varchar		2. Externer Schlüssel (Unterschlüssel) zur Identifikation des Wertesprungs aus einer <Fremdquelle	96	A
external_key3	varchar		3. Externer Schlüssel (Unter-Unterschlüssel) zur Identifikation des Wertesprungs aus einer Fremdquelle Dieser Schlüssel ist relevant für Boden-Pflanze-Paare	75	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
external_key4	varchar		4. Externer Schlüssel (Unter-Unter-Unterschlüssel) zur Identifikation des Wertesprungs aus einer Fremdquelle	28	C
extrakt_ido	int	extrakte_boden.ido		35	A
various_text	varchar		Dient der Aufnahme von Bedeutung und Typ her variierender Informationen zur Reinhaltung des Datenbankschemas als Text der Form <pseudoattributname>:<wert>'	75	U
various_json	jsonb		Dient der Aufnahme von Bedeutung und Typ her variierender Informationen zur Reinhaltung des Datenbankschemas als JSONB. Bitte various_text und various_json immer synchron füllen!	75	U
vollstaendigkeit_ido	int	vollstaendigkeiten.ido	Referenz auf die Liste der Messungsvollständigkeits (hier leider am Wert, da Messungen auf Weisung durch UBA-MA entfernt wurden). AHU-Attribut Qualität	28	U
ableitung_wv_pumw	bool		Verwendung dieses Werts für Wertevorschläge zur Ableitung von Prüf- und Maßnahmenwerten	47	U

A.2 Dioxindatenbank

Tabelle 15: Übersicht T_ANALYSEN_ERGEBNIS

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
ID	NUMBER	Primärschlüssel	N	
PROBE_ID	NUMBER	ID der Probe (Schlüssel)	A	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
STOFF_SPEKTRUM_ID	NUMBER	ID des Stoffes (Schlüssel)	A	m_bodenmesswerte:messgroesse_ido → lists_messgroessen
M01	NUMBER	Messwert	A	m_bodenmesswerte:messwert
B01	NUMBER	Bestimmungsgrenze (oder ggf. Nachweisgrenze) ¹⁷	A	m_bodenmesswerte:messwert, falls Kennzeichen „n.b.“ m_bodenmesswerte:zusatz_ido → lists.messwertzusaetze
ANM01	VARCHAR2		N	
VOLLSTAENDIGKEIT	NUMBER		N	
KENNZEICHEN	VARCHAR2	n.b. – unterhalb der BG n.n. – unterhalb der NG n.a. – nicht analysiert (ohne Kennzeichnung)	A	Es werden nur Messwerte ohne Kennzeichnung (= regulärer Messwert) oder mit der Kennzeichnung n.b. (= nicht bestimmbar unter entsprechender Kennzeichnung als Bestimmungsgrenze) übernommen
QS	NUMBER		N	
M02	NUMBER		N	
B02	NUMBER		N	
KENNZEICHNUNG_MESSWERT	NUMBER	Schlüssel für t_kennzeichnung_messwerte	N	

Tabelle 16: Übersicht T_PROBE

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
ID	NUMBER		N	

¹⁷ Da in der Datenbank und den verwendeten Erfassungsformularen kein Datenfeld für die Nachweisgrenze existiert, kann es sein, dass in einzelnen Fällen das Feld zur Angabe der Bestimmungsgrenze dafür genutzt wurde. Hier kann nur eine Einzelwertprüfung abschließende Klarheit bringen.

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
PROBENAHRME_ID	NUMBER	Schlüssel für t_probenahme	N	
PROBE_AUFB_HUM_M1_ID	NUMBER		N	
PROBE_AUFB_HUM_M2_ID	NUMBER		N	
PROBE_NR	VARCHAR2		A	m_bodenmesswerte:probenummer
BODENHORIZONT_ID	VARCHAR2	Schlüssel für t_bodenhorizont	A	h_kennzeichnung lists.horizonthauptsymbole, lists.horizont_nachsymbole, lists.horizont_vorsymbole
BODENART_ID	VARCHAR2	Schlüssel für t_bodenart	A	h_eigenschaften:bodenart_ka_5_ka_4_id,lists_bodenarten_bdf2
FARBE_ID	VARCHAR2		N	
DURCHWURZELUNG_ID	VARCHAR2	Schlüssel für t_durchwurzelung	A	h_eigenschaften:intensitaet_feinwurzeln_id, lists_feinwurzeln
DURCHFUECHTUNG_ID	VARCHAR2	Schlüssel für t_durchfeuchtung	N	n/a, da in DioxinDB keine Einteilung nach KA5
BIOTA_ID	NUMBER		N	
FUTTERMITTEL_ART_ID	VARCHAR2		N	
BIOTATEIL_ID	NUMBER		N	
SUBSTANZTYP_ID	NUMBER	Schlüssel für t_substanztypen (Es treten nur 90-Trockensubstanz und 99-unbekannt auf)	N	Auch aus Einheit der Messwerte ersichtlich
HUMAN_MATRIX_ID	VARCHAR2		N	
HUMAN_PROBEN_GEWICHT	NUMBER		N	
FETTGEHALT	NUMBER		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
GATTUNG	NUMBER		N	
GESCHLECHT	NUMBER		N	
ALTERSANGABE	NUMBER		N	
JMT	NUMBER		N	
TIEFE_BIS	NUMBER		A	m_bodenmesswerte:entnahmetiefe_uk
TIEFE_VON	NUMBER		A	m_bodenmesswerte:entnahmetiefe_ok
HUMUS	VARCHAR2		N	
BODENGEFUEGE	VARCHAR2		A	h_eigenschaften:gefuegeformen_ids; lists_gefuegeformen
TROCKENSUBSTANZ	NUMBER		N	
FRISCH_TROCKEN	NUMBER		N	
TROCKENMASSE	NUMBER		N	
HUMUSGEHALT	NUMBER		N	
P_H_WERT	NUMBER		A	m_bodenmesswerte:messwert messgroesse_id=227
KAKEFF	NUMBER		N	
WIEDERFINDUNGSRATE	NUMBER		N	
RUECKSICHT_WA_RATE	NUMBER		N	
ROHGAS_REINGAS	NUMBER		N	
BEZUGS_O2_GEHALT	NUMBER		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
GASPHASE	NUMBER		N	
SCHWEBSTAUB	NUMBER		N	
TROCKENE_DEPOSITION	NUMBER		N	
NASSE_DEPOSITION	NUMBER		N	
GESAMT_DEPOSITION	NUMBER		N	
REGENWASSER	NUMBER		N	
SCHNEE	NUMBER		N	
KRONENTRAUFDEPOSITION	NUMBER		N	
GROESSE_RECHTECK	NUMBER		N	
GROESSE_FLAECH	NUMBER		N	
MASSE	NUMBER		N	
STAUBART	NUMBER		N	
PHASE	NUMBER		N	
GRUNDWASSER	NUMBER		N	
PRODUKTNAME	VARCHAR2		N	
PRODUKTTYP	VARCHAR2		N	
HERSTELLERNAME	VARCHAR2		N	
ANSCHRIFT	VARCHAR2		N	
TEILKOMPARTIMENT	NUMBER		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
FUTTERMITTEL_ART_BEZ	VARCHAR2		N	
IST_TROCKENSUBSTANZ	NUMBER		N	
IST_ORGANISCHE_AUFLAGE	NUMBER		N	
BODENDICHTE	NUMBER		A	m_bodenmesswerte: messwert messgroesse_id=275 einheit_id=28
PROBENMENGE	NUMBER		N	
BEWERTUNG_BELASTUNG	VARCHAR2		A	uf_aufnahmesituation: bemerkung_aufnahmesituation
ID_BEWERTUNG	NUMBER		N	
DATUM	DATE	Unvollständig. (Datum entweder leer oder identisch mit t_probennahme.datum)	N	
SIEBFRAKTION	NUMBER		N	

Tabelle 17: Übersicht T_PROBENNAHME

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
ID	NUMBER	Primärschlüssel	N	
STANDORT_ID	NUMBER	Standort (Verweis)	(N)	Verknüpfen mit t_standort
BODENTYP_ID	VARCHAR2	t_bodentyp (Verweis)	A	uf_profilkennzeichnung:bodentyp_id, lists_bodentypen

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
				uf_profilkennzeichnung:bodensubtyp_id, lists_subbodentypen
PROBENAHRME_NR	VARCHAR2		A	uf_titeldaten:kennung
INSTITUTION_ID	NUMBER	Schlüssel zu t_institution (dort im Wesentlichen nur der Name der Institution)	A	(direkte Aufnahme in eBIS-Struktur nicht direkt möglich, da es sich hier um kaskadierende Quelleninformationen handelt)
ANSCHRIFT_ID	NUMBER		N	
ANSPRECHPARTNER_ID	NUMBER		N	
ENTNAHMEART_ID	VARCHAR2	Verweis auf www_entnahmeart_boden	B	m_bodenmesswerte:probennahmegeeraet_id, lists_probenahmegeeraete
DATUM	DATE	Datum der Probenahme	A	m_bodenmesswerte:probenahmedatum
VERPACKUNG_ART_ID	VARCHAR2		N	
DATUM_BIS	DATE		N	
WIEDERHOLUNG	DATE		N	
DAUER	NUMBER		N	
PROBENAHMESTRATEGIE	VARCHAR2		A	uf_titeldaten:bemerkung_titeldaten (mit Delimiter „;“ zusammenfassen)
		nähere Beschreibung zum beprobten Teilkompartment Boden 0 Keine Angabe 1 Gesamtboden 2 Bodenlösung 3 Sediment 4 Bodenmatrix 5 Bodenluft	A	uf_titeldaten:bemerkung_titeldaten (mit Delimiter „;“ zusammenfassen)
BESCHREIBUNG	NUMBER			

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
		6 Schwebstoffe		
FORSTNUTZUNG	NUMBER		A	uf_aufnahmesituation:nutzung_allgemein_ka5_id=5
GEWAESSERBEREICH	NUMBER		N	
ANDERE_LAGE	VARCHAR2		N	
ZUSTAND	VARCHAR2		A	uf_profilkennzeichnung:bemerkung_profilkennz
BODENMIKROORGANISMEN	NUMBER		N	
BODENTIER	NUMBER		N	
PFLANZE	NUMBER		N	
NAME	VARCHAR2		N	
KOERPER_BZW_PFLANZENTEIL	VARCHAR2		N	
ABFALLART	NUMBER		N	
FILTERART	NUMBER		N	
ANDERER_FILTER	VARCHAR2		N	
FAELLUNGSVERFAHREN	VARCHAR2		N	
ABGAS_ABWASSER_RESTSTOFF	NUMBER		N	
ANDERER_ABGAS_ABWASSER_REST	VARCHAR2		N	
ANDERER_INDUSTRIEABFALL_ID	VARCHAR2		N	
REINIGUNGSVERFAHREN	VARCHAR2		N	
SIEDLUNGSABFALL	NUMBER		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
ANDERER_SIEDLUNGSABFALL	VARCHAR2		N	
SONSTIGE_ABFALLARTEN	NUMBER		N	
ANDERE_ABFALLART	VARCHAR2		N	
SONSTIGE_ABWAESSER	NUMBER		N	
ANDERE_ABWAESSER	VARCHAR2		N	
ANFALLSTELLE	NUMBER		N	
VORBEHANDLUNG	NUMBER		N	
ZULAUF_ABLAUF	NUMBER		N	
NIEDERSCHLAGSMENGE	NUMBER		N	
LUFTFEUCHTIGKEIT	NUMBER		N	
LUFTTEMPERATUR	NUMBER		N	
WINDGESCHWINDIGKEIT	NUMBER		N	
PROBENART	NUMBER	-99 – keine Angabe → 39 1 – Einzelprobe → 10 2 – Referenzprobe → 39 3 – Mischprobe → 17	A	m_bodenmesswerte:beprobungsmethode_id
ANZAHL_ENTNAHMEPUNKTE	NUMBER		N	
ANZAHL_EINZELPROBEN	NUMBER		N	
VERDACHTSPROBE	NUMBER		N	
VERDACHTURSACHE	VARCHAR2		N	
VERFOLGUNGSPROBE	NUMBER		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
REFERENZPROBE	NUMBER		N	
POOLPROBE_KRITERIEN	VARCHAR2		N	
UNTERSUCHUNGLABOR	VARCHAR2		C	m_bodenmesswerte:labor
LABOR_NR	NUMBER		N	
PROBENMENGE	NUMBER		N	
PROBENAHPMEFLAECHE	NUMBER		N	
FLAECHE_EINZELPROBEN	NUMBER		N	
PROBENMENGE_REF_PROBE	NUMBER		N	
ENTNAHMEART	NUMBER		N	
GERAET	NUMBER		N	
ANDERES_GERAET	VARCHAR2		N	
NACH_RICHTLINIE	NUMBER		N	
BLATT_NR	NUMBER	Fremdschlüssel Probennahme	N	
BERGERHOFFGERAETE	NUMBER		N	
LIESEGANG_LOEBNER_GERAETE	NUMBER		N	
AUFSTELLUNGSGEOMETRIE	VARCHAR2		N	
SONSTIGES	VARCHAR2		N	
ANDERE_PROBENAHPMEVORSCHRIFT	VARCHAR2		N	
ABWEICHUNGEN	VARCHAR2		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
MATERIAL	NUMBER	Gefäßmaterial 0 – keine Angabe 1 – Glas → 14 2 – Edelstahl → 60	C	m_bodenmesswerte: probenahmegefaess_id → lists_probengefaessmaterialien
ANDERES_MATERIAL	VARCHAR2	„Aluminium“ → ohne Entsprechung „Kunststoffbeutel“ → ohne Entsprechung „Weithalsflaschen aus PE“ → 36	C	m_bodenmesswerte: probenahmegefaess_i d → lists_probengefaessmaterialien
KUEHLBEHAELTNIS	NUMBER		N	
TRANSPORTDAUER	NUMBER		N	
STD_TAGE	NUMBER		N	
NACH_EMPFEHLUNG	NUMBER		N	
STAUBART	NUMBER		N	
GES_FLAECH	NUMBER		N	
TEIL_FLAECH	NUMBER		N	
ANZ_RECHTECKE	NUMBER		N	
VERTEILUNG	NUMBER		N	
MASSE	NUMBER		N	
VOLUMEN	NUMBER		N	
TIEFE	NUMBER		N	
WASSERFUEHRUNG	NUMBER		N	
STUNDEN_MISCHPROBE	NUMBER		N	
MATERIAL_GERAETE	VARCHAR2		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
MATERIAL_SCHLAEUCHE	VARCHAR2		N	
VOLLSTAENDIGKEIT	NUMBER		N	
PROBENAHMESTOERUNG	NUMBER		N	
WASSERTIEFE	NUMBER		N	
PROBENAHMETIEFE	NUMBER		N	
TRANSPORTTEMPERATUR	NUMBER		N	
LAGERDAUER	NUMBER		N	
LAGERTEMPERATUR	NUMBER		N	
GEFAESS_ID	NUMBER		N	
ZEBS_WARENCODE	VARCHAR2		N	
VERARBEITUNGSZUSTAND_ID	VARCHAR2		N	
VERPACKUNG_LEB_ART_ID	VARCHAR2		N	
VERPACKUNG_SONSTIGE	VARCHAR2		N	
GESCHLECHT	NUMBER		N	
ALTER_SCHLACHTTIERE	NUMBER		N	
GEWICHT_FISCH	NUMBER		N	
KG_G	NUMBER		N	
LAENGE	NUMBER		N	
ERLAEUTERUNG	VARCHAR2		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
EINFLUSSFAKTOREN	VARCHAR2		N	
ART_HOLDUNG_ANBAU	VARCHAR2		N	
ZUSAMMENSETZUNG	VARCHAR2		N	

Tabelle 18: Übersicht T_STANDORT

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
ID	NUMBER	Primärschlüssel	N	
REG_NR_DS	NUMBER		N	
REG_NR_FN	NUMBER		N	
REG_NR_JAHR	NUMBER		N	
GEWAESSER_ID	NUMBER		N	
RESTSTOFF_ID	VARCHAR2		N	
BLM_SCH_V	VARCHAR2		N	
ENTNAHMESTELLE_ID	NUMBER		N	
KOMPARTIMENT_ID	VARCHAR2		N	
GEMEINDEKENNZIFFER	VARCHAR2		N	
STAAT_ID	VARCHAR2		N	
STANDORT_NR_INTERN	VARCHAR2		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
RECHTSWERT	NUMBER	Rechtswert, Gauss-Krüger, führende Streifenkennung	A	uf_titeldaten:rechswert
HOCHWERT	NUMBER	Hochwert, Gauss-Krüger	A	uf_titeldaten:hochwert
HOEHE_UE_NN	NUMBER	Geländehöhe über NN	A	uf_titeldaten:hoehe_ueber_nn
PROBENAHMESTELLE	NUMBER	Probenahmestelle befindet sich 0 - Keine Angabe 6 - Keine Besonderheiten 1 - in einem Naturschutzgebiet 2 - auf einem Truppenübungsplatz 3 - auf einem Gelände, das bereits als Altlast eingestuft wurde 4 - auf Bauerwartungsland 5 - in einem Rekultivierungsgebiet, unterliegt der Bergaufsicht	A	uf_titeldaten:bemerkung_titeldaten
FLUSSKILOMETER	NUMBER		N	
GROESSE	NUMBER		N	
INDUSTRIE	NUMBER		N	
ENTFERNUNG	NUMBER		N	
DIOXINRELEVANTE_ALTANLAGE	NUMBER		N	
VERKEHRSWEGE	NUMBER		N	
VERKEHRSAUFKOMMEN	NUMBER		N	
ALTLASTENVERDACHTSSTANDORT	NUMBER		N	
KLAERSCHLAMM	NUMBER		N	
VERFAHREN	VARCHAR2		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
BETREIBER	VARCHAR2		N	
ZIELPRODUKT	VARCHAR2		N	
NENNLEISTUNG	NUMBER		N	
EINHEIT1	VARCHAR2		N	
CHARGENZAHL	VARCHAR2		N	
LEISTUNG	NUMBER		N	
EINHEIT2	VARCHAR2		N	
BETRIEBSZUSTAND	NUMBER		N	
BETRIEBSTEMPERATUR	NUMBER		N	
VERWEILZEIT	NUMBER		N	
BETRIEB_O2	NUMBER		N	
BEZUG_O2	NUMBER		N	
ABFALLMENGE	NUMBER		N	
ABWASSERMENGE	NUMBER		N	
RESTSTOFFMENGE	NUMBER		N	
VORRICHTUNGEN	VARCHAR2		N	
ABGASVOLUMEN	NUMBER		N	
EINSATZSTOFF1	VARCHAR2		N	
GEHALT1	NUMBER		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
EINSATZSTOFF2	VARCHAR2		N	
GEHALT2	NUMBER		N	
EINSATZSTOFF3	VARCHAR2		N	
GEHALT3	NUMBER		N	
EINSATZSTOFF4	VARCHAR2		N	
GEHALT4	NUMBER		N	
EINSATZSTOFF5	VARCHAR2		N	
GEHALT5	NUMBER		N	
EINSATZSTOFF6	VARCHAR2		N	
GEHALT6	NUMBER		N	
SONST_STOFF	VARCHAR2		N	
VDI2448	NUMBER		N	
SONST_RICHT	VARCHAR2		N	
ART_GEBAEUDE	NUMBER		N	
SONST_ENTNOMMEN	VARCHAR2		N	
ART_RAUM	NUMBER		N	
SONST_RAUM	VARCHAR2		N	
NUTZUNG	VARCHAR2		N	
BRANDSCHADEN	NUMBER		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
BRANDJAHR	DATE		N	
BRAND_UMFANG	VARCHAR2		N	
BRANDSCHADEN_UMG	NUMBER		N	
BRANDJAHR_UMG	DATE		N	
BRAND_UMFANG_UMG	VARCHAR2		N	
HOLZSCHUTZMITTEL	NUMBER		N	
IN_WINDRICHTUNG	NUMBER		N	
LAGE_MESSTELLE	NUMBER		N	
ANDERE_LAGE	VARCHAR2		N	
GROESSE_EINZUGSGEBIET	NUMBER		N	
ABWASSEREINLEITUNG	NUMBER		N	
ART_PROBENAHRME	NUMBER		N	
ARBEITSPLATZ	NUMBER		N	
SONST_ARBEITSPLATZ	VARCHAR2		N	
ENTNOMMEN	NUMBER		N	
VOLLSTAENDIGKEIT	NUMBER		N	
SONSTIGE_AUFBRINGUNG	VARCHAR2		N	
GEMEINDEKENNZIFFER_HERKUNFT	VARCHAR2		N	
BINNENFISCHE_SEEFISCHE	NUMBER		N	

Feld	Datentyp	Bedeutung	Priorität	Entsprechung in eBIS
ZEBS_BINNENGEWAESSER_ID	VARCHAR2		N	
ZEBS_SEEGEWAESSER_ID	VARCHAR2		N	
GEMEINDEKENNZIFFER_GEWAESSER	VARCHAR2		N	
EMITTENTENNAEHE	NUMBER		N	
ENTNAHMESTELLE_LEB_ID	VARCHAR2		N	
INLAND	NUMBER		N	
ENTNAHMESTELLE	VARCHAR2		N	

A.3 Datenbanktabelle Bodendaten aus den bundesweiten Auen

Tabelle 19: Übersicht Datenbanktabelle Auen2016

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
ID	INTEGER	Primärschlüssel		A	uf_titeldaten:uba_kennung
Jahr	INTEGER	Auen-Projektjahr (2012 und 2016; Aue I und Aue II)		A	--
ProbenIdentNr	INTEGER	Probenidentifikationsnummer		U	m_bodenmesswerte:probenummer
StandortIdentNr	INTEGER	Standortidentifikationsnummer		U	--
Vollstaendig_alle	TEXT	Kriterium für statische Auswertungen		N	--
Vollstaendig_Organika	TEXT	Kriterium für statische Auswertungen		N	--
Vollstaendig_KW	TEXT	Kriterium für statische Auswertungen		N	--
Datenquelle	TEXT	Ursprungsdatei/Datenherkunft		A	projekte:name
Institution	TEXT	Datenhaltende Institution		A	--
NRBF	INTEGER			N	--
X_UTM	DOUBLE	UTM32 Koordinate (x) In Verbindung mit lists_koordinatensysteme, dort ido 2 (ETRS89, UTM Zone 32)		A	uf_titeldaten (rechtswert, hochwert, koordinatensystem_ido); lists_koordinatensysteme

Feld	Datentyp	Bedeutung	Vollständigkei t [0..100%]	Prioritä t	Entsprechung eBIS
Y_UTM	DOUBLE	UTM32 Koordinate (y) In Verbindung mit lists_koordinatensysteme, dort ido 2 (ETRS89, UTM Zone 32)		A	uf_titeldaten (rechtswert, hochwert, koordinatensystem_id); lists_koordinatensysteme
KoordX2	TEXT	UTM32 Koordinate (x) als Text		N	uf_titeldaten (rechtswert, hochwert, koordinatensystem_id); lists_koordinatensysteme
KoordY2	TEXT	UTM32 Koordinate (y) als Text		N	uf_titeldaten (rechtswert, hochwert, koordinatensystem_id); lists_koordinatensysteme
KoordX2Y2	TEXT	Zusammengesetzt aus KoordX2 und KoordY2 für Auswertungszwecke im Projekt		N	--
X_UTM_m	DOUBLE	Gerundete UTM32 Koordinate (x)		N	uf_titeldaten (rechtswert, hochwert, koordinatensystem_id); lists_koordinatensysteme
Y_UTM_m	DOUBLE	Gerundete UTM32 Koordinate (y)		N	uf_titeldaten (rechtswert, hochwert, koordinatensystem_id); lists_koordinatensysteme
Messstelle	TEXT	Messstelle zur Standortkennung		U	--
Fluss_Detail	TEXT	Name oder Abschnitt des Flusses nach Hydrologischer Atlas Deutschland (HAD) zur Standortkennung		C	--
EZG	TEXT	Einzugsgebiet (EZG) nach HAD aggregiert zur Standortkennung		C	--

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
EZG2	TEXT	Einzugsgebiet (EZG) weiter aggregiert zur Standortkennung		C	--
Testgebiet	TEXT	Zugehörigkeit zum Testgebiet aus erstem Auenprojekt		C	--
Profil_ID	TEXT			U	uf_profilkennzeichnung:profil_nummer
Standort_ID	TEXT	ähnlich „Messstelle“, aber mit weiteren Einträgen		B	uf_titeldaten:name
Proben_ID	TEXT			U	uf_titeldaten:probenummer
Bundesland	TEXT	Kürzel des Bundeslandes		A	uf_titeldaten:bundesland_ido → lists_bundeslaender
DienststellenNr	DOUBLE			N	--
Dienststelle	TEXT			N	--
zustaendDienststelle	TEXT			N	--
Projekt_Nr	DOUBLE			N	--
Projekt	TEXT	Projekt, aus dem Daten stammen		C	projekte:projekt
Probenahmedatum	DATETIME	Datum der Probenahme		A	m_bodenmesswerte;probenahmedatum
StandortNr	DOUBLE			N	--
Standortname	TEXT	Name des Standorts, z.B. auch BDF Standort		A	uf_titeldaten:beschreibung

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Nutzung	TEXT	Nutzung, z.B. Acker oder Grünland		A	uf_aufnahmesituation:nutzung_allgemein_ka5_ido → lists_nutzungsarten_ka5 uf_aufnahmesituation:nutzung_speziell_ka5_ido → lists.nutzungsarten_ka5_speziell
Vegetation	TEXT	Vegetation		A	uf_aufnahmesituation:vegetation_ka5_ido → lists.vegetation_ka5 uf_aufnahmesituation:vegetation_ka4_ido → lists.vegetationen_ka4_speziell
Bemerkung	TEXT	Bemerkungen zur Probe/Standort		N	--
Bodentyp_Name	TEXT	Name des Bodentyps/Bodenform		A	m_bodenmesswerte:bemerkung
Bodentyp	TEXT	Kürzel des Bodentyps		A	uf_profilkennzeichnung:bodentyp_ido → lists.bodentypen uf_profilkennzeichnung:bodensubtyp_ido → lists.bodensubtypen; uf_profilkennzeichnung:bemerkung_profilkennz
Substrattyp	TEXT	Substrattyp		A	uf_profil_substratsystematik_typ:substratgenese_gruppe_ido → lists.substratgenese_gruppen
Bodensubtyp	TEXT	Bodensubtyp		N ¹⁸	--
HORIZONT_NR	DOUBLE			N	--
Horizontbezeichnung	TEXT	Horizontbezeichnung nach KA 5		A	Aufteilung in Haupt-, Vor- und Nachsymbole
HorTiefe_von	DOUBLE	Horizonttiefe Oberkante in cm		A	h_kennzeichnung:horizont_obergrenze

¹⁸ Ohne Inhalt, Subtypen sind in der Spalte „Bodentyp“ erfasst.

Feld	Datentyp	Bedeutung	Vollständigkei t [0..100%]	Prioritä t	Entsprechung eBIS
HorTiefe_bis	DOUBLE	Horizont-Unterkante in cm		A	h_kennzeichnung:horizont_untergrenze
Bodenschicht	TEXT			N	--
Humusgehalt	TEXT	Humusgehaltsklasse nach KA 5		A	h_eigenschaften:humusgehalt_ido → lists_humusanteile
Bodenart_orig	TEXT	Kürzel Bodenart		N	--
Bodenart_ber_orig	TEXT			N	--
Bodenart_KA5_orig	TEXT	Kürzel Bodenart		A	h_eigenschaften:bodenart_ka_5_ka_4_ido →lists_bodenarten_bdf2; h_eigenschaften:muddeart_ido → lists_muddearten h_eigenschaften: bodenart_reiner_sand_ka5_ido→ lists_reinsande; h_eigenschaften:torfartengruppen_ido → lists_torfartengruppen; h_eigenschaften:torfarten_ido → lists.torfarten; h_kennzeichnung;horizont_hauptsymbol_ido →lists.horizonthauptsymbole; uf_profil_substratsystematik_klasse:bodenarten_hauptgruppe_id o → lists.substratbodenart_hauptgruppen; uf_profil_substratsystematik_subtyp:bodenarten_gruppe_ido → lists.substratbodenart_untergruppen
Hauptbodenart_orig	TEXT	Hauptbodenart		N	--
Bodenartengruppe_orig	TEXT			N	--

Feld	Datentyp	Bedeutung	Vollständigkei t [0..100%]	Prioritä t	Entsprechung eBIS
Parameter_orig	TEXT	Gemessener Parameter mit Einheit		N	
Param_tmp	TEXT			N	--
Extrakt_tmp	TEXT			N	--
Wert_orig	DOUBLE	Gemessener Wert		N	--
Messwert_tmp	DOUBLE			N	--
Einheit_alt	TEXT			N	--
LD_Text	TEXT	Lagerungsdichte klassiert nach KA 5		A	h_eigenschaften:effektive_lagerungsdichte_ido → lists_lagerungsdichte; h_eigenschaften:substanzanteil_ido → lists.substanzvolumina
Bodenart_KA5	TEXT	Bodenart		N	--
Bodarthauptgr_KA5	TEXT	Bodenartenhauptgruppe nach KA 5		N	--
Bodartgr_KA5	TEXT	Bodenartengruppe nach KA 5		N	--
Bodartgr_KA5_abgeleitet	TEXT			N	--
Bodart_herkunft	TEXT			N	--
Hauptbodenart	TEXT	Hauptbodenart		N	--
Carbonatgehalt	TEXT	CaCO3 Gehaltsklasse nach KA 5		A	h_eigenschaften:carbonargehalt_ido; lists_carbonatanteile

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Stoffgruppe	TEXT	Alle Stoffgruppen		A	(implizit über datTEF_PCDDDF_dIPCB)
Stoffgruppe2	TEXT	Stoffgruppe Dioxine/Furane		A	(implizit über datTEF_PCDDDF_dIPCB)
Extrakt	TEXT	Extraktionsmittel		A	m_bodenmesswerte:extrakt_ido → lists.extrakte_boden
Parameter	TEXT	Gemessener Parameter		A	m_bodenmesswerte:messgroesse_ido → lists.messgroessen
Methode	TEXT			A	m_bodenmesswerte:messverfahren_ido → lists.messverfahren m_bodenmesswerte: labormethode_ido → lists_labormethoden
Zusatz	TEXT	Zur Charakterisierung der Messwerte (z.B. bedeutet < → kleiner Bestimmungsgrenze)		A	m_bodenmesswerte:zusatz_ido → lists.messwertzusaetze
Messwert	DOUBLE	Messwert		A	m_bodenmesswerte:messwert
Messwert_BG0	DOUBLE	Messwert, unterhalb der Bestimmungsgrenze auf 0 gesetzt, wenn Zusatz → < (=lower bound Ansatz)		N	--
Einheit	TEXT	Einheit der Messgröße/gemessener Parameter		A	m_bodenmesswert:einheit_ido → lists.einheiten (Einheiten für Toxizitätsäquivalente weglassen)
Skelett	TEXT	Bodenskelettanteil		N	--
16_PAK	INTEGER	Oben schon vorhanden		N	--
6_PCB	INTEGER	Oben schon vorhanden		N	--
OTIEFE	DOUBLE	Oben schon vorhanden		N	--
PAK16_Sum	DOUBLE	Oben schon vorhanden		N	--

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
PCB6_Sum	DOUBLE	Oben schon vorhanden		N	--
BaP	DOUBLE	Oben schon vorhanden		N	--
EZG_HAD	TEXT	Oben schon vorhanden		N	--
KoordX_m	TEXT	Oben schon vorhanden		N	--
KoordY_m	TEXT	Oben schon vorhanden		N	--
Koord_XY_m	INTEGER	zusammengesetzte Koordinaten, die den Schlüssel für weitere Auswertungsdateien bilden		A	uf_titeldaten → uba_kennung

A.4 Datenbanktabelle Bodenzustandserhebung Wald II

Tabelle 20: Übersicht Tabellen BZE_Organika_0_5_neu, BZE_Organika_5_10_neu, BZE_Organika_Auflage_neu

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
BFHNR	double		(Bundes-)Kennnummer des BZE-Punktes; BFH = Bundesforschungsanstalt für Forst- und		A	uf_titeldaten:uba_kennung

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
			Holzwirtschaft (jetzt Thünen-Institut)			
LCode	double		Codiertes Bundesland		N	
Land	varchar		Kürzel Bundesland		A	uf_titeldaten:bundesland_id → lists_bundeslaender
PName	varchar		Probenname (vom Labor vergeben)		A	m_bodenmesswert:probenummer
erhebjahr_2	double		Jahr der Profilaufnahme		B	m_bodenmesswert:probenahmedatum (01.01. des jeweiligen Jahres)
Rechts	double		Gauß-Krüger Rechtswert		A	uf_titeldaten.rechtswert
Hoch	double		Gauß-Krüger Hochwert		A	uf_titeldaten:hochwert
Horizont_Tiefenstufe	varchar		Tiefenstufe der Beprobung (Auflage, 0-5 cm, >5-10 cm)		A	m_bodenmesswerte:entnahmetiefe_ok/entnahmetiefe_uk
NAPH	double		Naphthalin (µg/kg)		A	m_bodenmesswerte:messwert messgroesse_id: 170 einheit_id: 6
ACY	double		Acenaphthylen (µg/kg)		A	m_bodenmesswerte:messwert messgroesse_id: 34 einheit_id: 6
ACE	double		Acenaphthen (µg/kg)		A	m_bodenmesswerte:messwert messgroesse_id: 33

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						einheit_ido: 6
FLOU	double		Fluoren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 104 einheit_ido: 6
PHE	double		Phenanthren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 224 einheit_ido: 6
ANT	double		Anthracen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 40 einheit_ido: 6
FLA	double		Fluoranthren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 103 einheit_ido: 6
PYR	double		Pyren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 236 einheit_ido: 6
BAANT	double		Benzo[a]anthracen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 48 einheit_ido: 6
CHRY	double		Chrysen (µg/kg)		A	m_bodenmesswerte:

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						messwert messgroesse_ido: 65 einheit_ido: 6
BBFLOU	double		Benzo[b]fluoranthen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 50 einheit_ido: 6
BKFLOU	double		Benzo[k]fluoranthen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 52 einheit_ido: 6
BAP	double		Benzo[a]pyren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 49 einheit_ido: 6
IND	double		Indeno[1,2,3-cd]pyren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 126 einheit_ido: 6
DIBENZO	double		Dibenzo[a,h]anthracen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 85 einheit_ido: 6
BGHIPER	double		Benzo[ghi]perylen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 51 einheit_ido: 6

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
PCB28	double		Polychlorierte Biphenyle 28 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 212 einheit_ido: 6
PCB52	double		Polychlorierte Biphenyle 52 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 214 einheit_ido: 6
PCB101	double		Polychlorierte Biphenyle 101 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 196 einheit_ido: 6
PCB138	double		Polychlorierte Biphenyle 138 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 202 einheit_ido: 6
PCB153	double		Polychlorierte Biphenyle 153 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 203 einheit_ido: 6
PCB180	double		Polychlorierte Biphenyle 180 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 209 einheit_ido: 6
Aldrin	double		Aldrin (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 37

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						einheit_ido: 6
Dieldrin	double		Dieldrin (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 88 einheit_ido: 6
HCB	double		Hexachlorbenzol (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 121 einheit_ido: 6
aHCH	double		alpha Hexachlorcyclohexan (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 39 einheit_ido: 6
bHCH	double		beta Hexachlorcyclohexan (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 54 einheit_ido: 6
gHCH	double		gamma Hexachlorcyclohexan (Lindan) (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 110 einheit_ido: 6
24DDE	double		24 Dichlordiphenyldichlorethen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 188 einheit_ido: 6

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
44DDE	double		44 Dichlordiphenyldichlorethen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 231 einheit_ido: 6
24DDD	double		24 Dichlordiphenyldichlorethan (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 187 einheit_ido: 6
44DDD	double		44 Dichlordiphenyldichlorethan (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 230 einheit_ido: 6
24DDT	double		24 Dichlordiphenyltrichlorethan (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 189 einheit_ido: 6
44DDT	double		44 Dichlordiphenyltrichlorethan (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 232 einheit_ido: 6
SumPAH	double		Summe Polycyclische Aromatische Kohlenwasserstoffe (µg/kg)		N	
SumHCH	double		Summe Hexachlorcyclohexan (µg/kg)		N	

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
SumPCB	double		Summe Polychlorierte Biphenyle (6 Stück) (µg/kg)		N	
SumDDX	double		Summe DDE, DDD, DDT (µg/kg)		N	
TS	double		Trockensubstanz (%)		B	m_bodenmesswerte: messwert messgroesse_ido: 301 einheit_ido: 11
TOC	double		Total organic carbon (organischer Gesamtkohlenstoff) (%)		A	m_bodenmesswerte: messwert messgroesse_ido: 72 einheit_ido: 11
Humus_M%	double		Humusgehalt (Masse-%)		A	m_bodenmesswerte: messwert messgroesse_ido: 123 einheit_ido: 3
Humusklasse_HGW	varchar		Humusklasse für Hintergrundwerte-Auswertung ¹⁹		N	
Humusklasse_KA5	varchar		Humusklasse nach KA 5 (Bodenkundliche Kartieranleitung)		A	h_eigenschaften:humusgehalt_ido → lists.humusanteile (WHERE humusklasse_ka5 = lists.humusanteile.kurzwert)
Feldhöhe	double		Höhe über NN		C	uf_titeldaten:hoehe_ueber_nn

¹⁹ In BZE_Organika_5_10_neu und BZE_Organika_Auflage_neu nicht vorhanden

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
BestTyp1	double		Codierter Dominierender Bestockungstyp am BZE-Punkt		N	--
BestTyp kurz_1	varchar		Abkürzung Dominierender Bestockungstyp am BZE-Punkt		N	--
BestTyp lang_2	varchar		Dominierender Bestockungstyp am BZE-Punkt		B	-- (inhaltlich passen in uf_aufnahmesituation, dort gibt es aber kein passendes Feld für den „Bestockungstyp“)
BestTyp Sort	double		Codierte Baumart		N	--
Baumart	varchar		Baumart		B	-- (inhaltlich passen in uf_aufnahmesituation, dort gibt es aber kein passendes Feld für die „Baumart“)
Waldtyp Nr_2	double		Codierter Waldtyp		N	--
Waldtyp_2	varchar		Waldtyp		A	uf_aufnahmesituation:vegetation_ka5_id → lists_vegetationen_ka5
Bodentyp	varchar		Abkürzung Bodentyp		A	uf_profilkennzeichnung:bodensuptyp_id → lists_bodensubtypen uf_profilkennzeichnung:bodentyp_id → lists_bodentypen
Bodentyp lang	varchar		Bodentyp		A	uf_profilkennzeichnung:bemerkung_profilkennz
Hu Form lang_2	varchar		Humusform		C	--
Hu Form kurz_2	varchar		Abkürzung Humusform		B	uf_profilkennzeichnung:humusform_detailliert_id → lists.humusformen

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Hu Klasse	varchar		Humusklasse		N	--
SHN Haupteinheiten	varchar		Codierte Haupteinheit		B	uf_titeldaten:naturraeumliche_haupteinheitensgruppe_ido → lists.naturraeumliche_gliederungen_gruppen (1:1-Relation zu naturraeumliche_gliederungen_gruppen.schluesel)
Haupteinheit lang	varchar		Naturräumliche Haupteinheit		N	(vgl. SHN Haupteinheiten)
Naturraum Kennung	double		Codierter Naturraum		N	(vgl. Naturraum)
Naturraum	varchar		Name Naturraum		C	uf_titeldaten:naturraeumliche_haupteinheit_ido → lists.naturraeumliche_gliederungen_grob
Verdichtungsraum	varchar		Verdichtungsraum (ja/nein)		A	-- (keine Entsprechung)
Name_des_Verdichtungsraums	varchar		Name Verdichtungsraum		A	-- (keine Entsprechung, da eBIS feinere Granularität fordert)
Auenstandort	varchar		Auenstandort (ja/nein) ²⁰		A	--
Auentyp	varchar		Auentyp alt/rezent) ²¹		A	--
Niederschlag_DWD71_mm_1991_2000	double		Niederschlag (mm/a)		C	uf_klima: jahresmittel_niederschlag uf_klima:jm_niederschlag_periode_ido: 2
TRD	double		Trockenrohdichte (g/cm3)		A	m_bodenmesswerte: messwert

²⁰ In BZE_Organika_5_10_neu und BZE_Organika_Auflage_neu nicht vorhanden

²¹ In BZE_Organika_5_10_neu und BZE_Organika_Auflage_neu nicht vorhanden

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						messgroesse_ido: 275 einheit_ido: 28
Bodenart	varchar		Kürzel Bodenart		A	Doppelt schlüsseln: h_eigenschaften:bodenart_ka_5_ka_4_ido → lists.bodenarten_bdf2 h_eigenschaften: bodenart_reiner_sand_ka5_ido → lists.reinsande
Sand	double		Sand (%)		A	m_bodenmesswerte: messwert messgroesse_ido: 254 einheit_ido: 11
Schluff	double		Schluff (%)		A	m_bodenmesswerte: messwert messgroesse_ido: 280 einheit_ido: 11
Ton	double		Ton (%)		A	m_bodenmesswerte: messwert messgroesse_ido: 276 einheit_ido: 11
WE	double		West-Ost („alte“ und „neue“ Bundesländer)		N	-- (implizit)

Tabelle 21: Übersicht Tabelle PCDDF_dl_PCB_Wald_Auflage²²

Feld	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
PCB 77	PCB 77		A	m_bodenmesswerte: messwert messgroesse_ido: 215 einheit_ido: 112
PCB 81	PCB 81		A	m_bodenmesswerte: messwert messgroesse_ido: 217 einheit_ido: 112
PCB 126	PCB 126		A	m_bodenmesswerte: messwert messgroesse_ido: 201 einheit_ido: 112
PCB 169	PCB 169		A	m_bodenmesswerte: messwert messgroesse_ido: 207 einheit_ido: 112
PCB 105	PCB 105		A	m_bodenmesswerte: messwert messgroesse_ido: 197 einheit_ido: 112
PCB 114	PCB 114		A	m_bodenmesswerte: messwert messgroesse_ido: 198 einheit_ido: 112

²² Überführung in eine normalisierte Lookup-Tabelle, die mit eBIS-Stoffnamen kompatibel ist

Feld	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
PCB 118	PCB 118		A	m_bodenmesswerte: messwert messgroesse_ido: 199 einheit_ido: 112
PCB 123	PCB 123		A	m_bodenmesswerte: messwert messgroesse_ido: 200 einheit_ido: 112
PCB 156	PCB 156		A	m_bodenmesswerte: messwert messgroesse_ido: 204 einheit_ido: 112
PCB 157	PCB 157		A	m_bodenmesswerte: messwert messgroesse_ido: 205 einheit_ido: 6
PCB 167	PCB 167		A	m_bodenmesswerte: messwert messgroesse_ido: 206 einheit_ido: 112
PCB 189	PCB 189		A	m_bodenmesswerte: messwert messgroesse_ido: 210 einheit_ido: 112
2,3,7,8-TCDD	2,3,7,8-TCDD		A	m_bodenmesswerte: messwert messgroesse_ido: 17 einheit_ido: 112

Feld	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
1,2,3,7,8-PeCDD	1,2,3,7,8-PeCDD		A	m_bodenmesswerte: messwert messgroesse_ido: 12 einheit_ido: 112
1,2,3,4,7,8-HxCDD	1,2,3,4,7,8-HxCDD		A	m_bodenmesswerte: messwert messgroesse_ido: 4 einheit_ido: 112
1,2,3,6,7,8-HxCDD	1,2,3,6,7,8-HxCDD		A	m_bodenmesswerte: messwert messgroesse_ido: 8 einheit_ido: 112
1,2,3,7,8,9-HxCDD	1,2,3,7,8,9-HxCDD		A	m_bodenmesswerte: messwert messgroesse_ido: 10 einheit_ido: 112
1,2,3,4,6,7,8-HpCDD	1,2,3,4,6,7,8-HpCDD		A	m_bodenmesswerte: messwert messgroesse_ido: 1 einheit_ido: 112
OCDD	OCDD		A	m_bodenmesswerte: messwert messgroesse_ido: 184 einheit_ido: 112
2,3,7,8-TCDF	2,3,7,8-TCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 18 einheit_ido: 112

Feld	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
1,2,3,7,8-PeCDF	1,2,3,7,8-PeCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 13 einheit_ido: 112
2,3,4,7,8-PeCDF	2,3,4,7,8-PeCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 16 einheit_ido: 112
1,2,3,4,7,8-HxCDF	1,2,3,4,7,8-HxCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 5 einheit_ido: 112
1,2,3,6,7,8-HxCDF	1,2,3,6,7,8-HxCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 9 einheit_ido: 112
1,2,3,7,8,9-HxCDF	1,2,3,7,8,9-HxCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 11 einheit_ido: 112
2,3,4,6,7,8-HxCDF	2,3,4,6,7,8-HxCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 15 einheit_ido: 112
1,2,3,4,6,7,8-HpCDF	1,2,3,4,6,7,8-HpCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 2 einheit_ido: 112

Feld	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
1,2,3,4,7,8,9-HpCDF	1,2,3,4,7,8,9-HpCDF		A	m_bodenmesswerte: messwert messgroesse_ido: 3 einheit_ido: 112
OCDF	OCDF (Stoffname)		A	m_bodenmesswerte: messwert messgroesse_ido: 185 einheit_ido: 112
PCB 77_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 81_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 126_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 169_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 105_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 114_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 118_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 123_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 156_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	

Feld	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
PCB 157_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 167_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCB 189_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
2,3,7,8-TCDD_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,7,8-PeCDD_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,4,7,8-HxCDD_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,6,7,8-HxCDD_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,7,8,9-HxCDD_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,4,6,7,8-HpCDD_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
OCDD_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
2,3,7,8-TCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,7,8-PeCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	

Feld	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
2,3,4,7,8-PeCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,4,7,8-HxCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,6,7,8-HxCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,7,8,9-HxCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
2,3,4,6,7,8-HxCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,4,6,7,8-HpCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
1,2,3,4,7,8,9-HpCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
OCDF_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
dI_PCB_TEQ_2005	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	
PCDD_F_TEQ05	TEQ2005 = Toxizitätsäquivalent aus dem Jahr 2005 für den Stoff		N	

A.5 Datenbanktabellen Organische und Anorganische Hintergrundgehalte

Tabelle 22: Übersicht Tabelle Access_LW

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Standort_ID	VARCHAR		Eindeutige ID für die Zuordnung zu eBIS		A	uf_titeldaten:uba_kennung
Bundesland	VARCHAR		Bundesland		A	uf_titeldaten:bundesland_id → lists_bundeslaender
Großlandschaft Kennung	DOUBLE		Codierung der Großlandschaft		N	(redundant zu GL-Name)
Großlandschaften Name	VARCHAR		Geographische Landschaftsbezeichnung		C	uf_titeldaten:naturraeumliche_grossregion_id → lists_naturraeumliche_grossregionen
SHN Haupteinheiten	VARCHAR		Codierte Haupteinheit		B	uf_titeldaten:naturraeumliche_haupteinheitengruppe_id → lists_naturraeumliche_gliederungen_gruppen
Haupteinheiten lang	VARCHAR		Naturräumliche Haupteinheit		N	(Redundant zu codierter Haupteinheit)
Naturraum Kennung	DOUBLE		Codierter Naturraum		N	(redundant zu Naturraum)
Naturraum	VARCHAR		Name Naturraum		C	uf_titeldaten:naturraeumliche_haupteinheit_id → lists_naturraeumliche_gliederungen_grob
Niederschlag (mm J)	DOUBLE		Durchschnittlicher Jahresniederschlag :		N	(uf_klima – Bezugszeitraum unklar, daher nicht verwenden))

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Hoehe_NN	DOUBLE		Höhe über NN		C	uf_titeldaten:hoehe_ueber_nn ²³
Datenherkunft	VARCHAR		Projekt aus dem die Daten stammen		A	projekte:projekt
Kontakt	VARCHAR		Betreuung der Flächen		A	(direkte Aufnahme in eBIS-Struktur nicht direkt möglich, da es sich hier um kaskadierende Quelleninformationen handelt)
Probenahmedatum	TIMESTAMP		Probenahmedatum		A	uf_bodenmesswerte:probenahmedatum
Beprobungsmethode	VARCHAR		Kennzeichnet wie die Proben über die Fläche genommen wurden (Satellitenmethode, Quadratmetermethode)		N	(m_bodenmesswerte:beprobungsmethode_id → lists.beprobungsmethoden_mixe)
Standort_lang	VARCHAR		Standortbeschreibung		B	uf_titeldaten:name_bzw_ort
Verdichtungsraum	VARCHAR		Verdichtungsraum ja/nein		A	-- (nicht abbildbar, eBIS benötigt den konkreten Verdichtungsraum)

²³ Es wird die Nutzung eines Digitalen Höhenmodells empfohlen

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Nr_Verdichtungsraum	DOUBLE		Codierung des Verdichtungsraumes		N	--
Name_Verdichtungsraum	VARCHAR		Name Verdichtungsraum		A	-- (nicht abbildbar, eBIS verlangt eine feinere Schlüsselung)
Breitengrad	DOUBLE		Standortkoordinaten		A	uf_titeldaten:hochwert → koordinatensystem_ido=4 (WGS84)
Längengrad	DOUBLE		Standortkoordinaten		A	uf_titeldaten:rechtswert → koordinatensystem_ido=4 (WGS84)
Nutzung_allgemein	VARCHAR		Landnutzung (Acker/Grünland)		A	uf_aufnahmesituation:nutzung_allgemein_ka5_id → lists_nutzungsarten_ka5 (WHERE Nutzung_allgemein = schluessel)
Auflage-Horizontbezeichnung	VARCHAR		Bezeichnung der Oberboden-Horizonte (z.B. Acker Ap)		A	h_kennzeichnung: (Zerlegung in Vor-, Haupt- und Nachsymbole) horizontsymbolstring
OK Auflage Horizont (cm)	DOUBLE		Horizonttiefe Oberkante in cm		A	uf_bodenmesswerte:entnahmetiefe_ok
UK Auflage Horizont	VARCHAR		Horizonttiefe Unterkante in cm		A	uf_bodenmesswerte:entnahmetiefe_uk (negative Werte nicht übernehmen – wahrscheinlich ist das versehentlich enthalten in dem Sinne z.B. „bis 7 cm“ zu verstehen) (Einzelwert „+“ nicht übernehmen)
Bodenartenhauptgruppe	VARCHAR		Bodenartenhauptgruppe nach KA 5		A	h_eigenschaften:bodenartenhauptgruppe_id → lists.substratbodenart_hauptgruppen

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Bodenart (KA5)	VARCHAR		Bodenart nach KA5		A	h_eigenschaften:bodenart_ka_5_ka_4_ido → lists_bodenarten_bdf2
Humusklasse	VARCHAR		Humusklasse		A	h_eigenschaften:humusgehalt_id → lists.humusanteile (where kurzwert=humusklasse)
Carbonatgehalt (KA5)	VARCHAR		CaCO ₃ Gehaltsklasse nach KA 5		A	h_eigenschaften:carbonargehalt_id → lists.carbonatanteile (where carbonatgehalt_ka5=kurzwert)
Humus_Ma-%	DOUBLE		Humusgehalt, Masseprozent		A	m_bodenmesswerte: messwert messgroesse_ido: 123 einheit_ido: 3
Tongehalt_Ma%	DOUBLE		Tongehalt, Masseprozent		A	m_bodenmesswerte: messwert messgroesse_ido: 276 einheit_ido: 3
Schluffgehalt_Ma%	DOUBLE		Schluffgehalt, Masseprozent		A	m_bodenmesswerte: messwert messgroesse_ido: 280 einheit_ido: 3
Sandgehalt_Ma%	DOUBLE		Sandgehalt, Masseprozent		A	m_bodenmesswerte: messwert messgroesse_ido: 254 einheit_ido: 3
Trockenrohdichte_ g cm ³	VARCHAR		Trockenrohdichte des Bodens		A	m_bodenmesswerte: messwert

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						messgroesse_ido: 275 einheit_ido: 28
pH_CaCl2	DOUBLE		pH in CaCl2 gemessen		A	m_bodenmesswerte: messwert messgroesse_ido: 227 einheit_ido: NULL (Einheit wird bei pH nicht genannt) extrakt_id: 56 (lists_extrakte_boden)
Acenaphten_µg kg	DOUBLE		Acenaphten (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 33 einheit_ido: 6
Acenaphten_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	m_bodenmesswerte: messwert messgroesse_ido: 33 einheit_ido: 6 zusatz_ido: 2
Acenaphtylen_µg kg	DOUBLE		Acenaphtylen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 34 einheit_ido: 6
Acenaphtylen_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Anthracen_µg kg	DOUBLE		Anthracen (µg/kg)		A	m_bodenmesswerte: messwert

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						messgroesse_ido: 40 einheit_ido: 6
Anthracen_μg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Benzo(a)anthracen_μg kg	DOUBLE		Benzo(a)anthracen (μg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 48 einheit_ido: 6
Benzo(a)anthracen_μg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Benzo(a)pyren_μg kg	DOUBLE		Benzo(a)pyren (μg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 49 einheit_ido: 6
Benzo(a)pyren_μg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Benzo(b)fluoranthen_μg kg	DOUBLE		Benzo(b)fluoranthen (μg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 50 einheit_ido: 6

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Benzo(b)fluoranthen_μg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Benzo(g,h,i)perylen_μg kg	DOUBLE		Benzo(g,h,i)perylen (μg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 51 einheit_ido: 6
Benzo(g,h,i)perylen_μg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Benzo(k)fluoranthen_μg kg	DOUBLE		Benzo(k)fluoranthen (μg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 52 einheit_ido: 6
Benzo(k)fluoranthen_μg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Chrysen_μg kg	DOUBLE		Chrysen (μg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 65 einheit_ido: 6
Chrysen_μg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Dibenz(a,h)anthracen_µg kg	DOUBLE		Dibenz(a,h)anthracen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 85 einheit_ido: 6
Dibenz(a,h)anthracen_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Fluoranthen_µg kg	DOUBLE		Fluoranthen (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 103 einheit_ido: 6
Fluoranthen_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Fluoren_µg kg	DOUBLE		Fluoren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 104 einheit_ido: 6
Fluoren_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Indeno(1,2,3,-cd)pyren_µg kg	DOUBLE		Indeno(1,2,3,-cd)pyren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 126 einheit_ido: 6

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Indeno(1,2,3,-cd)pyren_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Naphtalin_µg kg	DOUBLE		Naphtalin (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 170 einheit_ido: 6
Naphtalin_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Phenantren_µg kg	DOUBLE		Phenantren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 224 einheit_ido: 6
Phenantren_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Pyren_µg kg	DOUBLE		Pyren (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 236 einheit_ido: 6
Pyren_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Summe PAK Lowerbound	DOUBLE		Summe der PAK16 mit lowerbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf 0 (Null) gesetzt)		N	
Summe PAK Upperbound	DOUBLE		Summe der PAK16 mit upperbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf Bestimmungsgrenze gesetzt)		N	
PCB101_µg kg	DOUBLE		Polychlorierte Biphenyle 101 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 196 einheit_ido: 6
PCB101_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
PCB138_µg kg	DOUBLE		Polychlorierte Biphenyle 138 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 202 einheit_ido: 6

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständig keit [0..100%]	Priorität	Entsprechung eBIS
PCB138_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB153_µg kg	DOUBLE		Polychlorierte Biphenyle 153 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 203 einheit_ido: 6
PCB153_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB180_µg kg	DOUBLE		Polychlorierte Biphenyle 180 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 209 einheit_ido: 6
PCB180_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB28_µg kg	DOUBLE		Polychlorierte Biphenyle 28 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 212 einheit_ido: 6
PCB28_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
PCB52_µg kg	DOUBLE		Polychlorierte Biphenyle 52 (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 214 einheit_ido: 6
PCB52_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Summe PCB Lowerbound	DOUBLE		Summe der PCB6 mit lowerbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf 0 (Null) gesetzt)		N	
Summe PCB Upperbound	DOUBLE		Summe der PCB6 mit upperbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf Bestimmungsgrenze gesetzt)		N	
alpha_HCH_µg kg	DOUBLE		Alpha-HCH (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 39 einheit_ido: 6

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
alpha_HCH_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
beta_HCH_µg kg	DOUBLE		Beta-HCH (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 54 einheit_ido: 6
beta_HCH_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
gamma_HCH_µg kg	DOUBLE		Gamma-HCH (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 110 einheit_ido: 6
gamma_HCH_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
delta_HCH_µg kg	DOUBLE		Delta-HCH (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 83 einheit_ido: 6
delta_HCH_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
epsilon-HCH_µg kg	DOUBLE		Epsilon-HCH (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 97 einheit_ido: 6
epsilon-HCH_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Summe HCH Lowerbound	DOUBLE		Summe der HCH mit lowerbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf 0 (Null) gesetzt)		A	
Summe HCH Upperbound	DOUBLE		Summe der HCH mit upperbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf Bestimmungsgrenze gesetzt)		A	
HCB_µg kg	DOUBLE		HCB (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 121 einheit_ido: 6

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
HCB_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
op-DDE_µg kg	DOUBLE		op-DDE (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 188 einheit_ido: 6
op-DDE_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
pp-DDE_µg kg	DOUBLE		pp-DDE (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 231 einheit_ido: 6
pp-DDE_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
op-DDD_µg kg	DOUBLE		op-DDD (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 187 einheit_ido: 6
op-DDD_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
pp-DDD_µg kg	DOUBLE		pp-DDD (µg/kg)		A	m_bodenmesswerte:

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						messwert messgroesse_ido: 230 einheit_ido: 6
pp-DDD_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
op-DDT_µg kg	DOUBLE		op-DDT (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 189 einheit_ido: 6
op-DDT_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
pp-DDT_µg kg	DOUBLE		pp-DDT (µg/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 232 einheit_ido: 6
pp-DDT_µg kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Summe DDX Lowerbound	DOUBLE		Summe der DDX mit lowerbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze)		N	

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
Summe DDX Upperbound	DOUBLE		e auf 0 (Null) gesetzt) Summe der DDX mit upperbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf Bestimmungsgrenze gesetzt)		N	
2,3,7,8 Tetra-CDD_ng kg	DOUBLE		2,3,7,8 Tetra-CDD (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 17 einheit_ido: 112
2,3,7,8 Tetra-CDD_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
1,2,3,7,8 Penta-CDD_ng kg	DOUBLE		1,2,3,7,8 Penta-CDD (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 12 einheit_ido: 112
1,2,3,7,8 Penta-CDD_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
1,2,3,4,7,8 Hexa-CDD_ng kg	DOUBLE		1,2,3,4,7,8 Hexa-CDD (ng/kg)		A	m_bodenmesswerte: messwert

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
1,2,3,4,7,8 Hexa-CDD_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	messgroesse_ido: 4 einheit_ido: 112 <i>zusatz_ido: 2</i>
1,2,3,6,7,8 Hexa-CDD_ng kg	DOUBLE		1,2,3,6,7,8 Hexa-CDD (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 8 einheit_ido: 112
1,2,3,6,7,8 Hexa-CDD_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
1,2,3,7,8,9 Hexa-CDD_ng kg	DOUBLE		1,2,3,7,8,9 Hexa-CDD (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 10 einheit_ido: 112
1,2,3,7,8,9 Hexa-CDD_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
1,2,3,4,6,7,8 Hepta-CDD_ng kg	DOUBLE		1,2,3,4,6,7,8 Hepta-CDD (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 1 einheit_ido: 112

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
1,2,3,4,6,7,8 Hepta-CDD_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Octa-CDD_ng kg	DOUBLE		Octa-CDD (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 184 einheit_ido: 112
Octa-CDD_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
2,3,7,8 Tetra-CDF_ng kg	DOUBLE		2,3,7,8 Tetra-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 18 einheit_ido: 112
2,3,7,8 Tetra-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
1,2,3,7,8 Penta-CDF_ng kg	DOUBLE		1,2,3,7,8 Penta-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 13 einheit_ido: 112
1,2,3,7,8 Penta-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
2,3,4,7,8 Penta-CDF_ng kg	DOUBLE		2,3,4,7,8 Penta-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 16 einheit_ido: 112
2,3,4,7,8 Penta-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
1,2,3,4,7,8 Hexa-CDF_ng kg	DOUBLE		1,2,3,4,7,8 Hexa-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 5 einheit_ido: 112
1,2,3,4,7,8 Hexa-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
1,2,3,6,7,8 Hexa-CDF_ng kg	DOUBLE		1,2,3,6,7,8 Hexa-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 9 einheit_ido: 112
1,2,3,6,7,8 Hexa-CDF_ng kg BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
1,2,3,7,8,9 Hexa-CDF_ng kg	DOUBLE		1,2,3,7,8,9 Hexa-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 11 einheit_ido: 112

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
1,2,3,7,8,9 Hexa-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
2,3,4,6,7,8 Hexa-CDF_ng kg	DOUBLE		2,3,4,6,7,8 Hexa-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 15 einheit_ido: 112
2,3,4,6,7,8 Hexa-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
1,2,3,4,6,7,8 Hepta-CDF_ng kg	DOUBLE		1,2,3,4,6,7,8 Hepta-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 2 einheit_ido: 112
1,2,3,4,6,7,8 Hepta-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
1,2,3,4,7,8,9 Hepta-CDF_ng kg	DOUBLE		1,2,3,4,7,8,9 Hepta-CDF (ng/kg)		A	m_bodenmesswerte: messwert messgroesse_ido: 3 einheit_ido: 112
1,2,3,4,7,8,9 Hepta-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Octa-CDF_ng kg	DOUBLE		Octa-CDF (ng/kg)		A	m_bodenmesswerte:

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						messwert messgroesse_ido: 185 einheit_ido: 112
Octa-CDF_ng kg_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
Summe PCDD PCDF Lowerbound	DOUBLE		Summe der Dioxine und Furane mit lowerbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf 0 (Null) gesetzt)		N	
Summe PCDD PCDF Upperbound	DOUBLE		Summe der Dioxine und Furane mit upperbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf Bestimmungsgrenze gesetzt)		N	
PCB 77	DOUBLE		PCB 77 in ng/kg		A	m_bodenmesswerte: messwert

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						messgroesse_ido: 215 einheit_ido: 112
PCB 77_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
PCB 81	DOUBLE		PCB 81 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 217 einheit_ido: 112
PCB 81_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
PCB 126	DOUBLE		PCB 126 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 201 einheit_ido: 112
PCB 126_BG	DOUBLE		BG = Bestimmungsgrenze		A	zusatz_ido: 2
PCB 169	DOUBLE		PCB 169 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 207 einheit_ido: 112

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
PCB 169_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB 105	DOUBLE		PCB 105 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 197 einheit_ido: 112
PCB 105_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB 114	DOUBLE		PCB 114 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 198 einheit_ido: 112
PCB 114_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB 118	DOUBLE		PCB 118 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 199 einheit_ido: 112
PCB 118_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB 123	DOUBLE		PCB 123 in ng/kg		A	m_bodenmesswerte:

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
						messwert messgroesse_ido: 200 einheit_ido: 112
PCB 123_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB 156	DOUBLE		PCB 156 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 204 einheit_ido: 112
PCB 156_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB 157	DOUBLE		PCB 157 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 205 einheit_ido: 6
PCB 157_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB 167	DOUBLE		PCB 167 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 206 einheit_ido: 112

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung eBIS
PCB 167_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
PCB 189	DOUBLE		PCB 189 in ng/kg		A	m_bodenmesswerte: messwert messgroesse_ido: 210 einheit_ido: 112
PCB 189_BG	DOUBLE		BG = Bestimmungsgrenze		A	<i>zusatz_ido: 2</i>
Summe PCB coplanar Lowerbound	DOUBLE		Summe der 12 dl-PCB mit lowerbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf 0 (Null) gesetzt) in ng/kg		N	
Summe PCB coplanar upperbound	DOUBLE		Summe der 12 dl-PCB mit upperbound Ansatz bestimmt (Werte unterhalb Bestimmungsgrenze auf Bestimmungsgrenze gesetzt) in ng/kg		N	

A.6 Tabellenblätter Umweltprobenbank

Tabelle 23: Übersicht Tabellen AggregierteWerte_Boden_Analysen

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
Aggregat-ID		Interne Aggregation Identifikation		B	-- (nicht abbildbar, da keine direkte Entsprechung in eBIS)
Type		Parameter Typ		N	
Probenahme Jahr		Jahr der Probenahme		A	m_bodenmesswerte:probenahmedatum
Probenmischung?		0=nein (Einzelprobe); 1=ja (Mischprobe);		N	
Mischungsnummer		Nummer der Probenmischung		N	
Probenzustand		0 : Mischprobe; 1: Einzelprobe		N	
Geschlecht		Bezeichnung des Geschlechts; wenn leer dann nicht zutreffend bzw. ohne Geschlechterunterscheidung		N	
Geschlecht-UPBCode		UPBCode des Geschlechts; wenn leer dann nicht zutreffend bzw. ohne Geschlechterunterscheidung		N	
Ort		Name des Ortsbezugs		A	uf_titeldaten:name_bzw_ort
Ort-UPBCode		UPB-Code des Ortsbezugs		A	uf_titeldaten:uba_kennung
Probenart		Probenart (Im Abzug immer Boden)		N	
Probenart-UPBCode		UPBCode der Probenart		N	

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
Matrix		Matrix		A	uf_profilkennzeichnung:bemerkung_profilkennz
Probenart-Matrix-UPBCode		UPB-Code der Probenart-Matrix		A	uf_titeldaten:kennung
Analysemethode		Analysenmethode: Bei analysierten Messwerten		A	m_bodenmesswerte:messverfahren_ido → lists_messverfahren
Extraktionsmethode		Extraktionsmethode: Bei Chemischen Belastungswerten		A	m_bodenmesswerte:extrakt_ido → lists_extrakte_boden
Parameter		Name des Parameters		A	m_bodenmesswerte:messgroesse_ido → lists_messgroessen - pH-Wert (Calciumchlorid) - pH-Wert (Kaliumchlorid) - pH-Wert (Wasser) erhalten den Schlüssel 227 in Verbindung mit dem jeweiligen Extraktionsmittel
Parameter-UPBCode		UPBCode des Parameters		N	
Probenzusatz		0 : Routineprobe; 8 : Screening; 9 : Sonderversuch; Kombinationen werden absteigend sortiert zusammengehängt (z.B. 80, 90, 98)		N	
RTM		Die RTM-Nummer der Probenahme (real time monitoring)		N	
Probenahme Monat		Monat der Probenahme		A	m_bodenmesswerte:probennahmedatum
Institut		Probenahme- oder Analyseinstitut		C	uf_titeldaten:behoerde_organisation
Institut-UPBCode		UPBCode des Instituts		N	

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
Listeneintrag		Name des Listeneintrags eines kategorischen Parameters		A	uf_profilkennzeichnung:bemerkung_profilkennz
Listeneintrag-UPBCode		UPBCode des Listeneintrag eines kategorischen Parameters		N	
Anteil Listeneintrag		Anteil des Listeneintrags am Gesamtergebnis		N	
Anzahl Individuen		Anzahl zugrundeliegender Individuen bzw. Anzahl paralleler Messreihen		N	
Anzahl Messungen		Die zugrundeliegende Fallzahl einzelner Individuen bzw. die Anzahl untersuchter Teilproben von Mischproben		A	m_bodenmesswerte:anzahl_einzelwerte
Anzahl Messungen ≥ LOQ		Anzahl Messungen größer oder gleich der Bestimmungsgrenze (LOQ) für Analysen		N	
Anzahl Messungen > Referenz Max. 1		Anzahl Messungen größer oder gleich des Referenz-Maximums 1		N	
Name Referenz Max. 1		Name des Referenz-Maximums 1		N	
Anzahl Messungen > Referenz Max. 2		Anzahl Messungen größer oder gleich des Referenz-Maximums 2		N	
Name Referenz Max. 2		Name des Referenz-Maximums 2		N	
Anzahl gelieferter LOQs		Anzahl der gelieferten LOQs		N	
Ältestes Analysedatum		Ältestes Analysedatum der zugrundeliegenden Messwerte		B	--
Jüngstes Analysedatum		Jüngstes Analysedatum der zugrundeliegenden Messwerte		B	m_bodenmesswerte:analysedatum

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
Letzte Wertänderung		Letzte Änderung des Werts		N	
Sex list item		Interne Identifikation des Listeneintrags		N	
Plan		Zugrundeliegender Probenahme- oder Analysenplan		N	
Einheit		Messeinheit		A	m_bodenmesswerte:einheit_ido → lists.einheiten
Ab Freigabelevel		Freigabelevel der Daten		N	
Export min aliquots stddev		Aliquot-Standardabweichung min		N	
Export max aliquots stddev		Aliquot-Standardabweichung max		N	
Export loq min		Kleinster LOQ: die kleinste ermittelte Bestimmungsgrenze im Verlauf der Analysen		A	-- (Bestimmungsgrenze kann in eBIS nicht explizit hinterlegt werden. Insbesondere bei gemittelten Werten aus mehreren Proben, die mit unterschiedlichen Analysemethoden gewonnen wurden, stößt die eBIS-Struktur an ihre Grenzen)
Export loq max		Größter LOQ: die größte ermittelte Bestimmungsgrenze im Verlauf der Analysen		A	-- (vgl. Export loq min)
Export am		Arithmetischer Mittelwert		A	m_bodenmesswerte:mittelwert
Export stddev		Standardabweichung		N	
Export gm		Geometrischer Mittelwert		N	
Export gm ci lb		Geom. 95%-KI (u. Gr.): 95% Konfidenzintervall für Geometrischer Mittelwert - Untere Grenze		N	

Feld	Datentyp	Bedeutung	Vollständigkeit [0..100%]	Priorität	Entsprechung in eBIS
Export gm ci ub		Geom. 95%-KI (o. Gr.): 95% Konfidenzintervall für Geometrischer Mittelwert - Obere Grenze		N	
Export min		Minimum		A	m_bodenmesswerte:min
Export p5		5. Perzentil		N	
Export p10		10. Perzentil		N	
Export p25		25. Perzentil		N	
Export p50		50. Perzentil (Median)		A	m_bodenmesswerte:median
Export p75		75. Perzentil		N	
Export p90		90. Perzentil		N	
Export p95		95. Perzentil		N	
Export max		Maximum		A	m_bodenmesswerte:max
Export ref max 1		Wert des Referenz-Maximums 1		N	
Export ref max 2		Wert des Referenz-Maximums 2		N	

A.7 Tabellenblätter Sonstige – Sickerwasser

Nachfolgende Tabelle fasst die Definitionen wiederkehrender Datenfelder in den nachfolgenden Tabellen zusammen.

Tabelle 24: Übersicht „Kopfdaten“

Feld	Bedeutung	Priorität	Abbildung nach eBIS
Standort Nr.	ID des Standortes	A	uf_titeldaten:kennung
Land	Nationalstaat (DE)	N	-
Bundesland		A	uf_titeldaten:bundesland_id
Standort	Name des Standortes	A	uf_titeldaten:name_bzw_ort
Rechtswert	Rechtswert GK3	A	uf_titeldaten:rechtswert
Hochwert	Hochwert GK3	A	uf_titeldaten:hochwert
Profiltyp		C	m_bodenmesswerte:probengewinnung_id → lists_probengewinnungen
Nutzung		A	uf_aufnahmesituation:nutzung_allgemein_ka5_id → lists_nutzungsarten_ka5
Bodentyp		A	uf_profilkennzeichnung:bodensubtyp_id → lists_bodensubtypen; uf_profilkennzeichnung:bodentyp_id → lists_bodentypen
Bodenausgangsgesteinsgruppe		A	uf_profil_substratsystematik:bodenausgangsgestein_typ_id → lists_bodenausgangsgesteinsgruppe_fein
Substratsystematik		A	uf_profilkennzeichnung:bemerkung_profilkennz

Tabelle 25: Übersicht „KGV_BO“

Feld	Bedeutung	Priorität	Abbildung nach eBIS
Standort Nr	ID Standort		
Standort/ Profil Nr.		N	
Horizont			

Feld	Bedeutung	Priorität	Abbildung nach eBIS
LIMS/ ext. Proben-Nr.			
Beprobungstiefe von	(Oberkante)		
Beprobungstiefe bis	(Unterkante)		
Ton	Tongehalt	A	messgroesse:276 einheit:31
fU		A	messgroesse:107 einheit:31
mU		A	messgroesse:168 einheit:31
gU		A	messgroesse:117 einheit:31
T+U		N	
Schluff	Schluffgehalt	A	messgroesse:280 einheit:31
ffS		A	messgroesse:102 einheit:31
fS		A	messgroesse:105 einheit:31
gfS		N	-
mS		A	messgroesse:166 einheit:31
gS		A	messgroesse:115 einheit:31
Sand	Sandgehalt	A	messgroesse:254 einheit:31

Tabelle 26: Übersicht „KW_BO“

Feld	Bedeutung	Priorität	Abbildung nach eBIS
Standort Nr.	ID Standort		
Standort/ Profil Nr.			
Horizont			
LIMS/ext.Proben-Nr.			

Feld	Bedeutung	Priorität	Abbildung nach eBIS
Mo	Stoffkonzentration in mg/kg	A	messgroesse:165 einheit: 1
Al	Stoffkonzentration in mg/kg	A	messgroesse:36 einheit: 1
U	Stoffkonzentration in mg/kg	A	messgroesse:278 einheit: 1
As	Stoffkonzentration in mg/kg	A	messgroesse:43 einheit: 1
B	Stoffkonzentration in mg/kg	A	messgroesse:45 einheit: 1
Ba	Stoffkonzentration in mg/kg	A	messgroesse:46 einheit: 1
Ca	Stoffkonzentration in mg/kg	A	messgroesse:59 einheit: 1
Cd	Stoffkonzentration in mg/kg	A	messgroesse:61 einheit: 1
Co	Stoffkonzentration in mg/kg	A	messgroesse:69 einheit: 1
Cr	Stoffkonzentration in mg/kg	A	messgroesse:74 einheit: 1
Cu	Stoffkonzentration in mg/kg	A	messgroesse:78
Fe	Stoffkonzentration in mg/kg	A	messgroesse:99 einheit: 1
K	Stoffkonzentration in mg/kg	A	messgroesse:128 einheit: 1
Mg	Stoffkonzentration in mg/kg	A	messgroesse:155 einheit: 1
Mn	Stoffkonzentration in mg/kg	A	messgroesse:164 einheit: 1
Na	Stoffkonzentration in mg/kg	A	messgroesse:169 einheit: 1
Ni	Stoffkonzentration in mg/kg	A	messgroesse:178 einheit: 1
P	Stoffkonzentration in mg/kg	A	messgroesse:190 einheit: 1
Pb	Stoffkonzentration in mg/kg	A	messgroesse:192 einheit: 1
S	Stoffkonzentration in mg/kg	A	messgroesse:248 einheit: 1
Sb	Stoffkonzentration in mg/kg	A	messgroesse:244

Feld	Bedeutung	Priorität	Abbildung nach eBIS
			einheit: 1
Se	Stoffkonzentration in mg/kg	A	messgroesse:247 einheit: 1
Sn	Stoffkonzentration in mg/kg	A	messgroesse:191 einheit: 1
Tl	Stoffkonzentration in mg/kg	A	messgroesse:299 einheit: 1
V	Stoffkonzentration in mg/kg	A	messgroesse:281 einheit: 1
Zn	Stoffkonzentration in mg/kg	A	messgroesse:251 einheit: 1

Tabelle 27: Übersicht „KAK_BO“

Feld ²⁴	Bedeutung	Priorität	Abbildung nach eBIS
Standort Nr.			
Standort		A	
Herkunft Probennahmepunkt		A	
Horizont		A	
LIMS/ext. Proben Nr.			
Beprobungstiefe von		A	
Beprobungstiefe bis		A	
Ca *			
H *			
K *			
KAKp *		A	messgroesse:135
Mg *			
Na *			
Al **			
Ca **			

²⁴ * KAKp nach Mehlich | ** KAKe mit BaCl₂ ungepuffert

Feld ²⁴	Bedeutung	Priorität	Abbildung nach eBIS
Fe **			
H **			
K **			
KAKe **		A	messgroesse:132
Mg **			
Mn **			
Na **			

A.8 Tabellenblätter Bodendaten Klimafolgenforschung / Erosionsdaten

Tabelle 28: Datentabelle Abtrag_Ist_Naturraum_UBA – Datenfelder

Feld	Bedeutung	Priorität
Naturraum ID	Id des Naturraums, codiert von BfN nach Meyen-Schmithüsen	C
Fläche [ha]	Fläche, die der Naturraum einnimmt in Hektar	B
Name	Name des Naturraums (BfN)	A
Haupteinheit	Name der übergeordneten Haupteinheit (BfN)	A
Naturraum_Zwischenebene Name2	Name des Naturraums (untergeordnet, BfN)	B
Bundesland (Hauptanteil) Land	Name des Bundeslandes, das den größten Anteil am jeweiligen Naturraum hat	B
Bemerkung BEMERKUNG	-	B
Fläche(ha) LW (AL+GL)	Flächenanteil Landwirtschaft (Ackerland und Grünland) am Naturraum	B
Acker (AL)	Flächenanteil Ackerland am Naturraum	B
Anteil (%) LW (AL+GL)	Prozentualer Flächenanteil Landwirtschaft (Ackerland und Grünland) am Naturraum	B
Acker (AL)	Prozentualer Flächenanteil Ackerland am Naturraum	B
C- Faktoren für verschiedene Annahmen der Bodenbearbeitung (Ist-Zustand)		

Feld	Bedeutung	Priorität
100% konventionell	Bewirtschaftungs/Bodenbedeckungsfaktor (C-Faktor) aus der ABAG Gleichung für konventionelle Landwirtschaft	C
25% konservierend	Bewirtschaftungs/Bodenbedeckungsfaktor (C-Faktor) aus der ABAG Gleichung für 25 % konservierende Bodenbearbeitung	C
50% konservierend	Bewirtschaftungs/Bodenbedeckungsfaktor (C-Faktor) aus der ABAG Gleichung für 50 % konservierende Bodenbearbeitung	C
100% konservierend	Bewirtschaftungs/Bodenbedeckungsfaktor (C-Faktor) aus der ABAG Gleichung für 100 % konservierende Bodenbearbeitung	C

Mittlerer Bodenabtrag auf LW-Fläche im Naturraum (RxKxS) [t/ha*a]

Abtrag mit R-Faktor aus Sommer-N	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Sommer Niederschlag und den angegebenen Faktoren	A
Abtrag mit R-Faktor aus Jahres-N	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Jahres Niederschlag und den angegebenen Faktoren	A

Mittlerer Bodenabtrag auf LW-Fläche im Naturraum (RxKxLxS) [t/ha*a]

Abtrag mit R-Faktor aus Sommer-N	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Sommer Niederschlag und den angegebenen Faktoren	A
Abtrag mit R-Faktor aus Jahres-N	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Jahres Niederschlag und den angegebenen Faktoren	A

Mittlerer Bodenabtrag auf Ackerfläche im Naturraum (CxRxKxLxS) [t/ha*a] (R-Faktor aus Sommer-N)

100% konventionell	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Sommer Niederschlag und den angegebenen Faktoren sowie der angegebene Bewirtschaftungsweise	A
25% konservierend	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Sommer Niederschlag und den angegebenen	A

Feld	Bedeutung	Priorität
	Faktoren sowie der angegebene Bewirtschaftungsweise	
50% konservierend	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Sommer Niederschlag und den angegebenen Faktoren sowie der angegebene Bewirtschaftungsweise	A
100% konservierend	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Sommer Niederschlag und den angegebenen Faktoren sowie der angegebene Bewirtschaftungsweise	A
Mittlerer Bodenabtrag auf Ackerfläche im Naturraum (CxRxKxLxS) [t/ha*a] (R-Faktor aus Jahres-N)		
100% konventionell	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Jahres Niederschlag und den angegebenen Faktoren sowie der angegebene Bewirtschaftungsweise	A
25% konservierend	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Jahres Niederschlag und den angegebenen Faktoren sowie der angegebene Bewirtschaftungsweise	A
50% konservierend	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Jahres Niederschlag und den angegebenen Faktoren sowie der angegebene Bewirtschaftungsweise	A
100% konservierend	Abtrag auf Landwirtschaftsflächen errechnet mit Erosivität der Niederschläge (R-Faktor der ABAG) für den Jahres Niederschlag und den angegebenen Faktoren sowie der angegebene Bewirtschaftungsweise	A

A.9 Datenbanktabelle Boden – Bodenorganismen

Tabelle 29: Übersicht Tabelle „datenbank“

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
ID	int		Unique ID	100	A
ID_ERSTELLUNG	timestamp		Datensatz: Erstellung / Datum des Eintrags (dd.mm.yyyy)	100	A
ID_AENDERUNG	timestamp		Datensatz: Änderung der ID_Erstellung (dd.mm.yyyy)	100	A
ID_BEARBEITER	varchar		Datensatz: Person, die Eintrag erstellt hat	100	A
Ref_AUTOR	varchar		Literaturzitat: Autor(en) der Literaturangabe	100	A
Ref_JAHR1	varchar		Literaturzitat: Jahr der Literaturangabe	100	A
Ref_JAHR2	int	Ref_JAHR2.id	Literaturzitat: Zusatz zum Jahr der Literaturangabe zur Unterscheidung, wenn Literatur desselben Autors im selben Jahr mehrfach genannt ist	41	A
Ref_k11_Nr#	varchar		Literaturzitat: Verknüpfung zu Reference Manager Literaturverzeichnis; Kombination aus Autor und Jahr	100	A
Ref_TITEL	varchar		Literaturzitat: Titel der Literaturangabe	99	A
Ref_MEDIUM	varchar		Literaturzitat: Herkunft der Literaturangabe (Buch, Zeitschrift, usw.)	100	A
Ref_BAND	varchar		Literaturzitat: Band der Literaturangabe (bei Buchreihen)	86	A
Ref_HEFT	varchar		Literaturzitat: Heft der Literaturangabe (bei Zeitschriften)	50	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
Ref_STARTSEITE	varchar		Literaturzitat: Erste Seite der Literaturangabe	97	A
Ref_ENDSEITE	varchar		Literaturzitat: Letzte Seite der Literaturangabe	97	A
Org_TESTORGANISMUS	varchar		Testorganismus: Organismen, die in der Literaturangabe getestet sind	100	A
Org_KLASSE	varchar		Testorganismus: Organismenklasse zu der der Testorganismus gehört	75	A
Org_TAXLEVEL	varchar		Testorganismus: Taxonomische Ebene, zu der der Testorganismus gehört	100	A
Org_GRUPPE	int	Org_GRUPPE.ID	Testorganismus: Gruppe (Pflanzen, Tiere, Mikroorganismen) zu der der Testorganismus gehört	100	A
Org_LEBENSFORM	varchar		Testorganismus: Lebensformtyp/Lebensraum des Testorganismus	99	A
Org_TROPHIESTUFE	varchar		Testorganismus: Trophiestufe (Ernährung) des Testorganismus	96	A
Stoff_BEZEICHNUNG	varchar		Schadstoff: Name des untersuchten/zugesetzten Wirkstoffes	100	A
Stoff_BEZCAS	varchar		Schadstoff: CAS (Chemical Abstracts Service) Nummer des Wirkstoffes	38	A
Stoff_SUBSTANZ	varchar		Schadstoff: Applizierte Verbindung der Wirksubstanz	100	A
Stoff_SUBSTCAS	varchar		Schadstoff: CAS (Chemical Abstracts Service) Nummer der Verbindung	38	A
Stoff_GRUPPE	int	Stoff_GRUPPE.ID	Schadstoff: Stoffgruppe (Schwermetall, Pestizide, usw.) der Wirksubstanz	100	A
Meth_VERSUCHSANSATZ	int	Meth_VERSUCHSANSATZ.ID	Untersuchungsmethode: Beschreibung Versuchsansatz; Freiland oder Laborversuch	95	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
Meth_RICHTLINIE	int	Meth_RICHTLINIE.ID	Untersuchungsmethode: Richtlinie (Versuchsansatz) der eingesetzten Untersuchungsmethode	64	A
Meth_TESTMEDIUM	int	Meth_TESTMEDIUM.ID	Untersuchungsmethode: Testmedium/Untersuchungsmedium: Streu, Boden, Kunstaboden	95	A
Meth_BODEN1	int	Meth_BODEN1.ID	Untersuchungsmethode: Spezifizierung des Untersuchungsmediums	94	A
Meth_BODEN2	varchar		Untersuchungsmethode: Weitere Spezifizierung des Untersuchungsmediums	98	A
Meth_AUFNAHME	int	Meth_AUFNAHME.ID	Untersuchungsmethode: Aufnahme des Wirkstoffes in den Organismus	38	A
Meth_APPLIKATION	varchar		Untersuchungsmethode: Art der Applikation des Wirkstoffes (Applikationsform)	73	A
Faktor_BODENPH	varchar		Toxizität beeinflussende Faktoren: Bodeneigenschaft: pH	97	A
Faktor_PHMETHODE	varchar		Toxizität beeinflussende Faktoren: Extraktionsmittel für Boden pH Messung	72	A
Faktor_BODENKAK	varchar		Toxizität beeinflussende Faktoren: Bodeneigenschaft: Kationenaustauschkapazität (KAK)	53	A
Faktor_KAKMETHODE	varchar		Toxizität beeinflussende Faktoren: Einheit KAK	53	A
Faktor_BODENC	varchar		Toxizität beeinflussende Faktoren: Bodeneigenschaft: Kohlenstoff; Einheit nicht überall angegeben	95	A
Faktor_BODENCN	varchar		Toxizität beeinflussende Faktoren: C: Stickstoff; Einheit nicht überall angegeben	39	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
Faktor_CN METHODE	varchar		Toxizität beeinflussende Faktoren: Angaben zur C oder N Bestimmungsmethode und um welche C oder N Form es sich handelt (organisch/anorganisch/gesamt)	93	A
Faktor_BODENART	varchar		Toxizität beeinflussende Faktoren: Bodeneigenschaft: Textur	73	A
Faktor_BODENANGABEN	varchar		Toxizität beeinflussende Faktoren: Angaben zum Boden: spezifische Angaben zur Textur; Bodentyp; Bodenhorizont; Auflage; Besonderheiten	82	A
Faktor_KONTDAUER	varchar		Toxizität beeinflussende Faktoren: Dauer der Kontamination	46	A
Faktor_EINHEITKONT	int	Faktor_EINHEITKONT.ID	Toxizität beeinflussende Faktoren: Einheit: Dauer der Kontamination	47	A
Faktor_EXPOSZEIT	varchar		Toxizität beeinflussende Faktoren: Dauer der Expositionszeit (Organismus-Wirkstoff)	39	A
Faktor_EINHEITEXPOS	int	Faktor_EINHEITEXPOS.ID	Toxizität beeinflussende Faktoren: Einheit: Dauer der Expositionszeit	39	A
Test_EBENE	int	Test_EBENE.ID	Testspezifikationen: Art des Tests (z.B. Langzeit)	95	A
Test_EFFEK TALLG	int	Test_EFFEK TALLG.ID	Testspezifikationen: Art des Effekts (Hemmung oder Förderung des Organismus)	100	A
Test_EFFEK TSPEZ	varchar		Testspezifikationen: Spezifizierung des Effektes auf den Organismus	100	A
Test_ENDPUNKT	varchar		Testspezifikationen: Endpunkt; Methode der Messung am Endpunkt	99	A
Test_METHENDPUNKT	varchar		Testspezifikationen: Spezifizierung der Methode am Endpunkt	38	A
Test_KONZLOW	varchar		Testspezifikationen: Niedrigste Konzentration des Wirkstoffes (von...)	76	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
Test_KONZHIGH	varchar		Testspezifikationen: Höchste Konzentration des Wirkstoffes (...bis)	76	A
Test_KONZEINHEIT	varchar		Testspezifikationen: Einheit der Konzentration des Wirkstoffes	76	A
Test_KONZSTUFEN	varchar		Testspezifikationen: Anzahl der Konzentrationsstufen	76	A
Test_DAUER	varchar		Testspezifikationen: Versuchsdauer	93	A
Test_EINHEITDAUER	varchar		Testspezifikationen: Einheit der Versuchsdauer (Tage, Stunden)	93	A
Test_RELIABILITY	int	Test_RELIABILITY.ID	Testspezifikationen: Datenverlässlichkeit nach Klimisch et al. 1997	100	A
Erg_NOECGRKL	varchar		Ergebnis: Größer oder kleiner als No observed effect concentration = Höchste Konzentration ohne beobachtete toxische Wirkung	42	A
Erg_NOECWERT	varchar		Ergebnis: No observed effect concentration = Höchste Konzentration ohne beobachtete toxische Wirkung	55	A
Erg_NOECEINHEIT	varchar		Ergebnis: Einheit der No observed effect concentration = Höchste Konzentration ohne beobachtete toxische Wirkung	55	A
Erg_NOECBEZUG	varchar		Ergebnis: Bezugsmedium für NOEC	53	A
Erg_NOECEND	varchar			55	A
Erg_NOECANMERK	varchar		Ergebnis: Anmerkung zur No observed effect concentration	37	A
Erg_LOECGRKL	varchar		Ergebnis: Größer oder kleiner als Lowest Observed Effect Concentration	40	A
Erg_LOECWERT	varchar		Ergebnis: Lowest Observed Effect Concentration	50	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
Erg_LOECEINHEIT	varchar		Ergebnis: Einheit der Lowest Observed Effect Concentration	50	A
Erg_LOECBEZUG	varchar		Ergebnis: Bezugsmedium für Lowest Observed Effect Concentration	49	A
Erg_LOECEND	varchar			51	A
Erg_LOECANMERK	varchar		Ergebnis: Anmerkung zu Lowest Observed Effect Concentration	38	A
Erg_EC20GRKL	varchar		Ergebnis: Kleiner oder größer als Effect concentration, bei der 20 % der Testorganismen betroffen sind	38	A
Erg_EC20WERT	varchar		Ergebnis: Effect concentration, bei der 20 % der Testorganismen betroffen sind	48	A
Erg_EC20EINHEIT	varchar		Ergebnis: Einheit von als Effect concentration, bei der 20 % der Testorganismen betroffen sind	48	A
Erg_EC20BEZUG	varchar		Ergebnis: Bezugsmedium für Effect concentration, bei der 20 % der Testorganismen betroffen sind	46	A
Erg_EC20END	varchar			48	A
Erg_EC20ANMERK	varchar		Ergebnis: Anmerkung Effect concentration, bei der 20 % der Testorganismen betroffen sind	38	A
Erg_EC50GRKL	varchar		Ergebnis: Kleiner oder größer als Effect concentration, bei der 50 % der Testorganismen betroffen sind	42	A
Erg_EC50WERT	varchar		Ergebnis: Effect concentration, bei der 50 % der Testorganismen betroffen sind	73	A
Erg_EC50EINHEIT	varchar		Ergebnis: Einheit von Effect concentration, bei der 50 % der Testorganismen betroffen sind	73	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
Erg_EC50BEZUG	varchar		Ergebnis: Bezugsmedium für Effect concentration, bei der 50 % der Testorganismen betroffen sind	66	A
Erg_EC50END	varchar			73	A
Erg_EC50ANMERK	varchar		Ergebnis: Anmerkung Effect concentration, bei der 50 % der Testorganismen betroffen sind	38	A
Erg_EC70GRKL	varchar		Ergebnis: Kleiner oder größer als Effect concentration, bei der 70 % der Testorganismen betroffen sind	38	A
Erg_EC70WERT	varchar		Ergebnis: Effect concentration, bei der 70 % der Testorganismen betroffen sind	40	A
Erg_EC70EINHEIT	varchar		Ergebnis: Einheit von Effect concentration, bei der 70 % der Testorganismen betroffen sind	40	A
Erg_EC70BEZUG	varchar		Ergebnis: Bezugsmedium für Effect concentration, bei der 70 % der Testorganismen betroffen sind	38	A
Erg_EC70END	varchar			40	A
Erg_EC70ANMERK	varchar		Ergebnis: Anmerkung Effect concentration, bei der 70 % der Testorganismen betroffen sind	38	A
Erg_LCXGRKL	varchar		Ergebnis: Kleiner oder größer als Lethal concentration = Konzentration, die die Abtötung (von x %) der Testorganismen bewirkt	40	A
Erg_LCXWERT	varchar		Ergebnis: Lethal concentration = Konzentration, die die Abtötung (von x %) der Testorganismen bewirkt	55	A
Erg_LCXEINHEIT	varchar		Ergebnis: Einheit von Lethal concentration = Konzentration, die die Abtötung (von x %) der Testorganismen bewirkt	55	A

Feld	Datentyp	Verweis (foreign key)	Bedeutung	Vollständigkeit [0..100%]	Priorität
Erg_LCXBEZUG	varchar		Ergebnis: Bezugsmedium von Lethal concentration = Konzentration, die die Abtötung (von x %) der Testorganismen bewirkt	54	A
Erg_LCXEND	varchar		Ergebnis:?	55	A
Erg_LCXANMERK	varchar		Ergebnis: Anmerkung Lethal concentration; Spezifikation von x %, z.B. LC 50= Konzentration, die die Abtötung von 50 %) der Testorganismen bewirkt	38	A
Bem_BEMERKUNG	varchar		Bemerkungen: Bemerkungen zur Studie	51	A

Für weitergehende Informationen: <https://cfpub.epa.gov/ecotox/help.cfm?sub=wi-definitions>

B Sonstige Anlagen

B.1 Checkliste der Projektfragen

Welche Stoffdaten und Begleitparameter sind in den Quelldatenbanken vorhanden?

✓ *Die zu integrierenden Datenbanken wurden entsprechend untersucht und die gefundenen Stoffe und Begleitparameter dokumentiert.*

Wie liegen die Daten vor (Access, Excel, txt, csv, XML ...)?

✓ *Die Daten liegen in unterschiedlichen Formaten vor. Die für die Übernahme am besten geeigneten Dateiformate sind in den Beschreibungen der Quelldatenbanken dokumentiert.*

Sind die Datenquellen untereinander austauschbar und harmonisiert?

✓ *Grundsätzlich ist eine Integration über eine kleinteilige Transformation der Datenschemata möglich. Die Transformation ist jedoch teilweise verlustbehaftet, da einige Zusatzinformationen nicht im eBIS-Schema abbildbar sind. Um Informationsverluste zu vermeiden, werden die Quelldaten in eigenen Schemata unverändert importiert und können zur Klärung von Detailfragen und Unklarheiten in der gleichen technischen Umgebung ausgewertet werden.*

Inwieweit sind die Daten qualitätsgesichert, gibt es z.B. doppelte Datensätze?

Allgemeine Aussagen zur Qualitätssicherung sind den Datensatzdokumentationen nicht zu entnehmen. Die Identifikation von Duplikaten nach der Zusammenführung der Datenbanken anhand von Ähnlichkeitsabfragen lieferte vielfältige Hinweise auf Dubletten – sowohl in eBIS selbst als auch im integrierten Datenbestand.

Müssen Schlüssellisten erstellt werden, um die Daten der einzelnen Quellen untereinander leicht und verlustfrei zusammenzuführen?

✓ *Ja. Dieser Vorgang muss manuell durchgeführt werden, ist grundsätzlich fehleranfällig und beansprucht sehr viel Zeit in der Zusammenstellung und Kontrolle.*

Sind die Datenmodelle von allen Quellen vorhanden?

✓ *Für die dauerhaft gepflegten Datenbanken sind Dokumentationen und Datenmodelle vorhanden. Einige der Datenmodelle, insbesondere solche mit einmaligem Projektbezug, mussten manuell via „reverse engineering“ in Zusammenarbeit mit dem UBA erstellt werden. Vielfach wird auch mit schwach getypten Datenmodellen gearbeitet, was die Fehleranfälligkeit erhöht (z.B. wenn Datumsangaben oder numerische Werte als Text codiert sind oder unklare Erfassungsvorschriften genutzt werden).*

Wie werden verschiedene Horizontmächtigkeiten in den Datenquellen abgebildet?

✓ *Unterschiedliche Horizontmächtigkeiten werden einheitlich als Ober-/Unterkante der Entnahmetiefe in cm abgebildet*

Wie werden unterschiedliche Einheiten geführt?

✓ *Vorkommende Einheiten können in den vorhandenen Schlüsseltabellen abgebildet und ggf. ergänzt werden.*

Wie kann sicher zwischen dem Bezug auf Frischgewicht/Trockengewicht unterschieden werden?

Die Bezugsgröße ist Teil der Einheitsdefinition.

Wie können unterschiedliche Probenahmezeitpunkte unterschieden werden?

✓ Probenahmezeitpunkte sind in den standortscharfen Proben Datenbanken vorhanden. Die Granularität ist allerdings unterschiedlich. Einige Datenbanken bilden die Probenahmezeitpunkte lediglich jahresscharf ab. Die Relevanz eines genauen Probenahmedatums richtet sich nach dem Gegenstand der Auswertung. Insbesondere für Abfragen in der Boden-Pflanze-Paarung ist das genaue Abfragedatum sehr wichtig. Für allgemeine Auswertezwecke reicht häufig das Jahr der Probenahme.

Der Datenbestand enthält mehrere räumliche und/oder zeitliche Wiederholungen bei der Probenahme. Wie können diese kenntlich gemacht werden? Wie können beispielsweise Mischprobe/Einzelprobe/Teilflächenbeprobung aus eBIS kenntlich gemacht werden?

✓ Einzelproben werden in eBIS als einzelner Messwert, Mischproben als Mittelwert gekennzeichnet. Eine explizite Relation zwischen Einzel- und Mischproben (i.S. einer Teil-Ganzes-Beziehung) ist nicht möglich. Eine eindeutige Identifikation von Teilflächenbeprobungen ist strukturell nicht möglich.

Wie können Dopplungen aus verschiedenen Datenquelle kenntlich gemacht werden und wie sorgt man dafür, dass sie in der Auswertung nur einmal auftreten?

✓ Für die Dublettendetektion wurden entsprechende Abfragefunktionen und Kontrolltabellen hinterlegt. Die Nachkontrolle und Deduplikation selbst muss manuell erfolgen. Nur so kann sichergestellt werden, dass Proben nicht mehrfach in der Auswertung auftauchen.

Müssen Schlüssellisten für Tabellenbeziehungen zur Datenharmonisierung erstellt werden?

✓ Ja. Die Erstellung dieser Schlüssellisten muss manuell anhand der vorhandenen Dokumentationsunterlagen erfolgen oder im Rahmen eines Reverse Engineerings aus den vorhandenen Quelldaten selbst abgeleitet werden.

Wie werden die Probezeitpunkte für Zeitreihen harmonisieren bzw. kenntlich gemacht?

✓ Echte Zeitreihen kann das eBIS-Datenmodell strukturell nicht abbilden. Probenahmezeitpunkte sind explizit hinterlegt und können in den Abfragen entsprechend parametrisiert werden.

Mit welchen verschiedenen Analyseverfahren wurden die Messwerte erhoben?

✓ Analyseverfahren sind in einigen Fällen dokumentiert und wurden dann in das eBIS-Schema übernommen.

Wie sind die Bestimmungsgrenzen/Nachweisgrenzen zu den Messwerten?

✓ Explizite Informationen sind i.d.R. nicht vorhanden. Die Bestimmungsgrenze wird als Messwert mit entsprechendem Zusatz (kleiner als) geführt. Eine Nachweisgrenze wird nicht mitgeführt. Ein nicht nachgewiesener Wert sollte in eBIS als mathematische „0“ geführt werden.

Wie können angegebene Messunsicherheiten und zusätzlichen Kontextangaben für die Auswertung bewahrt werden?

✓ Messunsicherheiten werden in den Quelldatensätzen i.d.R. nicht dokumentiert. Unterscheidung zwischen gemessenen und geschätzten Werten ist oft manuell möglich. Um möglichst alle Daten zu erhalten, werden diese inhaltlich unverändert in separate Schemata importiert, sodass die Quelldaten grundsätzlich erhalten bleiben.

Wie ist die Schreibweise der Stoffe in den verschiedenen Datenquellen? Werden allgemeine Klassifizierungen der Stoffe (z. B. „PAK“ als Stoffbegriff/Stoffgruppe) einheitlich verwendet?

✓ *Die Schreibweisen der Stoffe sind höchst unterschiedlich. Klassifizierungen werden – auch innerhalb der Quelldatensätze – nicht immer konsequent genutzt oder durchgehalten.*

Kann die Software „FME“ alle Strukturen problemlos erkennen und die Daten zusammenführen (müssen weitere Schlüssellisten erstellt werden?)

✓ *Die Software kann mit allen Quell- und Zieldatenformaten umgehen. Schlüssellisten müssen größtenteils manuell erstellt werden. Transformationsregeln müssen manuell definiert werden, da vollständig Schlüsselwerte und Bezeichner nahezu komplett uneinheitlich sind.*

Welche Filter müssen in der Software „FME“ zur Datenharmonisierung (Prüfung und Bereinigung) gesetzt werden?

✓ *FME wird für dem Import der Daten und die Abfragen genutzt. Die Harmonisierung der Datenbestände und die Anwendung zusätzlicher Schlüssel Tabellen erfolgt aus Gründen der Übersichtlichkeit und Typsicherheit mit SQL direkt auf der Datenbank.*

Wo können die Daten effizient zwischengespeichert werden, bevor sie ausgewertet werden können (ggf. gesonderte Datenbank)?

✓ *Alle Daten werden in einer Postgresql-Datenbank auf Basis von eBIS zusammengeführt. Die dafür benötigten Workflows zur Datentransformation und die integrierte Datenbank sind für den Einsatz auf Arbeitsplatzrechnern des UBA vorbereitet.*